

Índices de performance, gestión activa y eficiencia. Un análisis de sensibilidad y del fenómeno de la persistencia

El presente artículo profundiza en el denominado «fenómeno de persistencia» en el comportamiento económico de los fondos de inversión y profundiza en las conclusiones de un trabajo previo en el que se detectaron que algunas medidas de performance (tanto clásicas -índices de Sharpe o de Treynor-, como otras más modernas, -TRIP o el Ratio de Información-) aplicadas a una amplia muestra en un periodo bajista como el 99-02, resultaban útiles para tratar de batir al mercado en un periodo posterior de carácter alcista (03-05). Para ello se realiza un análisis sobre la robustez de las conclusiones (ante cambios en el título sin riesgo, el periodo básico de análisis o el benchmark utilizado) y se aplican los tests estadísticos más utilizados en el análisis de la persistencia. La conclusión más relevante se refiere al hecho de que, en el periodo estudiado y con la muestra considerada, se acepta la existencia del fenómeno de persistencia, lo que justificaría la utilización de las medidas de performance en la gestión activa de fondos; todo lo cual podría interpretarse como una evidencia en contra de la hipótesis de eficiencia débil del mercado.

Artikulu honek funtsen jokaeran «iraunkortasun fenomenoa» delakoaren azterketan sakondu egiten du. Aurretiko lan batean antzeman genuen performance deritzonaren neurri batzuk (bai klasikoak -Sharpe edo Treynor indizeak- bai beste batzuk berriagoak -TRIP edo Informazioaren Ratioa-) 99-02 bezalako beheranzko aldi batean lagin zabal bati aplikatuta erabilgarriak zirela merkatua gainditzen ahalegintzeko goranzko izaera zuen ondorengo aldian (03-05); oraingo lan honek ondorio horietan sakondu egiten du, horien sendotasunaren azterketa bat eginez (arrisku gabeko tituluan, analisiaren oinarrizko aldian edo erabilitako benchmark delakoan aldaketan aurrean), baita aipatutako iraunkortasun fenomenoaren azterketan gehien erabilitako test estatistikoak aplikatuz ere. Ondorio nabarmenena iraunkortasun fenomenoaren existitzen dela, aztertutako aldian eta kontuan hartutako laginarekin, onartzeari dagokio eta, horrek, performance delakoaren neurriak funtsen kudeaketa aktiboan erabiltzea justifikatuko luke; hori guztia merkatuaren eraginkortasun ahularen hipotesiaren aurkako ebidentziatzat uler liteke.

The paper deepens in the analysis of the so called «Persistence phenomenon» related to the performance of funds. In previous studies, we found that some performance measures (both classical -Sharpe, Treynor, ...- and new ones -Penalized Internal Rate of Return, Information Ratio-) calculated on a wide sample of European Funds in a bearish market (99-02) were useful in order to beat market performance in a subsequent bullish period (03-05). In this paper we deepen in the main conclusions of the previous study, by the means of analysing their robustness (when using a different risk-free rate, working on weekly basis -instead of monthly basis- or taking a different benchmark) and using several statistical tests in order to study the Persistence Hypothesis. Results show that in the period and sample studied the Persistence Hypothesis should be accepted, so performance measures could be helpful in order to beat market performance on an active management basis; which could be interpreted as evidence against the Weak Market Efficiency Hypothesis.

ÍNDICE

1. Introducción
 2. Método y principales conclusiones de estudios anteriores
 3. Un primer análisis de la robustez de los resultados obtenidos ante cambios en el título sin riesgo o en el período básico de análisis utilizados
 4. Estudio del fenómeno de la persistencia
 5. Conclusiones
- Referencias bibliográficas
Anexo

Palabras clave: medidas de performance, fondos, eficiencia, persistencia, empírico.

Keywords: performance measures, funds, efficiency, persistence, empiric.

N.º de clasificación JEL: G11, G14, G23

1. INTRODUCCIÓN¹

Hace ya bastantes años que existe un gran interés por conocer la calidad en la gestión de las inversiones en bolsa: fondos de inversión, carteras gestionadas, etc. A la hora de calificar como buena o mala su gestión, supuesta la liquidez de estas inversiones, nos deberemos fijar en su rentabilidad y en su riesgo. Todavía existen inversores, e incluso algunos profesionales, que se fijan fundamentalmente en la rentabilidad, y esto no es correcto en un contexto de ene-

migos del riesgo: dado que somos enemigos del riesgo (tal como ha supuesto desde hace años la literatura financiera) deberemos comparar la rentabilidad obtenida con el riesgo soportado por el inversor. Esto es lo que hacen los índices de *performance*, y por eso son interesantes para inversores, gestores de carteras, etc.

Una pregunta razonable es qué índice utilizar, dado que hay muchos posibles. En este trabajo comprobaremos cómo los cinco utilizados llegan a resultados similares. Otro tema interesante es el denominado fenómeno de la persistencia, o «manos calientes»; según esta hipótesis los fondos bien gestionados en un periodo, lo suelen ser también en el siguiente, y al

¹ Los autores quieren agradecer los comentarios y sugerencias realizadas por el equipo de evaluadores internos y externos de la revista *Ekonomiaz*. Cualquier error u omisión es responsabilidad estricta de los autores.

revés. Si esto fuera así, el saberlo tendría un interés indudable para los inversores y para la industria de fondos en general.

También se plantea con frecuencia si pequeños cambios en las carteras tomadas como referencia alterarán las conclusiones. Este tipo de preguntas se las hacen con frecuencia los profesionales de la gestión de carteras, y en este trabajo tratamos de colaborar a su contestación.

El interés por comparar los índices de *performance* ha sido clásico en la literatura sobre el tema. Más recientemente se han intensificado los trabajos sobre el fenómeno de la persistencia, con estudios ya clásicos como el de Malkiel (1995) u otros más recientes como Busse y Irvine (2006) o Kosowski, Timmermann, Wermers y White (2006).

En Gómez-Bezares *et al.* (2007) se analizaba la consistencia de los principales índices utilizados en el análisis de la *performance* de títulos y carteras en bolsa (tanto la que los distintos índices presentan entre sí en un determinado periodo, como la que cada uno muestra en dos periodos consecutivos), y se estudiaba también las posibilidades de batir al mercado con carteras diseñadas a partir de los mencionados índices. Los resultados obtenidos apuntaban entonces a la existencia de una elevada coherencia de las diferentes medidas de *performance*, así como a la posibilidad de batir al mercado utilizando estrategias basadas en la información proporcionada por los índices estudiados (conclusión ésta última que podría interpretarse como una evidencia en contra de la hipótesis de eficiencia débil del mercado, en la medida en que sería posible obtener rentabilidades extraordinarias partiendo de información histórica).

En el presente artículo nos proponemos profundizar algo más en el análisis de las conclusiones apuntadas (que pueden con-

siderarse como sorprendentes), lo que haremos desde una doble perspectiva. Por un lado, se analiza la sensibilidad de los resultados obtenidos (en lo que se refiere al tipo de interés sin riesgo empleado; a la utilización de índices de mercado como *benchmarks* o carteras a batir –en lugar de la cartera equiponderada formada por todos los fondos intervinientes–; o a la consideración de la semana en lugar del mes como periodo básico de análisis). El segundo enfoque se refiere a un análisis riguroso desde un punto de vista estadístico de lo que se conoce como «fenómeno de persistencia» (es decir, el hecho de que el comportamiento presentado por los fondos en un periodo tienda a mantenerse en el siguiente), comprobando hasta qué punto los resultados obtenidos en el trabajo anterior pudieran deberse al azar.

La estructura del artículo es la siguiente: en el segundo apartado resumiremos brevemente los aspectos más relevantes de los estudios anteriores (Gómez-Bezares *et al.*, 2007), tanto en lo que se refiere a la descripción de la base de datos utilizada y al método empleado como a las conclusiones más destacables; en el tercer apartado se realiza un análisis de sensibilidad de los resultados obtenidos previamente, replicando íntegramente el trabajo original con rentabilidades semanales y sustituyendo el título sin riesgo empleado por otro distinto, de menor rendimiento; el cuarto apartado recoge las principales conclusiones obtenidas en el estudio del fenómeno de la persistencia, realizándose también en este caso un análisis de sensibilidad relativo al *benchmark* utilizado; finalmente, en el quinto y último apartado mostraremos un resumen de las principales conclusiones alcanzadas.

La justificación teórica e interpretación de los índices se encuentra en Gómez-Bezares *et al.*, 2003, 2004, 2007.

2. EL MÉTODO Y PRINCIPALES CONCLUSIONES DE ESTUDIOS ANTERIORES

2.1. Descripción de la base de datos y metodología utilizada

La base de datos original con la que se realizó el estudio estaba constituida por las rentabilidades semanales (calculadas en euros) correspondientes a 983 fondos de renta variable comercializados en Europa en el periodo comprendido entre febrero de 1999 y octubre de 2005 (muestra que puede considerarse como extraordinariamente representativa, al incluir la práctica totalidad de los fondos de renta variable comercializados en Bélgica, Francia, Alemania, Irlanda, Italia, Luxemburgo, Holanda, España, Suiza y Reino Unido que presentaban la antigüedad indicada, una vez eliminados algunos que mostraban «huecos» en la información disponible). En cuanto a la alternativa de inversión considerada como suficientemente segura (la que actuaría como «título sin riesgo»), nos decantamos por el índice confeccionado por Analistas Financieros Internacionales (AFI) como media de los fondos comercializados en España que invierten su dinero en activos del mercado monetario en la zona euro durante el mismo periodo de tiempo considerado (también semanal). Finalmente, tomamos como «*benchmark*» (o alternativa a batir mediante una gestión activa) una «cartera de mercado» confeccionada como media equiponderada de las rentabilidades de los 983 fondos considerados².

Partiendo de lo anterior, y a la vista de que los estudios clásicos muestran una mejor adaptación de los modelos que constituyen la base de los principales índices de *performance* utilizados si se trabaja en base

mensual³, procedimos a obtener las rentabilidades mensuales correspondientes (de los 983 fondos originales, así como de la cartera de mercado —a la que llamaremos R_m — y del título sin riesgo —al que nos referiremos como R_0 —), para lo que supusimos meses de cuatro semanas (lo que permite obtener 13 rentabilidades mensuales por fondo o cartera y año considerados).

El siguiente paso consistía en dividir el periodo total en dos subperiodos: el comprendido entre 1999 y 2002; y el que abarca los años 2003 a 2005. Ello permitiría posteriormente estudiar dos cosas: la coherencia que un mismo índice de *performance* muestra entre los dos periodos analizados; y la posibilidad de utilizar la información obtenida en el primer periodo para construir carteras que batan al mercado en el segundo.

Con carácter previo, se realizó un estudio de la normalidad de las rentabilidades mensuales obtenidas, así como de la estabilidad de las betas utilizadas (que como es sabido, son la medida de riesgo sistemático utilizada por algunos de los índices de *performance* considerados). Una vez eliminados aquellos fondos que no cumplían con las características impuestas (menos del 10% del total, como veremos después), procedimos a calcular, para cada periodo, el valor que los distintos fondos presentaban en cada uno de los índices de *performance* considerados en el estudio:

- Índice de Sharpe (S_i). Premio de rentabilidad por unidad de riesgo total (medido con la desviación típica) ofrecido por el fondo⁴.
- Índice de Treynor (T_i). Premio por unidad de riesgo sistemático (medido por beta) ofrecido por el fondo⁵.

² De esta manera usamos, tanto como tipo sin riesgo, como para *benchmark*, medias de fondos, relacionando cada fondo con el conjunto de la industria.

³ Véase Fama (1976) o Gómez-Bezares, Madariaga y Santibáñez (1994).

⁴ Véase Sharpe (1966). Puede verse un planteamiento alternativo en Modigliani y Modigliani (1997).

- Índice de Jensen (J_j). Diferencia entre el premio por riesgo obtenido por el fondo y el que debería haber dado según el *Capital Asset Pricing Model* (CAPM)⁶.
- TRIP coherente con Sharpe (TRIP- S_j)⁷. Rentabilidad equivalente cierta obtenida por el fondo «i» sobre la base de la Línea del Mercado de Capitales (LMC)⁸.
- Ratio de Información (RI_i) entendido como cociente entre el promedio y el riesgo del *tracking error* (TE, diferencia, en cada mes, entre la rentabilidad obtenida por el fondo y la cartera que actúa como *benchmark*)⁹.

Obtenidos los valores que los fondos estudiados alcanzaban en cada índice (y en cada uno de los dos subperiodos considerados) procedimos a confeccionar los *rankings* correspondientes (es decir, el número de orden que los fondos ocupaban en cada índice y en cada periodo). Ello nos permitía realizar el análisis de la coherencia de las medidas de *performance* utilizadas desde una doble perspectiva y a través de dos caminos distintos:

- Coherencia entre los distintos índices de *performance* en un mismo periodo:
 - Análisis del valor, signo y significación estadística del coeficiente de correla-

ción de Pearson obtenido entre los valores en cada par de índices.

- Análisis del valor, signo y significación estadística del coeficiente de correlación de rango de Spearman obtenido entre los *rankings* correspondientes a cada par de índices.
 - Coherencia de cada índice de *performance* en los dos subperiodos analizados:
 - Análisis del valor, signo y significación estadística del coeficiente de correlación de Pearson calculado con las puntuaciones de los fondos en un mismo índice en cada uno de los dos periodos considerados.
 - Análisis del valor, signo y significación estadística del coeficiente de correlación de rango de Spearman calculado entre los *rankings* obtenidos con un mismo índice en los dos periodos estudiados.

La segunda parte de lo anterior constituía así una primera aproximación al estudio del «fenómeno de la persistencia». Para profundizar en el estudio del mencionado fenómeno construimos, para cada uno de los cinco índices considerados, y partiendo de los *rankings* obtenidos en el primer subperiodo, cuatro carteras:

- Cartera T1. Constituida por los cincuenta mejores fondos en cada índice (los situados en los puestos 1 a 50 en el *ranking* correspondiente).
- Cartera B1. Constituida por los cincuenta peores fondos (los que ocupan las cincuenta últimas posiciones en el *ranking* correspondiente).
- Cartera T2. Constituida por los cincuenta mejores fondos una vez eliminados los cincuenta primeros (es decir, los fondos clasificados como T1).
- Cartera B2. Formada por los cincuenta fondos peor situados en cada índice.

⁵ Véase Treynor (1965). Puede verse un planteamiento alternativo en Modigliani (1997).

⁶ Véanse Jensen (1968 y 1969). Pueden verse planteamientos alternativos a los índices clásicos de Sharpe, Treynor y Jensen, en Ferruz y Sarto (1997).

⁷ Cuya formulación incluimos aquí por tratarse del índice de *performance* menos conocido de los utilizados:

$$\text{TRIP-}S_i = \mu_i - \frac{\mu_m - R_o}{\sigma_m} \sigma_i$$

⁸ Véanse Gómez-Bezares, Madariaga y Santibáñez (2003 y 2004).

⁹ Puede verse formulado en Modigliani y Modigliani (1997). Sobre el TE pueden verse también Roll (1992) ó Muralidhar (2000 y 2001).

ce después de eliminar los que constituyen la cartera B1.

Nuestra intención era la de estudiar el comportamiento en el segundo subperiodo de las carteras formadas por los fondos que aparecían como «mejores» y «peores» en el primero (siempre con cada uno de los cinco índices de *performance* considerados), analizando si existían diferencias entre ellas en lo que se refiere a su capacidad de batir a la cartera utilizada como *benchmark* (R_m , tal como hemos indicado).

2.2. Principales conclusiones obtenidas

Con carácter previo, y para comprender mejor el alcance de algunos de los elementos que se destacan después, presentamos en el cuadro n.º 1 el comportamiento de la cartera de mercado y del título sin riesgo en cada uno de los dos subperiodos analizados: como puede verse, en el primero de ellos asistimos a un mercado bajista (con una rentabilidad media negativa), mientras que en el segundo la cartera de mercado presenta un rendimiento positivo y superior al del título sin riesgo (como es lógico esperar a priori, supuesta aversión al riesgo). Ello hace que la conclusión de que puede batirse al mercado en un periodo utilizando la información correspondiente al anterior sea en nuestro caso aún más sorprendente, en la medida en que supondría que los buenos gestores de fondos en periodos bajistas también lo son en periodos alcistas.

Las principales conclusiones alcanzadas en el trabajo previo pueden resumirse en los siguientes términos:

- La hipótesis de normalidad de las rentabilidades mensuales (contrastada con el test de Jarque-Bera) se acepta en la inmensa mayoría de los casos (lo que es coherente con los estudios clásicos): para un error alfa del 1% sólo

85 de los 983 fondos originales no pasan la prueba de normalidad.

- La hipótesis de estabilidad de beta (que contrastamos con el test de Chow) se acepta en 977 de los 983 fondos originales analizados (siempre para un error alfa del 1%). Así, el número total de fondos que superan simultáneamente las pruebas de normalidad y estabilidad de beta es de 893, que fueron los que consideramos finalmente en el resto del estudio¹⁰.
- La coherencia entre los *rankings* obtenidos en un mismo periodo con los distintos índices de *performance* utilizados es muy elevada (correlaciones de rango de Spearman en todos los casos superiores a 0,89 en el periodo 99-02; y mayores que 0,92 –salvo con el Ratio de Información, con valores siempre superiores a 0,74– en 03-05; valores todos ellos estadísticamente significativos).
- También la coherencia entre los valores obtenidos por los fondos en los distintos índices de *performance* analizados dentro de un mismo periodo es muy alta: si prescindimos del Ratio de Información, los coeficientes de correlación de Pearson en los periodos 99-02 y 03-05 son superiores a 0,94 y 0,87, respectivamente; y mayores que 0,86 y 0,74 si consideramos dicho índice; valores todos ellos nuevamente significativos desde un punto de vista estadístico.
- En cuanto a la consistencia que un mismo índice presenta en los dos periodos analizados, ésta es también alta, aunque algo menor. Así, por ejemplo, el coeficiente de correlación

¹⁰ La razón de eliminar los fondos en los que variaba beta se debe a que entendimos que tales fondos habían cambiado su «vocación».

de rango de Spearman más elevado es el asociado a los *rankings* obtenidos con el Ratio de Información (0,51), siendo el más bajo el asociado al índice TRIP-Sharpe (0,41); conclusiones similares se obtienen al analizar los coeficientes de correlación de Pearson entre los valores obtenidos por los fondos con un mismo índice en los dos periodos (si bien el valor correspondiente al índice TRIP-Sharpe baja hasta 0,35), resultando estadísticamente significativos todos los valores obtenidos.

- En lo que se refiere al comportamiento que los 100 mejores y los 100 peores fondos del periodo 99-02 muestran en 03-05 (con cada uno de los índices analizados y repartidos en carteras de 50 fondos) los resultados obtenidos se presentan de manera resumida en los cuadros n.º 2 y 3. Como puede verse en el primer, el número de fondos que «repiten grupo» en los dos periodos es elevado entre los mejor clasificados, y baja significativamente entre los últimos; por otro lado, y en lo referente a la posibilidad de batir al mercado en el segundo periodo, los fondos mejor clasificados en el primero muestran un alto grado de capacidad en este sentido, que baja también significativamente en el caso de los fondos peor clasificados. En cualquier caso, y tal como puede comprobarse en el cuadro n.º 3, las carteras T1 y T2 baten siempre al mercado en el segundo periodo (con todos los índices), mientras que las carteras B1 y B2 no lo hacen en ningún caso; lo que podría considerarse como una evidencia a favor de la utilidad de los índices estudiados a la hora de gestionar carteras (y en contra, tal como hemos dicho, de la hipótesis de eficiencia débil del mercado).

Conviene recordar que la propia metodología utilizada supone trabajar con rentabilidades de fondos que incluyen, en su caso, reinversión de dividendos (o de cualquier otra posible fuente de rentabilidad, como la venta de derechos de suscripción preferente, etc.), así como la corrección correspondiente a las comisiones de gestión. De esta manera, el índice de mercado elegido como *benchmark* (la media no ponderada de las rentabilidades de todos los fondos de renta variable considerados) incluiría también, tanto la corrección por dividendos, como las comisiones correspondientes. Es por ello que nos planteamos si la utilización de un *benchmark* distinto (por ejemplo, un índice de mercado suficientemente representativo del conjunto, a elegir de entre los más frecuentemente empleados) no podría tal vez influir en algunas de las conclusiones alcanzadas; como veremos más tarde, todo el fenómeno de la persistencia lo analizamos sobre la base de los resultados obtenidos con tres *benchmarks* distintos.

3. UN PRIMER ANÁLISIS DE LA ROBUSTEZ DE LOS RESULTADOS OBTENIDOS ANTE CAMBIOS EN EL TÍTULO SIN RIESGO O EN EL PERIODO BÁSICO DE ANÁLISIS UTILIZADOS

Como hemos indicado anteriormente, y con independencia de que continuemos después estudiando la sensibilidad del fenómeno de la persistencia ante cambios en el *benchmark* utilizado, decidimos en primer lugar estudiar si las principales conclusiones alcanzadas cambiarían al optar por un título sin riesgo distinto del considerado en el trabajo original, así como si se supone un periodo básico de análisis diferente al inicialmente empleado. De hecho, una parte importante de las críticas que suelen plantearse a la hora de valorar los resultados de este tipo de estudios se refieren a aspectos

metodológicos de partida (siempre discutibles) y que podrían matizar la validez de algunas de sus conclusiones.

Así, replicamos en primer lugar todo el análisis con un tipo de interés sin riesgo sensiblemente inferior al original, y que podría asociarse al rendimiento ofrecido por un depósito de baja remuneración; y posteriormente realizamos el trabajo completo con el tipo de interés original, pero en base semanal. Aunque en este último caso hay diferencias de importancia en cuanto a la aceptación de la hipótesis de normalidad (lo que también es coherente con los estudios clásicos en la materia), decidimos en el caso de rentabilidades semanales no eliminar ningún fondo (tampoco por falta de estabilidad en las betas, en cuyo caso las diferencias son mucho menores), con la intención de ir comprobando también hasta qué punto las conclusiones son o no sensibles a la eliminación de fondos por los motivos apuntados (aunque nuevamente ello presente también ventajas e inconvenientes).

Las diferencias causadas por la utilización de un tipo de interés distinto del original son muy pequeñas en lo relativo a la coherencia de los índices (tanto entre sí —en un mismo periodo— como entre periodos). Así, las diferencias en los coeficientes de correlación obtenidos entre los distintos índices se producen en muchos casos en el tercer decimal, apareciendo la mayor diferencia en el coeficiente de correlación de Pearson que relaciona los índices TRIP-Sharpe y Ratio de Información en el periodo 2003-05 (una diferencia que no llega a las 7 centésimas); en cualquier caso, todos los coeficientes obtenidos son estadísticamente significativos. También las tendencias se mantienen en lo relativo a la coherencia que muestra cada índice entre periodos, apareciendo sólo diferencias de cierta importancia en el índice TRIP-Sharpe (que ve cómo el coeficiente de correlación de Pearson cae

hasta en 6 centésimas, siendo ya bajo el valor de partida; con todo, resulta también significativo). En cuanto a los fondos que repiten grupo, la máxima diferencia es de tres fondos; y todavía menores son las discrepancias en lo que se refiere al número de fondos que baten al mercado en el segundo periodo. Finalmente, y en lo que hace referencia al comportamiento en 2003-05 de las carteras formadas con los mejores y peores fondos en el periodo 99-02, hay que decir que se mantiene el hecho de que todas las carteras que incluyen los 100 mejores fondos baten al mercado y todas las formadas por los 100 peores son batidas por el mismo (aunque con algunas diferencias de orden entre unas y otras según la aproximación al tipo de interés sin riesgo utilizada).

Algunas diferencias de mayor magnitud aparecen cuando cambiamos el periodo básico de análisis; de destacar serían las relativas al número de fondos que repiten grupo (que baja significativamente) y a la cantidad de fondos que baten al mercado en el segundo periodo (que aumenta de manera importante, incluso entre los fondos peor clasificados en el primero). También resulta reseñable el hecho de que todas las carteras caracterizadas como T1 ó T2 en el primer subperiodo baten al mercado en el segundo (al igual que en los casos anteriores), aunque en este caso también algunas caracterizadas como B1 (con Jensen y Treynor) lo hacen con algunos índices.

Como resumen de todo lo comentado, diremos que las conclusiones no parecen verse afectadas de manera importante por la utilización de una aproximación diferente al tipo de interés sin riesgo, mientras que el empleo de la semana como periodo básico de análisis arroja resultados similares, aunque algo más confusos (lo que nos remite nuevamente a los estudios clásicos, que apuntan ventajas asociadas a la utilización de rentabilidades mensuales).

4. ESTUDIO DEL FENÓMENO DE LA PERSISTENCIA

4.1. Aspectos metodológicos

Como indicábamos al principio del artículo, en este apartado profundizaremos en el estudio de la coherencia que cada índice presenta en dos periodos distintos (prescindiendo, por tanto, de las relaciones que en un mismo periodo pudieran existir entre los diferentes índices considerados).

Volvamos, en una primera aproximación, a lo realizado en Gómez-Bezares et al, 2007 en relación al fenómeno de la persistencia. Como se ha indicado, en su momento confeccionamos, para cada uno de los índices de *performance* considerados, el *ranking* de fondos correspondiente al periodo 99-02. Así, por ejemplo, tomando el índice de Sharpe, llegábamos a la siguiente información:

Nº orden	Fondo	Sharpe (99-02)
1	X	T1
...	...	T1
50	Y	T1
51	Z	T2
...	...	T2
100	A	T2
101	B	-
...	...	-
...	...	-
793	C	-
794	D	B2
...	...	B2
843	E	B2
844	F	B1
...	...	B1
893	G	B1

Partiendo de la información anterior, y realizando un proceso idéntico para el periodo 03-05, analizábamos si los fondos clasificados en cada grupo en el primer periodo tendían a serlo también en el segundo (es decir, si, por ejemplo, un fondo clasificado como T1 en 99-02 volvía a serlo en 03-05); y analizábamos también cuántos de los fondos clasificados en cada grupo en el primer periodo batían al mercado en el segundo. Los resultados obtenidos entonces eran interesantes, pudiendo percibirse una mayor tendencia a «repetir comportamiento» en los fondos mejor clasificados, en proporciones que intuitivamente parecían claramente superiores a las que pudiera provocar el azar (cosa que no ocurría en los fondos peor clasificados). Por otro lado, la capacidad de los fondos mejor clasificados en el primer periodo de batir al mercado en el segundo parecía clara, resultando la proporción de los que lo conseguían claramente superior a la que pudiera deberse al azar. Y todo ello venía avalado por el comportamiento mostrado por las carteras formadas por los fondos mejor clasificados en el primer periodo, que batían en todos los casos al mercado en el segundo (cosa que no ocurría nunca con las carteras formadas por los fondos peor clasificados).

En este trabajo trataremos ahora de:

- Confirmar estadísticamente la existencia del mencionado fenómeno de persistencia (para lo que utilizaremos dos tests distintos, construidos ambos sobre la base de las tablas de contingencia correspondientes).
- Analizar la robustez de los resultados obtenidos, utilizando como *benchmark* un índice de mercado diferente al empleado hasta el momento, y considerando ahora la totalidad de los fondos originales (sin prescindir de ninguno por razones de normalidad o de estabilidad de las betas; lo que pre-

Tabla n.º 1

Persistencia en la pertenencia al grupo T1 (los 50 mejores)

Periodo 03-05				
	Frec. Abs. Obs.	Grupo T1	Resto	Suma filas
Periodo	T1	$T1_{99-02} - T1_{03-05}$	$T1_{99-02} - resto_{03-05}$	50
99-02	Resto	$resto_{99-02} - T1_{03-05}$	$resto_{99-02} - resto_{03-05}$	933
	Suma columnas	50	933	983

senta la ventaja de eliminar posibles sesgos).

Con relación a este último punto, hay que indicar que la utilización de índices de mercado como *benchmark* presentaría a su vez ventajas e inconvenientes:

- Como ventaja, el índice no se vería penalizado por las comisiones (como ocurre en el caso de utilizar la cartera de mercado R_M , cuya rentabilidad se obtenía a partir de las de los fondos, donde están descontadas dichas comisiones). Se evitaría así dar una ventaja artificial a la gestión activa frente a la pasiva.
- Como inconveniente, los índices al uso no suelen considerar la reinversión de los dividendos (que sí es implícitamente considerada al utilizar la cartera R_M). Para evitar este problema, utilizamos un índice formado por los 600 títulos más importantes a nivel europeo, convenientemente corregido por la reinversión de los dividendos correspondientes.

Así las cosas, decidimos realizar esta parte del estudio tomando tres posibles *benchmarks* a batir: la cartera de mercado calculada como media no ponderada de todos los fondos considerados (la misma que se empleó en el trabajo anterior); el lla-

mado SXXP Index, que aglutina a las 600 empresas más grandes de Europa, corrigiendo con dividendos (y que en opinión de los profesionales de la gestión de fondos consultados sería el *benchmark* pasivo ideal en el contexto de la Teoría de cartera y el CAPM); y el SX5P INDEX (corregido nuevamente con dividendos), que es un índice similar al EUROSTOXX, pero que incluye algunos de los títulos más importantes de la zona no euro (y que es claramente peor que el anterior, en la medida en que considera un número de títulos muy inferior, pero que tomamos de alguna manera como contraste). Y todo ello con los 983 fondos y el título sin riesgo originalmente considerado y en base mensual.

Evidentemente, la utilización de *benchmarks* diferentes al empleado originalmente supone que las betas (y todos los índices en los que se consideren éstas) cambian¹¹. Aunque puede demostrarse que los cambios no serían demasiado importantes (de hecho los coeficientes de correlación obtenidos entre betas son elevadísimos), decidimos trabajar con los valores recalculados para cada *benchmark*.

¹¹ Al usar el propio *benchmark* como cartera de mercado.

De esta manera, para cada índice de *performance* y para cada uno de los tres *benchmarks* considerados, se construyeron cuatro tablas de contingencia (todas ellas con las frecuencias absolutas observadas).

Como puede verse en la tabla n.º 1, cada una de las cuatro celdillas de la tabla indica las posibles combinaciones en el comportamiento de un fondo en el transcurso de los dos periodos: así, $T1_{99-02}-T1_{03-05}$ indica el número de fondos clasificados en el grupo T1 en los dos subperiodos considerados; $T1_{99-02}-\text{resto}_{03-05}$ hace referencia al de los que aparecen clasificados entre los cincuenta primeros en el periodo 99-02 pero no lo hacen en 03-05; $\text{resto}_{99-02}-T1_{03-05}$ es el número de fondos clasificados como T1 en el periodo 03-05 pero no en 99-02; y

$\text{resto}_{99-02}-\text{resto}_{03-05}$ se refiere al de los fondos que no quedan clasificados en el grupo de los cincuenta mejores en ninguno de los dos periodos. Así, la primera fila debe sumar necesariamente 50 (ya que los fondos clasificados como T1 en el primer periodo deben repartirse entre T1 y resto en el segundo; y lo mismo ocurre con la suma de la primera columna, que indica el comportamiento en 99-02 de los fondos clasificados como T1 en 03-05 (nuevamente 50). Y dado que el total de fondos analizados es 983, conocida la primera celdilla pueden obtenerse todas las demás.

En la tabla n.º 2 correspondiente a cada índice de *performance* (y a cada *benchmark*) se realiza un estudio similar al anterior pero referido a los fondos clasificados en

Tabla n.º 2

Persistencia en la pertenencia a los grupos T1 ó T2 (los 100 mejores)

Periodo 03-05				
	Frec. Abs. Obs.	Grupo T12	Resto	Suma filas
Periodo	T12	$T12_{99-02} - T12_{03-05}$	$T12_{99-02} - \text{resto}_{03-05}$	100
99-02	Resto	$\text{Resto}_{99-02} - T12_{03-05}$	$\text{Resto}_{99-02} - \text{resto}_{03-05}$	883
	Suma columnas	100	883	983

Tabla n.º 3

Relación entre pertenecer o no a T1 (los 50 mejores) en el primer periodo y batir o no al *benchmark* en el segundo

Periodo 03-05				
	Frec. Abs. Obs.	Bate	No bate	Suma filas
Periodo	T1	$T1_{99-02} - \text{Bate}_{03-05}$	$T1_{99-02} - \text{No bate}_{03-05}$	50
99-02	Resto	$\text{Resto}_{99-02} - \text{Bate}_{03-05}$	$\text{Resto}_{99-02} - \text{No bate}_{03-05}$	933
	Suma columnas	X1	X2	983

cualquiera de los dos primeros grupos (T1 ó T2). Así, en la primera celdilla aparece el número de fondos clasificados como T1 ó T2 en los dos periodos considerados; la celdilla correspondiente a $T12_{99-02}$ -resto $_{03-05}$ indica el número de los que, clasificados como T1 ó T2 en el primer periodo, no lo son en el segundo; $resto_{99-02}$ - $T12_{03-05}$ recoge el número de fondos no clasificados entre los 100 mejores en el periodo 99-02 que sí aparecen como tales en 03-05; y los fondos $resto_{99-02}$ -resto $_{03-05}$ son aquellos que no aparecen entre los 100 mejores en ninguno de los dos periodos considerados. El lector puede comprobar que ahora la suma de la primera fila (lo mismo que la primera columna) debe necesariamente ser 100; con lo que nuevamente, el conocimiento de una celdilla cualquiera de la tabla permitiría calcular las cifras correspondientes al resto.

En la tabla n.º 3 correspondiente a un índice y un *benchmark* concretos analizaremos la relación existente entre «aparecer (o no) clasificado en el grupo de los 50 mejores fondos en el periodo 99-02» y «batir (o no) al mercado en el 03-05». En este caso, puede verse que la suma de la primera fila debe necesariamente ser igual a 50 (por lo que la segunda debe sumar 933), pero la suma de la primera columna no es conoci-

da a priori (ya que se refiere al número de fondos que baten al *benchmark* en el periodo 03-05, lo que no está fijado de antemano como en tablas anteriores). En resumen, X_1 es el número de fondos que baten al *benchmark* en el periodo 03-05 (siendo X_2 el de los que no lo hacen, es decir, $983 - X_1$)

La tabla n.º 4 correspondiente a cada índice (y *benchmark*) estudia la posible relación existente entre los atributos «pertener o no al grupo T1 ó T2 en el primer periodo» y «batir o no al *benchmark* en el segundo periodo». Nuevamente son conocidas a priori las sumas correspondientes a las filas (100 y 883, respectivamente), pero no a las columnas (aunque, evidentemente, las sumas de estas últimas deben coincidir para un mismo índice y *benchmark* en las tablas n.º 3 y 4).

Con cada una de las tablas presentadas se realizan varios tests (en alguno de los casos por caminos diferentes), con la intención de estudiar el fenómeno aludido de la persistencia. En concreto, las tablas n.º 1 y 2 tienen evidentemente que ver con dicho fenómeno, en la medida en que estudian si los fondos tienden a «repetir comportamiento» (es decir, permiten analizar si saber que un fondo pertenece a un determinado grupo en el primer periodo constituye o no una información relevante a la hora de pre-

Tabla n.º 4

Relación entre pertenecer o no a T1 ó T2 (los 100 mejores) en el primer periodo y batir o no al *benchmark* en el segundo

		Periodo 03-05			
		Frec. Abs. Obs.	Bate	No bate	Suma filas
Periodo 99-02	T12		$T12_{99-02} - Bate_{03-05}$	$T12_{99-02} - No\ bate_{03-05}$	100
	Resto		$Resto_{99-02} - Bate_{03-05}$	$Resto_{99-02} - No\ bate_{03-05}$	883
Suma columnas			X_1	X_2	983

Tabla n.º 5

Periodo 03-05				
	Frec. Abs. Obs.	Bate	No bate	Suma filas
Periodo	T1	O_{11}	O_{12}	50
99-02	Resto	O_{21}	O_{22}	933
	Suma columnas	X_1	X_2	983

Tabla n.º 6

Periodo 03-05				
	Frec. Abs. Est.	Bate	No bate	Suma filas
Periodo	T1	$E_{11} = p_1 \cdot n_1$	$E_{12} = p_2 \cdot n_1$	n_1
99-02	Resto	$E_{21} = p_1 \cdot n_2$	$E_{22} = p_2 \cdot n_2$	n_2
	Suma columnas	X_1	X_2	N

ver cómo quedará clasificado en el segundo); y en lo relativo a las tablas 3 y 4, lo que se analiza no es estrictamente el fenómeno de la persistencia, pero sí algo relacionado con ella (ya que nos indican si los fondos bien clasificados en el primer periodo tienden a batir al mercado en el segundo).

Tomando como ejemplo la tabla n.º 3 (queda algún mayor juego, al presentar atributos distintos en filas y columnas), los test utilizados son los siguientes:

1. El que algunos autores llaman test de la «Z de Malkiel»¹². En la tabla n.º 3 original tenemos las frecuencias absolutas «observadas», por lo que podemos utilizar la nomenclatura que se define en la tabla n.º 5.

Si no existiera relación entre los dos atributos considerados («pertenecer o no al grupo T1 en el primer periodo» y «batir o no

al mercado en el segundo»), lo que interpretaríamos como un indicio de falta de persistencia, las proporciones que X_1 y X_2 representan sobre el total (probabilidad de batir o no al *benchmark* en el periodo 03-05) deberían mantenerse de idéntica forma tanto en los fondos clasificados como «T1» o como «resto» en el periodo 99-02. Así, si llamamos:

$$p_1 = \frac{X_1}{N} \quad y \quad p_2 = \frac{X_2}{N}$$

(donde $N=983$), debería ocurrir (en ausencia de relación) que las frecuencias absolutas observadas respondieran al patrón «estimado» (tabla n.º 6), donde $n_1=50$ y $n_2=933$ (siempre que nos refiramos a la información

¹² Malkiel (1995).

recogida en la tabla 3). Insistimos una vez más en el hecho de que éstas son las frecuencias absolutas conjuntas estimadas bajo la hipótesis nula de no relación entre «pertenencia o no al grupo T1 en el primer periodo» y «batir o no al *benchmark* en 03-05».

Por otro lado, el estadístico:

$$Z = \frac{Y - np}{\sqrt{np(1-p)}}$$

que sigue aproximadamente una distribución normal tipificada (0,1), es el que utilizaremos para aceptar o rechazar la hipótesis nula planteada. En esto se basa la Z de Malkiel, que en nuestro caso llevaría a la siguiente expresión asociada a cada posible celdilla de la tabla anterior:

$$Z_{ij} = \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}(1-p_j)}}$$

La fórmula propuesta proporciona el valor experimental, para el que posteriormente se obtendría la cola que deja a su izquierda (o derecha). Lógicamente, la prueba debe realizarse a dos colas, por lo que si deseamos trabajar con un error alfa del 5% (de rechazar la hipótesis nula cuando es cierta) la aceptación de dicha hipótesis exigiría que dicha cola fuera igual o mayor que el 2,5%.

El lector puede comprobar que, en realidad, lo que se contrasta con el test propuesto es si las probabilidades p_1 y p_2 (frecuencias marginales de las columnas) se mantienen en todas las categorías del atributo representado en las filas (lo que sería un indicio de falta de relación entre los atributos recogidos en filas y columnas). Así, es evidente que el test es redundante en cada fila, es decir, que Z_{11} y Z_{12} (al igual que Z_{21} y Z_{22}) son valores totalmente relacionados (de hecho, presentan siempre valores idénticos, pero cambiados de signo). Y también

podría realizarse todo el proceso propuesto calculando las probabilidades a partir de las frecuencias de las filas (en lugar de las de las columnas; pero no entraremos en ello).

2. Test de la Chi cuadrado. Como es sabido, la relación entre dos atributos puede analizarse también mediante este test. En nuestro caso, el estadístico a estudiar sería el siguiente:

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

que sigue aproximadamente una distribución Chi cuadrado con $k = (f-1)(c-1)$ grados de libertad (donde f y c son, respectivamente, las categorías posibles en los atributos presentados en filas y columnas —en nuestro caso, $f=c=2$; y por tanto, $k=1$ grados de libertad—). Finalmente, diremos que, evidentemente, sea cual sea el camino elegido para realizar el test de la Z de Malkiel, la prueba de la Chi cuadrado es la misma (al serlo las tablas de frecuencias observadas y estimadas).

4.2. **Discusión de los resultados obtenidos**

Dado el ingente volumen de datos generados por la combinación de índices de *performance* considerados, *benchmark* elegidos y pruebas a realizar, trataremos de resumir la información relevante en unos cuadros sencillos con la estructura que se presenta en la tabla n.º7.

Como puede verse, en la segunda columna se recogen las dos posibles categorías que puede tomar el atributo considerado en el primer periodo (99-02). Así, por ejemplo, si nos centramos en la primera fila, la Z de Malkiel contrasta la siguiente hipótesis nula: «se mantienen en el grupo 'T1' del periodo 99-02 las frecuencias marginales de las columnas»; mientras que en la segunda fila, con la Z de Malkiel se contrasta la siguiente hipótesis nula: «se mantienen en

Tabla n.º 7

Índice: Sharpe
 Benchmark: Cartera equiponderada R_m

Nº de tabla	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	6,730	0,00%	SI	47,726	0,00%	SI
	resto	1,558	5,96%	NO			
Tabla 2	T1 ó T2	8,544	0,00%	SI	81,262	0,00%	SI
	resto	2,875	0,20%	SI			
Tabla 3	T1	5,476	0,00%	SI	31,594	0,00%	SI
	resto	1,268	10,25%	NO			
Tabla 4	T1 ó T2	7,744	0,00%	SI	66,766	0,00%	SI
	resto	2,606	0,46%	SI			

el grupo ‘resto’ del periodo 99-02 las frecuencias marginales de las columnas». En la cuarta columna se ofrece la cola que queda a la derecha del valor de Z obtenido (o a la izquierda, ya que el valor experimental de Z se muestra en valor absoluto): si es menor que el 2,5% deberemos rechazar la hipótesis nula (lo que interpretaremos como indicio de relación entre los atributos considerados). En realidad, el tener un valor alto de Z en valor absoluto (o de Chi cuadrado), indica «relación», pero no necesariamente persistencia; efectivamente, podría darse también el fenómeno contrario: que los mejores del primer periodo fueran los peores del siguiente, y eso daría también valores altos en los estadísticos, indicando relación. Sin embargo, en nuestro caso los valores de Z (con su signo) indican en todos los casos que esa relación es la coherente con el fenómeno de la persistencia.

La prueba de la Chi cuadrado contrasta directamente si hay relación entre los dos atributos considerados. Así, por ejemplo, la

hipótesis nula que se contrasta en la tabla 1 es: «No existe relación entre los atributos ‘ser T1 o resto en el primer periodo’ y ‘ser T1 o resto en el segundo periodo’». Es claro que se analiza algo muy parecido, pero la Z de Malkiel permite discernir si hay diferencias en la relación dependiendo de la categoría a la que hagamos referencia con respecto al primer periodo.

En los cuadros nº 4 a 8 se muestran los resultados obtenidos, para un error alfa del 5%. De su observación pueden extraerse las siguientes conclusiones:

—Análisis del fenómeno de persistencia en sentido estricto: tablas n.º 1 y 2.

Con la Z de Malkiel, hay que indicar que se rechaza en todos los casos la hipótesis nula en la primera fila de cada tabla. Dicho de otra manera, aceptamos, con todos los índices de *performance*, y con cualquiera de los *benchmark* estudiados, que se da el fenómeno de persistencia en relación a los fondos mejor clasificados. Algo muy pareci-

do ocurre también con los fondos clasificados como «resto» en el primer periodo, aunque en este caso, los índices de Sharpe, Treynor y el Ratio de Información presentan con algunos *benchmarks* resultados contradictorios. De cualquier modo, con el test de la Chi cuadrado, que contrasta directamente la relación entre los dos atributos estudiados, se acepta en todos los casos la existencia de dicha relación.

—Análisis de la relación existente entre «estar o no entre los 50 mejores en el primer periodo» y «batir o no al *benchmark* en el segundo periodo»: Tabla n.º 3.

Con el test de la Z de Malkiel se rechaza la hipótesis nula de que se mantengan las frecuencias marginales de las columnas en la primera fila de cada tabla con todos los índices y *benchmark* considerados (a excepción del tercer *benchmark*, que ya indicábamos que era claramente el peor, con el que se aceptaría dicha hipótesis con los índices de Treynor, Jensen y R.I.). En lo que se refiere a los fondos clasificados en el primer periodo como «resto», se acepta la hipótesis nula. El test de la Chi cuadrado corrobora la existencia de relación entre los dos atributos analizados en todos los casos, salvo con el tercer *benchmark* y con los índices de Treynor, Jensen y R.I.

—Análisis de la relación existente entre «estar o no entre los 100 mejores fondos en el primer periodo» (es decir, estar clasificado como T1 ó T2) y «batir o no al *benchmark* en el segundo periodo»: Tabla n.º 4.

Los resultados vuelven a ser bastante parecidos, rechazándose la hipótesis nula en la primera fila de cada tabla (con todos los índices y *benchmark* estudiados); mientras que en la segunda fila sólo se aceptaría en casos puntuales (en todos los índices, con el tercer *benchmark*; y en el índice R.I. con el segundo y el tercero). En cualquier caso, por

el test de la Chi cuadrado se acepta la existencia de relación entre los dos atributos considerados en todos los casos.

Finalmente, ofrecemos en los cuadros n.º 9 y 10 una información resumida que puede resultar de interés para el gestor de carteras y fondos. Así, en el cuadro n.º 9 se muestra el número de fondos clasificados como T1 y T2 en el periodo 99-02 que batían al *benchmark* en 03-05, junto con el número total de fondos (de los 983 considerados) que batían también a dicho *benchmark*; todo ello para cada índice de *performance* estudiado, y considerando como *benchmark* el que, en opinión de los expertos consultados, resulta más adecuado en este contexto: el SXXP Index (que, recordemos, aglutinaba a las 600 empresas más grandes de Europa), convenientemente corregido con dividendos. En el cuadro n.º 10 se muestra un resumen de los resultados que se obtienen en las tablas n.º 3 y 4 con la Z de Malkiel, empleando ahora un valor $p=0,5$ (ya que, a priori, cabría pensar que la mitad de los fondos quedarán por encima del *benchmark*). Como puede verse, se aceptaría la hipótesis de que la utilización de estrategias basadas en la selección de fondos clasificados como T1 y/o T2 en el periodo 99-02 permite batir al mercado en 03-05 con una probabilidad superior al 50%.

5. CONCLUSIONES

En un trabajo previo, detectamos que algunas medidas de *performance* (tanto clásicas –índices de Sharpe o de Treynor–, como otras más modernas –TRIP o el Ratio de Información–) aplicadas a una amplia muestra de fondos de inversión en un periodo bajista como el 99-02, resultaban útiles para tratar de batir al mercado en un periodo posterior de carácter alcista (03-05). A la vista de los resultados obtenidos en este nuevo trabajo, debemos decir que, realizado un análisis de robustez (o sensibilidad)

de los resultados alcanzados en el trabajo aludido referido a la posibilidad de utilizar un tipo de interés sin riesgo distinto del original (el asociado a depósitos de baja remuneración, en lugar del índice confeccionado por AFI como media de los fondos comercializados en España que invierten su dinero en activos del mercado monetario en la zona euro), un periodo básico de análisis diferente (la semana en lugar del mes) e incluso un *benchmark* alternativo (un índice de mercado que corrija con dividendos, pero no considere las comisiones, en lugar de la cartera equiponderada compuesta por todos los fondos analizados), las conclusiones fundamentales se mantienen: existe una cierta persistencia, al menos en lo que se refiere a los fondos que quedan mejor clasificados en el primer periodo, que tienden también a serlo en el segundo (recuérdese que el primero es un periodo bajista y el segundo es alcista); y por otro lado, se acepta la hipóte-

sis de existencia de relación entre quedar clasificado entre los mejores fondos en el primer periodo y batir al mercado en el segundo (lo que justificaría la utilización de los *ranking* obtenidos con los índices de *performance* en el primer periodo para formar carteras que batan a la gestión pasiva en el segundo). Todo lo cual presenta una significación estadística suficiente.

Por otro lado, podemos también indicar que, como era de esperar, los resultados son algo más confusos al utilizar un *benchmark* como el tercero (SX5P INDEX, que es un índice claramente menos representativo que el segundo -SXXP INDEX, al recoger una parte muy inferior de los títulos cotizados), no existiendo una diferencia importante entre utilizar una cartera de mercado que incluya todos los fondos equiponderados o un índice suficientemente representativo del conjunto.

REFERENCIAS BIBLIOGRÁFICAS

- BUSSE, J.A. y IRVINE, P.J. (2006): «Bayesian Alphas and Mutual Fund Persistence», *The Journal of Finance*, Vol. LXI, 5, October: 2251-2288.
- FAMA, E.F. (1976): *Foundations of Finance*, Basic Books, Nueva York.
- FERRUZ, L. y SARTO, J.L. (1997): «Revisión crítica de las medidas clásicas de performance de carteras y propuesta de índices alternativos. Aplicación a fondos de inversión españoles (1990-1995)», *Boletín de Estudios Económicos*, Vol. LII, 162, Diciembre: 549-573.
- GÓMEZ-BEZARES, F.; MADARIAGA, J.A. y SANTIBÁÑEZ, J. (1994): *Valoración de acciones en la Bolsa Española. Un análisis de la relación entre la rentabilidad y el riesgo*, Desclée de Brouwer, Bilbao.
- (2003): «Medidas de performance: Algunos índices clásicos y relación de la TRIP con la teoría de cartera», *Análisis Financiero Internacional*, 113, Tercer Trimestre: 5-19.
- (2004): «Performance ajustada al riesgo: Índices clásicos y nuevas medidas», *Análisis Financiero*, 93, Primer Cuatrimestre: 6-16.
- GÓMEZ-BEZARES, F.; MADARIAGA, J.A., SANTIBÁÑEZ, J. y APRAIZ, A. (2007): «Índices de performance, gestión activa y eficiencia. Un análisis empírico», *Revista Europea de Dirección y Economía de la Empresa*, vol. 16, 2: 21-40.
- JENSEN, M.C. (1968): «The performance of mutual funds in the period 1945-1964», *Journal of finance*, Mayo: 389-416.
- (1969): «Risk, the pricing of capital assets, and the evaluation of investment portfolios», *Journal of business*, Abril: 67-247.
- KOSOWSKI, R.; TIMMERMANN, A., WERMERS, R. y WHITE, H. (2006): «Can Mutual Fund 'Stars' Really Pick Stocks? New Evidence from a Bootstrap Analysis», *The Journal of Finance*, Vol. LXI, 6, December: 2551-2595.
- MALKIEL, B.G. (1995): «Returns from Investing in Equity Mutual Funds 1971 to 1991», *The Journal of Finance*, Vol. L, 2, June: 549-572.
- MODIGLIANI, L. (1997): «Don't Pick Your Managers By Their Alphas», *U.S. Investment Research, U.S Strategy*, Morgan Stanley Dean Witter, June.
- MODIGLIANI, F. y MODIGLIANI, L. (1997): «Risk-Adjusted Performance», *Journal of Portfolio Management*, vol. 23, 2, Winter: 45-54.
- MURALIDHAR, A. (2000): «Risk-Adjusted Performance: The Correlation Correction», *Financial Analysts Journal*, vol. 56, 5, September/ October: 63-71.
- (2001): «Optimal Risk-Adjusted Portfolios with Multiple Managers», *Journal of Portfolio Management*, vol. 27, 3, Spring: 97-104.
- ROLL, R. (1992): «A Mean/Variance Analysis of Tracking Error», *Journal of Portfolio Management*, vol. 18, 4, Summer:13-22.
- SHARPE, W.F. (1966): «Mutual fund performance», *Journal of business*, Enero: 119-138.
- TREYNOR, J.L. (1965): «How to rate management of investment funds», *Harvard business review*, Enero-Febrero: 63-75.

ANEXO

Cuadro nº. 1

(Rentabilidades mensuales, AFI FIAMM). Comportamiento, en cada periodo, del índice de mercado y del título sin riesgo

Periodo 99-02		μ	σ	β	μ_{TE}	σ_{TE}
Rm	Media aritmética no pon	-0,29%	4,98%	1	0,00%	0,00%
R0	AFI Media FIAMM	0,28%	0,11%	-0,008	0,58%	5,02%
Periodo 03-05		μ	σ	β	μ_{TE}	σ_{TE}
Rm	Media aritmética no pon	1,24%	2,67%	1	0,00%	0,00%
R0	AFI Media FIAMM	0,17%	0,06%	-0,001	-1,07%	2,67%

Cuadro nº. 2

(Rentabilidades mensuales, AFI FIAMM). Resumen estadístico: comportamiento en 2003-2005 de las carteras formadas en función del comportamiento en los distintos índices en 1999-2002

Nº de fondos que repiten comportamiento	S	T	J	TRIPS	RI
Nº de fondos T1 (en 99-02) que repiten grupo T1 en 03-05	13	17	19	16	9
Nº de fondos T2 (en 99-02) que repiten grupo T2 en 03-05	4	16	13	6	7
Nº de fondos B2 (en 99-02) que repiten grupo B2 en 03-05	6	6	4	5	11
Nº de fondos B1 (en 99-02) que repiten grupo B1 en 03-05	6	6	9	7	7
Nº de fondos que baten al mercado	S	T	J	TRIPS	RI
Nº de fondos T1 (en 99-02) que baten al mercado en 03-05	27	42	42	25	40
Nº de fondos T2 (en 99-02) que baten al mercado en 03-05	23	39	37	24	36
Nº de fondos B2 (en 99-02) que baten al mercado en 03-05	3	10	13	4	10
Nº de fondos B1 (en 99-02) que baten al mercado en 03-05	4	14	8	3	7
¿Baten las carteras a la cartera de mercado?	S	T	J	TRIPS	RI
La cartera T1 (según 99-02) bate al mercado en 03-05	SI	SI	SI	SI	SI
La cartera T2 (según 99-02) bate al mercado en 03-05	SI	SI	SI	SI	SI
La cartera B2 (según 99-02) bate al mercado en 03-05	NO	NO	NO	NO	NO
La cartera B1 (según 99-02) bate al mercado en 03-05	NO	NO	NO	NO	NO

Cuadro n°. 3

(Rentabilidades mensuales, AFI FIAMM). Resumen estadístico: comportamiento en 2003-2005 de las carteras formadas en función del comportamiento en los distintos índices en 1999-2002 (en relación a la cartera de mercado y el título sin riesgo)

nº	S	S	T	T	J	J	TRIP-S	TRIP-S	R.I.	R.I.
1	S	50 T1	S	50 T1	TRIP-S	50 T1	S	50 T1	R.I.	50 T2
2	T	50 T1	T	50 T1	J	50 T1	T	50 T1	TRIP-S	50 T2
3	R.I.	50 T2	J	50 T1	S	50 T1	J	50 T1	T	50 T1
4	J	50 T1	TRIP-S	50 T1	T	50 T1	TRIP-S	50 T1	S	50 T1
5	TRIP-S	50 T2	R.I.	50 T2	TRIP-S	50 T2	R.I.	50 T2	J	50 T1
6	TRIP-S	50 T1	TRIP-S	50 T2	T	50 T2	TRIP-S	50 T2	TRIP-S	50 T1
7	S	50 T2	T	50 T2	S	50 T2	T	50 T2	S	50 T2
8	T	50 T2	S	50 T2	R.I.	50 T2	S	50 T2	T	50 T2
9	J	50 T2	R.I.	50 T1	R.I.	50 T1	J	50 T2	J	50 T2
10	R.I.	50 T1	J	50 T2	J	50 T2	R.I.	50 T1	R.I.	50 T1
11	R _m	R _m	R _m	R _m	R0	R0	R _m	R _m	R _m	R _m
12	T	50 B1	T	50 B1	R _m	R _m	R0	R0	R.I.	50 B2
13	TRIP-S	50 B2	R.I.	50 B2	T	50 B1	T	50 B1	J	50 B1
14	R.I.	50 B2	TRIP-S	50 B2	TRIP-S	50 B2	TRIP-S	50 B2	R0	R0
15	S	50 B1	S	50 B1	J	50 B2	J	50 B2	TRIP-S	50 B1
16	J	50 B2	J	50 B2	S	50 B1	S	50 B1	R.I.	50 B1
17	S	50 B2	TRIP-S	50 B2						
18	J	50 B1	J	50 B1	R.I.	50 B2	R.I.	50 B2	T	50 B1
19	TRIP-S	50 B1	TRIP-S	50 B1	T	50 B2	T	50 B2	S	50 B2
20	T	50 B2	T	50 B2	J	50 B1	J	50 B1	S	50 B1
21	R.I.	50 B1	R.I.	50 B1	TRIP-S	50 B1	TRIP-S	50 B1	J	50 B2
22	R0	R0	R0	R0	R.I.	50 B1	R.I.	50 B1	T	50 B2

Cuadro nº. 4

Estudio de la persistencia con el Índice de Sharpe (Rentabilidades mensuales, AFI FIAMM; para un error alfa del 5%)

Benchmark: Cartera equiponderada R_m

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	6,730	0,00%	SI	47,726	0,00%	SI
	resto	1,558	5,96%	NO			
Tabla 2	T1 ó T2	8,544	0,00%	SI	81,262	0,00%	SI
	resto	2,875	0,20%	SI			
Tabla 3	T1	5,476	0,00%	SI	31,594	0,00%	SI
	resto	1,268	10,25%	NO			
Tabla 4	T1 ó T2	7,744	0,00%	SI	66,766	0,00%	SI
	resto	2,606	0,46%	SI			

Benchmark: SXXP Index

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	6,730	0,00%	SI	47,726	0,00%	SI
	resto	1,558	5,96%	NO			
Tabla 2	T1 ó T2	8,544	0,00%	SI	81,262	0,00%	SI
	resto	2,875	0,20%	SI			
Tabla 3	T1	5,349	0,00%	SI	30,146	0,00%	SI
	resto	1,238	10,78%	NO			
Tabla 4	T1 ó T2	7,565	0,00%	SI	63,707	0,00%	SI
	resto	2,546	0,55%	SI			

Benchmark: SX5P INDEX (50 ZONA EURO+GB+...)

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	6,730	0,00%	SI	47,726	0,00%	SI
	resto	1,558	5,96%	NO			
Tabla 2	T1 ó T2	8,544	0,00%	SI	81,262	0,00%	SI
	resto	2,875	0,20%	SI			
Tabla 3	T1	3,571	0,02%	SI	13,437	0,02%	SI
	resto	0,827	20,42%	NO			
Tabla 4	T1 ó T2	4,388	0,00%	SI	21,440	0,00%	SI
	resto	1,477	6,99%	NO			

Cuadro n°. 5

Estudio de la persistencia con el Índice de Treynor (Rentabilidades mensuales, AFI FIAMM; para un error alfa del 5%)

Benchmark: Cartera equiponderada R_m

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	7,374	0,00%	SI	57,290	0,00%	SI
	resto	1,707	4,39%	NO			
Tabla 2	T1 ó T2	10,859	0,00%	SI	131,281	0,00%	SI
	resto	3,654	0,01%	SI			
Tabla 3	T1	5,900	0,00%	SI	36,672	0,00%	SI
	resto	1,366	8,60%	NO			
Tabla 4	T1 ó T2	7,131	0,00%	SI	56,603	0,00%	SI
	resto	2,400	0,82%	SI			

Benchmark: SXXP Index

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	9,305	0,00%	SI	91,222	0,00%	SI
	resto	2,154	1,56%	SI			
Tabla 2	T1 ó T2	11,852	0,00%	SI	156,373	0,00%	SI
	resto	3,988	0,00%	SI			
Tabla 3	T1	5,185	0,00%	SI	28,327	0,00%	SI
	resto	1,200	11,50%	NO			
Tabla 4	T1 ó T2	6,726	0,00%	SI	50,355	0,00%	SI
	resto	2,263	1,18%	SI			

Benchmark: SX5P INDEX (50 ZONA EURO+GB+...)

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	8,661	0,00%	SI	79,038	0,00%	SI
	resto	2,005	2,25%	SI			
Tabla 2	T1 ó T2	12,513	0,00%	SI	174,319	0,00%	SI
	resto	4,211	0,00%	SI			
Tabla 3	T1	1,391	8,21%	NO	2,039	15,33%	NO
	resto	0,322	37,37%	NO			
Tabla 4	T1 ó T2	2,638	0,42%	SI	7,748	0,54%	SI
	resto	0,888	18,73%	NO			

Cuadro nº. 6

Estudio de la persistencia con el Índice de Jensen (Rentabilidades mensuales, AFI FIAMM; para un error alfa del 5%)

Benchmark: Cartera equiponderada R_m

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	9,305	0,00%	SI	91,222	0,00%	SI
	resto	2,154	1,56%	SI			
Tabla 2	T1 ó T2	12,183	0,00%	SI	165,224	0,00%	SI
	resto	4,100	0,00%	SI			
Tabla 3	T1	6,186	0,00%	SI	40,312	0,00%	SI
	resto	1,432	7,61%	NO			
Tabla 4	T1 ó T2	7,333	0,00%	SI	59,858	0,00%	SI
	resto	2,468	0,68%	SI			

Benchmark: SXXP Index

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	10,592	0,00%	SI	118,208	0,00%	SI
	resto	2,452	0,71%	SI			
Tabla 2	T1 ó T2	12,844	0,00%	SI	183,657	0,00%	SI
	resto	4,322	0,00%	SI			
Tabla 3	T1	5,471	0,00%	SI	31,541	0,00%	SI
	resto	1,267	10,26%	NO			
Tabla 4	T1 ó T2	6,321	0,00%	SI	44,474	0,00%	SI
	resto	2,127	1,67%	SI			

Benchmark: SX5P INDEX (50 ZONA EURO+GB+...)

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	10,592	0,00%	SI	118,208	0,00%	SI
	resto	2,452	0,71%	SI			
Tabla 2	T1 ó T2	12,513	0,00%	SI	174,319	0,00%	SI
	resto	4,211	0,00%	SI			
Tabla 3	T1	1,865	3,11%	NO	3,666	5,55%	NO
	resto	0,432	33,29%	NO			
Tabla 4	T1 ó T2	2,638	0,42%	SI	7,748	0,54%	SI
	resto	0,888	18,73%	NO			

Cuadro nº. 7

Estudio de la persistencia con el Índice de TRIP-Sharpe (Rentabilidades mensuales, AFI FIAMM; para un error alfa del 5%)

Benchmark: Cartera equiponderada R_m

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	8,661	0,00%	SI	79,038	0,00%	SI
	resto	2,005	2,25%	SI			
Tabla 2	T1 ó T2	9,205	0,00%	SI	94,335	0,00%	SI
	resto	3,098	0,10%	SI			
Tabla 3	T1	5,476	0,00%	SI	31,594	0,00%	SI
	resto	1,268	10,25%	NO			
Tabla 4	T1 ó T2	6,756	0,00%	SI	50,815	0,00%	SI
	resto	2,274	1,15%	SI			

Benchmark: SXXP Index

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	9,305	0,00%	SI	91,222	0,00%	SI
	resto	2,154	1,56%	SI			
Tabla 2	T1 ó T2	9,867	0,00%	SI	108,383	0,00%	SI
	resto	3,320	0,04%	SI			
Tabla 3	T1	5,953	0,00%	SI	37,334	0,00%	SI
	resto	1,378	8,41%	NO			
Tabla 4	T1 ó T2	6,498	0,00%	SI	47,002	0,00%	SI
	resto	2,187	1,44%	SI			

Benchmark: SX5P INDEX (50 ZONA EURO+GB+...)

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	9,949	0,00%	SI	104,278	0,00%	SI
	resto	2,303	1,06%	SI			
Tabla 2	T1 ó T2	11,521	0,00%	SI	147,765	0,00%	SI
	resto	3,877	0,01%	SI			
Tabla 3	T1	3,259	0,06%	SI	11,191	0,08%	SI
	resto	0,754	22,53%	NO			
Tabla 4	T1 ó T2	3,947	0,00%	SI	17,345	0,00%	SI
	resto	1,328	9,20%	NO			

Cuadro nº. 8

Estudio de la persistencia con el Ratio de Información (Rentabilidades mensuales, AFI FIAMM; para un error alfa del 5%)

Benchmark: Cartera equiponderada R_m

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	2,869	0,21%	SI	8,670	0,32%	SI
	resto	0,664	25,33%	NO			
Tabla 2	T1 ó T2	5,897	0,00%	SI	38,717	0,00%	SI
	resto	1,985	2,36%	SI			
Tabla 3	T1	6,365	0,00%	SI	42,690	0,00%	SI
	resto	1,474	7,03%	NO			
Tabla 4	T1 ó T2	7,562	0,00%	SI	63,655	0,00%	SI
	resto	2,545	0,55%	SI			

Benchmark: SXXP Index

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	2,225	1,30%	SI	5,216	2,24%	SI
	resto	0,515	30,33%	NO			
Tabla 2	T1 ó T2	2,258	1,20%	SI	5,678	1,72%	SI
	resto	0,760	22,36%	NO			
Tabla 3	T1	3,262	0,06%	SI	11,212	0,08%	SI
	resto	0,755	22,51%	NO			
Tabla 4	T1 ó T2	4,411	0,00%	SI	21,663	0,00%	SI
	resto	1,484	6,88%	NO			

Benchmark: SX5P INDEX (50 ZONA EURO+GB+...)

	Z Malkiel Grupo 99-02	Z Malkiel Valor exp	Z Malkiel Cola	Z Malkiel ¿Relación?	Chi Cuad Valor exp	Chi Cuad Cola	Chi Cuad ¿Relación?
Tabla 1	T1	2,225	1,30%	SI	5,216	2,24%	SI
	resto	0,515	30,33%	NO			
Tabla 2	T1 ó T2	2,589	0,48%	SI	7,463	0,63%	SI
	resto	0,871	19,18%	NO			
Tabla 3	T1	1,573	5,79%	NO	2,606	10,65%	NO
	resto	0,364	35,79%	NO			
Tabla 4	T1 ó T2	2,516	0,59%	SI	7,045	0,79%	SI
	resto	0,847	19,86%	NO			

Cuadro n°. 9

Número de fondos clasificados como T1 ó T2 en el periodo 99-02 que baten al mercado en 03-05, tomando como benchmark el SXXP Index (Rentabilidades mensuales, AFI FIAMM)

Nº de fondos que baten al mercado	S	T	J	TRIPS	RI
Nº de fondos T1 (en 99-02) que baten al mercado en 03-05	34	47	48	36	40
Nº de fondos T2 (en 99-02) que baten al mercado en 03-05	34	44	41	27	39
Nº total de fondos (de 983) que baten al mercado en 03-05	320	568	568	320	562

Cuadro n°. 10

Estudio de la persistencia con los distintos índices de performance (a través de las tablas 3 y 4: relación entre ‘pertenencia a los grupos T1 ó T2 en 99-02’ y ‘batir o no al benchmark en 03-05’), tomando como benchmark el SXXP Index (Rentabilidades mensuales, AFI FIAMM) y considerando un valor $p=0,5$ (significación estadística para un error alfa del 5%)

Grupo	Índice	Observados	Estimados	Z	Probabilidad	¿Relación Significativa?
T1	SHARPE	34	25	2,5456	0,55%	SI
T1	TREYNOR	47	25	6,2225	0,00%	SI
T1	JENSEN	48	25	6,5054	0,00%	SI
T1	TRIP-S	36	25	3,1113	0,09%	SI
T1	R.I.	40	25	4,2426	0,00%	SI
T12	SHARPE	68	50	3,6000	0,02%	SI
T12	TREYNOR	91	50	8,2000	0,00%	SI
T12	JENSEN	89	50	7,8000	0,00%	SI
T12	TRIP-S	63	50	2,6000	0,47%	SI
T12	R.I.	79	50	5,8000	0,00%	SI