FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

Tomaz Hamdan Melo Coelho

# Testing the efficient market hypothesis using financially illiterate investors

SÃO PAULO
2021

Tomaz Hamdan Melo Coelho

# Testing the efficient market hypothesis using financially illiterate investors

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.
Orientador: Prof. Dr. Bruno Cara Giovannetti
Coorientador: Prof. Dr. Fernando Chague

SÃO PAULO
2021

Coelho, Tomaz Hamdan Melo.

Testing the efficient market hypothesis using financially illiterate investors / Tomaz Hamdan Melo Coelho. - 2021.

55 f.

Orientador: Bruno Cara Giovannetti.
Co-orientador: Fernando Daniel Chague.

Dissertação (mestrado CMEE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Teoria do mercado eficiente. 2. Investimentos - Análise. 3. Processo decisório. 4. Investidores (Finanças) - Conduta. I. Giovannetti, Bruno Cara. II. Chague, Fernando Daniel. III. Dissertação (mestrado CMEE) – Escola de Economia de São Paulo. IV. Fundação Getulio Vargas. V. Título.

CDU 336.767

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O
Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

Tomaz Hamdan Melo Coelho

# Testing the efficient market hypothesis using financially illiterate investors

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Data de aprovação: __/__/____

Banca Examinadora:

---

Prof. Dr. Bruno Cara Giovannetti
FGV EESP (Orientador)

---

Prof. Dr. Fernando Chague
FGV EESP (Co-orientador)

---

Prof. Prof. Dr. Rodrigo De-Losso
USP-FEA

# Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Quaisquer falhas e omissões são de minha inteira responsabilidade.

"This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001"

# Resumo

Nosso trabalho propõe um teste para a hipótese de mercados eficientes. Propomos um teste simples: colocamos investidores individuais que não sabem nada sobre finanças para negociar no mercado de ações e os comparamos com o resto dos investidores individuais. De acordo com a HME, o desempenho desses investidores antes dos custos de transação e ponderados pelo risco não deve ser diferente. Utilizamos a compra de produtos estruturados comercializados no Brasil que proporcionavam retornos abaixo do ativo livre de risco para classificar os investidores como analfabetos financeiros. Esses investidores possuem um desempenho pior de suas compras. Parte deste resultado é explicado por vieses comportamentais.

**Palavras-chave:**: Hipótese do Mercado Eficiente; Decisões de Investimento; Finanças Comportamentais

# Abstract

Our paper proposes a test for the efficient markets hypothesis. We propose a simple test: we put individual investors who know nothing about finance to trade in the stock market and compare them to the rest of the individual investors. According to EMH the performance of these investors before transaction costs and weighted by risk should not be different. We used the purchase of structured products marketed in Brazil that provided returns below the risk-free asset to classify investors as financially illiterate. Financially illiterate investors show a worse performance of their purchases. Part of this result is explained by behavioral biases.



**Keywords**: Efficient Market Hypothesis; Investment Decisions; Behavioral Finance

# List of Figures


1 **COE Description** . . . . . 16

2 **Financially illiterate - Occupation** . . . . . 25

3 **All investors - Occupation** . . . . . 26

4 **Financially illiterate investors - Diversification Ration** . . . . . 33

5 **Financially illiterate investors - Stocks** . . . . . 34

6 **Financially illiterate investors - Purchases** . . . . . 34

7 **All investors - Diversification Ratio** . . . . . 49

8 **All investors - Purchases** . . . . . 49

9 **Daily ranking** . . . . . 50

# List of Tables


1 **COE - Volume** . . . . . 22

2 **COE - Investors** . . . . . 24

3 **Descriptive statistics** . . . . . 25

4 **Performance - purchases** . . . . . 28

5 **Performance Simulation** . . . . . 30

6 **Diversification Ratio** . . . . . 35

7 **Preference for lottery-like stocks** . . . . . 37

8 **Extrapolation** . . . . . 39

9 **Preference for salient stocks** . . . . . 41

10 **Bias and Performance** . . . . . 42

11 **COE payoff** . . . . . 48

12 **Herfindahl-Hirschman Index** . . . . . 51

13 **Performance Simulation t=1** . . . . . 52

14 **Performance Simulation t=3** . . . . . 53

15 **Performance Simulation t=6** . . . . . 54

16 **Performance Simulation t=12** . . . . . 55

# Contents


**Page**

**1 Introduction** **7**

**2 Literature Review** **10**

2.1 Efficient Market Hypothesis . . . . . 10

2.2 Financially illiterate investors . . . . . 12

2.3 Certificado de Operações Estruturadas . . . . . 13

**3 Methodology** **15**

**4 Testing the EMH** **18**

4.1 Theoretic model . . . . . 20

**5 Data** **22**

**6 Experiment** **27**

6.1 Performance Simulation . . . . . 29

**7 Investment Biases** **31**

7.1 Overconfidence and Sensation Seeking . . . . . 31

7.2 Under-diversification . . . . . 32

7.3 Lottery-like stocks . . . . . 36

7.4 Extrapolation . . . . . 38

7.5 Salient Stocks . . . . . 40

**8 Bias and Performance** **42**

**9 Conclusion** **43**

**References** **44**

**10 Appendix** **48**

# 1 Introduction

Are the tests that attempt to verify the validity of the efficient markets hypothesis (EMH) done until today correct? According to several authors, the answer to this question is no. Until today, the most usual way to test the validity of EMH is based on the search for some strategy that over time could confer superior returns over the market. According to Harvey (2017) when we provide researchers a large amount of data and incentive to find a result, it will be very difficult not to find something. Malkiel (2003) argues that the patterns found are either not robust over time or simply reflect risk factors not previously considered.

Our paper proposes a test for EMH where it will not be necessary to find some strategy in a large amount of data. We propose a simple test: we put individual investors who know nothing about finance to trade in the stock market and compare them to the rest of the individual investors. According to EMH, the performance of these investors before transaction costs and weighted by risk should not be different.

Despite the simplicity of the test some questions proved challenging. First we needed to find investors who know very little about finance, and once found it was necessary that these investors operate in the stock market. To this end, we proposed a clear rule to classify an investor as financially illiterate. We checked an ex-ante action taken by an investor that clearly demonstrated total ignorance of basic concepts of finance. Our classification rule should be clear, not correlated to the object of study, without measurement error and unaffected by the luck factor.

In order to classify investors as financially illiterate we use a financial instrument traded in Brazil. *Certificados de Operações Estruturadas* (COE) are financial assets documented by the literature as good for issuers and bad for investors. Henderson *et al.* (2020) points to evidence of market manipulation in the pricing of some assets. Bitu *et al.* (2021) demonstrated that most of the COEs traded in Brazil present negative Sharpe.

The simple purchase of a financial asset considered bad is not enough to classify an investor as financially illiterate. However, some of these assets provided returns lower than the risk free asset. Thus, we classify the investors who purchased these assets as financial illiterates.

For the purpose of testing the EMH, we tracked the performance of all stocks purchased

between 2016 and 2018 at different time horizons. We noticed that at all time horizons the performance of stocks purchased by investors considered financially illiterate was inferior to the rest of individual investors, even before transaction costs. This result is not in line with the EMH.

Thus, we attempt to understand why the EMH is not valid in this case. The behavioral finance literature points out that cognitive biases are responsible for a worse performance of investors. Odean (1999) demonstrates that investors who performed trades in an excessive manner obtain a worse return due to the transaction costs incurred. Grinblatt & Keloharju (2009) relates the large number of trades carried out by some investors with sensation seeking. Barber & Odean (2000) points out that lack of diversification may be related to risk-seeking or lack of knowledge. Eraker & Ready (2015) demonstrates that investors overpay for stocks with lottery-like payoffs.

In order to check the diversification of the investor's portfolios, we use the Diversification Ratio ($DR$) proposed by Choueifaty & Coignard (2008). This measure indicates the level of diversification of a portfolio. Financially illiterate investors diversify poorly, but we cannot conclude that the lack of diversification explains our result.

We checked out if investors had preferences for lottery-like stocks. We use the methodology proposed by Kumar (2009). Our results point out that financially illiterate investors purchase more lottery-like stocks. We use the methodology proposed by Birru *et al.* (2020) and notice that financially illiterate investors tend to buy stocks with high past returns. This behavior indicates extrapolation bias. We also note that financially illiterate investors conduct more trades than the rest of individual investors.

Finally we find that part of the investor's performance can be explained by the behavioral bias measures used throughout the paper. This result indicates that the non-validity of the EMH may be linked to behavioral biases.

The paper proceeds as follows. Section 2 is composed of a review of the literature on the efficient market hypothesis, financial illiteracy and *Certificado de Operações Estruturadas*. Section 3 presents the methodology applied in the study. Section 4 explains the validity of our test. Section 5 presents the data collected. Section 6 presents the experiment performed.

Section 7 presents the behavioral biases studied. Section 8 presents the relationship between performance and behavioral biases. Section 9 concludes.

# 2 Literature Review

## 2.1 Efficient Market Hypothesis

The idea that the market would be efficient in incorporating new information and translating it accurately into stock prices was widely accepted for a long period of time. The article by Fama (1970) was the initial benchmark for the market efficiency hypothesis to become widespread among economists.

The EMH is associated with the idea that stock prices behave like a random walk. The non-predictability of new news makes prices random. Thus, even an uninformed investor should have a performance on average close to a knowledgeable investor once the risk is adjusted. We can then understand market efficiency as the inability to consistently obtain above average returns once risk is taken into account.

According to Malkiel (2003) at the beginning of the 20th century, the consensus as to the validity of EMH began to be questioned. Some works tried to show that there could be a certain predictability in stock prices. In general, the researchers that tried to question the EMH demonstrated the existence of a certain pattern of stock predictability by means of different approaches or showed that in some periods of time the market behavior deviated from what should be considered efficient.

French (1980) demonstrated that the day of the week could have an impact on stock returns. According to the author, the first day of the week showed higher returns than the rest of the days. Lee *et al.* (1998) shows that January presents a high return in relation to the rest of the months.

Some studies attempt to demonstrate the predictability of stocks through valuation parameters.Campbell & Shiller (1988a) shows that the dividend yield proves to be a good predictor of future stock returns. The same type of analysis was performed by Campbell & Shiller (1988b) with the price-earnings ratio. The predictability in this case points out that stocks with lower price-earnings ratios confer a higher future return.

The predictability was also checked according to the size of the company. Keim (1983) shows that smaller firms give higher returns than larger firms. Some studies, done by Nicholson (1960),

Ball (1978) and Basu (1983) demonstrate that value stocks produced above average returns[1].

Some authors have demonstrated market inefficiency through event studies. For example, Shiller (2000) demonstrates that there was an Internet bubble and at that moment rationality was questionable. Lamont & Thaler (2003) studies the case of companies (Palm and 3Com) that were considered twins but had disconnected prices.

As we can notice, numerous works have been done in order to demonstrate directly or indirectly that the efficiency of markets may not be valid. The controversy continues to this day, and other assets have also been subjected to such a test. Urquhart (2016) and Shen *et al.* (2019) have tested EMH through the predictability of Bitcoin returns.

However, the work done to date has not yet ended the controversy surrounding the EMH. Those who defend the validity of market efficiency believe that it is not yet possible to disprove such a hypothesis. Malkiel (2003) argues that the patterns found are not robust over time or simply reflect risk factors not previously considered. According to the author, even if some patterns may exist, it is very likely that once these are documented, the market dynamics itself will make them disappear.

Many studies have been done in order to counter the possible evidence of non-validity of the EMH. Fama & French (1993) demonstrated that size and price equity value variables can be understood as proxy for risk and not as variables capable of predicting stock returns. Fluck *et al.* (1997) checked whether in fact investors are able to exploit possible predictability patterns present in the market.According to the researchers, patterns such as the price-earnings ratio and dividend yield are not able to provide real opportunities for investors.

Thus, those who seek to question the validity of the EMH look for some strategy that is supposed to generate positive returns for investors. Those who refute this approach point out that perhaps the researchers have found some new risk factor or the results point to statistical rather than economic significance.

The validity of market efficiency is still an unresolved question. One may wonder if the way it has been tested to date is the right one. According to Harvey (2017), no. The author points out that when we give researchers a large amount of data and the incentive to find a result, it

[1]Stocks considered to be value are those with a low price to earnings ratio or a low price to book value ratio.

will be very difficult not to find something. Thus, finding a strategy that possibly generated positive feedback is not the most appropriate way to test the EMH.

In this way, corroborating the criticism made by Harvey (2017), we will not seek a strategy that has produced positive feedback and we will not need any ad hoc questions. Our proposal is simple. We will observe individual investors with no knowledge of the financial market and check their performance even before the transaction cost relative to other individual investors. According to the EMH, we should see no significant difference between these two groups of investors. It is worth noting that we are comparing individual investors. In this way we avoid contaminating our results by possible professional investors. Our challenge then remains to find investors who know nothing about the financial market and still have a presence in it.

## 2.2 Financially illiterate investors

Lusardi & Mitchell (2011) points out that with the shift of the retirement regime, previously captained by governments, into the hands of individuals has made the ability of individuals to understand concepts of finance relevant to researchers and governments. According to Lusardi & Mitchell (2007) individuals with little knowledge in finance save less. Hilgert *et al.* (2003) suggests that knowledge of concepts about credit, savings, and investment are positively correlated with good management of these indicators.

Thus, in the face of the new need to save for retirement it has become even more important that individuals master basic concepts of finance. Faced with the impact of financial education on individual's lives, the OECD has set out to mitigate this problem. According to the organization, financial education should be understood as:

*The process by which financial consumers/investor improve their understanding of financial products and concepts and, through information, instruction, and/or objective advice, develop the skills and confidence to become more aware of financial risks and opportunities to make informed choices, to know where to go for help, and to take other effective actions to improve their financial well-being."*

According to Lusardi & Mitchell (2011), the capabilities described above can be summarized into basic skills that individuals should have. Three are the main concepts to check an individual's ability to have financial capability: i) compound interest, ii) inflation and iii) risk.

Although these concepts seem simple, Lusardi & Mitchell (2007) points out that the lack of knowledge about finance is recurrent in the general population. Even in developed countries the lack of knowledge of these three basic concepts is present. There is also the challenge of correctly measuring these skills. Usually, researchers submit a sample to a series of questions in order to measure each individual's ability. Lusardi & Mitchell (2011) points out that this mechanism contains measurement error since people can simply guess the answers or even not understand the questions as they are asked. The researchers show that by simply reversing the words on the questionnaire, the individual's ability to get it right doubled.

Thus, we realize that our experiment is challenging. We want to be able to point out that an individual has low knowledge in finance. The usual way of applying a questionnaire, besides being a costly undertaking, is ineffective because it contains measurement error. A possible solution is to verify some action of an individual that would point out a clear lack of knowledge of basic concepts of finance.

## 2.3 Certificado de Operações Estruturadas

Structured products are financial assets that have a payoff function which varies automatically and nonlinearly from one or more underlying assets. This type of financial product has emerged in Europe but is also popular in Asia and in the U.S. Brazil regulated this financial instrument in 2013 in the form of the financial product known as *Certificado de Operações Estruturadas* (COE).

This type of financial product is usually structured with derivatives and allows the investor to be exposed to many different assets: stocks, indexes, commodities, foreign exchange. The way this instrument is structured allows the investor to generally know his return in the best and worst case scenario.

Although this instrument is relatively new, some works has already shown that this financial instrument is often constructed in a way that prejudices the investor in favor of the

issuer. Célérier & Vallée (2017) shows that issuers seek to capture investor's desire for return. Furthermore, the authors show that there is a positive correlation between product complexity and risk. In this case, riskier products proved to be more profitable for issuers.

Given the great asymmetry of information between the parties involved, the COE often allows the issuer to obtain unfair benefits. Henderson *et al.* (2020) points to evidence of market manipulation in the pricing of some assets. For some structured products the closing price of some underlying asset on some specific days influences the payoff of investor's returns. For such securities, the authors demonstrate that there is abnormal order flow near the close on the relevant days.

Wilkens *et al.* (2003) demonstrates that structured products are priced above their theoretical value. The study shows that there is a relationship between overpricing and product complexity. Rathgeber & Wang (2011) points out that less transparent markets tend to have products with higher overpricing.

Despite having been instituted in Brazil in 2013, it was only in 2015 that CVM, through instruction 569, regulated how the distribution of these assets should be conducted. From 2015 to the end of 2019, more than R$39 billion of these assets have been distributed by 20 issuers to 315,434 investors.

A few studies were conducted in Brazil about this financial instrument. The study done by Bitu *et al.* (2021), through the simulation of returns by the Monte Carlos method, demonstrated that most of the COEs traded in Brazil present negative Sharpe. When the analysis is carried out ex-post, the results also show that the performance of COES in Brazil, on average, is poor. De Genaro (2021) demonstrates that the average profitability of COES, when weighted by the volume of issue, is 91 Bps below the CDI.

# 3 Methodology

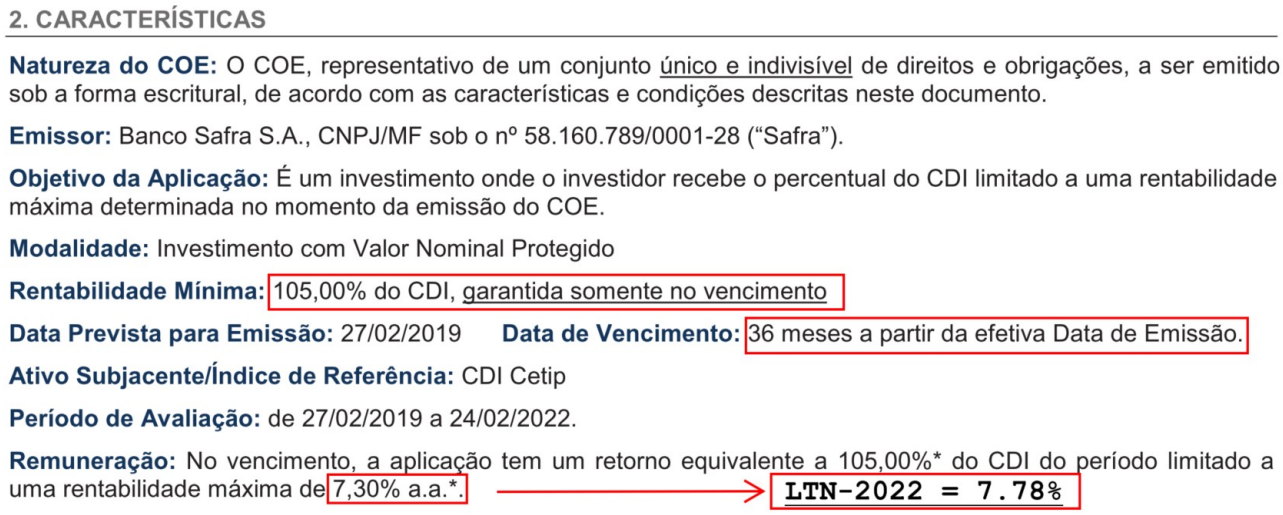
In order to perform the desired experiment we need to find people with low knowledge of finance who trade on the stock market. Finding such people proves to be a challenging task. As specified by Lusardi & Mitchell (2011) the application of questionnaires to measure knowledge in finance contains measurement error. Another possible alternative would be to search for the worst performing investors in the stock market over a period of time. However, this alternative also contains problems. The first is that past investor performance would be correlated with our object of study. The second problem consists in the fact that this investor may have good knowledge in finance, but may have been unlucky.

Thus, we look for some past action of the investor that clearly demonstrates a lack of knowledge of basic aspects of finance and that does not incur in the problems mentioned above. One of the basic aspects of finance defined by Markowitz (1952) is that the investor should seek to maximize his expected return and minimize risk. The author defines in this case the risk free asset as the most conservative decision from the portfolio selection point of view. Thus, choosing an asset that provides a return lower than the risk free asset in all possible scenarios would be an irrational choice by the investor. This choice would be even more serious if such asset had a return lower than the risk free asset even though it has a higher implicit risk.

In this case, we need to find some financial asset that provides a lower return than the risk-free asset no matter what the scenario is. So we need to find an investor who made this bad investment decision ex-ante. That is, this investor certainly did not have a bad performance by a random question. Once we find this individual, it is important for us to conclude our experiment that this individual operates in the stock market.

We have already seen that the performance of COES is poor regardless of whether we perform an ex-post or ex-ante analysis. However, it is still not possible to state categorically that an investor who buys this type of product is a person with little knowledge of finance. The ideal would be to find an investment that ex-ante is terrible and that the luck factor has no influence on the final result.

Figure 1: **COE Description**




The figure 1 is part of the description of a COE marketed in Brazil in 2019. For a number of reasons this product can be considered a very bad investment. We can see that this COE is linked to the CDI, but its return cannot be higher than 7.30%. Although this investment fits into the category of capital protection[2], it is important to point out that this investment does not have the protection of the *Fundo Garantidor de Crédito* (FGC) and has no liquidity in the secondary market.

Therefore, we have an investment that will provide the investor with a maximum return of 7.30% with a maturity of three years, but that does not count on FGC protection and does not count on liquidity in the secondary market. It is common practice in investment analysis to compare the expected return of a given investment with the risk free asset asset. Thus, we verify that on the same day of the purchase of this COE the investor could have acquired a *Letra do Tesouro Nacional* (LTN) with maturity very close to 3 years and which would provide a certain return of 7.78%. .

Assets with negative Sharpe can be considered bad investments. However, one can argue that ex-post the investor can obtain a return above the risk free rate. However, we are facing an investment that in no possible scenario will have a return higher than the risk-free asset

[2]Most of the COES marketed in Brazil are in the protected capital category, but the investor only receives the principal of the investment upon maturity. However, this capital is not protected against inflation. In many cases the investor has a negative real return.

besides having a higher risk. Add this issue to the lack of liquidity in the secondary market and the absence of the FGC's guarantee and we can claim that only an illiterate investor would make this investment.

Notice that we then have a clear identification method that we believe has little noise to measure whether an investor knows the basics of finance. We do not need to submit questionnaires or check performance ex-post. We will only check if the investor has made a bad investment ex-ante. For the conclusion of our experiment it is important that these investors also trade on the stock market.

# 4 Testing the EMH

The idea that prices fully reflect available information implies that successive changes in asset prices are independent. Moreover, returns are assumed to be identically distributed. Putting these two characteristics together implies that stock returns follow a random walk. Thus, the non-rejection of the hypothesis that prices follow a random walk is associated with the non-rejection of the EMH. Stated formally, given a time series $\{y_t\}_{t=1}^{T}$, the random walk hypothesis is formed when $\theta = 1$ by the $AR(1)$ model[3]:

$$y_t = \mu + \theta y_{t-1} + \epsilon_t$$

Given that the current price of the asset in an efficient market condenses all available information about its foundation value, and is therefore the best estimate of this value. Therefore the best estimate of this value, changes in this price occur only when new relevant information arises that affects expectations about future returns and the expectations about future returns and the perception of risk and liquidity of the asset.

Fama (2021) also describes the sufficient conditions for market efficiency: (i) there are no transactions costs in trading securities, (ii) all available information is costlessly available to all market participants, and (iii) all agree on the implications of current information for the current price and distributions of future prices of each security.

The HME can be scaled into categories according to the level and the speed of information incorporation by prices. In the literature there are three main forms: (i) weak, in which current prices reflect all available past public information; (ii) semi-strong, which resembles the weak form, but there is the fact that prices instantaneously adjust to new public information; and (iii) strong, which in addition to the characteristics of the semi-strong form, considers that prices already reflect even secret information.

Fama (1970) presented how the be to identify the efficiency of markets. For each form, a different type of test was devised. To identify the weak form, the tests seek to measure how well past returns predict future returns. If correlations of prices with any type of variable can

[3]When the error terms are not an i.i.d. sequence, the random walk is called a martingale process

be found, then the market is considered to be inefficient in its weak form.

For the semi-strong form, the tests seek to specify how quickly asset prices reflect public information. The faster the price adjustment as a result of a particular event, the more efficient efficient market, as it offers less opportunity for opportunity to make abnormal profits by exploiting this information. exploitation of that information.

Finally, there are the tests to evaluate efficiency in the strong form, in which prices reflect not only public information, but all information that can be obtained, including so-called privileged information. The tests of this form of efficiency efficiency tests seek whether any investor possesses privileged information, which is not fully reflected in prices, and whether they could benefit from such information, obtaining abnormal profits.

In his most recent work Fama (2021) provides a historical survey of the testing performed until today. According to the author, only the strongest form of the EMH has been effectively challenged.

*At the moment, however, corporate insiders and specialists are the only two groups whose monopolistic access to information has been documented. There is no evidence that deviations from the strong form of the efficient markets model permeate down any further through the investment community. For the purposes of most investors the efficient markets model seems a good first (and second) approximation to reality*

That said, why can our test be considered valid? We consider the criticism made by Malkiel (2003) that the tests so far would not be robust over time and that differences in performance can be explained by the risk propensity of investors. We consider in our experiment the risk propensity of the investor since we will control for ticker fixed effects. We also control for possible time fixed effects. Thus, our results reflect only the investors' ability to buy, regardless of which asset is being purchased.

Our results are also not influenced by transaction costs and the comparison we perform is only for individual investors. In this way, we want to show that before transaction costs and without the influence of risk factors a specific group of individual investors consistently under-performs the rest of the investors.

It is important to note that Grinblatt *et al.* (2009) also found a difference in performance

with respect to buying stocks among different investors. The author shows that high-IQ investors outperform low-IQ investors by 45 basis points after one month of purchase. In this case, the author argues that EMH violation may occur if low IQ investors make frequent purchases. Otherwise the performance difference is not significant enough to violate the EMH.

## 4.1 Theoretic model

To better understand the validity of our test we also propose a simple theoretical model. We will use a discrete theoretic model with 2 different agents and with n discrete periods. Denote:

$$\begin{aligned} t_n &= n\Delta t = \text{ discrete times} \\ x_n &= \text{ position at time } t_n \\ d_n &= \text{ step at time } t_n \\ &= x_n - x_{n-1}. \end{aligned}$$

We assume that the initial position is $x_0 = 0$ and that $d_n$ is simple binomial, i.e.

$$d_n = \begin{cases} 1 & \text{prob } = \theta \\ -1 & \text{prob } = 1-\theta \end{cases}$$

Note that if $\theta = \frac{1}{2}$ then $\bar{d} = E(d) = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$. In this case we are closer to the environment of the EMH. We assume that we have $H$ agents, where $H_{fi}$ indicates the financially illiterate investors and $H_{ri}$ indicates all other investors ($H = H_{fi} + H_{ri}$). We also assume that:

$$d_n^{fi} = \begin{cases} 1 & \text{prob } = \theta^{fi} \\ -1 & \text{prob } = 1-\theta^{fi} \end{cases}$$

We may prove that if $\sum_{i=1}^{n}\sum_{j=1}^{H_{fi}}(d_i^{fi}) < 0$ and $\sum_{i=1}^{n}\sum_{j=1}^{H_{ri}}(d_i^{ri}) > 0$ then $\theta^{fi} \neq \theta^{ri} \neq \frac{1}{2}$. That is, if financially illiterate investors obtain negative results and the rest of the investors obtain positive results, then we are not dealing with a random walk with an expected result equal to zero.

Now observe that $x_n = \sum_{i=1}^{n} d_i = \#\left(d_i = 1\right) - \#\left(d_i = -1\right) = \#\left(d_i = 1\right) - \left(n - \#\left(d_i = 1\right)\right) = 2\#\left(d_i = 1\right) - n = 2k - n$. Where $k = \#\left(d_i = 1, i \leq n\right)$. Therefore, for a random walk with steps $\pm 1$ with probabilities $(\theta, 1-\theta)$, we may have the following probability:

$$p\left(x_n = 2k - n\right) = 2^{-n}\theta^k(1-\theta)^{(n-k)} \begin{pmatrix} n \\ k \end{pmatrix} \text{ for } 0 \leq k \leq n$$

Note that the expected value is given by : $\bar{d} = E(d) = \theta + (1-\theta)(-1) = 2\theta - 1$. We now may apply the expectation operator in our first inequality. Then we have the following condition:

$$\mathbb{E}[\sum_{i=1}^{n}\sum_{j=1}^{H_{fi}}(d_i^{fi})] < \mathbb{E}(0)$$

$$H_{fi}(2\theta_{fi} - 1) < 0$$

Then we get that $\theta_{fi} < \frac{1}{2}$. We also get that $\theta_{ri} > \frac{1}{2}$.

# 5 Data

To perform this work two databases provided by the *Comissão de Valores Mobiliários* (CVM) were employed. The first database contains information on all COES that were traded in Brazil from October 2015 to December 2019. This database contains information both at the level of each product issued and also at the investor level.

Table 1: **COE - Volume**

This table presents an overview of the sample characteristics. The database consists of all the Coes that were issued in Brazil between October 2015 and December 2019. Each panel presents the 5 most common characteristics grouped by issuer, year, type and asset. The volume is presented in millions of reais and the percentage represents the share in relation to the total sample. The total volume issued in the period was R$ 39.47 billion.

| | Volume ( millions) | (%) | | Volume ( millions) | (%) |
|---|---|---|---|---|---|
| **Painel A: By Issuer** | | | **Painel C: By type** | | |
| Banco Santander | 8,494.64 | 21.52% | Call | 14,438.77 | 36.57% |
| Banco Safra | 6,990.36 | 17.77% | Digital Call | 6,273.14 | 15.89% |
| Banco ItaŁ | 5,909.85 | 14.97% | Call Spread | 5,998.53 | 15.19% |
| Banco Morgan Stanley | 5,673.38 | 14.37% | Call KO | 3,299.59 | 8.35% |
| Banco BNP Paribas | 3,588.84 | 9.09% | Digital Put | 1,533.62 | 3.88% |
| **Painel B: By year** | | | **Painel D: By asset** | | |
| 2015 | 1,305.12 | 3.30% | National index | 19,246.72 | 48.75% |
| 2016 | 7,243.43 | 18.35% | Exchange | 7,909.73 | 20.03% |
| 2017 | 9,549.96 | 24.19% | Basket | 5,932.18 | 15.02% |
| 2018 | 11,062.24 | 28.02% | International index | 4,583.00 | 11.61% |
| 2019 | 10312.41455 | 0.2612 | International stock | 1,197.42 | 3.03% |

We observe from the table above that the market is concentrated in a few issuers. In addition, different types of products were issued. The types of each product are related to their payoff function[4].

We perform the same analysis at the investor level. It is important to point out that despite

[4]In the appendix we present a table with a brief description of the COE types with the highest volume of issuance.

issuing COES, most of these products are sold by brokerage firms. Until December 2018, the broker that carried out most of the sales of the COES could not issue them. In this way, the concentration of a large number of investors in Morgan Stanley does not mean that the bank sold directly to its clients.

As shown in table 1, we have noticed an increasing number of investors in this type of product over the years. According to B3, in 2019 the Brazilian stock market had about 1.6 million active investors. Thus, the COES market represents around 20%, in terms of number of investors, of the Brazilian stock market. When compared with the number of active investors in the *Tesouro Direto*, this percentage is similar[5].

[5]According to the Ministry of Economy in January 2019 Brazil had about 1.4 million active investors in *Tesouro Direto*.

Table 2: **COE - Investors**

This table presents an overview of the sample characteristics. The database consists of all Coes that were issued in Brazil between October 2015 and December 2019. Each panel presents the 4 most common characteristics grouped by issuer, year and type . The database is formed by 315,434 investors.

| | Investors | | Investors |
|---|---|---|---|
| **Painel A: By Issuer** | | **Painel B: By Year** | |
| Banco Morgan Stanley | 126,851 | 2015 | 2,418 |
| Banco BNP Paribas | 63,876 | 2016 | 60,868 |
| Banco Santander | 56,667 | 2017 | 97,724 |
| Banco Citibank | 28,229 | 2018 | 139,123 |
| Banco BTG Pactual | 25,591 | 2019 | 107,594 |
| **Painel C: By type** | | **Painel D: By Asset** | |
| Call | 169,705 | National index | 157,785 |
| Digital Call | 126,169 | Basket | 119,300 |
| Call Spread | 37,920 | International index | 88,765 |
| Call KO | 22,385 | Exchange | 34,711 |
| Digital Put | 21,040 | International stock | 8,770 |

In the table below we verify the profile of the investors that are in our treatment group. We can see that the investors who are considered financially illiterate are slightly older than the rest of the investors and have been in the stock market longer[6]. We can also note that these investors performed a larger number of operations.

[6]We define experience as the difference between the time of the investor's registration and the deadline of our analysis

Table 3: **Descriptive statistics**

| | Mean | Std.Dev | Median | 1st Qu. | 3rd Qu. | N |
|---|---|---|---|---|---|---|
| **Panel A: Financially Illiterate Investors** | | | | | | |
| Age | 59.1 | 13.6 | 59 | 49 | 69 | 1,262 |
| Experience | 8 | 5.37 | 7 | 4 | 12 | 1,262 |
| Number of trades (Average) | 37 | 66 | 13 | 3 | 53.8 | 1,262 |
| **Panel B: All Investors** | | | | | | |
| Age | 44.2 | 13.4 | 41 | 34 | 52 | 480,343 |
| Experience | 6.1 | 4.4 | 4 | 3 | 8 | 480,343 |
| Number of trades (Average) | 12.1 | 51.3 | 2.7 | 0.6 | 9.3 | 480,343 |

Our database also allows us to verify the professional occupation of the investors. We can note that the investors in the treatment group do not have a predominance in a certain area of activity. We also noticed that there is no significant difference among the occupations of the control and treatment groups.

Figure 2: **Financially illiterate - Occupation**

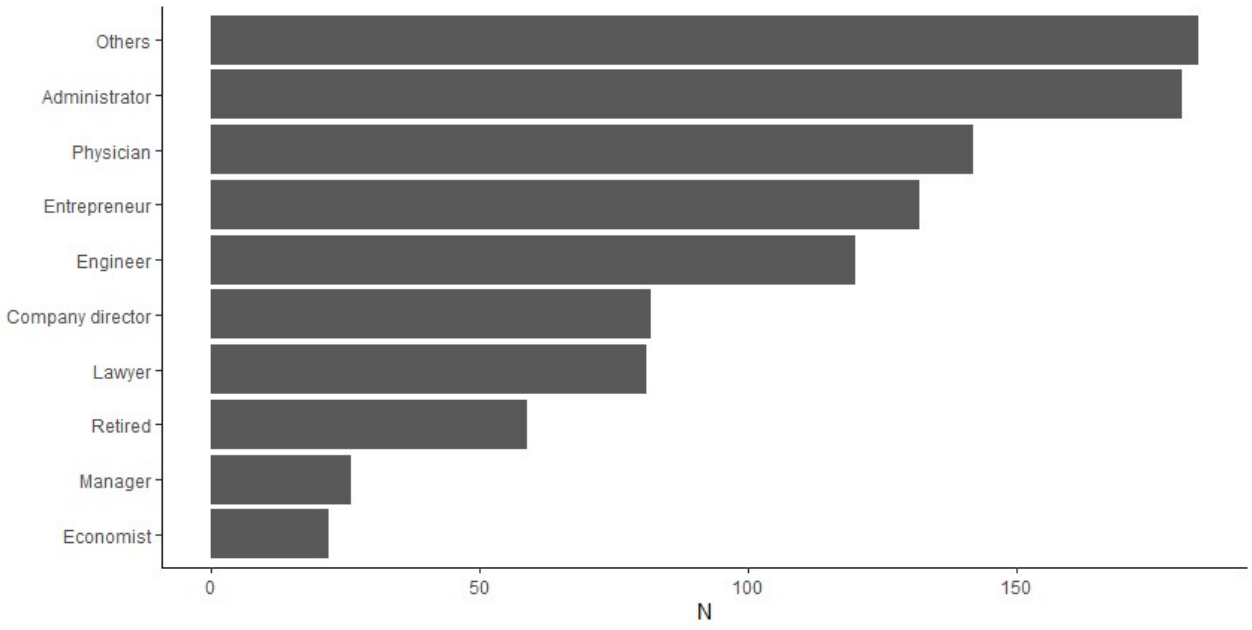
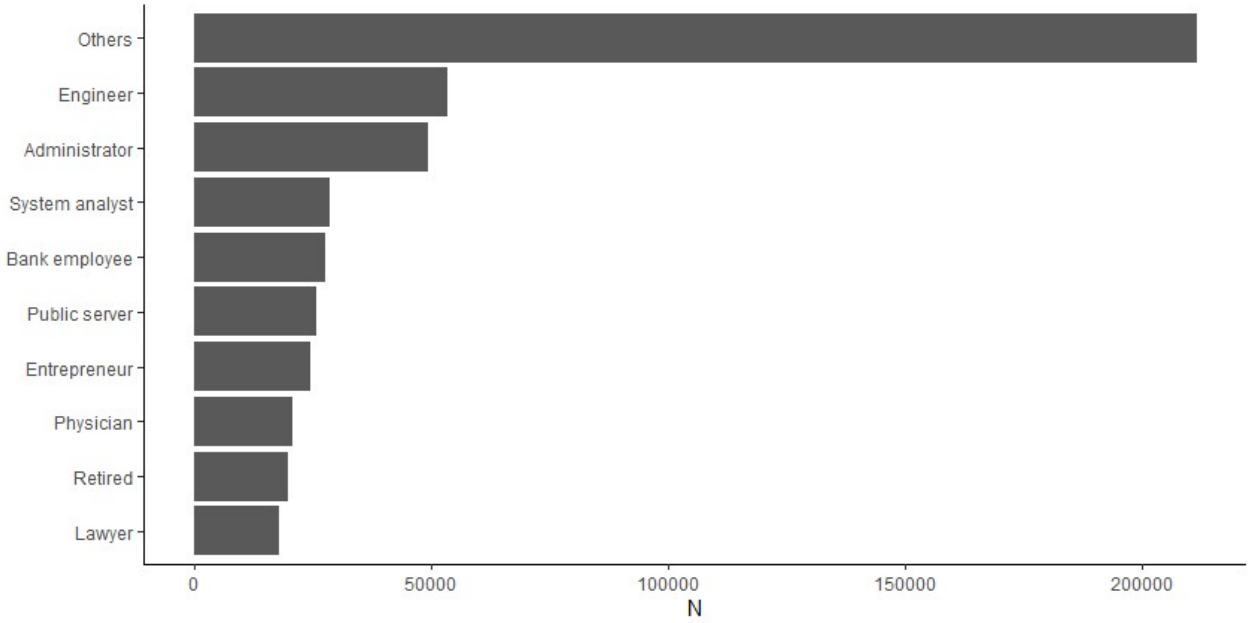


Figure 3: **All investors - Occupation**



# 6 Experiment

Our experiment consists in following the behavior of these investors defined as financial illiterate in the stock market. We will verify the asset selection capacity of such investors. Whenever they make a purchase on the stock exchange, we will follow up the performance of these assets in different periods. Thus, we have defined the following indicators:

$$Rent_{i,j,t+60} = \left( \frac{\sum_n^N Q_n^{i,j,t+60} \bar{P}^{j,t+60}}{\sum_n^N Q_n^{i,j,t} \widetilde{P}^{i,j,t}} \right) - 1 \tag{1}$$

$$Rent_{i,j,t+120} = \left( \frac{\sum_n^N Q_n^{i,j,t+120} \bar{P}^{j,t+120}}{\sum_n^N Q_n^{i,j,t} \widetilde{P}^{i,j,t}} \right) - 1 \tag{2}$$

$$Rent_{i,j,t+180} = \left( \frac{\sum_n^N Q_n^{i,j,t+180} \bar{P}^{j,t+180}}{\sum_n^N Q_n^{i,j,t} \widetilde{P}^{i,j,t}} \right) - 1 \tag{3}$$

$$Rent_{i,j,t+240} = \left( \frac{\sum_n^N Q_n^{i,j,t+240} \bar{P}^{j,t+240}}{\sum_n^N Q_n^{i,j,t} \widetilde{P}^{i,j,t}} \right) - 1 \tag{4}$$

Where, $Q_n^{i,j,t}$ represents all possible $N$ purchases made by individual $i$ of asset $j$ on day $t$, $\bar{P}^{j,t+60}$ is defined as closing price 60 days after the purchase of asset $j$ and $\widetilde{P}^{i,j,t}$ is defined as the average purchase price of all assets $j$ purchased by individual $i$ on day $t$. Thus, we check the profitability at different time horizons of each asset purchased by each individual.

Therefore, our econometric analysis measures whether there is a difference in performance between the financially illiterate investors and the control group. It is important that we are able to mitigate possible time fixed effects and also possible ticker effects. Thus, we estimate the following equations[7]:

$$Rent60_{itj} = \alpha + \beta_1 Financially\_Illiterate_i + \delta_t + \gamma_j + \varepsilon_{ijt} \tag{5}$$

$$Rent120_{itj} = \alpha + \beta_1 Financially\_Illiterate_i + \delta_t + \gamma_j + \varepsilon_{ijt} \tag{6}$$

[7]The errors were clustered at the time and ticker level.

$$Rent180_{itj} = \alpha + \beta_1 Financially\_Illiterate_i + \delta_t + \gamma_j + \varepsilon_{ijt} \quad (7)$$

$$Rent240_{itj} = \alpha + \beta_1 Financially\_Illiterate_i + \delta_t + \gamma_j + \varepsilon_{ijt} \quad (8)$$

The table below provides the data from the estimation performed. We can observe that over time the performance of investors considered financially illiterate is significantly lower than that of the rest of the investors. This result is not consistent with EMH. Even before transaction costs, investors with low knowledge in finance obtained a worse performance in relation to their purchases.

Table 4: **Performance - purchases**

This table presents the results of the difference in performance, at different time horizons, in relation to the purchases by investors considered financially illiterate and the rest of the investors. The estimation considers time and ticker fixed effects. The period covered is 2016 to 2018.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Rent60 | Rent120 | Rent180 | Rent240 |
| *Financially Illiterate* | -0.0105 | -0.0315** | -0.0547** | -0.0578** |
| | (0.00754) | (0.0137) | (0.0211) | (0.0284) |
| *Constant* | 0.101*** | 0.191*** | 0.278*** | 0.355*** |
| | (1.65e-05) | (2.95e-05) | (5.83e-05) | (0.000104) |
| *Date fixed effects* | YES | YES | YES | YES |
| *Ticker fixed effects* | YES | YES | YES | YES |
| *Observations* | 9,045,004 | 9,042,004 | 9,040,033 | 9,028,619 |
| *R-squared* | 0.420 | 0.451 | 0.464 | 0.430 |

*Robust standard errors in parentheses, clustered at the time and ticker level*

**** p<0.01, ** p<0.05, * p<0.1*

We compare the performance of individual investors who are considered financially illiterate by means of a direct choice criterion with other individual investors. Our result performed this comparison controlling for time and ticker fixed effects so that our result reflects only the asset choice ability of different investors. The result in is not affected by possible transaction costs incurred by investors. Thus, we performed an experiment that is based on important points: i) clear choice criteria, ii) absence of time dependence and possible risk factors, iii) no cherry picking.

The experiment therefore attempted to take into consideration the criticisms that have been made of the works that attempted to test the EMH. Our result points to the non-validity of the EMH. We shall try to understand why this hypothesis is not valid.

## 6.1 Performance Simulation

To validate the result obtained in Table 4, we propose an investment strategy based on the behavior of financially illiterate investors. In this way, we want to demonstrate that it is possible to obtain abnormal returns based on the behavior of a set of investors. The portfolios were built according to different time frames and refreshed monthly.

We checked whether in a given time frame the financially illiterate investors were net buyers or net sellers of a given stock. When the stock position of a given stock was long for this set of investors the portfolio was composed of a short position in this stock. We used the same logic when investors were short a certain stock. Once the portfolio was constructed, we calculated the weight of each stock to obtain an equal-weighted portfolio and a value-weighted portfolio.

Finally, we use the Fama-French Three-factor Model to check whether the strategy provide abnormal returns. As we can notice, Table 5 indicates that for all time frames the strategy produces abnormal returns[8].

[8]The data was obtained from Nefin. The complete results can be found in the appendix.

Table 5: **Performance Simulation**

This table presents the alpha of each portfolio for different time horizons. The portfolios were constructed to simulate the opposite behavior of financially illiterate investors. The portfolios are refreshed monthly and the parameter $t$ indicates the time frame used to construct the portfolio.

| | Equal Weitghted | Value Weighted | Equal Weitghted | Value Weighted |
|---|---|---|---|---|
| | t=1 | | t=6 | |
| Alpha | 0.01 | 0.01 | 0.008* | 0.005 |
| | (0.007) | (0.007) | (0.005) | (0.005) |
| | t=3 | | t=12 | |
| Alpha | 0.005 | 0.003 | 0.008* | 0.007 |
| | (0.005) | (0.005) | (0.005) | (0.006) |

# 7 Investment Biases

Several studies have documented behavioral biases practiced by investors that cause them to perform worse in the financial market. Our goal now is to try to relate the inferior performance of investors classified as financial illiterate to possible behavioral biases. The work done by Birru *et al.* (2020) has already shown that investor's decisions considered sub-optimal ex-ante have a direct relationship with several behavioral biases. The authors demonstrate that investors who have made non-efficient decisions in the financial market are affected by several behavioral biases: disposition effect, under diversification, preference for lottery-like stocks, attention grabbing and extrapolation.

It is worth mentioning, however, that some biases, despite being relevant to explain the investor's performance, were not relevant to explain the results found in this experiment, since we only verified the return on stocks bought by investors at different time horizons and also controlled for ticker and time fixed effects. In order to check each of the biases we estimate the following equation:

$$Bias = \beta_0 + \beta_1 FinanciallyIlliterate + \beta_2 age + \beta_3 sqrage + \beta_4 exp + \beta_5 day + \beta_6 month + \varepsilon_i \quad (9)$$

Where $Financiallyilliterate$ equals one according to our classification, $age$ is the investor's age, $sqrage$ is the investor's square age, $exp$ is the investor's experience in the financial market, $day$ is the number of days that the investor has traded in the financial market and $month$ is the number of months that the investor has traded in the financial market.

## 7.1 Overconfidence and Sensation Seeking

One of the first efforts to try to understand possible failures of investor behavior was carried out by Odean (1999). The author demonstrates that investors who performed trades in an excessive manner obtain a worse return due to the transaction costs incurred. According to the author, the large number of trades carried out by some investors is linked to the overconfidence of these investors. Grinblatt & Keloharju (2009) also relates the large number of trades carried

out by some investors with sensation seeking.

When we compare the average number of trades, we can verify that investors considered financially illiterate perform more trades. We conducted a mean difference test and found in fact that investors considered financially illiterate perform more trades at the significance level of 1%. However, excessive trading is not able to directly explain the difference in performance since we do not include transaction costs in our experiment.

## 7.2 Under-diversification

The importance of diversification in investor's portfolios is widely accepted in the finance literature. Markowitz (1952) demonstrates that by combining different negatively correlated assets the investor is able to obtain higher returns with less variance. The ability of the investor to understand the importance of diversification proved to be relevant to the extent that this is one of the three necessary skills defined by the OECD for an investor to be considered capable of understanding finance.

It has already been well documented in the literature that the individual investor tends to have undiversified portfolios. Barber & Odean (2000) points out that lack of diversification may be related to risk-seeking or lack of knowledge. Goetzmann & Kumar (2008) point out that the lack of diversification is more present in investors with the following characteristics: younger, low-income, less educated and less sophisticated. According to these authors the lack of diversification is also related to the following behavioral biases: overconfidence, trend-following behavior and local bias. Thus, we use the Diversification Ratio ($DR$) proposed by Choueifaty & Coignard (2008). This measure indicates the level of diversification of a portfolio. Given a portfolio consisting of $N$ assets and with vector $(\omega_1, \ldots, \omega_N)$ of weights such that $\sum_{i=1}^{N} \omega_i = 1$ , we can define DR as follows:

$$DR = \frac{\sum_{i=1}^{N} \omega_i \sigma_i}{\sigma_P} \tag{10}$$

Where $\sigma_i$ is the volatility of asset $i$ and $\sigma_P$ the portfolio volatility. Choueifaty *et al.* (2013)

proposes a method of decomposition of the DR to better understand how it works.

$$DR(\mathbf{w}) = [\rho(\mathbf{w})(1 - CR(\mathbf{w})) + CR(\mathbf{w})]^{-1/2} \tag{11}$$

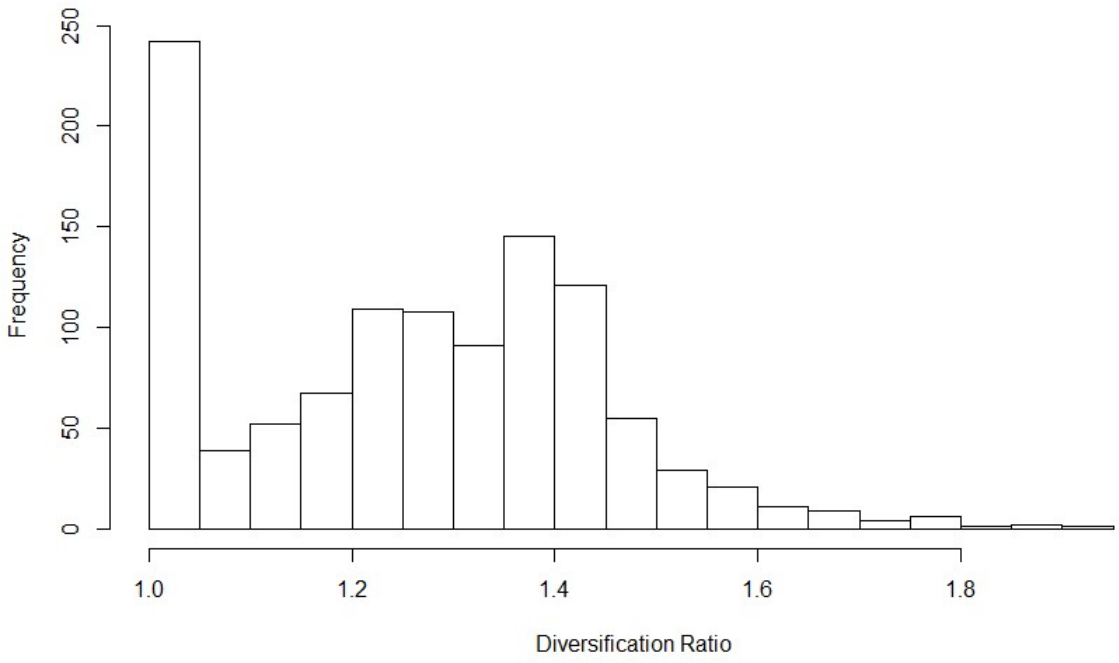
Where $\rho(\mathbf{w})$ is the volatility-weighted average correlation of the assets in the portfolio and $CR(\mathbf{w})$ is the volatility-weighted concentration ratio of the portfolio. Formally both are defined as follows:

$$\rho(\mathbf{w}) = \frac{\sum_{i \neq j}^{N} (\omega_i \sigma_i \omega_j \sigma_j) \rho_{ij}}{\sum_{i \neq j}^{N} (\omega_i \sigma_i \omega_j \sigma_j)} \tag{12}$$

$$CR(\mathbf{w}) = \frac{\sum_{i=1}^{N} (\omega_i \sigma_i)^2}{\left(\sum_{i=1}^{N} \omega_i \sigma_i\right)^2} \tag{13}$$

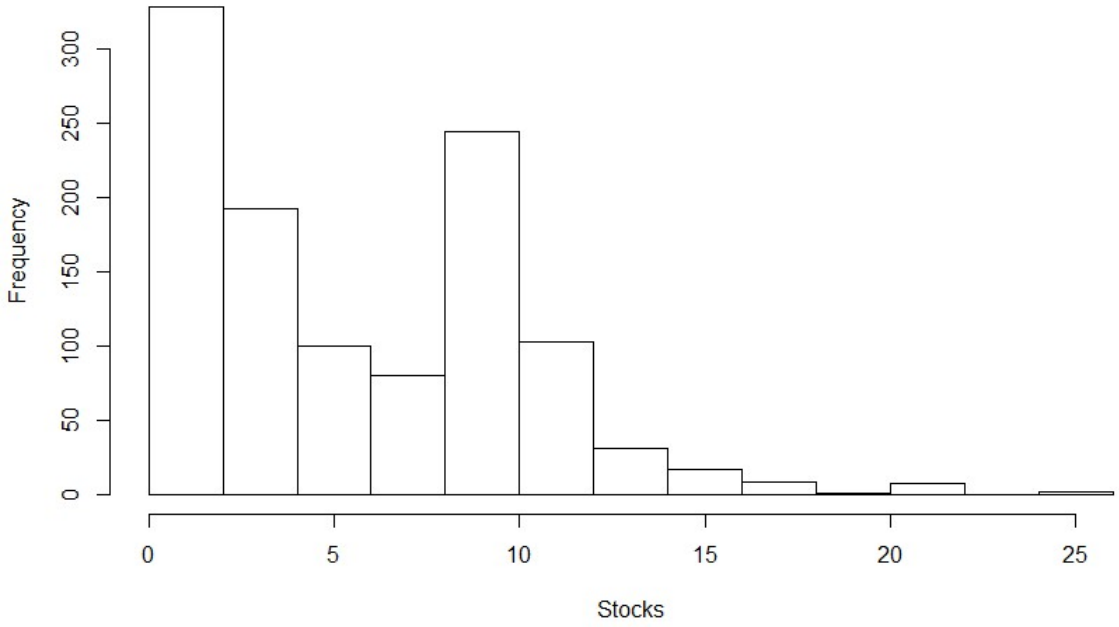
A portfolio concentrated in only one asset will have $DR$ equal to 1. According to the decomposition we can notice that very concentrated portfolios or with assets with high correlation will also present $DR$ equal to 1. From the figure below we can notice that a large part of the investors considered to be financially illiterate hold portfolios that are not very diversified.

Figure 4: **Financially illiterate investors - Diversification Ration**



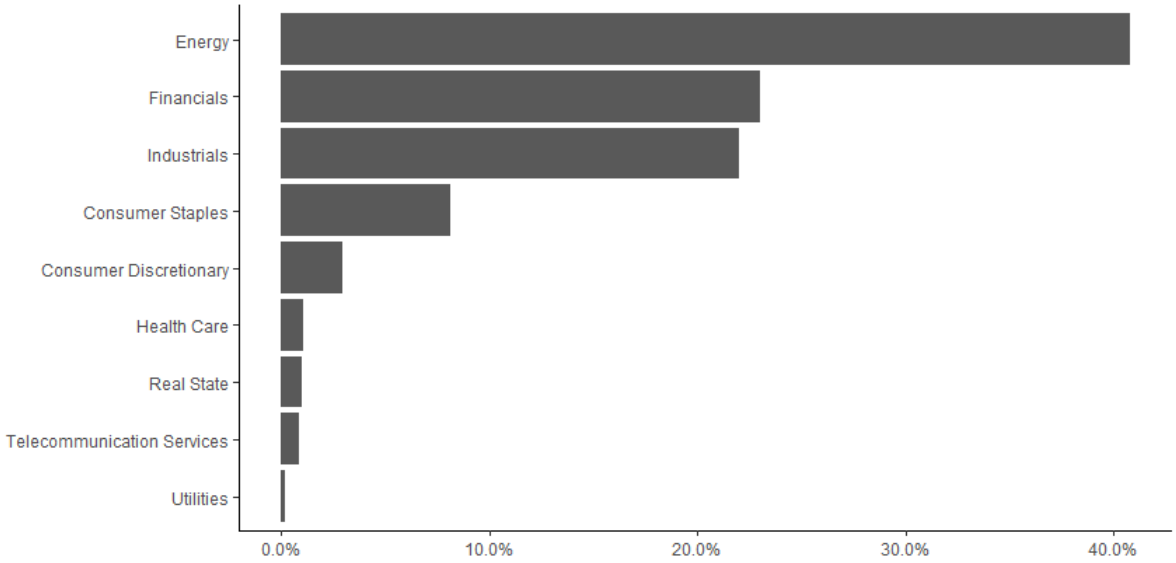
This low diversification can be explained according to the $DR$ decomposition. First the investors hold a small number of stocks. Over half of the investors hold less than 5 stocks in their portfolio.

Figure 5: **Financially illiterate investors - Stocks**




Second, investors own shares concentrated in only a few sectors of the economy. More than 80% of the stocks purchased are concentrated in only 3 sectors of the economy.

Figure 6: **Financially illiterate investors - Purchases**



From the table below we noticed a positive correlation between financially illiterate investors and DR. Thus, although financially illiterate investors diversify poorly, we cannot conclude that the lack of diversification explains our result[9].

Table 6: **Diversification Ratio**

This table presents the result of the diversification of the individuals. The diversification ratio indicates the level of diversification of a portfolio. We check the net position of each investor at the end of our sample. Portfolios that are concentrated or with highly correlated assets present DR close to one. We use controls *age*, *sqrage*, *exp*, *days* amd *months*. Where *age* is the investor's age, *sqrage* is the investor's square age, *exp* is the investor's experience in the financial market, *days* is the number of days that the investor has traded in the financial market and *months* is the number of months that the investor has traded in the financial market.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | DR | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Financially_Illiterate | 0.067*** | 0.060*** | 0.063*** | 0.063*** | 0.042*** |
| | (0.006) | (0.006) | (0.006) | (0.006) | (0.005) |
| age | | 0.0005*** | 0.005*** | 0.005*** | 0.003*** |
| | | (0.00002) | (0.0001) | (0.0001) | (0.0001) |
| sqrage | | | −0.00005*** | −0.00005*** | −0.00003*** |
| | | | (0.00000) | (0.00000) | (0.00000) |
| exp | | | | 0.0002** | −0.003*** |
| | | | | (0.0001) | (0.0001) |
| days | | | | | −0.0003*** |
| | | | | | (0.00001) |
| months | | | | | 0.008*** |
| | | | | | (0.0001) |
| Constant | 1.179*** | 1.158*** | 1.056*** | 1.056*** | 1.058*** |
| | (0.0003) | (0.001) | (0.003) | (0.003) | (0.003) |
| Observations | 318,729 | 318,729 | 318,729 | 318,729 | 318,729 |
| $R^2$ | 0.0004 | 0.002 | 0.006 | 0.006 | 0.082 |
| Adjusted $R^2$ | 0.0004 | 0.002 | 0.006 | 0.006 | 0.082 |
| Residual Std. Error | 0.193 (df = 318727) | 0.193 (df = 318726) | 0.192 (df = 318725) | 0.192 (df = 318724) | 0.185 (df = 318722) |
| F Statistic | 140.189*** (df = 1; 318727) | 272.263*** (df = 2; 318726) | 588.829*** (df = 3; 318725) | 442.935*** (df = 4; 318724) | 4,738.887*** (df = 6; 318722) |
| *Note:* | | | | | *p<0.1; **p<0.05; ***p<0.01 |

[9]The appendix shows the distribution of the DR for all investors as well as the number of stocks in the portfolio and the sectors of purchase.

## 7.3 Lottery-like stocks

The preference for lottery-like stocks is one of the behavioral biases that affect investor performance. Eraker & Ready (2015) demonstrates that investors overpay for stocks with lottery-like payoffs. Kumar (2009) shows that investor's propensity to gamble is positively correlated with their preference for lottery-like stocks.

To define a stock as lottery-like we use the methodology proposed by Kumar (2009). The author considers three characteristics so that a stock can be perceived by the investor as a lottery ticket: i) idiosyncratic volatility, ii) idiosyncratic skewness and iii) price. To find the idiosyncratic volatility we estimate for each month $t$ the following regression[10]:

$$R_{i,t} - Rf_t = a_i + b_i\left(Rm_t - Rf_t\right) + s_i\left(SMB_t\right) + h_i\left(HML_t\right) + w_i\left(WML_t\right) + e_{i,t} \tag{14}$$

The measure of idiosyncratic volatility is provided by the variance of the residuals. To obtain the idiosyncratic skewness, we perform the estimation of the following regression:

$$R_{i,t} - Rf_t = a_i + b_i\left(Rm_t - Rf_t\right) + c_i\left(Rm_t - Rf_t\right)^2 + e_{i,t} \tag{15}$$

The idiosyncratic skewness is provided by the third moment of the residuals. The stock price is defined as the last stock price in the past month. Thus, for each month we defined the stocks considered lottery-like those that were simultaneously in the first tercile of prices and in the last tercile of the idiosyncratic volatility and the idiosyncratic skewness. As Birru *et al.* (2020), we calculate the percentage of lottery-like stocks in relation to the total number of stocks purchased.

[10]For each month, we use data from six past months to perform the estimation.

Table 7: **Preference for lottery-like stocks**

This table presents data on individual's preferences for lottery-like stocks. The dependent variable is the fraction of purchases of lottery-like stocks. In order to classify a lottery-like stock we considered for each month the stocks that were in the first tercile of the price and also in the third tercile of the indiosyncratic volatility and skewness. We use controls *age*, *sqrage*, *exp*, *days* and *months*. Where *age* is the investor's age, *sqrage* is the investor's square age, *exp* is the investor's experience in the financial market, *days* is the number of days that the investor has traded in the financial market and *months* is the number of months that the investor has traded in the financial market.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Lo_Ratio | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Financially_Illiterate | 0.782 | 1.768*** | 1.808*** | 1.797*** | 1.508** |
| | (0.621) | (0.621) | (0.621) | (0.621) | (0.617) |
| age | | −0.066*** | −0.010 | −0.027** | −0.109*** |
| | | (0.002) | (0.013) | (0.013) | (0.013) |
| sqrage | | | −0.001*** | −0.001*** | 0.00004 |
| | | | (0.0001) | (0.0001) | (0.0001) |
| exp | | | | 0.105*** | −0.021*** |
| | | | | (0.008) | (0.008) |
| days | | | | | 0.036*** |
| | | | | | (0.001) |
| months | | | | | 0.185*** |
| | | | | | (0.007) |
| Constant | 12.214*** | 14.980*** | 13.776*** | 14.008*** | 14.573*** |
| | (0.032) | (0.106) | (0.298) | (0.298) | (0.297) |
| Observations | 480,723 | 480,723 | 480,723 | 480,723 | 480,723 |
| $R^2$ | 0.00000 | 0.002 | 0.002 | 0.002 | 0.016 |
| Adjusted $R^2$ | 0.00000 | 0.002 | 0.002 | 0.002 | 0.016 |
| Residual Std. Error | 22.409 (df = 480721) | 22.391 (df = 480720) | 22.391 (df = 480719) | 22.387 (df = 480718) | 22.229 (df = 480716) |
| F Statistic | 1.587 (df = 1; 480721) | 375.141*** (df = 2; 480720) | 256.329*** (df = 3; 480719) | 235.946*** (df = 4; 480718) | 1,303.721*** (df = 6; 480716) |

*Note:* *p<0.1; **p<0.05; ***p<0.01

We notice that financially illiterate investors tend to buy more lottery-like stocks. The purchases of financially illiterate investors tend to have on average a higher composition of lottery-like stocks. According to column (5), we find that financially illiterate investors purchase

10.3% (16.081%/14.573%) more lottery-like stocks. As Eraker & Ready (2015) demonstrates that investors overpay for stocks with lottery-like payoff, this result helps us understand the difference in performance of these investors.

## 7.4 Extrapolation

Tversky & Kahneman (1974) documented that individuals tend to have misperceptions about the concept of chance. When individuals observe a sequence of independent random events, the outcome of past realizations of that event influences the perception of future realizations. Several studies show that this lack of rationality is also present in the financial market. Patel *et al.* (1991) points out that investors tend to project the future performance of an asset based on its past performance. Benartzi (2001) demonstrated that individual's stock selection is positively correlated with the past performance of these assets. However, the future performance of these assets was not consistent with their past performance.

In order to check whether financially illiterate investors suffer from this behavioral bias, we applied the methodology proposed by Birru *et al.* (2020). Thus, we construct for each investor a measure of extrapolation. We check the percentage of stocks bought by each investor that obtained a high past return. In order for the stock to be considered in this criterion on the day of the purchase, its performance in the last 20 days should be higher than 16.92%[11].

[11] 90th percentile

Table 8: **Extrapolation**

This table presents data on individual's preferences for extrapolative stocks. The dependent variable is the fraction of "purchases by extrapolation" . A purchase by extrapolation is the purchase of a stock whose past 20-day returns is above 16.92%, the 90th percentile in our sample (2016-2018). We use controls *age*, *sqrage*, *exp*, *days* and *months*. Where *age* is the investor's age, *sqrage* is the investor's square age, *exp* is the investor's experience in the financial market, *days* is the number of days that the investor has traded in the financial market and *months* is the number of months that the investor has traded in the financial market.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Ex_Ratio | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Financially_Illiterate | 1.146* | 1.320** | 1.397** | 1.388** | 1.436** |
| | (0.628) | (0.629) | (0.629) | (0.629) | (0.628) |
| age | | −0.012*** | 0.095*** | 0.081*** | 0.052*** |
| | | (0.002) | (0.013) | (0.013) | (0.013) |
| sqrage | | | −0.001*** | −0.001*** | −0.001*** |
| | | | (0.0001) | (0.0001) | (0.0001) |
| exp | | | | 0.079*** | 0.039*** |
| | | | | (0.008) | (0.008) |
| days | | | | | 0.023*** |
| | | | | | (0.001) |
| months | | | | | 0.025*** |
| | | | | | (0.007) |
| Constant | 14.037*** | 14.526*** | 12.236*** | 12.412*** | 12.714*** |
| | (0.033) | (0.107) | (0.302) | (0.302) | (0.302) |
| Observations | 480,723 | 480,723 | 480,723 | 480,723 | 480,723 |
| $R^2$ | 0.00001 | 0.0001 | 0.0002 | 0.0004 | 0.003 |
| Adjusted $R^2$ | 0.00000 | 0.0001 | 0.0002 | 0.0004 | 0.003 |
| Residual Std. Error | 22.663 (df = 480721) | 22.663 (df = 480720) | 22.661 (df = 480719) | 22.659 (df = 480718) | 22.633 (df = 480716) |
| F Statistic | 3.333* (df = 1; 480721) | 13.078*** (df = 2; 480720) | 30.724*** (df = 3; 480719) | 47.459*** (df = 4; 480718) | 215.089*** (df = 6; 480716) |

*Note:* *p<0.1; **p<0.05; ***p<0.01

We notice that financially illiterate investors buy more stocks that have performed well in the past. According to column (5), financially illiterate investors buy 11% (14.15%)/12.714$) more stocks that have performed well in the past than the rest of the investors.

## 7.5 Salient Stocks

Ninio & Kahneman (1974) points out that attention demands effort from the individual and for this reason be considered a scarce resource. Barber & Odean (2000) documents that the investor tends to limit their stock buying set to the ones that catch their attention. According to the authors, individual investors tend to buy stocks that appear in the media. Since investors avoid trading short, the salience is restricted to purchases. Fedyk (2018) points out that stocks that receive more exposure in the media have higher trading volume and larger price changes.

This behavior can be deleterious to the investor's performance. Once an investor notices a news story about a particular stock, it is likely that this information is already incorporated into the price. Overconfident investors believe that the price does not fully reflect the new information.

A common practice of the media is to present the stocks with the worst and best daily performance (see Figure 9). Kumar *et al.* (2021) documents that stocks that appear in the rankings receive more attention from investors. Because of the buying pressure these assets suffer from overpricing. Such stocks subsequently underperform. To check whether illiterate investors suffer from the salience bias, we compute for each investor the percentage of salience stock purchases. We considered a stock to be salient if on the day of purchase at some point the stock was present as either the five best performers or the five worst performers.

The result of the table below indicates that financially illiterate investors tend to buy more stocks that are highlighted. About 21,314% of these investor's purchases are made up of stocks that are in the spotlight. According to column (5), financially illiterate investors buy 8.8% (21.314%/19.590%) more lottery-like stocks compared to the rest of the investors.

Table 9: **Preference for salient stocks**

This table presents data on individual's preferences for salient stocks. The dependent variable is the fraction of purchases of salient stocks. A stock is considered to be salient if on the day of its purchase it was at some point in the top five or five worst performers. We use controls *age*, *sqrage*, *exp*, *days* and *months*. Where *age* is the investor's age, *sqrage* is the investor's square age, *exp* is the investor's experience in the financial market, *days* is the number of days that the investor has traded in the financial market and *months* is the number of months that the investor has traded in the financial market.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | SA_Ratio | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Financially_Illiterate | 1.579** | 1.494** | 1.602** | 1.590** | 1.724** |
| | (0.734) | (0.736) | (0.736) | (0.735) | (0.736) |
| age | | 0.006** | 0.156*** | 0.138*** | 0.129*** |
| | | (0.003) | (0.016) | (0.016) | (0.016) |
| sqrage | | | −0.002*** | −0.002*** | −0.001*** |
| | | | (0.0002) | (0.0002) | (0.0002) |
| exp | | | | 0.107*** | 0.100*** |
| | | | | (0.009) | (0.010) |
| days | | | | | 0.014*** |
| | | | | | (0.001) |
| months | | | | | −0.026*** |
| | | | | | (0.008) |
| Constant | 22.665*** | 22.426*** | 19.190*** | 19.427*** | 19.590*** |
| | (0.038) | (0.126) | (0.353) | (0.353) | (0.354) |
| Observations | 480,723 | 480,723 | 480,723 | 480,723 | 480,723 |
| $R^2$ | 0.00001 | 0.00002 | 0.0002 | 0.0005 | 0.001 |
| Adjusted $R^2$ | 0.00001 | 0.00001 | 0.0002 | 0.0005 | 0.001 |
| Residual Std. Error | 26.511 (df = 480721) | 26.511 (df = 480720) | 26.508 (df = 480719) | 26.505 (df = 480718) | 26.501 (df = 480716) |
| F Statistic | 4.625** (df = 1; 480721) | 4.309** (df = 2; 480720) | 34.989*** (df = 3; 480719) | 58.892*** (df = 4; 480718) | 62.870*** (df = 6; 480716) |

*Note:* *p<0.1; **p<0.05; ***p<0.01

# 8 Bias and Performance

We also check whether the performance of financially illiterate investors is affected by biases. Thus, we estimate equations 5, 6, 7 and 8 controlling for the measures of biases calculated for each investor. As we can notice, when we control for the bias measures, the performance of the financially illiterate investor improves. Column (8) indicates an 18% improvement in the performance of this investor. The bias coefficients are as expected. This result is an evidence that part of the underperformance of financially illiterate investors might be explained by behavioral biases.

Table 10: **Bias and Performance**

This table presents the results of the difference in performance, at different time horizons, in relation to the purchases by investors considered financially illiterate and the rest of the investors. The estimation considers time and ticker fixed effects. The period covered is 2016 to 2018. We use controls $Extrapolation$, $Salience$, $Lottery$, $Diversification$.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Rent60 | Rent120 | Rent180 | Rent240 | Rent60 | Rent120 | Rent180 | Rent240 |
| *Financially Illiterate* | -0.0105 | -0.0315** | -0.0547** | -0.0578** | -0.00759 | -0.0259** | -0.0456** | -0.0478* |
| | (0.00754) | (0.0137) | (0.0211) | (0.0284) | (0.00718) | (0.0127) | (0.0197) | (0.0265) |
| *Extrapolation* | | | | | -0.000143 | -0.000248 | -0.000251 | -0.000614** |
| | | | | | (9.49e-05) | (0.000161) | (0.000216) | (0.000275) |
| *Salience* | | | | | -3.53e-05 | -7.18e-05 | -6.78e-05 | -0.000107 |
| | | | | | (5.09e-05) | (0.000102) | (0.000145) | (0.000180) |
| *Lottery* | | | | | 0.000335*** | 0.000621** | 0.000950** | 0.00123** |
| | | | | | (0.000108) | (0.000273) | (0.000417) | (0.000546) |
| *Diversification* | | | | | 0.000233*** | 0.000439*** | 0.000675*** | 0.000798*** |
| | | | | | (4.22e-05) | (7.05e-05) | (9.21e-05) | (0.000134) |
| *Constant* | 0.101*** | 0.191*** | 0.278*** | 0.355*** | 0.0836*** | 0.157*** | 0.223*** | 0.294*** |
| | (1.65e-05) | (2.95e-05) | (5.83e-05) | (0.000104) | (0.00350) | (0.00789) | (0.0108) | (0.0158) |
| *Date fixed effects* | YES | YES | YES | YES | YES | YES | YES | YES |
| *Ticker fixed effects* | YES | YES | YES | YES | YES | YES | YES | YES |
| *Observations* | 9,045,004 | 9,042,004 | 9,040,033 | 9,028,619 | 9,045,004 | 9,042,004 | 9,040,033 | 9,028,619 |
| *R-squared* | 0.420 | 0.451 | 0.464 | 0.430 | 0.421 | 0.452 | 0.465 | 0.432 |

*Robust standard errors in parentheses, clustered at the time and ticker level*

**** p<0.01, ** p<0.05, * p<0.1*

# 9 Conclusion

The tests on the efficient markets hypothesis conducted to date have suffered from numerous criticisms. Our paper sought to incorporate the discussions on this subject and propose a new type of test. We do not use a large amount of data to propose a new investment strategy. We propose a simple test: we put individual investors who know nothing about finance to trade in the stock market and compare them to the rest of the individual investors.

We note that stock purchases by financially illiterate investors perform worse than the rest of the investors. We relate this worse performance to behavioral biases. Financially illiterate investors tend to buy lottery-like stocks and salient stocks, overtrade and extrapolate future returns. Financially illiterate investors diversify poorly, however, this result failed to explain our results. Part of the underperformance of financially illiterate investors is explained by behavioral biases.

# References

Ball, Ray. 1978. Anomalies in relationships between securities' yields and yield-surrogates. *Journal of financial economics*, **6**(2-3), 103–126.

Barber, Brad M, & Odean, Terrance. 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance*, **55**(2), 773–806.

Basu, Sanjoy. 1983. The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of financial economics*, **12**(1), 129–156.

Benartzi, Shlomo. 2001. Excessive extrapolation and the allocation of 401 (k) accounts to company stock. *The Journal of Finance*, **56**(5), 1747–1764.

Birru, Justin, Chague, Fernando, De-Losso, Rodrigo, & Giovannetti, Bruno. 2020. Sub-optimal trading decisions and investment biases. *Fisher College of Business Working Paper*, 022.

Bitu, Otavio, Chague, Fernando, Giovanetti, Bruno, & Hamdan, Tomaz. 2021. O retorno esperado dos COEs. *Working Paper*.

Campbell, John Y, & Shiller, Robert J. 1988a. The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, **1**(3), 195–228.

Campbell, John Y, & Shiller, Robert J. 1988b. Stock prices, earnings, and expected dividends. *the Journal of Finance*, **43**(3), 661–676.

Célérier, Claire, & Vallée, Boris. 2017. Catering to investors through security design: Headline rate and complexity. *The Quarterly Journal of Economics*, **132**(3), 1469–1508.

Choueifaty, Yves, & Coignard, Yves. 2008. Toward maximum diversification. *The Journal of Portfolio Management*, **35**(1), 40–51.

Choueifaty, Yves, Froidure, Tristan, & Reynier, Julien. 2013. Properties of the most diversified portfolio. *Journal of investment strategies*, **2**(2), 49–70.

De Genaro, Alan. 2021. Afinal, produtos estruturados são bons para quem? *Working Paper*.

Eraker, Bjørn, & Ready, Mark. 2015. Do investors overpay for stocks with lottery-like payoffs? An examination of the returns of OTC stocks. *Journal of Financial Economics*, **115**(3), 486–504.

Fama, Eugene F. 1970. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, **25**(2), 383–417.

Fama, Eugene F. 2021. Efficient capital markets a review of theory and empirical work. *The Fama Portfolio*, 76–121.

Fama, Eugene F, & French, Kenneth R. 1993. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, **33**(1), 3–56.

Fedyk, Anastassia. 2018. *Front page news: The effect of news positioning on financial markets.* Tech. rept. working paper.

Fluck, Zsuzsanna, Malkiel, Burton G, & Quandt, Richard E. 1997. The predictability of stock returns: A cross-sectional simulation. *Review of Economics and Statistics*, **79**(2), 176–183.

French, Kenneth R. 1980. Stock returns and the weekend effect. *Journal of financial economics*, **8**(1), 55–69.

Goetzmann, William N, & Kumar, Alok. 2008. Equity portfolio diversification. *Review of Finance*, **12**(3), 433–463.

Grinblatt, Mark, & Keloharju, Matti. 2009. Sensation seeking, overconfidence, and trading activity. *The Journal of Finance*, **64**(2), 549–578.

Grinblatt, Mark, Keloharju, Matti, & Linnainmaa, Juhani. 2009. Do smart investors outperform dumb investors. *Chicago Booth Research Paper*, **9**(33), 46.

Harvey, Campbell R. 2017. Presidential address: The scientific outlook in financial economics. *The Journal of Finance*, **72**(4), 1399–1440.

Henderson, Brian J, Pearson, Neil D, & Wang, Li. 2020. Pre-trade hedging: Evidence from the issuance of retail structured products. *Journal of Financial Economics*, **137**(1), 108–128.

Hilgert, Marianne A, Hogarth, Jeanne M, & Beverly, Sondra G. 2003. Household financial management: The connection between knowledge and behavior. *Fed. Res. Bull.*, **89**, 309.

Keim, Donald B. 1983. Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of financial economics*, **12**(1), 13–32.

Kumar, Alok. 2009. Who gambles in the stock market? *The Journal of Finance*, **64**(4), 1889–1933.

Kumar, Alok, Ruenzi, Stefan, & Ungeheuer, Michael. 2021. Daily winners and losers. *Available at SSRN 2931545*.

Lamont, Owen A, & Thaler, Richard H. 2003. Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy*, **111**(2), 227–268.

Lee, Cheng-few, Porter, David C, & Weaver, Daniel G. 1998. Indirect tests of the Haugen-Lakonishok small-firm/January effect hypotheses: Window dressing versus performance hedging. *Financial Review*, **33**(2), 177–194.

Lusardi, Annamaria, & Mitchell, Olivia S. 2007. Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of monetary Economics*, **54**(1), 205–224.

Lusardi, Annamaria, & Mitchell, Olivia S. 2011. Financial literacy around the world: an overview. *National Bureau of Economic Research Working Paper Series*.

Malkiel, Burton G. 2003. The efficient market hypothesis and its critics. *Journal of economic perspectives*, **17**(1), 59–82.

Markowitz, Harry. 1952. Portfolio Selection. *The Journal of Finance*, **7**(1), 77–91.

Nicholson, S Francis. 1960. Price-earnings ratios. *Financial Analysts Journal*, **16**(4), 43–45.

Ninio, Anat, & Kahneman, Daniel. 1974. Reaction time in focused and in divided attention. *Journal of Experimental Psychology*, **103**(3), 394.

Odean, Terrance. 1999. Do investors trade too much? *American economic review*, **89**(5), 1279–1298.

Patel, Jayendu, Zeckhauser, Richard, & Hendricks, Darryll. 1991. The rationality struggle: Illustrations from financial markets. *The American Economic Review*, **81**(2), 232–236.

Rathgeber, Andreas W, & Wang, Yun. 2011. Market pricing of credit-linked notes: the case of retail structured products in Germany.

Shen, Dehua, Urquhart, Andrew, & Wang, Pengfei. 2019. Does twitter predict Bitcoin? *Economics Letters*, **174**, 118–122.

Shiller, Robert C. 2000. Irrational exuberance. *Philosophy and Public Policy Quarterly*, **20**(1), 18–23.

Tversky, Amos, & Kahneman, Daniel. 1974. Judgment under uncertainty: Heuristics and biases. *science*, **185**(4157), 1124–1131.

Urquhart, Andrew. 2016. The inefficiency of Bitcoin. *Economics Letters*, **148**, 80–82.

Wilkens, Sascha, Erner, Carsten, & Röder, Klaus. 2003. The pricing of structured products in Germany. *The Journal of Derivatives*, **11**(1), 55–69.

# 10 Appendix

Table 11: **COE payoff**

This table presents the payoffs of the COE types with the highest volume of issuance. The descriptions were obtained from B3. There are 61 types of COE marketed in Brazil.

| Types | Payoff |
|---|---|
| Call | The investor gains proportionally as the reference asset rises. |
| Put | The investor gains proportionally with the fall of the reference asset |
| Call KO | The investor earns proportionally as the reference asset rises to a certain limit, which, if reached, receives a fixed rate. |
| Put KO | The investor earns proportionally as the reference asset falls to a certain limit, which, if reached, receives a fixed rate. |
| Call Spread | The investor earns proportionally as the reference asset rises up to a certain limit, obtaining a maximum return. |
| Put Spread | The investor gains proportionally as the reference asset falls to a certain limit. |
| Straddle | The investor gains proportionally with the rise or fall of the reference asset. |
| Straddle KO Call | Gains proportionally with the rise of the asset's price, up to a certain limit. The investor also gains with the fall of the price, without limit. |
| Digital Call | The investor earns a fixed tax if the price of the reference asset is higher than a fixed price. |
| Digital Put | The investor earns a fixed tax if the price of the reference asset is lower than a fixed price. |

Figure 7: **All investors - Diversification Ratio**

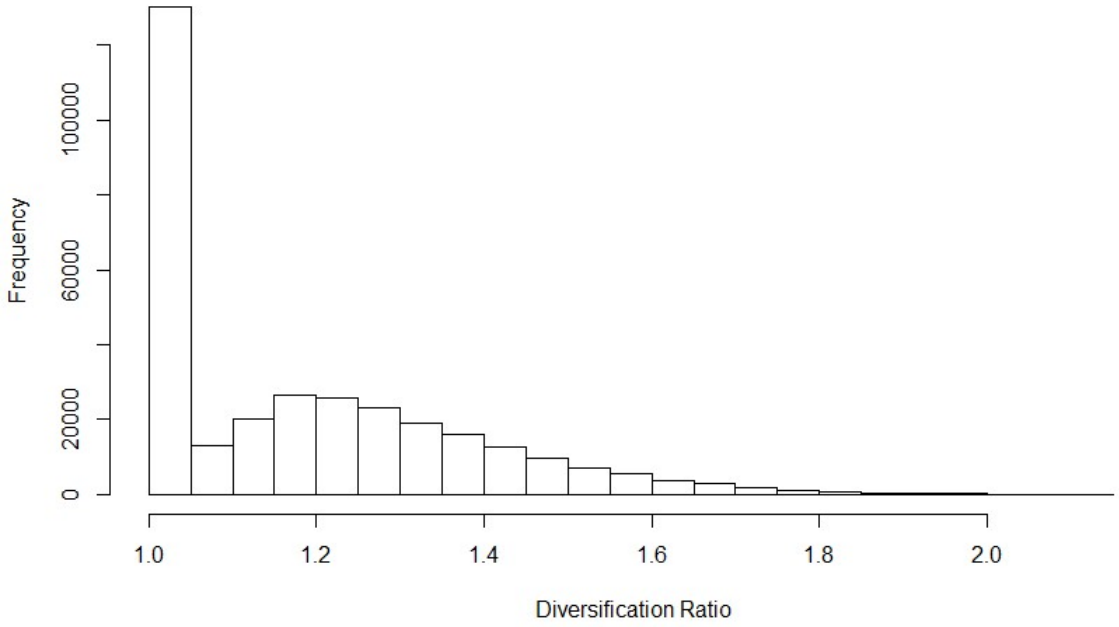



Figure 8: **All investors - Purchases**

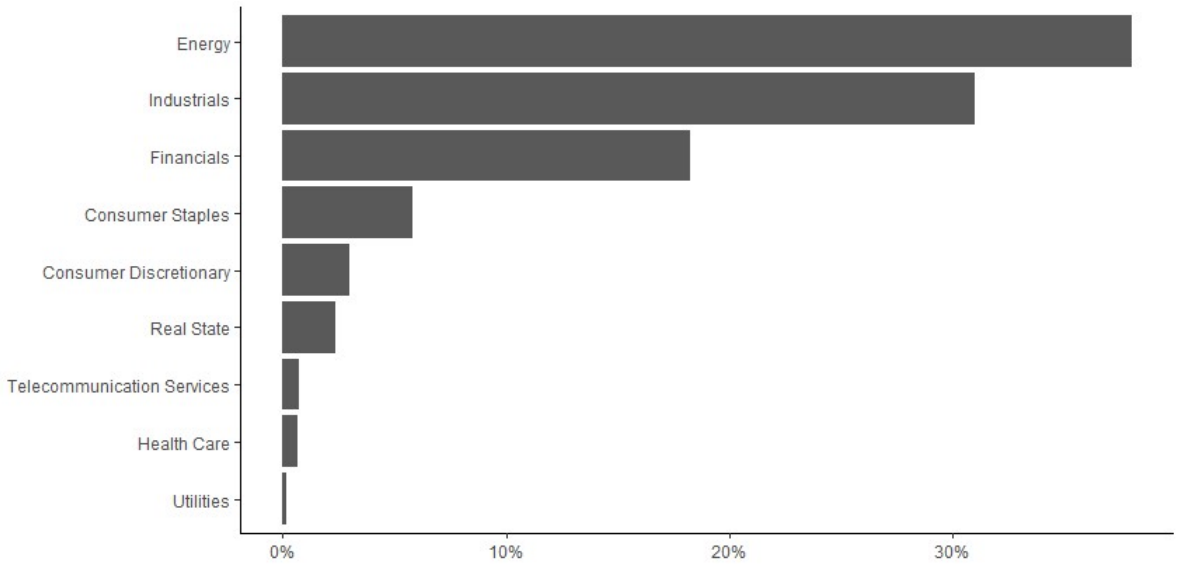
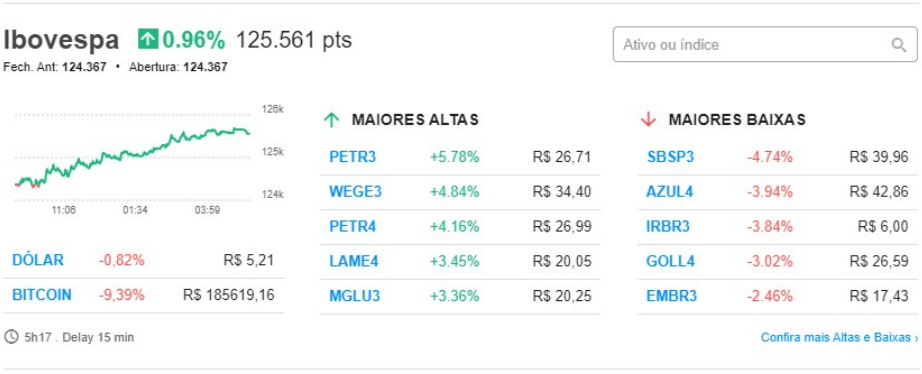


Figure 9: **Daily ranking**




We use an alternative measure for diversification,Birru *et al.* (2020). We classify each traded stock according to the sector of activity defined by B3. We calculate each month the concentration of purchases of each investor by means of the Herfindahl-Hirschman Index ($HHI$. We then calculate the average of this index for each investor.

$$HHI = \sum_{i=1}^{N} s_i^2 \tag{16}$$

Where $s_i^2$ is the market share of each sector in the purchases of the investor and $N$ represents the number of sectors.

Table 12: **Herfindahl-Hirschman Index**

This table presents the result of the diversification of the individuals. The dependent variable is the Herfindahl- Hirschman index based on the volume invested per stock in each month. We use the average of this monthly index. We use controls *age*, *sqrage*, *experience days*, *months*. Where *age* is the investor's age, *sqrage* is the investor's square age,*exp* is the investor's experience in the financial market, *day* is the number of days that the investor has traded in the financial market and *month* is the number of months that the investor has traded in the financial market.

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | HHI | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Financially_Illiterate | −0.129*** | −0.124*** | −0.128*** | −0.128*** | −0.141*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| age | | −0.0003*** | −0.005*** | −0.006*** | −0.004*** |
| | | (0.00002) | (0.0001) | (0.0001) | (0.0001) |
| sqrage | | | 0.00005*** | 0.00005*** | 0.00004*** |
| | | | (0.00000) | (0.00000) | (0.00000) |
| exp | | | | 0.005*** | 0.006*** |
| | | | | (0.0001) | (0.0001) |
| days | | | | | −0.002*** |
| | | | | | (0.00001) |
| months | | | | | 0.002*** |
| | | | | | (0.0001) |
| Constant | 0.823*** | 0.836*** | 0.932*** | 0.943*** | 0.921*** |
| | (0.0003) | (0.001) | (0.002) | (0.002) | (0.002) |
| Observations | 480,723 | 480,723 | 480,723 | 480,723 | 480,723 |
| $R^2$ | 0.001 | 0.002 | 0.005 | 0.016 | 0.138 |
| Adjusted $R^2$ | 0.001 | 0.002 | 0.005 | 0.016 | 0.138 |
| Residual Std. Error | 0.188 (df = 480721) | 0.188 (df = 480720) | 0.188 (df = 480719) | 0.187 (df = 480718) | 0.175 (df = 480716) |
| F Statistic | 615.610*** (df = 1; 480721) | 434.478*** (df = 2; 480720) | 848.444*** (df = 3; 480719) | 1,974.400*** (df = 4; 480718) | 12,846.450*** (df = 6; 480716) |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Table 13: **Performance Simulation t=1**

This table presents the result of the strategy that consists in operating contrary to the behavior of financially illiterate investors. We present the excess return of a long-short portfolio. The portfolios are renewed on a monthly basis. The portfolio was built with a time horizon of one month.

| | *Dependent variable:* | |
|---|---|---|
| | Excess return (Equal Weighted) | Excess return (Value Weighted) |
| | (1) | (2) |
| HML | −0.104 | −0.095 |
| | (0.168) | (0.160) |
| SMB | 0.400* | 0.334* |
| | (0.199) | (0.188) |
| Rm_minus_Rf | −0.266* | −0.226 |
| | (0.146) | (0.139) |
| Constant | 0.010 | 0.010 |
| | (0.007) | (0.007) |
| Observations | 36 | 36 |
| $R^2$ | 0.180 | 0.149 |
| Adjusted $R^2$ | 0.100 | 0.067 |
| Residual Std. Error (df = 32) | 0.038 | 0.036 |
| F Statistic (df = 3; 32) | 2.265 | 1.815 |
| *Note:* | | *p<0.1; **p<0.05; ***p<0.01 |

Table 14: **Performance Simulation t=3**

This table presents the result of the strategy that consists in operating contrary to the behavior of financially illiterate investors. We present the excess return of a long-short portfolio. The portfolios are renewed on a monthly basis. The portfolio was built with a time horizon of three months.

| | *Dependent variable:* | |
|---|---|---|
| | Excess return (Equal Weighted) | Excess return (Value Weighted) |
| HML | 0.127 | 0.116 |
| | (0.129) | (0.117) |
| SMB | −0.202 | −0.134 |
| | (0.152) | (0.138) |
| Rm_minus_Rf | −0.161 | −0.179 |
| | (0.120) | (0.109) |
| Constant | 0.005 | 0.003 |
| | (0.005) | (0.005) |
| Observations | 33 | 33 |
| $R^2$ | 0.155 | 0.153 |
| Adjusted $R^2$ | 0.068 | 0.065 |
| Residual Std. Error (df = 29) | 0.028 | 0.025 |
| F Statistic (df = 3; 29) | 1.779 | 1.748 |
| *Note:* | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

Table 15: **Performance Simulation t=6**

This table presents the result of the strategy that consists in operating contrary to the behavior of financially illiterate investors. We present the excess return of a long-short portfolio. The portfolios are renewed on a monthly basis. The portfolio was built with a time horizon of six months.

| | *Dependent variable:* | |
|---|---|---|
| | Excess return (Equal Weighted) | Excess return (Value Weighted) |
| HML | 0.028 | 0.074 |
| | (0.119) | (0.138) |
| SMB | −0.131 | −0.125 |
| | (0.143) | (0.166) |
| Rm_minus_Rf | −0.232** | −0.312** |
| | (0.109) | (0.127) |
| Constant | 0.008* | 0.005 |
| | (0.005) | (0.005) |
| Observations | 30 | 30 |
| $R^2$ | 0.308 | 0.308 |
| Adjusted $R^2$ | 0.228 | 0.229 |
| Residual Std. Error (df = 26) | 0.024 | 0.028 |
| F Statistic (df = 3; 26) | 3.849** | 3.865** |
| *Note:* | | *p<0.1; **p<0.05; ***p<0.01 |

Table 16: **Performance Simulation t=12**

This table presents the result of the strategy that consists in operating contrary to the behavior of financially illiterate investors. We present the excess return of a long-short portfolio. The portfolios are renewed on a monthly basis. The portfolio was built with a time horizon of twelve months.

| | *Dependent variable:* | |
|---|---|---|
| | Excess return (Equal Weighted) | Excess return (Value Weighted) |
| HML | −0.148 | −0.162 |
| | (0.117) | (0.140) |
| SMB | −0.263* | −0.198 |
| | (0.142) | (0.170) |
| Rm_minus_Rf | −0.028 | −0.024 |
| | (0.106) | (0.127) |
| Constant | 0.008* | 0.007 |
| | (0.005) | (0.006) |
| Observations | 24 | 24 |
| $R^2$ | 0.405 | 0.273 |
| Adjusted $R^2$ | 0.316 | 0.164 |
| Residual Std. Error (df = 20) | 0.022 | 0.026 |
| F Statistic (df = 3; 20) | 4.535** | 2.505* |
| *Note:* | | *p<0.1; **p<0.05; ***p<0.01 |