UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

VITOR BRAGA DINIZ

ANALYZING STOCK FUNDS METRICS BY MEANS OF CAUSAL RELATIONSHIPS WITH THEIR MANAGEMENT SKILLS

CAMPINA GRANDE - PB

2021

VITOR BRAGA DINIZ

ANALYZING STOCK FUNDS METRICS BY MEANS OF CAUSAL RELATIONSHIPS WITH THEIR MANAGEMENT SKILLS

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Professor Dr. Herman Martins Gomes
Co-orientador: Professor Dr. Luiz Felipe Pontes de Araújo Girão

CAMPINA GRANDE - PB

2021

D585a Diniz, Vitor Braga.
Analyzing stock funds metrics by means of causal relationships with their management skills. / Vitor Braga Diniz. – 2021.

13 f.

Orientador: Prof. Dr. Herman Martins Gomes; Co-orientador: Prof. Dr. Luiz Felipe Pontes de Araújo Girão.
Trabalho de Conclusão de Curso - Artigo (Curso de Bacharelado em Ciência da Computação) - Universidade Federal de Campina Grande; Centro de Engenharia Elétrica e Informática.

1. Métricas em fundos de investimento. 2. Fundos de investimento. 3. Habilidade de gestão – fundos de investimento. 4. Alfa de Jensen. I. Gomes, Herman Martins. II. Girão, Luiz Felipe Pontes de Araújo. III. Título.

CDU:004(045)

Elaboração da Ficha Catalográfica:

Johnny Rodrigues Barbosa
Bibliotecário-Documentalista
CRB-15/626

VITOR BRAGA DINIZ

ANALYZING STOCK FUNDS METRICS BY MEANS OF CAUSAL RELATIONSHIPS WITH THEIR MANAGEMENT SKILLS

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

BANCA EXAMINADORA:

Professor Dr. Herman Martins Gomes
Orientador – UASC/CEEI/UFCG

Professor Dr. Pedro Sérgio Nicoletti
Examinador – UASC/CEEI/UFCG

Professor Dr. Tiago Lima Massoni
Professor da Disciplina TCC – UASC/CEEI/UFCG

Trabalho aprovado em: 25/05/2021

CAMPINA GRANDE - PB

RESUMO

Com a forte alta do número de fundos de investimento no Brasil, há uma necessidade urgente de identificar aqueles que têm melhores habilidades de gestão. Dessa forma, este estudo foi desenvolvido com o objetivo de descobrir quais das métricas mais comuns presentes em fundos de ações brasileiros têm uma relação de causalidade com suas habilidades de gestão, medidas pelo Alfa de Jensen. Para tanto, selecionamos 14 métricas de fundos, a fim de verificar uma relação causal existente entre cada uma e o Alpha. Por fim, indicamos seis métricas que são capazes de ser utilizadas como proxy para a geração alfa.

Analyzing Stock Funds Metrics by Means of Causal Relationships with their Management Skills

Vítor Braga Diniz
Federal University of Campina Grande
Campina Grande, Brazil
vitor.diniz@ccc.ufcg.edu.br

Herman Martins Gomes
Federal University of Campina Grande
Campina Grande, Brazil
hmg@computacao.ufcg.edu.br

Luiz Felipe Pontes de Araújo Girão
TC TradersClub S.A.
São Paulo, Brazil
felipe.pontes@tc.com.br

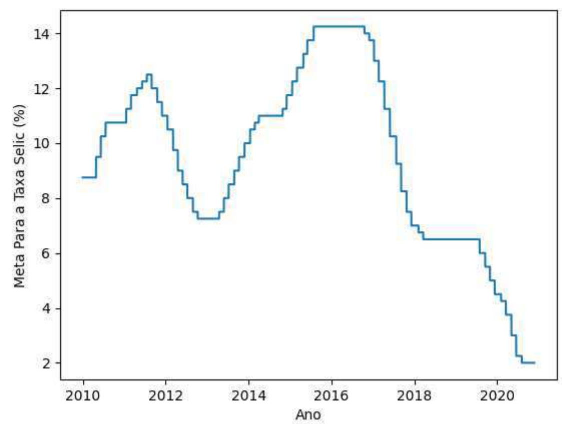
ABSTRACT

With the steep rise in the number of shares of investment funds in Brazil, there is an urgent need to identify those which have better management skills. Thus, this study was developed aiming to discover which common metrics present in Brazilian stock funds have a causal relationship with their management skills, measured by the Jensen's Alpha. For that purpose, we selected 14 stock fund metrics in order to verify an existing causal relationship between each one and the Alpha. At last, we indicate six metrics which are able to be used as proxy for the alpha generation.

ACM Reference Format:
Vítor Braga Diniz, Herman Martins Gomes, and Luiz Felipe Pontes de Araújo Girão. 2021. Analyzing Stock Funds Metrics by Means of Causal Relationships with their Management Skills. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

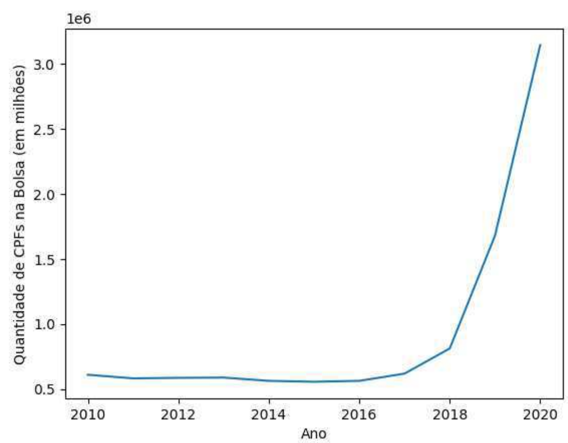
1 INTRODUCTION

After the 2015 Brazilian economical crisis, the Government and the Central Bank of Brazil's efforts to control the fiscal policy allowed the interest rate to be reduced from 14.25% to 6.5% in 2018, reaching the historical minimum level of 2%, in 2020, as shown in Figure 1 [3]. Thereby, investments in fixed income, became less attractive to common investors, who started to migrate to variable income security markets, such as the stock market (Figure 2).



Source: Own Author + [3]

Figure 1: Brazilian Interest Rate Evolution



Source: [6]

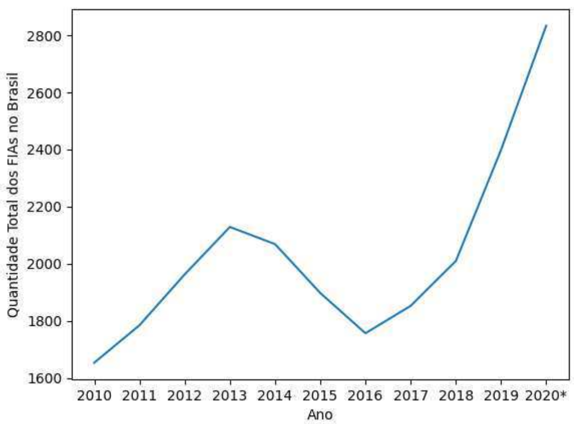
Figure 2: People Registered on Brazilian Stock Exchange (B3)

Moreover, not only have the investors sought to invest directly in public companies, but also through equity funds (stock funds), companies responsible for managing financial assets of shareholders.

Conference'17, July 2017, Washington, DC, USA

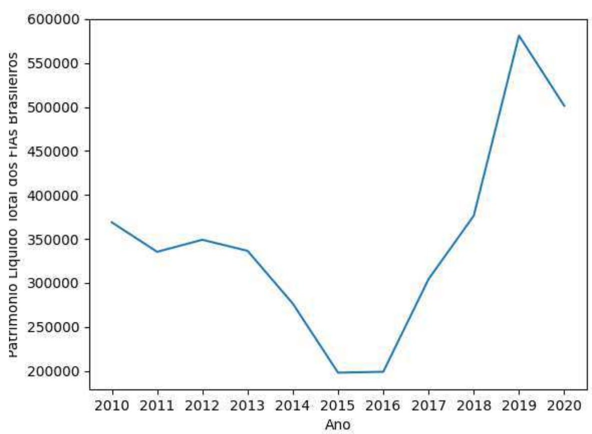
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Hence, this movement generated a strong demand for the service and stimulated both the growth of existing funds and the creation of new funds, as shown in Figures 3 and 4.



Source: [1]

Figure 3: Number of Stock Funds in Brazil



Source: [1]

Figure 4: Total Brazilian Stock Funds Net Worth

However, due to the increase in the number of stock funds, the task to verify and select skilled funds, those which are capable of performing returns above the expected when compared to their risk levels, requires much more effort. In addition, the Jensen's Alpha – metric traditionally used to measure fund management skills, as it will be discussed in Section 2.1 – is not only unusually revealed from those companies, but also it is difficult to be calculated.

Therefore, in order to assist the common investors to make their investment decisions, this study aims to identify which of some of the most common stock fund metrics – those more often displayed or more acknowledged by common investors – can be used to infer if a given fund generates alpha and, as a result, has good management skills.

2 THEORETICAL BACKGROUND

In this section, we discuss the financial and mathematical background necessary to develop this study. In Section 2.1, we understand the regression model we use to calculate Jensen's Alpha. In Section 2.2, we analyse the Autoregressive models used in the ensuing section, in which we discuss the Granger Causality definition and algorithm, used to verify if a given fund metric has a causal relationship with Alpha.

2.1 Capital Asset Pricing Model (CAPM) and Variations

In order to solve the problem of performance evaluation on financial portfolios, Sharpe [29], Lintner [22] and Treynor [33] defined the portfolio risk premium – difference between the expected return and a risk free rate ($E(R_{p,t}) - R_{f,t}$) – as the product of its exposure to the systematic risk and the market risk premium ($E(R_{m,t}) - R_{f,t}$), as the Equation 1 formalizes.

$$E(R_{p,t}) - R_{f,t} = \beta[E(R_{m,t}) - R_{f,t}] + \epsilon_t \tag{1}$$

Where ϵ_t means the regression error.

Nonetheless, Jensen [17] discovered empirically a linear coefficient, called Jensen's α, or simply α (Equation 2), which introduces an inefficiency to the Efficient Market Hypothesis, discussed by Fama [11], which states the reward-to-risk relation. Having a non-null α means that a portfolio may have lower risks and higher expected returns, this non-explicable behavior is the reason why it is also referred as abnormal return.

$$E(R_{p,t}) - R_{f,t} = \alpha + \beta[E(R_{m,t}) - R_{f,t}] + \epsilon \tag{2}$$

Furthermore, [17] also stated that skilled asset managers tend to have a positive α coefficient. Therefore, the investment industry began to use it as a skill metric.

From those definitions, [12] included two extra risk premia (size and value) to the model (Equation 3) and [23] added two more (liquidity and moment risk premia), in a five-factor model (Equation 4). After these modifications, each factor explains a fraction of the portfolio return. The alpha, however, remains as the abnormal portion, not explained by any factors, as noted by Frazzini [13] when calculating the Birkshire Hathaway alpha.

$$E(R_{c,t}) - R_{f,t} = \alpha + \beta[E(R_{m,t}) - R_{f,t}] + \gamma(SMB)_t + \delta(HML)_t + \epsilon_t \tag{3}$$

$$\begin{aligned} E(R_{c,t}) - R_{f,t} = \alpha + \beta[E(R_{m,t}) - R_{f,t}] + \gamma(SMB)_t + \delta(HML)_t \\ + \omega(MOM)_t + \nu(LIQ)_t + \epsilon_t \end{aligned} \tag{4}$$

Where α is the portfolio's alpha; $R_{c,t}$ is the return on the portfolio in period t; $R_{f,t}$ is the risk-free rate in period t; $R_{c,t} - R_{f,t}$ is the excess return on the portfolio in t; $R_{m,t}$ is the market return in t; $R_{m,t} - R_{f,t}$ is the market risk premium in t; SMB_t is the premium for the size factor in t; HML_t is the premium for the risk factor value in t; MOM_t is the premium for the risk factor moment in t; LIQ_t is the premium for the liquidity risk factor in t; ϵ_t is the regression error.

We used the Equation 4 to calculate Jensen's Alpha in this study.

2.2 Autoregressive Models

In this subsection, we discuss the Autoregressive Integrated Moving Average and Vector Autoregression models, which are used to verify a causal relationship between two time series.

2.2.1 ***Autoregressive Integrated Moving Average (ARIMA) Model***. The ARIMA model is formed by the integration of two parts: autoregressive (AR) and moving averages (MA). The AR model, according to Tsay [34], corresponds to statistical models which uses only past data to perform a forecast of future data. In other words, considering a time series X_n, an AR model of X_n uses only the X_{n-i} data, with n being the last value of the series and $i \geq 0$ to determine X_{n+j}, for $j > 0$), according to Equation 5.

$$X_n = \phi_0 + \sum_{i=1}^{p} \phi_i X_{n-i} + \epsilon_n \tag{5}$$

Where ϕ_0 and each ϕ_i are the autoregression coefficients of order p and ϵ_n means the model error (white noise) with zero mean and σ_ϵ^2 variance.

The MA model, nonetheless, assumes that X_n can be defined as the linear combination of white noises, as described in Equation 6 [31].

$$X_n = \epsilon_n + \sum_{i=1}^{q} \theta_i \epsilon_{n-i} \tag{6}$$

Where each θ_i is a non-null parameter resulting from q order modeling.

Thus, combining those definitions, we have the ARMA model (Autoregressive Moving Average), as describes the Equation 7.

$$X_n = \phi_0 + \sum_{i=1}^{p} \phi_i X_{n-i} - \sum_{i=1}^{q} \theta_i \epsilon_{n-i} \tag{7}$$

According to this equation, X_n is an ARMA model and it may be stated as $X_n = \frac{\partial}{\partial n} Z_n$, where Z_n means a non-stationary time series and X_n the differential of first order of Z_n. Generalizing to d differentiations, we obtain:

$$X_n = \frac{\partial^{(d)}}{\partial n^d} Z_n \tag{8}$$

Similarly, we may state Z_n as the $d-th$ integral of X_n, as illustrated in Equation 9 . Therefore, Z_n is defined as an ARIMA model of order $< p, d, q >$.

$$Z_n = \int^d X_n dn^d \tag{9}$$

2.2.2 ***Vector Autorregression (VAR) Model***. Tsay [34] considers a time series r_t as VAR of order 1, VAR (1), if it can be expressed by:

$$r_t = \phi_0 + \Phi r_{t-1} + a_t, \tag{10}$$

Where ϕ_0 is a k-dimensional vector, Φ is a $k \times k$ matrix and a_t is a sequence of uncorrelated vectors of zero mean and positive and defined covariance matrix Σ.

In this study, we apply, in Section 2.3, a bivariate VAR model, since we use a pair <fund metric, Alpha> to understand Alpha's behavior. Thus, considering this case, if we attribute $k = 2$, $r_t = (r_{1t}, r_{2t})$ and $a_t = (a_{1t}, a_{2t})$, the VAR(1) model consists of:

$$r_{1t} = \phi_{10} + \Phi_{11} r_{1,t-1} + \Phi_{12} r_{2,t-1} + a_{1t} \tag{11}$$

$$r_{2t} = \phi_{20} + \Phi_{21} r_{1,t-1} + \Phi_{22} r_{2,t-1} + a_{2t} \tag{12}$$

Where Φ_{nm} corresponds to the n-th and the m-th term of the matrix Φ, respectively, and ϕ_{i0} is the i-th element of ϕ_0 vector. Considering this, we notice that the values Φ_{nm} indicate if a linear dependence between the variables r_{it} and the past values $r_{j,t-1}$ exists. If $\Phi_{12} = 0$ and $\Phi_{21} \neq 0$, then we state a unidirectional relationship from r_1 to r_2; if $\Phi_{12} = 0$ and $\Phi_{21} = 0$, then the series are dissociated; if $\Phi_{12} \neq 0$ and $\Phi_{21} \neq 0$, exists, therefore, a feedback relationship between them.

Hence, it is possible to create a model generalization of order p, VAR (p), defined by Equation 13.

$$r_t = a_t + \phi_0 + \sum_{i=1}^{p} \Phi_i r_{t-i}. \tag{13}$$

Implementing the backshift operator B, we are able to rewrite Equation 13 as:

$$(I - \sum_{i=0}^{p} \Phi_i B^i) r_t = \phi_0 + a_t, \tag{14}$$

Where I corresponds to the identity matrix $k \times k$ and B^i, to the i-th backshift of r_t, or r_{t-i}. Considering $\Phi(B) = I - \sum_{i=0}^{p} \Phi_i B^i$, then the Equation 14 can be reduced to:

$$\Phi(B) r_t = \phi_0 + a_t, \tag{15}$$

Thereby, we have enough means to estimate the Φ_i parameters of the general model using the method of ordinary least squares (OLS). Considering the n-th equation of VAR (n) model, the OLS estimate for each parameter ($\Phi_i^{(n)}$) generates a residual value described as:

$$a_t^{(n)} = r_t - \phi_0^{(n)} \sum_{i=1}^{n} \phi_i^{(n)} r_{t-i}, \tag{16}$$

Which produces the ensuing covariance matrix:

$$\Sigma_n = \frac{1}{T - 2n - 1} \sum_{t=i+1}^{T} a_t^{(n)} [a_t^{(n)}]', n \geq 0. \tag{17}$$

Thus, the general $VAR(p)$ model is calculated and qualified to be used in Granger Causality Tests, as discussed in Section 2.3, presented next.

2.3 Granger Causality

The general definition of causality proposed by Granger [14] states that a vector Y causes another vector X if Y contains information about X that does not exist in a large set of information (A), as described in Equation 18.

$$Prob(X_{n+1} \in A|\Omega_n) \neq Prob(X_{n+1} \in A|\Omega_n - Y_n) \quad (18)$$

However, this definition is not operational, which means it is not able to be used with real data sets. For that purpose, Granger [14] performs a sequence of restrictions based on a restructuring of the general definition. Given that, the author considered the existence of a set of information J_n at time n of all data available about X_t and a third vector Z_t such that $X_t \subset Z_t$ and $Y_t \cap Z_t = \emptyset$.

Granger also defined a set J'_n consisting of J_n plus the elements of Y_{n-j}. Finally, $F(X_{n+1}|J_n)$ denotes the conditional probability distribution function of X_{n+1} for a given J_n, which has an average of $E[X_{n+1}|J_n]$. Considering that and starting from the equation 18, the author stated the ensuing definitions:

Definition 1: Y_n does not causes X_{n+1} with respect to J'_n, if:

$$F(X_{n+1}|J_n) = F(X_{n+1}|J'_n) \quad (19)$$

Definition 2: Y_n causes X_{n+1}, if $J'_n \equiv \Omega_n$ and:

$$F(X_{n+1}|\Omega_n) = F(X_{n+1}|\Omega_n - Y_n) \quad (20)$$

Definition 3: Y_n is *prima facie* cause of X_{n+1} with respect to J'_n, if:

$$F(X_{n+1}|J_n) \neq F(X_{n+1}|J'_n) \quad (21)$$

Definition 4: Y_n does not causes X_{n+1} in mean with regard to a J'_n, if 22 it is identically zero.

$$\delta_{n+1}(J'_n) = E[X_{n+1}|J_n] \quad (22)$$

Definition 5: Y_n causes X_{n+1} in mean, if $\delta_{n+1}(\Omega_n) \neq 0$.

Definition 6: Y_n is *prima facie* cause of X_{n+1} in mean with respect to J'_n, if $\delta_{n+1}(J'_n)$ is not identically zero.

From these definitions, [14], [15] and [16] apply the following algorithm to identify a causal relationship between two stationary vectors Y_t and X_t.

(1) Re-sample Y_t and X_t and build ARIMA models to obtain u_t and v_t vectors;
(2) Examine the cross correlations $\rho_{uv}(k)$ in order to identify evidence of any possible causal relations;
(3) For each possible causal relation, build a bivariate VAR model with u_t e v_t. If a unidirectional cause is suggested, then use the transfer function methods from [5]. If a bidirectional cause is suggested, then use [16]'s method;
(4) Evaluate ARIMA and VAR models to suggest a new model for the original data set.
(5) Evaluate and compare the prediction skills of ARIMA and VAR models in terms of the mean square error (MSE). If the VAR model is significantly better, then the null hypothesis can be discarded and evidence of a causal relationship is found.

Finally, the operational definition of causality between time series states that a vector Y_t causes another vector X_t if it may be used to predict values of X_t.

3 METHODOLOGY

In this section, we explain the origin of our data set (Subsection 3.1) and how we applied the theoretical background to calculate Jensen's Alphas (Subsection 3.2) and the Granger Tests (Subsection 3.4) in order to achieve our objective.

3.1 Data Collection

We divided the data sample in two sets: risk factors and equity funds data. The first one consists of market, size, value, momentum and liquidity risk factors time series extracted from a database provided by the University of Sao Paulo [7]. The second set consists of metrics of Brazilian stock funds which were at least 5 years of lifetime. Those are either extracted from Economatica's database [10] or calculated from those which are, and inspired in the method for variable selection proposed by Mendonça, Campani and Leal [24]. The metrics are: portfolios' returns, fundraising, withdraws, liquid fundraising, tax refund, net worth, lifetime, investors quantity, Sharpe-ratio, Treynor-ratio, Information-Ratio, standard deviation and downside deviation, as detailed in section 3.3.

All time series are in monthly frequency, in the closed interval of January 2^{nd}, 2001 (or minimum available) to December 31^{st}, 2020 (or maximum available) and they were calculated in 12, 24, 36, 48 and 60-month rolling windows.

3.2 Jensen's Alpha

After we formed the database with all necessary data, we calculated Jensen's Alpha based in the 5-factor model (Equation 4) following Algorithm 1:

Algorithm 1: Jensen's Alpha

```
input  : Funds' returns: DataFrame,
         Risk factors: DataFrame,
         Rolling Window: Integer
output : Funds' alpha series: DataFrame
begin
    foreach fundReturns in allFundsReturns do
        data ← innerJoin(fundReturns, riskFactors)
        data ← removeOutliers(data)
        for (i = rollingWindow to data.nrows) do
            rollingData ← data[(i − rollingWindow) : i]
            alphaData ← linearRegression(rollingData)
```

We implemented the linear regression using the Ordinary Least Squares (OLS) method from Seabold and Perktold's library [28], we also performed a 5-fold cross validation for each iteration using a cross_validate function proposed by Pedregosa's [26], in order to reduce bias. Finally, we selected the model with highest accuracy.

3.3 Fund Metrics

As we cited in Section 3.1, this study analyses 14 metrics, of which we use 9 with minimal processing and 5 (Sharpe-ratio, Treynor-ratio, Information-Ratio, standard deviation and downside deviation) we generate from the data extracted. Additional details are provided next:

- Sharpe-Ratio (SR) [30]: is one of the most popular portfolio evaluation metrics, defined by the ratio between the portfolio risk premium ($R_p - R_f$) and its standard deviation (σ_p), as shown in Equation 23:

$$SR = \frac{R_p - R_f}{\sigma_p} \tag{23}$$

- Treynor-Ratio (TR) [32]: is a portfolio evaluation metric defined by the ratio between the portfolio risk premium ($R_p - R_f$) and its systematic risk (β from Equation 4) , as shown in Equation 24.

$$TR = \frac{R_p - R_f}{\beta_p} \tag{24}$$

- Information-Ratio (IR) [2]: is another portfolio evaluation metric, defined by the ratio between the active return (difference between the portfolio's returns and the market's returns – $R_p - R_m$) and the tracking error (standard deviation of the active return – σ_{pm}).

$$IR = \frac{R_p - R_m}{\sigma_{pm}} \tag{25}$$

- Standard deviation(σ) [4]: is a basic statistics error measure, which is often used as risk measurement into the financial environment. It is calculated with the portfolio's returns as the equation 26 defines.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(R_{pi} - \bar{R}_p)^2} \tag{26}$$

  Where N means the number of elements in the sample; R_{pi} means i-th portfolio return and $\bar{R}_p$ means the average return.
- Downside deviation [25]: is a portfolio risk measure similar to the standard deviation, nonetheless, it only considers the values below average in the algorithm 2.

Algorithm 2: Downside Deviation

```
input  : data: List of numbers
output : Downside Deviation: Decimal
begin
    average ← mean(data)
    downside ← 0
    foreach element in data do
        if element < average then
            downside ← downside + (element − average)^2
    downside ← sqrt(downside/data.length)
```

Furthermore, we worked with the remaining metrics with minimal processing, for instance, re-sampling the data from daily to monthly frequency and re-arranging data in rolling windows. These metrics are:

- Lifetime: age in days of a fund;
- Returns: monthly portfolios' returns;
- Fundraising: New capital invested per month in Brazilian Real (BRL);
- Withdraws: Capital withdrew per month in BRL;
- Liquid Fundraising: difference between fundraising and withdraws;
- Tax refund;
- Net worth: difference between assets and liabilities in BRL;
- Investors quantity: number of investors in a fund at the end of a month.

3.4 Granger Causality Tests

For the application of the Granger causality tests, we used the algorithm described in Section 2.3 in the pair alphas and metrics time series of each fund. For that purpose, however, it is necessary to implement a stationary test for each data, since Granger [14] uses stationary time series in his method.

Hence, we applied the Augmented Dickey-Fuller (ADF) test [9], which consists in a study about the Equation 27.

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta y_{t-i} + \epsilon_i \tag{27}$$

Where Δ means the first difference operator; β_1 means the intersection point or drift; β_2 means the trend coefficient; δ means the unit root presence coefficient; m means the number of lags; y_t means the series t-th value.

Thereby, we perform a null hypothesis test, given by $\delta = 0$, where there is a unitary root and the series is not stationary. Thus, when the result (*p-value*) is lower then 0.05, then the null hypothesis may be safely discarded and we are able to apply the Granger test. Otherwise, the ADF test is redone with the series i-th difference (Algorithm 3), for $1 \leq i \leq 5$. Since the difference operator does not guarantee a stationary result, if, after the 5^{th} iteration, y_t persists non-stationary, it will be discarded, as describes Algorithm 4.

Algorithm 3: First Difference

```
input  : data: list
output : dataFirstDifference: list
begin
    dataFirstDifference[0] ← None
    for i = 1 to data.length do
        dataFirstDifference[i] ← data[i]/data[i − 1] − 1
```

After we obtained the stationary time series, we started the causality verification algorithm, discussed in Section 2.3 and implemented by Seabold and Perktold's library [28], which is executed for 1 to the maximum number of lags allowed by the library, according to each alpha-metric pair of series. For each lag, if the p-value

Algorithm 4: Stationary Transform

```
input  : data: array
output : stationaryData: array /* if passed in test */
         None ;                /* otherwise */
begin
    counter ← 0
    isStationary ← False
    while counter < 5 and not isStationary do
        isStationary ← (ADF(data) <= 0.05)
        if not isStationary then
            data ← firstDifference(data)
            counter ← counter + 1
    if isStationary then
        stationaryData ← data
    else
        stationaryData ← None
```

resulting is lower than 0.05, then it is considered statistically significant and that lag is marked as true (strong evidence that the metric causes alpha; otherwise as false (causality may not be confirmed).

Thus, we combined all funds results into frequency scores, which are in a range of [0,1] and are separated per lag and per metric as described in Algorithm 5.

Algorithm 5: Score Formation

```
input  : grangerResults: dictionary ;  /* dictionary of
         dataframes containing results of each
         fund */
output : scores: DataFrame
begin
    scores ← DataFrame() ; /* DF: metrics X lags */
    fundsNum ← fundResults.length
    foreach fundResults in grangerResults do
        foreach metric in fundResults.index do
            foreach lag in fundResults.columns do
                S ← fundResults[metric][lag]
                scores[metric][lag]+ = S
    scores[metric][lag]/= fundsNum
```

Finally, in order to compare the results, we applied the Simpson's rule [8] to calculate the area below each graphic formed by the score curves, as illustrated in Figure 8.

4 RESULTS

Succeeding the method's execution, we generated the scores illustrated in Figure 7 on Appendix A, which displays the fraction of funds, for each metric, where the causal relation may be found.

Thereby, we noted that the chosen metrics effects are not instantaneous, which means the metrics, on average, tend to cause alpha more frequently with some delay.

A second observation was that, when comparing peaks withing the same rolling window, some metrics causes alpha a lot more frequently than others, with highest differences between peaks of 0.715 points (in the 12-month rolling window case).

We also stated a parabolic behavior in all curves, which means there seem to be an optimal point to search causation in each curve and in each window case. We may also infer that metrics not only do not cause alpha instantaneously, but also they do not maintain the causal relationship after a large number of lags.

The fourth observation is an inversely proportional relationship between the window width and the highest scores, as well as the number of lags to the peak occurrence, as illustrated in figures 5 and 6. In other words, as the window width increases, the peak not only tends to happen in a lower delay, but also its score is inclined to decrease. That occurs due to the fact that ancient data tend to cause alpha in a lower frequency, as result of the parabolic behavior.

Thereby, to test a particular metric in 12-month window data, we use up to 12-month-old elements plus the number of lags, since the peak occurs in a lag of 22 months, this means we use up to 34-month-old data. On the other hand, analysing the 60-month window case, we already use 60 months (without any lags) to calculate the metric, which means we consider elements which did not cause alpha in the first measurement, disturbing the results in wider windows.

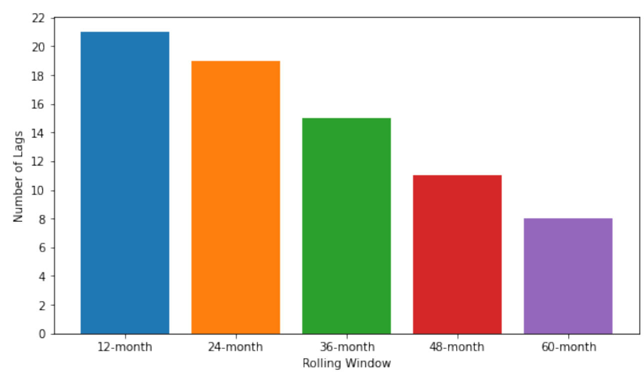


Figure 5: Average Peak Dislocation Between Windows

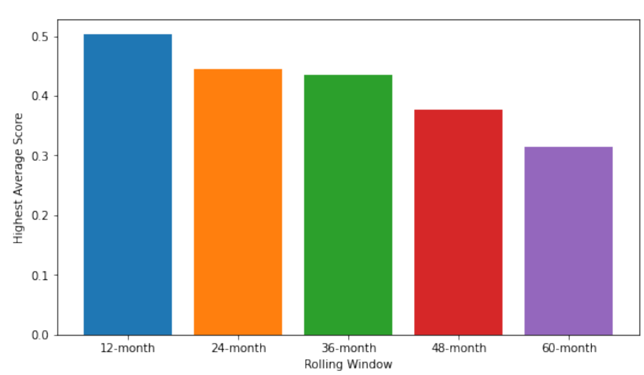


Figure 6: Average Peak Variation Between Windows

After the curve analysis, we gathered each curve into a single area-score, result of each curve area, so we could be able to compare all metrics in a window at once. Thus, we observed that all metrics above average are portfolio evaluation measures, as illustrated in Figure 8 on Appendix B, which means that the main item to analyse in a fund is its historical performance.

5 CONCLUSION

Thus, we are able to state that the these six metrics: Returns, Sharpe-ratio, Treynor-ratio, Information-Ratio, standard deviation and downside deviation are the most indicated variables to be analysed in funds for inferring management skills and alpha, since they present scores above average in all six windows cases.

Therefore, since these metrics are easier be be calculated by the common investors and these are more displayed in funds websites, we achieved our objective to identify variables capable of inferring a fund's Alpha.

A suggestion as sequel for this work is to make Engle-Granger cointegration tests between Jensen's Alpha these six variables to measure the short term values linkages, as realized by [27], and support (or reject) the current results.

After that, another portfolio evaluation metrics could be evaluated, such as the mean–variance–skewness [18], Value at Risk [21], Shortfall Risk [20] and even a portfolio comparison to blindfolded monkeys [19].

The last step would be to define a new Machine Learning model capable of predicting Alpha with an alternative form, beyond the traditional CAPM.

ACKNOWLEDGEMENTS

We thank Dr. Herman Gomes and Dr. Felipe Pontes for the vital guidance on this project; We also thank TC Traders Club, specially the TC-Matrix and TC-School teams, for the immeasurable support.

REFERENCES

[1] Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais. [n.d.]. Consolidado Histórico de Fundos de Investimento. https://www.anbima.com.br/pt_br/informar/estatisticas/fundos-de-investimento/fi-consolidado-historico.htm
[2] Carl R Bacon. 2012. *Practical risk-adjusted performance measurement.* John Wiley & Sons.
[3] Banco Central do Brasil. [n.d.]. Taxas de juros básicas – Histórico. https://www.bcb.gov.br/controleinflacao/historicotaxasjuros
[4] J Martin Bland and Douglas G Altman. 1996. Statistics notes: measurement error. *Bmj* 312, 7047 (1996), 1654.
[5] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. 2015. *Time series analysis: forecasting and control.* John Wiley & Sons.
[6] Brasil Bolsa Balcão. [n.d.]. Quadro De Investidores Pessoa Física. http://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/consultas/mercado-a-vista/historico-pessoas-fisicas/
[7] Brazilian Center for Research in Financial Economics of the University of São Paulo. [n.d.]. Data: Risk Factors. http://nefin.com.br/risk_factors.html
[8] D Corbit. 1996. Numerical integration: From trapezoids to RMS-Object-oriented numerical integration. *DR DOBBS JOURNAL* 21, 10 (1996), 117–+.
[9] David A Dickey and Wayne A Fuller. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association* 74, 366a (1979), 427–431.
[10] Economática. [n.d.]. Economática. https://economatica.com/
[11] Eugene F Fama. 1970. Efficient capital markets: A review of theory and empirical work. *The journal of Finance* 25, 2 (1970), 383–417.
[12] Eugene F Fama and Kenneth R French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* (1993).
[13] Andrea Frazzini, David Kabiller, and Lasse H Pedersen. 2013. *Buffett's alpha.* Technical Report. National Bureau of Economic Research.
[14] Clive WJ Granger. 1980. Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and control* 2 (1980), 329–352.
[15] Clive WJ Granger and Jin-Lung Lin. 1995. Causality in the long run. *Econometric theory* (1995), 530–536.
[16] Clive William John Granger and Paul Newbold. 2014. *Forecasting economic time series.* Academic Press.
[17] Michael C Jensen. 1968. The performance of mutual funds in the period 1945-1964. *The Journal of finance* 23, 2 (1968), 389–416.
[18] Tarja Joro and Paul Na. 2006. Portfolio performance evaluation in a mean–variance–skewness framework. *European Journal of Operational Research* 175, 1 (2006), 446–461.
[19] Yongjae Lee, Do-Gyun Kwon, Woo Chang Kim, and Frank J Fabozzi. 2018. An alternative approach for portfolio performance evaluation: enabling fund evaluation relative to peer group via Malkiel's monkey. *Applied Economics* 50, 40 (2018), 4318–4327.
[20] Martin L Leibowitz and Terence C Langetieg. 1989. Shortfall Risk And The Asset Allocation Decision: A Simulat. *Journal of Portfolio Management* 16, 1 (1989), 61.
[21] Thomas J Linsmeier and Neil D Pearson. 2000. Value at risk. *Financial Analysts Journal* 56, 2 (2000), 47–67.
[22] John Lintner. 1965. Security prices, risk, and maximal gains from diversification. *The journal of finance* 20, 4 (1965), 587–615.
[23] Weimin Liu. 2006. A liquidity-augmented capital asset pricing model. *Journal of Financial Economics* 82, 3 (2006), 631–671.
[24] João Antonio de Mendonça Júnior, Carlos Heitor Campani, and Ricardo Pereira Câmara Leal. 2017. A escolha de fundos de ações e o investidor individual. *Revista de Administração Contemporânea* 21, SPE (2017), 41–62.
[25] David N Nawrocki. 1999. A brief history of downside risk measures. *The Journal of Investing* 8, 3 (1999), 9–25.
[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
[27] Sasipa Pojanavatee. 2013. *An analysis of Australian mutual fund performance and market relationships.* Ph.D. Dissertation. Curtin University.
[28] Skipper Seabold and Josef Perktold. 2010. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference.*
[29] William F Sharpe. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance* 19, 3 (1964), 425–442.
[30] William F Sharpe. 1966. Mutual fund performance. *The Journal of business* 39, 1 (1966), 119–138.
[31] Robert H Shumway and David S Stoffer. 2017. *Time series analysis and its applications: with R examples.* Springer.
[32] Keith V Smith and Dennis A Tito. 1969. Risk-return measures of ex post portfolio performance. *Journal of Financial and Quantitative Analysis* (1969), 449–471.
[33] Jack L Treynor. 1962. Jack Treynor's' Toward a Theory of Market Value of Risky Assets'. *Available at SSRN 628187* (1962).
[34] Ruey S Tsay. 2005. *Analysis of financial time series.* Vol. 543. John wiley & sons.

A CAUSALITY SCORES

On Figure 7, we display our curve scores in graphics plotted in Cartesian planes generated by the axis **Number of Lags** and **Scores**.

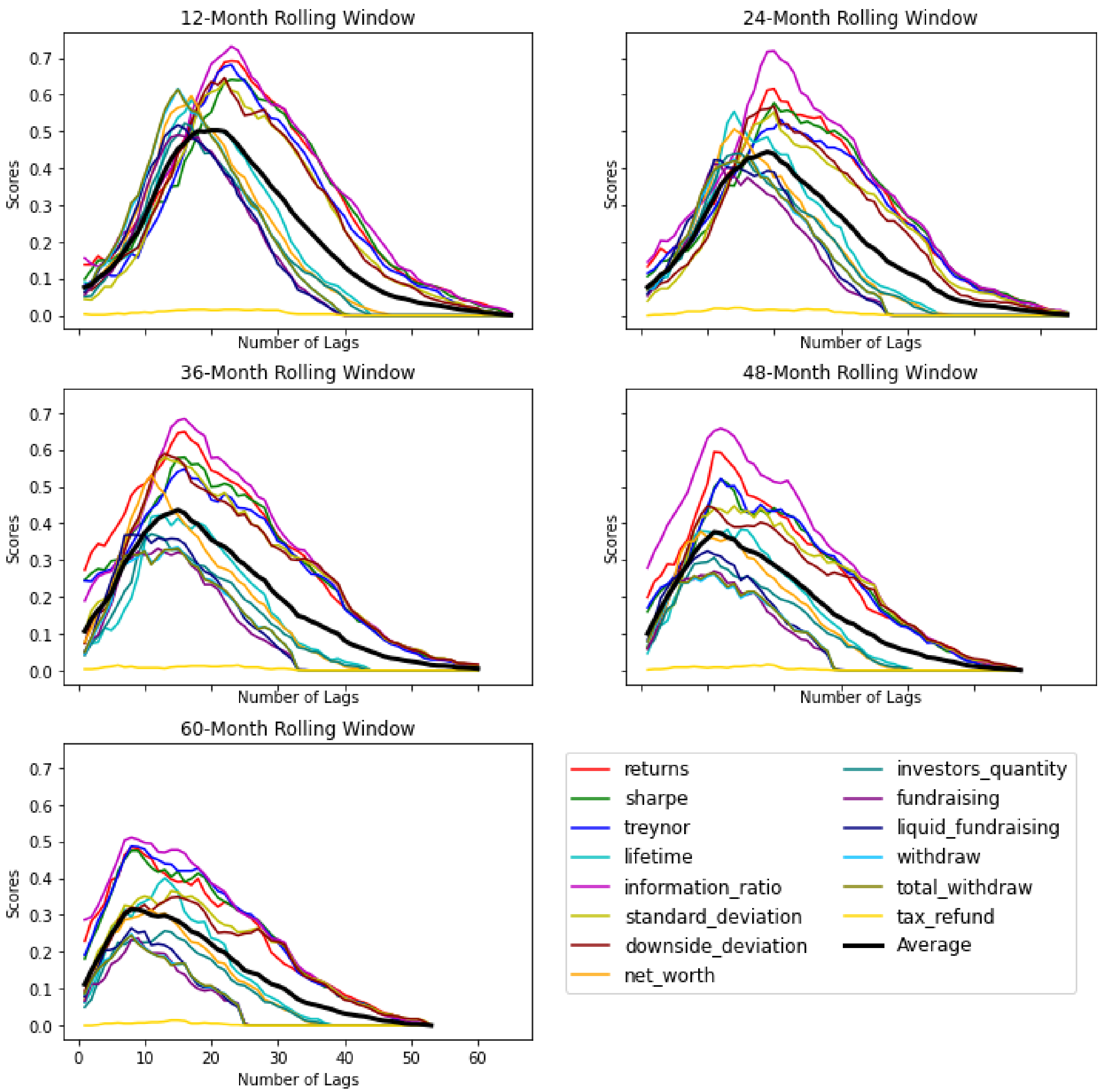


Figure 7: Scores

B AREAS SCORES

On Figure 8, we display our area scores in bar graphics, in which each bar consists in the area below its corresponding metric graphic for each window in Figure 7.

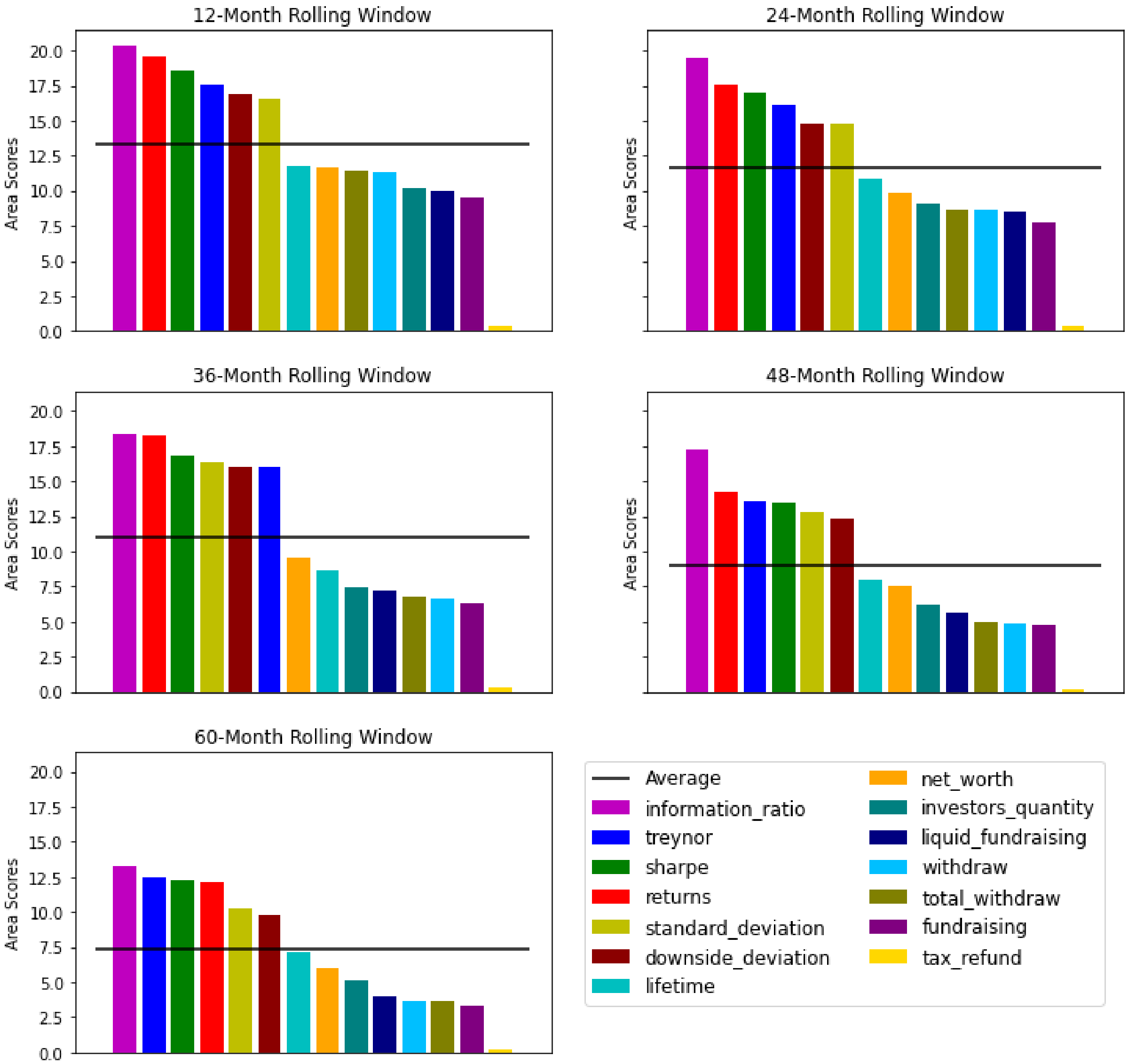


Figure 8: Area-Scores