Physica A 505 (2018) 1139–1149

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Do sentiments influence market dynamics? A reconstruction of the Brazilian stock market and its mood

Tanya Araújo [a,b,*], Samuel Eleutério [c], Francisco Louçã [a,b]

[a] ISEG (School of Economics and Management), University of Lisbon, Rua do Quelhas, 6 1200-781 Lisboa, Portugal
[b] UECE - Research Unit on Complexity and Economics, Rua do Quelhas, 6 1200-781 Lisboa, Portugal
[c] Instituto Superior Técnico, University of Lisbon, Rovisco Pais, 1 1049-001 Lisboa, Portugal

HIGHLIGHTS

- The Bovespa networks of stocks are induced from daily data, from 2006 to 2015.
- The crises of 2008, 2011 and 2014 shape the Minimal Spanning Trees of those networks.
- For the same time interval, the evolution of the TRMI-BR50 sentiment index is analysed.
- The co-joint evolution of the TRMI-BR50 sentiment index and the Bovespa is analysed.
- It provides a fair indicator of the dynamics of the Brazilian economy during that period.

ARTICLE INFO

Article history:
Received 15 January 2018
Received in revised form 27 February 2018
Available online 14 April 2018

Keywords:
Market sentiments
Stock markets
Bovespa
Brazil
MarketPsych

ABSTRACT

Sentiments play an important role in justifying economic actions and are typically presented as being a modern incarnation of expectations that influence financial markets, whether they be of a Keynesian or other type. For the case of the São Paulo Stock Market Index (IBovespa), this paper investigates whether sentiments, as publically expressed in specialised media, represent a covariate variable which influences stock market returns, and also how market dynamics evolve through time, especially in times of major shocks or recessions. In this study we use a network approach to relate the evolution of asset returns to a sentiments index. Daily data from IBovespa and a Thomson Reuters MarketPsych index are used as fair indicators of the evolution of the Brazilian economy from 2007 to 2015. We prove that changes in market prices affect news more than the reverse.

1. Introduction

Previous economists discussed the role of sentiments as they modelled or referred to expectations. One famous example is that of Keynes, who mentioned the "animal spirits" in his magnum opus: "Even apart from the instability due to speculation, there is the instability due to the characteristic of human nature that a large proportion of our positive activities depend on spontaneous optimism rather than mathematical expectations, whether moral or hedonistic or economic. Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as the result of animal spirits-a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities" [1]. As it is obvious, Keynes challenged the probabilistic representation of expectations and that is why he chose an ill-defined term such as "animal spirits" to

* Correspondence to: R. Miguel Lupi, 20, 1249-078 Lisboa, Portugal.
E-mail address: tanya@iseg.utl.pt (T. Araújo).

https://doi.org/10.1016/j.physa.2018.04.045
0378-4371/

represent the perceptions and decisions of agents faced with uncertainty and the future effect of their decisions. This radically subjective approach of possibilities and their interpretation by agents [2] is to be distinguished from the contemporary concept of sentiments, to be measured as statistically measurable signs such as the use of certain words, whose meaning is taken as objectively defined.

A new trend of literature over the most recent years has described the results of automated forecasting of financial markets using text-based information [3,4]. The availability of on-line text data, in the form of news and social media outlets, has experienced enormous growth. Relevant fluctuations in market volatility were shown to be correlated to the co-occurrence of keywords in on-line news [3], as discussed below.

The notion that markets react to sentiments, or that they generate sentiments is current in financial analysis [4]. However, as fluctuations in large stocks tend to have greater impacts, the analysis of these sentiments has become a pertinent research theme for various agencies and trading firms [5].

The analysis of data regarding sentiments is a recent phenomenon. It came to the fore in 2010, with the publication of the results of a team of researchers from Indiana University, who used a large data set of millions of tweet feeds to predict the behaviour of the stock market three days later, and announced an 87% accuracy level for their predictions [6].

This approach for gathering and analysing data on sentiments was used again in June 25th 2012, when Thomson Reuters published 18,864 new indices, which were updated minute by minute. The data was compiled by MarketPsych, a California company, and it analyses up to approximately 55,000 news sites, and 4.5 million social media sites, blogs and tweets. The availability of big data does not imply accuracy in all cases — in fact, Derwent Capital in London, a trader that had used Twitter signals for its market analysis since 2011, closed its fund, as did an internal fund that had been created by MarketPsych itself [7]. Nevertheless, Reuters proceeded with a project that makes available some of its data for scientific research, such as that used in this paper. Bloomberg created a similar project for sentiment analysis, using heavy computational capacities based on machine learning and natural language processing [8]. In these cases, relating co-occurring words in social media documents has also been a source of sentiment analysis [9], in which opinions with respect to any market item of interest (i.e. an asset) are transformed into quantities that can be re-scaled to fit in a bounded interval in order to define an index of market sentiment.

The enormous growth of available online text data in the form of statements gathered from on-line investment communities and other information increases the viability of these market indexes. The use of the information encoded in the indexes of market sentiments can be twofold: they can both be seen as a consequence and influence of the decisions of market investors. In both cases, gains and losses of financial markets would be closely related to the indexes behaviour. The Brazilian stock market was highly attractive to investors during the period under consideration, representing one of the world's fastest growing economies. Despite the many differences existing between the economies of emerging and developed countries during the last financial crisis (2007–2008), emerging and developed financial markets displayed similar volatility [10]. In the case of Brazil, its stock market exhibited a long memory [11], similar to the US and other developed markets. The main questions addressed in this paper are:

1. The induction of the Bovespa minimal spanning tree (MST) from the available data: We hypothesised that by relating Bovespa stock returns by a network approach, the dynamics of the Bovespa MST of stocks is shaped by the occurrence of bubbles and crises.
2. The analysis of the co-joint evolution of the TRMI-BR50 sentiment index, the IBovespa one-day returns, and the Bovespa MST: We verify whether relevant changes occurring in the Bovespa MST provide a fair indicator of the dynamics of the IBovespa, which is uncovered in the correlation between the sentiment index and the IBovespa one-day returns.

We envisage that during critical periods, an increasing correlation between IBovespa and TRMI-BR50 is the correlate of changes occurring in the Bovespa network of stocks. Accordingly, relevant changes occurring in the Bovespa network of stocks would provide a fair indicator of the dynamics of the Bovespa stock market and its behaviour would be a symptom of the evolution of the Brazilian economy during the period under consideration, as would be the correlation between IBovespa and TRMI-BR50.

To this end, the paper is organised as follows: the next section presents the method used to induce the networks of stocks, then Section 3 describes the empirical data: the Bovespa stocks, the IBovespa index and TRMI data, as well as the interactions between Bovespa and some TRMI sentiments. In Section 4 we discuss the implications. Section 5 concludes and outlines future work.

2. Method

Using Thomson Reuters MarketPsych Indices (TRMI) for the Brazilian Bovespa BR50 daily index for a period of nine years, from January 2007 to December 2015, we analyse the joint-evolution of TRMI-BR50 and IBovespa using a network approach. Network-based approaches are nowadays quite common in the analysis of the structural and behavioural aspects of stock markets [12].

In contrast to the dominant analysis of these markets, which describe the movements in returns as the result of a stochastic process, which is tainted by short-lived memory, our approach looks for evidence of structure as a consequence of perturbations. Furthermore, we prove that an empirically-based interpretation of these dynamics is possible and is effective in detecting patterns of change.

There may be many ways in which the elementary units and the links between them are defined. Here we induce networks where similarities between each pair of Bovespa stocks are used to define the existence of every link in each of those networks. Such a similarity is measured by the correlation coefficient of two (return) time series $\overrightarrow{r}(k)$ and $\overrightarrow{r}(l)$ computed over a time window t.

$$C_{kl} = \frac{\left\langle \overrightarrow{r}(k)\overrightarrow{r}(l)\right\rangle - \left\langle \overrightarrow{r}(k)\right\rangle\left\langle \overrightarrow{r}(l)\right\rangle}{\sqrt{\left(\left\langle \overrightarrow{r}^{2}(k)\right\rangle - \left\langle \overrightarrow{r}(k)\right\rangle^{2}\right)\left(\left\langle \overrightarrow{r}^{2}(l)\right\rangle - \left\langle \overrightarrow{r}(l)\right\rangle^{2}\right)}} \qquad (1)$$

being

$$r_t(k) = \log(p_t(k)) - \log(p_{t-1}(k)) \qquad (2)$$

the one-day return of stock k.

The network of stocks where links between each pair of stocks are defined by the correlation coefficient computed on a time window t are weighted graphs where the weight of each link corresponds to the intensity of the similarity between the linked pair of stocks. These weighted networks are further analysed through the construction of their corresponding minimal spanning trees (MST). In so doing, we are able to emphasise the main topological patterns that emerge from the network representations and to discuss their interpretation.

The distances d_{kl} are defined using the correlation coefficients C_{kl}.

$$D_{kl} = \sqrt{2\left(1 - C_{kl}\right)} \qquad (3)$$

From the $N \times N$ distance matrix D, a hierarchical clustering is then performed using the *nearest neighbour* method. Initially N clusters corresponding to the N stocks are considered. Then, at each step, two clusters c_i and c_j are clumped into a single cluster if

$$d\{c_i, c_j\} = \min\{d\{c_i, c_j\}\}$$

with the distance between clusters being defined by

$$d\{c_i, c_j\} = \min\{d_{pq}\} \text{ with } p \in c_i \text{ and } q \in c_j$$

This process is continued until there is a single cluster. This clustering process is also known as the *single link method*, being the method by which one obtains the minimal spanning tree (MST) of a graph.

In a connected graph, the MST is a tree of $N - 1$ edges that minimises the sum of the edge distances. In a network with N nodes, the hierarchical clustering process takes $N - 1$ steps to be completed, and uses, at each step, a particular distance $d_{i,j} \in D$ to clump two clusters into a single one.

The same calculation is performed for time permuted data. Time-permuted data were generated by permuting each stock (one-day return data) randomly over time. As each stock is independently permuted, time correlations among stocks disappear, while the resulting surrogate data preserve the mean and variance that characterise actual data.

3. The data

The data we work with is comprised of two data sources: (1) the 35 time series with daily data of 35 stocks of the Bovespa, the Sao Paulo Stock Exchange Index (IBovespa), and the time series with TRMI-BR50 daily data. For the first data source, only stocks with data for the full period were considered. One time series is also included with the daily (closing) prices of IBovespa itself.

3.1. Bovespa data

The IBovespa comprises the most representative companies of the Brazilian stock market, which is the largest stock exchange in Latin America. The data we considered comprises 35 stocks in the IBovespa from January 2007 through to December 2015. We considered the six sectors in which the following 35 companies of the IBovespa are grouped: Ambev (ABE), Banco do Brasil (BBA), Banco Bradesco (BBD3 e BBD4), Bradespar (BRA), Braskem (BRK), Grupo CCR (CCR), Centrais Elétricas de São Paulo – CESP (CES), Cia Energética de Minas Gerais — CEMIG (CMI), Companhia Paulista de Força e Luz – CPFL (CPF), Companhia Paranaense de Energia-Copel (CPL), Souza Cruz (CRU), Companhia Siderúrgica Nacional (CSN), Centrais Elétricas Brasileiras – Eletrobrás (ELE3 and ELE6), Embraer (EMB), Gafisa (GFS), Siderúrgica Gerdau (GGB), Metalúrgica Gerdau (GOA), Gol Linhas Aéreas Inteligentes (GOL), Itaúsa (ITS), Itaú Unibanco (ITU), Klabin (KLB), Lojas Americanas (LAM), Light (LIG), Lojas Renner (LRE), Natura (NAT), Pão de Açúcar (PCA), Petrobrás (PET3 and PET4), Localiza Rent a Car (REN), Tractebel (TBL), Usiminas (USI), and Vale (VAL3 and VAL5). We have analysed the set of daily data of the Bovespa stocks for a time period from January 2007 to December 2015. Table 1 show the 35 companies grouped in six different sectors: Energy, Finance, Basic Materials, Consumer Goods, Utilities and Industrials.

Table 1
The 35 Bovespa companies grouped in six sectors.

Energy	Financials	Materials	Utilities	Consumer	Industrials
PET3	BBA	BRK	CES	ABE	CCR
PET4	BBD3	BRA	CMI	GOL	CRU
	BBD4	CSN	CPF	LAM	EMB
	ITU	GGB	CPL	LRE	GFS
		GOA	ELE3	NAT	ITS
		USI	ELE6	PCA	KLB
		VAL3	LIG	REN	
		VAL5	TBL		

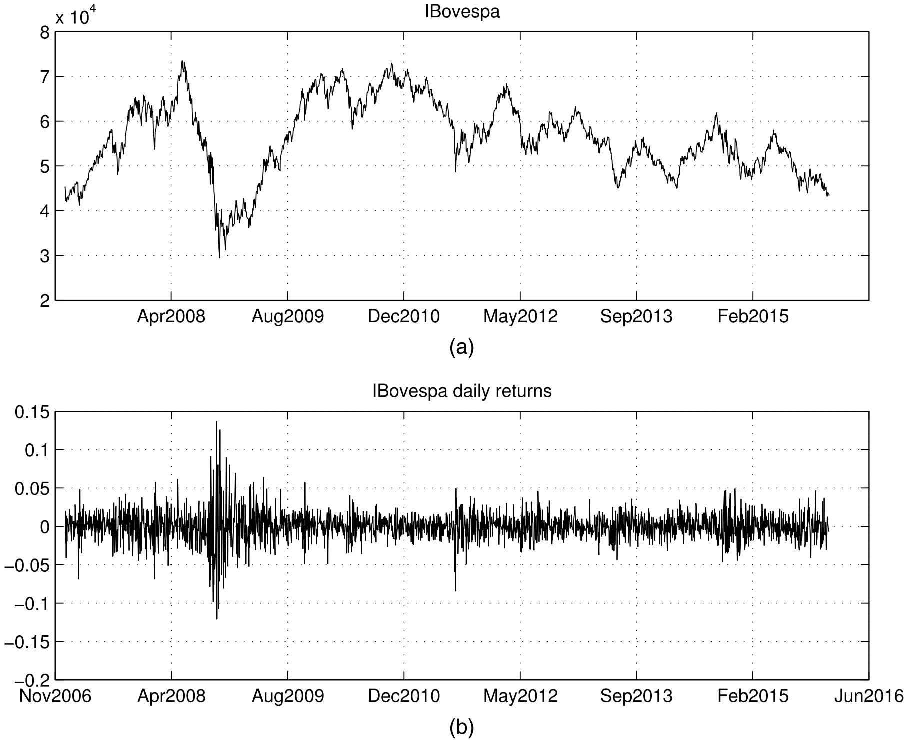


Fig. 1. (a) The evolution of the IBovespa (p_t^*) from 2007 to 2015 and (b) The evolution of IBovespa one-day returns (r_t^*) in the same time interval.

We presume that this population of firms provides a fair indicator of the dynamics of the global stock market and that its behaviour is a symptom of the evolution of the Brazilian economy during the period under consideration. The first plot in Fig. 1 shows the evolution of the IBovespa index (p_t^*) from January 2007 to December 2015. The second plot shows the evolution of the returns of the IBovespa index (r_t^*) for the same time interval.

In the next section, we investigate the dynamics of the impact of the sub-prime crisis of 2008, and that of the global commodity prices crisis of 2011, and also the consumption crisis, since 2014.

3.2. *Thomson Reuters MarketPsych indexes (TRMI)*

TRMI indexes result from the quantification of meaning originated by a huge amount of textual data from financial news and social media published since 1998. Indexes are processed daily and are updated every minute. MarketPsych-specific sources of text include The Wall Street Journal, The Financial Times, The New York Times, Seeking Alpha, and a dozen more sources that are consulted by professional investors [13]. Lexical ambiguity is one of the biggest difficulties for a correct validation of the entities being identified. In fact, homography, or compound words, using as a reference place names, etc., requires the use of filters and synonyms for a more correct identification of entities. To improve entity name disambiguation, MarketPsych uses supervised machine learning to identify, correlate, and anti-correlate words in proximity of ambiguous entity Ref. [13].

The TRMI consists of a set of variables normalised for the interval [0,1] or [−1,1]. For example, "optimism" can be represented by values between −1 and +1, according to whether the predominant feelings are optimistic or pessimistic.

Table 2
The 31 TRMI daily indexes for the Bovespa (TRMI-BR50).

Sentiment	Optimism	Fear	Joy
Trust	Violence	Conflict	Gloom
Stress	timeUrgency	Uncertainty	emotionVsFact
longShort	longShortForecast	priceDirection	priceForecast
Volatility	loveHate	Anger	debtDefault
Innovation	marketRisk	analystRating	Dividends
earningsForecast	fundamentalStrength	Layoffs	Litigation
managementChange	managementTrust	Mergers	

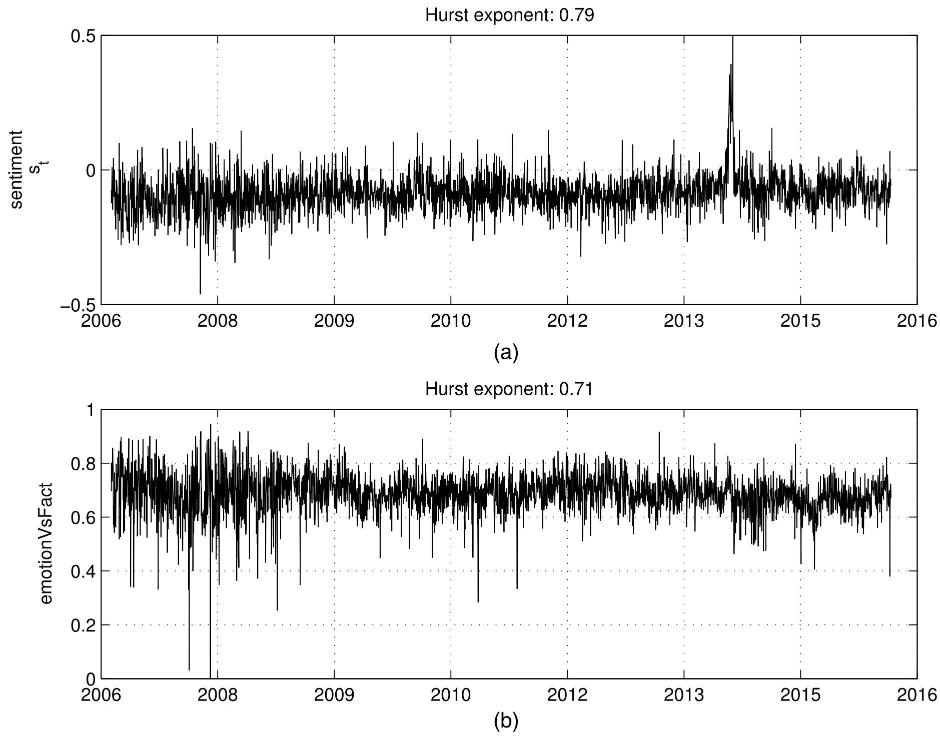


Fig. 2. The evolution of (a) *sentiment* index and (b) *emotionVsFacts* indexes.

"Fear" provides information about of fear and anxiety, and therefore is represented by values of between 0 and 1. 31 indexes for the companies and equity index asset classes are used. Each TRMI is composed of a combination of variables (Vars) for all constituents ($c \in C$) over the past 24 h. The sum of the absolute values of all contributing Vars (V) for all constituents of an asset yields is the Buzz index of this asset [13].

Table 2 shows the entire set of 31 TRMI daily indexes for the Bovespa (TRMI-BR50).

Fig. 2 shows two examples of the evolution of two TRMI: Sentiment and Emotion Vs Facts.

Complementary evidence is provided by the computation of the Hurst exponent for the period under consideration, indicating evidence of long term memory for the time interval. Two examples of these Hurst exponents are represented in Fig. 2. As the Hurst exponent computed for each of these signals indicates, many news sentiments exhibit a large degree of auto-correlation, suggesting the existence of strong effects of long memory [14].

4. Results

4.1. Critical periods from Bovespa data

Using the whole data for the nine years, the calculations described in the last section were performed for the actual returns data, and also for the time-permuted data. Fig. 3 shows the MST of the 35 Bovespa stocks obtained from time-permuted data and using the entire time period ($t = 2250$). As in the following figures, the 35 companies are coloured according to the sector of activity that they belong to.

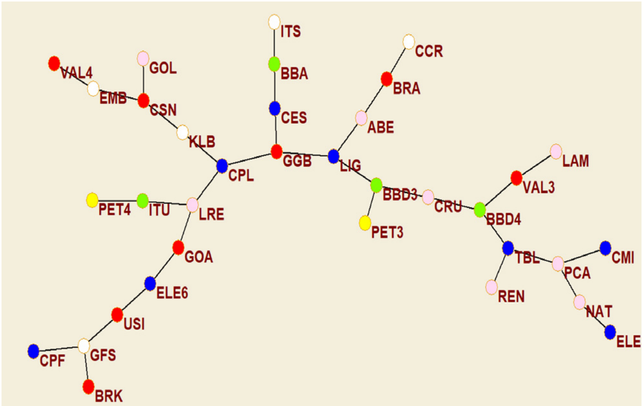


Fig. 3. The MST of the 35 Bovespa stocks obtained from surrogate (time-permuted) data.

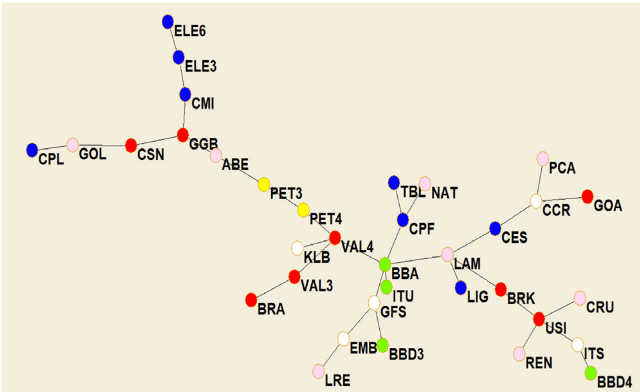


Fig. 4. The MST of the 35 Bovespa stocks obtained from actual data and using the entire time period 2007–2015.

The organisation of the 35 Bovespa stocks on the tree in Fig. 3 (obtained from surrogate data) does not show any clustering trend. All the branches combine companies belonging to different sectors of activity suggesting a situation where *business-as-usual* prevails. Conversely, Fig. 4 shows the MST of the 35 Bovespa stocks obtained from actual data, also using the entire time period ($t = 2250$). During these nine years (from 2007 to 2015), the empirical data shows some clustering trend, which seems to be restricted to the companies in the Materials sector (VAL3, VAL5 and BRA). On the contrary, stocks belonging to the Financial sector (BBD3, BBD4 and BBA) occupy different branches on the tree, suggesting that this sector is the less synchronous one in the network of stocks during the nine-year period.

Next, we divided the data into 9 chronologically-successive batches of 12 months ($t = 250$). The idea is to investigate the consequences of some particularly turbulent periods for the distribution of the stocks on the branches of the trees. The first MST presented in Fig. 5 shows the tree of the Bovespa stocks obtained for 2008 while the second tree shows the MST obtained from 2009 data. While in the first tree, the 35 Bovespa stocks seem to organise themselves according to some clustering trend, the way companies are organised in the second tree of Fig. 5 suggests that in this year *business-as-usual* prevails. There (the second tree in Fig. 5), the companies in the Energy (PET3 and PET4) sector occupy far positions on the tree, which seems to be a consequence of government intervention in the electricity sector in 2008, which ended up scaring off investment in the energy sector and caused a sharp reduction in demand for stocks of energy companies.

The 2009 Annual Report of the Central Bank of Brazil argues that this continuous growth of the Brazilian stock market was the result of the action of foreign institutional investors. Economic stabilisation, as well as a less tight monetary policy helped the performance of the capital market [15]. This is a significant result, especially when we take into account the turmoil that affected financial markets worldwide during the second half of the year, following on from the initial effects of the sub-prime crisis in the United States.

The branches of companies on the first tree (a) presented in Fig. 5 display a different situation, which arose from the impact of the global financial crisis on the Brazilian economy. There, the stocks of the Materials sector (VAL3, VAL5, GGB, CSN and BRA) display synchronous behaviour, occupying close positions on the tree obtained from the empirical daily data of 2008. Such a synchronous behaviour seems to be due to a credit expansion in the Brazilian financial system from January 2008 through March 2009, which led to a significant appreciation of stocks in the Materials sector [16].

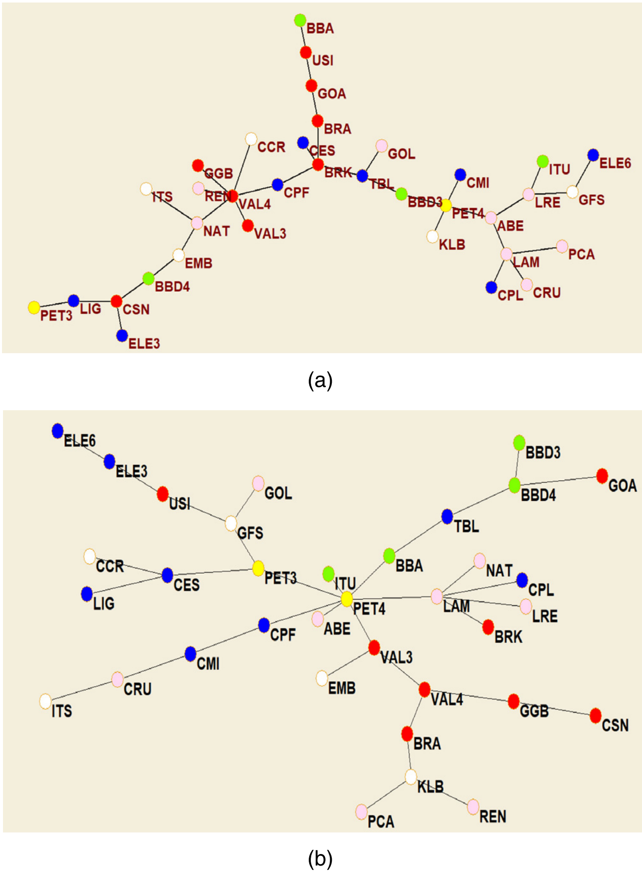


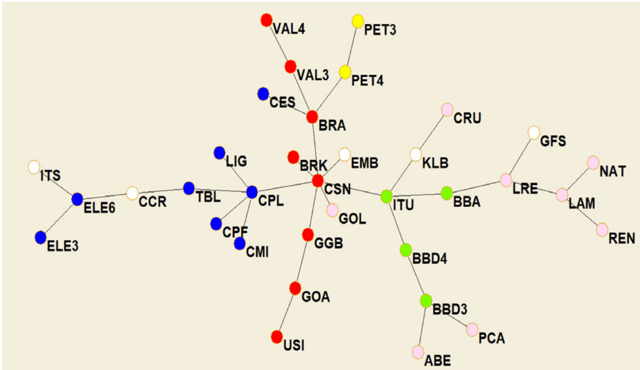
Fig. 5. The MST of the 35 Bovespa stocks obtained from actual data for (a) the critical year of 2008 and (b) the *business-as-usual* year of 2009.

The trees presented in Fig. 6 describe the behaviour of this population of firms during the recessions that occurred in 2011 and 2014. In both trees one sees a strong clustering of stocks belonging to the same sector. The first (a) MST in Fig. 6 shows the critical period of 2011, when the commodity crisis took place. There, the stocks of the Materials sector (red) occupy the most central branch of the tree, tying together every single stock in this sector. Likewise, the companies belonging to the Utilities sector (blue) are placed together on the left side of the MST. The Financial sector (green) is also organised on a single branch of the tree. The second tree in Fig. 6 shows the consequences of the domestic consumption crisis on the Bovespa stock market. As the financial crises proceed, the typical trajectory of firms is a type of herd behaviour, forming clusters that entirely separate stocks by sector. Again, this can be especially noted in the case of companies in the Financial (green), Materials (red), Utilities sectors. Moreover, the stocks in the Industrials sector (white) are also placed together, while the same happens to the stocks in the Energy sector (yellow). That amazing separation of stocks by sector shows that during critical periods, firms display herd behaviour, global and local structures are reinforced, being described by such a synchronous convergence on the MST.

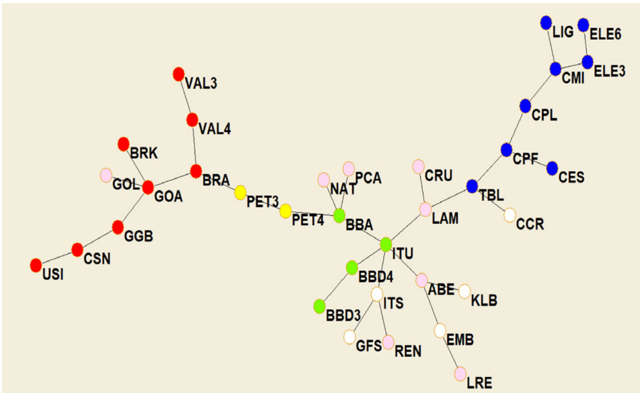
4.2. Looking for interactions among IBovespa and financial news sentiments

A basic illustration of a relationship between the behaviour of the financial news sentiment and the IBovespa prices is provided by plotting their moving averages.

Fig. 7 shows the evolution of three moving averages. The first is the moving average of the values of the sentiment index computed from an 80-days (short) moving window, the second is the same moving average computed from a 200-days (long) moving window, and the third is the moving average of the daily values of IBovespa. When the moving averages of the sentiment index of combined news about IBovespa is plotted against the value of IBovespa itself, a clear relationship emerges, as Fig. 7 shows. In this plot, the interval between the first and the second series is filled with dark (or light) shades, depending on whether long-term values are greater than (or less than) short-term ones. Looking at the simultaneous evolution of these three moving averages, we verify that when the short-term falls below the value of long-term sentiments (dark shades), there seems to be a negative pressure on IBovespa prices. Conversely, when short-term values rise above long-term sentiments (light shades), a positive pressure on IBovespa prices occurs.



(a)



(b)

Fig. 6. The MST of the 35 Bovespa stocks obtained from actual data for the critical years of (a) 2011 and (b) 2014. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

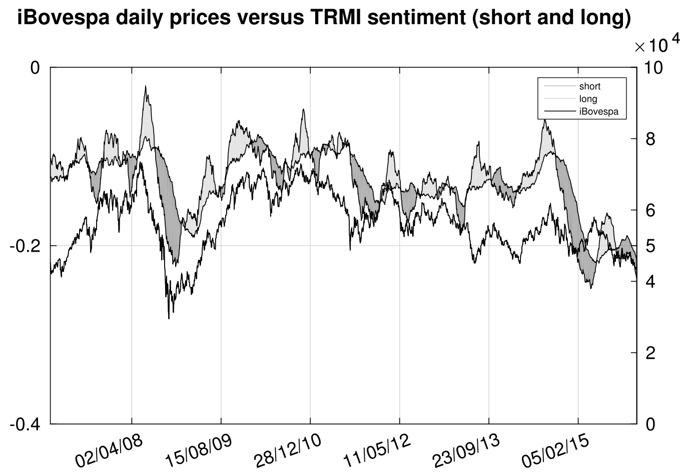


Fig. 7. Short (80-days) and Long (200-days) *sentiment* moving averages and IBovespa.

The rationale for this may be the notion that the financial market predominantly looks for immediate expectations that are formed based on the short term performance of stocks and their returns. Therefore, volatility may be a convenient indicator, as memory is a short term feature and behaviour changes rapidly. This confirms our finding that structure is created essentially when the market crashes or is subject to major recessions. Furthermore, the plot in Fig. 7 shows that the larger negative pressures (dark shades) on IBovespa prices coincide with the recessive periods of 2008, 2011 and 2014.

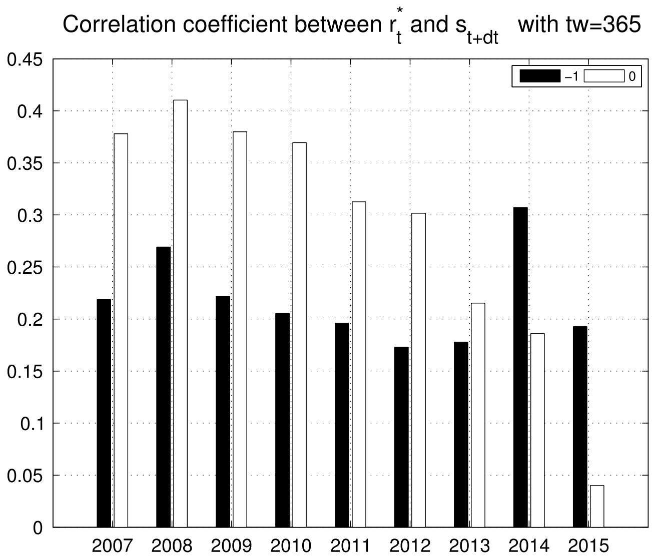


Fig. 8. Time lagged (dt) correlation coefficients between *sentiment* and IBovespa one-day returns (r^*_t).

Moreover, the plot in Fig. 7 shows that the larger negative pressures (dark shades) on IBovespa prices coincide with the recessive periods in 2008, 2011 and 2014.

4.3. Time-lagged correlations

Much attention has been given to understanding the way that stock markets behave during crises [17–20]. An increase in correlations has been associated with periods of economy instability, while normal or growth periods display low correlations between stocks [21]. Stocks tend to be more correlated as a consequence of the interdependencies of the companies with one another. In addition, an increase in correlations can also be due to a feedback effect emanating from a collective variable, such as the one represented by a market index comprised mostly of market stocks. In this last case, the index might represent unquantifiable economic and psychological factors [10]. In order to investigate the predictive power of the sentiment index, below we use a time lag (dt) which corresponds to the number of days between the observation of the media sentiment and the observation of market behaviour, as reflected in IBovespa prices. By doing this, and since the time lag (dt) can assume positive and negative values, we investigate the extent to which news leads to financial market movements, and conversely, to what extent do markets create news. When the time lag is set with negative values ($dt = -1, -2, -3$) we consider news during the days after, i.e., as a consequence of the market. On the contrary, when the time lag is set with positive values ($dt = 1, 2, 3$) we consider just news that was published during the previous days, which is an anticipation of market behaviour. In both cases, news are those conveyed in the sentiment index. Fig. 8 shows the correlation coefficient between the values of the sentiment index and IBovespa one-day returns, computed with a time window of 180 days, for different time lags (dt), from 2007 to 2015.

Along the whole eight years period, the highest values of lagged correlations between daily financial news sentiment and daily returns of IBovespa are reached at $dt = -1$, meaning that IBovespa returns tend to be significantly correlated with news published the day after, as a consequence of market behaviour. Not surprisingly, reactions captured in the news sentiment index respond to market behaviour, whereas a related response by the news to the market, i.e., an anticipation of market behaviour, does not seem to take place.

Furthermore, the lagged correlations between the daily financial news sentiment and daily returns of IBovespa reaches the highest value in 2008, with the occurrence of the global financial crisis, and then decreases from 2009 to 2013, showing an important increase in 2014, when the commodity crisis took place.

From 2007 to 2014, the second highest values of correlations between news sentiment and daily returns occur at $dt = 0$, and thus we decided to look at correlation values computed from multiple time windows ($tw = 30, 60, 90, 180$ and 365) at zero time lag. Fig. 9 shows the evolution of the non-lagged correlations at different time windows, from one month ($tw = 30$) to one year. The results presented in Fig. 9 correspond to the accumulated value ($\Sigma(C(i, s)_{dt=0})$) of the correlation coefficient between news sentiment and daily returns, measured with different time windows ($tw = 30, 60, 90, 180$ and 365) and for each year (t) in between 2007 and 2015.

Fig. 9 shows that the increased values of the correlation coefficient (in 2008, 2011 and 2014), measured with several time windows, correspond to the three relevant economic periods in Brazil: in 2008 it shows the impact of the global financial crisis on the Brazilian economy; in 2011 it reflects the commodity crisis, and; in 2014 the domestic consumption crisis has an

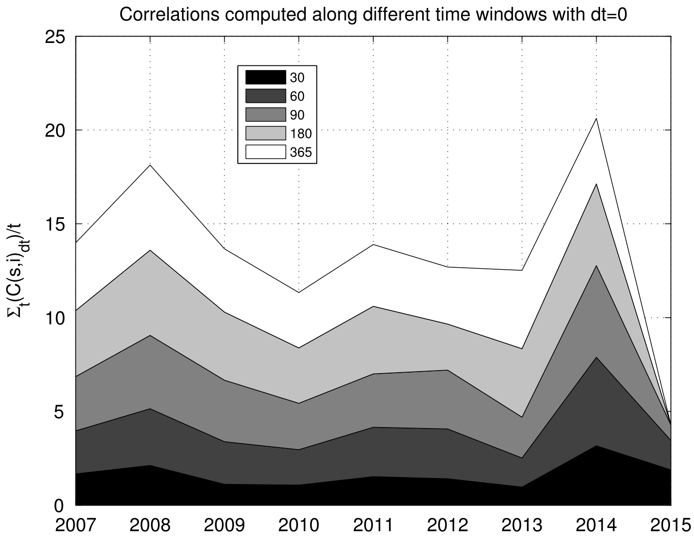


Fig. 9. Zero time lagged correlation coefficients between *sentiment* and IBovespa one-day returns measured with multiple time windows.

impact. The evolution of non-lagged correlations over time reinforces the correspondence between three important critical economic periods in Brazil, while the increase in the correlation between IBovespa and the market sentiment index in 2008, 2011 and 2014 reflects the correlation of the changes that occurred in the Bovespa network of stocks.

5. Conclusions

In this paper, by carrying out an un-biased analysis of the data, we have identified some properties of the Bovespa stock market. Market sentiments, when expressed by big data measuring the intensity of words revealing market perceptions by a multitude of agents, were found not to be a representation of expectations in the traditional Keynesian meaning, since the former convey a current representation of signs and attitudes related to the present, and not, as in the Keynesian framework, as the decision-making process related to the future impacts of economic activity. As previously stated, we distinguish the Keynesian view of expectations and the contemporary use of sentiments as expressed through measurable signs.

Nevertheless, we found that these sentiments are predominantly a consequence of changes in returns, and that they follow, but typically do not anticipate, changes in the geometric object describing the empirical structure of the market. Our main conclusions can be summarised as follows:

1. The Bovespa networks of stocks are induced from daily data, from 2006 to 2015.
2. The induction of the Bovespa network of stocks from the one-day returns of Bovespa prices enables the identification of important changes that occurred in the minimal spanning trees of these networks in 2008, 2011 and 2014. These results are in line with those presented in the reference literature [10], where an increase was found in all correlations between IBovespa stocks at the end of 2008.
3. These changes provide a fair indicator of the dynamics of the Bovespa stock market during critical periods.
4. The analysis of the joint-evolution of the sentiment index and IBovespa daily returns shows that when correlations are computed from prices and news sentiments (i), there is a negative pressure on prices when short-term values of the sentiment index fall below long-term ones, and, (ii) conversely, there is a positive pressure on prices when short-term values rise above the long-term values of the sentiment index.
5. The analysis of time-lagged correlations shows that news sentiment items affect financial market movements much less significantly than market prices affect news.
6. News sentiment signals exhibit a large degree of auto-correlation.
7. Financial markets anticipate news much more substantially than news items anticipate market movements. Therefore, these results refute the conclusion obtained in Ref. [6], and they suggest that media sentiments are not a predictor of market dynamics.
8. The important increases in the correlation between IBovespa and the sentiment index in 2008, 2011 and 2014 are the correlate of the changes that occurred in the Bovespa network of stocks, given that the financial crises occurring during those years.

Therefore, we surmise that the representation of sentiments, as obtained by intense computation from big data of published comments, articles, tweets, and the cloud of words and signals in the specialised and influential media is not adequate, at least in the present forms, to represent a quantitative measure of expectations. It is decisions and changes in market behaviour that drive sentiments, as expressed by these measures, rather than the other way round.

Acknowledgement

Financial support by FCT (Fundação para a Ciência e a Tecnologia), Portugal is gratefully acknowledged. This article is part of the Strategic Project: UID/ECO/00436/2013.

References

[1] J.M. Keynes, The General Theory of Employment, Interest, and Money, Harcourt, New York, 1936.
[2] G.A. Akerlof, R.J. Shiller, Animal Spirits: How Human Psychology Drives the Economy and Why It Matters for Global Capitalism, Princeton University Press, Princeton, 2009.
[3] M. Piskorec, N. Antulov-Fantulin, P.K. Novak, I. Mozetia, M. Grar, I. Vodenska, T. Smuc, Cohesiveness in financial news and its relation to market volatility, Nature Sci. Rep. 4 (2014).
[4] C. Curme, H.E. Stanley, I. Vodenska, Coupled network approach to predictability of financial market returns and news sentiments, Int. J. Theor. Appl. Finance 18 (2015).
[5] Financial Times 2016. Weak US services data weigh on sentiment, Financial Times, September 6 2016, available at: www.ft.com/content/5d98775c-73df-11e6-bf48-b372cdb1043a.
[6] J. Bollen, H. Mao, X. Zeng, Twitter mood predicts the stock market, J. Comput. Sci. 2 (2011) 18.
[7] The Economist 2012. The mood of the market The Economist, June 28. http://www.economist.com/blogs/graphicdetail/2012/06/tracking-social-media.
[8] Bloomberg, available at: www.bloomberglabs.com/data-science/projects/sentiment-analysis-financial-news-social-media/.
[9] S. He, X. Zheng, D. Zeng, C. Luo, Z. Zhang, Exploring entrainment patterns of human emotion in social media, PLoS One 11 (2016).
[10] S.P. Leahy, S.L. Carciente, H.E. Stanley, D.Y. Kenett, Structure and Dynamics of the Brazilian Stock Market: A Correlation Analysis, 2014. ssrn.com/abstract=2484648 or http://dx.doi.org/10.2139/ssrn.2484648.
[11] J. Cavalcante, A. Assaf, Long range dependence in the returns and volatility of the brazilian stock market, Eur. Rev. Econ. Finance 3 (2004) 22.
[12] T. Araújo, S. Banisch, Multidimensional analysis of linguistic networks, in: Towards a Theory of Complex Linguistic Networks, Springer, Berlin, 2014.
[13] R.L. Peterson, Trading on Sentiment: The Power of Minds over Markets, John Wiley & Sons, 2016.
[14] T. Araújo, F. Louçã, The seismography of crashes in financial markets, Phys. Lett. A 372 (2008) 429–434.
[15] Central Bank of Brazil 2009. Annual Report, Publications office of the CBB, Brasília. https://www.bcb.gov.br/RED-ANNUALREPORT.
[16] P. Andersen, The crisis in brazil, Lond. Rev. Books 38 (2016) 15–22. http://www.lrb.co.uk/v38/n08/perry-anderson/crisis-in-brazil.
[17] B. Podobnik, H.E. Stanley, Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series, Phys. Rev. Lett. 100 (2008) 84102.
[18] P. Gopikrishnan, V. Plerou, Y. Liu, L.N. Amaral, X. Gabaix, H.E. Stanley, Scaling and correlation in financial time series, Physica A 287 (2000) 362–373.
[19] D.Y. Kenett, T. Preis, G. Gur-Gershgoren, E. Ben-Jacob, Quantifying meta-correlations in financial markets, Europhys. Lett. 99 (2012) 38001.
[20] A.M. Petersen, F. Wang, S. Havlin, H.E. Stanley, Market dynamics immediately before and after financial shocks: Quantifying the omori, productivity, and bath laws, Phys. Rev. E 82 (2010) 036114.
[21] T. Preis, D.Y. Kenett, H.E. Stanley, D. Helbing, E. Ben-Jacob, Quantifying the behavior of stock correlations under market stress, Nature Sci. Rep. 2 (2012).