



Revista Brasileira de Finanças

ISSN: 1679-0731

rbfin@fgv.br

Sociedade Brasileira de Finanças

Brasil

Wu, Xiaoyan; Salazar Soares, Vasco; Dias Pacheco, Luís; Oliveira Tavares, Fernando  
Medidas de Avaliação de Desempenho: Os Fundos de Ações Europeias no Período 2001  
a 2015

Revista Brasileira de Finanças, vol. 15, núm. 3, julio-septiembre, 2017, pp. 403-433  
Sociedade Brasileira de Finanças  
Rio de Janeiro, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=305855644003>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

 redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal  
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# **Medidas de Avaliação de Desempenho: Os Fundos de Ações Europeias no Período 2001 a 2015**

**(Performance Measures: European Mutual Funds in 2001-2015)**

**Xiaoyan Wu\***

**Vasco Salazar Soares\*\***

**Luís Dias Pacheco\*\*\***

**Fernando Oliveira Tavares#**

## **Resumo**

As medidas de avaliação de desempenho dos fundos de investimento permitem estabelecer um ranking e têm um papel essencial para que os investidores tomem decisões de investimento. A escolha da medida indicada deve ter em consideração a preferência de risco por parte de investidores. Neste trabalho comparam-se as diferentes medidas ajustadas ao risco, tais como: índices de Sharpe, MM, Treynor, Jensen, RORAC, Informação, IVaR, Sortino, UPR, Omega, Fouse e Alfa de Sharpe. Os objetivos deste trabalho passam por testar a consistência entre as medidas, estudar a correlação das medidas com as preferências de risco e procurar as medidas que simulam o *Morningstar rating*. Para efetuar este estudo foram obtidos dados da base de dados da *Yahoo finance*, sendo que a amostra é constituída por 28 fundos de ações europeias, entre setembro de 2001 e setembro de 2015. Conclui-se que as medidas produzem rankings idênticos, pelo que existe uma correlação elevada entre elas, excluindo RORAC e o Índice de Informação. Este estudo mostra que a classificação dos fundos de investimento varia consoante o indicador possa assumir ou não elevada volatilidade de rendimentos. Os índices de Sharpe e MM evidenciaram maior capacidade explicativa em relação ao

---

Submetido em 13 de novembro de 2016. Reformulado em 26 de junho de 2017.

Aceito em 11 de março de 2018. Publicado on-line em 19 de junho de 2018. O artigo foi avaliado segundo o processo de duplo anonimato além de ser avaliado pelo editor. Editor responsável: Marcio Laurini.

\* Universidade Portucalense Infante D. Henrique. Porto, Portugal. E-mail:  
Lucia0088@msn.com

\*\* ISVOUGA- Instituto Superior de Entre Douro e Vouga. Santa Maria da Feira,  
Portugal. E-mail: v.soares@doc.isvouga.pt

\*\*\* Universidade Portucalense Infante D. Henrique Porto, Portugal. E-mail:  
luisp@upt.pt  
# Universidade Portucalense Infante D. Henrique. Porto, Portugal. E-mail:  
ftavares@upt.pt

*Rev. Bras. Finanças (Online), Rio de Janeiro, Vol. 15, No. 3, September 2017, pp. 403-433*  
ISSN 1679-0731, ISSN online 1984-5146

©2017 Sociedade Brasileira de Finanças, under a Creative Commons Attribution 3.0 license -  
<http://creativecommons.org/licenses/by/3.0>

*Morningstar rating* e apresentam também uma correlação elevada com a função utilidade quadrática e com a função valor da teoria da perspectiva. O grau de correlação diminui à medida que aumenta o coeficiente de aversão ao risco, sendo que as medidas baseadas no *downside risk* (índice de Sortino e Omega) mantêm uma boa correlação com todos os níveis de preferência, tanto no período global como nos subperíodos.

**Keywords:** Medidas ajustadas ao risco, fundos de investimento, preferência de risco, ranking de desempenho, *Morningstar ratings*.

**JEL Codes:** F65, G11, G12, G14, G15, N24.

### **Abstract**

Mutual funds performance evaluation measures allow to establish rankings and play an essential role for investors who want to make investment decisions. The choice of suitable measure should take into account the risk preference of investors. This paper compares the different risk-adjusted measures, such Sharpe, MM, Treynor, Jensen, RORAC, information, Ivar, Sortino, UPR, Omega, Fouse and Alfa Sharpe indexes. The objectives of this study are to test the consistency between the measurements, to study the correlation between measures and the risk preferences and look for measures that simulate the Morningstar rating. To perform this study the data was obtained at Yahoo finance, with the sample consisting of 28 funds of European shares, from September 2001 to September 2015. It was concluded that the measures produce identical rankings, so there is a high correlation between them excluding RORAC and the Information indexes. This study shows that the classification of investment funds varies depending on whether the indicator assumes high income volatility. Sharpe and MM indexes showed greater explanatory power in relation to Morningstar rating and have a high correlation with the quadratic utility function and the value function from prospect theory. The degree of correlation decreases as increases the risk aversion coefficient, with measures based on downside risk (Sortino ratio and omega) maintaining a good correlation with all preference levels, both in the overall period and in sub-periods.

**Keywords:** Risk-adjusted measures, Investment funds, Risk preference, Performance rankings, *Morningstar ratings*.



## 1. Introdução

A indústria de fundos de investimento tem assistido a um crescimento global intenso ao longo das últimas décadas. No final de junho de 2015, o valor líquido global cifrou-se em 34.216 mil milhões de euros e quase 43,72% deste valor está associado a fundos em ações. Destes fundos, 53% são geridos nos Estados Unidos o que evidencia uma elevada concentração ([www.iifa.ca](http://www.iifa.ca)).

Esta crescente importância dos fundos de investimento aumentou o interesse pela avaliação do seu desempenho, pelo que foram desenvolvidas medidas de avaliação de desempenho que permitem estabelecer rankings dos fundos e selecionar as oportunidades de investimento. Diversos estudos não encontraram convergência entre rankings de fundos calculados de acordo com várias medidas (Plantinga e Groot, 2001). Por exemplo, Eling e Schuhmacher (2007) concluem que existe convergência entre os rankings produzidos por várias medidas e identificam o índice de Sharpe como exibindo superioridade no estabelecimento de rankings.

Os métodos de classificação diferem fundamentalmente em razão de formas alternativas para as medidas de retorno e risco, mas também quanto ao modo de ajuste para o risco. Por outro lado, as medidas de desempenho ajustadas ao risco supõem que os investidores são avessos ao risco e necessitam de uma compensação pela sua exposição ao risco. Neste sentido, o principal contributo deste artigo, consiste em verificar até que ponto a classificação dos fundos de investimento varia consoante o indicador assume elevada volatilidade dos rendimentos ou não.

O presente artigo realiza uma análise comparativa das diferentes medidas de desempenho ajustadas ao risco. A análise incide sobre o desempenho dos fundos de ações europeias no período de 2001 a 2015, procurando verificar, através de testes estatísticos, o nível de consistência entre os índices de avaliação de desempenho de uma amostra significativa de fundos de investimento de ações. Na próxima secção faz-se uma revisão da literatura sobre as medidas de avaliação de desempenho, sendo apresentadas as principais vantagens e desvantagens de cada indicador e as hipóteses que serão testadas neste artigo. Na secção 3 são apresentados os



dados e na secção 4 apresentam-se e analisam-se os resultados obtidos. A última secção apresenta as principais conclusões.

## 2. Medidas De Avaliação Do Desempenho

### 2.1. Referencial teórico

A literatura apresenta uma diversidade de medidas de avaliação do desempenho de carteiras, podendo as mesmas ser classificadas nas seguintes categorias: i) medidas clássicas ajustadas ao risco; ii) extensões das medidas clássicas; iii) medidas de *market timing*; iv) medidas baseadas em riscos assimétricos; v) medidas baseadas em VaR; e vi) medidas baseadas na preferência. A Tabela 1 apresenta uma lista não exaustiva de algumas das principais medidas descritas na literatura.

As medidas tradicionais têm sobretudo a vantagem de serem de fácil interpretação e de serem ainda amplamente utilizadas, mas recebem críticas negativas pelas suas limitações (Friend e Blume, 1970; Levy, 1981; Roll, 1980; Lee e Jen, 1978; Ferguson, 1980 e 1986; Brown e Brown, 1987; Milan e Júnior, 2014; Matos, Silva e Silva, 2015). Os rácios modificados corrigem algumas anomalias, permitindo classificar de forma consistente os fundos. As medidas baseadas no *market timing* também sofrem de limitações relevantes, por não considerarem a variabilidade do risco ao longo do tempo, pelo que a especificação de Ferson e Schadt (1996) considera a variabilidade do risco ao longo do tempo, em função de variáveis de informações públicas que são utilizadas para captar o estado da economia. Na medida em que os indicadores com base no desvio padrão não permitem saber se as diferenças em comparação com a média foram produzidas acima ou abaixo desta foram introduzidas as medidas baseadas em riscos assimétricos. A noção de *downside risk* traz uma solução para este problema, tendo em consideração o risco de assimetria.



**Tabela 1:** Principais medidas de avaliação do desempenho de carteiras:

Designação	Fórmula de cálculo	Autor(es) e ano
<b>i) Medidas clássicas ajustadas ao risco</b>		
Índice de Sharpe	$\text{índice de Sharpe} = \frac{E(R_p) - R_f}{\sigma(R_p)}$	Sharpe (1966)
Índice de Treynor	$\text{índice de Treynor} = \frac{E(R_p) - R_f}{\beta(R_p)}$	Treynor (1965)
Alfa de Jensen	$\text{alfa de Jensen} = E(R_p) - [R_f + \beta * (E(R_m) - R_f)]$	Jensen (1968)
<b>ii) Extensões das medidas clássicas</b>		
Medida de Modigliani e Modigliani	$\text{índice de MM} = \frac{\sigma R_m}{\sigma R_p} E(R_p) + \left(1 - \frac{\sigma R_m}{\sigma R_p}\right) R_f$	Modigliani e Modigliani (1997)
Índice de informação	$\text{índice de informação} = \frac{E(R_p) - E(R_m)}{\sigma R_p - \sigma R_m}$	Grinold (1989)
Índice de Sharpe modificado	$\begin{aligned} \text{índice de sharpe modificado} \\ = \frac{E(R_p) - E(R_f)}{\sigma(R_p)^{(E(R_p) - E(R_f)) / (\text{abs}(E(R_p) - E(R_f))}} \end{aligned}$	Israelsen (2005)
Índice de informação modificado	$\begin{aligned} \text{índice de infor.mod.} \\ = \frac{E(R_p) - E(R_f)}{\sigma(R_p - R_m)^{(E(R_p) - E(R_f)) / (\text{abs}(E(R_p) - E(R_f))}} \end{aligned}$	Israelsen (2005)
<b>iii) Medidas de market timing</b>		
Modelo de Treynor e Mazuy	$E(R_p) - E(R_f) = \alpha + \beta_1[E(R_m) - E(R_f)] + \beta_2[E(R_m) - E(R_f)]^2$	Treynor e Mazuy (1966)
Modelo de Henriksson e Merton	$E(R_p) - E(R_f) = \alpha + \beta_1 [E(R_m) - E(R_f)] + \beta_2[\max(0, R_f - R_m)]$	Henriksson e Merton (1981)
Modelo de três fatores de Fama e French	$E(R_p) - E(R_f) = \alpha + \beta_1[E(R_m) - E(R_f)] + \beta_2(SMB) + \beta_3(HML)$	Fama e French (1993)
Modelo de quatro fatores de Carhart	$E(R_p) - E(R_f) = \alpha + \beta_1[E(R_m) - E(R_f)] + \beta_2(SMB) + \beta_3(HML) + \beta_4(MOM)$	Carhart (1997)
Medida de Ferson e Schadt <sup>(1)</sup>	$E(R_p) - E(R_f) = \alpha + \beta_1[E(R_m) - E(R_f)]^2 + \beta_2/Z_{1/[E(R_m) - E(R_f)]}$	Ferson e Schadt (1996)
<b>iv) Medidas baseadas em riscos assimétricos</b>		
Índice de Sortino	$\text{índice de sortino} = \frac{E(R_p) - R_{ma}}{\text{Downside risk}}$	Sortino e Price (1994)
UpSide Potencial Ratio (UPR) <sup>(2)</sup>	$UPR = \frac{\sum_{i=1}^m i + \min[0; (R_p - R_{ma})]/m}{\sum_{i=1}^m i - \min[0; (R_p - R_{ma})]^2/m}$	Sortino e Plantinga (1999)
Índice de Omega <sup>(3)</sup>	$\text{índice de Omega} = \frac{\int_a^b (1 - F(x)) dx}{\int_a^{rma} F(x) dx}$	Keating e Shadwick (2002)
<b>v) Medidas baseadas em VaR</b>		
Índice de Dowd	$\text{índice baseado no VaR} = \frac{E(R_p) - rf}{VaR\%}$	Dowd (1999)
Return on Risk Adjusted Capital (RORAC)	$\text{Return on Risk Adjusted Capital} = \frac{E(R_p)}{VaR\%}$	Bankers Trust (década de 1970)

**vi) Medidas baseadas na preferência**

Alfa de Sharpe	$\text{Alfa de sharpe} = E(R_p) - R_f - A * \sigma^2$	Plantinga e Groot (2001)
Índice de Fouse	$\text{Índice de Fouse} = E(R_p) - A * (dk)^2$	Sortino e Price (1994)

NOTAS: Elaborado pelos autores

 $E(R_p)$  = taxa de retorno esperada da carteira p $R_f$  = taxa de retorno do ativo sem risco $\sigma(R_p)$  = desvio padrão da taxa de retorno da carteira p $\beta(R_p)$  = beta da carteira p $E(R_m)$  = taxa de retorno esperada da carteira de mercado $\sigma R_m$  = desvio-padrão (risco) da taxa de retorno da carteira de mercado $\sigma R_p - \sigma R_m$  = tracking error $E(R_p) - E(R_f)$  = prémio de risco em valor absoluto $\beta_i$  = coeficientes de sensibilidade da carteira a cada um dos fatores de risco $SMB$  (*Small minus big*) = diferença entre o retorno de uma carteira de ações com baixa capitalização e com elevada capitalização $HML$  (*High minus low*) = diferença entre o retorno de uma carteira de ações com elevado *book-to-market* e com baixo *book-to-market* $MOM$  (*momentum*) = diferença entre o retorno de uma carteira de ações com os retornos mais elevados e mais baixos nos últimos onze meses, com um mês desfasado<sup>(1)</sup>  $\beta_1$  = Beta condicional médio, que representa a média dos betas condicionais $\beta_2$  = vetor que quantifica a relação entre o beta condicional e as variáveis de informação $Z_{t-1}$  = vetor dos desvios de  $Z_{t-1}$  em relação aos seus valores médios, ou seja,  $Z_{t-1} - E(Z)$ <sup>(2)</sup> m = número de períodos na amostra $R_p$  = retorno de um investimento num período $R_{ma}$  = taxa de retorno mínima aceitável pelo investidor $i^+$  = quando  $R_p > R_{ma}$  $i^-$  = quando  $R_p \leq R_{ma}$ <sup>(3)</sup> X = distribuição sob análise $F(x)$  = distribuição de frequências acumuladas de X

VaR % = VaR do portfólio medido em termos percentuais

A = parâmetro que representa o nível de aversão ao risco

 $Dk^2$  = semi-variância

Uma importante vantagem de utilizar o *Upside Potential Ratio* (UPR) em vez do índice de Sortino é a consistência na utilização de uma taxa de referência para a avaliação tanto nos lucros como nas perdas. O UPR oferece, portanto, o desempenho que se adeque aos desejos dos investidores: obter aumento potencial enquanto protege contra perdas.

O CAPM pressupõe que todos os retornos de ativos ou carteira de ativos têm distribuição normal e simétrica, sendo que os investidores têm preferências de média-variância, mas ignoram a assimetria e curtoza. Leland (1999) demonstra que o CAPM e as suas medidas de risco são inválidas, para além de que o alfa do CAPM não estima corretamente o valor acrescentado por gestores de investimento, tornando a carteira de



mercado de média-variância ineficiente. Entretanto, é de notar que uma função quadrática é inconsistente com o comportamento do investidor e que impede a distribuição dos retornos dos fundos de ser normalmente distribuída. Para suprir essa lacuna, foi introduzido por Keating e Shadwick (2002) uma nova medida, denominada índice de Omega. Esta medida tem em consideração tanto o terceiro quanto o quarto momento da distribuição, assimetria e curtose, respectivamente. No caso de avaliar os portfólios com distribuições normais, a otimização pela Média-Variância fornece os mesmos resultados do que a otimização através da maximização da medida Omega, porque nesses casos as distribuições são definidas apenas pelos dois primeiros momentos. Portanto, Omega caracteriza-se por ser mais geral e aplicável a qualquer situação. A métrica Omega tem duas grandes vantagens sobre as medidas tradicionais: i) capta todas as informações sobre o risco e retornos de uma carteira; ii) tem em conta o perfil de risco de cada investidor. Numa escolha entre carteiras com o mesmo retorno previsto, os investidores devem favorecer a carteira com o maior índice de Omega. Isto maximiza o potencial do nível de retorno desejado, e minimiza a probabilidade de perdas extremas.

Finalmente, a maioria das medidas anteriores pressupõe que a função utilidade dos investidores é côncava, exibindo assim uma atitude de aversão ao risco. No entanto, o comportamento implícito nessas funções de utilidade pode ser inconsistente na realidade. Por exemplo, Tversky e Kahneman (1992) desenvolveram uma função valor que captura as principais características da teoria da perspectiva de tomada de decisões sob incerteza (Kahneman e Tversky, 1979). Aqui utilizaremos uma versão linear da função valor, concretamente, uma função bilinear que é composta por dois segmentos lineares que distinguem perdas e ganhos. Tendo a taxa livre de risco como um ponto de referência, a perda é determinada pelo retorno abaixo de  $r_f$ , e acima desse ponto de referência considera-se o ganho. Constata-se uma maior inclinação para a esquerda de  $r_f$  do que para a direita. Tal função traduz que a aversão ao risco é maior quando está associado às perdas. O modelo de alfa Sharpe é uma derivação do índice de Sharpe mencionado por Plantinga e Groot (2001). A medida é frequentemente utilizada como uma representação alternativa da função utilidade quadrática. Esta medida tem em consideração o perfil dos investidores o que é uma vantagem, mas também uma desvantagem,



porque dois investidores terão dois rankings diferentes, por isso é difícil comparar a qualidade desta medida com outra.

## 2.2. Hipóteses a testar

O objetivo fundamental deste artigo é apresentar uma análise do desempenho dos fundos e o ranking daí resultante. O desempenho dos fundos é calculado segundo as equações apresentadas no ponto anterior, sendo estabelecidos os rankings dos fundos. Entre as medidas de avaliação de desempenho, são escolhidas doze medidas ajustadas ao risco para examinar as correlações entre elas. A classificação de estrelas da *Morningstar* é seguida por muitos investidores, pois possui um poder de influência na decisão de investimento. Torna-se assim relevante verificar a consistência da classificação da *Morningstar* com as medidas aplicadas neste estudo.

A metodologia utilizada foi escolhida tendo presente o objetivo de proceder a uma análise comparativa das medidas de avaliação de desempenho dos fundos. Consequentemente, é formulado o seguinte conjunto de seis hipóteses:

H1: a correlação entre as medidas de desempenho que assumem retornos com baixa dispersão e as medidas de desempenho de elevada volatilidade é nula;

H2: existe uma relação entre índices de desempenho ajustados pelo risco, em que os modelos de avaliação de desempenho são substitutos uns aos outros, sem prejuízo de análise;

H3: de acordo com as medidas de avaliação de desempenho dos fundos, aumentando a volatilidade aumenta o retorno ajustado ao risco;

H4: considerando que a *Morningstar rating* segue uma função utilidade, será de esperar que exiba uma maior correlação com as medidas de Fouse e alfa de Sharpe;

H5: os índices de desempenho considerados conseguem explicar as classificações de risco da *Morningstar*;

H6: existe correlação entre a função utilidade esperada e as medidas de avaliação de desempenho, verificando-se que o



aumento do coeficiente de aversão ao risco de 1 para 5, mantém a correlação positiva;

H7: a correlação entre as medidas de avaliação de desempenho e a função bilinear diminui, mas mantém-se positiva, à medida que aumenta o coeficiente de aversão ao risco.

À luz do estudo de Plantinga e Groot (2001), é replicada a análise e, para além das medidas aplicadas por aqueles autores - índices de Sharpe, alpha de Sharpe, Sortino, Fouse e *Upside Potential Ratio* - vamos acrescentar mais medidas ajustados ao risco, nomeadamente: índices de Treynor, Jensen, Modigliani-Modigliani (MM), rácio de informação (IF), Omega, RORAC e IVaR. É igualmente testada a correlação entre as diferentes medidas com a função utilidade quadrática. O mesmo teste é efetuado com a função valor de Tversky e Kahneman (1992), com diferentes valores para o coeficiente de aversão ao risco. Uma vez que os investidores se comportam de forma distinta quando o retorno de um fundo se situa acima e abaixo da taxa sem risco ( $r_f$ ), é calculada uma função bilinear para diferentes níveis de aversão ao risco.

### **3. Descrição Dos Dados**

A amostra utilizada neste estudo é constituída por 28 fundos mútuos de ações europeias, com um histórico operacional de, pelo menos, 10 anos. A base de dados foi obtida através de *Yahoo Finance*, para o período compreendido entre setembro de 2001 e setembro de 2015, correspondendo a um total de 170 dados mensais. Como a seguir é referido, este período caracteriza-se por diversos ciclos de expansão/contração, o que aumenta o interesse da análise. Refira-se também que diversos estudos similares ao aqui desenvolvido recorreram aquela base de dados (e.g., Lin e Yung, 2004; Wong e Shum, 2010; Sabbaghi, 2011). No âmbito da classe *european equity mutual funds*, os fundos selecionados foram aqueles que apresentavam um período mais longo de dados disponíveis. A base de dados é de frequência mensal, sendo que uma frequência demasiada alta nem sempre se traduz em melhores resultados, tornando as informações extraídas difíceis de interpretar. Vários estudos ilustram que a imperfeição

dos pressupostos em relação aos dados de retorno trimestral ou mensal é pequena, enquanto que para dados diários estes pressupostos são rejeitados (Sourd, 2007).

Os retornos foram calculados baseados nos preços ajustados aos dividendos. No cálculo do retorno mensal é considerado o valor das unidades de participação em USD. Este retorno mensal é calculado através da seguinte expressão:

$$Rp.t = \frac{Upt + dt}{Upt - 1} - 1 \quad (1)$$

Rp.t= retorno do fundo p no mês t;

Upt = valor da unidade de participação no mês t;

Upt-1 = valor da unidade de participação no mês t-1;

dt= dividendos.

Para o cálculo do retorno do mercado é utilizado o índice *Eurostoxx 50* com o mesmo período de observação e a mesma fórmula de cálculo. Refira-se que o período analisado é caracterizado por: i) uma primeira fase, entre 2001 a 2003, onde se registou uma queda acentuada no mercado acionista, influenciada em grande medida pelo rebentar da bolha das empresas tecnológicas; ii) uma segunda fase, entre 2004 e 2008, caracterizada pela melhoria do sentimento dos investidores e empresas, alimentada pela desregulamentação dos mercados e forte crescimento do crédito, que culminou com a crise do *subprime*; iii) como consequência, entra-se numa terceira fase caracterizada pela “Grande Recessão”, com a necessidade da intervenção direta dos governos para salvar bancos e outras instituições financeiras de grande dimensão, prolongando-se os efeitos negativos no mercado até 2011 e, finalmente, iv) uma última fase com algum crescimento relativo nos mercados acionistas.

A análise do desempenho dos fundos é realizada sobre o período global e três subperíodos (2001 a 2003, 2004 a 2008 e 2009 a 2015), com objetivos de evidenciar diferenças entre esses subperíodos e o impacto de crise no desempenho dos fundos. Como aproximação à taxa isenta de risco é utilizada a *yield* das obrigações alemãs de 10 anos. A taxa equivalente mensal foi calculada para servir de base à taxa sem risco (Tabela 2).



**Tabela 2:** Retorno esperado média mensal do Eurostock e do ativo sem risco

	2001-2015	2009-2015	2004-2008	2001-2003
Rm	0,10%	0,40%	-0,11%	-0,36%
Rf	0,25%	0,18%	0,32%	0,36%

Fonte: Elaboração própria com dados obtidos através de *Yahoo finance*(<http://finance.yahoo.com>)

A Tabela 3 apresenta uma descrição dos fundos constantes da amostra.

Apesar desses resultados não serem apresentados, o retorno médio mensal alcançado pelos fundos é de 0,75%, com valores máximos e mínimos no período de, respectivamente, 1,42% e 0,34% e um erro-padrão de 0,27%. De uma forma geral, os fundos alcançaram o melhor retorno no período entre 2001 a 2003, por outro lado, os retornos dos fundos são fortemente influenciados pela queda acentuada das cotações provocadas pela crise econômica global de 2008, com uma levea recuperação no período seguinte. Com o teste de Kruskal-Wallis prova-se a hipótese de que as distribuições de retornos entre diferentes períodos são diferentes. O resultado é com 95% de intervalo de confiança e estaticamente significativo, pois valor de p é muito inferior a 5%. Como as medidas de avaliação do desempenho com base na média-variância que serão utilizadas requerem, implicitamente, que os retornos dos fundos se distribuam normalmente, assume extrema importância testar a distribuição de retornos dos fundos. Para o efeito, utilizamos o teste de Kolmogorov-Smirnov, para um nível de significância de 5%, cujos resultados evidenciam que, no período global, a maioria dos fundos segue uma distribuição normal. Quanto à simetria da distribuição, os resultados apontam para a presença de assimetria negativa, ou seja, a cauda da distribuição é alongada à esquerda. Lembramos que, no período de 2004-2008, os retornos dos fundos foram mais baixos do que outros períodos, logo, é de esperar que o coeficiente de assimetria neste período se encontre mais baixo do que em outros períodos. Quanto à curtose, a maioria das distribuições apresenta valores abaixo de três, pelo que temos uma



distribuição platicúrtica. Apenas no período entre 2004-2008, dada a forte queda de 2008, é que a distribuição é leptocúrtica.



**Tabela 3:** Síntese das características dos fundos presentes na amostra

Fundo	Ações europeias	Ativos líquidos (milhões USD)	Data de criação do fundo	Rendibilidade anual (year-to-date)	Rating Morningstar	Rating de risco Morningstar
AEDAX	Invesco European Growth A	1520	03/11/97	2.80%	3	1
AEDBX	Invesco European Growth B	1520	03/11/97	2.28%	3	2
AEDCX	Invesco European Growth C	1520	03/11/97	2.31%	4	2
AEBBX	Columbia European Equity B	577	26/06/00	2.71%	4	3
ESMAX	Invesco European Small Company A	442	31/08/00	5.50%	4	3
EUGAX	Morgan Stanley European Equity A	148	28/07/97	-0.06%	4	2
EUGBX	Morgan Stanley European Equity B	148	28/07/97	-0.06%	3	2
EUOCX	Morgan Stanley European Equity C	148	28/07/97	-0.41%	2	2
EUGDX	Morgan Stanley European Equity D	148	28/07/97	0.16%	3	3
EUGIX	Shelton European Growth & Income Direct	10	18/01/00	-0.69%	2	3
FIEOX	Fidelity Europe	1450	01/10/86	3.97%	3	3
HFEAX	Henderson European Focus A	3540	31/08/01	4.92%	4	5
HFEBX	Henderson European Focus B	3540	31/08/01	4.33%	3	5
IEOAX	Ivy European Opportunities A	213	04/05/99	3.43%	2	2
IFEXX	JP Morgan Intelligent European Sel	1000	10/09/01	3.03%	4	4
MAEFX	Black Rock Euro Fund Instl	458	26/10/88	2.70%	3	4
MCERFX	Black Rock Euro Fund Inv C	458	21/10/94	1.98%	2	3
MDERFX	Black Rock Euro Fund Inv D	458	21/10/94	2.54%	2	4
MEURFX	Franklin Mutual European Z	3290	03/07/96	2.78%	5	1
PRESSX	T. Rowe Price European Stock	1620	28/02/90	2.27%	4	3
RIECX	Columbia European Equity C	577	26/06/00	2.92%	4	2
TEMIX	Franklin Mutual European A	3290	01/11/95	2.61%	4	1
TEURX	Franklin Mutual European C	3290	01/11/95	2.11%	5	1
UEPFX	Pro Funds Europe 30 Inv	4690	15/03/99	-7.56%	1	3
UEPSSX	Pro Funds Europe 30 Svc	4690	15/03/99	-7.56%	1	3
VESIX	Vanguard European Stock Index I	20810	15/05/00	0.11%	3	3
VEURX	Vanguard European Stock Index Inv	20810	18/06/00	0.00%	3	3
VEUSX	Vanguard European Stock Index Adm	20810	13/08/01	0.08%	3	3

Fonte: Elaboração própria com informações obtidas através de Yahoo finance (<http://finance.yahoo.com/>)

## 4. Resultados

#### 4.1. Desempenho e ranking dos fundos

Na presente secção, estuda-se o desempenho dos fundos segundo um conjunto de doze medidas ajustadas ao risco, nomeadamente: Sharpe, Alfa de Sharpe, Treynor, Jensen, Modigliani-Modigliani (MM), Índice de informação (IF), RORAC, IVaR, Sortino, UPR, Omega e índice de Fouse. Posteriormente, é estabelecido o ranking dos fundos de acordo com os desempenhos tendo o período global como referência (por economia de espaço, apenas são apresentados os dados para o período global).

A Tabela 4 apresenta, respetivamente, o desempenho dos 28 fundos no período 2001-2015 e o ranking daí resultante (onde se evidenciam os cinco primeiros classificados).

Note-se que, de um modo geral, o desempenho dos fundos foi muito abalado pela crise de 2008, acabando muitos destes fundos por apresentar retornos negativos, sendo o retorno médio inferior à taxa de retorno das obrigações.

De acordo com o índice de Jensen, os fundos obtiveram um valor positivo, evidenciando um desempenho superior face ao retorno estimado pelo CAPM. No entanto, para alguns fundos que obtiveram resultados negativos, significa que esses gestores não conseguiram obter retornos “anormais”, ou seja, para além do retorno esperado teórico do mercado em equilíbrio. Em geral, os fundos demonstraram melhor desempenho no primeiro período, 2001-2003 e foram quase todos afetados pela crise, mostrando uma recuperação a partir de 2009. O valor do índice de Modigliani (MM) indica o excesso de retorno do fundo face ao *benchmark*, após ajustamento das diferenças no nível de risco total, revelando valores positivos durante todos os períodos sob observação, embora no período 2004-2008 evidencie um desempenho mais fraco. Este índice é proporcional ao índice de Sharpe, pelo que o ranking produzido é naturalmente idêntico.

O índice de informação (IF) dá-nos o rácio da diferença entre o retorno do fundo e o retorno do *benchmark* pelo *tracking error*. É muitas



vezes usado para medir a capacidade dos gestores de fundos, pela sua estratégia de gestão, quanto ao valor que adicionam a alfa. Um elevado valor deste rácio pode ser explicado por um elevado retorno do fundo, por um retorno reduzido *do benchmark* ou um *tracking error* baixo. Este índice sofre a mesma limitação que o de Sharpe quando apresenta valores negativos.

Refira-se que, no primeiro período, todos os fundos obtiveram valores negativos, o que aconteceu com quase metade dos fundos no período global. No estudo de Israelsen (2005), é evidenciado que, no ranking segundo o índice de informação, os fundos com maior *tracking error* e menor retorno excedente são colocados acima da maioria dos fundos. É o que acontece no período de 2001-2003, portanto sendo muito diferente dos outros períodos. É também de esperar que os rankings sejam muito diferentes comparativamente com as outras medidas.

A medida *Return on Risk Adjusted Capital* (RORAC) é calculada com base no VaR. Segundo o teste de normalidade, percebeu-se que nem todos os retornos dos fundos seguem uma distribuição normal. Neste sentido, calcular o VaR com a simulação analítica, violaria o pressuposto de distribuição normal dos retornos, pelo que a estimativa de VaR foi obtida por simulação histórica para contornar o problema. Os resultados indicam-nos o retorno ajustado à máxima perda no intervalo de confiança de 95%. Os fundos que têm uma maior *Value at risk* são mais penalizados em termos de ranking. Quando observamos os retornos ajustados ao VaR vemos uma diferença para as outras medidas, pois o período pós 2009 é o que evidencia o pior desempenho. Quanto ao ranking, observamos que a ordenação dos fundos no último período de observação se encontra exatamente igual ao período global, enquanto que nos dois primeiros períodos o ranking é muito diferente.



	<b>Sharpe</b>	<b>α de</b>	<b>Trevino</b>	<b>Lense</b>	<b>MM</b>	<b>IF</b>	<b>RORA</b>	<b>Var</b>	<b>Sortin</b>	<b>UPR</b>	<b>Oneig</b>	<b>Fouse</b>	
AEDAX	13.51	6	0.57%	5	0.77%	1	0.79%	4	0.58	1	-1.57	2	2.69%
AEBBX	12.29	8	0.61%	7	0.99%	7	0.73%	7	0.92	7	-1.45	2	2.70%
AEDCK	12.29	9	0.61%	8	0.99%	8	0.73%	8	0.92	8	-1.44	1	2.70%
AEEBK	7.65%	1	0.38%	1	0.72%	1	0.53%	1	0.57	1	-1.73	2	2.62%
ESMAX	<b>20.66</b>	<b>1</b>	<b>1.09%</b>	<b>1</b>	<b>2.06%</b>	<b>4</b>	<b>1.25%</b>	<b>2</b>	<b>1.37</b>	<b>1</b>	<b>10.7</b>	<b>1</b>	<b>2.36%</b>
EUGAX	5.60%	2	0.27%	2	0.61%	2	0.43%	2	0.55	2	-2.80	2	2.35%
EUBBX	5.50%	2	0.27%	2	0.61%	2	0.43%	2	0.55	2	-2.68	2	2.31%
EUGCX	4.47%	2	0.21%	2	0.55%	2	0.49	2	2.36	2	2.36%	1	2.20%
EUGDX	5.48%	2	0.26%	2	0.61%	2	0.43%	2	0.25	2	-3.09	2	2.36%
EUGIX	4.30%	2	0.20%	2	0.54%	2	0.48	2	-2.12	2	2.38%	1	2.12%
FIEUX	9.97%	1	0.49%	1	1.06%	6	0.71%	1	0.79	1	1.66	1	2.24%
HFEAX	18.28	2	1.01%	2	2.18%	2	1.27%	1	1.24	3	1.35	1	2.30%
HFEBX	17.28	3	0.95%	3	2.08%	3	1.20%	3	1.19	5	1.29	1	2.29%
IEOAX	10.94	1	0.55%	1	4.23%	1	0.64%	1	0.84	1	5.15	3	2.35%
JFESX	11.85	1	0.60%	9	1.17%	5	0.78%	5	0.59	9	6.02	2	2.46%
MAEFX	7.65%	1	0.36%	1	0.86%	1	0.56%	1	0.57	1	1.78	1	2.14%
MCEFX	6.15%	2	0.28%	2	0.75%	1	0.48%	2	0.58	2	1.54	1	2.12%
MDEFX	7.35%	1	0.34%	1	0.83%	1	0.55%	1	0.55	1	1.75	1	2.14%
MEURX	15.34	4	0.68%	4	0.63%	1	0.75%	6	1.08	2	-0.40	1	2.91%
PRESX	9.58%	1	0.48%	1	0.96%	9	0.66%	1	0.77	1	4.32	4	2.20%
RECK	7.57%	1	0.38%	1	0.72%	1	0.53%	1	0.57	1	-1.76	2	2.59%
TEMIX	<b>14.61</b>	<b>5</b>	<b>0.65%</b>	<b>6</b>	<b>0.61%</b>	<b>2</b>	<b>0.73%</b>	<b>9</b>	<b>1.04</b>	<b>4</b>	<b>-0.38</b>	<b>1</b>	<b>2.90%</b>
TEURX	12.98	7	0.59%	1	0.56%	2	0.67%	1	0.95	6	-0.35	1	2.96%
UEIX	2.11%	2	0.06%	2	0.46%	2	0.24%	2	0.37	2	2.31	8	2.31%
UEPSX	1.60%	2	0.03%	2	0.43%	2	0.21%	2	0.34	2	1.85	9	2.51%
VESIX	7.26%	1	0.34%	1	0.80%	1	0.53%	1	0.64	1	2.78	7	2.31%
VEURX	7.55%	1	0.36%	1	0.82%	1	0.55%	1	0.56	1	3.08	5	2.32%
VEUSX	6.97%	2	0.33%	2	0.77%	1	0.52%	2	0.83	2	2.95	6	2.32%
Média	9.53%		0.47%		1.01%		0.53%		0.77		0.94		2.44%
													4.83%
													13.51
													0.5
													1.29
													0.61

Tabela 4: Desempenho dos fundos de investimento e respectivo ranking (2001-2015)



O índice baseado no VaR, como o nome indica, é outra medida que toma o VaR como indicador de risco. O denominador deste rácio é igual ao índice de Sharpe e os resultados indicam-nos o retorno em excesso à taxa sem risco ajustada à perda máxima no intervalo de confiança de 95%. O desempenho dos fundos no período de crise foi mais reduzido e os valores negativos geram o mesmo problema como o índice de Sharpe na sua interpretação. Em média, os fundos obtiveram um retorno ajustado ao risco de 7.87%, tendo no período seguinte uma quebra acentuada para 0,14%. No último período, recuperou-se essa perda, fazendo com que a média global atingisse os 4.83%. O ranking segundo esta medida é semelhante às outras medidas que têm o mesmo numerador.

Os valores positivos obtidos segundo o Índice de Sortino demonstram os retornos excedentes dos fundos e podem ser explicados por decisões inteligentes tomadas pelos gestores. Inicialmente, os fundos apresentam desempenhos muito favoráveis, destacando-se os três melhores fundos, ESMAX, HFEAX e HFEBX, que apresentam retornos excedentes ajustados ao *downside risk* extremamente elevados, 214.28%, 198.92% e 192.37%, respectivamente. Tal também promoveu que a queda em 2004-2008 fosse muito mais acentuada, mas, a partir de 2009, os fundos começaram a recuperar da crise, contudo não conseguiram atingir o patamar inicial.

O *Upside Potential Ratio* (UPR) aplica também o *downside risk* no numerador do rácio e este indica-nos o retorno de um fundo relativamente ao retorno mínimo exigido pelo investidor. Para haver consistência com as outras medidas, como taxa mínima aceitável na estimação do rácio é adotada a taxa do ativo sem risco, concretamente, a taxa de retorno das obrigações alemãs de 10 anos.

Esta medida ajustada ao risco mostra-nos a mesma evolução de desempenho dos fundos antes e depois da crise econômica global. Os fundos tiveram um bom desempenho no início do período de observação e pioraram na época de crise e, finalmente, demonstram melhorias após a crise.

O índice de Omega é calculado através da criação de uma partição na distribuição cumulativa de retornos, a fim de criar uma área de perdas e uma área para ganhos relativos à taxa mínima exigida por investidores

(RMA). Neste sentido, caso o valor obtido por esta medida seja inferior a 1, tal significa que a área dos retornos acima de RMA é menor que a área dos retornos abaixo de RMA. Na observação dos resultados constata-se que só alguns fundos tiveram este rácio inferior a 1 no período entre 2004-2008. Em média, no período 2001-2003, os ganhos dos fundos são 1.6 vezes maiores do que as perdas, tendo RMA como ponto de referência. A razão entre os ganhos e as perdas diminui com a crise, descendo para 1.02 em 2004-2008 e subindo para 1.44 em 2009-2015. Na perspectiva global, os ganhos são 1.3 vezes maiores que as perdas. Obviamente, este rácio depende essencialmente da taxa mínima aceitável pelos investidores.

O índice de Fouse difere das outras medidas ajustadas ao risco porque tem em conta o grau de aversão ao risco dos investidores. Na comparação com o alfa de Sharpe, este índice utiliza a semivariância em lugar da variância. O coeficiente de aversão ao risco nos cálculos apresentados na Tabela 4 foi 1. É de esperar que, à medida que aumentasse este parâmetro, os resultados tenderiam a ser mais negativos.

Após a análise individual das medidas, e fazendo aqui apenas uma síntese, pode-se concluir que os fundos apresentam o melhor desempenho no período de 2001-2003. A crise econômica global teve um impacto muito relevante no desempenho dos fundos. O período 2004-2008 foi o de pior desempenho e o ano de 2009 dá início à recuperação dos desempenhos. Os fundos que têm a maior volatilidade sofreram mais com o impacto da crise econômica. Em termos de ranking, verificamos que o comportamento dos fundos é similar em todos os períodos, exceto no período 2004-2008 que apresenta resultados muito distintos. As medidas nos diferentes períodos não apresentam a mesma média de performance dos fundos nem a sua forma de distribuição, tal como se demonstra pelo teste de Kruskal-Wallis. Com um nível de significância de 0,00, rejeitamos a hipótese nula, concluindo que a distribuição não é igual nos diferentes períodos.

#### **4.2. Correlação entre os rankings dos fundos**

Seguidamente, iremos verificar, através de testes estatísticos, o nível de consistência entre os índices de avaliação de desempenho dos fundos de investimento (Tabela 5).



Tabela 5: Correlação de ranking dos fundos para 2001-2015

	Sharpe	$\alpha$ de Sharp	Jensen	MM	IvaR	Sortino	Omega	UPR	Fouse	Trevnor	IF	RAROC
Sharpe	1.00											
<del>alpha</del> deSharpe	.995**	1.00										
Jensen	.965**	.974**	1.00									
MM	1.00**	.973**	.935**	1.00								
IvaR	.978**	.989**	.974**	.951**	1.00							
Sortino	.999**	.995**	.967**	.980**	.976**	1.00						
Omega	.998**	.994**	.970**	.978**	.974**	.998**	1.00					
UPR	.975**	.969**	.950**	.974**	.947**	.979**	.975**	1.00				
Fouse	.974**	.985**	.991**	.945**	.985**	.976**	.977**	.957**	1.00			
Trevnor	.652**	.683**	.750**	.647**	.730**	.654**	.661**	.681**	.757**	1.00		
IF	.199	.203	.284	.238	.216	.209	.213	.277	.275	.506**	1.00	
RAROC	.294	.287	.176	.313	.281	.291	.273	.227	.213	-.304	-.352*	1.00

Nota: \*\* - correlação significativa a 1%; \* - correlação significativa a 5%.



O índice MM é diretamente proporcional ao índice de Sharpe, estando este fortemente correlacionado com os índices de Sortino e Omega, podendo estes ser considerados seus substitutos. Já os índices de informação e RORAC não evidenciam uma correlação muito significativa com as outras medidas, nem entre eles, o que nos indica que estes indicadores não são substitutos dos outros, pelo que não podemos aceitar a hipótese H2 atrás formulada.

Calculando as correlações nos diferentes sub-períodos, nota-se que não existem grandes diferenças nos rankings dos indicadores. Apenas no período de 2004-2008, o nível de correlação é mais reduzido em relação aos outros períodos. Por exemplo, o RORAC apresenta uma correlação ainda menor comparado com os outros índices, o que pode ser explicado pelo fato do RORAC depender do VaR, sofrendo as mesmas limitações deste. Como o VaR depende de dados históricos, que se baseiam em condições normais do mercado, torna-se assim impossível prever variações extraordinárias nos mercados financeiros, como a crise econômica global.

#### **4.3. Análise De Correlações**

Conforme crítica apontada por Friend e Blume (1970), nas medidas tradicionais de avaliação existem possíveis correlações com as medidas de risco. Hipoteticamente, esta correlação não deveria existir, mas estudos de Klemkosky (1973), Ang e Chua (1979) e Chen e Lee (1981 e 1986) mostraram a existência de uma relação entre as medidas de avaliação do desempenho e as medidas de risco respetivas.

Realiza-se a seguir a análise de correlação para as medidas de avaliação de desempenho e medidas de risco, conforme apresentado na Tabela 6. No período global, as medidas que anteriormente mostravam uma correlação significativa entre elas não evidenciam uma correlação significativa com as medidas de risco. Apenas o RORAC mostrou uma correlação perfeita inversa com o *Value at Risk*. Os resultados observados para o RORAC não permitem rejeitar H1. Mostra-se assim que a classificação dos fundos de investimento varia consoante o indicador



assume ou não elevada volatilidade dos rendimentos. O índice de Treynor também evidências correlação negativa de 0.66. Assim, as correlações negativas informam-nos que o aumento de volatilidade reduz o retorno ajustado ao risco, portanto, podemos rejeitar a hipótese proposta para aquelas medidas que mostram uma correlação negativa, rejeitando H3. Apesar desses resultados não serem apresentados, quando se analisam os diferentes subperíodos, também somos conduzidos a rejeitar H3.

Como foi referido anteriormente, a classificação da *Morningstar* é basicamente uma indicação de comparação de um fundo de investimento com outros fundos similares. O sistema de estrelas é aconselhado pela *Morningstar* para ser usado como um sistema de alerta, o que sugere investigação mais aprofundada e não deverá constituir uma ferramenta de decisão final.

Nos fundos sob análise, são atribuídas cinco estrelas ao fundo MEURX, o que significa que pertence aos melhores 10% do seu grupo, em ações europeias. Os fundos da amostra que possuem 4 estrelas são ESMAX, HFEAX, JFESX, PREASX, REECX, TEMIX e TEURX. A *Morningstar* atribui uma estrela aos UEPIX e UEPSX, que são os fundos com o ranking mais baixo.

Pela análise da Tabela 6 confirmamos a hipótese H3 (considerando que a *Morningstar rating* segue uma função utilidade, será de esperar que exiba uma maior correlação com as medidas de Fouse e alfa de Sharpe). O índice de Sharpe, MM, Sortino e Omega apontam para uma correlação aceitável com o sistema da *Morningstar* em termos de longo prazo. Apesar de não ser apresentado, nota-se que os indicadores mostram uma maior correlação como sistema de estrela em 2004-2008, e menor correlação com o período de 2001-2003. Isto deve-se ao fato do cálculo da *Morningstar* ser uma ponderação dos ratings de diferentes prazos.



Tabela 6: Correlação das medidas de avaliação de desempenho com as medidas de risco

	<b>Sharpe</b>	<b><math>\alpha</math> de Sharpe</b>	<b>Treynor</b>	<b>Jensen</b>	<b>MM</b>	<b>IF</b>	<b>RORAC</b>	<b>IVaR</b>	<b>Sortino</b>	<b>UPR</b>	<b>Omega</b>	<b>Fouse</b>
<b>Erro-padrão</b>	-.100	-.086	.505	.043	-.118	.530	-.887	-.056	-.093	-.004	-.080	.005
<b>Beta</b>	-.145	-.172	-.663**	-.282	-.113	-.702**	.596**	-.219	-.152	-.216	-.161	-.280
<b>Variância</b>	-.100	-.086	.505**	.043	-.118	.530**	-.887**	-.056	-.093	-.004	-.080	.005
<b>Var%</b>	-.294	-.287	.304	-.176	-.313	.352*	-.100**	-.281	-.291	-.227	-.273	-.213
<b>Downside risk</b>	-.250	-.242	.351*	-.123	-.267	.444**	-.895**	-.222	-.245	-.169	-.232	-.161
<b>Morningstar rating</b>	.707**	.668**	.054	.617**	.707**	.117	.465**	.631*	.695**	.682**	.665**	.514**
<b>U1</b>	.995**	.999**	.682**	.974-	.974**	.206	.292	.990**	.995**	.968**	.992**	.985**
<b>U2</b>	.992**	.986**	.614**	.949**	.984**	.188	.352*	.966**	.991**	.968**	.989**	.958**
<b>U3</b>	.932**	.926**	.458**	.860**	.928**	.021	.541**	.909**	.930**	.894**	.922**	.882**
<b>U4</b>	.884**	.877**	.636*	.800**	.885**	-.100	.616**	.859**	.882**	.842**	.871**	.824**
<b>U5</b>	-.931**	-.945**	-.842**	-.978**	-.899**	-.357*	-.092	-.961**	-.933**	-.926**	-.937**	-.979**
<b>K1</b>	.953**	.956**	.529**	.908**	.959**	.111	.411*	.934**	.963**	.934**	.958**	.920**
<b>K2</b>	.914**	.907**	.407*	.839**	.914**	-.007	.582**	.889**	.912**	.867**	.903**	.856**
<b>K3</b>	.859**	.850**	.315	.767**	.854**	-.113	.639**	.833**	.854**	.797**	.842**	.791**
<b>K4</b>	.834**	.827**	.278	.733**	.829**	-.148	.653**	.807**	.831**	.764**	.820**	.761**
<b>K5</b>	.818**	.810**	.250	.715**	.810**	-.190	.663**	.789**	.814**	.745**	.804**	.743**

Nota: \*, \*\*, \*\*\* - valores significativos, respectivamente, a 10%, 5%, 1%.



Seguidamente, aplicamos a análise de regressão linear, no intuito de predizer o valor da nossa variável dependente (*Morningstar rating*) a partir de um conjunto de variáveis independentes (as medidas de desempenho). Com isto pretendemos testar H5, ou seja, se os índices de desempenho considerados têm o poder de explicar as classificações de risco da *Morningstar*.

**Tabela 7:** Regressão linear múltipla

	<b>Coef.</b>	<b>erro-padrão</b>
Constante	25.550	8.310 ***
Treynor	-77.545	54.700
Jensen	-407.200	1794.000
MM	18.010	4.861 ***
IF	.088	.051
RORAC	416.800	201.600 *
IVaR	-279.100	117.100
Omega	-33.089	8.346 **
UPR	-18.769	11.725
Fouse	40262.700	2401.800
 R <sup>2</sup>	0.816	
R <sup>2</sup> adj.	0.724	
ANOVA (F)	8.859	***
Durbin Watson	1.404	

\* p<0.10; \*\* p<0.05 e \*\*\* p<0.01

Sendo a *Morningstar rating* a variável dependente e Treynor, Jensen, MM, IF, RORAC, IvaR, Omega, UPR e Fouse os preditores, procedemos a esta análise através do software SPSS e obtemos um coeficiente de determinação (R<sup>2</sup>) de 81.6% e um R<sup>2</sup> ajustado de 72.4%, o que nos indica uma qualidade de ajustamento aceitável e que não devemos rejeitar H5.



O nível de significância 0,00% do teste de ANOVA diz-nos que os preditores são significativos em explicar a classificação da *Morningstar*. Lembramos que existe correlação muito elevada entre as medidas, o que viola os pressupostos de análise de regressão, auto-correlação e multicolinearidade. Portanto, o índice de Sharpe, alfa de Sharpe e índice de Sortino foram excluídos automaticamente pela análise. No entanto, os VIF da regressão apresentam-se inferiores a 2, excluindo, portanto, eventuais problemas de multicolinearidade.

Analizando a correlação das medidas com a função utilidade quadrática, em que o parâmetro de aversão ao risco vai alternando numa escala de 1 a 5, em que U1 representa o nível mais baixo de aversão ao risco, U2 e U3 são níveis intermédios e U4 e U5 níveis avançados, pretende-se testar a hipótese da correlação entre a função utilidade esperada e as medidas de avaliação de desempenho. À medida que aumenta o coeficiente de aversão ao risco de 1 para 5, o grau de correlação diminui, mas os resultados permitem-nos, globalmente, confirmar H6, que existe uma relação significativa entre a função utilidade esperada e as medidas de avaliação de desempenho, verificando-se que o aumento do coeficiente de aversão ao risco de 1 para 5, mantém a correlação positiva.

Finalmente, considera-se a função utilidade da Teoria da Perspectiva, onde  $\alpha=\beta=1$  para que a função seja linear, e o coeficiente de aversão ao risco é alterado de 1 a 5. A estimativa típica de  $k$  é de 2.25, indicando que as perdas são cerca de duas vezes mais dolorosas do que os ganhos. Para  $k = 1$ , os ganhos e as perdas teriam a mesma inclinação; para  $k$  maior que 1, as perdas pesam mais fortemente do que ganhos. Os estudos empíricos de Kahneman e Tversky estimam que  $\alpha$  é tipicamente igual a cerca de 0.88 e sempre menor do que 1. Quando o expoente  $\alpha$  é menor que 1, a curva vai acelerar negativamente, já para  $\alpha = 1.00$  a função seria linear e para  $\alpha$  maior que 1.00, aceleraria positivamente (Hastie e Dawes, 2001). Ressignando à Tabela 6, são aí apresentadas as correlações da função bilinear com as medidas de avaliação de desempenho, testando-se assim H7 (ou seja, a correlação entre as medidas de avaliação de desempenho e a função bilinear diminui, mas mantém-se positiva, à medida que aumenta o coeficiente de aversão ao risco). Pela análise da Tabela 6, podemos avaliar a correlação entre as medidas de desempenho ajustadas ao risco e a



função valor de teoria da perspectiva, à medida que se consideram variações no parâmetro do grau de aversão ao risco. Para o período global, o índice de Sharpe manifesta a maior correlação com todos os níveis de aversão ao risco. Diferentemente dos resultados obtidos com a função utilidade quadrática, o índice de Sharpe sobressai nos níveis intermédios. As restantes medidas de avaliação de desempenho estão também fortemente correlacionadas com a função bilinear, tirando o índice de Treynor e o índice de informação. O índice de Treynor só apresenta a correlação estatisticamente significativa com os dois primeiros níveis, no período global. O índice de informação não apresenta qualquer correlação com a função de valor da teoria da perspectiva. O grau de correlação manifesta uma tendência decrescente à medida que aumenta o coeficiente de aversão ao risco. Desta ordem, o índice de MM e índice de Omega mantêm uma boa correlação tanto para período global como os outros subperíodos. O índice de informação apresenta correlações negativas ao longo de tempo. Entre o período de 2004-2008, as correlações são muito distintas dos outros períodos em estudo, vários índices apresentaram correlação negativa com a função de preferência.

Para as medidas com maior consistência não podemos rejeitar a hipótese de a correlação entre as medidas de avaliação de desempenho e a função bilinear diminuir, mas mantendo-se positiva, à medida que aumenta o coeficiente de aversão ao risco.

## 5. Conclusão

Neste artigo foi efetuada uma análise comparativa entre as medidas de avaliação de desempenho de fundos de investimento em ações. Com base numa amostra de 28 fundos de ações europeus, e para o período 2001-2015, foram calculadas e comparadas doze medidas de desempenho tradicionais. Conclui-se que os diferentes índices estabelecem rankings idênticos, com exceção do índice de informação e do RORAC. Verifica-se que os fundos conseguiram melhores retornos no período de 2001-2003, mas aqueles foram fortemente penalizadas pela crise econômica global de 2008. A partir de 2009, os fundos mostram uma tendência de crescimento. Quanto à análise sobre o desempenho dos fundos em diferentes períodos, os retornos ajustados ao risco medido pelos índices de avaliação de desempenho mostram a mesma tendência.

Foram elaborados sete testes de hipóteses: i) não rejeitamos H1, o que evidencia que a classificação dos fundos de investimento varia consoante o indicador assume ou não elevada volatilidade dos rendimentos, pelo que se considera um resultado muito importante para a literatura; ii) não rejeitamos H2, pelo que existe correlação elevada entre as medidas em estudo, com exceção do RORAC e do índice de informação; iii) rejeitamos H3, porque a correlação entre as medidas de desempenho e as medidas de risco não são significativas, o que contradiz o estudo de Friend e Blume (1970); iv) rejeitamos H4, porque o índice de Sharpe e MM mostram maior correlação com a classificação da Morningstar; v) não rejeitamos a H5 pelo que o  $R^2$  ajustado obtido é de 72.4%, resultado que nos permite confirmar que encontramos medidas que replicam uma grande parte das variações de classificação da *Morningstar*; vi) rejeitamos H6, pelo que a correlação das medidas com a função quadrática é negativa quando o coeficiente de aversão ao risco é igual a 5. Em termos de correlação com a função utilidade quadrática, o alpha de Sharpe domina no nível mais baixo de aversão ao risco para todos os períodos. Isto não é surpreendente, porque o alfa de Sharpe é frequentemente utilizado como uma representação alternativa da função utilidade quadrática. O índice de Sharpe associa-se melhor com níveis intermédios de aversão ao risco. Para níveis mais avançados, o índice de MM e Jensen são mais indicados; vii) não rejeitamos H7 para as medidas com maior consistência, porque estas mantêm a correlação positiva com a função valor da teoria da perspectiva. Em relação ao grau de correlação com a função valor da teoria da perspectiva, o índice de Sharpe evidenciou uma correlação elevada com todos os níveis de função bilinear no período global e no período de 2009-2015. No período de 2004-2008, o índice de Omega correlaciona mais com nível mais baixo de aversão ao risco e o índice de MM domina nos níveis mais elevados, apenas no nível K3, o índice baseado no VaR sobressai mais. Já no período de 2001-2003, a maior correlação com nível k1 destaca-se no alfa de Sharpe e, nos níveis mais avançados, o índice de Omega é dominante. Apesar de consistência entre as medidas estar inversamente correlacionada com as medidas de risco, não deixa de ser curioso de verificar a relação em todos os subperíodos e período global das medidas baseadas em *downside risk* (Omega e Sortino) com a função valor da teoria da perspectiva.



Não devemos perder a vista que estamos a trabalhar com período limitado e uma amostra relativamente reduzida e sujeita ao enviesamento de sobrevivência, sendo uma limitação deste trabalho. Será interessante futuramente explorar a adaptação das medidas de desempenho a outros tipos de fundos, nomeadamente: fundos de obrigações, fundos mistos, etc.

Pensamos que este trabalho proporciona um contributo para um maior entendimento das diferentes medidas de desempenho de carteiras, podendo servir de ponto de partida para trabalhos futuros.

## Referências

- Ang, J. S., & Chua, J. H. (1979). Composite measures for the evaluation of investment performance. *Journal of Financial and Quantitative Analysis*, 14(02):361-384.
- Brown, K. C., & Brown, G. D. (1987). Does the composition of the market portfolio really matter?. *The Journal of Portfolio Management*, 13(2):26-32.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57-82.
- Chen, S. N., & Lee, C. F. (1981). The sampling relationship between Sharpe's performance measure and its risk proxy: sample size, investment horizon and market conditions. *Management Science*, 27(6):607-618.
- Chen, S. N., & Lee, C. F. (1986). The effects of the sample size, the investment horizon and market conditions on the validity of composite performance measures: A generalization. *Management Science*, 32:(11):1410-1421.
- Dowd, K. (1999). A value at risk approach to risk-return analysis. *The Journal of Portfolio Management*, 25(4):60-67.

- Eling, M., & Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds?. *Journal of Banking & Finance*, 31(9):2632-2647.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3-56.
- Ferguson, R. (1980). Performance measurement doesn't make sense. *Financial Analysts Journal*, 36(3):59-69.
- Ferguson, R. A. (1986). The trouble with performance measurement. *The Journal of Portfolio Management*, 12(3):4-9.
- Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *The Journal of Finance*, 51(2):425-461.
- Friend, I., & Blume, M. (1970). Measurement of portfolio performance under uncertainty. *The American Economic Review*, 60(4):561-575.
- Grinold, R. C. (1989). The fundamental law of active management. *The Journal of Portfolio Management*, 15(3):30-37.
- Hastie, R. (2001). Problems for judgment and decision making. *Annual Review of Psychology*, 5(1):653-683.
- Henriksson, R. D., & Merton, R. C. (1981). On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. *Journal of Business*, 513-533.
- Israelsen, C. (2005). A refinement to the Sharpe ratio and information ratio. *Journal of Asset Management*, 5(6):423-427.
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, 23(2):389-416.



- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.
- Keating, C., & Shadwick, W. F. (2002). A universal performance measure. *Journal of Performance Measurement*, 6(3):59-84.
- Klemkosky, R. C. (1973). The bias in composite performance measures. *Journal of Financial and Quantitative Analysis*, 8(3):505-514.
- Lee, C. F., & Jen, F. C. (1978). Effects of measurement errors on systematic risk and performance measure of a portfolio. *Journal of Financial and Quantitative Analysis*, 13(02):299-312.
- Leland, H. E. (1999). Beyond mean-variance: Performance measurement in a nonsymmetrical world (corrected). *Financial Analysts Journal*, 55(1):27-36.
- Levy, H. (1981). The CAPM and the investment horizon. *The Journal of Portfolio Management*, 7(2):32-40.
- Lin, C., & Yung, K. (2004). Real Estate Mutual Funds: Performance and Persistence. *Journal of Real Estate Research*, 26(1):69-94.
- Matos, P., Silva, W. C., & Silva, F. (2015). Há bons gestores de fundos de investimento em ações no Brasil? (Do Brazilian mutual stock fund managers have sufficient skill?). *Revista Brasileira de Finanças*, 13(2):325.
- Milan, P. L. A. B., & Junior, W. E. (2014). Elevada Rotatividade de Carteiras e o Desempenho dos Fundos de Investimento em Ações (High Portfolio Turnover and Performance of Equity Mutual Funds). *Revista Brasileira de Finanças*, 12(4):469.
- Modigliani, F., & Modigliani, L. (1997). Risk-adjusted performance. *The Journal of Portfolio Management*, 23(2):45-54.



- Plantinga, A., & Groot, J. S. (2001). *Risk-adjusted performance measures and implied risk-attitudes*. University of Groningen.
- Roll, R. (1980). Performance evaluation and benchmark errors (I)\*. *The Journal of Portfolio Management*, 6(4):5-12.
- Sabbaghi, O. (2011). The behavior of green exchange-traded funds. *Managerial Finance*, 37(5):426-441.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442.
- Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *The Journal of Portfolio Management*, 18(2):7-19.
- Sharpe, W. F. (1998). Morningstar's risk-adjusted ratings. *Financial Analysts Journal*, 54(4):21-33.
- Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing*, 3(3):59-64.
- Sortino, F. A., Van der Meer, R., & Plantinga, A. (1999). The Dutch Triangle. *Journal of Portfolio Management*, 26:50–58.
- Sourd, V. (2007). Performance measurement for traditional investment. *Financial Analysts Journal*, 58(4):36-52.
- Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review*, 43(1):63-75.
- Treynor, J., & Mazuy, K. (1966). Can mutual funds outguess the market. *Harvard Business Review*, 44(4):131-136.



Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297-323.

Wong, K., & Shum, W. (2010). Exchange-traded funds in bullish and bearish markets. *Applied Economic Letters*, 17(16):1615-1624.

