

# Problemas en la estimación y en el contraste de los rendimientos anormales a largo plazo: Estado de la cuestión<sup>1</sup>

JOSÉ E. FARINÓS

C. JOSÉ GARCÍA

ANA M.<sup>a</sup> IBÁÑEZ

*Universitat de València*

## Resumen:

*El paradigma de la eficiencia ha sido puesto en entredicho en las últimas décadas como consecuencia de la obtención de rendimientos anormales, estadística y económicamente significativos, durante amplios periodos de tiempo tras algunas importantes decisiones empresariales. No obstante, los problemas conceptuales y estadísticos que presenta la medición y contrastación de los rendimientos anormales a largo plazo ha supuesto que la evidencia obtenida pase a calificarse como anomalía. Dada la escasa proliferación de este tipo de estudios en nuestro mercado y el desafortunado desarrollo de algunos de los existentes, en este trabajo presentamos estos problemas y algunas de las soluciones que la literatura propone. Con ello pretendemos facilitar a los investigadores las herramientas necesarias para abordar con éxito este sugerente campo.*

## Palabras clave:

*Rendimientos anormales a largo plazo, problemas en la estimación y contraste, mercado español*

## Abstract:

*Market efficiency has been questioned in the last years as several papers found abnormal returns economically and statistically significant over long horizons after main firm decisions. Nevertheless, the conceptual and statistical problems related with the estimation and test of long-term returns have supposed that the evidence obtained pass to be qualified as anomaly. Given the scarce proliferation of this sort of studies in our market and the unfortunate development of some of them, we present these problems and some of the solutions that the literature proposes. In this way, we seek to facilitate to the investigators the necessary tools to approach with success in this suggestive topic.*

## Key words:

*Long-term abnormal returns, problems in estimation and testing, Spanish market*

---

<sup>1</sup> Los autores agradecen las sugerencias y comentarios de los asistentes al VI Congreso de Matemática Financiera y Actuarial (Valencia, junio de 2002), en especial a Rafael Santamaría, así como a un evaluador anónimo del IVIE. Este trabajo ha contado con el soporte financiero de la CICYT proyecto SEC2000-0773.

## 1. INTRODUCCIÓN

Un tópico recurrente en la literatura financiera de las últimas décadas ha sido la existencia de rendimientos anormales económica y estadísticamente significativos que se extienden durante amplios periodos de tiempo tras algunas importantes decisiones empresariales tales como fusiones [Agarawal *et al.* (1992), Loughran y Vijh (1997), Rau y Vermaelen (1998)], iniciación y omisión de dividendos [Michaely *et al.* (1995)], compra de acciones [Ikenberry *et al.* (1995)], *splits* [Dharan e Ikenberry (1995), Ikenberry *et al.* (1996)], emisión de deuda [Spiess y Affleck-Graves (1999)], emisión de acciones [Ritter (1991), Loughran y Ritter (1995), Spiess y Affleck-Graves (1995), Jegadeesh (2000), Baker y Wurgler (2000)], etc. Si el mercado es eficiente, el comportamiento anormal posterior de las empresas que son protagonistas de un determinado suceso debería ser neutral una vez que la información transmitida por el mismo hubiera sido incorporada por el mercado. El comportamiento anormal a largo plazo de las empresas que llevan a cabo decisiones empresariales relevantes detectado en la literatura pone en cuestión la noción de la eficiencia del mercado.

No obstante, la evidencia obtenida no ha reemplazado la teoría preexistente y el fenómeno de los rendimientos anormales a largo plazo ha pasado a ser catalogado como *anomalía*. Ello se ha debido, fundamentalmente, a tres razones. En primer lugar, porque la hipótesis alternativa es extremadamente vaga (la ineficiencia del mercado), no planteándose una hipótesis alternativa contrastable a la eficiencia del mercado<sup>2</sup>. El segundo motivo tiene que ver directamente con la propia naturaleza del concepto de *eficiencia del mercado* y su capacidad para ser contrastado. Aunque la idea de eficiencia es anterior<sup>3</sup>, Fama (1970) formaliza la teoría y enfatiza el hecho de que la eficiencia del mercado debe ser contrastada junto con un modelo generador de rendimientos esperados *normales*<sup>4</sup>. Y es aquí donde el problema se convierte en ubicuo. Como es sabido, todos los modelos, por definición, son descripciones incompletas de la realidad. Ello implica que los contrastes de la eficiencia del mercado van a estar siempre *contaminados* por el problema de la mala especificación del modelo, por lo que el rechazo de la hipótesis nula (la eficiencia del mercado) puede implicar tanto que los mercados son ineficientes como que el modelo es incorrecto. Un elemento que ha incidido negativamente es esta situación ha sido la escasa capacidad mostrada por los modelos económicos como el *Capital Asset Pricing Model* (CAPM) desarrollado por Sharpe (1964) y Lintner (1965) para explicar la sección cruzada de los rendimientos de los títulos y, en general, para determinar con exactitud las fuentes de riesgo, lo cual ha conducido a los investigadores a dirigir su atención hacia el desarrollo de modelos de valoración empíricos. Dado que no existe ningún tipo de guía en la construcción de estos últimos, el problema de la mala especificación y su crítica se han acentuado con su uso. No obstante, la incidencia de este pro-

---

<sup>2</sup> Puede encontrarse una excepción en los trabajos de Daniel *et al.* (1998) y Barberis *et al.* (1998) en los que se desarrollan dos modelos encuadrados en la denominada como *behavioral finance* (siguiendo a Villalba (2001), *finanzas conductistas*) que presentan hipótesis contrastables. Sin embargo, y como señala Fama (1998), ambos modelos adolecen del mismo problema: explican adecuadamente las anomalías para las cuales han sido desarrollados pero no son capaces de explicar lo que Fama denomina como *the big picture*.

<sup>3</sup> El interés por la eficiencia de los mercados se remonta a una casi olvidada tesis doctoral escrita en 1900 por Louis Bachelier.

<sup>4</sup> Véase también Fama (1976).

blema no puede generalizarse para todo tipo de análisis ya que depende directamente de la amplitud del horizonte estudiado. Así, Fama (1998) resalta que es menos serio en los estudios que se centran en ventanas pequeñas (de unos pocos días)<sup>5</sup>, agravándose a medida que la amplitud del horizonte estudiado se extiende dado que los errores en los rendimientos esperados inducidos por la mala especificación del modelo crecen más rápidamente que la volatilidad de los mismos, pudiendo producir rendimientos anormales estadísticamente significativos.

Estas consideraciones tienen su reflejo en la literatura con la detección de que las anomalías a largo plazo son sensibles a la metodología empleada [Brav *et al.* (2000) y Mitchell y Stafford (2000)], en el sentido de que diferentes modelos generadores de rendimientos esperados producen no sólo diferentes estimaciones de rendimientos anormales a largo plazo sino la desaparición de su significación cuando se emplean determinadas referencias o controles. Aunque, en principio, este hecho reforzaría la argumentación de que la anomalía detectada no supone una evidencia contra la eficiencia del mercado [Fama (1998) y Brav *et al.* (2000)], Loughran y Ritter (2000) contestan argumentando, por un lado, que si realmente el mercado no valora adecuadamente los títulos, entonces los rendimientos anormales *no deberían* ser robustos a metodologías alternativas. Ello sería consecuencia de que no todos los métodos tienen el mismo poder para detectar la mala valoración por parte del mercado<sup>6</sup>. Por otro lado, critican la inclusión de *proxies* de la mala valoración en lugar de verdaderos factores de riesgo como referencias, lo cual conduciría a estos trabajos a sesgar sus contrastes en contra de la detección de rendimientos anormales aun cuando éstos existieran. Desde nuestro punto de vista ésta es una cuestión importante, ya que tan difícilmente justificable es rechazar la eficiencia sin una hipótesis alternativa contrastable como rechazar la ineficiencia mediante el empleo de modelos *ad-hoc*.

Existe, asimismo, un elemento obviado en la anterior discusión acerca de la existencia de rendimientos anormales a largo plazo y la eficiencia. Se trata de la necesidad de un modelo económico (o de equilibrio, como el CAPM) para el contraste de la eficiencia del mercado. En este sentido, Loughran y Ritter (2000) señalan que si de lo que se hace uso es de un modelo positivo (construido empíricamente), entonces estos autores sostienen que no se está contrastando la eficiencia del mercado sino, simplemente, se contrasta si un determinado patrón en los rendimientos es independiente de otros patrones previamente documentados de forma empírica.

El tercer motivo que justifica la consideración como anomalía del comportamiento anormal encontrado está relacionado con la métrica y los problemas de inferencia estadística en los contrastes de rendimientos a largo plazo. Barber y Lyon (1997) y Kothari y Warner (1997) evidencian que la mala especificación está causada por el sesgo debido a la entrada en el mercado de nuevos títulos, el sesgo introducido por el reajuste de las carteras de control y el sesgo debido a la asimetría en la distribución de los rendimientos anormales. Mayor importancia tiene el problema que introduce en los contrastes la dependencia en sección cruzada de los rendimientos anormales a largo plazo [Mitchell y

---

<sup>5</sup> Fama (1998) argumenta que esto es debido a que los rendimientos diarios esperados son próximos a cero, por lo que la mala especificación del modelo generador tiene un impacto despreciable en la estimación de los rendimientos anormales (no esperados).

<sup>6</sup> El poder de un contraste es la probabilidad de rechazar la hipótesis nula cuando ésta es falsa.

Stafford (2000)] como consecuencia del solapamiento que se produce en el cálculo de los rendimientos anormales de las empresas individuales. Por consiguiente, se ha revelado como fundamental el estudio de cuestiones metodológicas relacionadas con la correcta medición y contraste de este tipo de rendimientos, produciéndose en la literatura un verdadero esfuerzo por solventar los anteriores inconvenientes [Lyon *et al.* (1999), Cowan y Sergeant (2001), Jakobsen y Sørensen (2001), Jakobsen y Voetmann (1999), Brav (2000), Mitchell y Stafford (2000)], aunque las soluciones aportadas no son definitivas y, por ende, la correcta especificación de contrastes estadísticos en el caso de los rendimientos anormales a largo plazo continúa siendo una cuestión abierta.

A diferencia de lo que sucede en otros mercados, los trabajos que estudian en el mercado español el comportamiento a largo plazo tras un determinado suceso son escasos y, en ocasiones, transmiten poca confianza. La mayor parte de los mismos estudian el comportamiento posterior de las empresas que realizan una oferta pública de venta u OPV [Ansótegui y Fabregat (1999), García *et al.* (1999), Álvarez (2001), Álvarez y González (2001), Farinós (1999, 2001), Farinós *et al.* (2002)], a excepción del trabajo de Martín y Pastor (2001) que toma como suceso de estudio las ampliaciones de capital. Por otra parte, y aunque no relacionado con un determinado suceso empresarial, encontramos los trabajos relativos al efecto sobre-reacción en el mercado español de Alonso y Rubio (1990) y Forner y Marhuenda (2001, 2002)].

El propósito de este trabajo es el de poner en evidencia las dificultades que entraña el estudio del comportamiento bursátil de las empresas a largo plazo y presentar algunas de las soluciones propuestas por la literatura. De esta forma, confiamos en que los investigadores interesados en este sugerente campo de investigación puedan contar con una referencia que les permita enfocar adecuadamente el problema.

En el desarrollo del presente estudio abordamos dos aspectos fundamentales que dificultan la aceptación de la ineficiencia del mercado. En primer lugar, tratamos el problema de la estimación de los rendimientos anormales e introducimos los dos métodos para su generación, a saber: la utilización de un modelo basado en las características específicas de las empresas (*firm-specific* o *characteristic-based model*); y el empleo de un modelo formal de valoración de activos. A continuación, abordamos el problema de la correcta medida y contraste de los rendimientos a largo plazo de forma separada de acuerdo con las tres aproximaciones generalmente empleadas por lo que se refiere a la estimación de los rendimientos a largo plazo, ya que cada una de éstas presenta dificultades conceptuales y estadísticas propias: la composición (multiplicación) de los rendimientos a corto plazo sobre el horizonte objeto de análisis, la acumulación de los rendimientos a corto plazo, y la utilización del método de las carteras de fecha de calendario (*calendar-time portfolios* o *rolling portfolios*). Finalmente, abordamos brevemente la controversia acerca del esquema de ponderación de los rendimientos apropiado, dado que se ha demostrado que la elección de uno u otro esquema (ponderación *versus* equiponderación) resulta no neutral.

El resto del trabajo se estructura como sigue. En la sección 2 se analizan los problemas teóricos relativos a la estimación de los rendimientos anormales e introducimos los métodos para su generación. El problema de la correcta medida y contraste de los rendimientos a largo plazo se aborda en la sección 3. La sección 4 está dedicada a la discusión del esquema de ponderación de los rendimientos en la medida del comportamiento a largo plazo. En la sección 5 se recogen las principales conclusiones.

## 2. PROBLEMAS RELATIVOS A LA CORRECTA ESTIMACIÓN DE LOS RENDIMIENTOS ANORMALES

La correcta estimación de los rendimientos anormales se encuadra, como señala Fama (1998), dentro del problema de la correcta especificación del modelo generador de rendimientos esperados. En este contexto, existen dos tipos de problemas relacionados con la adecuada especificación del modelo. El primero de ellos tiene que ver, como hemos comentado anteriormente, con la simplificación intrínseca de la realidad que implica cualquier modelo<sup>7</sup>, por lo que el modelo necesariamente no tiene porqué describir los rendimientos esperados. Este es el caso del primer modelo de valoración de activos, el CAPM de Sharpe (1964) y Lintner (1965). Los resultados de los contrastes efectuados sobre el mismo con datos de diferentes mercados internacionales indican que no es capaz de explicar los rendimientos esperados de las empresas pequeñas [Banz (1981)] o la relación existente entre la rentabilidad y algunos cocientes financieros, como el beneficio respecto del precio o el cociente valor contable/valor de mercado<sup>8</sup> [Basu (1983) y Rosenberg *et al.* (1985)]<sup>9</sup>. Por lo tanto, si la composición de la muestra que está siendo analizada está sesgada hacia empresas pequeñas o con unos determinados cocientes financieros, entonces la utilización del CAPM en la generación de rendimientos esperados puede producir rendimientos anormales espurios. El segundo tipo de problemas surge, aún en el caso de que se hiciera uso del verdadero modelo generador de rendimientos esperados, por el mero hecho de que en cualquier periodo de estudio la muestra presente algún patrón específico debido al azar que produzca la aparición de desviaciones sistemáticas de lo predicho por el modelo.

Existen dos aproximaciones posibles para limitar la aparición espuria de una anomalía. Una de ellas se sustenta en la utilización de modelos basados en las características de los títulos que están estudiándose. La estimación de los parámetros del modelo de mercado [Fama *et al.* (1969)] o del rendimiento medio en un periodo no contaminado por el suceso estudiado para la obtención del rendimiento esperado en el periodo del suceso [Masulis (1980)] han sido aproximaciones ampliamente empleadas en la literatura de sucesos<sup>10</sup>. En ambos casos, los rendimientos anormales generados dependen de las características de los propios títulos<sup>11</sup>.

En el caso del estudio del comportamiento a largo plazo de las OPV<sup>12</sup>, y especialmente de las OPI, estas aproximaciones para el cálculo de los rendimientos esperados no son viables, ya que la selección de un periodo normal (no contaminado) que permita la esti-

---

<sup>7</sup> En este sentido Lipsey (1971, pág. 14) afirma que «es importante recordar que cualquier teoría es una abstracción de la realidad. Si no utilizáramos la abstracción, lo único que haríamos sería duplicar la realidad, como si se tratase de una cámara fotográfica, no aumentando nuestra comprensión del mundo».

<sup>8</sup> En adelante, cociente VC/VM.

<sup>9</sup> Fama y French (1992) confirman la existencia de una relación negativa y significativa entre rentabilidad y tamaño y una relación positiva y significativa entre rentabilidad y el cociente VC/VM.

<sup>10</sup> Específicamente, el método desarrollado por Masulis (1980) se denomina *método del periodo de comparación*.

<sup>11</sup> Fama (1998) indica que estas aproximaciones pueden emplearse para estudiar la reacción de los precios ante un determinado suceso, pero son incapaces de identificar anomalías relacionadas con la sección cruzada de los rendimientos medios, como el efecto tamaño.

<sup>12</sup> Si se trata de una salida a bolsa, entonces la oferta recibe el nombre de oferta pública inicial (OPI). Si la OPV es realizada por una empresa ya cotizada, entonces la OPV recibe el nombre de oferta pública subsiguiente (OPS).

mación de los parámetros del modelo de mercado o que sirva como periodo de comparación es problemático. Por ello, en este tipo de estudios suele hacerse uso de aproximaciones que restringen la sección cruzada de los rendimientos esperados, aunque esto supone la descripción imperfecta de los rendimientos medios. El método más ampliamente empleado es el de emparejar los títulos de la muestra con una empresa o cartera no afectada por el suceso que presente unas características similares, características de las que se tenga constancia que están relacionadas con los rendimientos medios (*matching approach*)<sup>13</sup>. Así, es habitual emparejar las empresas de la muestra con empresas o carteras no afectadas por el suceso según su tamaño y cociente VC/VM dada la evidencia existente que ya hemos comentado. No obstante, y tal y como señala Fama (1998), es posible que ni siquiera el tamaño y cociente VC/VM sean capaces de describir por completo la variabilidad de la sección cruzada de los rendimientos medios. A este respecto, la evidencia en el mercado español no es concluyente. Así, si bien Rubio (1986) y Gómez y Marhuenda (1998) detectan una relación negativa y significativa entre tamaño y rendimiento, García-Ayuso y Rueda (1998) hallan que el cociente VC/VM no es significativo en la explicación de los rendimientos, mientras que Cuéllar y Lainez (1999) sí encuentran una relación significativa pero de signo contrario al de otros mercados. Más recientemente, Nieto (2001) obtiene resultados contradictorios respecto de los factores tamaño y cociente VC/VM: mientras que de los análisis en serie temporal se desprende que estos factores podrían considerarse como buenas aproximaciones de factores de riesgo comunes explicativos del rendimiento de los activos en el contexto del modelo de tres factores de Fama y French (1993), cuando analiza la capacidad de explicación de las betas de los tres factores en sección cruzada obtiene que ninguna de ellas se muestra como relevante.

Otro método de estimación de los rendimientos anormales es mediante el empleo de un modelo de valoración de activos. Los trabajos pioneros que emplearon rendimientos a largo plazo [Jaffe (1974), Mandelker (1974), Asquith (1983)] hicieron uso del CAPM. Sin embargo, las deficiencias detectadas en el modelo condujeron a que se considerara la necesidad de incluir más factores de riesgo sistemático que explicaran las variaciones de los rendimientos. De este modo, los trabajos más recientes emplean para la generación del rendimiento anormal el modelo de valoración por arbitraje (APT) desarrollado por Ross (1976) [Eckbo *et al.* (2000)], el modelo de tres factores de Fama y French (1993) [Loughran y Ritter (1995), Brav *et al.* (2000), Mitchell y Stafford (2000), Boehme y Sorescu (2002)] o la extensión que de éste hace Carhart (1997) [Brav *et al.* (2000)].

Existe un cierto debate en la literatura respecto de la conveniencia de emplear modelos de factores frente a modelos basados en el emparejamiento según las características de los títulos. Así, Daniel y Titman (1997) sugieren que la sección cruzada de los rendimientos de los títulos está más estrechamente relacionada con las características de los títulos que los factores de riesgo del modelo de Fama y French (1993). Por el contrario, Davis *et al.* (2000) discuten estas conclusiones tras analizar la sección cruzada de los rendimientos en un periodo de tiempo mucho mayor que el analizado por Daniel y Titman (1997)<sup>14</sup>.

---

<sup>13</sup> Existe evidencia empírica que sugiere que este procedimiento proporciona mejores previsiones *ex-ante* de la sección cruzada de los rendimientos futuros [véase Daniel y Titman (1997)].

<sup>14</sup> Davis *et al.* (2000) analizan un periodo muestral que se extiende desde julio de 1929 hasta junio de 1997 (68 años), mientras que Daniel y Titman (1997) analizan un periodo que abarca desde julio de 1973 hasta diciembre de 1993 (20,5 años).

Aunque la utilización de varios factores de riesgo parece coherente con la evidencia existente respecto de la sección cruzada de los rendimientos medios, ninguno de los anteriores modelos es capaz de ofrecer una explicación completa de los mismos. Como muestra Nieto (2001), el carácter estático de modelos como el CAPM tradicional o el modelo de Fama y French (1993) supone una restricción insalvable. Tanto es así que la autora afirma que «parece evidente que la extensa utilización de modelos incondicionales [en trabajos de impacto de sucesos, evaluación de fondos de inversión y otros] es ciertamente cuestionable»<sup>15</sup>. Por el contrario, los resultados que obtiene para un CAPM condicional basado en el modelo propuesto por Jagannathan y Wang (1996) y el modelo intertemporal de Campbell (1993) son «algo más alentadores».

En resumen, y como afirma Fama (1998), los problemas derivados de la mala especificación del modelo son (por el momento) inevitables, siendo éstos más serios cuando se utilizan en contrastes de rendimientos a largo plazo.

### 3. PROBLEMAS RELACIONADOS CON LA MÉTRICA Y EL CONTRASTE DEL COMPORTAMIENTO A LARGO PLAZO

En este apartado tratamos el problema de cómo medir y contrastar el comportamiento a largo plazo de un título o cartera. Como hemos comentado anteriormente, puede encontrarse en la literatura tres aproximaciones para ello, a saber: la composición (multiplicación) de los rendimientos a corto plazo sobre el horizonte objeto de análisis, la acumulación de los rendimientos a corto plazo, y la utilización del método de las carteras de fecha de calendario. Abordamos cada uno de ellos de forma separada, ya que presentan problemas conceptuales y estadísticos propios.

#### 3.1. Métrica y contraste mediante el método de los rendimientos compuestos

Aunque ya el artículo seminal de Ibbotson (1975) detecta rendimientos anormales a largo plazo posteriormente a la realización de una OPI, no es hasta el trabajo de Ritter (1991) cuando se suscita un verdadero interés en la literatura por este tipo de fenómenos. Tomando, pues, como punto de partida este trabajo, el modo más popular empleado para el cálculo del rendimiento a largo plazo ha sido el de componer los rendimientos a corto plazo (generalmente mensuales) para, de esta forma, obtener el rendimiento correspondiente al horizonte temporal o ventana que se desea estudiar. Los rendimientos calculados de esta forma se corresponderían con una estrategia consistente en *comprar y mantener* durante dicho horizonte, denominándose en la terminología anglosajona como *buy-and-hold returns* (BHR). Así, el rendimiento de la empresa  $i$  durante el periodo de inversión  $\tau$  se calcularía como sigue:

$$\text{BHR}_{i\tau} = \left[ \prod_{t=s}^{s+\tau} (1 + R_{it}) \right] - 1, \quad [1]$$

donde  $s$  es el mes del suceso y  $R_{it}$  es el rendimiento de la empresa  $i$  en el mes  $t$ .

<sup>15</sup> Nieto (2001), pág. 38.

A partir de su cálculo para las empresas de la muestra y de las referencias seleccionadas contra las que se compara el rendimiento de la muestra (índice de mercado o sectorial, empresa de control, etc.), se obtiene el rendimiento anormal (BHAR), esto es:

$$\text{BHAR}_{i\tau} = \text{BHR}_{i\tau} - \text{BHR}_{\text{CONTROL}, \tau}, \quad [2]$$

siendo la media muestral en sección cruzada de los  $\text{BHAR}_{i\tau}$  individuales ( $\overline{\text{BHAR}}_{\tau}$ ) el estimador empleado para medir el comportamiento de las empresas de la muestra:

$$\overline{\text{BHAR}}_{\tau} = \sum_{i=1}^N w_i \cdot \text{BHAR}_{i\tau}, \quad [3]$$

donde  $N$  es el número de sucesos en la muestra y  $w_i$  es el peso asignado a la empresa  $i$ . La práctica de emplear como peso el valor de mercado de la empresa en el momento del suceso presenta el inconveniente de sobreponderar con mayor intensidad las observaciones recientes frente a las iniciales. Para evitar este problema, común, por otro lado, cuando se realiza un estudio respecto de la fecha del suceso, se estandariza, antes del cálculo de los pesos en el cómputo de los rendimientos ponderados, el valor de mercado de la empresa correspondiente dividiéndolo por el nivel del índice de mercado [Mitchell y Stafford (2000)]. La hipótesis nula a contrastar consiste en que el rendimiento anormal medio en sección cruzada sea igual a cero para la muestra de  $N$  empresas en el horizonte  $\tau$ , siendo la hipótesis alternativa que dicha media sea distinta de cero, esto es:

$$H_0: \overline{\text{BHAR}}_{\tau} = 0, \quad \text{versus} \quad H_1: \overline{\text{BHAR}}_{\tau} \neq 0.$$

Alternativamente, son varios los trabajos [Ritter (1991), Loughran y Ritter (1995), Brav y Gompers (1997), Brav *et al.* (2000), entre otros] que expresan, siguiendo la metodología original de Ritter (1991), el comportamiento a largo plazo de las empresas de la muestra respecto de diversos controles mediante el cociente del rendimiento a largo plazo de éstos, denominándose en este caso el estimador como *riqueza relativa* ( $\text{RR}_{\tau}$ ):

$$-\text{RR}_{\tau} = \frac{1 + \overline{\text{BHR}}_{\text{MUESTRA}, \tau}}{1 + \overline{\text{BHR}}_{\text{CONTROL}, \tau}}, \quad [4]$$

de tal forma que una riqueza relativa inferior a la unidad puede interpretarse como que la muestra tiene un peor comportamiento que el respectivo control. El principal inconveniente de esta medida así definida radica en el hecho de que no permite la realización de contrastes estadísticos de hipótesis.

La inferencia estadística del rendimiento anormal a largo plazo en los primeros trabajos está condicionada por sostenerse en hipótesis que posteriormente fueron cuestionadas y por el escaso conocimiento de las propiedades de los BHR y de los contrastes estadísticos asociados<sup>16</sup>, lo cual condujo a que en estos trabajos se evitara la realización de inferencias estadísticas formales de los BHAR [Ritter (1991), Levis (1993), Loughran *et al.* (1994), Spiess y Affleck-Graves (1995), Loughran y Ritter (1995) o Ljungqvist (1997)].

<sup>16</sup> Kothari y Warner (1997) constatan el hecho de que pese a que los BHR son a menudo empleados en los estudios que se refieren al largo plazo, sus propiedades y contrastes asociados no habían sido estudiados en la literatura. No obstante, existían indicios de asimetría de los BHAR individuales documentados por Ibbotson (1975).



La primera propuesta dirigida a superar los problemas relacionados con la imposibilidad de aplicar los contrastes habituales en el caso de los BHAR es la realizada por Ikenberry *et al.* (1995) al introducir el procedimiento del *bootstrapping* en la realización de inferencias estadísticas. Mediante esta técnica, los autores generan la distribución empírica de los BHAR bajo la hipótesis nula, relajando las hipótesis de normalidad, estacionariedad e independencia temporal de las observaciones.

Posteriormente, Barber y Lyon (1997) y Kothari y Warner (1997) pusieron en cuestión los diferentes aspectos de la medida del comportamiento a largo plazo, llegando ambos trabajos a la conclusión de que los métodos habitualmente empleados para el cálculo de los rendimientos anormales a largo plazo [composición (BHR) y acumulación (CAR) de los rendimientos a corto plazo] son conceptualmente erróneos y/o conducen a contrastes mal especificados. Las simulaciones realizadas sugieren que estos métodos tienden a encontrar un comportamiento anormal positivo o negativo cuando éste no existe. Barber y Lyon (1997) identifican tres sesgos que afectan a la especificación de los contrastes estadísticos cuando se emplea como referencia carteras tales como un índice de mercado o carteras construidas por tamaño, sesgos que aparecen como consecuencia del listado de nuevos títulos, del reajuste de la cartera de referencia y de la asimetría de los rendimientos anormales multiperiodales. Una cuestión adicional tiene que ver con el poder de los contrastes en relación con la contaminación de las referencias con algunas de las empresas que se desea analizar.

Aunque estos sesgos afectan tanto a las estimaciones realizadas mediante la composición de los rendimientos como a través de su suma, Barber y Lyon (1997) prefieren la estimación del rendimiento a largo plazo mediante BHAR por dos motivos: por un lado, porque, como muestran, el rendimiento anormal acumulado es un estimador sesgado de los BHAR; y por otro lado, porque, aunque la inferencia realizada mediante rendimientos anormales acumulados fuera correcta, los BHAR «miden con precisión la experiencia del inversor».

El sesgo debido al listado de nuevos títulos se produce cuando la cartera que sirve de referencia incorpora empresas que se han incorporado al mercado con posterioridad a la fecha del suceso, mientras que, por definición, las empresas que conforman la muestra se negocian desde el día del suceso. Dado que la evidencia empírica muestra que las nuevas empresas que entran a cotizar presentan rendimientos a largo plazo inferiores a otros activos [Ritter (1991)], la utilización de carteras contaminadas por este tipo de empresas sesgaría al alza el rendimiento anormal de la muestra.

El sesgo debido al reajuste se deriva de la práctica de emplear un índice de mercado o una cartera de referencia equiponderada, ya que al componer los rendimientos se asume reajuste de los rendimientos periodales (mensuales) de todos los títulos que conforman el índice o la cartera<sup>17</sup>. Con el objeto de mantener las ponderaciones, los títulos que batan al mercado son vendidos mientras que los que son superados por éste son comprados. Este reajuste conduce a un sesgo en la media poblacional de los rendimientos anormales a largo plazo si los rendimientos mensuales consecutivos de los activos individuales están correlacionados. En concreto, se produce una sobreestimación del rendimiento a largo plazo de la cartera de referencia, lo que conduce a la obtención de rendimientos anormales sesgados a la baja<sup>18</sup>.

---

<sup>17</sup> Canina *et al.* (1998) documentan que la magnitud de este sesgo es más intensa cuando se emplean rendimientos diarios en lugar de mensuales.

<sup>18</sup> Para más detalles, véanse Barber y Lyon (1997) y Canina *et al.* (1998).

Por último, el sesgo que genera la asimetría de los rendimientos anormales a largo plazo es consecuencia de la observación empírica de que los activos individuales presentan observaciones de BHR extremas positivas mayores que observaciones extremas negativas, lo que conduce a una substancial asimetría positiva. Por su parte, los rendimientos de las carteras de referencia, al tratarse de medias, son menos asimétricos. Por consiguiente, los rendimientos anormales presentan asimetría positiva. Esta asimetría positiva se traduce en un sesgo negativo de los contrastes como consecuencia de la correlación positiva entre la media y la desviación típica de la muestra en distribuciones con asimetría positiva<sup>19</sup>.

Las propuestas de corrección de estos problemas han sido varias. Por lo que respecta al bajo poder del método para detectar un comportamiento anormal como consecuencia de la contaminación de la referencia, la solución habitual en la literatura es purgar las carteras de referencia, eliminando de su composición durante los  $\tau$  meses posteriores a la fecha del suceso (ventana de estudio) las empresas afectadas por el suceso que se pretenden estudiar. Por otra parte, Barber y Lyon (1997) señalan que la utilización de una empresa de control de tamaño y cociente VC/VM similar al de la empresa de la muestra produce contrastes bien especificados. Además de con este método, Lyon *et al.* (1999) obtienen contrastes bien especificados en muestras aleatorias mediante la construcción *cuidadosa* de las carteras y la realización de inferencias ajustando el estadístico  $t$  por asimetría vía *bootstrapping* y, por otro lado, mediante la generación empírica de la distribución de los BHAR de forma análoga a Ikenberry *et al.* (1995), entre otros.

En concreto, Lyon *et al.* (1999) sostienen que el método comúnmente empleado en la literatura para el cálculo del rendimiento a largo plazo de una cartera de control, esto es, la composición del rendimiento medio mensual de la cartera, no refleja de forma adecuada los rendimientos obtenidos en una estrategia pasiva de *comprar y mantener* consistente en invertir en los títulos que conforman la cartera. En la expresión [5] se recoge esta forma de cálculo.

$$\text{BHR}_{p\tau}^{\text{re}} = \left[ \prod_{t=s}^{s+\tau} (1 + R_{pt}) \right] - 1, \quad [5]$$

donde  $\text{BHR}_{p\tau}^{\text{re}}$  es el rendimiento de la cartera  $p$  sobre un horizonte temporal de  $\tau$  meses que comienza en el mes  $s$ , y  $R_{pt}$  es el rendimiento de la cartera  $p$  en el mes  $t$ . Esta forma de cálculo es inadecuada porque el rendimiento de esta cartera asume reajuste mensual para mantener las ponderaciones de los títulos que la componen<sup>20</sup>; y, por otro lado, porque incluye las empresas listadas subsiguientemente a la formación de la cartera en el periodo  $s$  que pudieran presentar rendimientos anormales negativos de acuerdo con la evidencia existente [Ritter (1991)].

Por el contrario, en el procedimiento propuesto por Lyon *et al.* (1999) se compone, en primer lugar, el rendimiento de los títulos que constituyen la cartera  $p$  en  $s$  y, a continuación, se calcula su media:

$$\text{BHR}_{p\tau}^{\text{cm}} = \sum_{i=1}^{n_s} w_i \left( \left[ \prod_{t=s}^{s+\tau} (1 + R_{it}) \right] - 1 \right), \quad [6]$$

<sup>19</sup> Para más detalles, véase Barber y Lyon (1997).

<sup>20</sup> El superíndice «re» hace referencia a este hecho.

donde  $n_s$  es el número de empresas correspondientes a la cartera  $p$  que se negocian en el mes  $s$ , y  $w_i$  es la ponderación de la empresa  $i$  de acuerdo con su valor de mercado. El rendimiento así calculado representa una estrategia de inversión pasiva consistente en invertir de forma ponderada en todos los títulos que componen la cartera de referencia en el mes  $s$ , esto es, representa una verdadera estrategia de *comprar y mantener*<sup>21</sup>.

De esta forma, se evita invertir en títulos que se incorporan al mercado tras el periodo  $s$ , y no se produce reajuste mensual de la cartera. Si una empresa de las que componen la cartera en  $s$  sale del mercado tras este mes y antes de  $s+\tau$ , es factible suponer que el inversor reinvierte el resultado de su inversión en esta empresa en un índice representativo del mercado [Cowan y Sergeant (2001)]. Esto significa que los rendimientos mensuales ausentes desde la fecha de la salida del mercado hasta  $s+\tau$  se rellenan con los rendimientos mensuales del índice de mercado correspondientes.

Por lo que respecta al posible sesgo por asimetría, si bien Barber y Lyon (1997) muestran que la utilización de la empresa de control anula este sesgo, Lyon *et al.* (1999) sostienen que para eliminarlo cuando se calculan los rendimientos anormales a largo plazo utilizando carteras de control es necesario el empleo del siguiente estadístico  $t$  ajustado por asimetría:

$$t_{asim} = \sqrt{N} \left( S + \frac{1}{3} \hat{\gamma} \cdot S^2 + \frac{1}{6N} \hat{\gamma} \right), \quad [7]$$

donde  $S = \frac{\overline{\text{BHAR}}_{\tau}}{\sigma(\text{BHAR}_{\tau})}$ ,  $\hat{\gamma} = \frac{\sum_{i=1}^N (\text{BHAR}_{i\tau} - \overline{\text{BHAR}}_{\tau})^3}{N\sigma^3(\text{BHAR})}$  y  $\sigma(\text{BHAR}_{\tau})$  es la desviación típica

en sección cruzada de los rendimientos anormales de la muestra.

Obsérvese que  $\hat{\gamma}$  es un estimador del coeficiente de asimetría y  $\sqrt{N} \cdot S$  es el estadístico  $t$  convencional recogido en la expresión [8]:

$$t = \frac{\overline{\text{BHAR}}_{\tau}}{\sigma(\text{BHAR}_{\tau})/\sqrt{N}}. \quad [8]$$

De acuerdo con Lyon *et al.* (1999), únicamente la aplicación *bootstrapped* del estadístico  $t$  ajustado por asimetría produce contrastes bien especificados. El proceso *bootstrap* consiste en extraer aleatoriamente  $b$  submuestras de tamaño  $N_b$  de la muestra original; calcular el estadístico ajustado por asimetría en cada una de las  $b$  submuestras aleatorias, y calcular los valores críticos del estadístico transformado a partir de los valores obtenidos del mismo en el proceso de remuestreo.

Fama (1998), no obstante, se manifiesta abiertamente en contra de la metodología BHAR por dos razones. Por un lado, porque los errores que se producen como consecuencia de la mala especificación del modelo generador de rendimientos esperados se componen con el cálculo de los rendimientos a largo plazo y, por otro lado, porque esta metodología ignora la dependencia en sección cruzada que se produce entre los rendimientos anormales de las empresas de la muestra que se solapan en el tiempo, tanto por

<sup>21</sup> Por este motivo denotamos el cálculo del rendimiento con el superíndice «cm».

lo que se refiere al solapamiento de los rendimientos de una misma empresa que está incluida en la muestra en diferentes momentos del tiempo como del solapamiento que se produce entre empresas similares, como aquéllas que pertenecen al mismo sector. Brav (2000) argumenta, en este sentido, que la dependencia en sección cruzada de las observaciones de la muestra afecta a la correcta especificación de los contrastes estadísticos.

Como señalan Mitchell y Stafford (2000), las decisiones empresariales relevantes (ofertas de adquisición, ampliaciones de capital u ofertas públicas de venta) no son sucesos aleatorios y, por tanto, no pueden tomarse como observaciones independientes. Dado que existe evidencia de que las decisiones empresariales se concentran por sectores y en el tiempo, la correlación de los residuos implica un serio problema para la metodología BHAR, ya que ésta supone la independencia de todas las observaciones. Estos autores muestran cómo la existencia de correlación en sección cruzada de los BHAR tiene un efecto trascendental sobre la credibilidad de las inferencias realizadas mediante la metodología tradicional, ya que al realizar una corrección aproximada de los estadísticos  $t$  para tres muestras distintas de sucesos (fusiones, ofertas públicas subsiguientes y recompras de acciones) obtienen que los estadísticos  $t$  pasan de  $-6,05$  (OPS) y  $4,86$  (recompras) a  $-1,49$  y  $1,91$ , respectivamente, indicando la inexistencia de rendimientos anormales estadísticamente significativos a largo plazo<sup>22</sup>.

Aunque buena parte de los sesgos que afectan a la estimación de  $\overline{\text{BHAR}}$  como consecuencia de las pobres propiedades estadísticas de los BHAR individuales [Barber y Lyon (1997), Kothari y Warner (1997) y Lyon *et al.* (1999)] pueden mitigarse con el empleo de muestras grandes y la cuidadosa construcción de las carteras de referencia, los problemas asociados con la correlación en sección cruzada de los BHAR en muestras no aleatorias son de difícil corrección<sup>23</sup>. En este sentido, Lyon *et al.* (1999), Mitchell y Stafford (2000) y Brav (2000) presentan metodologías que, aunque la reducen, no son capaces de eliminar la mala especificación de los contrastes.

Dado que los anteriores estadísticos presentados asumen independencia en sección cruzada de las observaciones de las muestras, si existe algún tipo de dependencia, éstos estarán mal especificados. Así, considérese la expresión del estadístico  $t$  convencional recogido en [8] en notación matricial:

$$t = \overline{\text{BHAR}}_t \left[ \frac{\mathbf{I}'\hat{\Sigma}\mathbf{I}}{N} \right]^{-1/2}, \quad [9]$$

donde  $\mathbf{I}$  es un vector  $N \times 1$  de unos. Si suponemos que las observaciones de la muestra son independientes en sección cruzada, entonces la matriz de varianzas-covarianzas de los rendimientos anormales ( $\Sigma$ ) es diagonal. En este caso, la desviación típica en sección cruzada de las observaciones de la muestra [ $\sigma(\text{BHAR}_t)$ ] es la estimación de los  $N$  elementos de la diagonal principal de la matriz  $\Sigma$ . Si, por el contrario, suponemos la existen-

<sup>22</sup> Las muestras de OPS y recompras consisten en 4.911 y 2.421 sucesos, respectivamente, ambas para el periodo 1958-1993. Para mayores detalles en relación con la incidencia de la correlación en sección cruzada, véase la tabla 6 (pág. 307) de su trabajo.

<sup>23</sup> Nos referimos al solapamiento que se produce entre empresas distintas, ya que, como señalan Lyon *et al.* (1999), cuando el solapamiento se da entre sucesos de una misma empresa la única solución posible es su eliminación de la muestra.

cia de algún tipo de dependencia en sección cruzada, entonces la matriz de varianzas-covarianzas de los rendimientos anormales deja de ser diagonal. En este caso, Lyon *et al.* (1999) estiman el elemento  $\sigma_{ij}$  de la matriz de varianzas-covarianzas de los rendimientos anormales como la covarianza entre los rendimientos anormales de las empresas  $i$  y  $j$  en el espacio de tiempo (de calendario) en el que están solapados<sup>24</sup>.

Aunque algunos autores han abogado por el método del *bootstrapping*, para la resolución del problema del problema de dependencia, éste tampoco es capaz de resolverlo. Como indican Mitchell y Stafford (2000), ello se debe a que la generación de una distribución empírica mediante submuestras aleatorias no es capaz de capturar la estructura de la matriz de varianzas-covarianzas de la muestra de empresas.

No obstante los problemas que hemos señalado asociados a la estimación del mal comportamiento a largo plazo mediante la estrategia de *comprar y mantener*, se continúan realizando esfuerzos en la literatura en favor de su superación. Así, Cowan y Sergeant (2001) estudian cómo los diferentes tipos de sesgos interactúan en la especificación de los contrastes en función del tamaño de la muestra y la amplitud de la ventana empleada. Como resultado de sus simulaciones, proponen para la mejora de la calidad de las inferencias la *winsorización* de las observaciones a tres veces su desviación típica<sup>25</sup>, el emparejamiento de las empresas de la muestra con carteras de referencia según las características tamaño y cociente VC/VM y el empleo de un contraste de diferencia de medias de dos grupos como el que se recoge en la expresión [10]:

$$Z = \text{BHAR}_\tau \frac{\overline{\text{BHAR}}}{\sqrt{\frac{\hat{\sigma}_{\text{MUESTRA}}^2}{N} + \frac{\hat{\sigma}_{\text{CONTROL}}^2}{N}}} \quad [10]$$

Esta sencilla metodología produce contrastes paramétricos mejor especificados y a menudo más potentes que los contrastes propuestos en la literatura anteriormente [Barber y Lyon (1997) y Lyon *et al.* (1999)]. Los autores reconocen, sin embargo, que el procedimiento propuesto no está perfectamente especificado en todos los casos y que el proceso de *winsorización* puede no ser del agrado de algunos investigadores al implicar la manipulación de algunas observaciones, aunque tan sólo sea para el cálculo del estadístico. Por el contrario, aducen en su defensa que permite, de una forma razonable, examinar la robustez de los resultados.

Jakobsen y Voetmann (1999), por otro lado, desarrollan una nueva aproximación de la riqueza relativa en la que emplean las propiedades de la distribución log-normal de ésta para la realización de contrastes estadísticos de hipótesis, ya que tal y como habi-

<sup>24</sup> Así, supóngase que la empresa  $i$  tiene un rendimiento anormal a largo plazo calculado desde el periodo  $s$  hasta  $s + \tau$  y la empresa  $j$  desde el periodo  $s + a$  hasta  $s + \tau + a$ . Si  $a < \tau$ , entonces el elemento  $\sigma_{ij}$  de la matriz de varianzas-covarianzas de rendimientos anormales se estima como sigue:

$$\frac{1}{\tau - a - 1} \sum_{t=s+a}^{s+\tau} (AR_{it} - \overline{AR}_i)(AR_{jt} - \overline{AR}_j),$$

donde  $AR_{it}$  es el rendimiento anormal de la empresa  $i$  en el mes natural  $t$ . Si  $a \geq \tau$ , entonces se supone que la covarianza es cero.

<sup>25</sup> Es decir, si el valor (absoluto) de la observación excede tres veces la desviación típica en sección cruzada de las observaciones de la muestra, entonces se le asigna el valor límite (con el signo correspondiente).

tualmente se había estado calculando desde Ritter (1991) no permitía la realización de contrastes estadísticos<sup>26</sup>. En concreto, estos autores no toman la media de los rendimientos compuestos en el numerador y denominador, tal y como se ha definido en [4], sino que calculan uno para cada empresa, esto es:

$$RR_{i\tau} = \frac{\prod_{t=s}^{s+\tau} (1 + R_{it})}{\prod_{t=s}^{s+\tau} (1 + R_{CONTROL,t})} \quad [11]$$

A través de la transformación logarítmica de la expresión [11], Jakobsen y Voetmann (1999) logran descomponer el valor medio esperado de la riqueza relativa en niveles en dos componentes (una media transformada y un componente de volatilidad), tal y como se muestra en la expresión [12].

$$E(RR)_{\tau} = e^{\mu_{\tau} \cdot \tau} = e^{\alpha_{\tau} \cdot \tau + \frac{1}{2} \sigma_{\tau}^2 \cdot \tau} = \underbrace{e^{\alpha_{\tau} \cdot \tau}}_{\substack{\text{Media} \\ \text{Transformada}}} \cdot \underbrace{e^{\frac{1}{2} \sigma_{\tau}^2 \cdot \tau}}_{\text{Volatilidad}} \quad [12]$$

Como puede observarse en la expresión [14] el componente de volatilidad introduce un sesgo al alza en el valor medio esperado de la riqueza relativa, lo cual hace imprescindible ajustar la media esperada en sección cruzada si se desea inferir correctamente el comportamiento a largo plazo. A partir de lo anterior, Jakobsen y Voetmann (1999) establecen que la riqueza relativa definida por ellos tiene una distribución log-normal, de lo que deducen que ésta sigue una distribución normal:

$$\text{Log}(RR_{i-\text{CONTROL}, \tau}) \sim N(\alpha_{\tau} \cdot \tau, \sigma_{\tau}^2 \cdot \tau) \quad [13]$$

Según Jakobsen y Voetmann (1999), una vez ajustada por la volatilidad, la media transformada es un estimador contrastable y correcto del comportamiento a largo plazo del título. Por tanto, la hipótesis nula a contrastar en este caso vendría dada porque la media transformada en sección cruzada es igual a cero en el horizonte  $\tau$ , siendo la hipótesis alternativa que ésta sea distinta de cero, esto es:

$$H_0: \alpha_{\tau} \cdot \tau = 0, \quad \text{versus} \quad H_1: \alpha_{\tau} \cdot \tau \neq 0.$$

La estimación de los parámetros correspondientes a la media transformada y la varianza puede realizarse de la siguiente forma:

$$\hat{\alpha}_{\tau} = \frac{1}{\tau \cdot N} \sum_{i=1}^N \text{Log}(RR_{i-\text{CONTROL}, \tau}), \quad [14]$$

$$\hat{\sigma}_{\tau}^2 = \frac{1}{\tau \cdot (N - 1)} \sum_{i=1}^N [\text{Log}(RR_{i-\text{CONTROL}, \tau}) - \hat{\alpha}_{\tau} \cdot \tau]^2 \quad [15]$$

Los estimadores  $\hat{\alpha}_{\tau}$  y  $\hat{\sigma}_{\tau}^2$  son estimadores marginales en cada momento del tiempo y están basados únicamente en la riqueza relativa esperada en sección cruzada en el mo-

<sup>26</sup> Véase también Jakobsen y Sørensen (2001).

mento  $\tau$ . De esta forma, estos estimadores pueden tratarse independientemente de estimadores asociados a otros horizontes temporales, por lo que están libres de correlación serial u otro tipo de dependencias en serie temporal. De acuerdo con Jakobsen y Voetmann (1999),  $\alpha_\tau \cdot \tau$  se distribuye en el momento  $\tau$  como una  $t$  de Student con  $N-1$  grados de libertad.

Haciendo uso de esta metodología, los autores estudian el comportamiento a largo plazo de una muestra de 142 OPI y 413 OPS realizadas en la Bolsa de Copenhague en el periodo 1983-1998. Sus resultados empleando la metodología tradicional indican que en un horizonte de 5 años las OPI y OPS obtienen un rendimiento 27,3% y 21,4% inferior al del mercado. Sin embargo, tras ajustar por el componente de volatilidad, el mal comportamiento se cifra en un 43,7% y un 38,1%, respectivamente.

Pese a los diversos problemas que plantea la metodología de los rendimientos compuestos que hemos sintetizado anteriormente, ésta ha sido ampliamente utilizada en la literatura. Una de las ventajas que presenta frente a su alternativa tradicional (la acumulación de los rendimientos) es que se muestra neutral por lo que respecta al sesgo al alza de los rendimientos de los títulos individuales calculados con precios de cierre, sesgo debido fundamentalmente al efecto *bid-ask* [Blume y Stambaugh (1983)]. Conrad y Kaul (1993) demuestran que con el método de composición de los rendimientos este sesgo permanece constante a lo largo del horizonte de estudio, mientras que, por el contrario, se acumula cuando el método empleado es el de acumulación, incrementándose, por tanto, a medida que se amplía la ventana de análisis.

Por otra parte, Barber y Lyon (1997) y Lyon *et al.* (1999) defienden la importancia de los BHAR argumentando que «miden de forma precisa la experiencia del inversor». Sin embargo, esta razón no puede ser suficiente para concentrar toda la atención en esta metodología. Evidentemente, la estrategia de *comprar y mantener* captura la experiencia del inversor cuando compra un título y lo mantiene en cartera durante tres o cinco años. Pero ésta no es el único tipo de experiencia posible. La obtención sistemática de rendimientos mensuales anormales por parte de las empresas de la muestra también es una cuestión relacionada, por ejemplo. Por otra parte, la composición de los rendimientos provoca que el comportamiento anormal crezca con el horizonte de estudio, aun cuando éste sólo se hubiera producido en una pequeña fracción del mismo<sup>27</sup>. Este hecho, aunque pueda parecer trivial, no lo es en absoluto, ya que la elección del horizonte de estudio es totalmente arbitraria, analizándose, generalmente, varios horizontes de diferente amplitud con el objeto de investigar hasta cuándo se extiende tras el suceso el comportamiento anormal. Por último, y mucho más importante, no puede obviarse los serios problemas estadísticos comentados. Estamos de acuerdo con Mitchell y Stafford (2000) cuando apuntan que «dado que nuestro objetivo es medir de forma fiable los rendimientos anormales, es imperativo que la metodología permita inferencias estadísticas fiables»<sup>28</sup>.

<sup>27</sup> Así, supóngase que un título genera mensualmente un rendimiento del 1.5%, al igual que una determinada referencia que se emplea de control, y que el horizonte de estudio es de 5 años (60 meses). Si el título presentara un mal comportamiento durante los primeros seis meses del horizonte, generando un rendimiento nulo, el BHAR para una ventana de 12 meses sería de -10,22%, mientras que al cabo del horizonte sería del -20,88%.

<sup>28</sup> Mitchell y Stafford (2000, pág. 296).

### 3.2. Métrica y contraste mediante el método de los rendimientos acumulados

Además de la utilización del rendimiento medio en sección cruzada de los BHAR individuales, la mayor parte de los trabajos iniciales emplea la adición de los rendimientos anormales calculados en el corto plazo (diarios o mensuales), metodología que en la literatura financiera se denomina como *rendimientos anormales acumulados* (CAR).

Aunque el contraste del mal comportamiento a largo plazo mediante CAR puede verse potencialmente afectado por los mismos tres sesgos identificados por Barber y Lyon (1997), comentados anteriormente, Lyon *et al.* (1999) reconocen que los elaborados métodos desarrollados para los BHAR aplicados a los CAR no producen inferencias más correctas que el estadístico *t* convencional. Ello es debido a que los CAR presentan una menor asimetría que la composición de rendimientos mensuales para la obtención de rendimientos a largo plazo, siempre y cuando las carteras de referencia se hayan construido evitando los sesgos inducidos comentados anteriormente. De esta forma, Lyon *et al.* (1999) calculan el CAR de la empresa *i* de la muestra en el periodo  $\tau$  como sigue:

$$CAR_{i\tau} = \sum_{t=s}^{s+\tau} \left[ R_{it} - \frac{1}{n_t^s} \sum_{j=1}^{n_t^s} R_{jt} \right], \quad [16]$$

donde  $R_{jt}$  es el rendimiento mensual para las  $j = 1, \dots, n_t^s$  empresas que componen la cartera de control de la empresa *i* de la muestra, y que se negocian tanto en *s* como en *t*. La hipótesis a contrastar sería, por tanto:

$$H_0: \overline{CAR}_\tau = 0, \quad \text{versus} \quad H_1: \overline{CAR}_\tau \neq 0.$$

Alternativamente, si las carteras han sido purgadas de nuevos listados y consideramos como despreciable el sesgo debido al reajuste de las carteras [Farinós *et al.* (2002)], entonces el rendimiento anormal acumulado en el periodo  $\tau$  correspondiente a la muestra ( $CAR_\tau$ ) puede calcularse acumulando el rendimiento anormal medio en sección cruzada en cada mes *t* tras el suceso ( $\overline{AR}_t$ ):

$$CAR_\tau = \sum_{t=1}^{\tau} \overline{AR}_t. \quad [17]$$

Asimismo, el rendimiento anormal medio en sección cruzada ( $\overline{AR}_t$ ) se calcula como sigue:

$$\overline{AR}_t = \sum_{i=1}^N w_i \cdot AR_{it}, \quad [18]$$

donde  $AR_{it}$  es el rendimiento anormal de la empresa muestral *i* en el mes *t* posterior al suceso, calculado como la diferencia entre el rendimiento de la empresa muestral y el rendimiento esperado correspondiente a una determinada referencia, y  $w_i$  es el peso asignado a la empresa *i*<sup>29</sup>.

<sup>29</sup> Si la ponderación se establece en función del valor de mercado, es aplicable aquí también la corrección propuesta para el método de los rendimientos compuestos.



La hipótesis de interés es determinar si el rendimiento anormal acumulado hasta el periodo  $\tau$  es significativamente distinto de cero:

$$H_0: CAR_\tau = 0, \quad \text{versus} \quad H_1: CAR_\tau \neq 0.$$

Aunque los rendimientos acumulados presentan una menor asimetría, se ven afectados, al igual que los BHAR, por el problema de la dependencia en sección cruzada. Espenlaub *et al.* (2000) proponen el empleo de el estadístico recogido en [19] que es a su vez una variante del procedimiento que Brown y Warner (1980) denominan como *Crude Dependence Adjustment test*, con el cual es posible corregir el problema de correlación en sección cruzada.

$$t = \frac{CAR_\tau}{\sqrt{\left( \tau \cdot \left( \sum_{t=1}^{\tau} \left( \overline{AR}_t - \frac{1}{\tau} \sum_{t=1}^{\tau} \overline{AR}_t \right)^2 \right) \right) / (\tau - 1)}} \quad [19]$$

Sin embargo, y a pesar del mejor y más amplio conocimiento que de los CAR se tenía, el sentimiento generalizado era que «los resultados obtenidos empleando los CAR deberían ser tomados como meramente descriptivos, ya que no representan una estrategia de inversión realista»<sup>30</sup>, en contraposición con los BHAR. Esta postura también es defendida por Barber y Lyon (1997) y Lyon *et al.* (1999), como ya hemos mencionado. En concreto, el apoyo de Barber y Lyon (1997) al empleo de los BHAR en la detección de rendimientos anormales a largo plazo, frente al uso de los CAR, se deriva del análisis que realizan de la relación entre BHAR y CAR, en el que obtienen que los CAR son un estimador sesgado de los BHAR.

### 3.3. Métrica y contraste mediante el método de las carteras de fecha de calendario

La tercera alternativa para la medición del comportamiento a largo plazo que consideramos es la denominada como *carteras de fecha de calendario*. Jaffe (1974) y Mandelker (1974) fueron los primeros en emplear esta aproximación, la cual es vehementemente defendida por Fama (1998) frente a la composición de los rendimientos (BHAR), e incluso al cálculo de rendimientos anormales acumulados (CAR).

Este método consiste en la construcción de una cartera compuesta cada mes natural por todas aquellas empresas que en los  $\tau$  meses anteriores hubieran protagonizado un determinado suceso, donde  $\tau$  hace referencia a la amplitud de la ventana u horizonte de estudio tras el suceso. Cada mes la cartera se reajusta, extrayendo de ella aquellas empresas que concluyen el periodo de análisis de  $\tau$  meses, e incorporando a la misma las empresas que hubieran realizado un suceso de estudio correspondiente en el mes natural anterior. Con la formación de esta cartera, la correlación en sección cruzada de los rendimientos individuales de las empresas del suceso es tenida en cuenta de forma automática en la varianza de la cartera en cada uno de los periodos.

<sup>30</sup> Ikenberry *et al.* (1995), pág. 188.

El comportamiento de la cartera de títulos afectados por el suceso se analiza en comparación con un modelo de valoración o con algún otro tipo de referencia, contrastándose si el rendimiento anormal medio periodal (mensual) de la cartera en la ventana de estudio es significativamente distinto de cero. Así, las dos variantes más comunes del método de las carteras de fecha de calendario en la literatura son (i) la regresión en serie temporal de una cartera compuesta por las empresas protagonistas del suceso con el modelo de tres factores desarrollado por Fama y French (1993); y (ii) la estimación del rendimiento anormal mensual promedio de la serie de rendimientos anormales mensuales respecto de alguna referencia o control ( $\overline{\text{MAR}}$ ).

En el primer caso, a partir de la formación de la cartera muestral en fecha de calendario, se calcula el exceso de rendimiento en cada mes natural. La serie de excesos de rendimientos se emplea para estimar la siguiente regresión en serie temporal del modelo de tres factores de Fama y French (1993):

$$R_{pt} - R_{ft} = \alpha_p + \beta_p (R_{mt} - R_{ft}) + s_p \text{SMB}_t + h_p \text{HML}_t + \varepsilon_{pt}, \quad [20]$$

donde  $R_{pt}$  es el rendimiento en el mes natural  $t$  de la cartera muestral,  $R_{ft}$  es el rendimiento en el mes  $t$  de las Letras del Tesoro,  $R_{mt}$  es el rendimiento mensual de un índice de mercado,  $\text{SMB}_t$  es la diferencia entre los rendimientos de carteras construidas con empresas pequeñas y grandes y  $\text{HML}_t$  es la diferencia entre los rendimientos de carteras formadas por empresas con altos y bajos cocientes VC/VM<sup>31</sup>. Si el comportamiento anormal de las empresas de la muestra es una mera manifestación de diferentes efectos tales como diferencias en riesgo, tamaño y cociente VC/VM, entonces la estimación de la constante del modelo de Fama y French (1993) en las diferentes regresiones no debería ser ni económica ni estadísticamente distinta de cero. Por consiguiente, dado el modelo, la estimación de la constante ( $\alpha_p$ ) permite contrastar la hipótesis nula de que el rendimiento anormal mensual medio de la cartera formada por empresas de la muestra es cero, indicando, por tanto, la ausencia de mal comportamiento<sup>32</sup>. Esto es:

$$H_0: \alpha_p = 0, \quad \text{versus} \quad H_1: \alpha_p \neq 0.$$

Si, por el contrario, el modelo de Fama y French (1993) no es capaz de describir correctamente los rendimientos esperados, entonces la constante refleja dos efectos combinados: el mal comportamiento de la muestra y la mala especificación del modelo. En este último caso, si la muestra presenta características que el modelo no puede valorar correctamente, entonces la hipótesis nula puede ser rechazada, indicando la existencia de comportamiento anormal a largo plazo, cuando no es así (error tipo I).

Loughran y Ritter (1995), Brav y Gompers (1997), Brav *et al.* (2000) y Mitchell y Stafford (2000) estudian el comportamiento a largo plazo de las OPV con la primera de las variantes. En todos los trabajos se hace referencia a la dificultad que introduce en el anterior contraste el problema de la mala especificación del modelo. Esta alusión es inevitable, ya que Fama y French (1993) muestran que su modelo de tres factores no es capaz de explicar completamente los rendimientos medios de carteras construidas según tamaño y cociente VC/VM, que son las dimensiones de los rendimientos medios para las

<sup>31</sup> La construcción de los factores SMB y HML se explica detalladamente en Fama y French (1992,1993).

<sup>32</sup> Esta medida tiene un papel análogo a la del *alfa de Jensen* en el contexto del CAPM [véase Jensen (1968)].

que el modelo está diseñado<sup>33</sup>. Este resultado sugiere que la hipótesis nula (constante igual a cero del modelo de regresión) puede ser problemática cuando la muestra presenta características que el modelo no es capaz de valorar correctamente. Es el caso de las OPI en el mercado norteamericano. El modelo de tres factores de Fama y French (1993) tiene problemas para valorar empresas pequeñas con cocientes VC/VM bajos [Fama y French (1993), Mitchell y Stafford (2000)]. Dado que la mayor parte de las empresas que realizan una OPI en el mercado norteamericano poseen las anteriores características, el mal comportamiento a largo plazo detectado puede argumentarse no como efecto del suceso *per se*, sino como un resultado espurio consecuencia de la mala especificación del modelo de valoración [Brav y Gompers (1997)].

Como respuesta a los problemas relacionados con la especificación del modelo de Fama y French (1993), Brav *et al.* (2000) y Eckbo *et al.* (2000) estudian el comportamiento a largo plazo de las emisiones de títulos haciendo uso de distintos modelos. Así, Brav *et al.* (2000) emplea una extensión del modelo de tres factores de Fama y French (1993) realizada por Carhart (1997), quien argumenta que un cuarto factor puede incrementar el poder explicativo del modelo de Fama y French (1993) para el caso de los rendimientos de los fondos de inversiones. El cuarto factor que introduce Carhart (1997), denominado *momentum*, se inspira en la observación empírica [Jegadeesh y Titman (1993)] de que las empresas que presentan elevados rendimientos un año, continúan teniéndolos al año siguiente. La extensión del modelo, no obstante, continúa presentando limitaciones para explicar los rendimientos de todas las carteras construidas según las características tamaño y cociente VC/VM. Cuando emplean este modelo en su muestra de OPI y OPS<sup>34</sup>, el mal comportamiento estadísticamente significativo detectado en los cinco años posteriores a la emisión que aparece con el modelo de Fama y French (1993) desaparece en el caso de las OPI, persistiendo, aunque en una menor cuantía y nivel de significación, en el caso de las OPS. El mal comportamiento de éstas últimas deja de ser estadísticamente significativo cuando purgan la construcción de los factores y reconstruyen el factor HML del modelo de Fama y French (1993) excluyendo las empresas del NYSE<sup>35</sup>.

Eckbo *et al.* (2000), por su parte, emplean un modelo de seis factores<sup>36</sup> consistente con el modelo APT de Ross (1976), Chamberlain (1988) y el modelo intertemporal de Merton (1973). Mediante el empleo de este modelo, no detectan rendimientos anormales estadísticamente significativos tanto para carteras equiponderadas como ponderadas de OPS en los cinco años posteriores a la emisión<sup>37</sup>, en contraposición con el modelo de Fama y French (1993) y el de Connor y Korajczyk (1988) quienes sí presentan, para su muestra, cierta evidencia de mal comportamiento en el caso de una cartera equipondera-

<sup>33</sup> En este sentido, véase también la tabla 7 (pág. 310) de Mitchell y Stafford (2000).

<sup>34</sup> La muestra de OPI está compuesta por 4.622 casos y la de OPS por 4.526 ofertas realizadas por 2.772 empresas. Ambas muestras abarcan el periodo 1975-1992.

<sup>35</sup> Véase Brav *et al.* (2000) para mayores detalles.

<sup>36</sup> Los factores de riesgo que incluye su modelo son un índice (ponderado) de mercado, el diferencial de tipos de interés entre los bonos del Tesoro a 20 años y 1 año, el diferencial de tipos de interés entre los pagarés del Tesoro a 90 y 30 días, el cambio en el consumo real per cápita de bienes no duraderos, el diferencial en el rendimiento mensual de bonos de empresa con calificación BAA y AAA, y, finalmente, la inflación no esperada. Véase Eckbo *et al.* (2000) para más detalles.

<sup>37</sup> La muestra que emplean está formada por 4.860 OPS y comprende el periodo 1964-1997.

da de emisiones realizadas por empresas del NASDAQ<sup>38</sup>. Eckbo *et al.* (2000) concluyen que su modelo aporta evidencia de que el mal comportamiento detectado tras las OPS por anteriores trabajos es debido a que en ellos no se controla de forma adecuada las fuentes de riesgo de las empresas que realizan una OPS. En concreto, señalan que el empleo de una empresa (cartera) de control para la estimación de los rendimientos anormales es inadecuada debido a que, por una combinación de efectos<sup>39</sup>, las empresas que realizan una OPS presentan un menor riesgo tras la emisión, lo cual conduce a una reducción de sus rendimientos esperados en relación con su control.

Mitchell y Stafford (2000) realizan, en este contexto, un experimento con el fin de obtener una visión más clara de cuál es el efecto que las deficiencias del modelo de tres factores de Fama y French (1993) tienen sobre las tres muestras que analizan en su trabajo. En concreto, descomponen la constante en dos componentes: por un lado, el comportamiento anormal esperado, de acuerdo con las características de la muestra (tamaño, cociente VC/VM y frecuencia del suceso en el tiempo); y por otro lado, la cantidad de comportamiento anormal atribuible a otras fuentes, incluido el suceso. La constante esperada, condicionada a la composición de la muestra, es estimada como la constante media de 1.000 regresiones en serie temporal de otras tantas muestras aleatorias compuestas por empresas similares a las del suceso. De esta forma estiman una nueva «constante ajustada» que mide la diferencia entre la constante estimada mediante la cartera de empresas sujetas al suceso y la anterior constante media. Sus resultados indican que una tercera parte del rendimiento anormal medio mensual estimado para las muestras de fusiones y OPS es consecuencia de la mala especificación del modelo y no del suceso. En cualquier caso, y aunque se produce una disminución de su cuantía y del valor de la *t* asociada, el mal comportamiento de adquisiciones y OPS continua siendo estadísticamente significativo tras realizar el mencionado ajuste.

Aunque el método de regresión en serie temporal resuelve el problema de la dependencia en sección cruzada, puede presentar algunos problemas que deben tenerse en cuenta. En primer lugar, este método asume que los parámetros estimados en las regresiones son constantes a lo largo del tiempo, lo cual resulta poco probable dado que la composición de la cartera cambia cada mes. Por otra parte, el hecho de que distintos sectores tengan estimaciones distintas de los parámetros [Fama y French (1997)] conduce a estimaciones sesgadas cuando la composición de la cartera muestral pasa de concentrarse en un sector a otro. En segundo lugar, como consecuencia del cambio a lo largo del tiempo de la composición de la cartera puede presentarse heteroscedasticidad en el rendimiento anormal, lo cual produce un estimador ineficiente, aunque no sesgado, cuando se estima por mínimos cuadrados ordinarios. En tercer lugar, Loughran y Ritter (2000) identifican tres razones por las que este método tiene un bajo poder para identificar rendimientos anormales. Por un lado, si el comportamiento anormal es más intenso en unos periodos que en otros, al ponderar todos los meses por igual, la regresión en serie tempo-

---

<sup>38</sup> El modelo de Connor y Korajczyk (1988) empleado por Eckbo *et al.* (2000) consta de cinco factores extraídos mediante componentes principales de la matriz de covarianzas de los rendimientos del NYSE/Amex para el período 1964-1997 y del NASDAQ para el período 1974-1997.

<sup>39</sup> Aunque las empresas que realizan una OPS presentan un riesgo de mercado un tanto mayor que empresas que no realizan una emisión, tras ésta se produce una disminución del riesgo de la empresa como consecuencia de la reducción del apalancamiento financiero y el incremento de su liquidez en el mercado (medido por el número de títulos en circulación).

ral del horizonte muestral completo<sup>40</sup> tendrá un bajo poder para detectar un comportamiento anormal en los periodos con una elevada concentración de sucesos frente a la ausencia de comportamiento anómalo en los periodos con un bajo número de sucesos. Por otro lado, si el mal comportamiento se da con mayor intensidad entre las empresas pequeñas que en las grandes, la construcción de la cartera muestral equiponderada debería generar una estimación del rendimiento anormal medio mensual mayor que si se emplea una cartera ponderada por el valor de mercado. Dado que existen buenas razones para esperar que la mala valoración sea más común y mayor entre empresas pequeñas que entre empresas grandes<sup>41</sup>, entonces la ponderación por el valor de mercado provoca un infraestimación de los rendimientos anormales. La tercera razón está directamente relacionada con el empleo de una referencia para la estimación de los rendimientos anormales que a su vez está contaminada con algunas de las empresas que se pretenden analizar. En este punto, Loughran y Ritter (2000) argumentan que si el modelo de Fama y French (1993) se considera como un modelo de equilibrio, como así lo hace Fama (1998), entonces no es apropiado purgar los factores tamaño y valor contable-valor de mercado de empresas de la muestra, como es el caso del factor cartera de mercado. Sin embargo, si no se acepta como modelo de equilibrio, entonces afirman que es legítimo purgar la construcción de los factores, al igual que es razonable purgar las referencias cuando se emplea el método del emparejamiento (*matching approach*), ya que cuando se pretende contrastar una anomalía es empíricamente lícito determinar si un comportamiento es distinto de otros comportamientos en sección cruzada detectados empíricamente o si bien es una simple manifestación de éstos.

Por lo que respecta al método del rendimiento anormal mensual medio, éste es similar al del método de la regresión en serie temporal al calcularse los rendimientos anormales para cada mes natural, incorporando la varianza de la cartera la correlación cruzada de los rendimientos anormales individuales de las empresas que la forman. Como antes, la cartera muestral se reajusta cada mes con el objeto de eliminar de ella aquellas empresas que concluyen el periodo de análisis de  $\tau$  meses. El rendimiento anormal ( $AR_{it}$ ) en el mes natural  $t$  de la empresa  $i$  de la muestra se calcula como sigue:

$$AR_{it} = R_{it} - E(R_{CONTROL,t}), \quad [21]$$

donde  $E(R_{CONTROL,t})$  es el rendimiento esperado en el mes natural  $t$  aproximado a través de una determinada referencia. En cada mes natural  $t$ , se calcula el rendimiento anormal medio ( $MAR_t$ ) de las empresas de la muestra:

$$MAR_t = \sum_{i=1}^{N_t} w_{it} \cdot AR_{it}, \quad [22]$$

donde  $N_t$  es el número de empresas en la cartera muestral en el mes  $t$  y  $w_{it}$  es el peso asignado a la empresa  $i$ . El peso  $w_{it}$  es  $1/N_t$  cuando los rendimientos anormales se calculan equiponderados, y  $VM_{it}/\sum_{i=1}^{N_t} VM_{it}$  cuando los rendimientos anormales se calculan ponderados por el valor de mercado, donde  $VM$  hace referencia al valor de mercado de la empresa  $i$

<sup>40</sup> Si el objetivo es analizar el comportamiento del título en los  $\tau$  meses posteriores al suceso, el horizonte muestral comienza en la fecha de la primera emisión en el calendario y termina  $\tau$  meses después de la última.

<sup>41</sup> En la próxima sección comentamos estas razones.

en el mes natural en curso  $t$ . Por último, se calcula la media de los rendimientos anormales medios mensuales ( $\overline{MAR}_\tau$ ) para la ventana de estudio de  $\tau$  meses como sigue:

$$\overline{MAR}_\tau = \frac{1}{T} \sum_{t=1}^T MAR_t, \quad [23]$$

donde  $T$  es el número total de meses naturales en los que existe observaciones de la cartera muestral.

La hipótesis a contrastar, como es habitual, es la existencia o no de un comportamiento anormal tras el suceso, esto es:

$$H_0: \overline{MAR}_\tau = 0, \quad \text{versus} \quad H_1: \overline{MAR}_\tau \neq 0.$$

Esta aproximación presenta dos ventajas a la hora de realizar el contraste estadístico de la existencia de rendimientos anormales significativos tras el suceso cuando los rendimientos anormales se estandarizan cada mes con una estimación de la desviación típica de la cartera en ese mes [Jaffe (1974) y Mandelker (1974)]: por un lado, la estandarización de los rendimientos anormales mensuales permite corregir la posible existencia de heteroscedasticidad inducida por el cambio en la composición de la cartera a lo largo del tiempo, y, por otro lado, al estandarizar se pondera con mayor peso los periodos de mayor concentración de sucesos que los periodos con baja actividad [Fama (1998), Mitchell y Stafford (2000)] como consecuencia de que la varianza residual de la cartera decrece con el tamaño de la misma, *ceteris paribus*. En consecuencia, el estadístico  $t$  convencional recogido en la expresión [8] se calcula a partir de la serie temporal de rendimientos anormales medios mensuales estandarizados.

Para finalizar, es necesario realizar un último apunte en relación con el bajo poder del método de las carteras de fecha de calendario argumentado por Loughran y Ritter (2000). Mitchell y Stafford (2000) aportan evidencia de que el método de las carteras de fecha de calendario tiene más poder para detectar mal comportamiento que la aproximación BHAR tras tomar en consideración la dependencia en sección cruzada de los rendimientos anormales individuales, en contraposición con lo que sostienen Loughran y Ritter (2000).

#### 4. EL ESQUEMA DE PONDERACIÓN DE LOS RENDIMIENTOS ANORMALES

Si la mala valoración tras el suceso tuviera la misma intensidad tanto para las empresas pequeñas como para las grandes, entonces la estimación del rendimiento anormal sería la misma bajo cualquiera de los dos esquemas: equiponderando los rendimientos o ponderándolos por el valor de mercado. Sin embargo, es habitual en la literatura que bajo un esquema de rendimientos equiponderados los resultados muestren rendimientos anormales significativos durante largos periodos tras un determinado suceso, mientras que la significatividad e incluso el signo del comportamiento desaparecen cuando el esquema es ponderado.

En este punto surge una disyuntiva. Desde un punto de vista económico existen razones que justifican el que la mala valoración sea más intensa entre las empresas pequeñas que en las grandes. Brav y Gompers (1997) argumentan que es más probable que las asi-

metrías informativas se den en las empresas pequeñas como consecuencia (i) del escaso seguimiento por parte de las firmas de bolsa de este tipo de empresas y la consiguiente falta de información<sup>42</sup>, y (ii) por el predominio de inversores individuales debido a la existencia de barreras para los inversores institucionales. Incluso en el caso de que la mala valoración afectara indiferentemente a las empresa pequeñas y grandes, el proceso de ajuste en el precio potencialmente se extendería más en el tiempo como consecuencia del predominio ya comentado de inversores individuales y las propias características en la negociación de las empresas pequeñas (como horquillas más amplias que las grandes empresas) que conllevan costes de negociación mayores [Hensler (1998), Loughran y Ritter (2000)].

Por otra parte, dado que un esquema de equiponderación da un mayor peso a las empresas pequeñas, las inferencias están expuestas con mayor intensidad al problema de la mala especificación del modelo generador de rendimientos normales. Además, Loughran y Ritter (2000) sostienen que la elección del esquema de ponderación es importante en relación con el poder del contraste estadístico. Como señalan Loughran y Ritter (2000), si lo que se intenta medir es el rendimiento anormal de la empresa media tras un determinado suceso, entonces cada empresa debería estar igualmente ponderada. Alternativamente, si el objetivo es cuantificar el cambio medio de la riqueza del inversor tras el suceso, Brav *et al.* (2000) apuntan que el método correcto es ponderar las empresas de la muestra por el valor de mercado.

En cualquier caso, el procedimiento habitual en la literatura es presentar los resultados bajo ambos esquemas, fundamentalmente con el objetivo de destacar las diferencias en sección cruzada que puedan existir.

## 5. CONCLUSIONES

En los últimos años se ha desarrollado una abundante literatura alrededor del estudio del comportamiento bursátil a largo plazo de las empresas tras determinados sucesos empresariales. Los resultados obtenidos se enfrentan con la visión clásica de la incorporación rápida y completa de la información pública por parte de los mercados. Por el contrario, muestran que los precios se ajustan lentamente ante la llegada de nueva información, sugiriendo, por tanto, que los mercados no son eficientes.

No obstante, los problemas conceptuales y estadísticos que presenta la medición y contrastación de los rendimientos anormales a largo plazo han supuesto (i) que parte de la evidencia haya sido puesta en tela de juicio y (ii) que se haya producido un verdadero esfuerzo en la literatura con el propósito de resolver los problemas suscitados.

En este contexto, el objetivo del trabajo es doble: por un lado, poner en evidencia las dificultades que supone este tipo de análisis; y, por otro lado, señalar algunas de las soluciones que se han aportado para superar las limitaciones detectadas. De esta forma, esperamos aportar de una forma clara y concisa los elementos necesarios para un fructífero desarrollo de los estudios a largo plazo en nuestro mercado.

Los problemas para la correcta estimación del rendimiento anormal se derivan del hecho de no disponer del modelo correcto de valoración de activos. En consecuencia, nos

---

<sup>42</sup> Véase, por ejemplo, *Expansión*, 3 de febrero de 2001.

encontramos con que son comunes para todo tipo de estudios, acentuándose, no obstante, en aquéllos con amplios horizontes de análisis. En cualquier caso, ésta continúa siendo una cuestión abierta.

Por lo que respecta a la métrica y contraste del comportamiento a largo plazo, tras analizar los diferentes métodos existentes en la literatura nos decantamos por el uso de carteras formadas según la fecha de calendario. En nuestra opinión, este método proporciona una elevada confiabilidad debido a que (i) resuelve el problema de la dependencia en sección cruzada, problema que, dada la evidencia obtenida en nuestro mercado por Farinós *et al.* (2002), puede estar presente en las estimaciones; (ii) se muestra poco sensible a los potenciales problemas a los que está expuesto; (iii) se demuestra como el más idóneo cuando el tamaño muestral es pequeño (caso este harto frecuente en nuestro mercado) frente a los métodos que estiman el mal comportamiento como la media del rendimiento anormal en sección cruzada de las empresas de la muestra respecto de la fecha del suceso; y (iv) porque no padece del problema de la magnificación de los rendimientos anormales que se produce a medida que el horizonte temporal sobre el que se estiman se amplía. Este hecho es especialmente relevante ya que la amplitud del horizonte de estudio es arbitraria y puede distorsionar gravemente los resultados.

## REFERENCIAS BIBLIOGRÁFICAS

- AGARAWAL, A., JAFFE, J.F. y MANDELKER, G.N. (1992): «The post-merger performance of acquiring firms: a re-examination of an anomaly», *Journal of Finance* 47, pp. 1605-1621.
- ALONSO, A. y RUBIO, G. (1990): «Overreaction in the Spanish equity market», *Journal of Banking and Finance* 14, pp. 469-481.
- ÁLVAREZ, S. (2001): «¿Son las OPIs malas inversiones a largo plazo?», *Actualidad Financiera*, marzo, pp. 21-35.
- ÁLVAREZ, S. y GONZÁLEZ, V.M. (2001): «El comportamiento a largo plazo de la Ofertas Públicas Iniciales (OPIs) en el mercado español de capitales», V Foro de Finanzas de Segovia (Workshop in Finance), julio.
- ANSOTEGUI, C. y FABREGAT, J. (1999): «Initial public offerings on the Spanish Stock Exchange», documento de trabajo, ESADE.
- ASQUITH, P. (1983): «Merger bids, uncertainty, and stockholder returns», *Journal of Financial Economics* 11, 51-83.
- BACHELIER, L. (1900): *Théorie de la Spéculation*, Gauthiers-Villars, París. [Citado en Brealey, R. y Myers, S. (1998): *Fundamentos de Financiación Empresarial*, McGraw-Hill, Madrid].
- BAKER, M. y WURGLER, J. (2000): «The equity share in new issues and aggregate stock returns», *Journal of Finance* 55, pp. 2219-2257.
- BANZ, R. (1981): «The relationship between return and market value of common stocks», *Journal of Financial Economics* 9, pp. 3-18.
- BARBER, B.M. y LYON, J.D. (1997): «Detecting long-run abnormal stock returns: the empirical power and specification of test statistics», *Journal of Financial Economics* 43, pp. 341-372.
- BARBERIS, N., SHLEIFER, A. y VISHNY, R. (1998): «A model of investor sentiment», *Journal of Financial Economics* 49, pp. 307-343.
- BASU, S. (1983): «The relationship between earnings yield, market value, and return for NYSE common stocks: further evidence», *Journal of Financial Economics* 12, pp. 129-156.
- BLUME, M. y STAMBAUGH, R. (1983): «Biases in computed returns: An application to the size effect», *Journal of Financial Economics* 12, pp. 387-404.



- BOEHME, R.D. y SORESCU, S.M. (2002): «The long-run performance following dividend initiations and resumpions: Underreaction or product of chance?», *Journal of Finance* 57, pp. 871-900.
- BRAV, A. (2000): «Inference in long-horizon event studies: A bayesian approach with application to Initial Public Offerings», *Journal of Finance* 55, pp. 1979-2016.
- BRAV, A. y GOMPERS, P.A. (1997): «Myth or reality? The long-run underperformance of initial public offerings: evidence from venture and nonventure capital-backed companies», *Journal of Finance* 52, pp. 1791-1821.
- BRAV, A., GECZY, C. y GOMPERS, P.A. (2000): «Is the abnormal return following equity issuances anomalous?», *Journal of Financial Economics* 56, pp. 209-249.
- BROWN, S.J. y WARNER, J.B. (1980): «Measuring security price performance», *Journal of Financial Economics* 8, pp. 205-258.
- CAMPBELL, J. (1993): «Intertemporal asset pricing without consumption data», *American Economic Review* 83, pp. 487-512.
- CANINA, L., MICHAELY, R., THALER, R. y WOMACK, K. (1998): «Caveat compounder: a warning about using the daily CRSP equal-weighted index to compute long-run excess returns», *Journal of Finance* 53, pp. 403-416.
- CARHART, M. (1997): «On persistence in mutual fund performance», *Journal of Finance* 52, pp. 57-82.
- CHAMBERLAIN, G. (1988): «Asset pricing in multiperiod securities markets», *Econometrica* 56, pp. 1283-1300.
- CONNOR, G. y KORAJCZYK, R. (1988): «Risk and return in an equilibrium APT: Application of a new test methodology», *Journal of Financial Economics* 21, pp. 255-289.
- CONRAD, J. y KAUL, G. (1993): «Long-term market overreaction or biases in computed returns», *Journal of Finance* 48, pp. 39-63.
- COWAN, A.R. y SERGEANT, M.A. (2001): «Interacting biases, non-normal returns distributions and the performance of test for long-horizon event studies», *Journal of Banking and Finance* 25, pp. 741-765.
- CUÉLLAR, B. y LAINEZ, J.A. (1999): «Relación de las variables fundamentales con la rentabilidad de los títulos», X Congreso de AECA, Zaragoza.
- DANIEL, K. y TITMAN, S. (1997): «Evidence on the characteristics of cross sectional variation in stock returns», *Journal of Finance* 52, pp. 1-33.
- DANIEL, K., HIRSHLEIFER, D. y SUBRAHMANYAM, A. (1998): «Investor psychology and security market under- and overreactions», *Journal of Finance* 53, pp. 1839-1885.
- DAVIS, J., FAMA, E.F. y FRENCH, K.R. (2000): «Covariances and average returns: 1929-1997», *Journal of Finance* 55, pp. 389-406.
- DHARAN, B. y IKENBERRY, D. (1995): «The long-run negative drift of post-listing stocks returns», *Journal of Finance* 50, pp. 1547-1574.
- ECKBO, B.E., MASULIS, R.W. y NORLI, Ø. (2000): «Seasoned public offerings: resolution of the 'new issues puzzle'», *Journal of Financial Economics* 56, pp. 251-291.
- ESPENLAUB, S., GREGORY, A. y TONKS, I. (2000): «Re-assessing the long-term underperformance of UK Initial Public Offerings», *European Financial Management* 6, pp. 319-342.
- FAMA, E.F. (1970): «Efficient capital markets: a review of theory and empirical work», *Journal of Finance* 25, pp. 383-417.
- FAMA, E.F. (1976): *The Foundations of Finance*, Basic Books.
- FAMA, E.F. (1998): «Market efficiency, long-term returns, and behavioral finance», *Journal of Financial Economics* 49, pp. 283-306.
- FAMA, E.F. y FRENCH, K.R. (1992): «The cross-section of expected stock returns», *Journal of Finance* 47, pp. 427-465.
- FAMA, E.F. y FRENCH, K.R. (1993): «Common risk factors in the returns on stocks and bonds», *Journal of Financial Economics* 33, pp. 3-56.

- FAMA, E.F. y FRENCH, K.R. (1997): «Industry costs of equity», *Journal of Financial Economics* 43, pp. 153-193.
- FAMA, E.F., FISHER, L., JENSEN, M. y ROLL, R. (1969): «The adjustment of stock prices to new information», *International Economic Review* 10, pp. 1-21.
- FARINÓS, J.E. (1999): «La post-performance de las OPVs en España», documento de trabajo, *Quaderns de Treball* 91 (nova època), Universitat de València.
- FARINÓS, J.E. (2001): «Rendimientos anormales de las OPV en España», *Investigaciones Económicas* 25, pp. 417-437.
- FARINÓS, J.E., GARCÍA, C.J. y IBÁÑEZ, A.M. (2002): «Comportamiento a largo plazo de las ofertas públicas de venta subsiguientes en el mercado español», IVIE (Institut Valencià d'Investigacions Econòmiques) WP-EC 2002-23.
- FORNER, C. y MARHUENDA, J. (2001): «¿Existe en el mercado español un efecto sobre-reacción?», *Revista Española de Financiación y Contabilidad* 107, pp. 39-66.
- FORNER, C. y MARHUENDA, J. (2002): «Evidencia adicional del efecto sobre-reacción en el mercado español de capitales», IVIE (Institut Valencià d'Investigacions Econòmiques) WP-EC 2002-02.
- GARCÍA, D., GARRIDO, J. y SÁNCHEZ, M.J. (1999): «OPVs en la Bolsa de Madrid, 1993-1999», *Revista de la Bolsa de Madrid* 78, pp. 4-8.
- GARCÍA-AYUSO, M. y RUEDA, J.A. (1998): «Anomalías en el mercado de capitales español», documento de trabajo, Universidad de Sevilla.
- GÓMEZ, J.C. y MARHUENDA, J. (1998): «La anomalía del tamaño en el mercado de capitales español», *Revista Española de Financiación y Contabilidad* 97, pp. 1033-1059.
- HENSLER, D.A. (1998): «The nature and persistence of initial public offering aftermarket returns predictability», *Review of Quantitative Finance and Accounting* 10, pp. 39-58.
- IBBOTSON, R.G. (1975): «Price performance of common stock new issues», *Journal of Financial Economics* 3, pp. 235-272.
- IKENBERRY, D., RANKINE, G. y STICE, E. (1996): «What do stock splits really signal?», *Journal of Financial and Quantitative Analysis* 31, pp. 357-377.
- IKENBERRY, D., LAKONISHOK, J. y VERMAELEN, T. (1995): «Market underreaction to open market share repurchases», *Journal of Financial Economics* 39, pp. 181-208.
- JAFFE, J.F. (1974): «Special information and insider trading», *Journal of Business* 47, pp. 410-428.
- JAGANNATHAN, R. y WANG, Z. (1996): «The conditional CAPM and the cross-section of expected returns», *Journal of Finance* 51, pp. 3-53.
- JAKOBSEN, J. y SØRENSEN, O. (2001): «Decomposing and testing long-term returns: An Application on Danish IPOs», *European Financial Management* 7, pp. 393-417.
- JAKOBSEN, J. y VOETMANN, T. (1999): «Volatility-Adjusted performance. An alternative approach to interpret long-run returns», documento de trabajo.
- JEGADEESH, N. (2000): «Long-term performance of seasoned equity offerings: Benchmark errors and biases in expectations», *Financial Management*, otoño, pp. 5-30.
- JEGADEESH, N. y TITMAN, S. (1993): «Returns to buying winners and selling losers: implications for stock market efficiency», *Journal of Finance* 48, pp. 65-91.
- JENSEN, M.C. (1968): «The performance of mutual funds in the period 1945-1964», *Journal of Finance* 23, pp. 389-416.
- KOTHARI, S.P. y WARNER, J.B. (1997): «Measuring long-horizon security price performance», *Journal of Financial Economics* 43, pp. 301-339.
- LEVIS, M. (1993): «The long-run performance of the initial public offerings: The UK experience 1980-1988», *Financial Management* 22, pp. 28-41.
- LINTNER, J. (1965): «The valuation of risk assets and the selection of risk investments in stock portfolios and capital budgets», *The Review of Economic and Statistics* 47, pp. 13-37.
- LIPSEY, R.G. (1971): *Introducción a la Economía Positiva*, Vicens Vives, Barcelona.

- LJUNGQVIST, A.P. (1997): «Pricing initial public offerings: Further evidence from Germany», *European Economic Review* 41, pp. 1309-1320.
- LOUGHRAN, T. y VIJH, A.M. (1997): «Do long-term shareholders benefit from corporate acquisitions?», *Journal of Finance* 52, pp. 1765-1790.
- LOUGHRAN, T. y RITTER, J.R. (1995): «The new issues puzzle», *Journal of Finance* 50, pp. 23-51.
- LOUGHRAN, T. y RITTER, J.R. (2000): «Uniformly least powerful tests of market efficiency», *Journal of Financial Economics* 55, pp. 361-389.
- LOUGHRAN, T., RITTER, J.R. y RYDQVIST, K. (1994): «Initial public offerings: International insights», *Pacific-Basin Finance Journal* 2, pp. 165-199.
- LYON, J.D., BARBER, B.M. y TSAI, C. (1999): «Improved methods for tests of long-run abnormal stocks returns», *Journal of Finance* 54, pp. 165-201.
- MANDELKER, G. (1974): «Risk and return: the case of merging firms», *Journal of Financial Economics* 1, pp. 303-335.
- MARTÍN, J.F. y PASTOR, M.J. (2001): «Efectos a largo plazo de las ampliaciones de capital en el mercado español», V Foro de Finanzas de Segovia (Workshop in Finance), julio.
- MASULIS, R. (1980): «Stock repurchase by tender offer», *Journal of Finance* 35, pp. 305-319.
- MERTON, R. (1973): «The theory of rational option pricing», *The Bell Journal of Economics and Management Science*, primavera, pp. 141-183.
- MICHAELY, R., THALER, R. y WOMACK, K. (1995): «Price reactions to dividend initiations and omissions», *Journal of Finance* 38, pp. 1597-1606.
- MITCHEL, M.L. y STAFFORD, E. (2000): «Managerial decisions and long-term stock price performance», *Journal of Business* 73, pp. 287-329.
- NIETO, B. (2001): «Los modelos multifactoriales de valoración de activos: Un análisis empírico», IX Foro de Finanzas, Pamplona.
- RAU, P.R. y VERMAELEN, T. (1998): «Glamour, value, and the post-acquisition performance of acquiring firms», *Journal of Financial Economics* 49, pp. 223-253.
- RITTER, J.R. (1991): «The long-run performance of initial public offerings», *Journal of Finance* 46, pp. 3-27.
- ROSENBERG, B., REID, K. y LANSTEIN, R. (1985): «Persuasive evidence of market inefficiency», *Journal of Portfolio Management* 11, pp. 9-17.
- ROSS, S.A. (1976): «The arbitrage theory of capital asset pricing», *Journal of Economic Theory* 13, pp. 341-360.
- RUBIO, G. (1986): «Análisis multivariante del cero-beta CAPM: el mercado español de capitales», *Revista Española de Economía* 3, pp. 343-365.
- SHARPE, W.F. (1964): «Capital asset prices: a theory of market equilibrium under conditions of risk», *Journal of Finance* 19, pp. 425-452.
- SPIESS, D.K. y AFFLECK-GRAVES, J. (1995): «Underperformance in long-run stock returns following seasoned equity offerings», *Journal of Financial Economics* 38, pp. 243-267.
- SPIESS, D.K. y AFFLECK-GRAVES, J. (1999): «The long-run performance following public bond issues», *Journal of Financial Economics* 54, pp. 243-267.
- VILLALBA, D. (2001): «Behavioral finance. Un paradigma emergente en finanzas», *Revista de Bolsa de Madrid* 102, pp. 57-61.