EL COLEGIO DE MÉXICO

CENTRO DE ESTUDIOS ECONÓMICOS

MAESTRÍA EN ECONOMÍA

TRABAJO DE INVESTIGACIÓN PARA OBTENER EL GRADO DE MAESTRO EN ECONOMÍA

TWITTER SINGS, THE STOCK MARKET LISTENS

JORGE LUIS SOSA GARCÍA

PROMOCIÓN 2016-2018

ASESOR:

DRA. LAURA JUÁREZ GONZÁLEZ

MAYO 2018

I want to thank…

El Colegio de México for changing the way I see the world, and more importantly, for changing the way I want to change the world.

My advisor, Laura Juárez, for believing and supporting this project before anyone else.

My professors, especially Raymundo Campos and Aurora Ramírez, for pushing me to reach my full potential.

Marco Gómez and Samuel Hernández, for their friendship and company in this journey.

Ivette Agundez, for all the incredible moments together. By her side I have been happier in the last two years than ever before.

To my parents and my brother, for their unconditional support and love, without which I would not be where I am and for which I am who I am.

Abstract

I provide evidence that social mood can influence the performance of the Mexican stock market. Using a Granger Causality test, I correlate different measurements of the collective mood derived from Twitter with the opening and closing values, as well as with daily volume of trading on the leading index of the Mexican stock market, the IPC, and the stock of AZTECA.

The results show that the total number of tweets posted, especially the positive ones, are Granger-causal of the daily volume of trading for both IPC and AZTECA, and of the opening values for AZTECA. The results are very similar if the Tweeter feeds considered are restricted only to the five states with the most Internet users.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the primary results in Behavioral Economics is that emotions can profoundly affect individual behavior and decision making. If this is the case then a natural follow up question is whether this holds for societies at large, i.e. can social mood affect their collective decision making? I will investigate whether measurements of social mood derived from Twitter feeds are correlated to the values of the leading index in the Mexican stock market (BMV) the Price and Quotations Index (IPC) and the stock of AZTECA (a large telecommunication firm in Mexico).

From a merely theoretical perspective, the socionomic hypothesis as explained by Nofsinger (2005), provides valuable insights into this matter. According to this hypothesis, human interactions spread emotions, which in turn characterize how people will act (what people think derives from how they feel). This way, if the mood is correlated across society, optimism or pessimism affects financial decisions that can lead to marketwide phenomena.

This means, regarding this investigation, that the social mood (as expressed in Twitter) has predictive power over the value of the IPC and the share of AZTECA.

However, it is just as natural to think that it is the stock market performance that, to some degree, determines social mood. Thus, it is possible for example, that when media reports a lousy day for the IPC both financial and non-financial agents take to Twitter to express more negative comments than usual.

Existing literature shows mixed results in regard to this. Earlier research shows that new information mainly drives the stock market, and since new information is unpredictable, the stock market cannot be predicted. Nonetheless, research done in the las decade provide evidence that the information coming from social media can provide leading indicators of commercial and economic activity, and therefore this information can help us predict the stock market.

In this research, I use a Machine Learning algorithm developed by the National Institute of Statistics and Geography (INEGI) to get a measurement of the collective social mood derived from Twitter activity in Mexico to find a correlation with the IPC over time.

I use tweets posted from Mexico between January 1st, 2016 and March 2nd, and the financial series of the IPC and AZTECA over the same period to construct Vector Autoregressive (VAR) models, which I then use to run a Granger Causality Test. This test informs me of the statistical significance of the data coming from Twitter to predict the Mexican stock market performance.

I find that the volume of tweets posted, particularly positive tweets, are predictive of the volume of daily transactions of the IPC, but none of the measures derived from Twitter proposed here seem to be useful for predicting either the opening or closing values for the IPC.

This research differs from existing literature in two key aspects. First, I consider tweets only from Mexico and use them to predict the IPC, the leading index in BMV, while previous literature used worldwide tweets to predict the Dow Jones Industrial Average (DJIA) or NASDAQ index. Second, I use a very simple categorization of social moods: positive or negative. To the best of my knowledge, the interaction between this simple categorization of social moods and the use of only national tweets has not been analyzed before.

In this work, I make three contributions to the existing literature. First, I show that the collective mood of a specific society does have an impact on the (economic) decisions it makes. Second, I indirectly show the importance of incorporating commonly ignored factors, such as mood, in the mainstream economic analysis. Finally, I show that using new technique such as Machine Learning as well as information from non-conventional sources like social media is a promising path for economic and financial research. Notably, using an adequate Machine Learning algorithm to extract data from social media can provide us with data that is not available anywhere else in the world.

The rest of the research is organized as follows: Chapter 2 offers a brief revision of the existing literature. Chapter 3 describes the Machine Learning algorithm used as well as the data considered. Chapter 4 states and discusses the main results, and Chapter 5 concludes.

Chapter 2

Literature Review

Stock market prediction has long been in the interest of academia. Early research like Fama et al. (1969), Fama (1970) and Fama (1991) was based on the Efficient Markey Hypothesis (EMH) developed in Fama (1965). According to the EMH, stock market process are primarily driven by new information, i.e. news, rather than past or present prices. Since news are unpredictable, the EMH implies that prices in the stock market follow a random walk; thus prices cannot be predicted.

Nonetheless, research on behavioral finance like Kavussanos and Dockery (2001) or Qian and Rasheed (2007) shows that stock market prices do not follow a random walk and can be predicted to some degree.

Although it is thought that news are the primary driver of stock market prices, public mood or sentiment may be just as important. We know from the groundbreaking work from Kahneman and Tversky (1979) that emotions, as well as information, play a significant role in human decision making. Therefore, it is reasonable to assume that collective mood can drive stock market value just as much as news.

With this in mind, recent research suggest that early indicators of collective mood can be extracted from online social media, i.e. Twitter feeds, blogs message boards, etc., to predict changes in economic and commercial outcomes. For example, Gruhl et al. (2005) analyzes internet chatter and show that it can help predict book sales on Amazon, whereas Lee et al. (2011) uses different measures of online activity on blogs to show that positive messages about a product can influence consumers' decision to shop online.

On the financial side, Tirunullai and Tellis (2012) show that the volume of chatter helps predict both trace volume and volatility of the returns for some stock. Antweiler and Frank (2004) study message boards on Yahoo! Finance and Raging Bull, and find that the volume of messages posted and the level of disagreement of said messages are predictive of the volume of transactions made on the following day. They conclude that the information contained in social media platforms can be relevant to predict stock market performance.

On the other hand, Tumarkin and Whitelaw (2001) study message boards on Raging Bull and find that days with positive returns precede days with more activity on the board, thus concluding that online information is not very useful and that news are still the primary driver of stock market prices. This research, while interesting, differs from my work in that the authors use information coming from comments exclusively related with the stocks they analyze, while I intend to use Tweeter, which contains information and comments related to a large variety of events and topics.

Even though there is no clear consensus on whether measurements of sentiments are helpful to forecast stock prices, research done in the last five years, especially Rechenthin et al. (2013), Luo and Zhang (2013), Luo et al. (2013), Nguyen et al. (2015), and Yu et al. (2013) seems to indicate that predictive models that include these measurements are indeed better than those using only past and present price information.

My work is more closely related to Zheludev et al. (2014), who find that tweets can help predict hourly returns of the Standard & Poor's 500 Index (S&P500). More specifically, the authors find that tweets posted in the last 22 hours contain relevant information to help predict the hourly returns of the S&P500.

Bollen et al. (2011) and Zhang et al. (2011) are also closely related to my work. Both provide evidence that when tweets are submitted to sentiment analysis and classification[1], the information coming from Twitter can help predict the DJIA, S&P500, and NASDAQ indices.

Bollen et al. (2011) provide evidence that tweets posted three of four days earlier that are classified as "Calm" are Granger-causal of the closing value of the DJIA, while Zhang et al. (2011) show that the percentage of tweets that can be classified as either having a "Fear" or "Hope" sentiment are negatively correlated with the Dow Jones, NASDAQ and S&P500.

[1] Bollen et al. (2011) use the Google Profile of Mood States (GPOMS) classification: Calm, Alert, Sure, Vital, Kind and Happy; while Zhang et al. (2011) use a simple Hope/Fear classification.

More complex analyses like Liu et al. (2015) show that Twitter-specific metrics are useful to measure comovement of shares prices even better than grouping based on industrial category and, Vu et al. (2012) uses not only sentiment analysis of tweets but also consumer confidence on the brand and change in share prices to predict the ups and downs in the share prices of Apple, Google, Microsoft and Amazon with at least 75% accuracy.

Altogether, existing literature indicates that even without a clear consensus on whether or not something else besides news can drive stock market prices, looking at social media platforms to extract leading indicators of commercial and economic activity may be something worth studying in more detail.

Whit this in mind, this work is, to the best of my knowledge, the first study of its kind for the BMV. Therefore, in addition to the contributions previously mentioned, this work will also expand the existing literature to the analysis of a developing country context.

Chapter 3

Data and Strategy

3.1 Sentiment Analysis

The sentiment analysis tool[2] used is called “Mexican tweeter’s mood” and was developed by INEGI. This tool is meant to serve as a first step towards using non-conventional data sources to generate experimental statistics. Being more specific, this tool classifies all tweets posted from Mexico[3] according to its sentiment. To do this, it uses a Machine Learning algorithm to classify text as either having a positive or negative attitude.

The tool feeds from tweets emitted after January 1st, 2016 up to the present day, and according to INEGI it successfully classifies[4] 80% of the tweets evaluated. This is a significant detail for this research because I want to capture the collective mood of Twitter; thus I require the tool to be able to accurately depict and adapt to social and cultural events in Mexico.

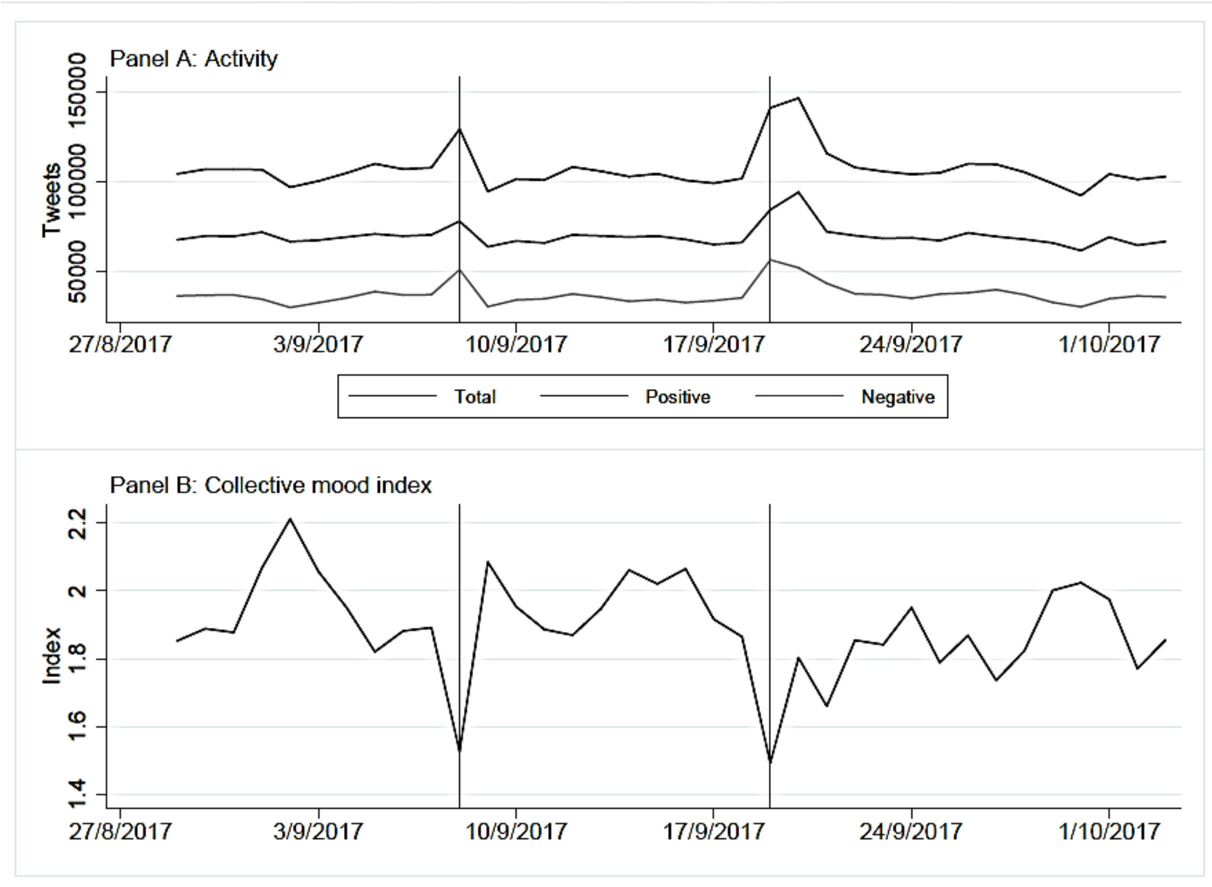
To verify this accuracy, I conduct an analysis of Twitter activity during September 2017. I choose this period because in this month Mexico experienced two strong earthquakes, which gives me an ideal scenario to test the ability of the Machine Learning algorithm to capture and reflect immediate and unpredictable changes in collective mood.

Figure 3.1 shows the overall activity of Mexican tweeters from August 29th to October 3rd, 2017. Panel A shows that both positive and especially negative tweets increased when the earthquakes happened.

---

[2] Available at www.beta.inegi.org.mx/app/animotuitero.
[3] All tweets whose location is identified as within Mexico.
[4] Makes the same classification a human would do.

Figure 3.1: Twitter activity during September, 2017.



Note: The vertical lines represent the September 7 and September 19 earthquakes.

Meanwhile, Panel B depicts the collective mood index, defined as the ratio of positive tweets to negative tweets. Here we observe that indeed Twitter turns more negative on September 7 and September 19, but after a couple days, the social mood seems to return to its normal levels.

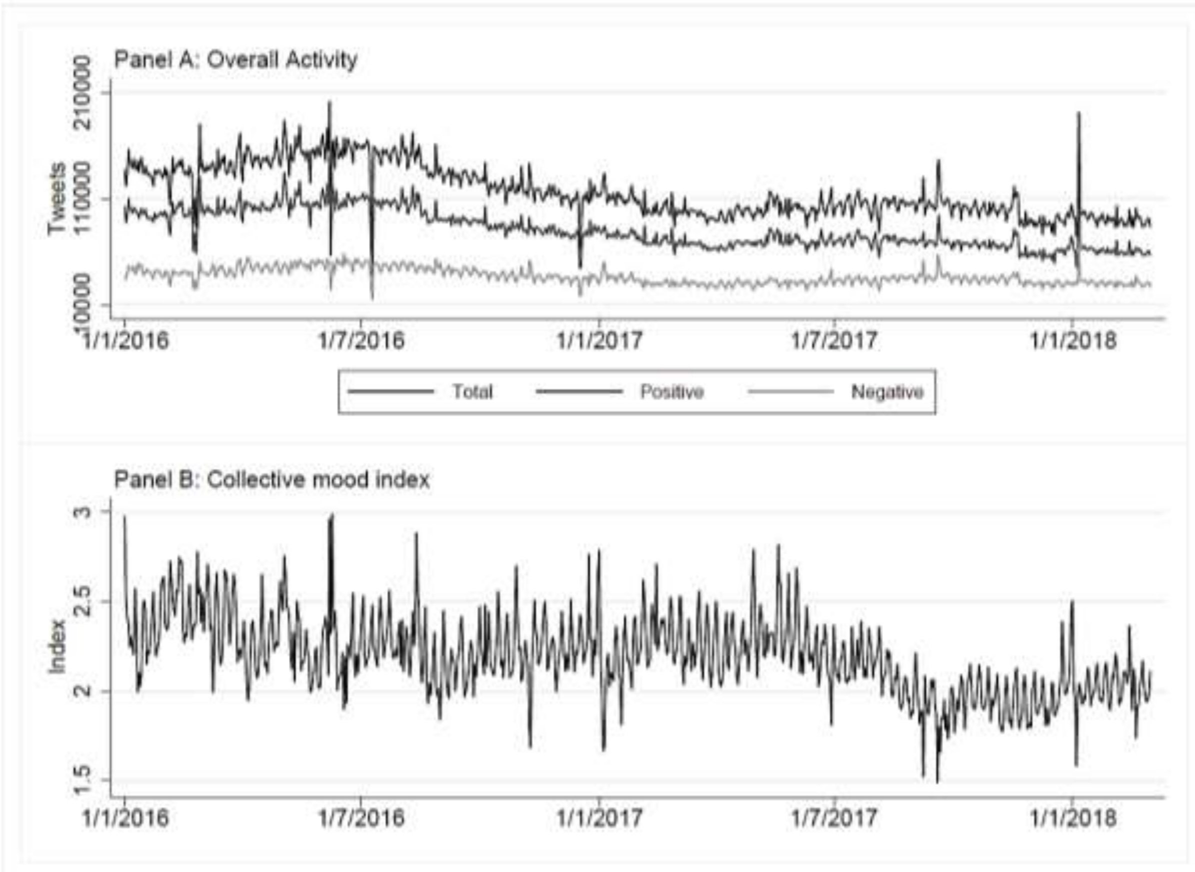
Table 3.1: Twitter Descriptive Statistics.

Variable	Mean	Std. Dev.	Minimum	P25	P50	P75	Maximum
Total	116,861.71	24,057.49	52,992	98,004	109,559.50	136,246	201,407
Positive	80,283.97	17,441.23	37,755	66,835	75,325.50	94,935.50	150,532
Negative	36,577.75	7,221.40	15,237	30,831.50	35,335	41,325	62,862
Index	2.197	0.220	1.492	2.057	2.188	2.341	2.987

Note: Twitter daily activity in Mexico. P25, P50 and P75 represent the 25th, 50th and 75th percentile, respectively.

In this research, I will use all tweets posted from Mexico between January 1st, 2016 and March 3rd, 2018. Table 3.1 shows the descriptive statistics of the data and Figure 3.2 depicts similar information to that in Figure 3.1 for the whole period considered.

Figure 3.2: Twitter Activity in Mexico.



Note that of all the tweets posted from Mexico, nearly two-thirds are classified as positive. Thus, on an average day, there are twice as many positive tweets as negative tweets. Panel B of Figure 3.2 confirms this as the collective mood index oscillates between 2 and 2.5 most of the time.

3.2 Defining Twitter variables

In order to use the information coming from Twitter, I must match it with the BMV schedule. To do this, I exploit the fact that INEGI's tool captures the hour range in which any tweet is posted. This way I can identify correctly all the information that may be relevant to the financial series.

The BMV operates from 8:30 a.m. to 3:00 p.m., but since the tool I am using is not able to capture the exact time a tweet is posted I will assume that the BMV starts operating at 8:00 a.m. This, the IPC and AZTECA closing values and volume of trading is determined at 3:00 p.m., and its opening values are determined at 8:00 a.m.

With this in mind, I spit any day into two blocks that contain all the relevant tweets for our variables: working house (WH, from 8:00 a.m. to 3:00 p.m.) and non-working hours (HWH, from 3:01 p.m. to 7:59 a.m.). Table 3.2 describes all the variables constructed for the analysis.

Table 3.2: Twitter's Variables definition.

Panel A: Close and Volume

Variable	Tweet type	Time window
Index WH	Total tweets	8:00 a.m. to 14:59:59 p.m.
Tweets WH	Total tweets	8:00 a.m. to 14:59:59 p.m.
Positive WH	Positive tweets	8:00 a.m. to 14:59:59 p.m.
Negative WH	Negative tweets	8:00 a.m. to 14:59:59 p.m.
Index 24	Total tweets	3:00 p.m. to 14:59:59 p.m.
Tweets 24	Total tweets	3:00 p.m. to 14:59:59 p.m.
Positive 24	Positive tweets	3:00 p.m. to 14:59:59 p.m.
Negative 24	Negative tweets	3:00 p.m. to 14:59:59 p.m.

Panel B: Open

Variable	Tweet type	Time window
Index NWH	Total tweets	3:00 p.m. to 7:59:59 a.m.
Tweets NWH	Total tweets	3:00 p.m. to 7:59:59 a.m.
Positive NWH	Positive tweets	3:00 p.m. to 7:59:59 a.m.
Negative NWH	Negative tweets	3:00 p.m. to 7:59:59 a.m.
Index 24	Total tweets	8:00 a.m. to 7:59:59 a.m.
Tweets 24	Total tweets	8:00 a.m. to 7:59:59 a.m.
Positive 24	Positive tweets	8:00 a.m. to 7:59:59 a.m.
Negative 24	Negative tweets	8:00 a.m. to 7:59:59 a.m.

Note: WH and NWH represent working and non-working hours of the BMV.

Notice that while it is possible to construct NWH variables to analyze the closing values and volume of trading, there is no reason to believe that the information contained in tweets posted in those hours will be more relevant than that contained in the WH variables. Something similar applies to WH variables for the opening values, where the content of the tweets posted in the 16 hours preceding the opening of the BMV (and following the close of the market in the previous day) should be more relevant than the content of the tweet of the last cycle of working hours.

3.3 The Mexican Stock Market

The IPC is the leading index in the Mexican stock market and is composed of the 35 most important stock in the BMV. Therefore, the IPC serves as an indicator of the overall performance of the Mexican stock market, just as de DJIA or the NASDAQ indices.

I will also analyze TV Azteca's share (AZTECA) performance on the BMV. I choose this firm because it is a large firm on the Mexican telecommunication industry (but not large enough to be included in the IPC[5]). Also since AZTECA is in the telecommunication industry, it has a strong presence on social media across all of the accounts related to the firm[6] with nearly 5.2 million followers. This is relevant because it is possible that the information contained in Twitter to be more relevant for firms with a strong presence on social media.

As with Twitter, I will consider daily data from January 1st, 2016 to March 3rd, 2018. I will focus on the opening and closing[7] values and on the volume of daily transactions of the IPC and AZTECA.

Table 3.3 shows the descriptive statistics of the relevant variables for both IPC and AZTECA, and Figures 3.3 and 3.4 depict the time series.

Note that for the IPC on an average day the closing value is higher than the opening value and that over the period considered the IPC had been steadily increasing in value. AZTECA displays a very different performance. On average, the opening value is higher than the closing value, and over the period considered the share of AZTECA has had relatively big increases in its price (like the one registered in July 2016) but also relatively big decreases in its price (like the one registered in early 2018.)

---

[5] AZTECA's daily volume of trading is generally not large enough to grant it a spot on the IPC.

[6] These accounts include the main account, but also separate accounts for their news and sports programs, both of the TV channels the firm owns, etc.

[7] Since the IPC do not pays dividends the closing value and adjusted closing values are equal, so I will consider regular closing value. For the AZTECA share I will consider adjusted closing values, but will refer to it simply as closing value.

Table 3.3: Financial Series Descriptive Statistics.

Panel A: IPC

Variable	Mean	Std. Dev.	Minimum	P25	P50	P75	Maximum
Open	47,461.38	2,376.12	40,254.50	45,705.52	47,650.24	49,231.61	51,590.48
Close	47,464.66	2,367.23	40,265.37	45,707.87	47,661.69	49,209.58	51,713.38
Volume	195.663	8.674	0	145.365	181.161	225.710	645.087

Panel B: AZTECA

Variable	Mean	Std. Dev.	Minimum	P25	P50	P75	Maximum
Open	3.06	0.48	2.03	2.70	3.23	3.44	3.97
Close	3.05	0.49	2.01	2.70	3.24	3.44	3.94
Volume	1.794	3.359	0.024	0.442	0.954	2.560	70.488

Note: P25, P50 and P75 represent the 25th, 50th and 75th percentile, respectively. Statistics for daily volume of trading are shown in millions.

Figure 3.3: IPC Time Series.

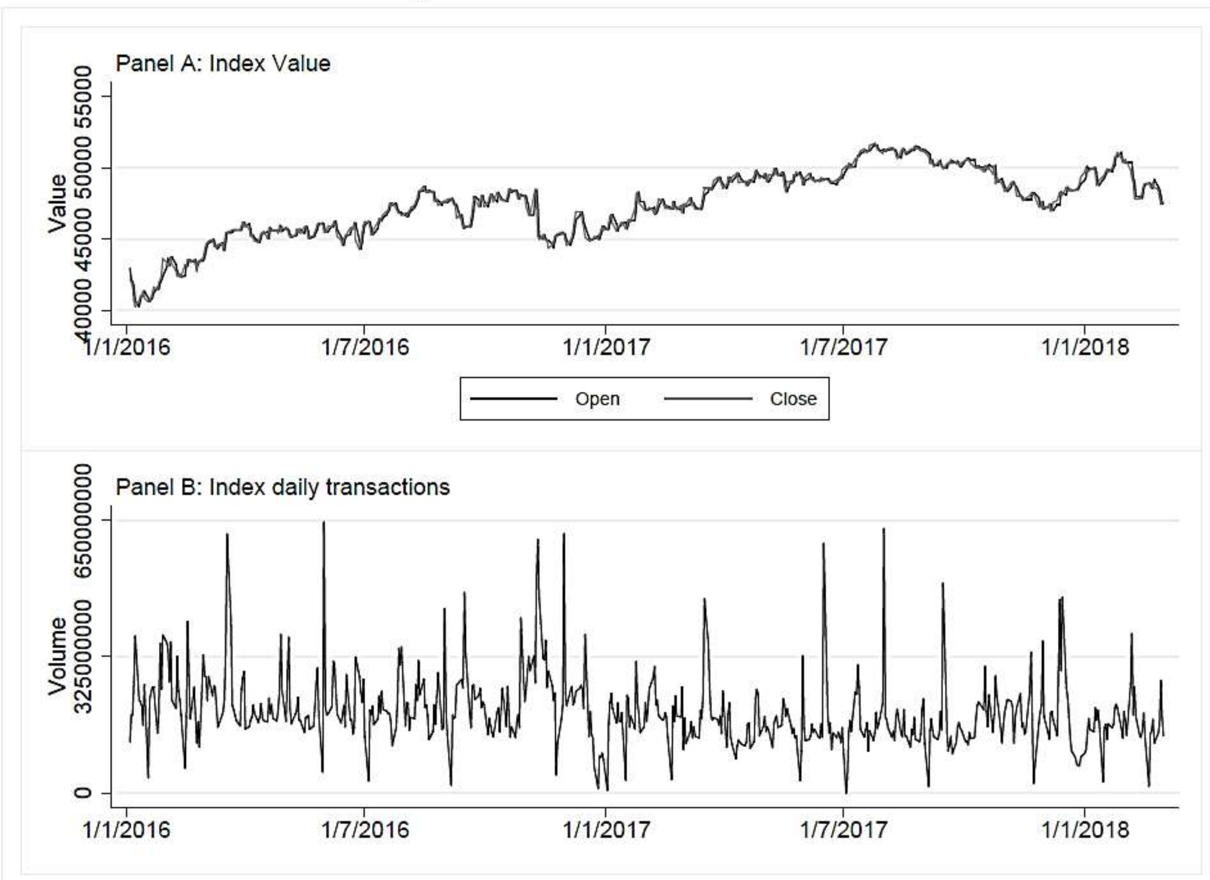


Figure 3.4: AZTECA Time Series.

3.4 Empirical Strategy

The econometric test I will use is the Granger Causality Test. This test allows me to determine whether X causes Y in the sense of how much of the current value of Y is explained by a regression of lagged values of Y and X versus a regression including only lagged values of Y.

Notice that a Granger causality does not mean that Y is a direct consequence of X. It is more a test of the relevance for Y of past information contained in X than a test for causality in a traditional sense.

To run the test one must first adjust the following two VAR models:

$$Y_t = \alpha_0 + \sum_{l=1}^{M} \alpha_l Y_{t-l} + \sum_{l=1}^{M} \beta_l X_{t-l} + \varepsilon_t \tag{3.1}$$

$$X_t = \alpha_0 + \sum_{l=1}^{M} \alpha_l X_{t-l} + \sum_{l=1}^{M} \beta_l Y_{t-l} + v_t \qquad (3.2)$$

The number of lags to be considered, M, is determined in accordance with the information criteria. Here I will use the Akaike (AIC), Hannan-Quinn (HQIC) and the Schwarz (SBIC) information criteria as guidelines, choosing the value that the majority of the criteria indicate[8].

Once both VAR models indicated in equations (3.1) and (3.2) are done, the next step is to run a Wald test for the joint hypothesis $\beta_1 = \beta_2 = \cdots = \beta_M = 0$ for both models, which will inform me of the statistical significance if the lagged values of the variables believed to be Granger-causal (the variables associated with the $\beta's$ coefficients on the VAR models) conditional on the lagged values of the other variable.

Note also that the test is to be run for "X is Granger causal of Y" as well for "Y is Granger causal of X". This is useful because, as stated before, the direction of the causality is not clear beforehand.

[8] If all the criteria indicate a different number of lags, then the SBIC lag value will be selected.

Chapter 4

Results

4.1 Granger Causality: Twitter and the BMV

Before running the VAR models, it must be verified that all the variables are stationary or integrated of the same order, and the number of lags to be considered in each model must also be determined.

To check for stationarity, I run a Dickey-Fuller unit root test (Results shown in Table A1 in the Appendix). All the variables considered are stationary at the 99% level, except for the opening and closing values of both IPC and AZTECA, so first differences are to be considered from now on for this variables.

As for the number of lags to be considered in each model, Tables A3 and A4 in the Appendix show the lags selected by the different criteria as well as the selected lag value. For example, row 2 in Panel A of Table A3 shows that for the VAR model that includes lagged values of the “Index 24” variable, AIC indicates four lags, whereas both HQIC and SBIC indicate two lags, so for the model I consider only two lags.

Similarly in row 2 in Panel B of Table A4, since all the criteria indicate a different number of lags to be selected, the SBIC value is chosen because this criteria is the one that penalized the most the inclusion of a higher number of lags.

Once every VAR model is specified (results not shown) a Wald Test is run to test whether or not the addition of the lagged values of a variable X(the variable X is the one coming from Twitter, but as mentioned before the test is performed for both directions of the causality) can improve the explicative power of a model for the current value of Y that contains only lagged values of itself.

4.1.1 Can Twitter help predict closing values in the BMV?

Tale 4.1 shoes the p-values of the Wald Test. Column (1) displays the results of the test applied to the VAR model specified in equation (3.1), i.e. this column shows the results of whether the indicated variable Granger-causes the closing values in the BMV. Whereas column (2) displays the results of the test applied to the VAR model specified in equation (3.2), i.e. this column shows the result of whether the closing values in the BMV Granger-causes the indicated variable.

Results show that there is no Granger-causality in either direction for any of the variables constructed. This may be due to several factors: for the variables containing all the tweets posted in the last twenty-four hours it may be possible that they are way too noisy to accurately reflect or mimic the performance of the IPC or AZTECA, as for the variables that contain all the tweets

Table 4.1: Granger causality for closing values in the BMV.

Panel A: IPC

Variable	Lags included	(1)	(2)
Index 24	1	0.51961	0.23243
Tweets 24	2	0.82053	0.70405
Positive 24	2	0.78503	0.73062
Negative 24	3	0.18574	0.43519
Index WH	1	0.46042	0.37892
Tweets WH	2	0.74931	0.65693
Positive WH	1	0.56580	0.96691
Negative WH	2	0.91751	0.23085

Panel B: AZTECA

Variable	Lags included	(1)	(2)
Index 24	2	0.40996	0.20063
Tweets 24	2	0.54448	0.72409
Positive 24	2	0.48153	0.91012
Negative 24	3	0.69504	0.68878
Index WH	3	0.13469	0.33732
Tweets WH	2	0.16532	0.55188
Positive WH	3	0.18475	0.87174
Negative WH	2	0.24627	0.16653

Note: p-values of the Wald Test. Column (1) are the results of the test applied to the VAR model specified on equation (3.1), and column (2) are the results of the test applied to the VAR model specified in equation (3.2). Lags included in accordance with tables A3 and A4. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

posted between 8:00 a.m. and 3:00 p.m. it is possible that they do not capture enough information on the social mood to reflect the closing values of the IPC or AZTECA.

Another possibility is that neither the information contained in tweets posted in working hours nor a full twenty-four hour window are adequate to capture the effect of the social mood on the closing values of the stock market, i.e. it is possible that the relevant tweets may be posted in a time window smaller than eight hour window considered in the WH variable or in a window somewhere in between eight and twenty-four hours.

Another possibility is that working with all the tweets posted from Mexico is not ideal. This may ve if, for example, tweets coming from states with social problems such as Oaxaca and Guerrero are far more negative than those coming from the rest of the country; or if tweets coming from touristic states are much more positive than those from other states. This possibility is explored latter in this chapter by restricting the Twitter data to the five states with the most Internet users.

4.1.2 Can Twitter help predict the volume of trading in the BMV?

Table 4.2 shows the results of the Wald Test conducted over daily volume of trading. Again, column (1) displays the results of the test applied to the VAR model specified in equation (3.1), and column (2) displays the results of the test applied to the VAR model specified on equation (3.2).

The first thing to note is that all the statistically significant Granger-causalities run from Twitter to the BMV and not in the opposite direction. Panel A shows that the total number of tweets posted in the las forty-eight hours can help in predicting the number if daily transactions of the IPC. Note also that the positive tweets posted in the last forty-eight hours are Granger-causal of the volume of trading, but negative tweets are not. This suggest that the explicative power of the tweets series is due to the information contained tweets classified as positive. This hypothesis is supported by the results for the WH variables, in which only the positive tweets posted between 8:00 a.m. and 3:00 p.m. seem to provide us with useful information.

Results for AZTECA (shown in Panel B) are very similar to those obtained for the IPC except that here both Index 24 and Index WH are also Granger-causal of the volume of trading. Once again this effect seems to be driven by the content of the positive tweets.

Table 4.2: Granger causality for volume of trading in the BMV.

Panel A: IPC

Variable	Lags included	(1)	(2)
Index 24	1	0.08074*	-
Tweets 24	2	0.00792***	0.35132
Positive 24	2	0.00990***	0.32927
Negative 24	3	0.30536	0.19434
Index WH	2	0.49702	-
Tweets WH	2	0.07749*	0.18092
Positive WH	1	0.00229***	0.14502
Negative WH	2	0.14322	0.13120

Panel B: AZTECA

Variable	Lags included	(1)	(2)
Index 24	1	0.00000***	0.08821*
Tweets 24	2	0.00674***	0.15820
Positive 24	2	0.00211***	0.06635*
Negative 24	3	0.92227	0.92408
Index WH	1	0.00000***	0.09309*
Tweets WH	3	0.39542	0.68345
Positive WH	1	0.00000***	0.23441
Negative WH	3	0.92810	0.70386

Note: p-values of the Wald Test. Column (1) are the results of the test applied to the VAR model specified on equation (3.1), and column (2) are the results of the test applied to the VAR model specified in equation (3.2). Lags included in accordance with tables A3 and A4. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

These findings are consistent with the literature mentioned in Chapter 2, providing more evidence that a simple collective mood measurement derived from social media can be useful for the economic and financial analysis to some extent.

Specifically, a binary categorization of collective mood like "Fear/Hope" classification of Zhang et al. (2011) or the "Positive/Negative" classification used in this work seems more than capable of providing useful information for predicting the volume of trading in the stock market. It is still possible, however, than a more complex classification like the one in Bollen et al. (2011) can be useful for more intricate analysis. This possibility could be explored in future research.

4.1.3 Can Twitter help predict opening values in the BMV?

Table 4.3 shows the results of the Wald Test for the opening values of the IPC and AZTECA. In Panel A, column (1) display the results of the test applied to the VAR model specified in equation (3.1), and column (2) display the results of the test applied to the VAR model specified in equation (3.2).

Table 4.3: Granger causality for opening values of the BMV.

Panel A: IPC

Variable	Lags included	(1)	(2)
Index 24	2	0.18027	0.06122*
Tweets 24	3	0.21685	0.35143
Positive 24	3	0.14967	0.31013
Negative 24	3	0.59483	0.12490
Index NWH	2	0.18145	0.06756*
Tweets NWH	3	0.13731	0.20742
Positive NWH	3	0.12127	0.20679
Negative NWH	3	0.39137	0.06553*

Panel B: AZTECA

Variable	Lags included	(1)	(2)
Index 24	2	0.31234	0.94426
Tweets 24	3	0.01693***	0.80448
Positive 24	3	0.01387***	0.78587
Negative 24	3	0.02446**	0.89080
Index NWH	2	0.60637	0.97691
Tweets NWH	2	0.29698	0.28617
Positive NWH	3	0.01509**	0.71983
Negative NWH	3	0.2392**	0.85993

Note: *p*-values of the Wald Test. Column (1) are the results of the test applied to the VAR model specified on equation (3.1), and column (2) are the results of the test applied to the VAR model specified in equation (3.2). Lags included in accordance with tables A3 and A4. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Notice that as with the closing values it is not possible to determine if there is any Granger-causality in either direction for the IPC.

For AZTECA the tweets posted in the last seventy-two hours are Granger-causal of the opening value, but this effect cannot be attributed exclusively to either positive or negative tweets since both variables are statistically significant when analyzed individually.

4.2 Restricting the information coming from Twitter

As mentioned before, it is possible that considering tweets from some states may make the results less accurate. O address this concern, I use the National Availability and Use of Information Technologies in Households (ENDUTIH) to determine the five states with the most Internet users in 2016 and 2017. These states are Mexico, Mexico City, Jalisco, Nuevo Leon and Veracruz for both years.

An important consequence of restricting our analysis to only five states is that, since Machine Learning algorithms can be very demanding in computing power, if the results are similar to those obtained with all the tweets, then this will be evidence that it suffices to focus only on the Twitter activity of these five states.

Results of the analysis for the IPC are shown in Table 4.4.

Panel A shows the results for the closing values. With the restricted sample it is not possible to say that any of the variables construed is Granger-causal of the closing values.

Panel B shows that all the variables (except the indices of collective mood) are Granger-causal of the daily volume of trading. This means that even when only tweets coming from the five states with the most Internet users are considered the results in Table 4.2 hold, except that now the explicative power cannot be solely attributed to the positive tweets since the negative tweets also show significant results.

Finally, Panel C shows the results for the opening values. In contrast with the results shown in Table 4.3, Table 4.4 show that both the Index of social mood and the Negative tweets posted during non-working hours are Granger-causal of the opening values. This may be due to the fact that the states being considered are not only the ones with the most Internet users, but also the one most commonly associated with economic activity in Mexico. Thus it is not so farfetched to think that the content of the negative tweets (which affects the social mood index) maw have an impact on opening values for the IPC.

Table 4.5 shows the results of the test (with the restricted sample) for AZTECA. Panel A shows that, as in Table 4.1, there are no statistically significant Granger-causalities for the closing values. As for Panel B, once again the tweets posted in the last forty-eight hours (especially the positive ones) and the tweets posted in the eight hours prior to the closing of the BMV are Granger-causal of the volume of trading. Panel C shows similar results to those on Table 4.3.

Table 4.4: Granger causality for the IPC (Restricted sample).

Panel A: Closing values (1st differences)

Variable	Lags included	(1)	(2)
Index 24	1	0.34544	0.12549
Tweets 24	2	0.79387	0.45770
Positive 24	2	0.80046	0.49968
Negative 24	2	0.68071	0.12220
Index WH	1	0.26400	0.53951
Tweets WH	1	0.84042	0.75639
Positive WH	1	0.72381	0.89245
Negative WH	2	0.74809	0.43116

Panel B: Volume of trading

Variable	Lags included	(1)	(2)
Index 24	1	0.09663*	-
Tweets 24	2	0.00394***	0.28394
Positive 24	2	0.00489***	0.26581
Negative 24	2	0.00402***	0.35263
Index WH	2	0.17556	-
Tweets WH	1	0.00179***	0.41638
Positive WH	1	0.00162***	0.26208
Negative WH	1	0.00707***	0.58534

Panel C: Opening values

Variable	Lags included	(1)	(2)
Index 24	2	0.13534	0.04532*
Tweets 24	3	0.29095	0.23026
Positive 24	3	0.20276	0.19917
Negative 24	3	0.73606	0.10741
Index NWH	2	0.18311	0.01281**
Tweets NWH	2	0.56241	0.54335
Positive NWH	3	0.15355	0.12891
Negative NWH	3	0.51835	0.04492**

Note: p-values of the Wald Test conducted considering only tweets coming from the 5 states with more Internet users: Mexico City, Jalisco, Mexico, Nuevo Leon, Veracruz. Column (1) are the results of the test applied to the VAR model specified on equation (3.1), and column (2) are the results of the test applied to the VAR model specified in equation (3.2). Lags included in accordance with table A5. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

We can also restrict the information coming from Twitter by exploiting a unique characteristic of the Machine Learning tool developed by INEGI. INEGI's algorithm is capable of identifying whether a tweeter is posting content from its own locality or if he/she is posting from a locality in

Table 4.5: Granger causality for AZTECA (Restricted sample).

Panel A: Closing values (1st differences)

Variable	Lags included	(1)	(2)
Index 24	1	0.08672*	0.82691
Tweets 24	2	0.73930	0.64147
Positive 24	2	0.64544	0.82541
Negative 24	3	0.77035	0.79648
Index WH	3	0.13739	0.59215
Tweets WH	2	0.21161	0.52267
Positive WH	2	0.25925	0.79638
Negative WH	2	0.23603	0.21243

Panel B: Volume of trading

Variable	Lags included	(1)	(2)
Index 24	1	0.00000*	0.06549*
Tweets 24	2	0.00941***	0.22734
Positive 24	2	0.00343***	0.10586
Negative 24	3	0.80141	0.88720
Index WH	1	0.00011***	0.13833
Tweets WH	1	0.00000***	0.32229
Positive WH	1	0.00000***	0.27727
Negative WH	3	0.79672	0.72487

Panel C: Opening values

Variable	Lags included	(1)	(2)
Index 24	2	0.17576	0.98843
Tweets 24	3	0.02301**	0.78264
Positive 24	3	0.01739**	0.73198
Negative 24	3	0.03804**	0.92589
Index NWH	2	0.42705	0.98402
Tweets NWH	2	0.40569	0.22113
Positive NWH	3	0.01867***	0.62248
Negative NWH	3	0.03775**	0.84342

Note: p-values of the Wald Test conducted considering only tweets coming from the 5 states with more Internet users: Mexico City, Jalisco, Mexico, Nuevo Leon, Veracruz. Column (1) are the results of the test applied to the VAR model specified on equation (3.1), and column (2) are the results of the test applied to the VAR model specified in equation (3.2). Lags included in accordance with table A6. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

which he/she is a tourist. This feature is useful because it is possible that tourists are consistently more positive than locals, thus biasing the collective mood upwards.

To address this possibility, I use only the tweets which are identified as posted by a local from August 1st, 2016 to March 2nd, 2018. The results (shown in Tables A7 y A8) are very similar to the ones obtained so far.

4.3 How useful is the information coming from Twitter?

So far I have provided evidence that the information coming from Twitter is useful to predict the performance of the BMV, but I have not said anything about how useful is this information. To asses this, a model that accurately captures how the collective mood state affects the BMV must be built. According to Bollen et al. (2011), the relationship between collective mood and stock market performance is not linear; thus a simple VAR model cannot accurately model this relationship.

Table 4.6: How useful is the information coming from Twitter?

Panel A: IPC

Variable	Lags	Difference in MAE	Difference in MAPE
Tweets 24	2	-2,386,028.6	-3.166%
Positive 24	2	-2,355,173.1	-3.120%
Positive WH	1	-631,363.8	-0.437%

Panel B: AZTECA

Variable	Lags	Difference in MAE	Difference in MAPE
Index 24	1	-76,258.54	-42.234%
Tweets 24	2	-20,909.16	-11.562%
Positive 24	2	-23,871.40	-13.391%
Index WH	1	-70,586.54	-41.138%
Positive WH	1	-56,684.5	-30.136%

Since the design and implementation of a model that can capture the true nature of this relationship, like a neural network, is beyond the scope of this research, Table 4.6 present the differences in Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) of the statistically significant models constructed to predict the daily volume of trading of the IPC and AZTECA using the full tweets sample when compared with a model that includes only lagged values of past prices. Again, this results do not reflect the actual usefulness of the date coming from Twitter but are presented as a simple benchmark.

Nota that, as expected, all the statistically significant models constructed have a smaller MAE and MAPE than the model that contains only past values of the daily volume of trading, thus confirming that the information coming from Twitter is indeed useful for stock market prediction.

Chapter 5

Conclusions

In this research, I use a simple classification of tweets as either positive or negative to construct several measurements of the social mood and relate said measures to the performance of the leading index in the Mexican stock market, the IPC, and the share of TV Azteca, a large telecommunication firm in Mexico.

Results show that the measurements of the social mood proposed here do not provide useful information to forecast the closing values of the IPC or AZTECA. Nonetheless, these measurements are instructive to forecast the daily volume of trading of both IPC and AZTECA and the opening values of AZTECA.

Being more specific, when I consider all the tweets posted from Mexico, I find that the total number of tweets posted in the forty-eight hours proper to the closing of the stock market, especially those classified as positive, are informative of the volume of trading observed for the IPC and AZTECA. This is also true if I restrict to the positive tweets posted during working hours, i.e. between 8:00 a.m. and 3:00 p.m.

Since Machine Learning algorithms can be quiet demanding in computing power, I explored the possibility of analyzing Twitter feed of only some states and not the whole country. Using the ENDUTIH, I selected the five states with the most Internet users and found that none of the social mood metrics constructed are Granger-causal of the closing value. With this restricted sample I also found that all tweets (positive and negative) posted in the forty-eight hours prior to the closing of the stock market, as well as tweets (positive and negative) posted during working hours, can help predicting the daily volume of trading.

In summary, this research presents evidence that collective mood may indeed affect the decisions made by a society. Since here I only present statistically significant results for the volume of trading, future research should focus on finding similar results for more economically significant variables like closing values or returns. This can be done using a more complex classification (non-binary) of the collective mood or using another social network, especially if said social network contains information related more directly to the financial series that is being analyzed.

Bibliography

Antweiler, W. and Frank, M. Z. (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boeards. *The Journal of Finance*, 5983): 1259-1294.

Bollen, J., Mao, H., and Zheng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1): 1-8.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1): 34-105.

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

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

Fama, E. F., Fisher, L., Jensen, M., and Roll, R. W. (1969). The adjustment of stock prices to new information. *International Economic Review*, 10: 1-21.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3): 424-438.

Gruhl, D., Guha, R., Kumar, R., Novak, J., and Tomkins, A. (2005). The predictive power of online chatter. In *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, KDD '05, pages 78-87, New York, NY, USA. ACM.

Instituto Nacional de Estadística y Geografía. (2007). Estado de ánimo de los tuiteros en los estados unidos mexicanos: documento metodológico. Technical report, INEGI.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2): 263-291.

Kavussanos, M. G. and Dockery, E. (2001). A multivariate test for market efficiency: the case of ASE. *Applied Financial Economics*, 11(5): 573-579.

Lee, m. K., Shi, N., Cheung, C. M., Lim, K. H., and Sia, C. L. (2011). Consumer's decision to shop online: The moderating role of positive informational social influence. *Information & Management*, 48(6): 185-191.

Li, X., Xie, H., Chen, L., Wang, J., and Deng, X. (2014). News impact on stock price return via sentiment analysis. *Knowledge-Based Systems*, 69: 14-23.

Liu, L., Wu, J., Li, P., and Li, Q. (2015). A social-media-based approach to predicting stock comovement. *Expert Systems with Applications*, 42(8): 3893-3901.

Luo, X. and Zhang, J. (2013). How do consumer buzz and traffic in social media marketing predict the value of the firm? *Journal of Management Information Systems*, 30(2): 213-238.

Luo, X. and Zhang, J., and Duan, W. (2013). Social media and firm equity value. *Information Systems Research*, 24(1): 146-163.

Nassirtoussi. A. K., Aghabozorgi, S., Wah, T. Y., and Ngo, D. C. L. (2014) Text mining for market prediction: A systematic review. *Expert Systems with Applications*, 41(16): 7653-7670.

Nguyen, T. H., Shirai, K., and Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert systems with Applications*, 42(24) 9603-9611.

Nofsinger, J. (2005). Social mood and financial economics. *Journal of Behavioral Finance*, 6: 144-160.

Pak, A. and Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. European Languages Resources Association (ELRA).

Qian, B. and Rasheed, K. (2007). Stock market prediction with multiple classifiers. *Applied Intelligence*, 26(1): 25-33.

Rechenthin, M., Street, N., and Srinivasan, P. (2013). Stock chatter: Using stock sentiment to predict price direction. *Algorithmic Finance*, 2: 169-196.

Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2): 41-49.

Schumaker, R. P. and Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The AZFinText System. *ACM Trans. Inf. Syst.*, 27(2): 1-19.

Sul, H., Denis, A. R., and Yuan, L. I. (2014). Trading on Twitter: The financial information content of emotion in social media. In *2014 47th Hawaii International Conference on System Sciences*, pages 806-815.

Tirunillai, S and Tellis, G. J. (2012). Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science*, 31(2): 195-215.

Tumarkin, R. and Whitelaw, R. F. (2001). News or Noise? Internet postings and stock prices. *Financial Analysts Journal*, 57(3): 41-51.

Vu, T. T., Chang, S. Y., Ha, Q., and Colloer, N. (2012). An experiment in integrating sentiment features for tech stock prediction in Twitter. In *Proceedings of the Workshop on Information Extraction and Entity Analysis on Social Media Data*, pages 23-38. Coling.

Yu, Y., Duan, W., and Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems*, 55(4): 919-926.

Zhang, X., Fuehres, H., and Gloor, P. A. (2011). Predicting stock market indicators through Twitter "I hope it is not as bad as I fear". *Procedia – Social and Behavioral Sciences*, 26: 55-62.

Zheludev, I., Smith, R., and Aste, T. (2014). When can social media lead financial markets? *Scientific Reports*, 4(4213): 1-12.

Appendix

Table A1: Unit Root Test.

Close and Volume Variable	DF Statistic	*Open* Variable	DF Statistic
IPC Close	-2.656*	IPC Open	-2.787*
IPC Diff. Close	-17.868***	IPC Diff. Open	-17.773***
IPC Volume	-12.896***	-	-
AZTECA Close	-2.509	AZTECA Open	-2.371
G AZTECA Diff. Close	-23.499***	AZTECA Diff. Open	-25.135***
AZTECA Volume	-42.632***	-	-
Index 24	-11.324***	Index 24	-23.752***
Tweets 24	-89.338***	Tweets 24	-88.926***
Positive 24	-87.302***	Positive 24	-57.427***
Negative 24	-88.402***	Negative 24	-86.965***
Index WH	-9.275***	Index NWH	-35.648***
Tweets WH	-5.657***	Close NWH	-120.720***
Positive WH	-5.228***	Positive NWH	-72.367***
Negative WH	-7.880***	Negative NWH	-116.151***

Note: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A2: Unit Root Test (Restricted Sample).

Close and Volume Variable	DF Statistic	*Open* Variable	DF Statistic
Index 24	-12.652***	Index 24	-21.512***
Tweets 24	-86.396***	Tweets 24	-85.693***
Positive 24	-86.193***	Positive 24	-56.668***
Negative 24	-80.999***	Negative 24	-89.180***
Index WH	-10.372***	Index NWH	-33.440***
Tweets WH	-5.563***	Close NWH	-120.090***
Positive WH	-5.134***	Positive NWH	-73.146***
Negative WH	-7.889***	Negative NWH	-109.173***

Note: Variables constructed using only tweets coming from the 5 states with more Internet users: Mexico City, Jalisco, Mexico, Nuevo Leon, Veracruz. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A3: Lag Selection for IPC's VAR Models.

Panel A: Close (1st differences)

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	2	1	1	1
Tweets 24	4	2	2	2
Positive 24	4	2	2	2
Negative 24	3	3	2	3
Index WH	4	1	1	1
Tweets WH	2	2	1	2
Positive WH	2	1	1	1
Negative WH	2	2	2	2

Panel B: Volume

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	1	1	1	1
Tweets 24	4	2	2	2
Positive 24	4	2	2	2
Negative 24	3	3	2	3
Index WH	3	2	2	2
Tweets WH	2	2	1	2
Positive WH	2	1	1	1
Negative WH	2	2	1	2

Panel C: Open

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	2	2	2	2
Tweets 24	3	3	3	3
Positive 24	3	3	2	3
Negative 24	3	3	2	3
Index NWH	2	2	2	2
Tweets NWH	3	3	2	3
Positive NWH	3	3	2	3
Negative NWH	3	3	3	3

Table A4: Lag Selection for AZTECA's VAR Models.

Panel A: Close (1st differences)

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	2	2	1	2
Tweets 24	3	2	2	2
Positive 24	4	2	2	2
Negative 24	3	3	2	3
Index WH	3	3	2	3
Tweets WH	2	2	2	2
Positive WH	2	2	1	2
Negative WH	2	2	2	2

Panel B: Volume

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	2	1	1	1
Tweets 24	4	3	2	2
Positive 24	4	3	2	2
Negative 24	3	3	3	3
Index WH	3	1	1	1
Tweets WH	3	3	2	3
Positive WH	3	2	1	1
Negative WH	3	3	2	3

Panel C: Open

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	2	2	2	2
Tweets 24	3	3	2	3
Positive 24	3	3	2	3
Negative 24	3	3	2	3
Index NWH	2	2	2	2
Tweets NWH	4	3	2	2
Positive NWH	3	3	2	3
Negative NWH	3	3	3	3

Table A5: Lag Selection for IPC's VAR Models (Restricted Sample).

Panel A: Close (1st differences)

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	1	1	1	1
Tweets 24	4	2	2	2
Positive 24	4	2	2	2
Negative 24	3	2	2	2
Index WH	4	1	1	1
Tweets WH	2	1	1	1
Positive WH	1	1	1	1
Negative WH	2	2	1	2

Panel B: Volume

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	1	1	1	1
Tweets 24	4	2	2	2
Positive 24	4	2	2	2
Negative 24	3	2	2	2
Index WH	3	2	2	2
Tweets WH	2	1	1	1
Positive WH	1	1	1	1
Negative WH	2	1	1	1

Panel C: Open

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	2	2	2	2
Tweets 24	3	3	3	3
Positive 24	3	3	2	3
Negative 24	3	3	2	3
Index NWH	2	2	2	2
Tweets NWH	4	3	2	2
Positive NWH	3	3	2	3
Negative NWH	3	3	2	3

Note: Models fitted using only tweets coming from the 5 states with more Internet users: Mexico City, Jalisco, Mexico, Nuevo Leon, Veracruz.

Table A6: Lag Selection for AZTECA's VAR Models (Restricted Sample).

Panel A: Close (1st differences)

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	3	1	1	1
Tweets 24	4	2	2	2
Positive 24	4	2	2	2
Negative 24	3	3	2	3
Index WH	3	3	2	3
Tweets WH	2	2	1	2
Positive WH	2	2	1	2
Negative WH	3	2	2	2

Panel B: Volume

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	2	1	1	1
Tweets 24	4	3	2	2
Positive 24	4	3	2	2
Negative 24	3	3	2	3
Index WH	3	1	1	1
Tweets WH	3	2	1	1
Positive WH	3	1	1	1
Negative WH	3	3	1	1

Panel C: Open

Variable	AIC	HQIC	SBIC	Lags selected
Index 24	2	2	2	2
Tweets 24	3	3	2	3
Positive 24	3	3	2	3
Negative 24	3	3	2	3
Index NWH	2	2	2	2
Tweets NWH	4	3	2	2
Positive NWH	3	3	2	3
Negative NWH	3	3	2	3

Note: Models fitted using only tweets coming from the 5 states with more Internet users: Mexico City, Jalisco, Mexico, Nuevo Leon, Veracruz.

Table A7: Granger causality for IPC: Only local tweeters.

Panel A: Closing values (1st differences)

Variable	Lags included	(1)	(2)
Index 24	2	0.37483	0.00196***
Tweets 24	1	0.96512	0.50048
Positive 24	1	0.99920	0.51794
Negative 24	1	0.88440	0.58200
Index WH	1	0.19044	0.09438*
Tweets WH	2	0.79780	0.09288*
Positive WH	1	0.65594	0.38515
Negative WH	2	0.53545	0.00421***

Panel B: Volume of trading

Variable	Lags included	(1)	(2)
Index 24	1	0.57048	-
Tweets 24	1	0.01504**	0.00007***
Positive 24	1	0.01697**	0.00004***
Negative 24	1	0.01245**	0.00089
Index WH	1	0.45563	-
Tweets WH	1	0.04096**	0.23396
Positive WH	1	0.03372**	0.14221
Negative WH	1	0.12851	0.56578

Panel C: Opening values

Variable	Lags included	(1)	(2)
Index 24	1	0.69885	0.17034
Tweets 24	3	0.97206	0.32180
Positive 24	3	0.89140	0.27619
Negative 24	1	0.03028**	0.25696
Index NWH	1	0.25572	0.14636
Tweets NWH	3	0.89234	0.21336
Positive NWH	3	0.88033	0.20325
Negative NWH	1	0.02528**	0.21737

Note: p-values of the Wald Test conducted considering only tweets coming from local tweeters. Column (1) are the results of the test applied to the VAR model specified on equation (3.1), and column (2) are the results of the test applied to the VAR model specified in equation (3.2). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A8: Granger causality for AZTECA: Only local tweeters.

Panel A: Closing values (1st differences)

Variable	Lags included	(1)	(2)
Index 24	2	0.71148	0.01381**
Tweets 24	1	0.88036	0.06366*
Positive 24	1	0.82429	0.04323**
Negative 24	1	0.98805	0.22305
Index WH	2	0.33020	0.22508
Tweets WH	2	0.58209	0.12925
Positive WH	1	0.10953	0.10438
Negative WH	2	0.55765	0.02466**

Panel B: Volume of trading

Variable	Lags included	(1)	(2)
Index 24	1	0.00487***	0.24245
Tweets 24	1	0.31506	0.96863
Positive 24	1	0.40460	0.70451
Negative 24	1	0.16617	0.51568
Index WH	2	0.01012**	0.25534
Tweets WH	2	0.74628	0.98108
Positive WH	1	0.12957	0.67371
Negative WH	2	0.30455	0.85447

Panel C: Opening values

Variable	Lags included	(1)	(2)
Index 24	2	0.63726	0.18622
Tweets 24	1	0.08686*	0.05573*
Positive 24	3	0.00840***	0.33783
Negative 24	1	0.16225	0.11447
Index NWH	1	0.46049	0.65432
Tweets NWH	3	0.01204**	0.40733
Positive NWH	3	0.01312**	0.39718
Negative NWH	1	0.15218	0.06068*

Note: p-values of the Wald Test conducted considering only tweets coming from local tweeters. Column (1) are the results of the test applied to the VAR model specified on equation (3.1), and column (2) are the results of the test applied to the VAR model specified in equation (3.2). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.