ISSN 1808-057X

Empirical Evidence: Arbitrage with Exchange-traded Funds (ETFs) on the Brazilian Market *

Yuri Sampaio Maluf
Specialist, Department of Business Administration, University of Brasilia
E-mail: yurimaluf@gmail.com

Pedro Henrique Melo Albuquerque
Ph.D., Department of Business Administration, University of Brasilia
E-mail: pedroa@unb.br

Received on 2.1.2012 - Accepted on 2.2.2012 - 4th version accepted on 12.20.2012

ABSTRACT

According to risk management literature, diversification helps mitigate risk. Index funds, known as exchange-traded funds (ETFs), which were recently introduced into the Brazilian market, make diversification straightforward to accomplish. This paper investigates the efficiency of the valuation process of the Ibovespa iShares with respect to the fair value of the shares. For this purpose, a high-frequency time series analysis of ETFs and Ibovespa was used, followed by strategy simulations that included goodwill and negative goodwill between asset sets with and without transaction costs. To avoid data-snooping effects on the transaction outcomes, a time series bootstrap was applied. The results initially indicated share-pricing inefficiency because the inclusion of goodwill and negative goodwill in the strategy resulted in returns of 172.5% above the fund's index. Additionally, it became apparent that even with the introduction of operating costs, the gains continued to exhibit inefficiency. However, after applying the bootstrap technique, the results did not suggest excess returns, which could be attributed to data snooping. Therefore, the results demonstrate the impossibility of agents earning abnormal returns from the differences between the values of the ETF share and its corresponding index, thereby indicating that the Ibovespa iShare fund pricing is efficient.

Keywords: ETF. Arbitrage. Market-efficiency hypothesis. Data snooping.

* Paper presented at the XIth Brazilian Finance Meeting, Rio de Janeiro, RJ, 2011.

1 INTRODUCTION

The emergence of the exchange-traded funds (ETFs) investment modality is a result of technological advances in capital-market products (Gastineau, 2001). In 1990, this development resulted in the creation of a new category of index funds, known as ETFs, on the Canadian stock market (the Toronto Stock Exchange). ETFs were introduced in the United States in 1993 with the aim of replicating the Standard and Poor's (S&P) 500 index (Guedj & Huang, 2009). In Brazil, the first ETF was created in 2004 with the goal of reproducing the profitability of the IBrX-50 (Brazil Index) (Farias, 2009).

The purpose of these funds is to seek profits corresponding to an index. However, unlike conventional index funds, ETF shares are traded on stock exchanges in the same way as stocks (Poterba & Shoven, 2002).

ETF funds replicate the same index portfolio to which they are linked. The fair or fundamental value is calculated using the equity value of the fund divided by the total number of shares and is published daily in the fund's financial statements. Consequently, fluctuations in secondary market prices should follow the same index returns. However, because the shares are traded on the stock exchange, there can be a disconnect between the market price and the fundamental value of the fund (Rompotis, 2007).

In general, there is a small difference between these values, termed the discount, that remains stable over time. According to Cherry (2004), operating strategies that consider discounts can generate predictive information on the future returns of ETFs. However, this possibility contradicts the assumption of an efficient market because ultimately transaction costs outweigh the expectation of abnormal earnings, which makes such strategies unfeasible.

Based on this hypothesis, asset prices are determined using information publicly available to all market investors (Fama, 1970). Consequently, asset values converge to their fundamental value, which eliminates the possibility of arbitrage on finding systematically over- or undervalued assets (Ross, Westerfield, & Jordan, 2008). According to Damodaran (2009), this market efficiency does not necessarily imply that asset prices remain static with respect to their values but that these fluctuations are configured as random walks.

In this regard, there have been several studies on anomalies. Bondt and Richard (1985) note the potential for systematic additional earnings through "loser" portfolios. That is, portfolios that have had the worst performances in previous years experience more than the market average gain in subsequent years. Studies by Basu (1983) and Jaffe, Keim, and Westerfield (1989) also note the possibility of obtaining above-average returns through directly observable data, such as the ratio between asset price and company earnings, which is known as the P/E ratio.

The origin of strategy formation is based on "bad" asset pricing, i.e., the separation of an asset's price from its real value. According to Dimson and Mussavian (2000), changes in prices occur randomly and with serially independent successive returns. However, Delong, Shleifer, Summers, and Waldmann (1990) and Bohl and Siklos (2004) argue that, in the short term, positive serial correlations occur in asset prices.

The existence of pricing errors in the market combined with the interdependent nature of assets means that agents are capable of predicting value movements. In this context, ETFs offer arbitrage opportunities. These opportunities occur because any decoupling of ETF shares relative to their fundamental value provides information to agents (Cherry, 2004).

Modern finance theory recommends diversification as a way to minimize the volatility of an investment portfolio. Thus, risk management can be performed using ETFs. Because of the idiosyncratic nature of ETFs and their recent arrival in the Brazilian market, this study proposes to examine the efficiency of share valuation by market agents. To this end, it is necessary to analyze whether shares reflect fair value, i.e., to verify the relationship between a fund's equity value and its market value. One possible outcome of this pricing inefficiency is the possibility of establishing strategies based on these differences that produce abnormal market returns. The present study addresses the Ibovespa iShare Index Fund ETF. This ETF is traded on the Brazilian Securities, Commodities and Futures Exchange (Bolsa de Valores, Mercadorias & Futuros de São Paulo - BM&FBOVESPA) under the code BOVA11, and its benchmark is the main Brazilian stock index, Ibovespa.

2 THEORETICAL FRAMEWORK

Although ETFs have only recently been introduced, they have been the focus of a substantial amount of international research. The literature discusses various issues, such as the efficiency and the pricing structure of index funds, arbitrage opportunities, ETF performance relative to conventional funds, and risk exposure through international index ETFs. However, in Brazil, studies on this instrument remain scarce.

In Asia, studies on ETFs, such as the study by Jares and Lavin (2004), have investigated Japanese and Hong

Kong indices and their traded shares on the U.S. market. In Taiwan, Lin, Chan and Hsu (2005) and Wang, Liau and Yang (2009) analyzed the dynamics of the relationships between ETFs and their respective indices. Despite instances of deviation from the fundamental value, the results did not reveal a capacity for systematic economic gains.

In the U.S. market, Dolvin (2009) and Marshall, Nguyen, and Visaltanachoti (2010) examined arbitrage margins for ETFs based on the S&P 500, the Standard and Poor's Depository Receipts (New York Stock Exchange: SPY), and the iShares S&P 500 Index Fund (New York Stock Exchange: IVV) indices. These studies indicate that arbitrage opportunities exist, particularly during periods of high volatility, based on the historical average. Furthermore, ETFs suggest a capacity as an indicator in the future market. However, Hasbrouck (2003) concluded that even with the inclusion of ETFs, the capacity to predict the net asset value (NAV) remains modest and is not significant compared with future contracts.

Simon and Sternberg (2005) discuss "overreactions" by the ETFs of European indices (the Deutsche Aktienindex [DAX], the Financial Times Stock Exchange Index [FTSE], and the Cotation Assistée en Continu [CAC]) traded in the U.S. After the closing of the European markets, the asset movements offer a margin of NAV prediction in subsequent periods. However, Kayali (2007) noted the infeasibility of arbitrage in the Turkish market with an ETF indexed to the Dow Jones Istanbul (DJIST).

Gallagher and Segara (2004) examined the performance of ETFs in the Oceania market and found a close fit between discounts and the NAV, with any decoupling quickly disappearing. According to the authors, this phenomenon implies an efficient Australian ETF market.

For the Brazilian market, Farias (2009) notes the influence of the PIBB (Papéis de Índice Brasil Bovespa [Brazil Bovespa Index Papers]) ETF on the liquidity of the IBrX-50. Farias concluded that there is no evidence that these papers will bring increased liquidity to the index. Borges, Eid, and Yoshinaga (2012) compared the Ibovespa iShare ETFs with other funds indexed on the same portfolio. They found that ETFs in Brazil display worse adherence compared to index funds. Yang, Cabrera, and Wang (2010) used various methods to investigate the return predictability of 18 ETFs of global indices traded in the U.S. The authors argued that considering the data-snooping effects, which were captured by bootstrapping techniques, it was impossible to predict the daily returns of the Brazilian index ETF.

Using a structural approach, Cherry (2004) investigated the excess volatility of ETFs relative to their benchmark and the volatility's determinants. The study demonstrated that this excess implies a high correlation between lagged discounts and futures returns. Thus, according to the author, strategies that consider such features are likely to generate abnormal returns. However, these approaches would display little profit in high-volume markets with high variation in discounts and in the case of international indices.

Although there are many studies on ETFs outside Brazil, many aspects have yet to be examined, particularly in emerging markets, such as Brazil, where these instruments are novel. Thus, the present article will analyze the relationships between ETF and NAV series and investigate whether EFTs create the opportunity for excess earnings while noting flaws in share pricing that do not conform to the efficient market hypothesis (EMH).

2.1 Exchanged-Traded Funds.

ETFs combine features of mutual funds and stocks. The primary goal of all ETFs, as is the case for index funds, is to offer the same profitability as a particular index through a passive investment strategy. However, unlike conventional funds, ETFs offer the same flexibility as a stock transaction because ETF shares are freely traded on a secondary market throughout the trading period of the stock exchange (Gastineau, 2001). Moreover, in Brazil, the shares of these funds and stock shares are redeemed according to D[1] + 0, whereas settlements are according to D + 3. In contrast, mutual funds are subject to D + 1 and D + 4, respectively (CVM [Comissão de Valores Mobiliários/Brazilian Securities and Exchanges Commission] Instruction, 2002).

The ETF investor can recreate a diversified portfolio that is equal to the index with only one asset. This fact confers two advantages for investors and portfolio managers:

1. There is a reduction in costs because there is no need for the spot purchase of the shares in the proportions of the index to which the fund is linked.
2. Increased acquisition speed occurs because a spot index can be acquired in only one transaction (Rompotis, 2005).

Mutual funds incur higher costs than ETFs. According to Wild (2007), in the U.S. market, conventional funds charge a management fee of approximately 1.67% per year. However, ETFs charge substantially lower management fees of approximately 0.2% per year, despite incurring brokerage and custody costs, because ETFs are traded as stocks. In Brazil, the fee charged by the Ibovespa iShare fund is 0.54%[2], whereas in 2010, the average fee for Brazilian stock funds was 2.20% (Anbima, 2011).

In short-term corporate finance, the protection of the value of the assets in the portfolio against market fluctuations is desirable. This strategy, known as hedging, can be used directly with ETF funds. The process is performed by borrowing shares referred to as short-position trades, which results in profits with a fall in asset prices (Meziani, 2007). Thus, profits obtained from the price decrease are offset against losses recorded in the price decrease of assets in the portfo-

[1] The term 'D' refers to the day in question.

[2] Available at http://br.iShares.com. Accessed on March 14, 2011.

lio. The sum of all of the traded fund shares does not necessarily reflect the equity value of the shares. There is a difference between the NAV and the total value of the secondary market shares. However, under normal circumstances, the returns are the same as those of the shares. Given their characteristics, ETFs may present certain disadvantages. For example, their performance may not fully reflect the index performance. The presence of a secondary market of shares with buying and selling forces can result in decoupling in relation to the NAV, which can cause momentary changes in discounts (Rompotis, 2007).

3 THE MODEL

The total value of the fund's assets per share should reflect the stock trade values after subtracting the transaction costs (Dolvin, 2009). In market equilibrium, ETF returns are equal to the index to which they are linked. Otherwise, the price disequilibrium would create arbitrage opportunities in the market.

The discounts are calculated as the differences between the *ln* share value of the ETF fund, traded on the market, and its NAV. Thus, at time *t*, the discount/premium variable d_t is defined as follows:

$$d_t = ln\left[\frac{P_t}{NAV_t}\right],$$

$$d_t = ln\,[P_t] - ln\,[NAV_t]. \quad \textbf{1.1}$$

Thus, large discounts correspond to high d_t values, albeit in absolute terms.

Arbitrage transactions can be performed when the value of the discount between the NAV and ETF values is large, which makes profitable redemption possible. However, this transaction type is inconvenient with respect to operating time and the redemption mechanism. According to iShare fund regulations, shares are delivered in minimum lots and with at least 95% of their value in assets in proportion to the index. In turn, the agent must sell these assets on the spot market to recover the fiscal investment. In addition, gains may be earned in the opposite direction through the creation of shares. However, a relatively large deviation between markets price of share and NAV are required to overcome costs (Dolvin, 2009).

An alternative redemption method from the fund is to buy and sell fund shares according to their discounts. This strategy entails purchasing ETF shares when the share value is well below the equilibrium level and selling later when balance has been restored.

Restoring the stability between ETF share values and the NAV intrinsically depends on the movement of these assets. The disequilibrium correction mechanism can be applied to both assets to gain stability. However, according to Cherry (2004), for this ETF operational strategy to be attractive, the main corrective movement must be via the ETF, not the NAV.

The discount value remains stable over time, which generates equal returns between the NAV and the ETF, leaving only random deviations. In principle, there is no predictive information on discounts that results in abnormal returns. The ETF shares returns can be broken down as follows:

$$R_t^{ETF} = R_t^{NAV} + \Delta d_t, \quad \textbf{1.2}$$

$$E\,[R^{ETF}] = E\,[R^{NAV}] + E\,[\Delta d_{t,\,t-1}], \quad \textbf{1.3}$$

where E[X] is the expectation operator and $\Delta d_{t,\,t-1}$ is d_t-d_{t-1}. However, the equilibrium condition implies that E[Δd_t]=0. Therefore, Eq. 1.3 yields

$$E\,[R^{ETF}] = E\,[R^{NAV}] \quad \textbf{1.4}$$

However, when breaking down E[Δd_t], it is possible to rearrange the relation of Eq. 1.3 as follows:

$$E\,[R^{ETF}] = E\,[R^{NAV}] + E\,[d_t] - E\,[d_{t-1}],$$
$$= E\,[R^{NAV}] + E\,[d_t] - d_{t-1}, \quad \textbf{1.5}$$

where E $[d_{t-1}]= d_{t-1}$ because the value of the discount in period *t* is already known. This formulation makes clear that fluctuations in past discounts result in different returns between the ETF and the NAV. All else being equal, an increase in the value of d_{t-1} results in a decrease in R^{ETF} or, conversely, generates increases in R_t^{NAV}.

According to Pontiff (1997), the necessary condition for the R^{ETF} to correlate more closely with discounts over time is an R^{ETF} volatility greater than R^{NAV}. Using variance as a proxy for volatility, if

$$Var\,(R^{ETF}) > Var\,(R^{NAV}), \quad \textbf{1.6}$$

then,

$$\frac{Cov\,(\Delta d_t, R^{NAV})}{Var(\Delta d_t)} > -\frac{1}{2}. \quad \textbf{1.7}$$

where $Cov\,[\Delta d_t, R^{NAV}]$ is the covariance and $Var[\Delta d_t]$ is the variance. Eq. 1.7 proposes that an ETF with excess volatility correlates more closely with the prior discounts compared with the NAV. The differences between R^{ETF} and R^{NAV} are represented by Δd_t. If, on average, R^{NAV} decreases by less than half of the Δd_t, as a consequence, R^{ETF} will increase to more than half of the Δd_t, which suggests that R^{ETF} is more volatile than R^{NAV}. The interpretation of Eq. 1.7 can be deduced by the regression of R_t^{NAV} against Δd_t. If the prior discounts offer information that is predictive of ETF returns, the strategy that considers this factor account is likely to obtain consistently abnormal gains.

Another key point is the analysis of the rate of convergence to equilibrium between the series. The slower the correction movements are, the larger the opportunity to capture these disequilibria and benefit from them. One possibility is to ascertain whether previous discounts and the disequilibrium correction rate offer feasible opportunities by performing a simulation. This technique can provide optimal values for the degree of decoupling between the ETF and the NAV series that will enable a profit on a market transaction.

4 METHOD

4.1 Excess Volatility.

As previously mentioned, it is of substantial importance to the operating strategy to analyze the difference between the variances of the series. For the ETF to correlate more closely with past discounts, it is necessary to establish whether $\sigma^2_{ETF} > \sigma^2_{NAV}$. Using the regression

$$R_t^{NAV} = \beta \Delta d_t + u_t \, , \tag{1.8}$$

where $u_t \sim N(0,\sigma^2)$, it is possible to ascertain the relationship between the variances. Thus, the interpretation of the coefficient is a logical extension of Eq. 1.7. If the regression coefficient presents β>-1/2, the correlation between R^{ETF} and the lagged discounts will be larger than that of R^{NAV}.

Once this relationship is identified, the second step is to investigate the opening of the operating margins. The fact that the ETF shares are designed to reproduce the Ibovespa index suggests strong cointegration. Therefore, the stationarity of the regression residuals is verified as follows:

$$\ln P_t^{ETF} = \alpha + \beta \ln P_t^{NAV} + u_t, \tag{1.9}$$

where P_t^{ETF} and P_t^{NAV} are the prices indexed at time t of the ETF share and the NAV, respectively. With the cointegration identified, it becomes necessary to combine short- and long-term relationships. This process is performed by embedding long-term disequilibrium as an error correction term. It is possible to understand the relationship structure from the regression

$$R_t^{ETF} = \omega + \zeta R_t^{NAV} + \varphi \, [\ln P_{t-1}^{ETF} - \beta \ln P_{t-1}^{NAV} - \alpha] + \epsilon_t \, ,$$
$$R_t^{ETF} = \omega + \zeta R_t^{NAV} + \varphi \, u_{t-1} + \epsilon_t \, . \tag{1.10}$$

where parameter φ indicates the rate of convergence toward equilibrium between the series. The value of this parameter is closely related to the operating margins because the strategy uses this convergence. The parameters ζ and β represent the short- and long-term relationships, respectively, between the ETF values and the NAV.

4.2 Simulation.

The formulation and assembly of the strategy discussed above, captured by the degree of decoupling between the ETF and the NAV series, were diagnosed by means of simulation. Thus, one can determine whether the strategy's ability to generate gains is feasible and consistent over time. Based on this technique, the optimal entry and exit points in terms of maximizing returns are estimated.

For the strategy adopted to be feasible, the returns on these transactions must not only be above market level but also exceed transaction costs. Strategies that exhibit the characteristic of generating average gains over market level are generally referred to as "winners".

The simulation includes the search for and the analysis of this strategy. The first step is to verify the hypothesis that the strategy outperforms the Ibovespa. The second step is to measure the degree of efficiency against the maximum return value. This strategy is expressed by a strategy function proposed by Cherry (2004) and consists of offering the formal rules of trading according to which the agent should buy an asset, keep it in the portfolio, or sell it.

Consider σ_d the standard deviation of discounts and $\bar{d}$ the value of the average discount between the ETF and the NAV. The input and output parameters of the transactions are defined by the degree of standard deviations from the mean. Thus, the strategy function E(t) is defined as follows:

$$E(d_t) = \begin{cases} 1, & \text{if } d_{t-1} \leq \bar{d} + B\sigma_d \\ 1, & \text{if } d_{t-1} \leq \bar{d} + A\sigma_d \text{ and } E(d_{t-1}) = 1 \\ 0, & \text{otherwise} \end{cases} \tag{1.11}$$

The returns F in period t are calculated as follows:

$$F[E(d_t)] = (1+R_{t+1}^{ETF})E(d_t) + [1-E(d_{t-1})] \, , \tag{1.12}$$

where the image of the function $E(d_t)$=1 represents the ETF asset in portfolio and $E(d_t)$=0 otherwise. Given the strategy function $E(d_t)$, the simulation aims to indicate the optimum "A" and "B" points that provide the highest total return and earnings statistics. To operationalize the generation of these results, the simulation process was as follows:

1. Establish a uniform random value for A and B.
2. Generate strategy function values.
3. Extract the values of the total returns.

The value of the total return is computed as follows:

$$\prod_{t=1}^{T} F[E(d_t)] = \prod_{t=1}^{T} \{(1+R_{t+1}^{ETF})E(d_t) + [1-E(d_t)]\} \, . \tag{1.13}$$

That is, capitalization is achieved at a negotiation frequency of five minutes. In other words, when the ETF strategy is not used, the interest earned is zero. Conversely, when the strategy is used, the transaction return is computed.

4.3 Bootstrapping.

The values achieved using the simulation refer to absolute and punctual performance and do not provide confidence intervals for performance inference. According to White, Sullivan, and Timmermann (1999), this problem occurs because the strategy is based on a single performance of the stochastic process of the movement of the ETF and the NAV assets. Thus, it is possible to capture any strategy that suggests a priori a higher predictive ability, when in fact this ability is only the result of chance, which is known as data snooping.

To avoid this problem, the resampling technique known as bootstrapping is used, which enables empirical distributions of the strategy results to be established through alternative asset realization. This method is important to the definition of confidence intervals to assess whether the performance obtained in the strategy simulation differs from average market gains.

The characteristic mode of the bootstrapping technique is successive resampling from the original sample. Bootstrapping considers the elements of an iid sample[3]. According to Brooks (2008), the observations of financial time series are not usually independent and identically distributed. To overcome this problem, bootstrapping is employed for time series using non-overlapping blocks (Hall, Horowitz, & Jing, 1995).

Block bootstrapping is primarily used in time series to construct samples with independent replacement that also retain the dependence of realizations. However, the quality of the technique is influenced by the choice of block size. To scale the blocks to maximize their efficiency, studies such as the one by Hall, Horowitz, and Jing (1995) describe the conditions and rules for optimal choice. Then, to estimate the distribution of returns, the size of the blocks (l) should follow $l \approx T^{1/4}$, where T is the sample size.

In this context, the sample generation technique employs the following procedure, where $S=\{X_i : i=1, \ldots T\}$ is the original series and the k-th block is formed by the following vector $B_k=\{X_w, \ldots X_{w+l-1}\}$, where $w=T/l\times k$, l is the block size and T is the original sample size. After h blocks are constituted, they are selected randomly with replacement to generate an alternative series as follows:

$$S^*=\{B^1_j, B^2_j \ldots B^h_j\} \\ =\{X^1_w \ldots X^1_{w+l-1}, X^2_w \ldots X^2_{w+l-1}, \ldots, X^h_w \ldots X^h_{w+l-1}\}, \quad \textbf{1.14}$$

where j is a random value between 1 and h.

Then, the parameter λ from the unknown distribution is estimated by $v(S)=\hat{\lambda}$, where S is the original series. However, there are no a priori indicators of the accuracy of the estimator v. To assess the standard error of $\hat{\lambda}$, estimates are made of the resampled series $v(S^*_b)=\hat{\lambda}^*_b$, where b is the b-th bootstrapping replication. According to Efron and Tibshirani (2000), the bootstrap estimate of standard error, $se(\hat{\lambda})$, is calculated as follows:

$$se_B=\left\{\sum_{b=1}^{B}[\hat{\lambda}^*_b-\hat{\lambda}^*(\cdot)]^2/(B-1)\right\}^{\frac{1}{2}} \quad \textbf{1.15}$$

where $\hat{\lambda}^*(\cdot)=\sum_{b=1}^{B}\hat{\lambda}^*/B$ and B is the total number of bootstrap samples. In this study, the parameter $\hat{\lambda}$ is estimated as follows:

$$v(S)=T^{-1}\sum_{d_t=1}^{T}F(d_t)-R^{NAV}_{t+1}=\bar{\lambda}. \quad \textbf{1.16}$$

The bootstrapping technique used to define the confidence intervals was that of percentiles extracted from the $\hat{\lambda}^*$ replications. According to Davison and Hinkley (2009), no transformation is required to use this method. The confidence intervals are calculated using the cumulative function $\hat{G}$ of $\hat{\lambda}^*$. Given the confidence intervals, the lower bands α and upper bands 1-α of the $\hat{G}$ percentiles are set at the 1-2α confidence level (Efron & Tibshirani, 2000). That is,

$$[\hat{\lambda}_{inf}, \hat{\lambda}_{sup}]=[\hat{G}^{-1}(\alpha), \hat{G}^{-1}(1-\alpha)], \quad \textbf{1.17}$$

where the inverse, $G^{-1}(\alpha)=\hat{\lambda}^{*(\alpha)}$, is the α-th percentile of the bootstrap distribution. Therefore, the confidence interval can be defined as follows:

$$[\hat{\lambda}_{inf}, \hat{\lambda}_{sup}]=[\hat{\lambda}^{*(\alpha)}, \hat{\lambda}^{*(1-\alpha)}]. \quad \textbf{1.18}$$

Bootstrapping enables us to establish estimates of the confidence intervals for the means of the amount of the returns of the strategy $\bar{\lambda}$. Therefore, it is possible to determine whether the transactions create feasible opportunities to outperform the market without the occurrence of data snooping.

5 EMPIRICAL RESULTS

The sample was extracted from the Consultoria, Métodos, Assessoria e Mercantil S/A [Consulting, Methods, Advisory and Mercantile] (CMA) group database. Approximately 26,400 ETF asset prices available under the code BOVA11 and the Ibovespa index score were collected. The asset series displays the closing of the respective assets at a frequency of five minutes. The period is from 05/04/2009 to 08/05/2010 throughout the opening hours of trading but excluding after-market prices. Any missing values that originate from the CMA database were interpolated linearly to form a continuous series.

5.1 Regression.

The iShare ETF fund ultimately replicates the Ibovespa. Figure 1 shows the evolution of the historical series of both assets, which enables this relationship to be verified.

To make the strategy that consider this feature viable, it is important that the BOVA11 shares are more volatile than the NAV. To satisfy that assumption, the regression $R^{NAV}_t=\beta\Delta d_t+u_t$ was used. Dividends and interest on equity were not incorporated in the return series because these amounts are reinvested in the fund.

In Table 1, the coefficient $\beta \geq -1/2$. As explained, this coefficient implies excess volatility of the ETF in relation to the NAV. The p-values marked by (*) indicate a high significance level because they are an order of magnitude less than 10^{-3}.

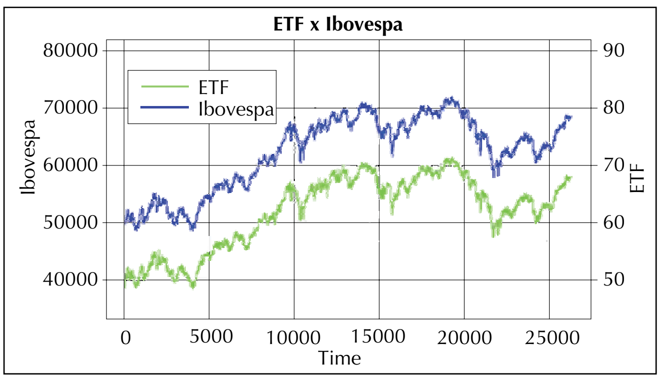

Figure 1 *ETF Historical Series and Ibovespa*

[3] The term refers to the distribution of random independent and identically distributed variables.

Table 1 *Excess Volatility Regression*

	Coefficient	Standard Error	Z-statistic	*p*-value
β	-8.02×10^{-5}	1.38×10^{-5}	-5.795	0.000*
	LogFV	131,614.4	AIC criterion	-10.1244

The returns of the above series are stationary and obtained using the Dickey-Fuller test at a significance level of 1%.

To check for cointegration, the stationarity of the residuals from Eq. 1.9 must be tested. However, the critical values used in the Dickey-Fuller test are not appropriate to test the hypothesis of the stationarity of residuals. Therefore, the Engle-Granger test was used with one lag to test for the occurrence of cointegration. The test indicated the occurrence of cointegration with a *p*-value of 0.0001 (Table 2).

Table 2 *Engle-Granger Cointegration Test*

Engle-Granger Test	Value	*p*-Value
τ - statistic	-55.9527	0,000*

To analyze the short-term relationship between the series, the regression from Eq. 1.10 was used. Because of the heteroskedasticity of the series, which was diagnosed by White's test, the GARCH (1.1) method was used. The model parameters were quantified using the maximum likelihood estimators (MLE), in which the following function was maximized:

$$L(\varepsilon_t|\Theta) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\sigma^2_t - \frac{1}{2}\sum_{t=1}^{T}\frac{(R_t^{ETF}-\omega-\zeta R_t^{NAV}-\varphi u_t)^2}{\sigma^2_t}, \quad (1.19)$$

where $\sigma^2_t = \varnothing_0 + \varnothing_1\varepsilon^2_{t-1} + \gamma\sigma^2_{t-1}$ and Θ is the parametric space.

According to Table 3, all of the coefficients were statistically significant at a level of 1%.

Table 3 *Regression of Equation 1.10*

	Coefficient	Standard Error	Z-statistic	*p*-value
	Mean Equation			
ζ	0.5237	0.0019	268.0967	0.000*
φ	-0.0038	4.06×10^{-5}	-94.6230	0.000*
ω	-4.01×10^{-5}	1.37×10^{-5}	-2.9205	0.0035
β	0.9857	4.39×10^{-5}	22449.60	0.000*
α	0.0514	0.0002	284.6514	0.000*
	GARCH (1,1) Variance Equation			
$\varnothing_0$	4.55×10^{-7}	3.7×10^{-9}	123.1097	0.000*
$\varnothing_1$	0.66	0.0057	115.9246	0.000*
γ	0.25	0.0045	55.2283	0.000*
	LogFV	134611.3	AIC criterion	-10.3548

As expected, the coefficient φ is negative, which indicates that discounts above the long-term equilibrium, i.e., $u_{t-1}<0$, exert a positive influence on returns in time *t*. In contrast, small discounts, i.e., $u_{t-1}>0$, imply market losses on the ETF in subsequent time *t*. The coefficient ζ was 0.52. In the long term, the series presents a narrow equilibrium with β approximately 0.98. From Table 3, it can be inferred at a 95% level that β does not differ from one, which indicates that parameter α represents the average discount between assets.

5.2 Simulation.

After verifying the relationship between the series, it is essential to analyze the possible occurrence of gains from the strategy function $E(d_t)$ for capturing lagged discounts. A sample size of 1,000 was used to determine the values of the A and B parameters. Initially, without a consideration of the transactions costs, one can infer from Figure 2 the relationship between the parameters and the total return, on the one hand, and the Sharpe ratio, on the other.

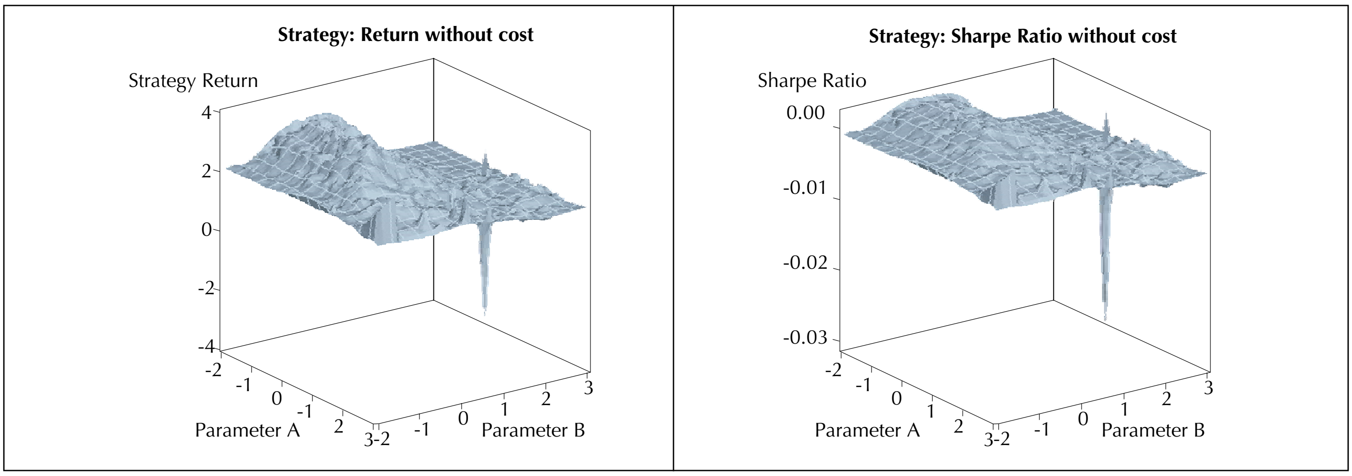

Figure 2 *Return Strategy (left) and Sharpe Ratio (right) in Relation to Parameters A and B*

Table 4 summarizes the strategy simulation for an ETF, which far outperformed the index to which the ETF is linked by achieving a 213% profit—40.5% above the market.

Table 4 *Strategy Simulation without Costs*

N Simulation	Maximum Return
1000	3.13
A	B
-0.1752	-0.2960

The figures provide an initial idea of the evolution of gains relative to the market. From the outset, it can be observed that the strategy consistently outperforms the market, although the first 2,000 observations fluctuate between gains and losses. The average return per operation every five minutes was 4.4×10^{-3}%, which offered an average gain above market level (excess) of approximately 3.11×10^{-3}% for the analyzed period.

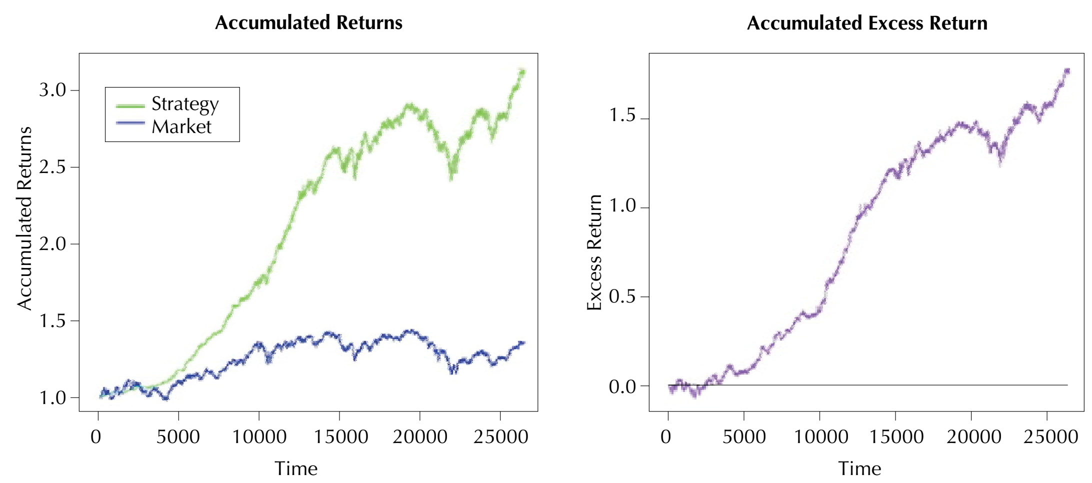

Figure 3 . *Evolution of the Strategy and Market (left) and Excess Return of the Strategy (right) without Cost*

Table 5 *Exploratory Statistic of Strategy without Cost*

	$\hat{\lambda}$	Geometric Mean	σ
Operation Gain	4.4×10^{-3}%	4.3×10^{-3}%	12×10^{-4}
Excess Gain	3.11×10^{-3}%	1.2×10^{-3}%	18.45×10^{-4}

Despite the large gains offered by the optimized strategy, the results do not strongly reflect reality because they do not include transaction costs. According to the Bovespa website, operating costs (exchange fees) for day trades are 0.025%. The brokerage costs charged by brokers and custody were disregarded because they can be fixed costs. Taxes are levied only in the case of capital gain. That is, a tax expense is only incurred when the strategy is profitable or when a brokerage service tax (Imposto sobre Serviços - ISS) of any type is charged.

To refine the strategy with costs of 0.025% per operation, a simulation of 1,000 was performed. The introduction of costs altered the gains pattern (Figure 4).

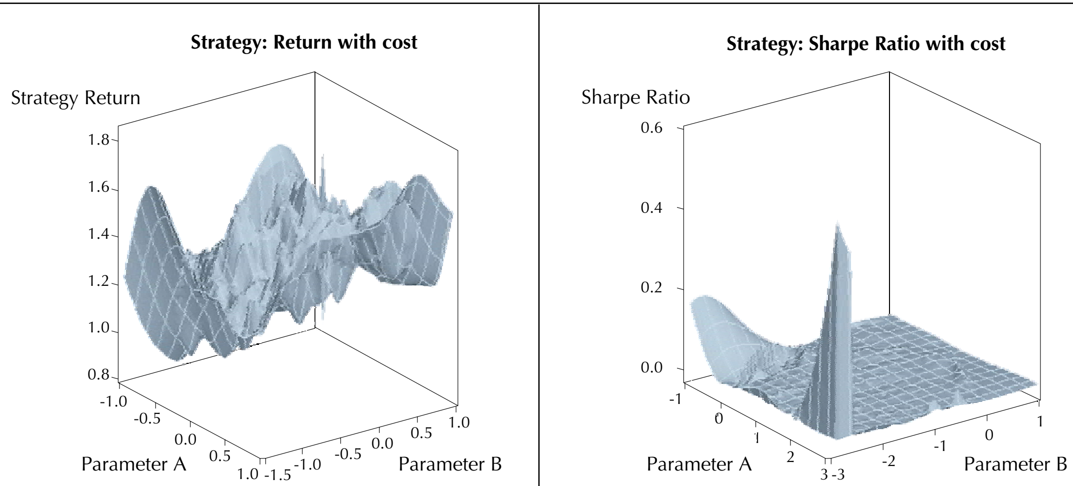

Figure 4 *Return Strategy (left) and Sharpe Ratio (right) in Relation to Parameters A and B*

When costs are included, the values of A and B become less narrow because small deviations that were previously profitable are now covered by costs. These results are shown in Table 6.

Table 6 *Strategy Simulation with Cost*

N Simulation	Maximum Return
1000	1.7585
A	B
0.2854	-0.3411

The profitability of the strategy decreased substantially to 76%. The effect of adding costs is shown in Figure 5. It can be observed that even with the introduction of operating costs, the strategy proves to be a "winner" by outperforming the market by approximately 40% over the analyzed period.

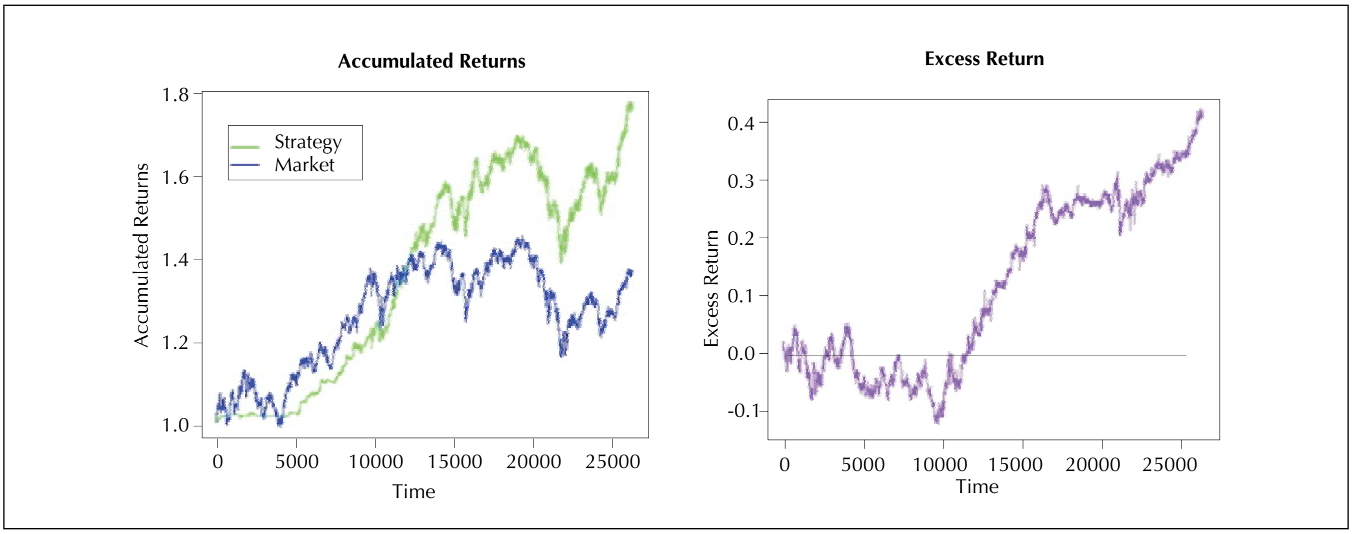

Figure 5 *Evolution of the strategy and Market (left) and Excess Return of the Strategy (right) with Cost*

However, this tactic only displays gains after 10,000 observations (Table 7).

Table 7 *Exploratory Statistic of Strategy with Cost*

	$\hat{\lambda}$	Geometric Mean	σ
Operation Gain	0.22×10^{-4}%	0.21×10^{-4}%	11×10^{-4}
Excess Gain	0.0931×10^{-4}%	0.13×10^{-4}%	17×10^{-4}

5.3 Bootstrapping.

The assessment of the simulation results included only one realization of the data generating process, which may cause problems for the analysis, such as the previously mentioned data snooping. Using the bootstrapping technique, the empirical distribution of the strategy's profitability $\hat{\lambda}$ was estimated for operations with and without transaction costs. A total of 1,000 alternative series were generated from the sample of 26,400 observations, each with 2,200 size-12 blocks. However, the following analyses differed with respect to the results of the transaction strategies for both sets of circumstances.

First, for the operations without costs, the average excess gains of the realizations were approximately 2.05×10^{-3}%, in contrast to 3.11×10^{-3}% for the original series. Figure 6 shows the distribution $\hat{\lambda}$, i.e., the average income generated by the strategy. Therefore, the results presented indicate a departure from data-snooping effects.

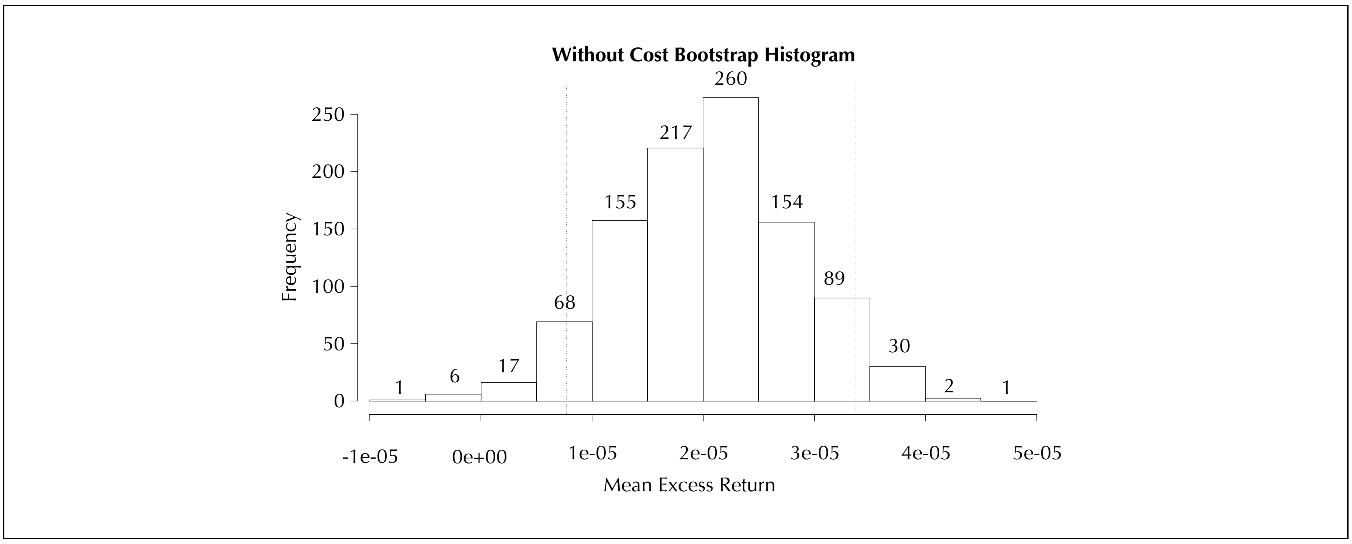

Figure 6 *Empirical Distribution of Excess Gain of the Strategy without Costs*

As shown in Table 8, the standard error se_B was 2.31×10^{-11}, which provided a safety margin. The parameter λ lies within the following confidence interval, which is equally distributed between the dashed lines in Figure 6.

Table 8 *Exploratory Statistics of the Empirical Distribution of the Strategy Return without Costs*

	$\hat{\lambda}^*(\cdot)$	Median	Standard Error	Kurtosis	Skewness
Excess Gain $\hat{\lambda}^*$	2.05×10^{-3}%	2.055×10^{-3}%	2.31×10^{-11}	0.0538	-0.0425

Therefore, according to the previously mentioned confidence interval for a confidence level of 95%, the hypothesis that the strategy outperforms the market can be accepted. The results endorse those results found in the first simulation. Thus, without including costs, the strategy $E(d_t)$ results in abnormal gains.

Finally, the effects of the transaction costs appear divergent for cost-free operations. Based on Eq. 1.16, the histogram in Figure 7 shows the empirical distribution of returns using the strategy. With the same sample and block configuration, the excess gain by operation $\hat{\lambda}^*(\cdot)$ was -1.44×10^{-3}%, as shown in Table 9.

Table 9 *Exploratory Statistic of the Empirical Distribution of the Strategy Return with Costs*

	$\hat{\lambda}^*(\cdot)$	Median	Standard Error	Kurtosis	Skewness
Excess Gain $\hat{\lambda}^*$	-1.44×10^{-3}%	-1.49×10^{-3}%	9.44×10^{-4}%	0.0079	0.1214

After providing inputs for the hypothesis test for the same level of confidence as above, the confidence interval is constructed. Figure 7 shows the range generated using the percentile technique.

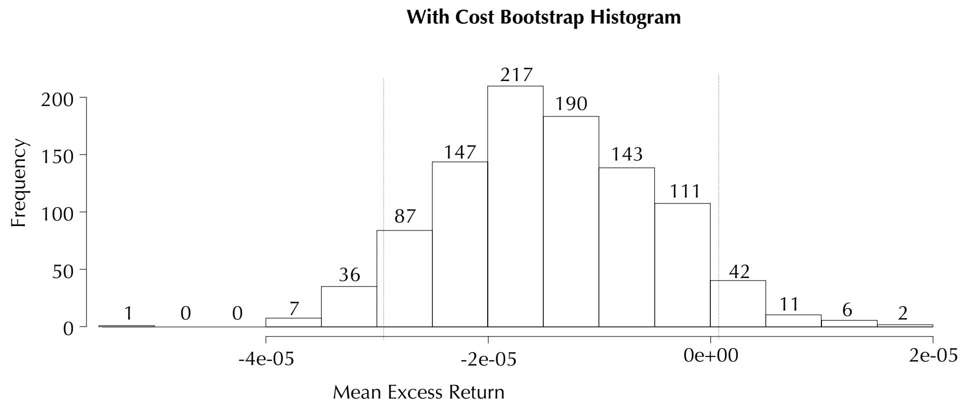

Figure 7 *Empirical Distribution of Excess Gain of the Strategy with Cost*

Despite the apparent outperforming of the market demonstrated in the simulation with transaction costs, the bootstrapping results indicate the data-snooping effects on the strategy. It can be observed that the null hypothesis of the strategy outperforming the market is rejected. In short, if the data-snooping issue is bracketed, there is no evidence that these transactions result in a "winner" strategy.

6 CONCLUSION

This study analyzed the behavior of the time series of the Ibovespa iShare ETF and its respective benchmark. Additionally, the study addressed the characteristics of this investment type and the arbitrage mechanism using decoupling between the ETF and the NAV. Moreover, through simulation, the optimal points of the strategy were examined, while incorporating possible momentary flaws in fund share pricing. To circumvent the effects of data snooping, the bootstrapping technique was applied.

The results show that the fund shares are more volatile than their benchmark asset. In the long term, the series proved to be strongly coupled and cointegrated. In the short term, the parameter $\varnothing$ of Eq. 1.10 suggests a good operating margin.

The strategy simulation signaled abnormal gains with and without operating costs with excess returns of approximately 176% and 36%, respectively. However, the results represent gross amounts and do not include capital gains taxes.

Although the results of the simulation are promising, the estimation of the empirical distributions of excess return per operation did not invite the same conclusion. The divergence of the results may be the result of the confidence intervals of the average return. The strategy without including the costs presents evidence of real gains. However, when operating costs are considered, the bootstrap estimator rejects the hypothesis of abnormal returns.

Based on the study's results, a strategy based on lagged discounts in the iShare fund is not able to outperform the market. The preliminary performances generated by the simulation can be credited to chance because the reproduction of the data-generation process indicated data-snooping effects. Thus, this technique fails as a "buy-and-hold" strategy. Therefore, the results enable the hypothesis regarding the efficiency of this market to be rejected.

References

Anbima. (2011). *Fundo de investimento: relatório de taxas de administração*. São Paulo: Associação Brasileira das Entidades dos Mercados Financeiros e de Capitais.

Basu, S. (1983). The relationship between earning`s yield, market value and return for NYSE common stocks: further evidence. *Journal of Financial Economics, 12* (1), 129-156.

Bohl, M., & Siklos, P. (2004). Empirical evidence on feedback trading in mature and emerging markets. *Applied Financial Economics, 18* (17), 1379-1389.

Bondt, W. F. M. D., & Richard, T. R. (1985). Does the stock market overreact? *The Journal of Finance, 40* (3), 793-805.

Borges, E. C., Eid Jr, W., & Yoshinaga, C. E. (2012). Exchange traded funds versus fundos indexados no Brasil. *Revista de Finanças Aplicadas, 1* (0), 1-18.

Brooks, C. (2008). *Introductory econometrics for finance*. (2nd ed.). South-Western Pub: Cambridge University Press.

Cherry, J. (2004). The limits of arbitrage: evidence from exchange traded funds. *Working Paper*, University of California-Berkeley.

Damodaran, A. (2009). *Avaliação de investimentos*. (2. ed.). Rio de Janeiro: Qualitymark.

Davison, A. C., & Hinkley, D. (2009). *Bootstrap method and their application*. (11th ed.). New York: Cambridge University Press.

Delong, J. B., Shleifer, A., Summers, L.H., & Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance, 45* (2), 379-395.

Dimson, E., & Mussavian, M. (2000). Market Efficiency. *The Current State of Business Disciplines, 3*, 959-970.

Dolvin, S. D. (2009). ETFs: Arbitrage opportunities and market forecasting. *Journal of Index Investing, 1* (1), 107-116.

Efron, B., & Tibshirani, R. (2000). *An introduction to the Bootstrap*. Florida: Hapman Hall/CRC.

Fama, E. F. (1970). Eficient capital markets: a review of theory and empirical work. *The Journal of Finance, 25* (2), 383-417.

Farias, M. A. (2009). ETFs, PIBB, e a liquidez das ações. *Caderno de Pesquisa Investimento*, Fundação Getúlio Vargas, *1*, 3-57.

Gallagher, D. R., & Segara, R. (2004). The performance and trading characteristics of exchange-traded funds. *Working Paper*, The University of New South Wales.

Gastineau, G. L. (2001). An introduction to exchange-traded funds (ETFs). *Journal of Portfolio Management and Economics, 27* (3), 88-96.

Guedj, I., & Huang, J. (2009). Are ETFs replacing index mutual funds? *Working Paper*, University of Texas.

Hall, P., Horowitz, J. L., & Jing, B. (1995). On blocking rules for bootstrap with dependent data. *Biometrika, 82* (3), 561-574.

Hasbrouck, J. (2003). Intraday price formation in us equity index markets. *Journal of Finance, 58* (6), 2375-2400.

Instrução Normativa CVM n. 359. (2002, janeiro). Brasil. Recuperado de http://www.cvm.gov.br/asp/cvmwww/atos/exiato.asp?File=%5Cinst%5Cinst359.htm.

Jaffe, J., Keim, B. D., & Westerfield, R. (1989). Earnings yields, market values, and stock returns. *The Journal of Finance, 44* (1), 135-148.

Jares, T. E., & Lavin, A. M. (2004). Japan and Hong-Kong exchange traded fund ETFs: disconts, return and trading strategies. *Journal of Financial Services Research, 25* (1), 57-69.

Kayali, M. M. (2007). Pricing efficiency of exchange traded funds in Turkey: early evidence from the Dow Jones Istanbul 20. *International Research Journal of Finance and Economics, 10*, 14-23.

Lin, C., Chan, S., & Hsu, H. (2005). Pricing efficiency of exchange traded funds in Taiwan. *Journal of Asset Management, 7* (1), 60-68.

Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2010). ETF arbitrage. *Working Paper*, New Zealand, University of Auckland.

Meziani, A. S. (Fall, 2007). Long-short investment strategies using ETFs. ETFs and Indexing, 2007 (1), 106-110.

Pontiff, J. (1997). Excess volatility and closed-end funds. *The American Economic Review, 87* (1), 155-169.

Poterba, J., & Shoven, J. B. (2002). Exchange traded funds: A new investment option for taxable investors. *Working Paper*.

Regulamento Ishare Ibovespa. (2009, novembro). Recuperado de: http://www.bmfbovespa.com.br/etf/download/RegulamentoiBovespa.pdf.

Rompotis, G. (2005). An empirical comparing investigation on exchange traded funds and index fund performance. *Working Paper*, Athens Greece, University of Athens.

Rompotis, G. (2007). An empirical look on exchange traded funds. *Working Paper*, Athens Greece, University of Athens.

Ross, S., Westerfield, R., & Jordan, B. (2008). *Fundamentals of corporate finance alternate value*. (8th ed.). New York: Mc Graw Hill.

Simon, D. P., & Sternberg, J. S. (2005). Overreaction and trading strategies in European ishare. *Journal of Alternative Investment, 8* (1), 29-41.

Wang, C. C., Liau, Y. S., & Yang, J. J. W. (2009). Information spillover in the spot and ETF indices in Taiwan. *Global Journal of Business Research, 3* (1), 117-131.

White, H., Sullivan, R., & Timmermann, A. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *The Journal of Finance, 54* (5), 1647-1691.

Wild, R. (2007). *Exchange-traded funds for dummies*. Indianapolis, Indiana: Wiley Publishing.

Yang, J., Cabrera, J., & Wang, T. (2010). Nonlinearity, data-snooping, and stock index ETF return predictability. *European Journal of Operational Research, 200* (2), 498-507.