A metodologia de Henriksson e Merton na detecção da *performance* de fundos de pensões

Elisabete Fátima Simões Vieira (*)
Manuel José da Rocha Armada (**)

1 — Introdução e revisão da literatura

A pesquisa efectuada no campo da avaliação do desempenho aplicada aos gestores dos fundos de pensões é ainda diminuta, quando comparada com a importância que estes fundos têm vindo a adquirir ao longo do tempo. Contudo, a tendência é para uma investigação crescente neste domínio.

Em particular, o Modelo de Equilíbrio dos Activos Financeiros (CAPM), desenvolvido por Sharpe (1964), Lintner (1965) e Mossin (1966), e em geral, a teoria sobre o mercado de capitais deram origem às medidas tradicionais de avaliação da performance de Treynor (1965), Sharpe (1966) e Jensen (1968). Além destas três medidas, duas outras, as quais também consideram a *performance* em termos globais, foram aplicadas ao estudo da *performance* dos fundos de pensões. São elas a medida de Markowitz (1952), que considera o retorno em excesso por unidade de risco limitado a um determinado nível [1], e a de Telser (1956) que maximiza a rendibilidade esperada de um título ou carteira, limitando o risco.

Apesar da sua contribuição, as medidas tradicionais de performance têm sido alvo de várias críticas, nomeadamente por Friend e Blume (1970), Klemkosky (1973), McDonald (1974) e Black, Jensen e Scholes (1972), os quais chegaram à conclusão que estas não são independentes da medida de risco utilizada. Ainda outra crítica apontada prende-se com o facto do horizonte temporal seleccionado poder enviesar as referidas medidas, opinião partilhada por Gaumnitz (1970), Friend e Blume (1970), Levy (1972, 1981, 1984), Ang e Chua (1979) e Kane e Marks (1988). Roll (1978), por sua vez, vem ilustrar a ambiguidade do critério da Linha de Mercado de Títulos, que tem como base o CAPM, defendendo que a utilização de diferentes índices como aproximações à carteira de mercado podem implicar resultados enviesados, originando diferentes ordenações das carteiras relativamente à sua *performance.* Esta crítica foi reforçada por Ferguson (1980, 1986) e Brown e Brown (1987). Além do mais, estas medidas dão-nos apenas estimativas da *performance* total.

Consequentemente, outras pesquisas têm vindo a ser efectuadas, as quais permitem estimar separadamente, pelo menos a nível teórico, as componentes que contribuem para a performance total, dando-se, assim, ênfase à selectividade e ao *timing.*

Treynor e Mazuy (1966) adicionaram um termo quadrático à relação de Jensen (1968), de modo a poderem testar a capacidade de *timing.* Jensen (1972) mostrou que esta capacidade pode ser medida pela correlação entre as previsões do gestor e as rendibilidades realizadas no mercado. No entanto, foi Fama (1972) o primeiro autor a propor, a um nível puramente teórico, a decomposição da rendibilidade global da carteira nas suas diferentes componentes de *timing* e selectividade. Grant (1977, 1978) dá ênfase, nos seus estudos, ao *timing,* mostrando que esta capacidade, no contexto da medida α_p de Jensen (1968), irá causar um enviesamento negativo das estimativas desta. Kon e Jen (1978, 1979) recorrem à técnica da *switching regression* a fim de examinar a possibilidade da existência de vários níveis de risco sistemático, tendo encontrado evidência de que muitos fundos têm alterações discretas ao nível deste risco, o que é consistente com o princípio de que os gestores tentam incorporar actividades de *timing* nas suas

(*) Equiparada a professora-adjunta do Instituto Superior de Contabilidade e Administração de Aveiro, 3810 Aveiro, Portugal.
(**) Professor associado da Escola de Economia e Gestão da Universidade do Minho, Gualtar, 4710 Braga.
[1] Esta medida foi posteriormente simplificada por Sharpe (1964).

decisões de investimento. Bhattacharya e Pfleiderer (1983) alargaram o trabalho de Jensen (1972) e mostraram que se pode recorrer a uma simples técnica de regressão para obter medidas separadas (pelo menos teoricamente) das capacidades de *timing* e selectividade, sendo o seu modelo um refinamento do de Treynor e Mazuy (1966). Hagigi e Kluger (1987) desenvolvem uma outra medida de *performance,* baseada no modelo de Telser (1956), tendo os resultados empíricos obtidos evidenciado uma baixa correlação entre as medidas de Sharpe (1966) e de Telser (1956) quando os fundos são considerados seguros [2] e uma correlação elevada nos casos em que os fundos são pouco seguros. Lehmann e Modest (1987) combinaram a Teoria de Equilíbrio por Arbitragem (APT) com a técnica de regressão de Treynor e Mazuy (1966), tendo constatado que as medidas de *performance* respectivas são sensíveis à carteira padrão escolhida. Coggin, Fabozzi e Rahman (1993) baseando-se no modelo de Treynor e Mazuy (1966) e no de Battacharya e Pfleiderer (1983), recorrem a testes de heterocedasticidade, adaptando-os, de modo a combater uma fraqueza daqueles dois modelos, que ignoram a capacidade de *timing* negativa.

Como podemos constatar, estudos mais recentes tentam analisar e estimar as capacidades de selectividade e *timing* de forma isolada. No entanto, Merton (1981), e pela primeira vez depois de Fama (1972), propõe uma estrutura teórica em que assume que a capacidade de previsão por parte dos gestores de carteiras pode ser dividida em duas componentes: selectividade e *timing* [3]. A principal contribuição centra-se no facto de o modelo de rendibilidades do conjunto de activos ter uma correspondência isomórfica com as rendibilidades derivadas de estratégias de investimento no domínio das opções, mais concretamente, em estratégias *protective put.* É a partir desta correspondência que Merton (1981) desenvolve um modelo de equilíbrio para testar as capacidades de *timing* dos gestores. Por outro lado, baseados no modelo anterior, Henriksson e Merton (1981) e Henriksson (1984) desenvolvem procedimentos estatísticos (i. e., testes paramétricos e não paramétricos) para detectar as capacidades de previsão por parte dos gestores, quer quanto ao *timing,* quer quanto à selectividade.

Neste contexto, e principalmente pelas razões já explicitadas no resumo deste trabalho, propomo-nos analisar as duas componentes da *performance,* tendo como base uma amostra de fundos de pensões em Portugal, e recorrendo às metodologias de Henriksson e Merton (1981).

Assim, na secção seguinte é descrita a metodologia adoptada, bem assim como os dados utilizados. Na secção 3 apresentamos os resultados obtidos, comparando-os com estudos anteriores. Finalmente, na secção 4, procuramos sintetizar as principais conclusões.

2 — Metodologia e descrição dos dados

Quando as previsões dos gestores são observáveis, são utilizados testes não paramétricos, que não requerem pressupostos adicionais sobre a distribuição das rendibilidades. No caso em que não são observáveis, Henriksson e Merton (1981) desenvolvem testes paramétricos baseados no CAPM, podendo deste modo analisar-se separadamente as duas componentes que contribuem para a performance total, e para as quais se requerem apenas as séries temporais das rendibilidades em excesso realizadas em relação à taxa isenta de risco, tanto no que se refere à carteira a analisar como ao mercado.

2.1 — Testes não paramétricos

No caso em que as previsões dos gestores são conhecidas, as probabilidades condicionais de uma previsão correcta são estatísticas suficientes para medir a capacidade de previsão, e não dependem da distribuição das rendibilidades do mercado, nem do CAPM ou de qualquer outro modelo de equilíbrio. As probabilidades de previsão dos gestores, condicionadas à rendibilidade do mercado, podem ser expressas do seguinte modo:

$$p_1(t) \equiv \text{prob}\ [\gamma\ (t) = 0 \mid R_{m,t} \leq R_{f,t}]$$
$$1 - p_1(t) = \text{prob}\ [\gamma\ (t) = 1 \mid R_{m,t} \leq R_{f,t}] \quad \textbf{(1.a)}$$
$$p_2(t) \equiv \text{prob}\ [\gamma\ (t) = 1 \mid R_{m,t} > R_{f,t}]$$
$$1 - p_2(t) = \text{prob}\ [\gamma\ (t) = 0 \mid R_{m,t} > R_{f,t}] \quad \textbf{(1.b)}$$

onde:

$R_{m,t}$ = rendibilidade da carteira de mercado no período t;

[2] O conceito de «seguro», subjacente no modelo de Telser (1956), é considerado como um «limite de risco».
[3] O autor utiliza também os termos microprevisão e macroprevisão para designar selectividade e *timing,* respectivamente.

$R_{f,t}$ = taxa isenta de risco no período t;
$\gamma(t) = 0$ é a previsão feita em t-1, para o período t, de que $R_{m,t} \leq R_{f,t}$;
$\gamma(t) = 1$ é a previsão feita em t-1, para o período t, de que $R_{m,t} > R_{f,t}$.

Consequentemente, $p_1(t)$ é a probabilidade condicional de uma previsão correcta, dado $R_{m,t} \leq R_{f,t}$ (ou seja, mercado em baixa) e $p_2(t)$ é a probabilidade condicional de uma previsão correcta, dado $R_{m,t} > R_{f,t}$ (ou seja, mercado em alta).

O teste está orientado para se determinar até que ponto se pode rejeitar a hipótese nula (*Ho*) de não existência de capacidade de previsão, sendo:

$$Ho\text{: } p_1(t) + p_2(t) = 1 \qquad \textbf{(2)}$$

não dependendo $p_1(t)$ nem $p_2(t)$ da magnitude de $|R_{m,t} - R_{f,t}|$.

Neste quadro, pretende-se determinar a probabilidade P de que a amostra derive de uma população que satisfaça a hipótese nula. Esta é definida pela distribuição hipergeométrica:

$$P(n_1 \mid N_1, N, n) = \left[\binom{N_1}{n_1}\binom{N_2}{n-n_1}\right] \Big/ \binom{N}{n} \qquad \textbf{(3)}$$

onde:

n_1 = número de previsões correctas, dado $R_{m,t} \leq R_{f,t}$;
n = número de previsões de que $R_{m,t} \leq R_{f,t}$;
N_1 = número de observações onde $R_{m,t} \leq R_{f,t}$;
N_2 = número de observações onde $R_{m,t} > R_{f,t}$;
$N = N_1 + N_2$ = número total de observações.

Dado que a distribuição é independente de p_1 e p_2, não é necessário estimar as probabilidades condicionais. Recorrendo a um teste de aceitação de uma ou de duas abas, com um determinado nível de confiança, chegamos à conclusão se devemos aceitar ou rejeitar a hipótese nula. Os autores argumentam que o teste de uma aba é mais apropriado dado que os gestores racionais nunca (ou raramente) terão uma previsão que torne verdadeira a seguinte expressão: $p_1(t) + p_2(t) < 1$. Contudo, em amostras pequenas, como é o nosso caso, tanto os testes de uma como de duas abas permitem resultados «íntegros». Igualmente, no nosso estudo não temos conhecimento das previsões, tendo que recorrer a uma aproximação [4], e então poderá ser possível que $p_1(t) + p_2(t) < 1$, pelo que vamos adoptar o teste de duas abas. Assim, dado N_1, N_2 e n, a distribuição de n_1 é determinada por **(3)** para a hipótese de não existir capacidade de previsão, onde a área de valores possíveis para n_1 é dada por:

$$n_1 \equiv \max(0, n\text{-}N_2) \leq n_1 \leq \min(N_1, n) \equiv \overline{n}_1 \qquad \textbf{(4)}$$

Estas duas últimas relações [**(3)** e **(4)**]) podem ser usadas para estabelecer os intervalos de confiança para testar a hipótese de não capacidade de *timing*. Para um teste de duas abas com um nível de confiança de c, podemos rejeitar a hipótese se $n_1 \geq \overline{x}(c)$ ou se $n_1 \leq \underline{x}(c)$, onde $\overline{x}$ e $\underline{x}$ são determinados pela solução de [5]:

$$\sum_{n_1 = x}^{\overline{n}_1} \left[\binom{N_1}{n_1}\binom{N_2}{n-n_1}\right] \Big/ \binom{N}{n} = (1 - c)/2 \qquad \textbf{(5a)}$$

e:

$$\sum_{n_1 = n_1}^{\underline{x}_1} \left[\binom{N_1}{n_1}\binom{N_2}{n-n_1}\right] \Big/ \binom{N}{n} = (1 - c)/2 \qquad \textbf{(5b)}$$

Como já referimos anteriormente, não temos conhecimento das previsões dos gestores, tendo que recorrer a uma *aproximação* para elaborar os testes não paramétricos. A fim de evitar erros de mensuração da aproximação [Henriksson (1984)], as rendibilidades actuais dos fundos de pensões vão ser utilizadas como medida de *performance*, sendo comparadas com a performance de uma estratégia passiva admissível. A hipótese nula, *Ho*: $p_1(t) + p_2(t) = 1$, reflecte agora a probabilidade dos fundos superarem a estratégia passiva, condicional ao facto desta exceder a rendibilidade dos títulos isentos de risco e a estratégia passiva é baseada numa carteira composta por títulos isentos de risco e pela carteira de mercado, onde a proporção investida nesta última é igual ao beta do fundo, sendo o beta determinado através da seguinte regressão:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p\,[R_{m,t} - R_{f,t}] + \varepsilon_{p,t} \qquad \textbf{(6)}$$

Os parâmetros necessários para o cálculo de $p_1(t) + p_2(t) = n_1/N_1 + (N_2 - n_2)/N_2$ são,

---

[4] Assim como Henriksson (1984).

[5] Dado que a distribuição hipergeométrica é discreta, as igualdades de **(5a)** e **(5b)** poderão não ser obtidas. Portanto, em **(5a)**, $\overline{x}$ deve ser interpretado como o valor mais baixo de x para o qual o somatório não excede (1-c)/2. Em **(5b)**, $\underline{x}$ deve ser considerado como o valor mais alto de x para o qual o somatório não excede (1-c)/2.

por analogia, definidos agora do seguinte modo:

n_1 = número de observações em que $R_{p,t} \leq R_{f,t}$ e $R_{e,t} \leq R_{f,t}$, sendo $R_{e,t}$ a rendibilidade da estratégia passiva;
n_2 = número de observações de que $R_{p,t} \leq R_{f,t}$ e $R_{e,t} > R_{f,t}$;
N_1 = número de observações onde $R_{e,t} \leq R_{f,t}$;
N_2 = número de observações onde $R_{e,t} > R_{f,t}$.

2.2 — Testes paramétricos

O caso que se verifica com maior frequência é o da não existência de previsões dos gestores, tendo os investigadores somente acesso às séries temporais das rendibilidades realizadas nas carteiras. Assim, são mais facilmente aplicados os testes paramétricos.

De acordo com Henriksson e Merton (1981), e de modo a separar, pelo menos teóricamente, os incrementos resultantes do *timing* e da selectividade, a estrutura teórica para este efeito pode ser apresentada do seguinte modo:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_{1,p}\, X_{p,t} + \beta_{2,p}\, Y_{p,t} + \varepsilon_{p,t} \quad (7)$$

onde $Y_{p,t} \equiv \max\,[0,\ R_{f,t} - R_{m,t}] = \max\,[0,\ -X_{p,t}]$, que reflecte o valor da rendibilidade de uma opção de venda (mais precisamente do tipo *protective put*) [6]. De acordo com os autores, $\hat{\beta}_{1,\phi}$ corresponde à fracção investida na carteira de mercado na estratégia de opções, $\hat{\beta}_{2,\phi}$ exprime a capacidade de *timing* do gestor [7] e, relacionando com a estratégia *put*, corresponde ao número de opções de venda no mercado derivadas da capacidade de *timing*. O incremento do desempenho derivado da selectividade pode igualmente ser medido usando a estimativa de α_p.

2.3 — Descrição dos dados

A amostra utilizada neste estudo consiste em 14 fundos de pensões de empresas portuguesas, cobrindo o período de Dezembro de 1987 a Dezembro de 1994 [8], tendo sido a dimensão da amostra e o período de tempo seleccionado condicionados sobretudo pela disponibilidade de dados [9]. A informação necessária para o cálculo das rendibilidades dos fundos de pensões foi obtida nos relatórios trimestrais dos fundos, elaborados pelas entidades gestoras destes. Em alguns dos casos os relatórios foram obtidos junto das empresas associadas e noutros nas próprias empresas gestoras. As rendibilidades foram calculadas em termos brutos, não deduzindo as despesas de gestão. Para o cálculo da rendibilidade do mercado foi utilizado um índice de aproximação ao mercado: o índice do BTA, ajustado a dividendos, o qual foi obtido junto desta instituição bancária. A taxa isenta de risco foi estimada com base na taxa de rendibilidade dos bilhetes do Tesouro a 91 dias, anualizadas. Consideramos a respectiva proporcionalidade para a obtenção das taxas em termos trimestrais. Esta informação foi obtida junto do Banco de Portugal.

2.4 — Hipóteses a testar

A fim de testar as duas metodologias descritas neste trabalho, pretendemos determinar até que ponto podemos rejeitar a hipótese nula (*Ho*), sendo esta definida do seguinte modo (testes não paramétricos):

Ho: os gestores de fundos de pensões não têm capacidade de previsão.

Consideramos ainda duas outras hipóteses, as quais permitirão detectar as duas componentes de *performance*, respectivamente, selectividade e *timing*. Assim, temos (testes paramétricos):

Ho_1: os gestores de fundos de pensões não têm capacidade de selectividade (α_p=0);

Ho_2: os gestores de fundos de pensões não têm capacidade de *timing* ($\beta_{2,p}$=0).

---

[6] Para uma análise mais profunda desta relação, v. Merton (1981) e, resumidamente, Henriksson e Merton (1981).

[7] O verdadeiro valor de $\beta_{2,\phi}$ será zero se o gestor não tiver capacidade de previsão dos movimentos do mercado ou se não actuar sobre esta previsão.

[8] Por motivos de promessa de sigilo às empresas que forneceram a informação, e pelo facto de se tratar de informação não pública, o nome dos fundos não será revelado.

[9] A amostra está isenta do enviesamento devido à sobrevivência dos fundos *(survivorship bias)*, uma vez que no período considerado não houve nenhuma fusão nem desaparecimento de qualquer fundo.

3 — Resultados empíricos

3.1 — Testes paramétricos

Começamos por elaborar os testes paramétricos, dado que a informação facultada pelas empresas nos permitiu, desde logo, obter as rendibilidades ao longo do período em análise, e não as previsões dos gestores.

3.1.1 — Análise para o período global

Os resultados obtidos para as estimativas de selectividade e *timing* que derivam da análise para o período global e da aplicação da expressão **(7)** apresentam-se, de seguida, na tabela 1.

TABELA 1

	Amostra para o período global	
Média (desvio padrão):		
α_p	0.002 09	(0.005 5)
$\beta_{1,p}$	0.056 94	(0.091 0)
$\beta_{2,p}$	0.031 42	(0.105 4)
Número de fundos:		
Rejeitam $\alpha_p = 0$ a 5% (1%)	2 +	(1+)
Rejeitam $\beta_{1,p} = 0$ a 5% (1%)	1 –	(0)
$\alpha_p > 0$	9	
$\beta_{2,p} > 0$	10	

Testes paramétricos: $R_{p,t} - R_{f,t} = \alpha_p + \beta_{1,p} X_{p,t} + \beta_{2,p} Y_{p,t} + \varepsilon_{p,t}$.
Nota. — + designa o número de fundos com estimativas significativamente positivas; — designa o número de fundos com estimativas significativamente negativas.

As estimativas do parâmetro α_p variam entre -0.0075 e 0.0114 [10], sendo positivo em nove fundos e negativo em cinco. Contudo, somente dois fundos exibem estimativas significativamente positivas para α_p com um nível de significância de 5% e um fundo quando este nível é de 1%. Neste contexto, os resultados permitem-nos sugerir que os gestores dos fundos de pensões não tiveram capacidade para prever os preços individuais dos títulos.

As estimativas do parâmetro $\beta_{1,p}$ apresentam valores baixos, sendo mesmo alguns negativos, variando de -0.011 a 0.340. A pequena sensibilidade dos fundos a variações do mercado cremos poder ser justificada pela composição das carteiras daqueles, uma vez que, e após análise cuidada destas para o respectivo período, a percentagem de acções é relativamente diminuta.

Os valores estimados para o parâmetro $\beta_{2,p}$ variam entre -0.154 e 0.289. A um nível de confiança de 95%, rejeitamos a hipótese de $\beta_{2,p} = 0$ apenas para um fundo, sendo o seu valor significativamente negativo, pelo que poderemos dizer que neste caso a capacidade de *timing* é negativa. Chang e Lewellen (1984) designaram este fenómeno de *timing* perverso. Quando aumentamos o nível de confiança para 99%, aceitamos a hipótese $\beta_{2,p} = 0$ para todos os fundos, pelo que, e de acordo com o modelo, os resultados nos indiciam que os gestores não têm capacidade de prever os movimentos gerais do mercado, adaptando de forma coerente a sua carteira. É curioso realçar que existe uma correlação negativa entre $\hat{\alpha}_p$ e $\hat{\beta}_{2,\phi}$, sendo o seu valor de -0.321. De facto, em 9 dos 14 fundos, ou seja em 64.3% dos casos, os parâmetros $\hat{\alpha}_p$ e $\hat{\beta}_{2,\phi}$, apresentam sinal contrário, sugerindo este fenómeno que fundos que obtêm rendibilidades superiores devido à sua capacidade de selectividade têm, ao mesmo tempo, uma capacidade de *timing* negativa (e o contrário). Este fenómeno verificou-se em aplicações anteriores, como por exemplo nos trabalhos pioneiros elaborados por Henriksson (1984) e Chang e Lewellen (1984) relativamente aos USA, Pink (1989) relativamente ao Canadá e Armada (1992) relativamente ao UK. Algumas explicações já foram dadas [v. g. Armada (1992)] embora continue a ser um *puzzle* na área das finanças empresariais.

3.1.2 — Análise para os subperíodos

Com o objectivo de verificar se os resultados mantêm consistência ao longo do período global de análise, dividimos a amostra em dois subperíodos de igual duração: o primeiro vai desde Dezembro de 1987 a Junho de 1991 e o segundo de Julho de 1991 a Dezembro de 1994. Os resultados encontram-se sumariados de seguida, na tabela 2.

[10] Todos os resultados podem ser obtidos mediante solicitação aos autores.

TABELA 2

Período	Dezembro de 1987/Junho de 1991		Julho de 1991/Dezembro de 1994	
Média (desvio padrão):				
α_p	– 0.000 37	(0.010 2)	– 0.000 38	(0.006 7)
$\beta_{1,p}$	0.037 54	(0.055 8)	0.137 56	(0.310 9)
$\beta_{2,p}$	0.026 14	(0.112 0)	0.133 67	(0.363 0)
Número de fundos:				
Rejeitam $\alpha_p = 0$ a 5% (1%)	1 +	(1 +)	1 +	(1 +)
Rejeitam $\beta_{2,p} = 0$ a 5% (1%)	0	(0)	0	(0)
$\alpha_p > 0$ ($\beta_{2,p} > 0$)	6	(9)	8	(9)

Testes paramétricos: $R_{p,t} - R_{f,t} = \alpha_p + \beta_{1,p} X_{p,t} + \beta_{2,p} Y_{p,t} + \varepsilon_{p,t}$, considerando os dois subperíodos.
Nota. — + designa o número de fundos com estimativas significativamente positivas.

Quando analisamos os resultados relativos aos dois subperíodos verificamos que somente um fundo apresenta estimativas significativamente positivas para α_p em ambos os subperíodos [11], diminuindo o número de fundos que apresenta este parâmetro positivo em relação ao período global. Nenhum fundo apresenta capacidade de *timing* significativa. A correlação entre $\hat{\alpha}_p$ e $\hat{\beta}_{2,\phi}$ continua a ser igualmente negativa, assumindo o valor de -0.92 no primeiro subperíodo e de -0.85 no segundo.

Se, adicionalmente, ignorarmos a existência de *timing* e analisarmos a regressão especificada por Jensen (1968) e explicitada em **(6)**, obtemos os resultados que constam da tabela 3, apresentada de seguida.

TABELA 3

	Amostra	
Média (desvio padrão):		
$\hat{\alpha}_p$	0.0037	(0.0068)
$\hat{\beta}_\phi$	0.0471	(0.0639)
Número de fundos:		
Rejeitam $\hat{\alpha}_p = 0$ a 5% (1%)	0	(0)
$\hat{\alpha}_p > 0$	10	

Testes paramétricos: $R_{p,t} - R_{f,t} = \alpha_p + b_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$, considerando o período global e ignorando a capacidade de *timing*.

Neste caso, e ao invés do que geralmente já foi detectado neste sentido noutros países (v. g. Henriksson (1984), onde, em média, os $\hat{\beta}_{2,\phi}$ eram negativos), as estimativas de α_p são ligeiramente mais altas, pois reflecte o facto de o coeficiente omitido ($\beta_{2,p}$) ser, em média, ligeiramente positivo para a especificação de regressão que considera a possibilidade de existência de *timing*. Verifica-se também que, embora existam 10 fundos com positivo, para nenhum deles é rejeitada a hipótese de $\alpha_p = 0$, pelo que poderemos afirmar que os resultados não nos sugerem a existência de capacidade de selectividade, de resto, conclusões estas que foram igualmente retiradas, por exemplo, por Jensen (1968), Henriksson (1984) e Chang e Lewellen (1984).

No anexo 1 apresentam-se os resultados da aplicação do modelo de Jensen para ambos os subperíodos. De realçar que relativamente a todos os fundos se aceita a hipótese de $\alpha_p = 0$, sendo no entanto a média deste parâmetro superior no segundo subperíodo, assim como o número de fundos que apresenta este parâmetro positivo. A correlação de α_p nos dois subperíodos é de -0.0483 [12]. De referir também que as estimativas do β_ϕ em qualquer dos períodos, e por razões já explicitadas, são muito baixas.

3.1.3 — Análise de acordo com a magnitude de $|X_{p,t}|$

Com o intuito de analisar se os gestores têm maior capacidade de prever grandes magnitudes de movimentos de mercado do que pequenas, dividimos a amostra em dois subgrupos, de acordo com a magnitude absoluta do valor de $X_{p,t}$ (tendo como referência de separação o valor mediano de $|X_{p,t}|$), e estudamos o modelo já exposto. Somente um fundo tem um valor significativamente positivo para a estimativa de α_p nas

[11] Sendo o resultado idêntico ao do período global!
[12] Usando o teste de Fisher (1970), nenhum dos fundos rejeita a hipótese de α_p ser igual nos dois subperíodos, tanto a um nível de confiança de 95 como 99%.

duas amostras, considerando um nível de 5% e somente no caso da amostra de pequena magnitude, e com um nível de 5%, existem fundos com valores significativos para a estimativa de $\beta_{2,p}$ (dois significativamente positivos e um negativo) [13]. Não foi encontrada evidência de que os gestores tenham mais sucesso a prever grandes variações no valor da carteira de mercado, do que variações menores. Assim, poderemos inferir que a magnitude absoluta da rendibilidade da carteira de mercado parece não ter influência na avaliação da *performance*.

3.1.4 — O efeito da possível existência de heterocedasticidade nos resultados

Se bem que os coeficientes estimados anteriormente nos dão as estimativas dos parâmetros de *performance* dos fundos analisados, os estimadores dos mínimos quadrados podem ser ineficientes, dado que um dos pressupostos do modelo de regressão (homocedasticidade) pode não se verificar, isto é, o pressuposto de que a variância de cada termo residual, condicional ao valor das variáveis explicativas, é uma constante [14]. Para verificar se rejeitamos ou não a hipótese de homocedasticidade, recorremos ao teste de White (1980) [15], tendo chegado à conclusão de que não se rejeita para nenhum fundo a hipótese de homocedasticidade, tanto para uma significância de 5% como 1%. Deste modo, não necessitamos fazer qualquer correcção para a heterocedasticidade, pelo que as nossas conclusões se mantêm válidas [16].

3.1.5 — Análise da estacionariedade dos parâmetros estimados

De modo a podermos verificar se a distribuição dos parâmetros é estacionária de um período para o seguinte, resolvemos proceder ao teste de Kruskal-Wallis (sugerido por Kolb e Rodriguez (1990)). Tanto a um nível de significância de 5% como de 1%, os valores desta estatística são menores do que o valor de $\chi^2_{\text{crítico}}$, pelo que a hipótese de existir estacionariedade da distribuição dos parâmetros ao longo dos dois subperíodos não é rejeitada para nenhum deles [17]. No entanto, Kolb e Rodriguez (1990) falam de estacionariedade pelo estudo da distribuição dos parâmetros, pois assumem que se a distribuição é idêntica, então os parâmetros são igualmente idênticos, ou seja, são estáveis.

Embora este fenómeno possa ser um indício de estabilidade, não podemos inferir que necessariamente assim aconteça. Para verificar se efectivamente a estabilidade se verifica, elaboramos o teste de Chow (1960), cujos resultados estão apresentados no anexo 3. Tanto a um nível de significância de 5% (F_{tabela}= 3.05) como 1% (F_{tabela}= 4.82), o valor de F tabelado é sempre maior que o F calculado através da amostra, pelo que não rejeitamos o facto dos parâmetros serem estáveis nos dois subperíodos para a nossa amostra.

Embora o teste de Chow (1960) nos indique se as regressões relativas aos subperíodos são ou não diferentes, não identifica a fonte desta diferença. Dado que o nosso interesse se prende essencialmente com o comportamento da selectividade e do *timing,* vamos usar para o efeito o teste *F* de Fisher (1970), que foi igualmente aplicado por Henriksson (1984). Os resultados podem ser vistos no anexo 4. Nenhum dos fundos rejeita a hipótese nula, tanto a um nível de significância de 5 como 1%, de que a selectividade e o *timing* são iguais nos dois subperíodos, o que parece implicar que as estimativas são geralmente estáveis ao longo do tempo.

Dado os resultados dos testes elaborados para avaliar se existe diferença (estacionariedade) nos parâmetros estimados entre os dois subperío-

---

[13] Os resultados encontram-se apresentados de forma sucinta no anexo 2.

[14] Henriksson e Merton (1981) defendem que a estimação dos mínimos quadrados pode ser ineficiente porque $\beta_{\phi,t}$ não é estacionário, o que causará um desvio em $\varepsilon_{\phi,t}$, sendo este uma função crescente de $|X_{p,t}|$. Deste modo, será necessário corrigir (caso exista) a heterocedasticidade.

[15] A vantagem do teste de White (1980) face a outros possíveis, é que não depende do pressuposto de normalidade.

[16] De igual modo, os resultados podem ser obtidos junto dos autores se solicitados para o efeito.

[17] O teste de Mann-Whitney foi também aplicado a todos os pares possíveis de períodos para determinar até que ponto os parâmetros têm a mesma distribuição em dois subperíodos. A hipótese nula a testar é que os parâmetros em dois períodos são distribuídos de forma idêntica. Os valores de Z_{amostra} para os parâmetros α_p, $\beta_{1,p}$ e $\beta_{2,p}$ são de, respectivamente, 0.075, 0.675 e 0.2251. Como $Z_{\text{amostra}} < Z_{\text{crítico}(0.025)}$ = 1,96 para todos os parâmetros, não rejeitamos a hipótese de que estes são distribuídos de forma idêntica nos dois subperíodos. Concluímos portanto, e em particular através destes resultados, que a hipótese da distribuição dos parâmetros α_p (selectividade) e $\beta_{2,p}$ *(timing)* ser estacionária entre os dois subperíodos não pode ser rejeitada, considerando o nível de significância de 5%.

dos [18], podemos concluir que não existem diferenças significativas nos resultados.

3.2 — Testes não paramétricos

Para verificar a coerência dos resultados obtidos aquando da aplicação dos testes paramétricos, nomeadamente pelo facto de existir uma correlação negativa entre $\hat{\alpha}_p$ e $\hat{\beta}_{2,p}$, quer para o período global quer para os subperíodos, aplicamos de seguida os testes não paramétricos. Como não temos as previsões dos gestores, usamos as rendibilidades actuais dos fundos como medida de *performance* e comparámo-las com a *performance* de uma estratégia passiva, aproximação esta também adoptada por Henriksson (1984). Passamos a examinar se as carteiras geridas activamente são ou não capazes de gerar rendibilidades em excesso face àquelas que são obtidas através de uma estratégia passiva, de acordo com a aproximação descrita na secção anterior.

As rendibilidades dos fundos e a estratégia passiva são comparadas, usando para tal o teste de independência descrito na secção 2, onde se aplicam as expressões **(4)**, **(5a)** e **(5b)**. Convém referir que, deste modo, estamos a analisar a *performance* global dos fundos de pensões e não as capacidades de *timing* e selectividade, separadamente.

3.2.1 — Análise para o período global

Os resultados obtidos através da aplicação dos testes não paramétricos para o período global encontram-se na tabela 4, apresentada de seguida.

TABELA 4

	Período global
Estimativa média ($p1 + p2$)	1.30159
Desvio padrão	0.23782
Número de fundos:	
Rejeitam *Ho* a 5% (1%)	1 + (1 +)
Estimativa ($p1 + p2 > 1$	13

Testes não paramétricos: período global; rendibilidade do fundo *vs* estratégia passiva, hipótese nula: $p_1 + p_2 = 1$.
Nota. — + designa o número de fundos com estimativas significativamente positivas.

Embora os fundos tenham, em média, um valor positivo para a estimativa de $p_1 + p_2$, que reflecte a capacidade de previsão dos gestores, estes valores não são significativamente positivos, o que é consistente com a hipótese considerada de não existência de capacidade de previsão. De facto, somente um fundo rejeita a hipótese nula ($p_1 + p_2 = 1$), tanto para um nível de significância de 5% como de 1%. Este fundo apresenta igualmente valores significativamente positivos para $\hat{\alpha}_p$ aquando da aplicação dos testes paramétricos.

3.2.2 — Análise para os subperíodos

De modo a avaliar a consistência dos resultados, dividimos a amostra em dois subperíodos já identificados, cujos resultados podem ser observados na tabela 5.

TABELA 5

Período	Dezembro de 1987 a Junho de 1991	Julho de 1991 a Dezembro de 1994
Estimativa média ($p1 + p2$)	1.298701	1.303571
Desvio padrão	0.218472	0.324576
Número de fundos:		
Rejeitam *Ho* a 5% (1%)	0 (0)	2 + (1 +)
Estimativa ($p1 + p2 > 1$	13	11

Testes não paramétricos: Dois subperíodos; rendibilidade do fundo *vs* estratégia passiva, hipótese nula: $p_1 + p_2 = 1$.
Nota. — + designa o número de fundos com estimativas significativamente positivas.

Os resultados mostram que, em média, a estimativa de ($p_1 + p_2$) é maior que 1 em ambos os subperíodos. Contudo, somente no segundo subperíodo existem fundos que rejeitam a hipótese de não existência de capacidade de previsão. A um nível de significância de 5% existem dois fundos e a 1% apenas um, coincidindo este com o fundo que rejeita *Ho* para o período global.

À semelhança do que fizemos nos testes paramétricos, dividimos também a amostra pela magnitude do valor absoluto da rendibilidade em excesso. Os resultados obtidos encontram-se resumidos na tabela 6, seguidamente apresentada.

[18] Devido à possibilidade de existência de autocorrelação entre os termos residuais da regressão, aplicamos o método de Durbin-Watson (1951). Dos 14 fundos, 11 aceitam a hipótese nula, ou seja, de que $\rho = 0$ e 3 são inconclusivos. Para os fundos cuja aplicação do teste foi inconclusiva, elaboramos o teste alternativo do multiplicador de Lagrange e, a um nível de significância de 5%, apenas um fundo rejeita a hipótese de não existência de autocorrelação. Quando utilizamos um nível de 1%, todos os fundos aceitam a hipótese nula, pelo que os resultados nos sugerem que, para a quase totalidade dos fundos, não existe autocorrelação.

TABELA 6

Magnitude	Menor do que a mediana	Maior do que a mediana
Estimativa média ($p1 + p2$)	1.222381	1.34643
Desvio padrão	0.19809	0.32961
Número de fundos:		
Rejeitam *Ho* a 5% (1%) ..	0 (0)	1 (1)
Estimativa ($p1 + p2 > 1$	12	11

Testes não paramétricos: Retorno dos fundos *vs.* estratégia passiva. Amostra dividida de acordo com a magnitude de $|X_{p,t}|$.

Nenhum dos fundos exibe *performances* significativamente superiores em ambos os períodos. Dos 14 fundos, 8 apresentam uma estimativa de ($p_1 + p_2$) superior para a amostra de grande amplitude de $|X_{p,t}|$ [19]. Contudo, não são valores suficientes para rejeitar a hipótese nula de não capacidade de previsão (somente um dos fundos a rejeita no caso desta última amostra). Assim, estes resultados também não mostram evidência de que os gestores são mais eficientes a prever grandes variações de mercado, em relação a pequenas variações.

4 — Conclusões

Os resultados obtidos neste trabalho sugerem-nos que, para o período de Dezembro de 1987 a Dezembro de 1994, os fundos de pensões da nossa amostra não evidenciam capacidades superiores na previsão dos preços dos títulos individuais (selectividade) nem na previsão e antecipação dos movimentos do mercado, ajustando a composição (risco) da carteira de modo a superar o mercado *(timing)*.

Quando aplicamos o modelo de Jensen (1968) verificamos que todos os fundos nos sugerem que não existe capacidade superior por parte dos gestores, dado que todos aceitam a hipótese de $\alpha_\phi = 0$. Em termos gerais, obtemos conclusões semelhantes quando aplicamos o modelo de Henriksson e Merton (1981), que detecta, separadamente, a capacidade de *timing* e a selec-tividade. De facto, considerando o período global, apenas dois fundos revelaram capacidade de selectividade significativa e um fundo apresentou uma capacidade de «timing» significativamente negativa [20]. Conclusões idênticas foram retiradas quando dividimos a amostra em dois subperíodos [21].

Uma das possíveis explicações para estes resultados, supomos poder estar relacionada com a composição dos fundos da amostra, os quais têm uma percentagem ínfima de acções. Assim, poder-se-á levantar a questão da escolha da carteira padrão *(benchmark)* apropriada. Como nos fundos de pensões temos investimentos em títulos com risco e títulos «isentos de risco», ou de relativamente pouco risco, tendo estes um peso consideravelmente maior nos fundos, assim como ainda imóveis, a carteira padrão talvez devesse consistir numa carteira ponderada deste tipo de activos. O índice que mais se aproximaria da composição dos fundos de pensões da amostra e para o período considerado, seria um índice de obrigações. Contudo, não se pode fazer a análise baseada neste índice, pois em Portugal ele não existe para a totalidade do período em questão [22].

Outras razões podem ainda explicar os nossos resultados, como sejam os elevados custos de transacção e gestão dos fundos, a diferente dimensão destes, que possibilitaria economias de escala para alguns deles, restrições de ordem legal, nomeadamente as limitações de investimento nos vários tipos de activos, a eventual existência de restrições ao nível da tomada de decisão por parte de quem gere os fundos de pensões (já que, por vezes, as decisões são tomadas a nível dos proprietários dos fundos e não dos seus gestores) e o comportamento aleatório do mercado, que dificulta a previsão dos seus movimentos.

Finalmente, interessaria investigar com profunfidade a(s) razão(ões) pela(s) qual(is) continua a observar-se uma correlação negativa entre $\hat{\alpha}_\phi$ e $\hat{\beta}_{2,\phi}$ no contexto do modelo de Henriksson e Merton (1981).

[19] Estes valores podem igualmente ser obtidos junto dos autores.

[20] Estes resultados são obtidos considerando um nível de significância de 5%. Para um nível de 1%, o número de fundos que apresenta valores significativos para a selectividade e o *timing* reduz-se para 1 e 0, respectivamente.

[21] Em ambos os subperíodos somente um fundo apresentou capacidade significativa de selectividade, não evidenciando nenhum deles capacidade de *timing,* tanto a um nível de significância de 5 como 1%.

[22] O índice de obrigações começou a ser calculado em Janeiro de 1993, e publicado somente a partir de 1994. E, mesmo assim, talvez que este índice também não fosse o adequado, pois os tipos de obrigações contidos nos fundos de pensões não são propriamente os contidos no índice.

ANEXO 1

Testes paramétricos: $R_{p,t} - R_{f,t} = \alpha_p + \beta_p (R_{m,t} - R_{f,t}) + \varepsilon_{p,t}$, considerando os dois subperíodos e ignorando a capacidade de *timing*

Período	Dezembro de 1987 a Junho de 1991	Julho de 1991 a Dezembro de 1994
Média (desvio padrão):		
$\hat{\alpha}_p$	0.0012 (0.0266)	0.0057 (0.0831)
$\hat{\beta}_p$	0.0043 (0.0464)	0.0118 (0.1444)
Número de fundos:		
Rejeitam $\hat{\alpha}_p = 0$ a 5%	0	0
Rejeitam $\hat{\alpha}_p = 0$ a 1%	0	0
$\hat{\alpha}_p > 0$	8	13

ANEXO 2

Testes paramétricos: $R_{p,t} - R_{f,t} = \alpha_p + \beta_{1,p} X_{p,t} + \beta_{2,p} Y_{p,t} + \varepsilon_{p,t}$. Amostra dividida de acordo com a magnitude de $|X_{p,t}|$

Amostra	Pequena magnitude	Grande magnitude
Média (desvio padrão):		
$\hat{\alpha}_p$	-0.0043 (0.0191)	0.0167 (0.0355)
$\hat{\beta}_{1,\phi}$	0.0311 (0.2364)	0.0093 (0.0549)
$\hat{\beta}_{2,\phi}$	0.0261 (0.3424)	-0.1022 (0.2439)
Número de fundos:		
Rejeitam $\hat{\alpha}_p = 0$ a 5%	1 +	1 +
Rejeitam $\hat{\alpha}_p = 0$ a 1%	1 +	0
Rejeitam $\hat{\beta}_{2,\phi} = 0$ a 5%	2 + 1 -	0
Rejeitam $\hat{\beta}_{2,\phi} = 0$ a 1%	0	0
$\hat{\alpha}_p > 0$	5	12
$\hat{\beta}_{2,\phi} > 0$	11	5

Nota. — + designa o número de fundos com estimativas significativamente positivas; - designa o número de fundos com estimativas significativamente negativas.

ANEXO 3

Teste de Chow à estabilidade dos parâmetros

Fundos	*SRQ1*	*SRQ2*	*SRQ3*	*F*-Chow
A	0.008567	0.005832	0.001784	0.916631
B	0.006198	0.004220	0.001403	0.749681
C	0.008456	0.005612	0.002660	0.162833
D	0.005472	0.001307	0.003767	0.574697
E	0.020294	0.008823	0.010180	0.498215
F	0.002853	0.001006	0.001568	0.794580
G	0.013261	0.010991	0.000920	0.830959
H	0.012414	0.004343	0.006028	1.444707
I	0.237665	0.019034	0.159974	2.402989
J	0.020694	0.018336	0.001896	0.167350
K	0.007558	0.005527	0.001473	0.584391
L	0.013002	0.008084	0.004471	0.297158
M	0.000038	0.000034	0.000002	0.391470
N	0.006214	0.003118	0.002507	0.768741

SRQ1: soma residual dos quadrados, considerando o período global. SRQ2: soma residual dos quadrados, considerando o primeiro subperíodo. SRQ3: soma residual dos quadrados, considerando o segundo subperíodo.

ANEXO 4

Teste de Fisher à igualdade de um subconjunto de parâmetros nas duas regressões

Fundos	SRQ_R	SRQ_U	*F*-Teste
A	0.008064	0.007615	0.393023
B	0.005632	0.005623	0.010623
C	0.008466	0.008272	0.156282
D	0.005518	0.005074	0.582600
E	0.020089	0.010003	0.380970
F	0.002716	0.002574	0.366158
G	0.012950	0.011911	0.581348
H	0.012157	0.010371	0.147911
I	0.216328	0.179008	0.389896
J	0.020467	0.020232	0.077370
K	0.008162	0.007000	0.106738
L	0.012534	0.012496	0.020295
M	0.000039	0.000036	0.551884
N	0.005986	0.005625	0.427755

SRQ_R: Soma residual dos quadrados da hipótese restritiva. SRQ_U: Soma residual dos quadrados da hipótese não restritiva.

Bibliografia

ANG, James S., e CHUA, Jess H. — «Composite Measures for the Evaluation of Investment Performance», *Journal of Financial and Quantitative Analysis*, June 1979, vol. XVI, n.º 2, pp. 361-384.

ARMADA, M. J. Rocha — «On the Investigation of Timing and Selectivity in Portfolio Management», PhD Dissertation, December 1992, Manchester Business School.

BHATTACHARYA, S., e PFLEIDERER, P. — «A Note on Performance Evaluation», *technical report*, 1983, 714, Stanford University.

BLACK, F., JENSEN, M. C., e SCHOLES, M. — «The Capital Asset Pricing Model: Some Empirical Tests», 1972, in Jensen, M. C., (ed.), *Studies in the Theory of Capital Markets*, New York, Praeger Publishers.

BROWN, Keith C., e BROWN, Gregory D. — «Does the Composition of the Market Portfolio Really Matter?», *The Journal of Portfolio Management*, Winter 1987, pp. 26-32.

CHANG, Eric C., e LEWELLEN, Wilbur, G. — «Market Timing and Mutual Fund Investment Performance», *The Journal of Business*, 1984, vol. 57, n.º 1, pp. 57-72.

CHOW, G. C. — «Tests os Equality between Sets of Coefficients in Two Linear Regressions», *Econometrica*, July 1960, vol. 28, pp. 591-605.

COGGIN, T. Daniel, FABOZZI, Frank J., e RAHMAN, Shafiqur — «The Investment Performance of U.S. Equity Pension Fund Managers: An Empirical Investigation», *The Journal of Finance*, July 1993, n.º 3, pp. 1039-1055.

DURBIN, J., e WATSON, G. S. — «Testing for Serial Correlation in Least-Squares Regression», *Biometrica*, 1951, vol. 38, pp. 159-177.

FAMA, E. — «Components of Investment Performance», *The Journal of Finance*, June 1972, vol. XXVII, n.º 3, pp. 551-567.

FERGUSON, Robert — «Performance Measurement Doesn´t Make Sense», *Financial Analysts Journal*, May-June 1980, pp. 59-69.

—— «The Trouble With Performance Measurement», *The Journal of Portfolio*, Spring 1986, pp. 4-9.

FISHER, F.M. — «Tests of Equality Between Sets of Coefficients in Two Linear Regressions: An Expository Note», *Econometrica*, March 1970, 38, pp. 361-366.

FRIEND, Irwin, e BLUME, Marshall — «Measurement of Portfolio Performance Under Uncertainty», *The American Economic Review*, September 1970, vol. 70, n.º 4, pp. 561-575.

GAUMNITZ, J. — «Appraising Performance of Investment Portfolios», *Journal of Finance*, June 1970, vol. 25, n.º 3, pp. 555-560.

GRANT, Dwight — «Portfolio Performance and the 'Cost' of Timing Decisions», *The Journal of Finance*, June 1977, vol. XXXII, n.º 3, pp. 837-846.

—— «Market Timing and Portfolio Management», *The Journal of Finance*, September 1978, vol. XXXIII, n.º 4, pp. 1119-1131.

HAGIGI, Moshe, e KLUGER, Brian — «Assessing Risk and Return of Pension Funds' Portfolios by the Telser Safety-First Approach», *Journal of Business Finance & Accounting*, Summer 1987, vol. 14, n.º 2, pp. 241-253.

HENRIKSSON, Roy D. — «Market Timing and Mutual Fund Performance: An Empirical Investigation», *The Journal of Business*, 1984, vol. 57, n.º 1, pp. 73-96.

HENRIKSSON, Roy D., e MERTON, Robert, C. — «On Market Timing and Investment Performance. II — Statistical Procedures for Evaluating Forecasting Skills», *The Journal of Business*, 1981, vol. 54, n.º 4, pp. 513-533.

JENSEN, Michael C. — «Problems in Selection of Security Portfolios — The Performance of Mutual Funds in the Period 1945--1964», *The Journal of Finance*, May 1968, nº23, pp. 389-416.

—— «Optimal Utilization of Market Forecasts and the Evaluation of Investment Performance», *Mathematical Models in Investment and Finance*, 1972, by Szego e Shell, North-Holland Press, pp. 310-335.

KANE, Alex, e MARKS, Stephen Gary — «Performance Evaluation of Market Timers: Theory and Evidence», *Journal of Financial and Quantitative Analysis*, December 1988, vol. 23, n.º 4, pp. 425-435.

KLEMKOSKY, R. — «The Bias in Composite Performance Measures», *Journal of Financial and Quantitative Analysis*, June 1973, pp. 505-514.

KOLB, Robert W., e RODRIGUEZ, Ricardo J. — «Is the Distribution of Betas Stationary?» *The Journal of Financial Research*, 1990, vol. XIII, n.º 4, Winter, pp. 279-283.

KON, Stanley J., e JEN, Frank C. — «Estimation of Time-varying Systematic Risk and Performance for Mutual Fund Portfolios: An Application of Swittching Regression», *The Journal of Finance*, May 1978, vol. XXXIII, n.º 2, pp. 457-475.

—— «The Investment Performance of Mutual Funds: An Empirical Investigation of Timing, Selectivity and Market Efficiency», *The Journal of Business*, April 1979, vol. 52, n.º 2, pp. 263-289.

LEHAMANN, B., e MODEST, D. — «Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons», *The Jounal of Finance*, 1987, n.º 42, pp. 233-265.

LEVY, Haim — «Portfolio Performance and the Investment Horizon», *Management Science*, August 1972, vol. 18, n.º 12, pp. 645-653.

—— «The CAPM and the Investment Horizon», *The Journal of Portfolio Management*, Winter 1981, pp. 32-40.

—— «Measuring Risk and Performance Over Alternative Investment Horizons», *Financial Analysts Journal*, March-April 1984, pp. 61-68.

LINTNER, J. — «The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets», *Review of Economics and Statistics*, February 1965, vol. 47, n.º 1, pp. 13-37.

MARKOWITZ, H. M. — «Portfolio Selection», *The Journal of Finance*, March 1952, vol. 8, n.º 1, pp. 77-91.

McDONALD, J. — «Objectives and Performance of Mutual Funds», *Journal of Financial and Quantitative Analysis*, June 1974, pp. 311-333.

MERTON, Robert, C. — «On Market Timing and Investment Performance. I. An Equilibrium Theory of Value for Market Forecasts», *The Journal of Business*, 1981, vol. 54, n.º 3, pp. 363-406.

MOSSIN, J. — «Equilibrium in a Capital Asset Market», *Econometrica*, October 1966, vol. XXXIV, pp. 768-783.

PINK, George H. — «Market Timing by Canadian Bond Funds», *Journal of Business Administration*, 1989, vol. 18, n.os 1 e 2, pp. 165-181.

ROLL, Richard — «Ambiguity When Performance is Measured by the Security Market Line», *The Journal of Finance*, September 1978, vol. XXXIII, n.º 4, pp. 1051-1069.
SHARPE, William — «Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk», *The Journal of Finance*, September 1964, vol. 19, n.º 3, pp. 425-444.
—— «Mutual Fund Performance», *The Journal of Business*, January 1966, vol. 39, n.º 1, pp. 119-138.
TELSER, L. G. — «Safety First and Hedging», *The Review of Economic Studies*, 1956, vol. XXIII, pp. 1-16.
TREYNOR, Jack L. — «How to Rate Management of Investment Funds», *Harvard Business Review*, January-February 1965, pp. 63-75.
TREYNOR, Jack L., e MAZUY, K. — «Can Mutual Funds Outguess the Market?», *Harvard Business Review*, July-August 1966, pp. 131-136.
WHITE, H. — «A Heteroscedasticity-Consistent Covariance Matrix Estimation and a Direct Test for Heteroscedasticity», *Econometrica*, May 1980, vol. 48, n.º 4, pp 817-838.