

Laes, Marco Antonio; da Silva, Marcos Eugênio

Article

Performance of mutual equity funds in Brazil – A bootstrap analysis

EconomiA

Provided in Cooperation with:

The Brazilian Association of Postgraduate Programs in Economics (ANPEC), Rio de Janeiro

Suggested Citation: Laes, Marco Antonio; da Silva, Marcos Eugênio (2014) : Performance of mutual equity funds in Brazil – A bootstrap analysis, EconomiA, ISSN 1517-7580, Elsevier, Amsterdam, Vol. 15, Iss. 3, pp. 294-306,
<https://doi.org/10.1016/j.econ.2014.08.002>

This Version is available at:

<https://hdl.handle.net/10419/179578>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/3.0/>

Performance of mutual equity funds in Brazil – A bootstrap analysis[☆]

Marco Antonio Laes^{*}, Marcos Eugênio da Silva

University of São Paulo, Brazil

Received 4 October 2013; received in revised form 30 July 2014; accepted 21 August 2014
Available online 16 October 2014

Abstract

This article reports a study on the performance of mutual equity funds in Brazil from January 2002 to August 2012. For the analyses, Carhart's four-factor model is used as the benchmark for performance, and bootstrap procedures are applied to separate skill from luck. The results show that returns of the best performers are more due to luck than skill of their managers. For the bottom ranked funds, on the contrary, there is statistical evidence that their poor performance is caused mainly by bad management, rather than by bad luck. It is also showed that the largest funds perform better than the small or middle-sized funds.

© 2014 National Association of Postgraduate Centers in Economics, ANPEC. Production and hosting by Elsevier B.V. All rights reserved.

JEL classification: C15; G11; G15; G23

Keywords: Equity funds; Performance analysis; Bootstrap

Resumo

Este artigo analisou o desempenho dos fundos de investimento em ações no Brasil no período de janeiro de 2002 a agosto de 2012. A base para as análises foi o modelo de quatro fatores de Carhart, sobre o qual foram aplicados procedimentos de *bootstrap* a fim de separar habilidade de sorte nas análises. Os resultados indicaram que os retornos dos fundos de melhor performance se deveram mais à sorte do que propriamente da habilidade de seus gestores. Para os fundos de pior performance, pelo contrário, há evidência estatística de que seu mau desempenho foi causado principalmente pela má gestão, e não por azar. Os resultados indicaram também que os maiores fundos apresentaram desempenho superior ao dos fundos de pequeno e médio porte no período.

© 2014 National Association of Postgraduate Centers in Economics, ANPEC. Production and hosting by Elsevier B.V. All rights reserved.

Palavras chave: Fundos de Ações; Análise de desempenho; *Bootstrap*

[☆] This paper is based on the Master Dissertation of the first author, under the supervision of the second author. A snapshot of the results was presented at the BoletimEconômico da FIPE (Laes, 2010).

^{*} Corresponding author. +55 11993810384.

E-mail addresses: marco_laes@hotmail.com (M.A. Laes), medsilva@usp.br (M.E.d. Silva).

Peer review under responsibility of National Association of Postgraduate Centers in Economics, ANPEC.

1. Introduction

In the previous decade, the combination of several factors, such as an international scenario of high liquidity and the relative strengthening of capital market institutions in Brazil, created the conditions for a vigorous expansion of the Brazilian stock market. During this period, mutual funds became a major conduit for private investment in the country: disregarding funds of funds, there were 6982 mutual funds in Brazil in August 2012, compared to 2483 in January 2002. In our sample, which considered only actively managed equity funds, the number of funds increased from 226 to 764 in the period, with a 510% increase in total assets under management, from R\$ 10 to R\$ 61 billion¹ – which represents approximately the value of all free float stocks in Bovespa.

The backdrop for the popularity of mutual funds is their presumed ability to provide professional management to the “uninformed” investors, generating higher returns. The natural question that arises is whether this professionalism really adds value to the investors. Since the seminal works of Treynor (1965), Sharpe (1966), and Jensen (1968), a large number of studies have focused on assessing whether portfolio managers actually obtain superior returns, with the majority of these studies (especially before the 2000s) centering their analysis in the significance of alphas from regressions of the CAPM or Fama–French models. The results are not uniform, but most of them show no clear evidence that mutual funds systematically achieve superior performance.

In Brazil, on the contrary, the results are more favorable to the existence of superior performance. Leusin and Brito (2008), using the CAPM plus the *market timing* term from Treynor and Mazuy as a benchmark, found a fairly high number of funds with positive and significant alphas (15 funds in a sample of 243) during the period 1998–2003. Gomes and Cresto (2010) analyze the performance of funds that employ the long-short strategy, using the same approach of Leusin and Brito, and find strong evidence of superior performance: 8 of 45 funds using the CAPM, and 17 of 45 when using the CAPM plus the market timing. Castro and Minardi (2009), using the Carhart model plus the *market timing* term, found evidence of superior performance in 4.8% of their sample, which comprised 626 equity funds in the 1996–2006 period.

In our study, we found a very different scenario. The approach adopted was similar to that employed by Kosowski et al. (2006) and Fama and French (2010), in which the alphas and the *t*-statistics from the “traditional” regressions (using the Carhart model as benchmark) are compared to their respective simulated cross-sectional distributions, obtained via bootstrap techniques, in which all funds are assumed to have zero alpha. This procedure allowed us to differentiate between funds with real superior performance from funds whose apparent superior performance had been achieved by mere luck. Analyzing the performance of the equity funds industry in Brazil, considering a universe of 1111 funds during the period 2002–2012, we found evidence of superior performance on only 19 funds and, more wearisome, there were strong signs of inferior performance in more than half of the funds.

Besides this introduction, the paper is divided into the following sections: Section 2 – Data, where we describe the funds that are part of our analysis; Section 3 – Model, where we deal with the methodology adopted and its advantages compared to more traditional analysis; Section 4 – Results, with a summary of our major findings; and Section 5 – Conclusion.

2. Data

The study focuses on Brazilian equity mutual funds, defined by CVM² Instruction no. 409 as those with more than 66.6% of their net asset value (NAV) invested in stocks. The data were obtained from Associação Brasileira das Entidades dos Mercados Financeiros e de Capitais (ANBID Information System, version 4.2³) and CVM, and included weekly NAV and NAV/share for 1111 equity funds between January 2002 and August 2012, totaling 553 weeks of data.

To track only funds with active management and overall strategies based on equities, we excluded from our sample funds of funds, tracker/index funds, sectorial funds, and funds with strategies based on specific segments, such as

¹ Source: ANBID.

² Comissão de Valores Mobiliários is the Brazilian federal agency responsible for discipline, rule, and the supervision of the activities of all market participants.

³ We thank Anbima for kindly having provided the data.

Table 1

Summary statistics of funds. This table presents size, annualized net returns (based on weekly observations) and number of equity mutual funds in Brazil in existence during the 2002 to August of 2012 period, according to our cutoff criteria. The 1111 funds were aggregated in three different groups, based on their average mean net asset value during the period.

	Funds between 2 and 20 million	Funds between 20 and 100 million	Funds above 100 million
Number of funds	557	403	151
Mean net asset value (millions)	8.7	45.7	316.0
Net returns – Mean (annual) (%)	5.8	13.5	13.4
Net returns – Max. (annual) (%)	94.8	77.3	54.6
Net returns – Min. (annual) (%)	–79.6	–49.8	–25.3
Net returns – Std. Dev. (%)	16.7	14.4	12.0
Non survivors	228	136	36

Source: tabulated by the authors with ANBIMA data.

companies with minimum governance standards. Although one could argue that other exclusions should be made, such as sorting out open and exclusive funds, for example (as they usually have different fees and objectives), the main idea was to include all funds that operate under the same investment strategies restrictions.

There were also two more cutoff criteria: to be part of the study, a fund had to have a history of at least 52 weeks between January 2002 and August 2012, and an average NAV of more than R\$ 2 million.

It is noteworthy that all funds that obeyed these criteria were retained in the analysis, including those that ceased to exist during the period, largely avoiding the so-called survivorship bias. These funds, in theory, should be amongst the worst performers, and had they been discarded from the sample, the survivorship bias could have arisen.

Even after excluding from our sample funds that presented an average NAV below R\$ 2 million in the period, there was still a wide difference between the size of the smallest and largest funds, with 168 funds presenting average NAV between R\$ 2 million and R\$ 5 million, and 18 funds with average NAV over R\$ 500 million. To assess the differences in management that should arise from operating at such different scales, we opted to divide the analyses into three ranges by average NAV: R\$ 2–20 million, R\$ 20–100 million, and over R\$ 100 million.

Table 1 has a descriptive summary of the funds included in our sample. We can observe very diverse characteristics among the three groups, with the smaller funds presenting the worst returns and the largest proportion of non-survivor funds.

3. The model

The intuition behind most models used in performance analysis can be described as follows⁴: assuming that the expected return of any asset (or portfolio of assets) is a linear function of K benchmark portfolios (represented below as F_t), its expected return can be written as:

$$E(R_{i,t} - R_{f,t} | \Omega_{t-1}) = E(\gamma_i + \beta_i F_t | \Omega_{t-1}) \quad (1)$$

The expected excess return over the risk free rate of an asset i at time t , conditional on the information set Ω available immediately in preceding period (Ω_{t-1}), is a linear function of the expected returns of the K portfolios plus an intercept, also conditional on the Ω_{t-1} set. If (1) represents an equilibrium model of asset pricing, under the hypothesis of a semi-strong efficient market, the expectation of γ_i is zero. Thus, to test the existence of differentiated performances, one could simply test the value and significance of the intercept in the equation:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i F_t + \varepsilon_{it} \quad (2)$$

Even if (2) is not a representation of an equilibrium model of asset pricing, the intercept would be the average return of the fund in excess of the fund's exposure to the factors – which is a performance measure.

⁴ This definition is presented in Baks et al. (1999).

3.1. Carhart's four-factor model

The benchmark chosen to evaluate the funds' performances was the Carhart's (1997) four-factor model. To construct the weekly factors, the following basic procedures were adopted: (i) to be part of the analysis in a determined quarter, the stock should obey a minimum criterion of liquidity⁵ and the company should not present negative equity in the period, and (ii) all portfolios were rebalanced each quarter, weighted by the market value of the stocks.

The factors' construction followed procedures similar to Fama and French (1993) and Carhart (1997). In each quarter,⁶ the stocks that passed the cutoff criteria were ranked according to their market value. Those that were above the median were classified as big (B), and those that were below as small (S).

The stocks were also ordered according to their book-to-market ratio, and divided into three groups, according to the following breakpoints: the firms with the highest 30% ratios were classified as high (H), those with the lowest 30% ratios were classified as low (L), and the remainder of the firms as medium (M). To obtain the book value of a stock, the net worth of the company was divided by the number of outstanding shares (this is the definition of book value per share according to the Law 6.404/1976). If a company possessed common and preferred stocks,⁷ the asset value classification was done for each paper separately, thus enabling the two types of stocks of the same company to be classified into different portfolios.

It should be noted that the portfolios H, M, and L are constructed to capture the relationship between market value and book value; therefore, we used the book-to-market ratio separately for common and preferred stocks, allowing us to differentiate them. As the portfolios B and S are constructed to capture the difference between the returns of large and small companies, the company's full market value was used for both ordinary and preferred shares simultaneously.

Subsequently, the Fama and French factors were constructed; the SMB (Small Minus Big) as the average return of the Small portfolio minus the average return of the Big portfolio, and HML (High Minus Low) as the average return of the High portfolio minus the average return of the Low portfolio. The factor momentum (MOM) was constructed similar to the HML; however, instead of ordering the stocks according to their book-to-market ratio, they were sorted according to their accrued returns over the last 6 months.

The market portfolio (MKT) was calculated as the return of the value-weighted portfolio made up of all securities traded in Bovespa that presented the minimum liquidity criterion in each quarter. The weekly accrued daily Interbank Deposit rate (DI), obtained from CETIP, was used as the risk-free rate.

3.2. Bootstrap simulations

After the construction of the Carhart factors, regression (3) was estimated for each fund:

$$r_{i,t} = \alpha_i + \beta_{1,i} MKT_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} MOM_t + \varepsilon_{i,t} \quad (3)$$

where, in each instant t ($t = 1$ to 553), r_i is the excess return of the fund i in relation to the risk free rate, MKT is the excess return on the market portfolio, SMB is the size returns, HML is the book-to-market ratio returns, MOM is the momentum returns, α_i is the intercept (which represents the excess return on the factors), and ε_i is the regression residual.

From these regressions, we started the implementation of the bootstrap procedures. As stated by Kosowski et al. (2006), the tendency of funds to display non-normality in the residuals, especially those with extreme performance (which are precisely those of greatest interest in the studies), and the nonnormalities that arise in a cross-sectional distribution of alphas (even if individual funds have normal returns), can hinder the traditional inferences. The bootstrap

⁵ The liquidity value used was the one reported by Economatica, which considers the number and value of transactions and the number of days in which the stock was negotiated. The minimum threshold utilized was 0.001, as suggested by Argolo (2008), a number that should retain in the sample stocks that represented, at least, 0.1% of the volume and transactions in the respective period. This approach resulted in samples with approximately 250 stocks per quarter.

⁶ The criterion of rebalancing the portfolios on a quarterly basis was chosen as the sample period is of a booming capital market in Brazil, with a large number of companies making their initial public offerings. We sought in this way to capture the market dynamics with the highest possible fidelity.

⁷ In Brazil, since the adoption of the Law 10.303/2001, up to 50% of the capital stock of a company may be composed of preferred stocks. The preferred stock will have at least one less right than the common stock (normally voting power), but will have preference in receiving dividends.

procedure helps circumvent these points, as it allowed us to evaluate the joint distribution of alphas, and to take into account the non-normality of their distributions.

The most important advantage of this approach, however, is the possibility of explicitly considering the “luck” factor in the funds’ performance, because in a given large collection of funds, it is expected that a portion of these will show superior performance simply by mere chance. As an example (from [Gorman and Weigand, 2007](#)), imagine a world where 1024 managers have zero-alpha management skills, but present positive or negative alphas because of luck (with a 50% chance for each state); after 10 years, it is expected (although with a large variance) that one manager will produce positive alphas every year, but with no real superior management skills.

The method chosen was the paired bootstrap, as in [Fama and French \(2010\)](#).⁸ The bootstrap simulations were performed using the estimated value of the alphas and their *t*-statistics. Although the latter have superior statistical qualities as a pivotal quantity,⁹ only the results for alphas are presented here, given their more intuitive economic meaning and the similarity in the results.

3.2.1. Paired bootstrap

The application of the paired bootstrap proposed by [Fama and French \(2010\)](#) is as follows: for each investment fund *i*, the Eq. (3) in matrix form, for *t* = 1 to 553, is:

$$\begin{bmatrix} r_{i,1} \\ r_{i,2} \\ \vdots \\ r_{i,t} \\ \vdots \\ r_{i,553} \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_i \\ \hat{\alpha}_i \\ \vdots \\ \hat{\alpha}_i \\ \vdots \\ \hat{\alpha}_i \end{bmatrix} + \begin{bmatrix} MKT_1 & SMB_1 & HML_1 & MOM_1 \\ MKT_2 & SMB_2 & HML_2 & MOM_2 \\ \vdots & \vdots & \vdots & \vdots \\ MKT_t & SMB_t & HML_t & MOM_t \\ \vdots & \vdots & \vdots & \vdots \\ MKT_{553} & SMB_{553} & HML_{553} & MOM_{553} \end{bmatrix} \begin{bmatrix} \hat{\beta}_{1,i} \\ \hat{\beta}_{2,i} \\ \hat{\beta}_{3,i} \\ \hat{\beta}_{4,i} \end{bmatrix} \begin{bmatrix} \hat{\epsilon}_{1,1} \\ \hat{\epsilon}_{1,2} \\ \vdots \\ \hat{\epsilon}_{1,t} \\ \vdots \\ \hat{\epsilon}_{1,553} \end{bmatrix} \quad (4)$$

For the simulations, a new set of returns is created by subtracting its estimated alpha value ($\hat{\alpha}_i$), making it zero by construction:

$$\begin{bmatrix} \tilde{r}_{i,1} \\ \tilde{r}_{i,2} \\ \vdots \\ \tilde{r}_{i,t} \\ \vdots \\ \tilde{r}_{i,553} \end{bmatrix} = \begin{bmatrix} r_{i,1} - \hat{\alpha}_i \\ r_{i,2} - \hat{\alpha}_i \\ \vdots \\ r_{i,t} - \hat{\alpha}_i \\ \vdots \\ r_{i,T} - \hat{\alpha}_i \end{bmatrix} + \begin{bmatrix} MKT_1 & SMB_1 & HML_1 & MOM_1 \\ MKT_2 & SMB_2 & HML_2 & MOM_2 \\ \vdots & \vdots & \vdots & \vdots \\ MKT_t & SMB_t & HML_t & MOM_t \\ \vdots & \vdots & \vdots & \vdots \\ MKT_{553} & SMB_{553} & HML_{553} & MOM_{553} \end{bmatrix} \begin{bmatrix} \hat{\beta}_{1,i} \\ \hat{\beta}_{2,i} \\ \hat{\beta}_{3,i} \\ \hat{\beta}_{4,i} \end{bmatrix} \begin{bmatrix} \hat{\epsilon}_{1,1} \\ \hat{\epsilon}_{1,2} \\ \vdots \\ \hat{\epsilon}_{1,t} \\ \vdots \\ \hat{\epsilon}_{1,553} \end{bmatrix} \quad (5)$$

These returns have the same statistical properties of actual fund returns, except that the intercept is set to zero. Then, 10,000 simulations are performed, resampling with replacement the 553 weeks of the study period (which is equivalent to jointly resampling returns and independent variables) and then estimating the Carhart model, thus obtaining 10,000 new values of $\tilde{\alpha}_i^j$ for each fund *i*, where *j* represents each simulation:

$$[\tilde{\alpha}_i^1 \quad \tilde{\alpha}_i^2 \quad \dots \quad \tilde{\alpha}_i^j \quad \dots \quad \tilde{\alpha}_i^{10,000}] \quad (6)$$

⁸ As a comparative, all the analyses were also made using the wild bootstrap procedure, proposed by [Wu \(1986\)](#). This method is very similar to the residual bootstrap used by [Kosowski et al. \(2006\)](#), but with minor corrections for the presence of heteroscedasticity in the data. The results were virtually identical to those presented here, and can be obtained by request.

⁹ The distribution of a pivotal quantity does not depend on the parameters of the sample, and thus could be used independently of them.

This procedure is repeated, separately for each size category, for all funds, resulting in a set with N series of 10,000 alpha values (N being the number of funds in that particular size category). For each set, the alpha values are ordered from largest to smallest in each of the simulations:

$$\begin{bmatrix} \tilde{\alpha}_{MAX}^1 & \tilde{\alpha}_{MAX}^2 & \dots & \tilde{\alpha}_{MAX}^j & \dots & \tilde{\alpha}_{MAX}^{10,000} \\ \tilde{\alpha}_{MAX-1}^1 & \tilde{\alpha}_{MAX-1}^2 & \dots & \tilde{\alpha}_{MAX-1}^j & \dots & \tilde{\alpha}_{MAX-1}^{10,000} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\alpha}_n^1 & \tilde{\alpha}_n^2 & \dots & \tilde{\alpha}_n^j & \dots & \tilde{\alpha}_n^{10,000} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\alpha}_{MIN-1}^1 & \tilde{\alpha}_{MIN-1}^2 & \dots & \tilde{\alpha}_{MIN-1}^j & \dots & \tilde{\alpha}_{MIN-1}^{10,000} \\ \tilde{\alpha}_{MIN}^1 & \tilde{\alpha}_{MIN}^2 & \dots & \tilde{\alpha}_{MIN}^j & \dots & \tilde{\alpha}_{MIN}^{10,000} \end{bmatrix} \quad (7)$$

where the superscripts represent the number of the simulation. Thus, for example, the theoretical distribution of α_{MAX} for funds with average NAV over \$100 million is a function of the 10,000 values of the bootstrapped $\tilde{\alpha}_{MAX}$. In the same way, the distribution of the alphas for the fund ranked as the second best is determined by the distribution of the second largest values of the alphas in each draw, and so on.

The maximum value of alpha in a particular simulation will not necessarily come from the fund that presented the largest alpha in the original estimations. Thus, this distribution uses information from all funds in the category, and not just from a single fund. The same is valid for other percentiles of returns.

Fama and French (2010) emphasize that, as the sequence of weeks for each of the 10,000 simulations is the same for all funds, this method should capture the cross-correlation between funds. As the returns and explanatory variables are jointly sampled, any possible correlated heteroscedasticity is also captured.

A setback that can arise from this approach is that only funds present in every week of the study will certainly have simulations with the same length as the original series. For a fund present in a period shorter than 553 weeks (the majority of funds), the simulations would not have the exact length of the series (since returns in the weeks that a fund did not exist are missing values). This should be a minor problem given the large number of simulations.

3.2.2. Inferences from the bootstrap simulations

The bootstrapped cross-sectional distributions for the three size categories were used as the basis for the inferences. Initially, the estimated alphas in the original regressions and the (average of) alphas obtained in the bootstrap analyses were grouped in different ranges of returns (similar to a histogram) and compared. Given the construction procedure of the bootstrap, approximately half of the bootstrapped alphas will be positive and half negative, and distributed in a fairly symmetrical way. If the original values have a distinct distribution (a larger number of positive or negative values, or some concentration around some specific value), we have an indication that the funds have returns different from the zero-alpha hypothesis.

Next, we present the inferences for individual funds. The idea is to compare the original alpha values for some selected percentiles and their respective empirical distribution of the same order – for example, we compare the larger original alpha with the bootstrap distribution of the largest alpha, and so on. Values within a certain confidence interval are indicative of zero alpha; values outside this confidence interval are considered real skill, or lack of it.

Before proceeding, it is important to discuss the possible limitations of the second approach, with a brief conceptual analysis of the distribution of alphas. Initially, it can be assumed that the excess return in relation to the Carhart factors (the alphas) can be decomposed into two components, skill and luck. The skill of the fund manager can be positive (generating a positive alpha), negative (generating a negative alpha), or neutral (the zero-alpha hypothesis). Luck also affects alpha, can be positive, null or negative, and has zero mean.

In a hypothetical population comprising a few funds with superior skill, some with inferior skill, and most with neutral skill, a symmetric distribution of alphas is expected, with a large mass close to zero. However, the distribution of funds would probably not be in the exact order of their skill (because of the presence of the luck factor). Thus, we cannot assume that the alphas obtained in the regressions are being compared to the exact same percentile of the empirical distributions: for example, the fund with the largest alpha in the original regressions possibly could not be the

most skilled fund, because of the luck factor. Nevertheless, if a large number of funds in a size category systematically present alphas close to the tails of the bootstrapped distributions, we have indication of superior or inferior performance.

4. Results

As already noted, all analyses were conducted using weekly the net returns of funds. Thus, the conclusions about superior or inferior performance are related to the returns that would be obtained by the shareholder, and not the actual returns obtained by the managers (the gross returns), where the costs are basically from transactions.¹⁰ All analyses were performed on a weekly basis, with the funds' return funds obtained from the NAV/share value on the last business day of the week; to facilitate the interpretation, results are reported in annualized terms.

4.1. Carhart's four factors

Table 2 shows the factors returns, by year:

Table 2
Carhart's factors – annual returns.

	SMB (%)	HML (%)	MOM (%)	MKT (%)
2002	10.6	1.2	60.0	−16.2
2003	−3.6	21.0	2.9	40.9
2004	22.1	−17.9	−14.9	14.8
2005	−13.3	−2.9	23.8	8.2
2006	8.3	11.1	−17.0	16.3
2007	−1.4	18.5	29.3	25.4
2008	−43.4	24.9	24.9	−52.3
2009	28.7	35.3	−0.1	58.5
2010	21.2	−9.6	−2.3	−5.0
2011	−6.8	−10.1	18.7	−20.1
2012 ^a	24.7	−21.2	−20.6	4.8

^a Annualized from the January to August period.

We can see that the SMB and MKT portfolio were the most affected by the crisis in 2008, and the MOM portfolio presented the best returns, especially due to a strong performance in 2002. Next, we have the correlation matrix (Table 3) of the factors. The results were similar to Chague (2007): the correlations between MKT–SMB, MKT–HML and SMB–HML were, respectively, −0.52, 0.12 and −0.35, where the values found by Chague between 1999 and 2007 were −0.40, 0.08 and −0.68.

4.2. Individual regressions

Table 4 presents a summary of the regressions for the 1111 funds. The excess market return (MKT) was significant for the vast majority of funds, and quite close to 1 on average.

The alphas were on average, negative for the smaller funds (−4.33%, in annualized terms), and close to zero for the larger funds (0.54%, also in annualized values). Table 5 gives a more detailed description of the alphas, in which the vast majority was negative, in particular in the smaller funds category. For funds with average NAV over 100 million, alphas were distributed in a fairly symmetrical way around zero.

Regarding the possible non-normality, the residuals of 613 funds (55% of the total) were shown to be non-normal by the Shapiro–Wilk test (see Table 6). This value is higher than the one to be found, for example, by Kosowski et al.

¹⁰ One could try to determine the gross returns considering fees and costs from all funds since, theoretically, this information is public. However, since many funds inform the maximum (and not the actual) fees charged, and we do not have the historical series of fees for the majority of funds, we could not proceed the analyses with the gross returns.

Table 3
Carhart's factors – correlation matrix.

	SMB	HML	MOM	MKT
SMB	1.000	–	–	–
HML	–0.352	–1.000	–	–
MOM	0.060	–0.246	1.000	–
MKT	–0.523	0.122	–0.239	1.000

Table 4

Individual regressions – summary statistics. In panels A, B and C we have, respectively for each size category, the summary statistics from the regressions of the excess return for each fund against Carhart's four factors. The "Significant" column gives the number of funds in that size category that presented significant coefficients (alphas and Carhart's factors), at a 5% significance level.

Coefficient	Significant ^a	Mean	Max.	Min.	Std. Dev.
Panel A – Fund between 2 and 20 million (557 funds)					
Alpha (annualized)	96	–0.0433	1.2347	–0.7703	0.1214
MKT	553	0.9581	2.2623	–0.6523	0.2457
SMB	336	0.2029	1.5853	–0.6050	0.2548
HML	212	0.0609	1.7115	–0.7991	0.2048
MOM	188	–0.0191	0.5948	–0.7790	0.1103
Panel B – Fund between 20 and 100 million (403 funds)					
Alfa (annualized)	86	–0.0109	0.5832	–0.4403	0.0848
MKT	395	0.9508	1.6870	–0.0724	0.2283
SMB	242	0.1589	1.2788	–0.8192	0.2251
HML	212	0.0264	1.3037	–0.5780	0.1631
MOM	166	–0.0216	0.4795	–0.4470	0.0804
Panel C – Fund above 100 million (151 funds)					
Alfa (annualized)	28	0.0054	0.2304	–0.2649	0.0708
MKT	148	0.9218	1.8952	0.0075	0.2451
SMB	87	0.1484	1.7698	–0.2316	0.2275
HML	75	0.0381	0.6713	–0.4727	0.1589
MOM	59	–0.0090	0.1639	–0.2694	0.0675

^a At 5% level.

Table 5
Alphas analysis.

	Total	Significant ^a
Panel A – Fund between 2 and 20 million (557 funds)		
Positives	150	12
Negatives	407	84
Panel B – Fund between 20 and 100 million (403 funds)		
Positives	160	30
Negatives	243	56
Panel C – Fund above 100 million (151 funds)		
Positives	68	14
Negatives	83	14

^a At 5% level.

Table 6
Shapiro–Wilk test.

	Normally distributed residuals (%)
Funds between 2 and 20 million	45.8
Funds between 20 and 100 million	44.9
Funds above 100 million	41.0
All funds	44.8

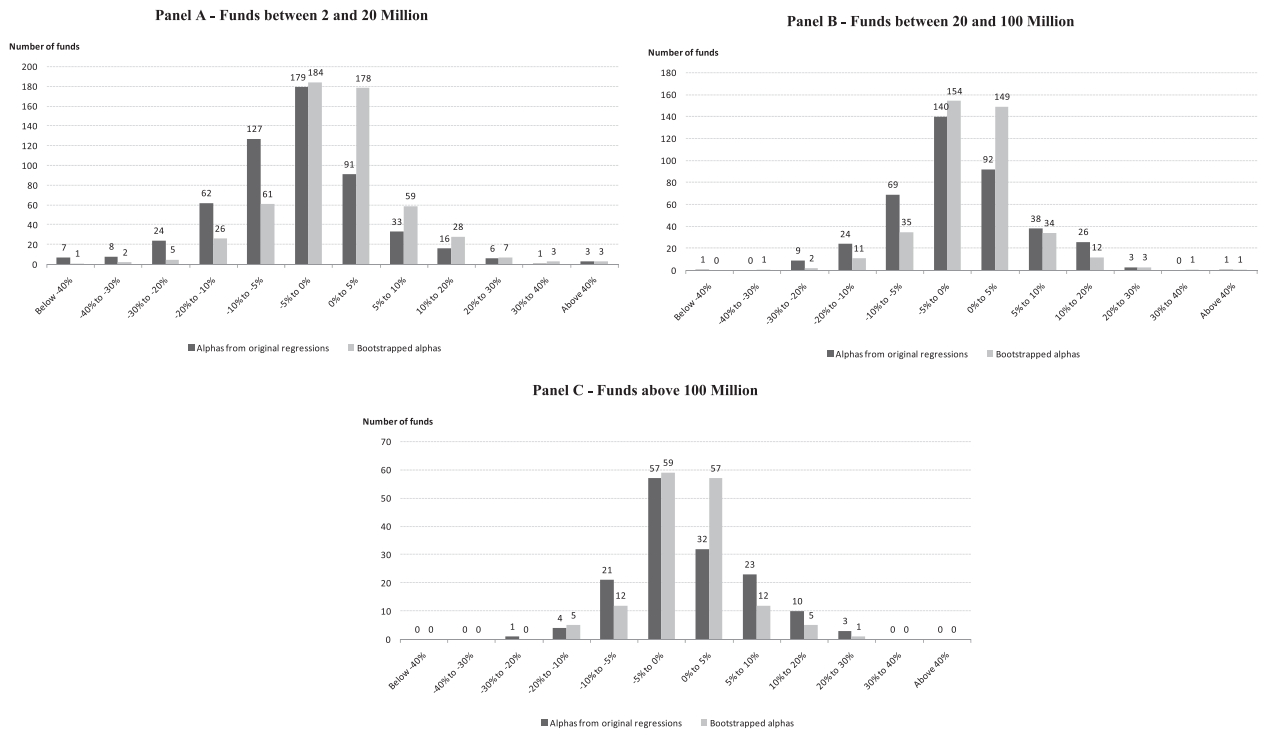


Fig. 1. Distribution of Alphas – original regressions and bootstrap simulations. In panels A, B, and C we have, for each size category, the alphas distributions from the original regressions and the average alpha from the bootstrap simulations. The darker bars of the graphs show the number of funds we have from the original regressions in a particular range of annualized alpha; the lighter bars show the number of funds from the bootstrap simulations there are in that same range of alpha.

(2006), in which 50% of funds were shown as non-normal. In addition, the largest funds had the lowest percentage of normally distributed funds.

4.3. Bootstrap simulations analysis

For each size category of funds, the bootstrap procedures were performed after the original regressions. The first analysis was a comparison between the average annualized alphas, obtained in the bootstrap (which gives the expected distribution of alphas under the zero alpha hypothesis), and the values from the original regressions. Fig. 1 shows the histogram for the original and the bootstrapped distributions of alphas in each size category.

The smaller funds systematically presented a worse performance than expected. For example, for the funds with average NAV between 2 and 20 million, under the assumption of zero-alpha, we expected 59 funds to present alphas between 5% and 10%, and 61 funds with alphas between -10% and -5%. In the original regressions, we observed only 33 funds in the first range and 127 funds in the second range. On the other hand, the distribution for the funds with NAV over 100 million was similar to that expected, but with heavier tails, indicating under and over performances (e.g., we expected 18 funds to have performance above 5%, and the actual number was 36 funds).

The next analysis was the comparison between the value of the original alpha, for each fund, and its respective simulated distribution. For example, if the funds ordered on a size category systematically present value of their regression $\hat{\alpha}$ above its respective 95% bootstrapped confidence interval, we have strong statistical evidence that there are funds in that category that possess real skill. Similarly, if a large number of negative values are found in the bottom of the simulations, we can reject the hypothesis of just bad luck and an indication of real underperformance. Values in-between indicate alphas different from zero purely by chance.

Table 7

Individual regressions and bootstrap simulations – selected percentiles. In panels A, B and C in the table we have, for each size category, the alphas from the original regressions for the best, median and worst funds, and for other selected percentiles, and their respective confidence interval obtained from the bootstrap simulations for the null hypothesis of zero alpha. The CI for the best fund in a determined size category was constructed from the distribution generated by the largest alphas in each of the 10,000 simulations for that category; the CI for the median fund came from the distribution generated by the median alphas in each of the 10,000 simulations, and so on. In panel D we have a summary of the results, with the number of funds, for each category, that lie above, below and within the CI.

	Original regression (%)	[95% Bootstrap confidence interval] (%)	
Panel A – Funds between 2 and 20 million (557 funds)			
Best fund	122.1	35.1	214.7
95th percentile	9.4	8.9	17.8
90th percentile	5.2	5.2	11.6
75th percentile	0.5	1.2	6.0
Median Fund	−3.7	−2.2	2.2
25th percentile	−8.6	−5.6	−1.1
10th percentile	−15.1	−10.2	−4.8
5th percentile	−22.3	−14.7	−7.9
Worst fund	−77.5	−71.5	−26.3
Panel B – Funds between 20 and 100 million (403 funds)			
Best fund	58.0	22.9	193.9
95th percentile	12.4	6.0	13.1
90th percentile	7.5	3.4	9.0
75th percentile	2.4	0.4	4.9
Median Fund	−1.4	−2.3	2.2
25th percentile	−5.1	−4.7	−0.3
10th percentile	−9.6	−8.2	−3.3
5th percentile	−12.9	−11.4	−5.5
Worst fund	−44.2	−66.4	−19.2
Panel C – Funds above 100 million (151 funds)			
Best fund	23.0	11.7	58.3
95th percentile	12.3	4.8	13.6
90th percentile	9.1	3.0	9.2
75th percentile	4.7	0.3	4.9
Median Fund	−0.6	−2.2	2.2
25th percentile	−3.9	−4.7	−0.2
10th percentile	−6.2	−8.1	−2.7
5th percentile	−7.4	−11.2	−4.4
Worst fund	−26.6	−36.0	−10.3
	Funds below CI	Funds within CI	Funds above CI
Panel D – Summary			
Funds between 2 and 20 million	504	53	0
Funds between 20 and 100 million	126	276	1
Funds above 100 million	0	133	18

In Table 7, we compare the values of the alphas (for the best and worst fund, and some selected percentiles) and the values found in the bootstrap simulations.¹¹ Exemplifying the results, in Fig. 2, we have a histogram of the bootstrapped alpha of the median fund in category of average NAV of more than 100 million. The median annualized alpha value in the original regression was −0.65%, whereas the two-tailed 95% confidence interval was [−2.3%, 2.4%]. Thus, considering this confidence interval, the original regression value is in line with the assumption of zero alpha.

¹¹ The comparison between the full cumulative distribution of the actual and simulated (median, 10th and 90th percentile) alphas can be seen in the Appendix.

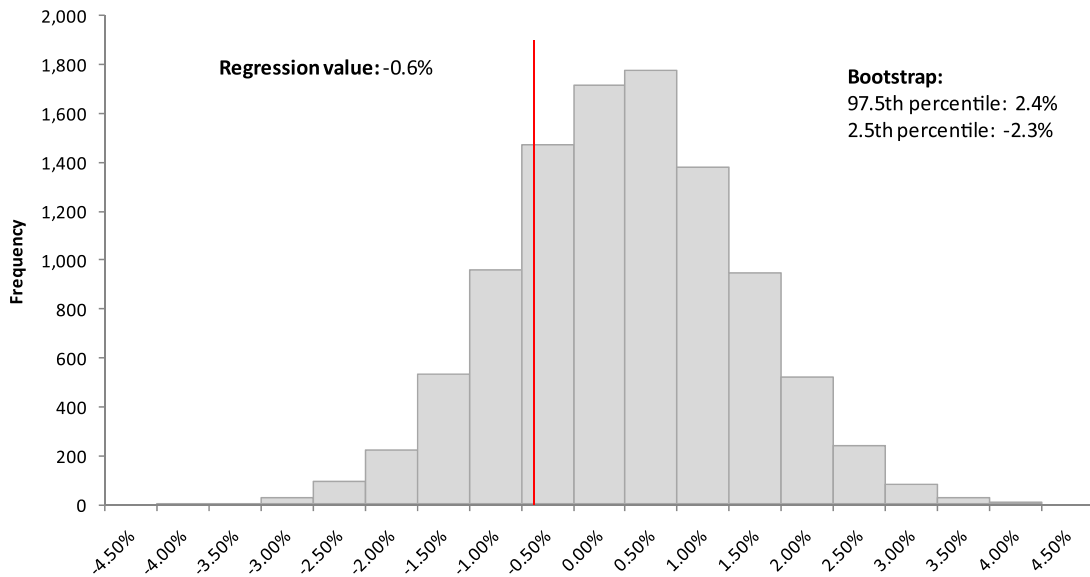


Fig. 2. Simulated distribution of the median fund – funds above 100 million.

4.3.1. Funds between 2 and 20 million

The smaller funds were the worst performers in the analysis. Almost all funds (504) had their original alphas below the 95% confidence interval of the simulated alphas, and none was above it, as exemplified in the Panel A of Table 7. These results give a clear indication that there are a large number of underperforming funds in this category, and no funds with superior performance.

They also illustrate the main difference in the results between the bootstrap approach and the traditional p value analysis, where the results were considerably less negative for the fund's management: in the original regressions for the size category, 12 funds presented positive and significant alphas, and only 84 funds presented negative and significant alphas.

4.3.2. Funds between 20 and 100 million

In this category, we have only one fund whose regression alpha was above the 95% confidence interval, well below the 30 funds that have shown positive and significant alphas in the original regression. As in the category of smaller funds, but to a lesser extent, we have a large number of funds (126 funds) that were below the bootstrapped confidence interval.

4.3.3. Funds above 100 million

Different from the other two categories, we have a large number of funds above the bootstrapped confidence interval (18 funds), indicating superior performance because of superior skill. On the other extreme, none of the funds were below the confidence interval. We have in this category, thus, a significant number (12%) of funds indicating the presence of superior management skills, and the others close to the underlying hypothesis of zero alphas.

4.4. Fama and French (2010)

Since the present study is based on the methodology used by Fama and French (2010), it is interesting to compare the results obtained by the authors in US market and those obtained in the present study.

Despite the same bootstrap approach to distinguish between skill and luck, and the analysis being made for groups of funds with different sizes, there are some differences in some procedures: the first is that the authors use the t statistics of the regressions, and not alphas, in the analyses. As our results with the t statistics yielded virtually the same results presented, there should be no problems with the comparative.

The other important difference is that the authors use the separation of funds by size in a clustered manner, i.e., the three size groups are defined only by the minimum value of asset under management of the funds, and not the maximum. As in our preliminary analysis, we observed that smaller funds had the worst performance; we choose to make the analysis of a non-nested way, precisely to highlight possible differences between size groups.¹²

The results presented by Fama and French (2010) are reasonable similar to ours: in that study, no fund had a *t* statistic above the average value of the simulations, a strong indication of inferior performance. For the funds in the higher percentiles, the results show a performance close to the simulations averages, indicating that some funds are able to cover their costs. Evidences are a little more favorable to the funds in the analyses made with their estimative of gross returns, where there is stronger evidence of manager skill, negative as well as positive.

5. Conclusion

This study analyzed the performance of the Brazilian equity funds industry. Compared to previous national studies, the major differences is the analysis using the empirical cross-sectional distribution of alphas obtained via bootstrap simulations, seeking with this approach to evaluate the influence of the luck in the analysis.

The basis for the evaluation of performance was the four-factor model of Carhart. In an analysis of the original estimated alphas, the results were in line with the traditional literature, with a large number of values not significant and, among the significant alphas, the majority presenting negative values, indicating thus that the professional management in general reduce rather than add value to the investors. Nonetheless, there are funds adding value (with positive and significant alphas) in the three size categories of funds analyzed.

The main criticism in using this approach, however, is that these results could be invalidated because of violation of several underlying assumptions, for example, the normal distribution of returns (only 44.8% of funds in this study). More importantly, this approach ignores the fact that in a large sample of funds, there is a likelihood of finding funds whose alphas are significant (positive or negative) by mere chance, without presenting actual differentiated performance.

To overcome these problems, we complemented the regression analysis with bootstrap procedures, simulating the empirical distribution of alphas, in an attempt to factor out the luck factor in the analysis. For this, a new set of returns was constructed, in which the values of the alphas of the original regressions were subtracted from the original returns. The underlying hypothesis to be tested with this procedure is that all funds had alpha equal to zero.

We then performed the paired bootstrap, with 10,000 simulations for each fund. In each simulation, a regression was performed for the new series of returns, ordering the bootstrapped alphas from largest to smallest, thus creating a cross-sectional series of alphas. To analyze the existence of differentiated performance, we compared the alphas obtained in the original regression with their respective simulated distributions.

Although several studies for the Brazilian market showed superior performance, we expected results similar to the ones found by Fama and French (2010) for the US market, where the results indicated that the vast majority of funds did not generate value for investors in the period, with performance insufficient to cover transaction costs and administration fees. This is exactly what we found: 56.7% of the funds in our analysis underperformed, showing negative alphas, that is, did not perform sufficiently enough to cover the costs charged to the investors.

Different from Fama and French (2010), where the results were evenly distributed among the different size classes, the negative results found here were most pronounced for the smaller funds, where 504 of 557 funds underperformed. Funds with the largest NAVs had the best results, with 18 funds presenting positive and significant alphas (which represented 12% of the largest funds, or 1.7% of the full sample), and the remaining showing results in line with the zero alpha hypothesis.

Appendix.

See Fig. 3.

¹² The downside of this approach arises for funds near the upper cutoff points, which could have behavior closer to funds in the larger group above. However, as our analyses are for the distribution of all funds, and not the results of individual funds, and since the same criticism can be made for funds next to the lower cutoff points, we opted for the non-nested approach.

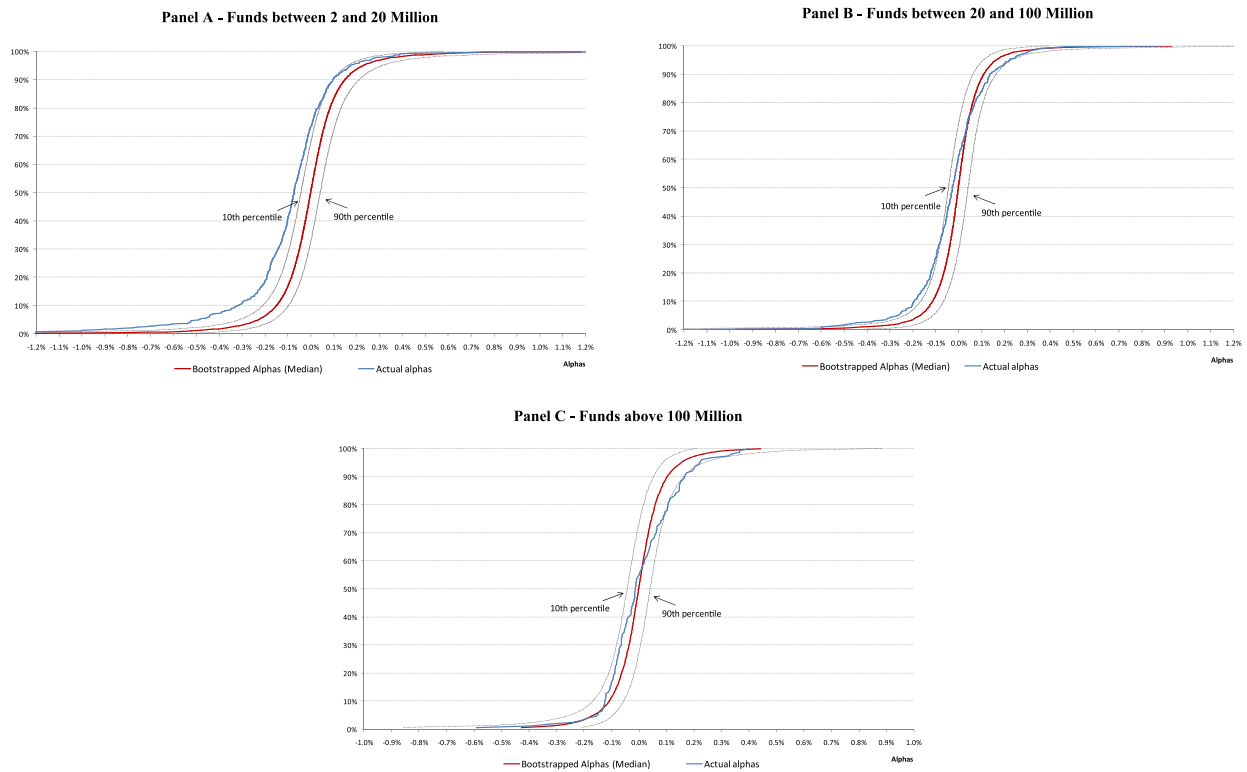


Fig. 3. Cumulative distribution of alphas – original regressions and bootstrap simulations. In panels A, B, and C we have, for each size category, the cumulative alphas distributions from the original regressions and the average alpha from the bootstrap simulations. The thinner lines represent the cumulative distribution of the 10th and 90th percentile of each size category, as a confidence interval.

References

- Argolo, É.F.B., (Dissertação (Mestrado em Administração)) 2008. *Implementação Prática do Modelo de Fama e French para o Cálculo do Custo de Capital Acionário no Brasil*. 2008. Universidade Federal do Rio de Janeiro, Instituto Coppead de Administração, Rio de Janeiro, pp. 95f.
- Baks, K., Metrick, A., Wachter, J., 1999. *Bayesian Performance Evaluation*. NBER Working Paper n. 7069.
- Brasil, 1976. Lei 6.404, de 15 de dezembro de 1976, Disponível em: http://www.planalto.gov.br/ccivil_03/Leis/L6404consol.htm (Último acesso em 25.10.2009).
- Carhart, Mark., 1997. *On persistence in mutual fund performance*. J. Fin. 52 (1), 57–82.
- Castro, B.R., Minardi, A.M.A.F., 2009. *Comparação do Desempenho dos Fundos de Ações Ativos e Passivos*. Rev. Bras. Fin. 7, 143–161.
- Chague, F., (Dissertação (Mestrado em Economia)) 2007. *The CAPM and Fama-French Models in Brazil: A Comparative Study*, 2007. Fundação Getúlio Vargas, Escola de Economia de São Paulo, São Paulo, pp. 39f.
- Fama, E.F., French, K.R., 1993. *Common risk factors in the returns on bonds and stocks*. J. Fin. Econ. 33 (1), 3–56.
- Fama, E.F., French, K.R., 2010. *Luck versus skill in the cross section of mutual fund returns*. J. Fin. 65 (5), 1915–1947.
- Gomes, F.A.R., Cresto, V., 2010. *Avaliação do Desempenho dos Fundos Long-Short no Brasil*. Braz. Rev. Fin. 8 (4), 505–529.
- Gorman, L.L., Weigand, R.A., 2007. *Measuring Alpha Based Performance: Implications for Alpha Focused, Structured Products*, Disponível em <http://allaboutalpha.com/blog/wp-content/uploads/2008/12/measuring-alpha-based-performance-white-paper-by-gorman-weigand.pdf> (Último acesso em 02.02.2013).
- Jensen, M.C., 1968. *The performance of mutual funds in the period 1945–1964*. J. Fin. 23 (2), 389–416.
- Kosowski, R., Timmermann, A., Wermers, R., White, H., 2006. *Can mutual funds “Stars” really pick stocks? New evidence from a bootstrap analysis*. J. Fin. 61 (6), 2551–2595.
- Laes, M.A., 2010. *Desempenho dos Fundos de Investimentos em Ações no Brasil: Mensurando Sorte e Azar Via Procedimentos de Bootstrap*, Temas de Economia Aplicada. Boletim Informações FIPE (Dezembro), 27–32.
- Leusin, L.M.C., Brito, R.D., 2008. *Market Timing e Avaliação de Desempenho dos Fundos Brasileiros*. Revista de Administração de Empresas 48, 22–36.
- Sharpe, W.F., 1966. *Mutual fund performance*. J. Bus. 34, 119–138.
- Treynor, J.L., 1965. *How to rate management of investment funds*. Harv. Bus. Rev. 43, 63–75.
- Treynor, J.L., Mazuy, K., 1966. *Can mutual funds outguess the market?* Harv. Bus. Rev. 44, 131–136.
- Wu, C.F.J., 1986. *Jackknife, bootstrap and other resampling methods in regression analysis*. Ann. Stat. 14, 1261–1350.