UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

MAICON DANIEL ROSA RIBEIRO

THE PRICE DISCOVERY PHENOMENON IN THE STOCK MARKET

Porto Alegre
2023

MAICON DANIEL ROSA RIBEIRO

THE PRICE DISCOVERY PHENOMENON IN THE STOCK MARKET

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título de Mestre em Economia.
Orientador: Prof. Dr. Flávio A. Ziegelmann

Porto Alegre
2023

CIP - Catalogação na Publicação

Rosa Ribeiro, Maicon Daniel
The price discovery phenomenon in the stock market / Maicon Daniel Rosa Ribeiro. -- 2023.
51 f.
Orientador: Flávio Augusto Ziegelmann.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, BR-RS, 2023.

1. Price Discovery. 2. Kalman Filter. I. Ziegelmann, Flávio Augusto, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

MAICON DANIEL ROSA RIBEIRO

THE PRICE DISCOVERY PHENOMENON IN THE STOCK MARKET

Dissertação submetida ao Programa de PósGraduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título de Mestre em Economia.

Aprovado em: Porto Alegre 21 de Dezembro de 2022.

BANCA EXAMINADORA:

Dr. Flávio A. Ziegelmann - Orientador
Universidade Federal do Rio Grande do Sul – UFRGS

Prof. Dr. Gustavo Fruet Dias
School of Economics, University of East Anglia and CREATES

Prof. Dr. Marcelo Perlin
Universidade Federal do Rio Grande do Sul - UFRGS

Prof. Dr. João Caldeira
Universidade Federal de Santa Catarina - UFSC

Table of Contents

Foreword

The road so far has been long... I am glad that the path has been full of wonderful people always willing to help.

The list is not short, which makes sense since the path was not short. I would like to thank my family (my mother, Patricia, for her kindness, my father, Daniel, for the incentive, my grandfather, Luis, for the temperance, and my grandmother, Marisa) and my friends (specially Mrs Fernandes, one of the few that I have the honour to call Friend).

Also I would like to thank Professor Flavio Ziegelmann, for his patience and aid in all the steps, Professor Gustavo Dias, for the insights in the modeling process, and Professor Marcelo Perlin, for his willingness to share his data.

This research had the CNPQ as the main financial contributor.

Abstract

This research aims to investigate the *Price Discovery* phenomenon between Brazilian and American assets. To do so one can use two different pairs of assets: S&P500 index and SPIXI11 ETF and Ibovespa index and EWZ ETF. With these pairs, one can investigate if the equilibrium price is more affected by American or Brazilian assets. The data is in high frequency, intraday data with 5, 10 and 15 frequencies. Using the Kalman Filter and the *Component Share* metric, the *Price Discovery Dynamic Contribution* is implemented to evaluate the phenomenon dynamics. As main results, it can be seen that the phenomenon occurs in the American market for Brazilian and American assets, i.e., the American market inputs more information in the equilibrium price and is, therefore, a *Satellite* market. The measures point out that, despite the metric not being constant over time, there is no change in the balance of power, i.e., the market that leads the process does that during the full time frame. And, as the final result, it can be seen that the exchange rate has a significant role in the *Price Discovery* phenomenon.

Keywords: Price Discovery, Kalman filter, Component Share.

Resumo

Essa pesquisa tem por objetivo investigar o fenômeno de *Descoberta de Preços* entre ativos Brasileiros e Americanos. Para realizr essa tarefa, é possível se utilizar dois pares diferentes de ativos: o índice S&P500 e o ETF SPXI11 e o índice Ibovespa e o ETF EWZ. Com esses pares, é possível se investigar se o preço de equilíbrio é mais afetado por ativos brasileiros ou americanos. Os dados estão em alta frequência, dados de frequência intradiária com 4, 10 e 15 minutos. Usando o Filtro de Kalman e a métrica *Component Share*, implementa-se a métrica *Price Discovery Dynamic Contribution* (Contribuição Dinâmica de Descoberta de Preços) com o intuito de se quantificar a dinâmica do fenômeno. Como resultados principais, verifica-se que o fenômeno ocorre no mercado americano tanto para ativos brasileiros quanto para ativos americanos. Isto é, o mercado americano imputa mais informação no preço de equilíbrio e é, por consequência, o mercado *Satélite*. As medidas apontam que, mesmo quando não há constância ao longo do tempo, não há mudança na balança de poder, ou seja, o mercado que guia o processo, o faz durante todos os intervalos de tempo. E, como resultado final, pode ser visto que a taxa de câmbio apresenta um importante pale no fenômeno da *Descoberta de Preços*.

Palavras-Chave: Descoberta de Preços, Filtro de Kalman, *Component Share*.

1 INTRODUCTION

With the enhanced globalization of financial markets, the number of non-US companies cross-listing shares, on a US exchange, has substantially increased (EUN; SABHERWAL, 2003). Many authors argue that there is evidence of reduction in the capital costs for companies that are cross-listed. There are two main explanations for this reduction in cost. One is related to the investor recognition and the other to the dissipation of the premium risk related to barriers between two markets, i.e., once there is no barrier the cost may fall (FOERSTER; KAROLYI, 1993; FOERSTER; KAROLYI, 1998; FOERSTER; KAROLYI, 1999). Since, in the international market, the number of companies cross-listed in Foreign exchanges has risen, it is essential and important to consider where information is impounded into prices (FRIJNS; GILBERT; TOURANI-RAD, 2010).

The study of *Price Discovery* relies on the implicit assumption that the price differentials between markets are bounded by arbitrage opportunities and, hence, the prices are cointegrated (FRIJNS; GILBERT; TOURANI-RAD, 2010). *Price Discovery* is, therefore, the search for an *Equilibrium Price*; which is in accordance with the definition (SCHREIBER; SCHWARTZ, 1986).

Once verified there is cointegration between two assets, it is conceivable to determine which market influences more the common price. It is important to emphasize that in *Price Discovery* the two prices are known *a priori*, which differs from a classical *Pairs Trading* strategy. The market that influences the common price more is called *Satellite* while the market that corrects its price, in function of the variations in the other market, is called *Peripheral*. It is important to emphasize that: the *Peripheral* market can influence the *Equilibrium Price* but with less influence power than the *Satellite* market (FRIJNS; GILBERT; TOURANI-RAD, 2010).

To determine which market is the *Satellite*, it is necessary to compute which market corrects its price slower to changes of the other market's price. Consequently, influencing more the common price. There are different metrics to evaluate which of the markets is the leading one, such as: the *Information Share* (IS), the *Component Share* (CS) and the *Price Discovery Dynamic Contribution* (PDDC) (HASBROUCK, 1995; GONZALO; GRANGER, 1995; CAPORALE; CIFERRI; GIRARDI, 2010).

Hasbrouck (1995) presented a paper showing the first measure of *Price Discovery*, the *Information Share* (IS), that measures the relative contribution of each market in the common price variance, there are some interesting properties about that metric in continuous time, see (DIAS; FERNANDES; SCHERRER, 2021). In the same sense, the

Component Share (CS) measure also was an innovation in the literature about this topic (GONZALO; GRANGER, 1995)

Using State Space models, along with the *Component Share* metric, Caporale, Ciferri and Girardi (2010) proposed the *Price Discovery Dynamic Contribution* (PDDC), a very attractive metric which evolves over time, allowing one to verify the phenomenon stability. In other words, if the market that leads the phenomenon *always* does this or *in most of the time* does this. It might be the case that *in most of the time* the market that is closer to the asset (the Domestic one) is the one that leads, but for a short period of time, when the markets are in stress, the Foreign market might be the one that leads. In the 2008's crisis, for example, the US market led the process in the Taiwanese-US stock-ADRs market, despite the Taiwanese market being the domestic one (WANG; WU, 2014).

The two markets under analysis not necessarily need to be in different geographic regions (as given in the example above); Patel et al. (2020), for example, apply *Price Discovery* metrics in the stock and derivative market in the United States of America. In the Brazilian market the phenomenon was studied between the spot and future prices of exchange rate, both in the same geographical region (SANTOS; GARCIA; MEDEIROS, 2014). Likewise, the phenomenon is not limited by only one class of asset, being studied in various classes such as: stocks (HARRIS et al., 1995; HASBROUCK, 1995), options (CHAKRAVARTY; GULEN; MAYHEW, 2004), credit spreads (FORTE; PENA, 2009), regional markets (MIZRACH; NEELY, 2008), floor versus electronic markets (MARTENS, 1998) and commodities (XU; RAO, 2018).

When an asset is traded in many markets, a natural question is raised: in what market the new information is absorbed faster? This dissertation main's objective is to verify whether the *Price Discovery* phenomenon exists between the following sources:

a) *Ibovespa* (the index for the Brazilian Stock Market, traded in the Brazilian Market) and the *EWZ* (an ETF that mimics the index for the Brazilian Stock Market, traded in the American Market);

b) S&P500 (the index for the American Stock Market, traded in the American Market) and the *SPXI11* (an ETF that mimics the index for the American Stock Market, traded in the Brazilian market).

If this phenomenon exists it is aimed to determine which market has more influence, whether the Domestic or the Foreign (as a thumb rule we will always name the Domestic Market the one that has the index) and, by doing so, which market absorbs the information quicker.

One of the main contributions is the use of a different metric to compute the phenomenon, the PDDC (which uses the Kalman filter algorithm). By computing this metric, it is possible to evaluate whether the phenomenon presents a static behavior or not. Besides, if there are situations in which the phenomenon switches, from a market leading price to another, one can capture it by using the PDCC. Since the PDCC is more similar to the CS metric, in this research the main focus will be on these two metrics.

This dissertation is divided into four chapters. The first describes the methodology employed, the VECM model and the metrics commonly used. The second regards the data used, whereas the third is devoted to the discussion and results. The last chapter shows the Conclusion.

2 METHODOLOGY

The first section aims to explain the VECM model, which will be extensively used to compute the *Price Discovery* metrics. The second section aims to show the *Price Discovery* metrics and their formulae. Despite the aim of this research being to investigate only the CS and the PDDC, it will be shown the other popular metric (the IS). The Kalman filter algorithm, used for compute the PDDC, is shown in Appendix A.

2.1 Vector Error Correction Model

The majority of the literature review, related to the *Price Discovery* methods, share the same common principle: to identify a common price (SANTOS; GARCIA; MEDEIROS, 2014). Using this principle, the asset prices may diverge in the short-run but will correct themselves in the long-run. In other words, they are cointegrated and share a long-run relation (FRIJNS; GILBERT; TOURANI-RAD, 2010). Cointegration implies that the dynamics of price changes can be described by an Error Correction Model:

$$\Delta Y_t = C + \Phi Y_{t-1} + \sum_{i=1}^{N} A_i \Delta Y_{t-i} + \varepsilon_t, \tag{2.1}$$

where C is a vector of constants and $\Phi = \alpha\beta'$. The α vector measures the speed of adjustment to the error correction term, Y_t is a vector of prices and A_i are matrices of AR coefficients; a key parameter, as stated before, is the α parameter[1]. The market that presents the bigger α, in absolute terms, is the one that corrects most its price and is, therefore, the one that influences least the common price (FRIJNS; GILBERT; TOURANI-RAD, 2010). In this research the α parameter is used directly to compute the CS metric.

Both, the *Information Share* (IS) and the *Component Share* (CS), might be computed using the VECM α parameters and the reduced form errors. These measures are improvements on prior lead-lag methods because they take into account the cointegration structure (PATEL et al., 2020). The structural VMA coefficients are obtained by computing the orthogonalized impulse response functions, from Equation 2.1, and the (contemporaneously uncorrelated) structural VMA errors ($\varepsilon_{i,t}$ to $\varepsilon_{j,t}$) by mapping their relation to the reduced form errors. The permanent price impacts of shocks to

[1] It measures the speed of adjustment of the price of one market in function of the other market

the prices are easily obtained from the structural VMA (PATEL et al., 2020).

$$\Delta y_{i,t} = \sum_{l=0}^{\infty} A_{i,l}\varepsilon_{i,t-l} + \sum_{l=1}^{\infty} A_{j,l}\varepsilon_{j,t-l} \tag{2.2}$$

$$\Delta y_{j,t} = \sum_{l=0}^{\infty} B_{i,l}\varepsilon_{i,t-l} + \sum_{l=0}^{\infty} B_{j,l}\varepsilon_{j,t-l} \tag{2.3}$$

where i stands for the first time series and j for the second.

For example, a unit shock in the $\Delta y_{i,t}$ ($\varepsilon_{i,t} = 1$) has a permanent effect on all variables equal to $\theta_{\varepsilon_i} = \sum_{l=0}^{\infty} A_{i,l}$, the permanent price impact of shocks in $\Delta y_{j,t}$ are given in a similar manner [2] (PATEL et al., 2020). In Hasbrouck's (1995) temporary-permanent decomposition, which is based on Stock and Watson's common trend representation (STOCK; WATSON, 1988), innovations in the permanent component (the efficient price, m_t) are given by

$$\Delta m_t = \theta_{\varepsilon_i}\varepsilon_{i,t} + \theta_{\varepsilon_j}\varepsilon_{j,t}. \tag{2.4}$$

The variance in the efficient price is

$$Var(\Delta m_t) = Var(\theta_{\varepsilon_i}\varepsilon_{i,t} + \theta_{\varepsilon_j}\varepsilon_{j,t}) = \theta_{\varepsilon_i}^2 Var(\varepsilon_{i,t}) + \theta_{\varepsilon_j}^2 Var(\varepsilon_{j,t}). \tag{2.5}$$

From Equation 2.1, The vector Y_t contains the data to compute the metrics wanted in this research. To compute them one needs to decompose the time series in a permanent and transitory components. Several decompositions have been proposed to decompose the series into transitory and permanent components. In general, the decomposition takes the following form:

$$Y_t = (c_\perp(\beta' c_\perp)^{-1}\beta' + \beta_\perp(c'\beta_\perp)^{-1}c)Y_t$$

The existence of this decomposition is not always garanteed, because the matrix $c'\beta_\perp$ might not have rank full. By Gonzalo and Granger (1995), one would use $c = \alpha_\perp$. another possibility, is to forecast each estimated point from the VECM, those values are the permanent component. The permanent component (P_t) can be writen by: $P_t = \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp Y_t$. Since $\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}$ is just a scalar, the part of interest is $\alpha'_\perp Y_t$ (TSAY, 2005). And, hence, the $\alpha_\perp$ is the most important part to compute the phenomenon.

[2] $\theta_{\varepsilon_j} = \sum_{l=1}^{\infty} A_{j,l}$. The permanent price impact is the same for all prices, $\theta_{\varepsilon_i} = \sum_{l=0}^{\infty} A_{i,l} = \sum_{l=0}^{\infty} B_{i,l}$, because permanent impacts are innovations in the efficient value and the efficient value is common to all prices as they refer to the same underlying asset (HASBROUCK, 1995)

2.2 *Price Discovery* Metrics

The *Information Share* (IS) is obtained as the price's contribution to the variance of the efficient price innovation:

$$IS_i = \frac{\theta_{\varepsilon_i}^2 Var(\varepsilon_{i,t})}{Var(\Delta m_t)} \tag{2.6}$$

The IS is not unique, because it depends on the ordering used in the estimation procedure [3]. Applying the standard approach used in the literature, one should estimate the IS for each price series under all time series orders. The final IS measure is the average, between the upper and lower bonds (BAILLIE et al., 2002).

This metric indicates which proportion of the equilibrium price variance is explained by each market; and thus, this metric might be used to determine which market moves first in the *Price Discovery* phenomenon. Since this metric was already used in intraday data (DIAS; FERNANDES; SCHERRER, 2021), my main focus in this paper will be in the CS and the PDDC metrics.

The Gonzalo and Granger (1995) *Component Share* (CS) is obtained by normalizing the permanent price impacts of each price series in the reduced form model

$$CS_i = \frac{\theta_{e_i}}{\theta_{e_i} + \theta_{e_j}} \tag{2.7}$$

The metric CS, in Equation (2.7), can be computed as in Baillie et al. (2002) and Jong (2002).

$$CS_i = \frac{\alpha_i^{\perp}}{\alpha_i^{\perp} + \alpha_j^{\perp}} = \frac{\theta_{e_i}}{\theta_{e_i} + \theta_{e_j}} \tag{2.8}$$

where the vector $(\alpha_i^{\perp}, \alpha_j^{\perp})$ is orthogonal to the vector α, that represents the adjustment speed (in the VECM). Note that the CS_i is simply the $\alpha^{\perp}$ normalized to sum one. To obtain the orthogonalized α vector, such as $\alpha^{\perp} 1_m = 1$, one can use

$$\alpha^{\star\perp} = \frac{1_m - \alpha(\alpha'\alpha)^{-1}\alpha' 1_m}{1_m' 1_m - 1_m'\alpha(\alpha'\alpha)^{-1}\alpha' 1_m} \tag{2.9}$$

Using Equation 2.9 along with Equation 2.8 one finds Equation 2.10 which is good when one deals with more than three times series but with a single stochastic trend (DIAS; FERNANDES; SCHERRER, 2022).

$$CS_i = \alpha_i^{\star\perp}. \tag{2.10}$$

By using this metric it is conceivable to determine which market influences more the common price. Small values of the parameter α indicate more influence in the common price. From this metric it is possible to extract which market leads the phenomenon.

[3] This assumption is implicit in procedures that use *Cholesky* factorization of the covariance matrix of reduced form errors (PATEL et al., 2020).

In terms of comparison one can see that while the *IS* is obtained from the innovations, *CS* is obtained from parameters in the VECM equation (SANTOS; GARCIA; MEDEIROS, 2014). Nevertheless, in the sense that they point to the same results, these two metrics are equivalent (HASBROUCK, 2002; LEHMANN, 2002; BAILLIE et al., 2002), (LEHMANN, 2002).

The Gonzalo-Granger decomposition presents two unwanted characteristics: the first one is related to the permanent component, that will not necessarily follow an I(1) process; and the second one is related to the transitory components, that show correlations among them (HASBROUCK, 2002). As pointed out by Santos, Garcia and Medeiros (2014), this might be the disagreement source of the majority of the results presented by different authors.

The last metric, proposed by Caporale, Ciferri and Girardi (2010), the *Price Discovery Dynamic Contribution* (PDDC), uses the Kalman filter, a recursive process for obtaining filtered state variables based on all available information at time t. The goal of the Kalman filter is to update knowledge of the state variables recursively when a new data point becomes available. Once the new observations are obtained, the Kalman filter can be used to continuously correct the estimate of the state vector (KALMAN, 1960; GE et al., 2019; HARVEY, 1990; TSAY, 2005).

To compute the PDDC, one uses a State-Space representation of the VECM form (Equation (2.1)). The estate space equations are random walks as in Equation (2.11):

$$\alpha_{i,t} = \alpha_{i,t-1} + \varepsilon_{i,t} \qquad \alpha_{j,t} = \alpha_{j,t-1} + \varepsilon_{j,t} \tag{2.11}$$

The PDDC, when there are only two time series, is given by

$$CS_{i,t} = \frac{|\alpha_{j,t}|}{|\alpha_{i,t}| + |\alpha_{j,t}|}, \tag{2.12}$$

where the $\alpha_{i,t}$ is the long-run error correction term from the VECM representation. Notice that Equation 2.12 can be expressed as

$$CS_{i,t} = \frac{\alpha_{j,t}}{\alpha_{j,t} - \alpha_{i,t}}, \qquad CS_{j,t} = -\frac{\alpha_{i,t}}{\alpha_{i,t} - \alpha_{j,t}} \tag{2.13}$$

which, in turn is the normalized $\alpha_t^{\perp}$ (see Equations 2.8, 2.9 and 2.10).

By using the PDDC it is possible to extract the dynamic *Price Discovery* measure over time. Note, also, that the PDDC metric is a dynamic CS metric. Again, using the VECM equation form (Equation 2.1), and allowing for three times series (this is important because in this paper the asset pairs are computed along side the exchange rate, to avoid bias in the VECM computation), one would have:

$$\alpha_{i,t} = \alpha_{i,t-1} + \varepsilon_{i,t}, \qquad \alpha_{j,t} = \alpha_{j,t-1} + \varepsilon_{j,t}, \qquad \alpha_{k,t} = \alpha_{k,t-1} + \varepsilon_{k,t}. \tag{2.14}$$

First of all, in this paper, with the asset pairs and the exchange rate, one has in the first regression three time series (the first pair and the exchange rate). That is, in this case one has three α's and two stochastic trends, the first concerning the assets' stochastic trend and the second the exchange rate. To compute the normalized orthogonal values of $\alpha_t = (\alpha_{i,t}, \alpha_{j,t}, \alpha_{k,t})$, such that the sum is equal to one, one would find the PDDC metric

$$CS_{i,t} = \alpha^{\star\perp}_{i,t} \tag{2.15}$$

which is the dynamic CS metric.

But, since there are two stochastic trends, there exists more than one vector of normalized α, three concerning the first stochastic trend and three concerning the other. Since the primary interest is in the asset's *common* stochastic trend, the $\alpha^{\perp}$ vector are those concerning the *common* trend. The common trend ($\alpha'_{\perp} Y_t$) is such that it cointegrates with the time series of interest (Y_t), to obtain the trend one uses the product between the orthogonalized α's and the time series (TSAY, 2005). The procedure to obtain *Dynamic Component Share* is:

1. Obtain the α's from the Kalman Filter VECM,

2. Find the orthogonalized α's, they are orthogonal to the α's, sum to one and, since in this paper there are more than one stochastic trends, one should use those that are cointegrated with the asset's common trend, to do so one could normalize the first orthogonal alpha to one find the others and then normalize again by the absolute sum of the three.

3 THE DATA

The B3[1] is the only stock exchange in Brazil and the leading one in Latin America (FERNANDES; SCHERRER, 2014). The B3 provides a central clearing for equity, commodities, derivatives, etc. It is among the 10 largest stock exchanges in the planet.

Among many sources to finance a company, it is common to categorize them in *Internal* and *External sources*. A firm can reinvest its profits in its activities (an example of internal source) or raise funds from outside the firm either by taken loans, from financial institutions, or by issuing bonds and commercial papers or by issuing the *Initial Public Offerings* (IPO) on stock exchanges[2]. By using the IPO mechanism, there is the possibility of Domestic companies issuing rights represented by their stocks in Foreign financial markets, which are known as *Depository Receipts* (DR) (SILVEIRA; MACIEL; BALLINI, 2014).

Depository Receipts traded in the US market are called *American Depository Receipts* (ADR) and *Depository Receipts* traded in Brazil are called *Brazilian Depository Receipts* (BDR). This kind of assets allows the investor to diversify its investments among different markets with lower costs, since the transactions occur in the Domestic market. For the case of Brazilian ADRs the mechanism behind it is: the ADRs are backed by brazilian shares, i.e., for each ADR there is a stock traded in the Brazilian market[3]. By the firm's view that is another way to capitalize itself.

An *Exchange Traded Fund* (ETF), as pointed out by Gerber (2011), is a mutual fund that trades on an exchange like a stock. From the stock perspective, ETFs are an exchange listed vehicle; this gives investors trading flexible. In its essence, an ETF follows an index, a good index accurately tracks its market or strategy.

The design of an index depends on its purpose. How to support this purpose or objective leads to the selection of its universe and its construction. What an index selects as its securities composes its universe. Some options include *style-pure* (growth, value, international, domestic, large or small capitalization), a blend, target segments of the market, or market themes. From there, selection criteria for that universe and frequency for rebalancing index gets established. Criteria may include: equal, float, capitalization or modified capitalization weight.

With purpose, securities, criteria, and frequency in place, the index is construc-

[1] *Bolsa Brasil Balcão*
[2] Those are some examples, just to point out.
[3] There are some cases in which one BDR or ADR represents more than one stock, that proportion varies from DR to DR.

ted and calculated. An index provider essentially calculates its index in three steps. First, daily calculation of an index involves updating the price of all component parts. Next, rebalancing (at predetermined times such as monthly, quarterly, or annually) allows the provider to make changes to its index. Finally, adjustments are also made for events such as stock splits, stock dividends, and mergers (GERBER, 2011; HEHN, 2006).

The most familiar type of index is the traditional. The traditional index tries to represent a particular market segment. It may be price, capitalization, equal or modified equal weight. The S&P500, the Dow Jones Industrial Average (DJIA) and the Ibovespa are examples of popularly known indexes. As an example of ETFs that follows an index is the *EWZ*, whose objective is to represent Brazilian Equities, and the *SPXI11*, whose objective is to represent American Equities.

This research focuses on the two pairs of assets. The first is composed by the American Market, S&P500 and the SPXI11, and the second pair by the Brazilian Market, the Ibovespa and the EWZ. The majority of studies in the *Price Discovery* literature use the American market as the Foreign, when there is a market not based in the US; although there are some of authors that use other markets (SU; CHONG, 2007; WANG; WU, 2014; DING et al., 1999; KADAPAKKAM; MISRA; TSE, 2003; LOK; KALEV, 2006). In this research, the words *Foreign* and *Domestic* designate the asset's origin. If the S&P500 has its origin in the US, and the SPXI11 has its origin in Brazil, then the Domestic market is the US and the Foreign is the Brazilian market. The same for the pair EWZ and Ibovespa.

However, since the time series are in a different scale and magnitude one should first normalize the data to view it along side. To do so one uses two different main approachs. The first one is to normalize the data by the first observation of each time series, by doing so every time series starts in one. The second approach uses the min max normalization, by doing so every time series will be between zero and one.

Equation 3.1 shows how to obtain the first normalization method. For every time series *P(t)*, one can normalize it by dividing all its terms by the first observation *P(1)*.

$$p(t) = \frac{P(t)}{P(1)} \tag{3.1}$$

Equation 3.2 shows how to obtain the second normalization method. For every time series *P(t)*, one can rescale it by subtracting the the time series minimum ($P(t)_{min}$) and diving the result by difference between the maximum and its minimum value.

The result will be a rescale time series with range between zero and one.

$$p(t) = \frac{P(t) - P(t)_{min}}{P(t)_{max} - P(t)_{min}} \tag{3.2}$$

3.1 Origins and details

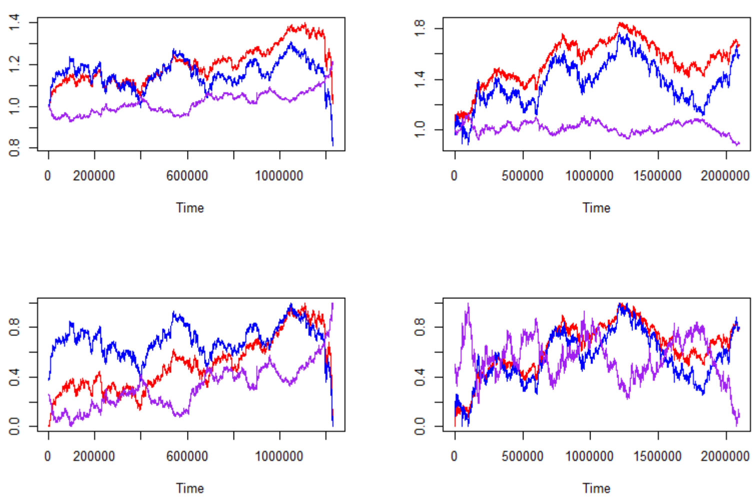
The data used to model the Brazilian assets (the Ibovespa and the EWZ) starts in 27th December 2018 at 10:42 am and ends in 8th April 2022 at 3:59 pm.

Table 1 - Number of observations by data frame

	Brazil		USA	
	Before	After	Before	After
5 Minutes	20517	34863	5873	26743
10 Minutes	10290	17444	4023	16575
15 Minutes	6865	11631	3238	11887

Fonte: Elaborated by the author

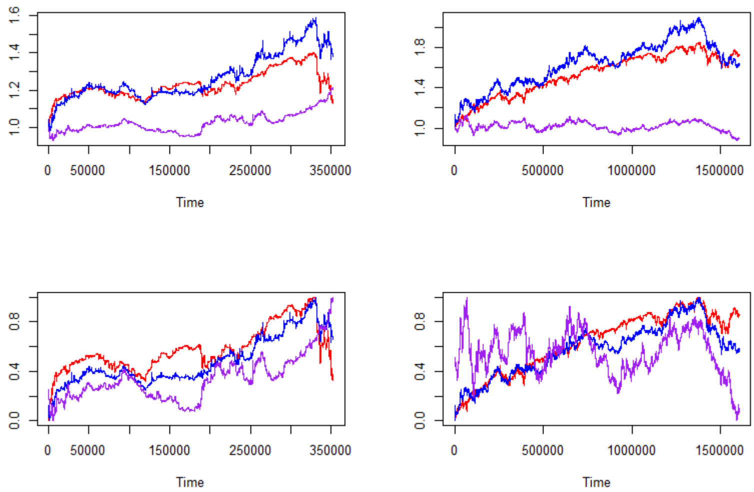
Figure 1 - Ibovespa, EWZ and Exchange rate 5 minutes before and after the covid



Fonte: Elaborated by the author

As can be seen from the figures, there seems to be some common trajectories in the data. To test this one uses the Johanson cointegration test. In Figure 1 the red series is the Ibovespa, the blue is the EWZ and the purple is the Exchange rate. The first line in the Figure represents the before and after the covid, in 5 minutes frequency, with normalization by the first observation (all series start in one). The second line in Figure one represents the same time series but using the MinMax approach. The Figure 2 shows the S&P500 in red, the SPXI11 in blue and the exchange rate in purple.

Figure 2 - S&P500, SPXI11 and Exchange rate 5 minutes before and after the covid



Fonte: Elaborated by the author

The data used to model the American assets (the S&P500 and the SPXI11) starts in 27th December 2018 at 1:02 pm and ends in 8th April 2022 at 4:20 pm. The data is divided in a *Before* Covid data and a *After* Covid data with frequencies of 5, 10 and 15 minutes for each pair of assets, Table 1 displays the observations in each time frame. The before Covid ends in 5th April 2020 at 10 am, the after Covid starts in 9th March 2020 at 4:00pm[4].

The main R packages used were the *tsDyn*, to compute VECM models, *quantmod*, to compute returns, *KFAS*, for the Kalman Filter, *xts*, for data organizing, and *vars*, to determine the lags in VECM. The R version used to compute the results is R version 4.2.0 (2022-04-22 ucrt) – "Vigorous Calisthenics".

The data scripts may be found in: https://github.com/PriceDiscovery/Disserta-o.git

[4] The intraday-data was acquired from the private company Enfoque. https://www.enfoque.com.br/

4 Discussion and Results

The *Component Share* is computed from the VECM structure. In this research, to determine how many lags will be used in the different time frames, the Schwarz information criterion was used. Since the data has different scales, i.e., some time series are in thousands, like the S&P500, and some are in units, like the Exchange rate, transformations were used before the regressions were applied. To verify the cointegration existence, the Johansen test is used. To verify the VECM's residuals the Ljung Box test is used. For the initial values, needed to initiate the Kalman filter, it was used the values of the static form, i.e., the VECM results. And a diffuse matrix for the *Price Discovery* metric.

If there is an asset priced in different currencies, most studies choose to treat the exchange rate as an exogenous variable, converting all prices into a common currency before testing for *Price Discovery*; some authors, for example, convert all the prices used in US dollars (LIEBERMAN; BEN-ZION; HAUSER, 1999). Other authors examine the impact of exchange rate in the phenomenon (GRAMMIG; MELVIN; SCHLAG, 2005)[1]. Treat the exchange rate as an exogenous variable might cause bias in the calculations because the dynamics in the exchange rate might be important to better understand the phenomenon. Indeed the results show that the exchange rate is always significant, in all time frames, to explain the phenomenon. As can be seen from Table 2, the component share for the exchange rate is always high. The results show that the exchange rate has its share of importance in explaining the phenomenon (GRAMMIG; MELVIN; SCHLAG, 2005).

Table 2 - Component Share for the Brazilian assets

	Ibovespa	EWZ	Exchange rate
Before the Covid-19			
5 minutes	3,4%	81,0%	15,7%
10 minutes	5,3%	71,9%	22,8%
15 minutes	7,2%	67,6%	25,2%
After the Covid-19			
5 minutes	0,3%	83,1%	16,6%
10 minutes	0,2%	72,8%	27,0%
15 minutes	8,2%	75,7%	16,1%

Fonte: Elaborated by the author

1 They applied to the German stock market.

Some authors argue that the *Price Discovery* phenomenon, when an asset is trading in different countries, occurs in the Domestic market (where the asset has its origin). In this situation the Foreign market corrects its prices in response to variations in the Domestic market (in short, for these authors, the Foreign market is the *Peripheral* and the Domestic is the *Satellite*). However, there are authors that argue just the opposite[2]: the domestic market leads the *Price Discovery* phenomenon (CAPORALE; GIRARDI, 2013). In short: there is no consensus, or general rule, among the authors. As an example of this ambiguity: in the Brazilian market, using the stocks and *American Depository Receipts* (ADRs) of Vale and Petrobras, the Foreign market lead the process (FERNANDES; SCHERRER, 2014); for the Spanish market, instead, the Foreign market was insignificant (PASCUAL; PASCUAL-FUSTER; CLIMENT, 2006).

Table 3 - Component Share for the American assets

	S&P500	SPXI11	Exchange rate
Before the Covid-19			
5 minutes	68,0%	16,0%	16,0%
10 minutes	70,2%	14,9%	14,9%
15 minutes	73,2%	13,4%	13,4%
After the Covid-19			
5 minutes	99,7%	0,2%	0,2%
10 minutes	94,5%	2,7%	2,7%
15 minutes	86,3%	6,9%	6,9%

Fonte: Elaborated by the author

The *Component Share* results (see Table 2 and Table 3) point in the direction that the American assets (those quoted in the American market) are the assets that guide the process, despite their location.

As for this research results, in the static *Component Share* metric, the foreign market was the *Satellite*, in accordance with Fernandes and Scherrer (2014). Despite the assets not being of the same class (they use Brazilian ADRs). Wang and Wu (2014) points out that most findings show that ADRs prices are more exposed to the home market than the Foreign market and that a few studies indicate that ADRs have a larger co-movement with the US market. The results, Table 2 and 3, show that there are co-moviments (the foreign market is important to explain the phenomenon) but Brazilian assets are more exposed to foreign market. In short: the foreign market

[2] In the derivative market some authors argue that the *future market* is the one that leads the *Price Discovery* phenomenon, that is the natural gas market situation (SCHULTZ; SWIERINGA, 2013). While for other authors there is no evidence that the future market leads(MURAVYEV; PEARSON; BROUSSARD, 2013).

inputs more information in the Brazilian assets. For the American assets it is the opposite: the American market is the one that inputs more information in the prices.

The *Component Share* interpretation for the Brazilian assets (see Table 2), is that the EWZ is the *Satellite* Market, i.e., the market that guides the *Price Discovery* process and, hence, the market that imputes more information in the common price. This happens for the periods before and after the Covid-19 pandemic and in different frequencies homogeneously. It is important to note that the exchange rate also plays an important role in explaining the phenomenon. And, despite not being the leading market, the Ibovespa has some influence in the common price.

The Component Share interpretation for the American market (see Table 3), is that for both periods the S&P500 is the *Satellite* market, i.e., the market that guides the price discovery process and, hence, the market that imputes more information in the common price. As for the exchange rate importance the results are the same. The S&P500 is the *Satellite* market and the SPXI11 is the *Peripheral*.

Table 4 - PDDC for the Brazilian market - before the covid-19 - 5 Minutes

	Ibovespa	EWZ	Exchange Rate
Mean	3.4%	80.9%	15.7%
Sd	0.003%	0.039%	0.04%
Max	3.4%	81.0%	15.9%
Min	3.3%	80.8%	15.6%
Median	3.3%	81.0%	15.7%

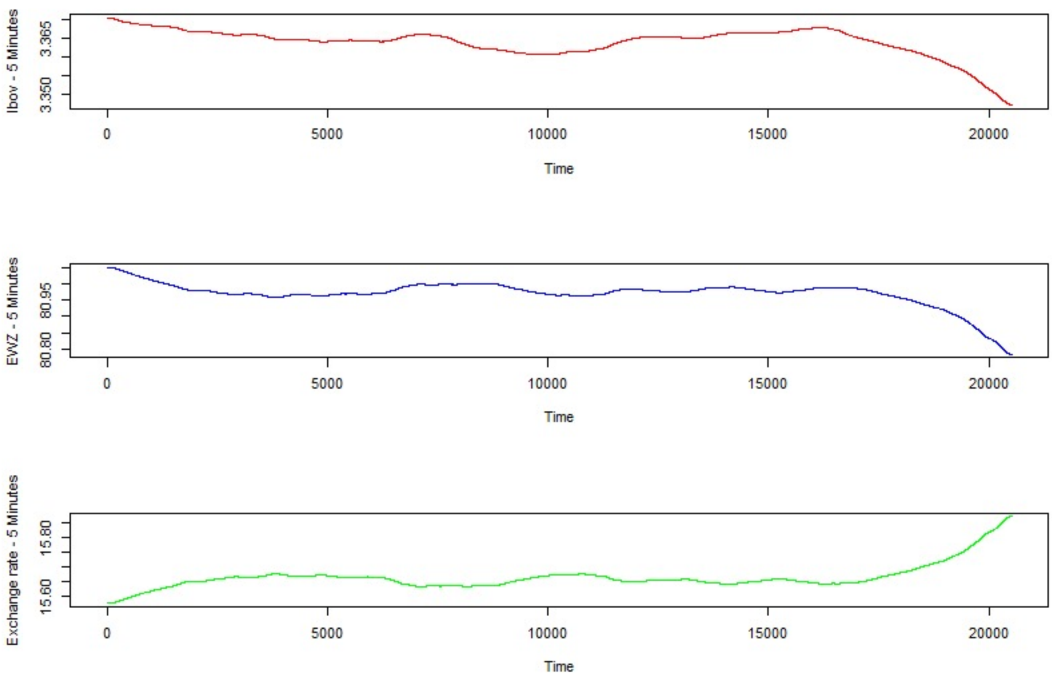
Fonte: Elaborated by the author

For the dynamic *Component Share* (PDDC results), our results confirm the results for the static CS. In other words the market that guides the phenomenon always does this despite the time frame. This research finds that the metric, indeed, varies over time. But its variation, across time, is not enough to change leadership. The cointegration results showing that those $\alpha^{\perp}$ used to compute the metric are right ones may be found the Appendix.

Tables 4 and Figure 3 show the *dynamic component share* for the Brazilian market in the 5 minutes time frame. Indeed the PDDC moves but not enough to change the balance of power. In the figure, the EWZ leads and the Exchange rate plays a major role in explaining the phenomenon.

As for the Brazilian assets in the 10 minutes time frame, see Table 5 and Figure 4, again the results point that the EWZ leads and the Exchange rate plays a major role. Even for PDDC extreme values the balance of power is the same. In the 15 minutes time frame, see Table 6 and Figure 5, the results are the same.

Figure 3 - PDDC for the Brazilian market before the Covid - 5 Minutes (in %)



Fonte: Elaborated by the author

Table 5 - PDDC for the Brazilian market - before the covid-19 - 10 Minutes

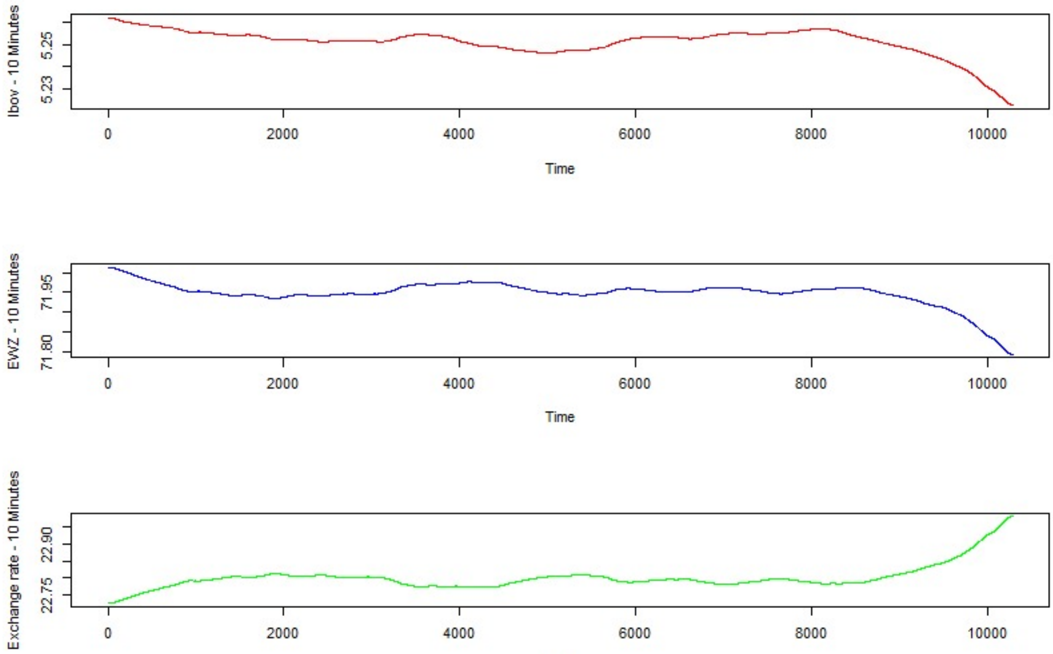
	Ibovespa	EWZ	Exchange Rate
Mean	5.3%	72.0%	23.0%
Sd	0.006%	0.031%	0.037%
Max	5.3%	72.0%	23.0%
Min	5.2%	72.0%	23.0%
Median	5.3%	72.0%	23%

Fonte: Elaborated by the author

About the Brazilian assets before the covid, by Figures 3, 4 and 5, one can see that near the covid beginning the dynamics tend to change. The exchange rate grows in importance to explain the phenomenon and the Ibovespa index falls in terms of dynamic PDDC. As for the EWZ ETF, with the higher frequencies (5, 10 and 15 minutes), the PDDC moves down in terms of explanation power. The results for the Brazilian market before the covid show that the EWZ is the leader and, therefore, the Brazilian market is not a *Satellite* market.

Table 7 show that, in mean, the Ibovespa had less influence in the *equilibrium price* after the covid in the 5 minutes time frame. Again the PDDC for the exchange rate show its influence in the *price equilibrium* (see Figure 6).

Figure 4 - PDDC for the Brazilian market Before the Covid - 10 Minutes (in %)



Fonte: Elaborated by the author

Table 6 - PDDC for the Brazilian market - before the covid-19 - 15 Minutes

	Ibovespa	EWZ	Exchange Rate
Mean	7.17%	67.6%	25.19%
Sd	0.018%	0.040%	0.058%
Max	7.2%	67.73%	25.51%
Min	7.0%	67.41%	25.05%
Median	7.18%	67.63%	25.19%

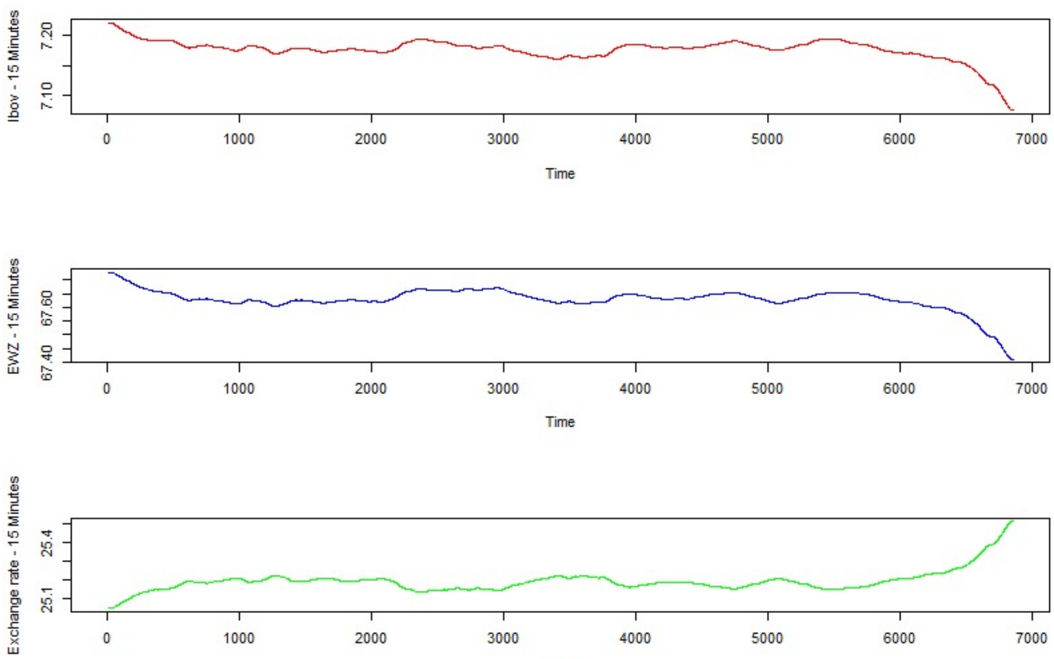
Fonte: Elaborated by the author

The results, in Table 8, 9 and 10, show that the EWZ is the leader (in the 5, 10 and 15 minutes frequencies, see Figure 6, 7 and 8).

For the Brazilian market, the main results are: the exchange rate plays a major role in phenomenon (if one chooses to transform the date in only one currency, indeed, one might have a biased result since the exchange rate's *Component Share* has high values) and the EWZ has more influence in the *equilibrium price*. In short: the Brazilian market is a *Peripheral* market.

Regarding the American market, see Table 10 and Figure 9, the S&P500 index is the most important, between the American and Brazilian assets, to explain the

Figure 5 - PDDC for the Brazilian market Before the Covid - 15 Minutes (in %)



Fonte: Elaborated by the author

Table 7 - PDDC for the Brazilian market - after the covid-19 - 5 Minutes

	Ibovespa	EWZ	Exchange Rate
Mean	0.229%	83.34%	16.4%
Sd	0.001%	0.074%	0.075%
Max	0.232 %	83.4 %	16.62%
Min	0.225%	83.15%	16.33%
Median	0.230%	83.3%	16.41%

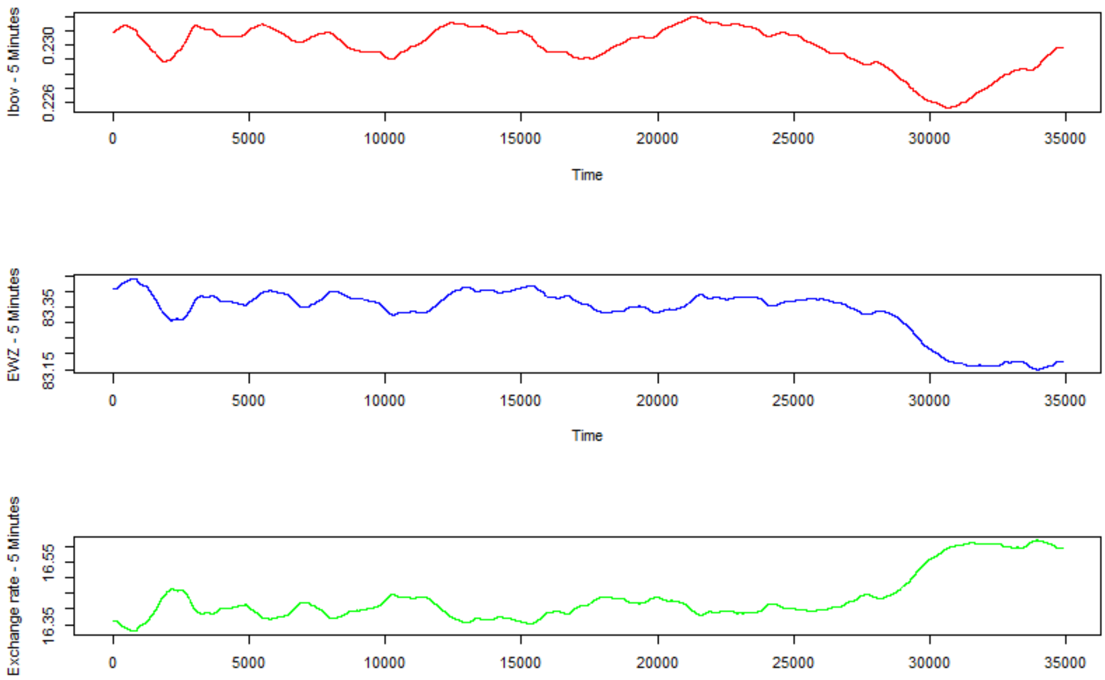
Fonte: Elaborated by the author

phenomenon. The exchange rate again plays an important role in the phenomenon. The American index seems to behave likewise the Brazilian index near the covid beginning, both seem to decreases their share in explaining the phenomenon. The exchange rate seems to differ, its PDDC decreases.

In Table 11 and Figure 8 one can see the PDDC for the 10 minutes time frame. The results are similar to those in the 5 minutes time frame.

As the main results for the pair S&P500 and SPXI11 before the covid, the exchange rate went down. In short: for the American pair, the S&P500 is the asset that

Figure 6 - PDDC for the Brazilian market after the Covid - 5 Minutes (in %)



Fonte: Elaborated by the author

Table 8 - PDDC for the Brazilian market - after the covid-19 - 10 Minutes

	Ibovespa	EWZ	Exchange Rate
Mean	0.237%	72.8%	26.96%
Sd	0.0014%	0.11%	0.111%
Max	0.239%	72.99%	27.14%
Min	0.236%	72.63%	26.77%
Median	0.237%	72.79%	26.97%

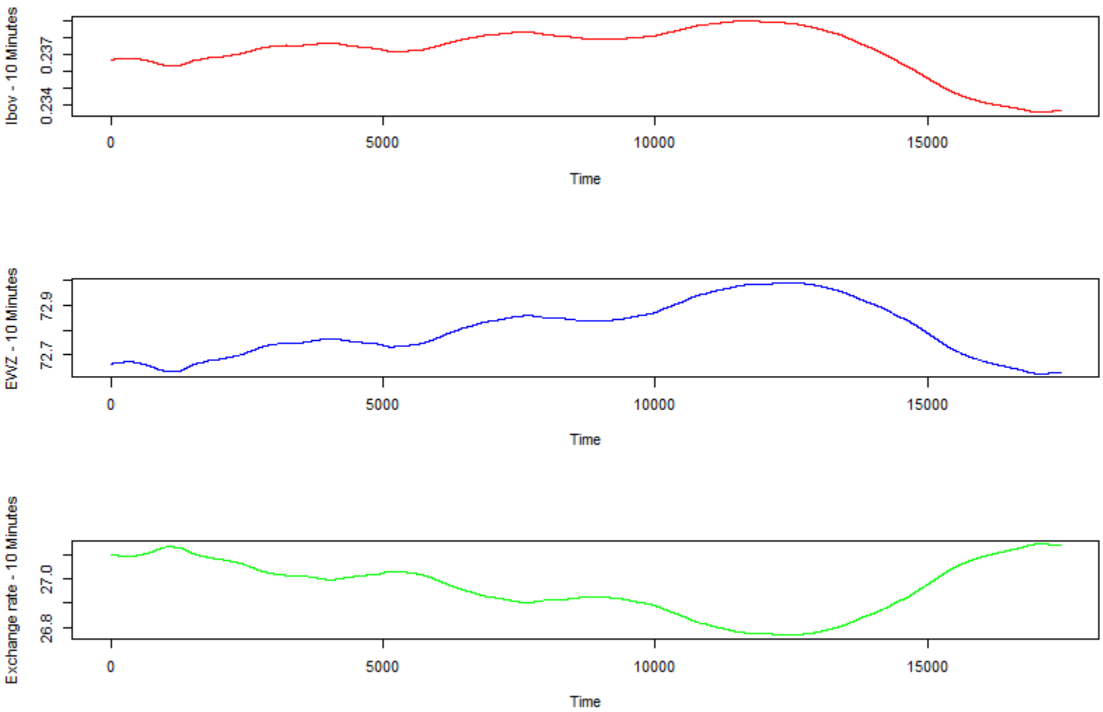
Fonte: Elaborated by the author

impounds more information in the *equilibrium price* and, hence, the American market is the *Satellite* market.

The results for the American pair after the covid are in Table 13 and Figure 9. In the short period after the pandemic shock the PDDC for all the time series were very volatile. Nonetheless, the S&P500 was the leading one in the process.

The last results concerning the American pair in the 10 and 15 minutes frequencies show the same behavior as the one in the 5 minutes frame (in the short period after the covid shock). The SPXI11 show little influence in the common price (see Table 14 and 15 and Figures 13 and 14).

Figure 7 - PDDC for the Brazilian market after the Covid - 10 Minutes (in %)



Fonte: Elaborated by the author

Table 9 - PDDC for the Brazilian market - after the covid-19 - 15 Minutes

	Ibovespa	EWZ	Exchange Rate
Mean	8.15 %	75.73%	16.12%
Sd	0.009%	0.040%	0.048%
Max	8.17 %	75.82 %	16.3%
Min	8.11%	75.53%	16.01%
Median	8.155%	75.73 %	16.11%

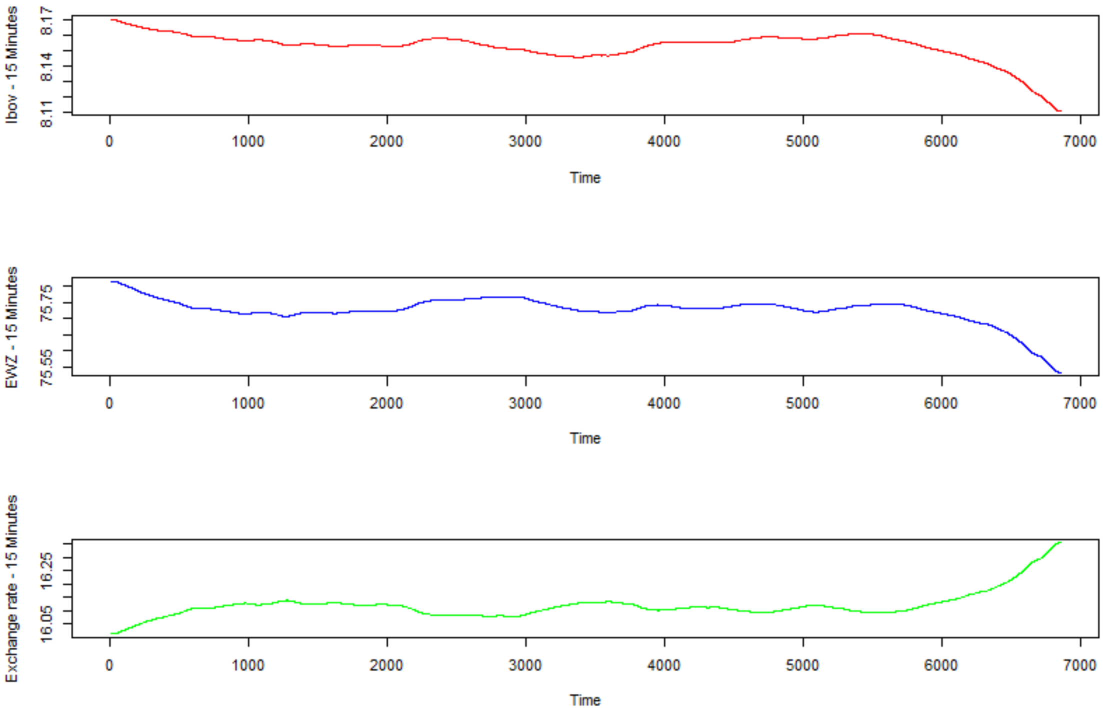
Fonte: Elaborated by the author

Table 10 - PDDC for the American market before the Covid - 5 Minutes

	S&P500	SPXI11	Exchange Rate
Mean	30.22%	12.18%	% 57.59
Sd	0.288%	0.839%	0.55%
Max	31.1%	15.89%	59.26%
Min	28.95%	9.63%	55.16%
Median	30.24%	12.128 %	57.63%

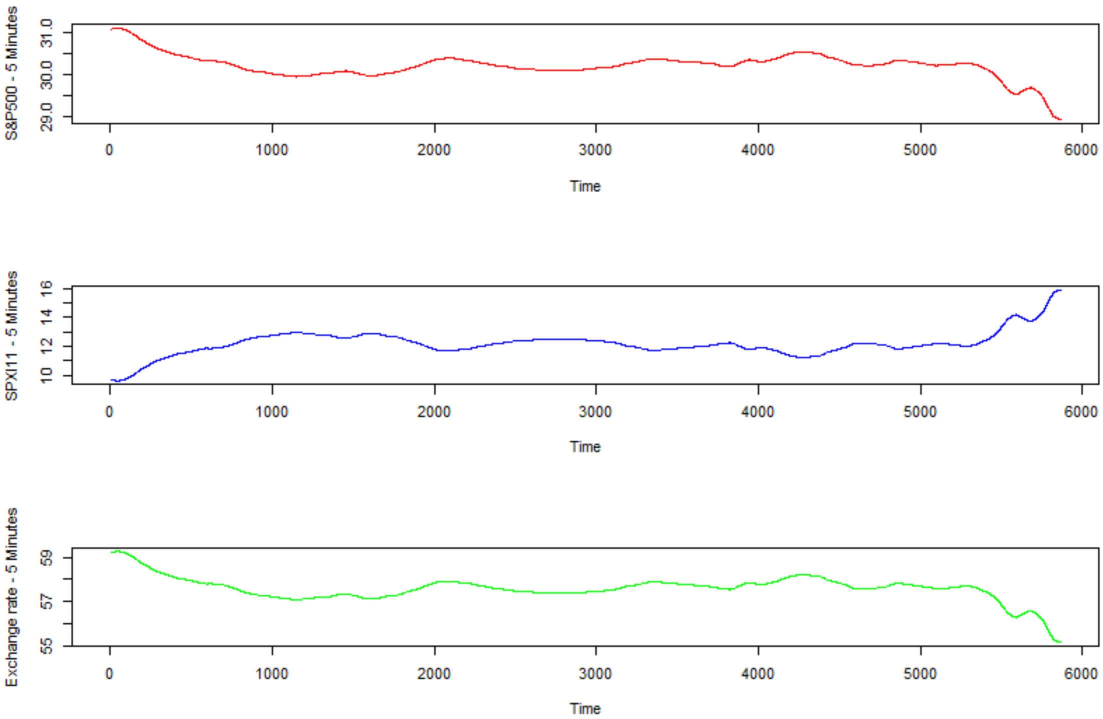
Fonte: Elaborated by the author

Figure 8 - PDDC for the Brazilian market after the Covid - 15 Minutes (in %)



Fonte: Elaborated by the author

Figure 9 - PDDC for the American market before the Covid - 5 Minutes (in %)



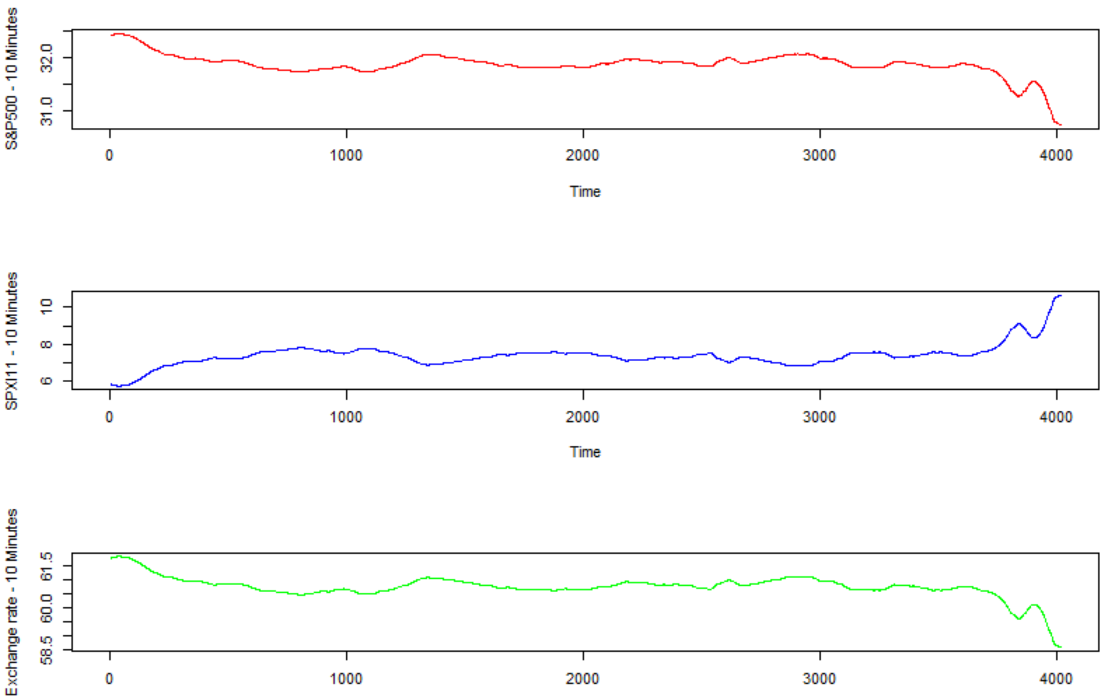
Fonte: Elaborated by the author

Table 11- PDDC for the American market before the Covid - 10 Minutes

	S&P500	SPXI11	Exchange Rate
Mean	31,87%	7.38%	60.74%
Sd	0.205%	0.597%	0.391%
Max	32.4 %	10.65 %	61.82%
Min	30.75%	5.73%	58.6%
Median	31.89%	7.34 %	60.77%

Fonte: Elaborated by the author

Figure 10 - PDDC for the American market before the Covid - 10 Minutes (in %)



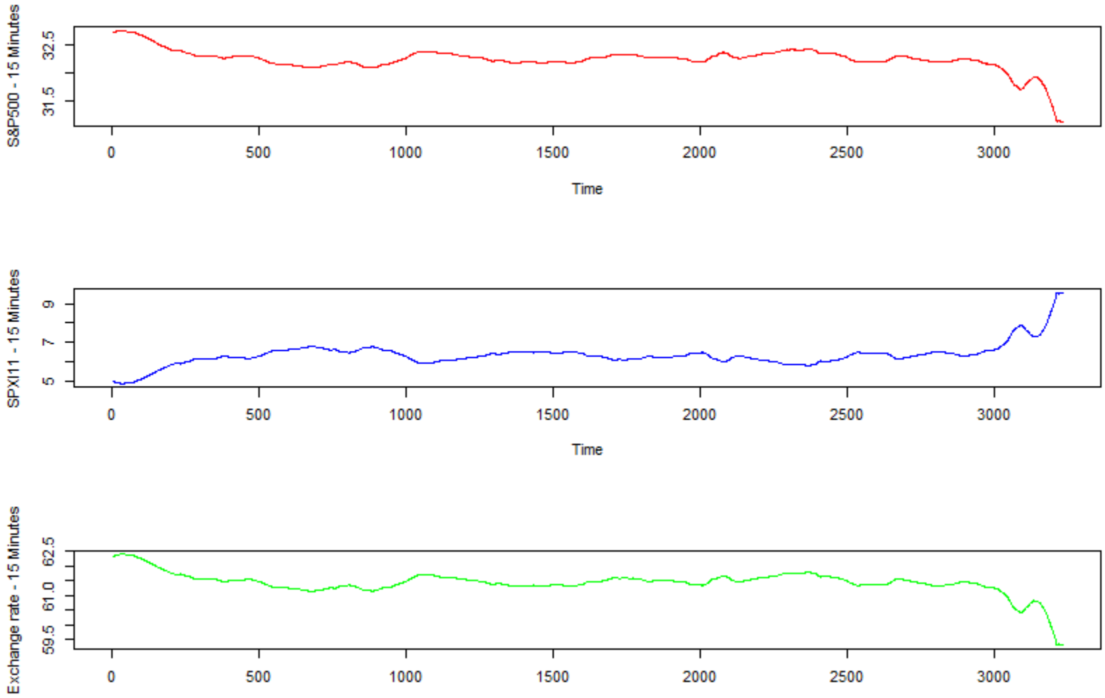
Fonte: Elaborated by the author

Table 12 - PDDC for the American market before the Covid - 15 Minutes

	S&P500	SPXI11	Exchange Rate
Mean	32.24%	6.32%	61.44%
Sd	0.196%	0.570%	0.373%
Max	32.74%	9.566%	62.4%
Min	31.12%	4.855%	59.31%
Median	32.26%	6.274%	61.47%

Fonte: Elaborated by the author

Figure 11 - PDDC for the American market before the Covid - 15 Minutes (in %)



Fonte: Elaborated by the author

Table 13 - PDDC for the American market after the Covid - 5 Minutes

	S&P500	SPXI11	Exchange Rate
Mean	34.1%	0.830%	65.04%
Sd	0,032%	0.093%	0.061%
Max	34.18%	1.34%	65.1%
Min	33.95%	0.676%	64.7%
Median	34.13%	0.82%	65.04%

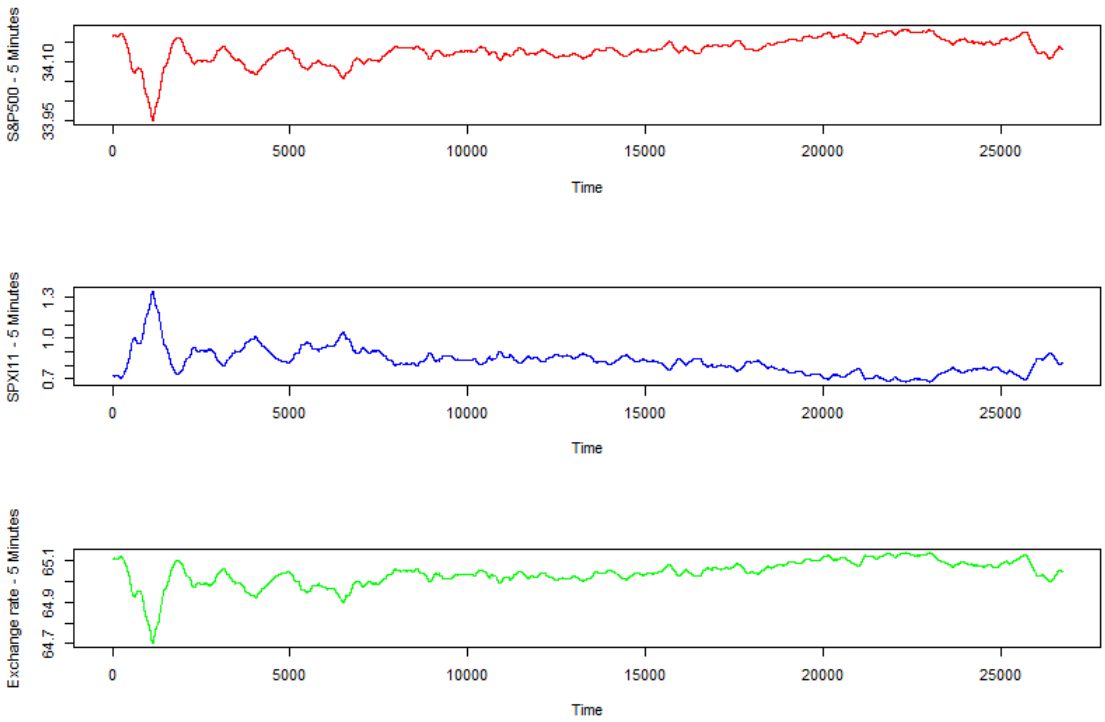
Fonte: Elaborated by the author

Table 14 - PDDC for the American market after the Covid - 10 Minutes

	S&P500	SPXI11	Exchange Rate
Mean	34.29%	0.355%	65.35%
Sd	0.012%	0.037%	0.02428%
Max	34.3%	0.525%	65.4%
Min	34.2%	0.287%	65.2%
Median	34.29%	0.357%	65.3%

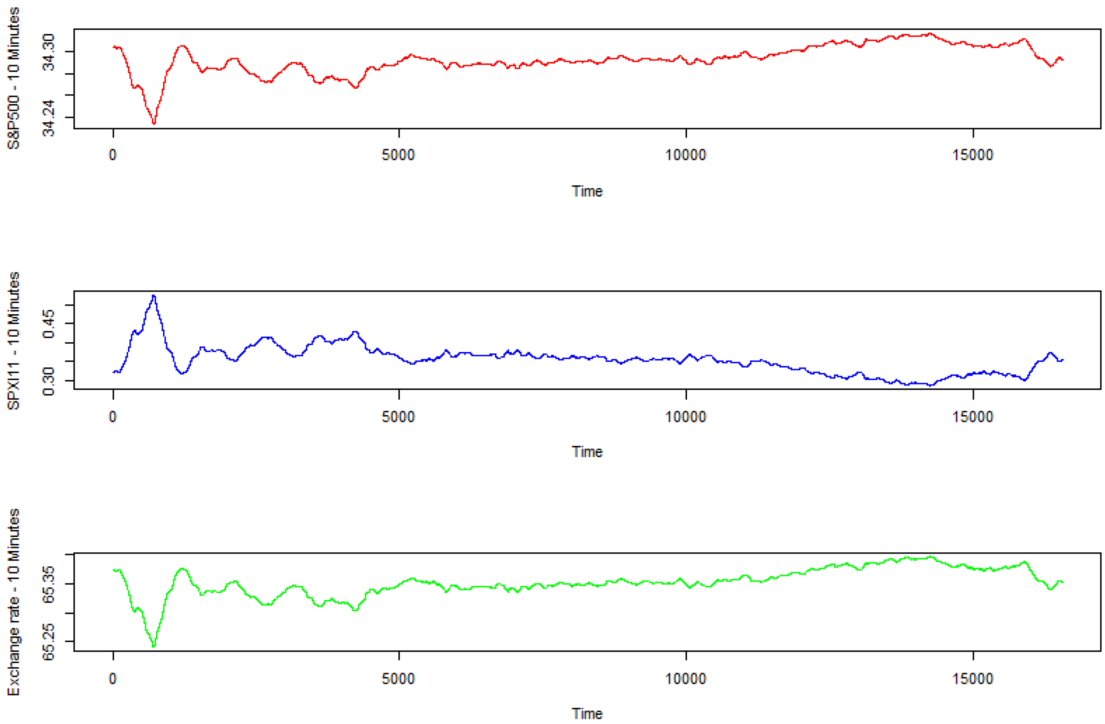
Fonte: Elaborated by the author

Figure 12 - PDDC for the American market after the Covid - 5 Minutes (in %)



Fonte: Elaborated by the author

Figure 13 - PDDC for the American market after the Covid - 10 Minutes (in %)



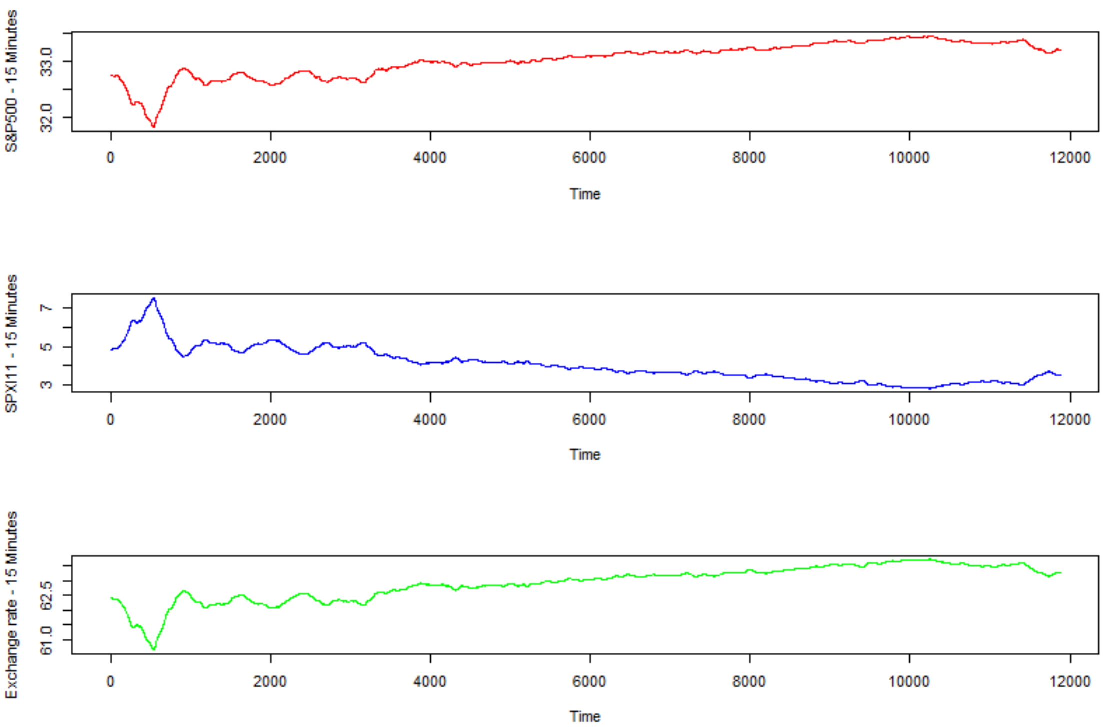
Fonte: Elaborated by the author

Table 15 - PDDC for the American market after the Covid - 15 Minutes

	S&P500	SPXI11	Exchange Rate
Mean	33.02%	4.05%	62.92%
Sd	0.3112%	0.904%	0.593%
Max	33.45%	7.55%	63.74%
Min	31.82%	2.812%	60.63%
Median	33.08%	3.865%	63.05%

Fonte: Elaborated by the author

Figure 14 - PDDC for the American market after the Covid - 15 Minutes (in %)



Fonte: Elaborated by the author

5 CONCLUSION

As the this research main conclusion, it was discovered that indeed the price discovery metric, *Component Share*, does not seem to be always constant over time. Nonetheless, this does not change the main results found when using only a static CS metric. The market that leads the *Price Discovery* phenomenon continues to lead the phenomenon despite the data time frame. Moreover, the market that guides the process is not necessarily the one that hosts the asset. The Exchange rate played a major role in determining the *price equilibrium*, in the most part it did show an important role in terms of PDDC, and CS, metric. Indeed by simply converting the times series to a same currency one might have a biased result.

In the pairs used, the American market assets were leaders in the *Price Discovery* process (*Satellite* market). The EWZ and the S&P500 were the assets that imputed more information to the common price. The SPXI11 and the Ibovespa were the assets (in the *Peripheral* market) that corrects its prices to change in the long run price. These results show that despite the time frame used to evaluate the *Price Discovery Phenomenon* the results does not change dramatically, that the Exchange rate plays a major role in determining the *equilibrium price* and that, at least for the Brazilian market, the American market plays a major role in determining the assets *equilibrium prices*.

References

BAILLIE, R. T. et al. Price discovery and common factor models. *Journal of financial markets*, Elsevier, v. 5, n. 3, p. 309–321, 2002.

CAPORALE, G. M.; CIFERRI, D.; GIRARDI, A. Time-varying spot and futures oil price dynamics. CESifo Working Paper Series, 2010.

CAPORALE, G. M.; GIRARDI, A. Price discovery and trade fragmentation in a multi-market environment: Evidence from the mts system. *Journal of Banking & Finance*, Elsevier, v. 37, n. 2, p. 227–240, 2013.

CHAKRAVARTY, S.; GULEN, H.; MAYHEW, S. Informed trading in stock and option markets. *The Journal of Finance*, Wiley Online Library, v. 59, n. 3, p. 1235–1257, 2004.

DIAS, G. F.; FERNANDES, M.; SCHERRER, C. M. Price discovery in a continuous-time setting. *Journal of Financial Econometrics*, Oxford University Press, v. 19, n. 5, p. 985–1008, 2021.

DIAS, G. F.; FERNANDES, M.; SCHERRER, C. M. Time-varying price discovery. *Journal of Financial Econometrics*, Oxford University Press, 2022.

DING, D. K. et al. An investigation of price discovery in informationally-linked markets: equity trading in malaysia and singapore. *Journal of Multinational Financial Management*, Elsevier, v. 9, n. 3-4, p. 317–329, 1999.

EUN, C. S.; SABHERWAL, S. Cross-border listings and price discovery: Evidence from us-listed canadian stocks. *The Journal of Finance*, Wiley Online Library, v. 58, n. 2, p. 549–575, 2003.

FERNANDES, M.; SCHERRER, C. *Price discovery on common and preferred shares across multiple markets*. [S.l.], 2014.

FOERSTER, S. R.; KAROLYI, G. A. International listings of stocks: The case of canada and the us. *Journal of International Business Studies*, Springer, v. 24, n. 4, p. 763–784, 1993.

FOERSTER, S. R.; KAROLYI, G. A. Multimarket trading and liquidity: a transaction data analysis of canada–us interlistings. *Journal of International Financial Markets, Institutions and Money*, Elsevier, v. 8, n. 3-4, p. 393–412, 1998.

FOERSTER, S. R.; KAROLYI, G. A. The effects of market segmentation and investor recognition on asset prices: Evidence from foreign stocks listing in the united states. *The Journal of Finance*, Wiley Online Library, v. 54, n. 3, p. 981–1013, 1999.

FORTE, S.; PENA, J. I. Credit spreads: An empirical analysis on the informational content of stocks, bonds, and cds. *Journal of Banking & Finance*, Elsevier, v. 33, n. 11, p. 2013–2025, 2009.

FRIJNS, B.; GILBERT, A.; TOURANI-RAD, A. The dynamics of price discovery for cross-listed shares: Evidence from australia and new zealand. *Journal of banking & finance*, Elsevier, v. 34, n. 3, p. 498–508, 2010.

GE, Y. et al. Does china's iron ore futures market have price discovery function? analysis based on vecm and state-space perspective. *Journal of Business Economics and Management*, v. 20, n. 6, p. 1083–1101, 2019.

GERBER, M. *How to create and manage a mutual fund or exchange-traded fund: a professional's guide*. [S.l.]: John Wiley & Sons, 2011. v. 404.

GONZALO, J.; GRANGER, C. Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 13, n. 1, p. 27–35, 1995.

GRAMMIG, J.; MELVIN, M.; SCHLAG, C. Internationally cross-listed stock prices during overlapping trading hours: price discovery and exchange rate effects. *Journal of Empirical Finance*, Elsevier, v. 12, n. 1, p. 139–164, 2005.

HAN, J. *Time-variant CAPM: learning about factor loading*. [S.l.], 2006.

HARRIS, F. H. d. et al. Cointegration, error correction, and price discovery on informationally linked security markets. *Journal of financial and quantitative analysis*, Cambridge University Press, v. 30, n. 4, p. 563–579, 1995.

HARVEY, A. C. Forecasting, structural time series models and the kalman filter. Cambridge university press, 1990.

HASBROUCK, J. One security, many markets: Determining the contributions to price discovery. *The journal of Finance*, Wiley Online Library, v. 50, n. 4, p. 1175–1199, 1995.

HASBROUCK, J. Stalking the "efficient price" in market microstructure specifications: an overview. *Journal of Financial Markets*, Elsevier, v. 5, n. 3, p. 329–339, 2002.

HEHN, E. *Exchange Traded Funds: Structure, regulation and application of a new fund class*. [S.l.]: Springer Science & Business Media, 2006.

JONG, F. D. Measures of contributions to price discovery: A comparison. *Journal of Financial markets*, Elsevier, v. 5, n. 3, p. 323–327, 2002.

KADAPAKKAM, P.-R.; MISRA, L.; TSE, Y. International price discovery for emerging market stocks: evidence from indian gdrs. *Review of Quantitative Finance and Accounting*, Springer, v. 21, n. 2, p. 179–199, 2003.

KALMAN, R. E. A new approach to linear filtering and prediction problems. 1960.

LEHMANN, B. N. Some desiderata for the measurement of price discovery across markets. *Journal of Financial Markets*, Elsevier, v. 5, n. 3, p. 259–276, 2002.

LIEBERMAN, O.; BEN-ZION, U.; HAUSER, S. A characterization of the price behavior of international dual stocks: an error correction approach. *Journal of International Money and Finance*, Elsevier, v. 18, n. 2, p. 289–304, 1999.

LOK, E.; KALEV, P. S. The intraday price behaviour of australian and new zealand cross-listed stocks. *International Review of Financial Analysis*, Elsevier, v. 15, n. 4-5, p. 377–397, 2006.

MARTENS, M. Price discovery in high and low volatility periods: open outcry versus electronic trading. *Journal of International Financial Markets, Institutions and Money*, Elsevier, v. 8, n. 3-4, p. 243–260, 1998.

MIZRACH, B.; NEELY, C. J. Information shares in the us treasury market. *Journal of Banking & Finance*, Elsevier, v. 32, n. 7, p. 1221–1233, 2008.

MURAVYEV, D.; PEARSON, N. D.; BROUSSARD, J. P. Is there price discovery in equity options? *Journal of Financial Economics*, Elsevier, v. 107, n. 2, p. 259–283, 2013.

PASCUAL, R.; PASCUAL-FUSTER, B.; CLIMENT, F. Cross-listing, price discovery and the informativeness of the trading process. *Journal of Financial Markets*, Elsevier, v. 9, n. 2, p. 144–161, 2006.

PATEL, V. et al. Price discovery in stock and options markets. *Journal of Financial Markets*, Elsevier, v. 47, p. 100524, 2020.

SANTOS, F. E. d. L. A.; GARCIA, M. G. P.; MEDEIROS, M. C. *Price discovery no mercado de câmbio Brasileiro: O preço é formado no mercado à vista ou futuro?* [S.l.], 2014.

SCHREIBER, P. S.; SCHWARTZ, R. A. Price discovery in securities markets. *The Journal of Portfolio Management*, Institutional Investor Journals Umbrella, v. 12, n. 4, p. 43–48, 1986.

SCHULTZ, E.; SWIERINGA, J. Price discovery in european natural gas markets. *Energy Policy*, Elsevier, v. 61, p. 628–634, 2013.

SHUMWAY, R. H.; STOFFER, D. S.; STOFFER, D. S. *Time series analysis and its applications*. [S.l.]: Springer, 2000. v. 3.

SILVEIRA, R. L. F. D.; MACIEL, L.; BALLINI, R. Cointegration and causality-in-mean and variance tests: Evidence of price discovery for brazilian cross-listed stocks. *Revista Brasileira de Economia de Empresas/Brazilian Journal of Business Economics*, v. 14, n. 1, 2014.

STOCK, J. H.; WATSON, M. W. Testing for common trends. *Journal of the American statistical Association*, Taylor & Francis, v. 83, n. 404, p. 1097–1107, 1988.

SU, Q.; CHONG, T. T.-L. Determining the contributions to price discovery for chinese cross-listed stocks. *Pacific-Basin Finance Journal*, Elsevier, v. 15, n. 2, p. 140–153, 2007.

TSAY, R. S. *Analysis of financial time series*. [S.l.]: John wiley & sons, 2005.

WANG, M.-C.; WU, Y.-C. Where does price discovery occur? an empirical study of taiwan's adrs and their underlying foreign stocks. *International Journal of Financial Research*, Citeseer, v. 5, n. 3, p. 43, 2014.

XU, C.; RAO, S. Discovery efficiency and dynamic changes of prices in china's iron ore futures market. *Journal of Jiangxi University of Finance and Economics*, v. 1, p. 20–29, 2018.

Appendix A - The Kalman Filter algorithm

Dynamic Linear Models (DLM) or *State Space Models* (SSM) where originally invented to describe the behavior of non-observed variables that are not constant over time. Those variables only reveal itself through the realizations of other observed variables. A system of equations for those two types of variables (observed and not-observed) is composed in two parts: the first one is called *State Equation* - that describes the transition process between $t = 0$ and $t = 1$ -, and the second one is called *Observation Equation*, that describes how do state variables relate to observed variables in time t (HAN, 2006; TSAY, 2005).

The *State Equation* can be described by

$$x_t = \Phi x_{t-1} + \omega_t, \tag{1}$$

where $\omega_t \sim$ N(0,Q), x_t is a vector $p \times 1$ of state variables and Φ is a non-time varying transition matrix.

The *Observation Equation* can be described by

$$y_t = H_t x_t + \nu_t, \tag{2}$$

where $\nu_t \sim$ N(0,R) and H_t is the matrix of observations.

Given the realizations $(y_1, y_2, \cdots, y_T)$, and using the filter, the state equation parameters can be found. The solution for the problem can be understood as a recursive process. Given some initials configurations (initial "guesses"for the parameters) in $t = 0$ the filtering process can be initiated. Using these initials values, in $t = 0$ predictions are made for $t = 1$ with the aim to obtain an *a priori* estimate of the state. While new information (observations) are being added to the model, estimates is being updated creating an *a posteriori* estimate.

The main objective in the filtering process is to find the estimates Φ, Q e R. From these parameters one can compute $x_1, x_2, \cdots\ x_T$.

There is not a closed formula to determine those parameters; it is required thus to resort in algorithms (SHUMWAY; STOFFER; STOFFER, 2000). The procedure is presented below:

- Choose initial values for the parameters ("guesses") $\Theta^0 = \{\Phi^0, Q^0, R^0\}$;

- Apply the filter using the initial parameters. Compute the errors $\varepsilon_1, \cdots, \varepsilon_T$ and the variances Σ;
- Using the past innovations apply numeric methods to find a estimate Θ^1 that maximize the maximum likelihood function (see Equation (3));
- Repeat the procedure until the parameters show some stability (here is necessary to define a stop criterion);

The problem maximum likelihood function is given for

$$-2\ln L_y(\Theta) = \sum_{t=1}^{T} \ln|\Sigma_t(\Theta)| + \sum_{t=1}^{T} \varepsilon_t(\Theta)'\Sigma_t(\Theta)^{-1}\varepsilon_t(\Theta). \tag{3}$$

If the initial values are not known it is advisable the use of a *diffuse matrix* (a matrix with high variance for the initial values).

The State Equation used in the research is

$$\begin{bmatrix} \alpha_{i,t} \\ \alpha_{j,t} \\ \alpha_{k,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_{i,t-1} \\ \alpha_{j,t-1} \\ \alpha_{k,t-1} \end{bmatrix} + \begin{bmatrix} \xi_t \\ v_t \\ \eta_t \end{bmatrix} \tag{4}$$

where $\xi_t \sim \mathrm{N}(0,\sigma^2_\xi)$, $v_t \sim \mathrm{N}(0,\sigma^2_v)$ and $\eta_t \sim \mathrm{N}(0,\sigma^2_\eta)$.

The only difference between the equation presented below, Equation 5, and the Equation 2.1 is the α parameter, not assumed constant anymore. The Observations Equation is

$$\Delta Y_t = C + \Phi_t Y_{t-1} + \sum_{i=1}^{N} A_i \Delta Y_{t-i} + \chi_t, \tag{5}$$

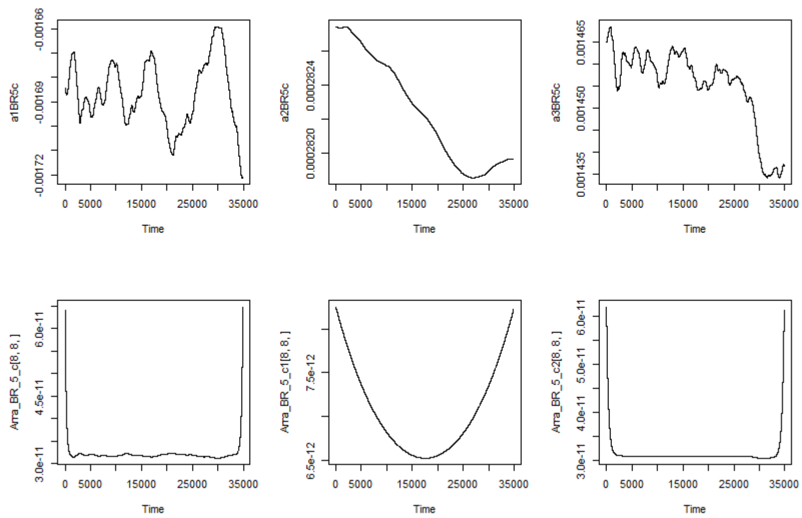
where $\chi_t \sim \mathrm{N}(0,\sigma^2_\chi)$, $\Phi_t = \alpha_t \beta'$ and $\alpha_t = \left[\alpha_{i,t}, \alpha_{j,t}, \alpha_{k,t}\right]$.

By using this configuration, it is possible to compute a dynamic measure of *Price Discovery*. If this measure is constant over time, the variance converges to zero; and, hence, the market that leads always leads the phenomenon.

Appendix B - Dynamic results

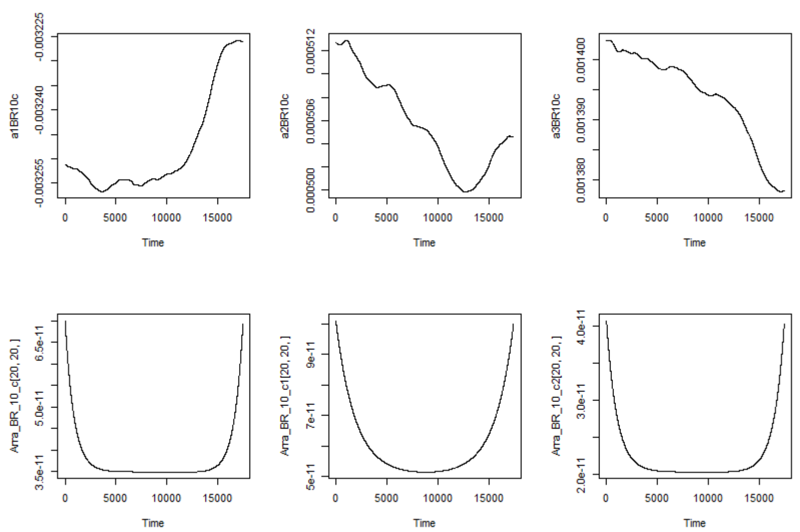
In the appendix chapter one can find the α's used to compute the PDDC. More details may be found the code script. The first line is always concerning the smoothed α the second line concerning it's standard deviations.

Figure B1 - Ibovespa, EWZ and Exchange rate 5 minutes before the covid α's



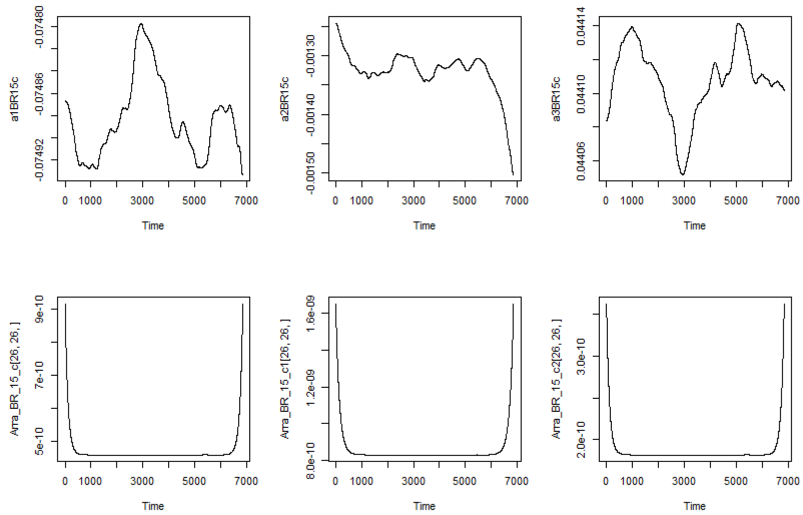
Fonte: Elaborated by the author

Figure B2 - Ibovespa, EWZ and Exchange rate 10 minutes before the covid α's



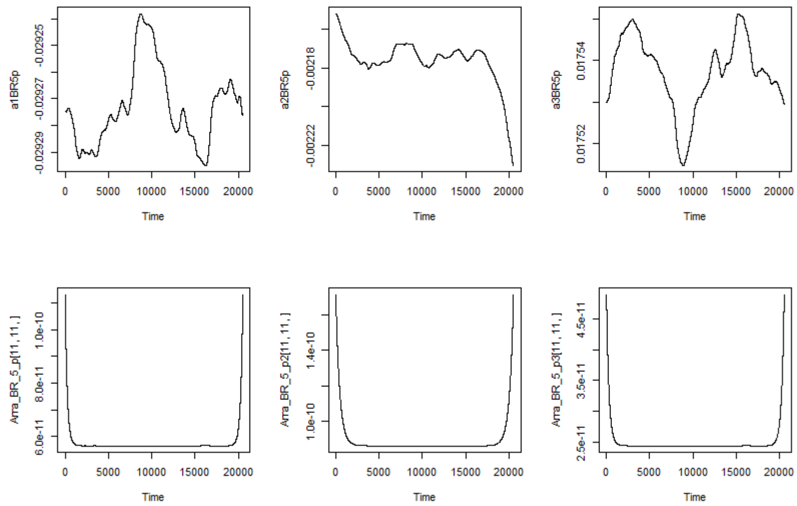
Fonte: Elaborated by the author

Figure B3 - Ibovespa, EWZ and Exchange rate 15 minutes before the covid α's



Fonte: Elaborated by the author

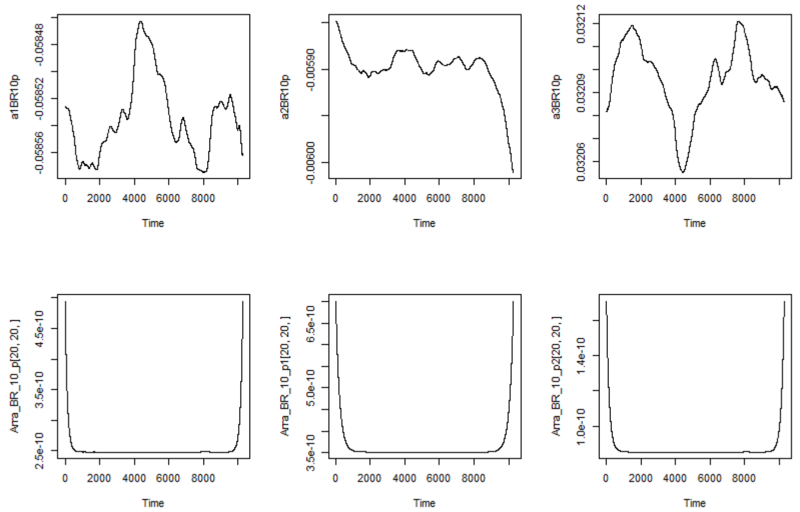
Figure B4 - Ibovespa, EWZ and Exchange rate 5 minutes after the covid α's



Fonte: Elaborated by the author

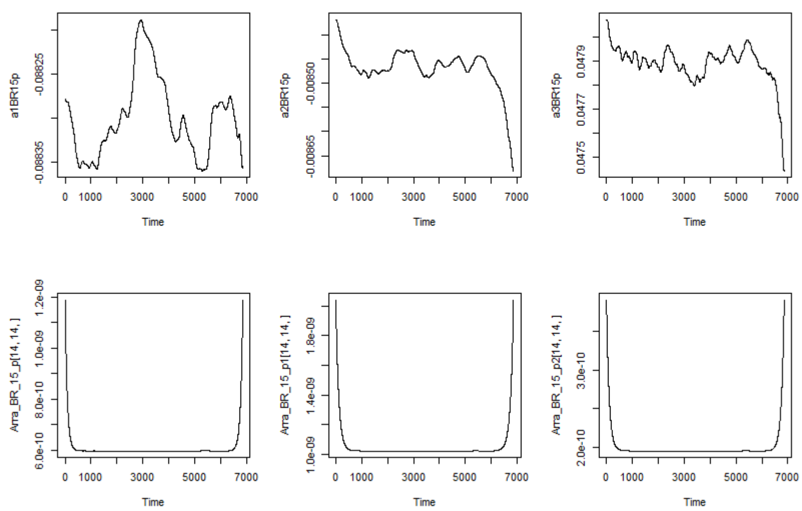
From Figures B7 to the end are the results concerning the US market.

Figure B5- Ibovespa, EWZ and Exchange rate 10 minutes after the covid α's



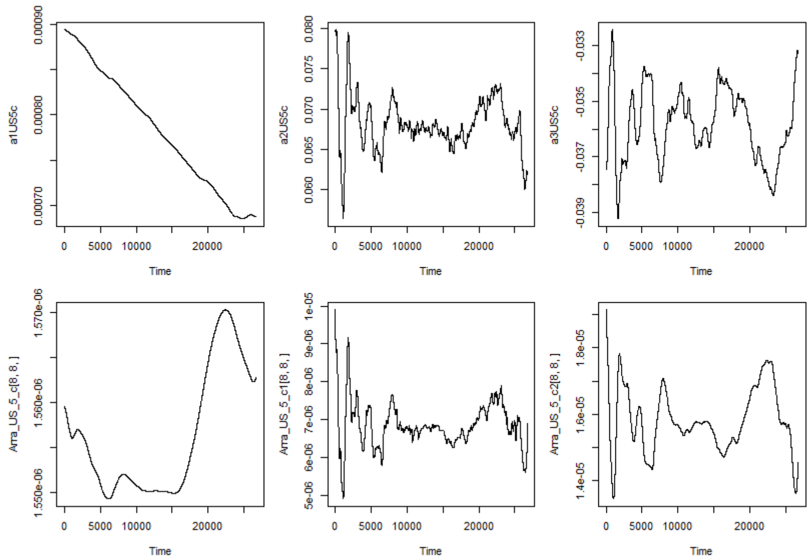
Fonte: Elaborated by the author

Figure B6 - Ibovespa, EWZ and Exchange rate 15 minutes after the covid α's



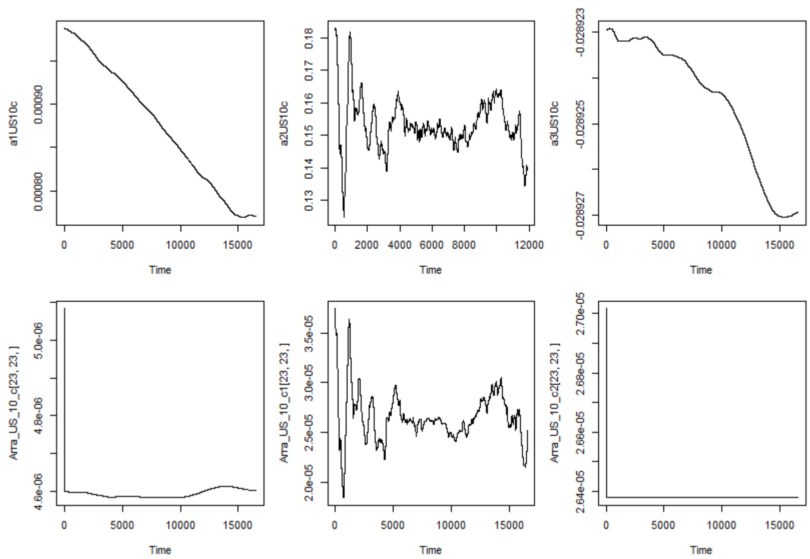
Fonte: Elaborated by the author

Figure B7 - S&P500, SPXI11 and Exchange rate 5 minutes before the covid α's



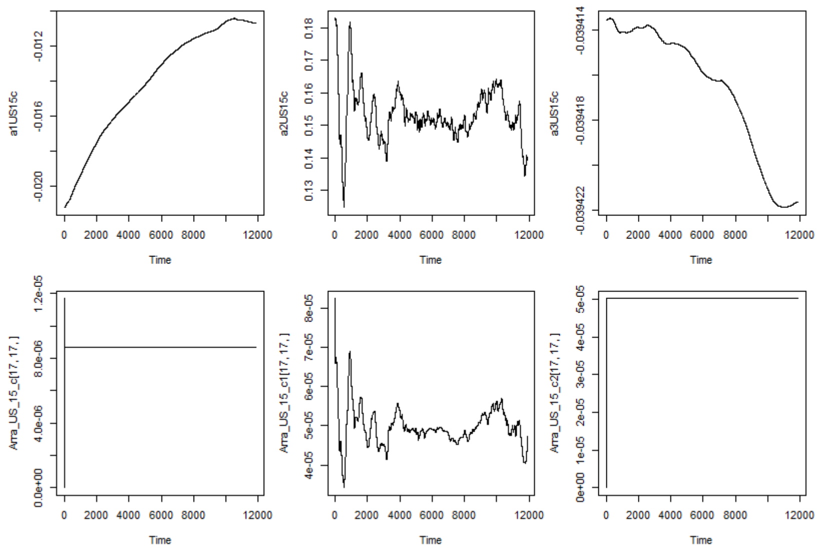
Fonte: Elaborated by the author

Figure B8 - S&P500, SPXI11 and Exchange rate 10 minutes before the covid α's



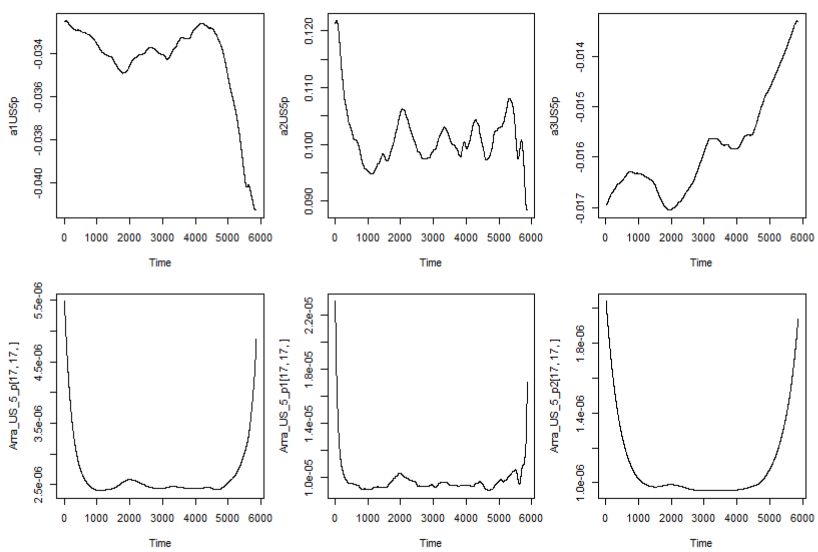
Fonte: Elaborated by the author

Figure B9 - S&P500, SPXI11 and Exchange rate 15 minutes before the covid α's



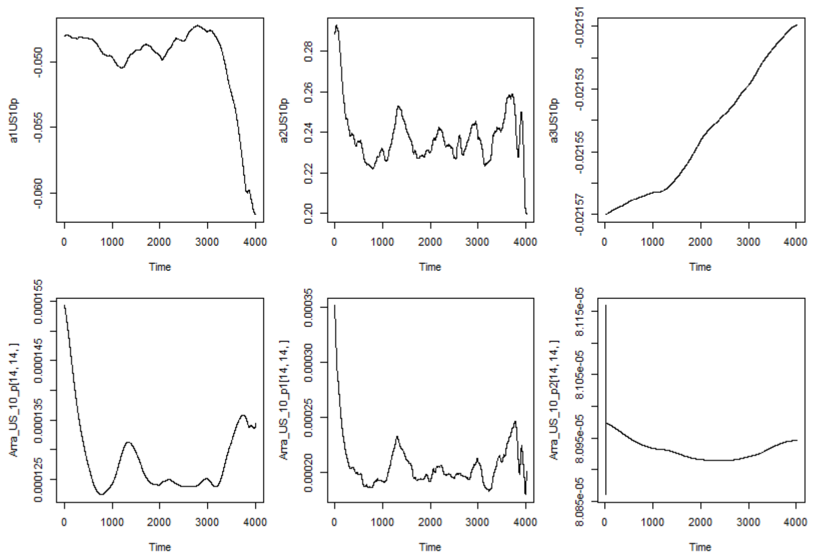
Fonte: Elaborated by the author

Figure B10 - S&P500, SPXI11 and Exchange rate 5 minutes after the covid α's



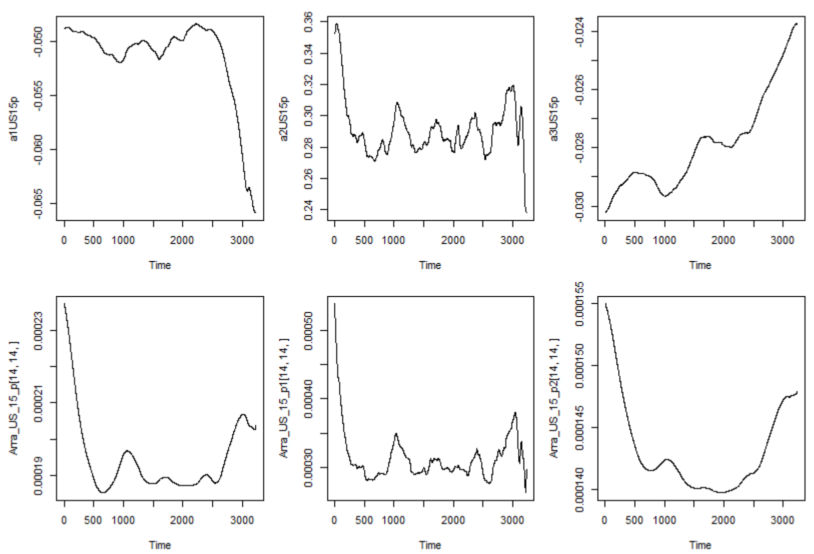
Fonte: Elaborated by the author

Figure B8 - S&P500, SPXI11 and Exchange rate 10 minutes after the covid α's



Fonte: Elaborated by the author

Figure B9 - S&P500, SPXI11 and Exchange rate 15 minutes after the covid α's

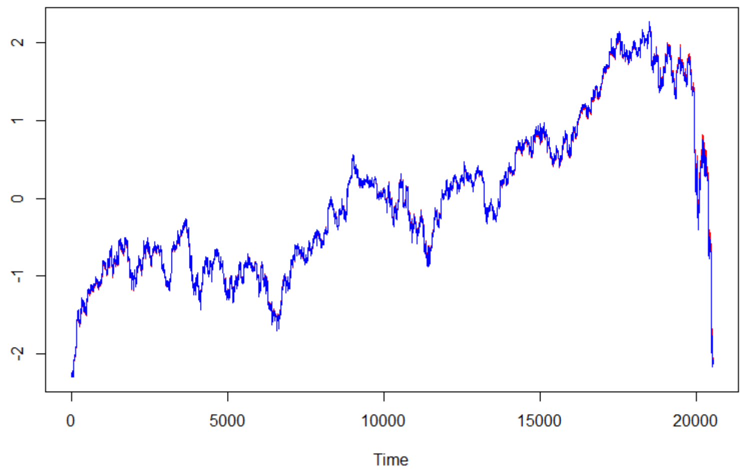


Fonte: Elaborated by the author

Appendix C - Stochastic trends and the data

In this appendix chapter one can find the stochastic trend used along side the data. One uses the $\alpha^{\perp}Y_t$ to find the stochastic trend (TSAY, 2005). In red there is the asset's time series (Ibovespa for Brazil and S&P500 for the American market), in blue one can find $\alpha^{\perp}Y_t$ concerningt the asset.

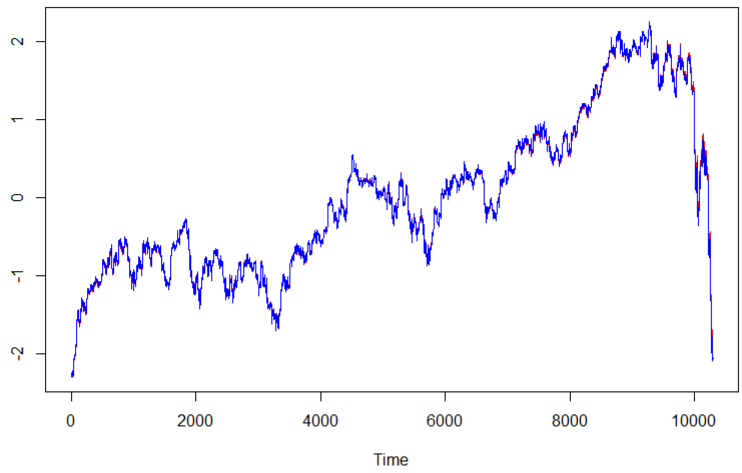
Figure C1 - Ibovespa 5 minutes before the covid



Fonte: Elaborated by the author

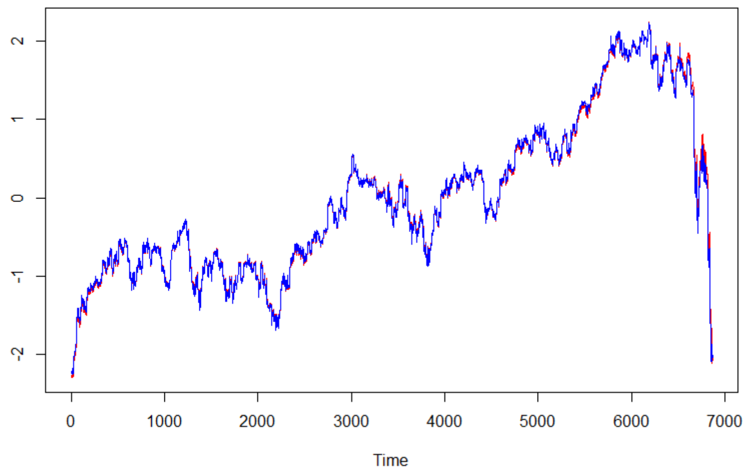
From Figures C7 to the end are the results concerning the US market.

Figure C2 - Ibovespa 10 minutes before the covid



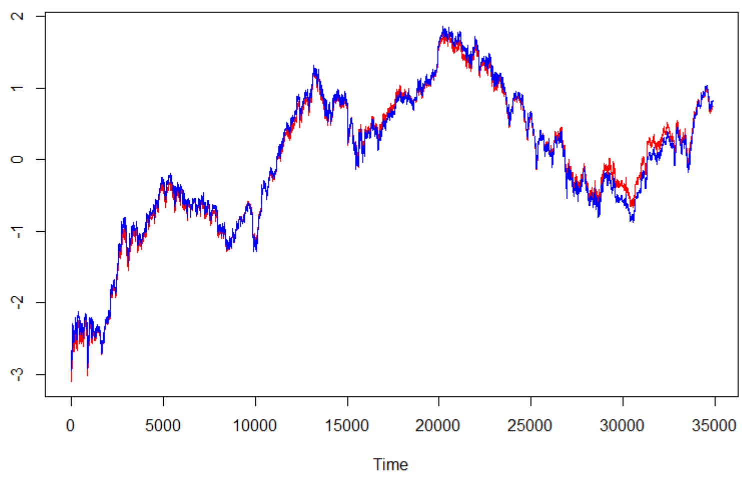
Fonte: Elaborated by the author

Figure C3 - Ibovespa 15 minutes before the covid



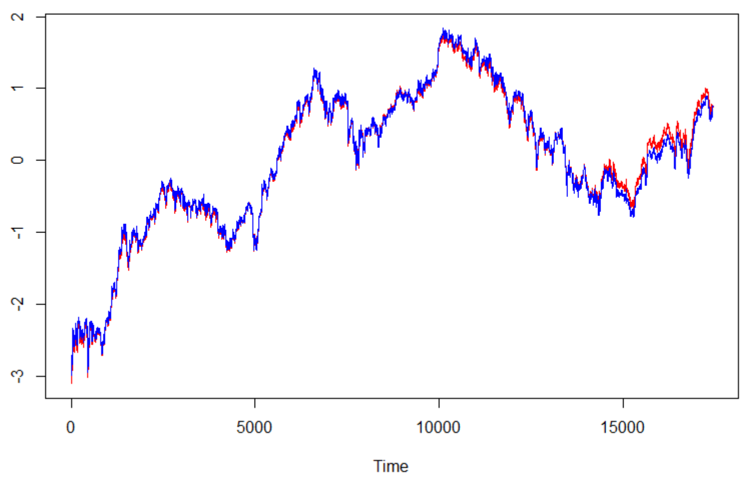
Fonte: Elaborated by the author

Figure C4 - Ibovespa 5 minutes after the covid



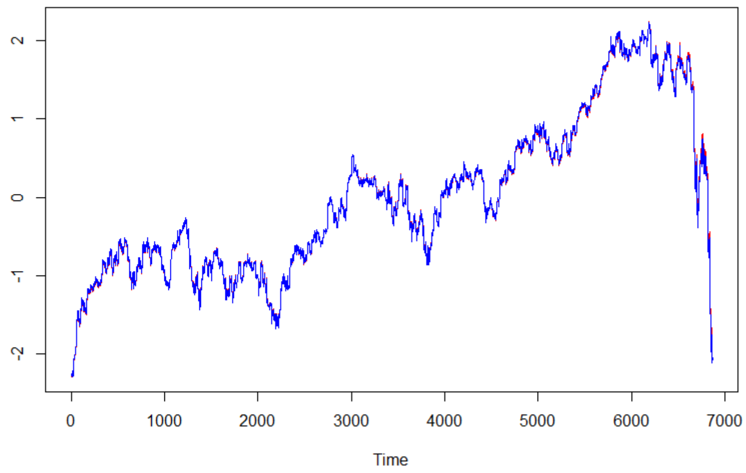
Fonte: Elaborated by the author

Figure C5 - Ibovespa 10 minutes after the covid



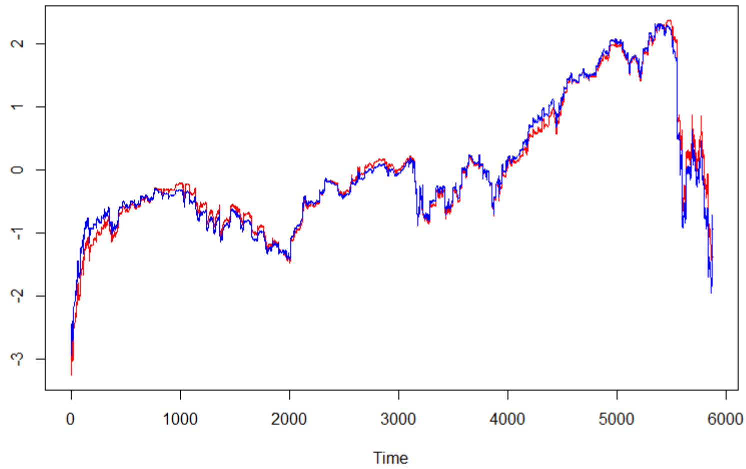
Fonte: Elaborated by the author

Figure C6 - Ibovespa 15 minutes after the covid



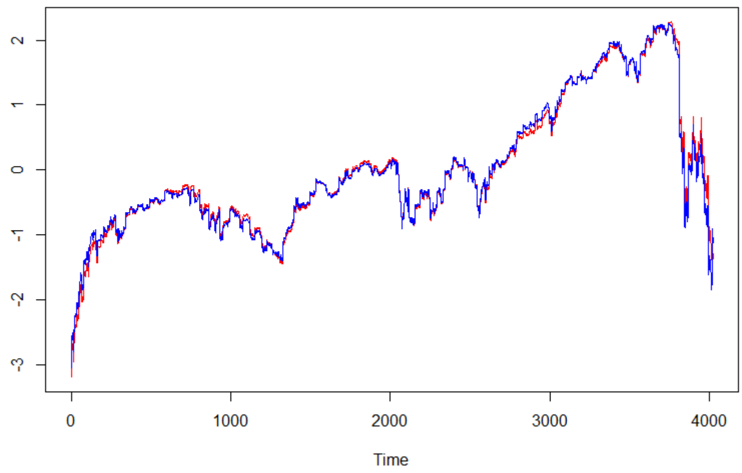
Fonte: Elaborated by the author

Figure C7 - S&P500 rate 5 minutes before the covid



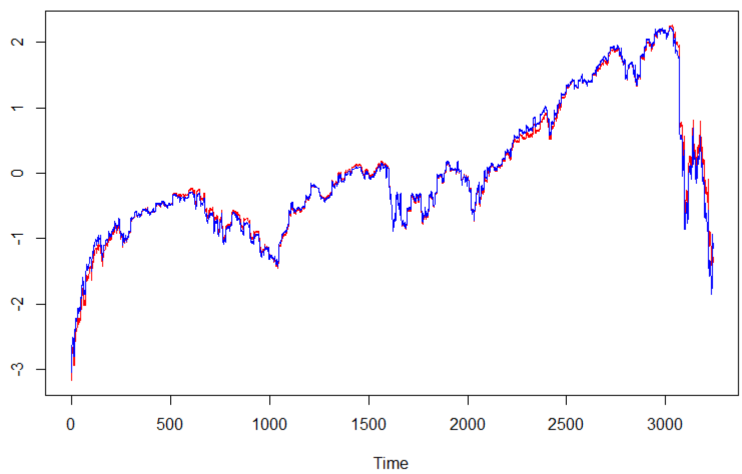
Fonte: Elaborated by the author

Figure C8 - S&P500 rate 10 minutes before the covid



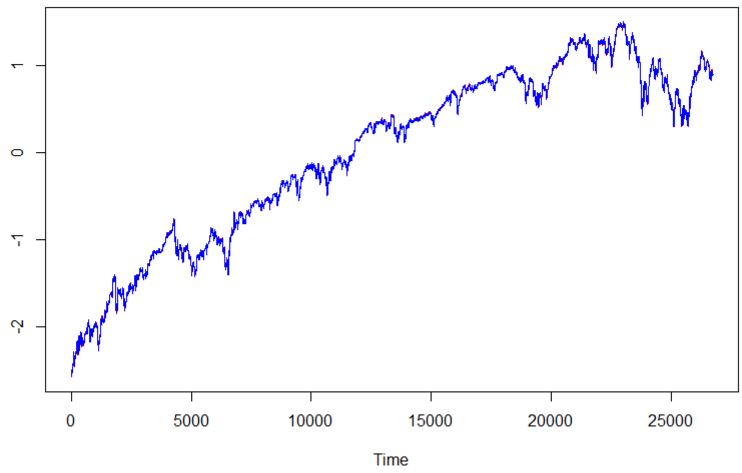
Fonte: Elaborated by the author

Figure C9 - S&P500 15 minutes before the covid



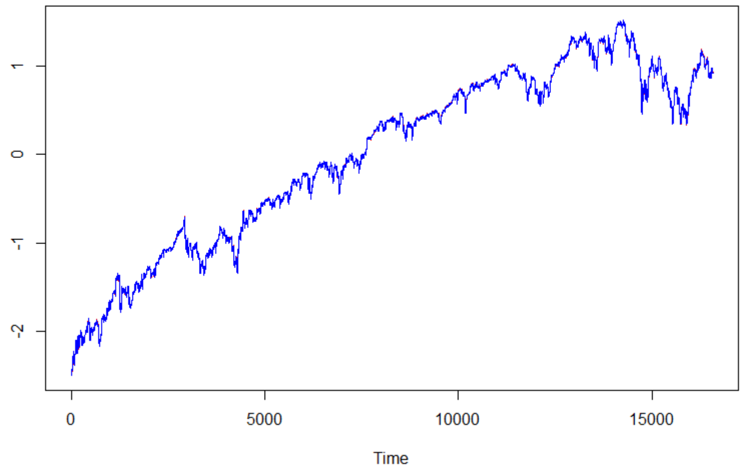
Fonte: Elaborated by the author

Figure C10 - S&P500 5 minutes after the covid



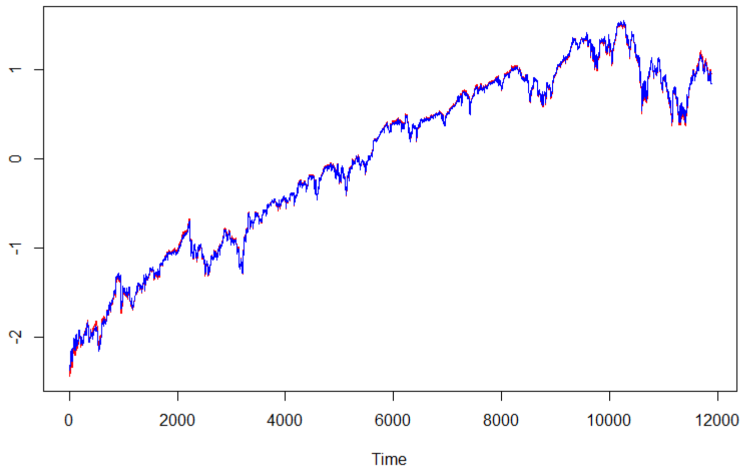
Fonte: Elaborated by the author

Figure C11 - S&P500 10 minutes after the covid



Fonte: Elaborated by the author

Figure C12 - S&P500 15 minutes after the covid α's



Fonte: Elaborated by the author