Economics Bulletin

Volume 37, Issue 4

The Gap Effect on the Brazilian Exchange

Paulo Sergio Ceretta
Federal University of Santa Maria

Alexandre Silva Da costa
Federal University of Santa Maria

Abstract

We spotted a market anomaly, related to the opening gap of three asset classes, the exchange rate, Bovespa blue chips and the Ibovespa, the major stock index in Brazil; and further investigated through algorithmic trading simulation to verify our initial hunch. Our assumption, that turned out to be correct, we called the Gap Effect, and it is that big slumps or spikes in the opening gap on the beginning of the trading day, tend to a reversal, or a significant come back in the first fifteen minutes of the trading day, creating great opportunities for intraday trading. Using a large dataset of tick-by-tick data, we found a pattern which can spot striking opportunities to develop algorithmic trading strategies (long or short), based on the early movements of a security. Moreover, we confirm through Data Panel with Thresholds that the larger is the opening gap (up or down), the larger is the chance to a price reversal in the early minutes of the trading day, just after the initial auction is over.

Citation: Paulo Sergio Ceretta and Alexandre Silva Da costa, (2017) "The Gap Effect on the Brazilian Exchange", *Economics Bulletin*, Volume 37, Issue 4, pages 2505-2516
Contact: Paulo Sergio Ceretta - ceretta10@gmail.com, Alexandre Silva Da costa - xandicosta@gmail.com.
Submitted: September 11, 2017. **Published:** November 19, 2017.

1. Introduction

Trying to make a quick buck betting on market inefficiencies is not unusual, since the beginning of speculative times investors and scholars studied market inefficiencies, perhaps Bachelier, in the beginning of the last century, is one of the pioneers. He stated that influences that determine the movements of the Stock Exchange are innumerable; events, past, present or even anticipated, often showing no apparent connections to price changes, yet have repercussions on its course and the fluctuations may not be foreseen. Some may be probable and this probability can be evaluated mathematically, although far from being an exact science, Bachelier (1900).

But let us go forward to 1965, when a different view started to gain momentum. Samuelson, posting a rather general stochastic model of price change, deduced a theorem in which next period´s price differences are likely to be uncorrelated if not completely independent to the previous one. Two lines of thinking have dominated the financial literature since then, the defenders that the market place is a fair and equal environment, where participants have the same chance of succeeding. And the others, who advocate pro a world in which chaos wreaks havoc.

Going along the history of this everlasting financial dichotomy, the EMH (efficient market hypothesis) started to be built up with Fama (1970); supporting fiercely the EMH and its notions of equality. The EMH in general terms, is that markets provide accurate signals for resource allocation: markets in which, firms can make investment decisions, and investors can choose among securities that represent ownership of firm´s activities under the assumption that prices reflect all available information of that security. So, markets that "fully reflect" all available information are called "efficient", Fama (1970).

A weaker, but economically more realistic, version of the hypothesis is that prices reflect information up to the point where the marginal benefits of acting on the information (the expected profits to be made) do not exceed the marginal costs of collecting it (Jensen 1978). After Jensen, comes Fama (1991) and by his own words, that he approaches market efficiency literature with "trepidation" and himself admits that since trading costs and positive information became an issue, the extreme version of the EMH is false. Grossman and Stiglitz (1980) hammered hard against EMH, stating that it is impossible to have a competitive equilibrium that all arbitrage costs are eliminated based mostly by differences of information. On the EMH side, perhaps the most important work is presented by Burton and Malkiel (2003) with strong and almost indisputable evidence of the EMH, with a prolonged study of the S&P500 and Wilshire stock indexes, where only a handful of investors outperform those indexes in the long run. In other words, median and large capitalization professionally managed equity funds has underperformed the S&P 500 index by almost two percentage points over the past, 10, 15 and 20 years. Summing it up, actively managed funds are regularly outperformed by the benchmark with equivalent risk.

Although we do not want to get involved in defending one or the other (one of the authors strongly believe on the EMH, and the other one, not so much), we want to present a small market inefficiency (or anomaly), discovered through algorithm trading on high frequency data, which can bring arbitrage profits.

Trading securities in the Brazilian Stock Market, with the naked eye, we noticed something unusual, stocks with big drops or big spikes (Gaps) in the opening bell, tended to reverse their momentum during the first hour of the trading day. Furthermore, we noticed some predictability of the closing price, based on that spike or slump. Then, we created an algorism to check in which time interval would be best to buy or sell, assets traded in the stock exchange that started the day with high volatility; or with big price difference compared to the last closing day.

Our first impression, was very true and throughout the paper we will show striking evidence of a large negative relationship between the opening gap and the price of the fifteenth minute of trading. As per the predictability of the closing price, we find some relationship, however that relationship was not strong enough to write about. But before we go ahead, we would like to give some definitions of the following topics:

1.1 Market Inefficiencies

The presence of calendar anomalies has been documented extensively for the last 3 decades in financial markets, Berument and Kiymaz (2001). The most common ones are related to the calendar such as *The January Effect*, *The Day of the Week Effect* or *The Weekend Effect* and so on. Cross (1973) studied the S&P 500 index and concluded that over the period of 1953 to 1970 that mean returns on Fridays are higher than mean returns on Mondays. Similar results are concurred by French (1980) and Gibbons and Hess (1981). Ariel (1987) identified anomalies in US stock prices varying from the beginning and end of the calendar month: on the last day of the month and the three following days the changes in prices are significantly more positive than negative (*The Turn of the Month Effect*).

This so-called *Effects* are anomalies that investors can bet on, long or short to produce abnormal returns, contradicting EMH. There are another kinds of anomalies, not only motivated by the calendar; Levis (1988) studied and proposed investment strategies based on PE ratios and dividend yield, generating significant results, concentrating the portfolio based on firm size in the London Stock Exchange. Barone (1990) studied the *Settlement Effect*, on the Italian stock market where on the settlement day (payment and delivery day), usually the last day of the calendar month, one can observe between spot and forward prices.

More related to our study, comes Niederhoffer and Osborne (1966) that pointed out that the accurate record of stock market ticker prices displays striking properties of dependence. They found that after a decline of 1/8 of a point between transactions, an advance on the next transaction is three times as likely as a decline. Moreover, they disclosed that after two prices changes in the same direction, the odds in favor of a continuation in that direction are almost twice as great as after two changes in contrary directions.

Krueger and Kennedy (1990), fiercely affirmed that the league affiliation of the Super Bowl winner, predicts market direction, and investors could outperform the market by reacting to Super Bowl game outcomes.
There is also a notion, that when an anomaly becomes public and investors start to bet on it, the market starts to understand it´s causes and automatically sorts out and levels off its efficiency, Jensen (1978).

1.2 Algorithmic Trading

The use of algorithmic trading, where computers make the decisions and manage the trading process, usually at a high frequency, became common in most major financial markets, since the late 1990s, Chaboud at al. (2014). Nowadays, high frequency trading (HFT) represents the majority of trading volume in today's markets; and HFT refers to orders submitted by algorithms, that emit orders or order cancellations in reaction to market updates or other events within milliseconds, Boehmer, Fong and Wu (2012).

Since its introduction there was a wide interest in understanding its potential and the impact on market dynamics, not without concern as its adverse selection costs and proximity to large cables network. Eichengreen, Lafarguette and Mehl (2016) identified exogenous technological changes by the connections of countries to submarine fiber optic cables used for electronic trading, so trading houses with more computer power and proximity to the exchange, could have benefits as their order would have an advantage over overseas and private investors. The design of trading algorithms requires sophisticated mathematical models, a solid analysis of financial data, and a deep understanding of how markets and exchanges function, Cartea, Jaimungal, Penalva (2015).

High speed and algorithmic trading can advantage of other less technological ways of placing orders by having a much faster response to information releases, Chordia, Green and Kottimukkalur (2015).

As everything in finance literature, there is no agreement on the effectiveness of algorithmic trading. Biais, Foucault and Moinas (2011) showed that electronic trading display advantage towards humans when it comes to speed as computers react rapidly to public information. Another view is presented by Foucault, Hombert and Rosu (2012) says that in a world of no asymmetric information, electronic trader´s speed advantage does not increase trading profits, it simply increases adverse selection costs. Moreover, Chaboud, Hjalmarsson and Vega (2015) states that HFT (High Frequency Trading) or algorithmic trading improve price efficiency through lower return autocorrelations and fewer arbitrage opportunities.

1.3 High Frequency Data

Because of fast-growing computer power, gathering financial data is easier than ever. Data are no longer recorded daily or weekly. Many large institutions began to collect so-called tick-by-tick exchange rates in the early eighties, Zhou (1996). In this study, we purchased intraday data from Cero Technologies, a leading technology and trading software provider to the Brazilian market. The data is *tick-by-tick* and in a *15 minutes* interval, from February 2012 to April 2015 of the most traded stocks, the dollar exchange rate and the major index in Brazil, the Ibovespa. According to Investopedia, a tick is a measure of the minimum upward or downward movement in the price of a security. A tick-by-tick refers to the change in the price of a security from trade to trade, in our case, within the same trading session, not daily.

1.4 Gap

A gap is a significant price movement of a security or commodity between two trading sessions (or any given interval), such that there is no overlap in the trading ranges for the two days; or sometimes, the second day's opening price is outside the first day's trading range, Downes and Goodman (2010). Another concept is that a gap is a break between prices on a chart that occurs when the price of a stock makes a sharp move up or down with no trading occurring in between. Gaps can be created by factors such as regular buying or selling pressure, earnings announcements, a change in an analyst's outlook or any other type of news release, Bulkowski, (2003).

To go further, one must understand the process of the opening bell and the initial auction in stock markets. Nowadays, with electronic trading everywhere, exchanges function very similarly, and we will describe briefly how is it in the Bm&fBovespa: At the beginning of the day and before the execution of the first trade for a specific security, the reference price will be equal to its previous day's closing price adjusted security, and this equivalent to the center of the bands from the intraday limits, the rejection bands, bands of static limits and auction bands. After the execution of the first trade, the security's reference price will be: (a) the price of the last trade of the security, which will be equivalent to the center of the rejection bands and bands auction; (b) the opening price, which will be equivalent to the center of the bands of static limits, or (c) the closing price will be equivalent to the center of the intraday limit bands.

2. Literature review

The majority of studies that deal with market anomalies is conducted with daily data, and we can highlight the work of Jagadeesh and Titman (1993). In their study, they affirm that strategies which buy stocks that have performed well in the past and sell stocks that have performed poorly in the past, generate significant positive returns over 3 to 12 month holding periods.

Jagaddeesh (1990), presented empirical evidence of predictability of individual stock returns. The negative first-order serial correlation in monthly stock returns was found highly significant, plus significant positive serial correlation is found at longer lags, and the 12-month serial correlation is strong.

As per intraday data, there is not much literature that investigate momentum, although we found some similarities and some pioneering work of Gao, et al. (2015), studied ETFs from 1993 to 2014 documented an intraday momentum pattern: the first half-hour on the market predicts the last half-hour returns. Although their statistical significance was small, they set the ground rules for our finds. Another similarity found in their work and ours is that in more volatile days such as economic news and data release days, predictability is stronger.

Lam et al. (2007) examined whether a day´s surge or plummet in stock price serve as market entry or exit signal for major world stock indices. They found that trading rules perform well in the Asian Indices but not in Europe´s or America´s. Venter (2009) studied ultra-short term return predictability based on intraday momentum and contrarian effects on the Johannesburg Stock Exchange. He found statistically significant return predictability in returns calculated from mid-quote prices. However, when returns are calculated under bid-asking price profitability and predictability disappear.

Grant, Wolf and YU (2005) gave a long-term assessment of intraday price reversals in the US stock index futures market following large price changes at market open. They found highly significant intraday reversals in yearly and day-of-the-week investigations. Similar to what we found, they found strong reversal signals when the market opens with large positive price changes. Assess, Moskowiz and Pedersen (2013), found consistent value and momentum return premium across eight diverse markets and asset classes, and a strong common factor structure among their returns. Value and momentum returns correlate more strongly across asset classes than passive exposures to the asset class, but value and momentum correlate negatively.

3. Data and method

Our dataset comprises 797 observations of 10 blue chips, 1 stock index and 1 exchange rate. The blue chips are the most traded Brazilian stocks by volume: ABEV3 (Ambev S.A.), BBDC4 (Banco Bradesco S.A. Preferencial), BRFS3 (Brasilfoods S.A.), CIEL3 (Cielo S.A.), ITSA4 (Investimentos Itau Preferencial), JBSS3 (JBS S.A.), PETR3 (Petrobras S.A. PN); PETR4 (Petrobras S.A. ON), UGPA3(Ultrapar Participações S.A.) and VALE5 (Vale do Rio Doce S.A. Preferencial); the main stock index in Brazil: INDM (Ibovespa); the Dollar exchange rate DOLM. From February 2012 to April 2015, in two databases – tick-by-tick and 15-minutes intraday data. We use the first one to spot the optimum interval to conduct our research as per the following formula:

$X_{i,d}$ Initial price of day *d*;

$X_{f,d}$ Final price of day *d*;

$X_{j,d}$ Price after *j* minutes of opening of day *d*;

$gap_d = \ln\left(\frac{X_{i,d}}{X_{f,d-1}}\right)$ day d gap; (1)

$r_{j,d} = \ln\left(\frac{X_{j,d}}{X_{i,d}}\right)$ return of *j* minutes of opening of day *d*; (2)

$gap_d \sim F$ gaps in the sample have a probability distribuition F;

$G = \left\{gap_d, d = 1, \dots, n : \ gap_d < F^{-1}\left(\frac{\alpha}{2}\right), gap_d > F^{-1}\left(1 - \frac{\alpha}{2}\right)\right\}$ Gaps that overstep; (3)

the limits stablished by the interval of confidence alfa% ;

$\widetilde{gap}_d = \{gap_d : \ gap_d \in G\}$ gaps that overstep the limits. (4)

Firstly, we sorted our data out to show a logarithm base of the gap between the closing of the market on the day before, and the opening of the current trading session on a daily basis. Then we correlated through computer bootstrap, that gap and different returns, on different times throughout the trading day, in an intraday, *tick-by-tick* basis. Our simulation reverted that the most significant and predictable return occurs for the entire sample we chose, around after 15 minutes of the opening, just after the end of the initial auction.

The next step is to estimate the econometric methodology of the panel with thresholds, explained in this section. The methodology follows three steps: 1) identification of the number of thresholds with the Likelihood-Ratio test (LR test); 2) identification of regimes in the dependent variable; 3) estimation of OLS regression considering the independent variables and the different regimes.

Hansen (1999) explains that threshold regression models specify that individual observations can be divided into classes based on the value of an observed variable. His method proposes the estimation of threshold and regression slopes using fixed-effects transformations.

Heterogeneity is a common research barrier faced by researchers when utilizing a method such as a panel data. Hansen (1999) questioned if the regression functions should be identic to all individuals of a certain sample. If each is different, the structural relationship may vary between them. The traditional approaches (fixed and random effects) allow to consider this heterogeneity with a somewhat partial approach. Hsiao (2014), discussed alternate options of those approaches.

So, the Threshold panel data, rises as an option that can accommodate in a more refined way, heterogeneity from the individuals of a given sample. The approach proposed by Hansen (1999), allows to describe jumps or structural breaks from different individuals, segmenting the sample based in a value of a determined variable. In this way, an initially heterogenic sample can me divided in one, two, three or four sub-samples, where a specific structural relationship between the variables can be identified. The model proposed by Hansen (1999), allow the regression equation coefficient to change its value depending on the sub-sample or regime that it is set. The model with two regimes, or single threshold can be described as (5).

$$y_{it} = \mu_i + x_{it} I(q_{it} \leq \gamma)\beta_1 + x_{it} I(q_{it} > \gamma)\beta_2 + \varepsilon_{it} \qquad (5)$$

I is an indicating function assuming values from $I = 1$, when $(q_{it} \leq \gamma)$ and 0, and $I = 1$, when $(q_{it} > \gamma)$ and 0; q_{it} is the threshold variable, γ the threshold parameter that divides the equation in two regimes with coefficients $\beta = (\beta_1, \beta_2)$; ε_{it} is the error term assumed to be independent and identically distributed (i.i.d.) with zero mean and finite variance that may have heteroscedasticity,

An alternate representation of (5) can be described by (6)

$$y_{it} = \mu_i + \beta z_{it}(\gamma) + \varepsilon_{it} \qquad (6)$$

In (6), $z_{it}(\gamma) = (x_{it} I(q_{it} \leq \gamma), x_{it} I(q_{it} > \gamma))$ and B = (β_1 and β_2). It is defined a sample $\Gamma = (\underline{\gamma}, \overline{\gamma})$, where $\underline{\gamma} > \min\{q_{it}\}$ and $\overline{\gamma} < \max\{q_{it}\}$. Note that for each value of $\gamma \in \Gamma$, the vector $z_{it}(\gamma)$ will assume an specific form. The coefficient estimation is through OLS and the selection is by grid search of the coefficient estimate which generate a least Sum of Squared Error (SSE), hence, for each value $\gamma \in \Gamma$, the coefficients and the Sum of Squared Error ($SSE_\gamma = \sum \sum \varepsilon_{it}^2$) are given, the most adequate estimates are the ones which minimize the function SSE_γ on the space Γ.

The model with three regimes, double threshold can be described as (7).

$$y_{it} = \mu_i + x_{it} I(q_{it} \leq \gamma_1)\beta_1 + x_{it} I(\gamma_1 < q_{it} \leq \gamma_2)\beta_2 + x_{it} I(\gamma_2 < q_{it})\beta_3 + \varepsilon_{it} \qquad (7)$$

A more intuitive way of describing the double-threshold is as in (8).

$$y_{it} = \begin{cases} \mu_i + \beta_1 x_{it} + \varepsilon_{it}, & q_{it} \leq \gamma_1, \\ \mu_i + \beta_2 x_{it} + \varepsilon_{it}, & \gamma_1 < q_{it} \leq \gamma_2, \\ \mu_i + \beta_3 x_{it} + \varepsilon_{it}, & \gamma_2 < q_{it}. \end{cases} \qquad (8)$$

In (8), the sample is divided into three regimes, depending only if the threshold variable is smaller, greater or it will be in an interval defined by the thresholds. By definition, this procedure guarantees more heterogeneity within each regime, as it contributes to more realistic coefficients. The Hansen (1999) model supports up to three thresholds. For a better, comprehension of the estimation, one should consider the alternative of equations (7) and (8), given by (9).

$$y_{it} = \mu_i + \beta z_{it}(\gamma_1, \gamma_2) + \varepsilon_i \tag{9}$$

In (9), $z_{it}(\gamma_1, \gamma_2) = (x_{it}I(q_{it} \leq \gamma_1), x_{it}I(\gamma_1 < q_{it} \leq \gamma_2), x_{it}I(\gamma_2 < q_{it}))$ and $B = (\beta_1, \beta_2$ and $\beta_3)$. Note that for each pair $(\gamma_1, \gamma_2) \in \Gamma x \Gamma$, the vector $z_{it} = (\gamma_1, \gamma_2)$ assumes and specific form. The coefficient estimation is through OLS and the selection is by grid search of the coefficient estimates which generate a least Sum of Squared Error, or, for each value of γ_1 and $\gamma_2 \in \Gamma x \Gamma$, the coefficients are obtained by the Sum of Squared Error ($SSE_{\gamma 1,\gamma 2} = \sum\sum \varepsilon_{it}^2(\gamma_1, \gamma_2)$), and the most adequate estimations are the ones which minimize the function in the space $\Gamma x \Gamma$.

For values of (γ_1, γ_2) the coefficients (β_1, β_2 and β_3) are linear and the estimation of OLS through grid search is adequate. The coefficients are the ones which minimize the Sum of Squared Error ($SSE_{\gamma 1,\gamma 2} = \sum\sum \varepsilon_{it}^2(\gamma_1, \gamma_2)$).

In this context, it is need to verify the threshold effect (γ) significance. The difference $\beta_1 - \beta_2$ has to be sufficiently large that (γ) is significant. Lagrange test (LR), proposed by Hansen (1999), is described by (10.a, 10.b and 10.c).

$$LR(\gamma) = (SSE(lm) - SSE(\gamma))/\sigma_\gamma^2 \tag{10.a}$$
$$LR(\gamma 1, \gamma 2) = (SSE(\gamma) - SSE(\gamma 1, \gamma 2))/\sigma_{\gamma 1,\gamma 2}^2 \tag{10.b}$$
$$LR(\gamma 1, \gamma 2, \gamma 3) = (SSE(\gamma 1, \gamma 2) - SSE(\gamma 1, \gamma 2, \gamma 3))/\sigma_{\gamma 1,\gamma 2,\gamma 3}^2 \tag{10.c}$$

The LR Test is robust and heteroscedasticity has its critical values determined by a bootstrap procedure. In (10.a), if the statistical value of LR overcomes the critical value, the function assumes two regimes, where the association between the independent and dependent variables are distinct, for at least for one of the variables. On the other hand, if the LR statistics does not overcame its critical value, the linear model with fixed effects is more adequate. The same goes to 10.b and 10.c, however the comparison should be 1vs2 threshold and 2vs3 threshold.

4. Results

The first run of the Data Panel should determine the number of thresholds. At this point one must understand the notation of a threshold, which can also be described as a *subsample*, Hansen (2000). So, the number of thresholds, regimes or subclasses are displayed on Table 1.

Results on Table 1, display the existence of the threshold, LR test for tripe threshold are 2427.058 (p-value < 0.000). Therefore, is presented a flexible model of four regimes, which is better adjusted to the generating data process, allowing for a significant improvement of the pseudo R2 from 0.112 to 0.690. The next step is to display the output of the non-dynamic effect with the thresholds. Noted that the VIF (Variance Inflation Factors) showed no multicollinearity between the variables.

Table 1. Test for determining the number of thresholds

Model	SSE	Threshold			LR Test	p-value	Pseudo
		1	2	3			R2
Zero Threshold	1.031						0.112
Single Threshold	0.555	0.021			8194.782	0.000	0.528
Double Threshold	0.457	0.003	0.021		2038.578	0.000	0.612
Triple Threshold	0.365	-0.025	0.003	0.021	2427.058	0.000	0.690

SSE = Sum of Squared Errors; LR Test = Test for threshold effect; Pseudo R2 = 1-(SSE/TSS).

Table 2. Threshold effects in non-dynamic panels, dependent variable log-return

Regimes	Variables independent	
	Gap[t]	Volatility[t-1]
First Regime	2.599	-0.038
	(18.624)	(-3.084)
Second Regime	0.425	-0.038
	(12.666)	(-3.084)
Third Regime	-0.629	-0.038
	(-22.407)	(-3.084)
Fourth Regime	-1.602	-0.038
	(-22.907)	(-3.084)

The main line is the coefficient estimate, and the t-test is in parentheses. We consider White-corrected standard errors for heteroscedasticity.

As per Table 2, one can identify that the strength of our assumption is on the extremes. When the market opens on the downside (first regime) returns on the 15-minute of trading produced the strongest coefficient, 2.599 to the positive side. Quite the contrary happens when the market opens with big positive gaps, represented by the fourth regime; the trend is that markets reverse to the negative side and the coefficient produced by our model is -1.602.

By the *Betas* on the regimes, being the first and second regime when the market opens with price drops, the Gap shows positive response, meaning that when the opening is negative, the prices should reverse in the first 15-minutes. Moreover, on the first regime, where prices dropped the most, it is more likely that the price gets to a reversal. Even though, on the second regime the variable is significant, on the first threshold the likelihood is much greater.

That itself creates a great arbitrage opportunity, one can bet against the gap while prices are unchanged during the initial auction. In this case, an order to buy a tanking market has great odds of succeeding, just after the initial auction is done. In other tests we conducted to this research, we separated the odd of a reversal according to the size of the Gap. It turned out, as proven by the threshold panel, that the highest the gap is, the highest is the chance to a reversal. In a choppy market it can create great buying points for day traders. The same happens when the gap is in the other direction. When there is a significant price spike, or a big gap, traders tend to position themselves in the selling side after the initial auction. Once again, the further along we go on the regimes, the higher the *betas*. That similarity in finance, which is higher dependence in extreme values, has been seen in many works in financial literature such as Righi and Ceretta (2011); Costa, Ceretta, Muller (2015) and Jondeau and Rockinger (2006).

An easier and more practical way to see the efficiency of our trading strategy is by calculating how many times, our theory behaved the way we thought it would. Therefore, we design a way to display the percentage of times a trader/stock market operator would be successful following our strategy. We divided it into 2 groups: when the opening gap is largely, and the other, the opposite, largely positive. We considered gaps bigger than 1% for the stocks; and 0.5% for Indexes (DOLM and INDM) – for the positive side. For the negative, were considered gaps lower than -1% for the stocks; and -0.5% for the indexes. We called a Hit when the asset went the way we predicted; a Zero when it did not change; and a Miss when it went the other way.

Table 3. Percentage of HIT-ZERO-MISS for negative openings gaps <1% for the stocks and <0.5% for the indexes.

	ABEV3	CIEL3	INDM	BRFS3	VALE5	JBBS3	DOLM	PETR4	ITSA4	UGAP3
Hit	65%	63%	63%	58%	56%	55%	53%	55%	55%	53%
Zero	0%	4%	5%	5%	6%	11%	10%	2%	14%	6%
Miss	35%	33%	33%	37%	39%	35%	37%	43%	31%	41%

On Table 3 one can note that when the openings are largely negative, all stocks changed direction in the first fifteen minutes of trading a greater number of times. Moreover, the three first assets behaved according to our strategy almost 2 times more than not. Chart 1 brings with easier notation that all assets studied, behaved more according to the way we described than in the wrong direction.

Table 4 brings the same results, but when the opening is too optimistic. For the Dollar (DOLM) our strategy worked almost four times than not. UGPA3, three times than not. When betting in the main index in the Brazilian market (INDM) the level of success is also very significant. When it is zero we did not consider a loss as the trading cost normally is too low or irrelevant.

Chart 1. Percentage of HIT-ZERO-MISS for negative openings gaps <1% for the stocks and <0.5% for the indexes.

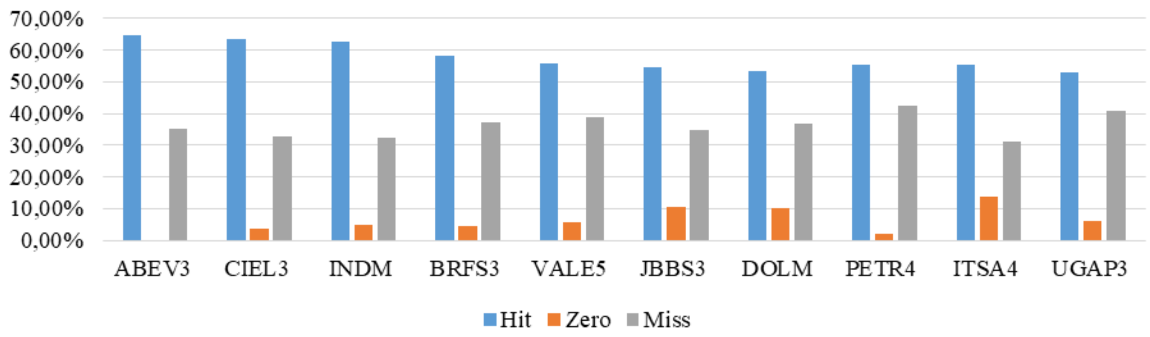
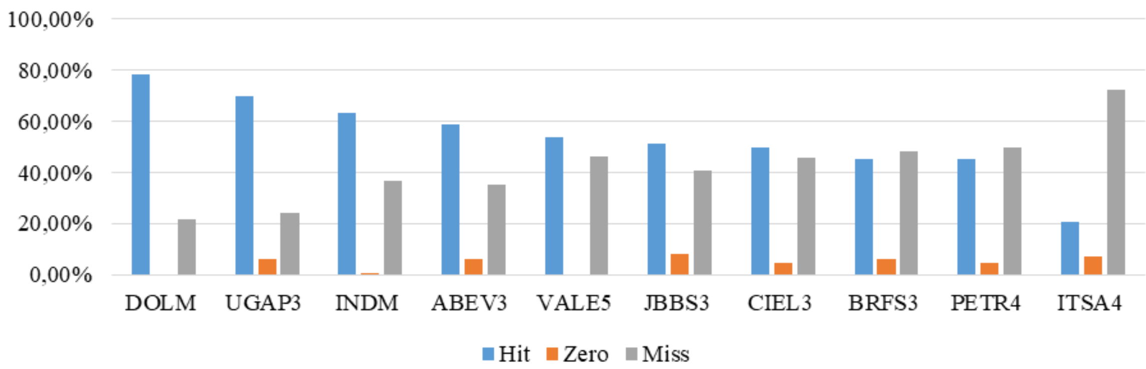


Table 4. Percentage of HIT-ZERO-MISS for positive openings gaps >1% for the stocks and >0.5% for the indexes.

	DOLM	UGAP3	INDM	ABEV3	VALE5	JBBS3	CIEL3	BRFS3	PETR4	ITSA4
Hit	78%	70%	63%	59%	54%	51%	50%	45%	45%	21%
Zero	0%	6%	0%	6%	0%	8%	4%	6%	5%	7%
Miss	22%	24%	37%	35%	46%	41%	46%	48%	50%	72%

Chart 2. shows that even though the figures tell a stronger reality according to our theory, some of the assets behaved differently, BRFS3, PETR4 and ITSA4 displayed different results we predicted. We had to discard the results for BBDC4 and PETR3 as the data seemed in disarray.

Chart 2. Percentage of HIT-ZERO-MISS for positive openings gaps >1% for the stocks and >0.5% for the indexes.



Conclusions

At the end of our research we gathered evidence of almost all we proposed, and with a statistical significance that exceed our initial expectations. At least, for Brazilian blue chips the pattern is clear, securities that start off the day with a big swing up or down, tend to revert, or lose steam of some of that trend in the next few minutes after the initial auction. The reason for that maybe psychological, behaviorist or else and it is not our field of expertise. We suggest that this behavior should be researched as a suggestion for future studies, by scholars whose area of expertise lies on comportment and psychological affairs.

Following this pattern, the likelihood of success in daytrading tends to increase. The idea is simple, when the market starts really strong, investors should sell short during the opening auction and buy back to cover the position after the market reverts which by our calculations may happen around the fifteenth minute after the market is open. The opposite is also valid, when the market is really weak, investors should buy stocks during the opening auction and sell after prices are being freely fluctuating.

As per our study, the main difficulty was to deal with intraday data which is frequently presented in an erratic data format. Although the creation of a good algorithm solved part of that hassle. Our next step is to evaluate if volume plays a role in that pattern, and we would like to test this anomaly in different stock markets.

References

Ariel, R. A. (1987) "A monthly effect in stock returns." *Journal of financial economics* **18**(1): 161-174.

Asness, C. S., Moskowitz, T. J., and Pedersen, L. H. (2013) "Value and momentum everywhere." *The journal of finance* **68**(3): 929-985.

Baltagi, B. (2008) "*Econometric analysis of panel data*". John Wiley & Sons.

Barone, E. (1990) "The Italian stock market: efficiency and calendar anomalies." *Journal of Banking & Finance* **14**(2): 483-510.

Berument, H. and Kiymaz, H. (2001) "The day of the week effect on stock market volatility." *Journal of economics and finance* **25**(2): 181-193.

Biais, B., Moinas S, and Foucault, T. (2011) "Equilibrium high frequency trading". *International Conference of the French Finance Association* (AFFI).

Boehmer, E., Fong, K., and Wu, J. (2012) "*International evidence on algorithmic trading*". AFA 2013 San Diego Meetings Paper.

Bulkowski, T. N. (2003) "*Trading classic chart patterns*". John Wiley & Sons.

Cartea, Á., Jaimungal, S., and Penalva, J. (2015) "*Algorithmic and high-frequency trading*". Cambridge University Press.

Chaboud, A., Hjalmarsson, E. and Vega, C. (2015) "The Need for Speed: Minimum Quote Life Rules and Algorithmic Trading." *Manuscript, Federal Reserve Board.*

Chaboud, A. P., Chiquoine, B., and Hjalmarsson, E. (2014) "Rise of the machines: Algorithmic trading in the foreign exchange market." *The journal of finance* **69**(5): 2045-2084.

Chordia, T., Green, T. C. and Kottimukkalur, B. (2015) "Do High Frequency Traders Need to be Regulated? Evidence from Algorithmic Trading on Macro News." *Working paper*. Goizueta Business School, Emory University. Atlanta.

Downes, J. and J. E. Goodman (2010) "*Dictionary of finance and investment terms*". Barron's educational series.

Eichengreen, B., Lafarguette, R., Mehl, A. (2016) Cables, sharks and servers: Technology and the geography of the foreign exchange market". *NBER Working Paper*. National Bureau of Economic Research.

Fama, E. F. (1991) "Efficient capital markets: II." *The journal of finance*, **46**(5): 1575-1617.

Foucault, T., Hombert, J., and Rosu, I. (2016) "News trading and speed." *The journal of finance* **71**(1): 335-382.

Gao, L., Han, Y., Li, S. Z., and Zhou, G. (2015) "Intraday momentum: The first half-hour return predicts the last half-hour return." *Available at SSRN 2440866.*

Gibbons, M. R. and P. Hess (1981) "Day of the week effects and asset returns." *Journal of business*: 579-596.

Grant, J. L., Wolf, A. and Yu, S. (2005) "Intraday price reversals in the US stock index futures market: A 15-year study." *Journal of Banking & Finance* **29**(5): 1311-1327.

Grossman, S. J. and Stiglitz, J. E. (1980) "On the Impossibility of Informationally Efficient Markets." *The American Economic Review* **70**(3): 393-408.

Hansen, B. E. (2000) "Sample splitting and threshold estimation." *Econometrica* **68**(3): 575-603.

Hsiao, C. (2014) "*Analysis of panel data*". Cambridge university press.

Im, K. S., Pesaran, M. H., and Shin, Y. (2003) "Testing for unit roots in heterogeneous panels." *Journal of econometrics* **115**(1): 53-74.

Jegadeesh, N. (1990) "Evidence of predictable behavior of security returns." *The journal of finance* **45**(3): 881-898.

Jegadeesh, N. and Titman, S. (1993) "Returns to buying winners and selling losers: Implications for stock market efficiency." *The journal of finance* **48**(1): 65-91.

Jensen, M. C. (1978) "Some anomalous evidence regarding market efficiency." *Journal of financial economics* **6**(2/3): 95-101.

Jensen, M. C. (1978) "Some anomalous evidence regarding market efficiency." *Journal of financial economics* **6**(2/3): 95-101.

Jondeau, E. and Rockinger, M. (2006) "The copula-garch model of conditional dependencies: An international stock market application." *Journal of international money and finance* **25**(5): 827-853.

Levin, A., Lin, C., and Chu, C. J. (2002) "Unit root tests in panel data: asymptotic and finite-sample properties." *Journal of econometrics* **108**(1): 1-24.

Levis, M. (1989) "Stock market anomalies: A re-assessment based on the UK evidence." *Journal of Banking & Finance* **13**(4): 675-696.

Malkiel, B. G. (2003) "The efficient market hypothesis and its critics." *The Journal of Economic Perspectives* **17**(1): 59-82.

Malkiel, B. G. and Fama, E. F. (1970) "Efficient capital markets: A review of theory and empirical work." *The journal of finance* **25**(2): 383-417.

Righi, M. B. and Ceretta, P. S. (2011) "Extreme values dependence of risk in Latin American markets." *Economics Bulletin* **31**(4): 2903-2914.

Samuelson, P. A. (1965) "Proof That Properly Anticipated Prices Fluctuate Randomly", *Industrial Management Review* 6:2 p.41

Silva da Costa, A., Ceretta, P. S., and Müller, F. M. (2015) "MARKET MICROSTRUCTURE–A HIGH FREQUENCY ANALYSIS OF VOLUME AND VOLATILITY INTRADAY PATTERNS ACROSS THE BRAZILIAN STOCK MARKET." *Brazilian Journal of Management/Revista de Administração da UFSM* **8**(3).

Venter, J. H. (2009) "Intraday momentum and contrarian effects on the JSE." *Investment Analysts Journal* **38**(70): 47-60.

Lam, V. W. S., Chong, T. T. L. and Wong, W. K. (2007) "Profitability of intraday and interday momentum strategies." *Applied Economics Letters* **14**(15): 1103-1108.

Zhou, B. (1996) "High-frequency data and volatility in foreign-exchange rates." *Journal of Business & Economic Statistics* **14**(1): 45-52.