EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

Stefano Borgioli, Giampiero M. Gallo, Chiara Ongari

Financial returns, sentiment and market volatility. A dynamic assessment.

No 2999

Abstract

In 1936, John Maynard Keynes proposed that emotions and instincts are pivotal in decision-making, particularly for investors. Both positive and negative moods can influence judgments and decisions, extending to economic and financial choices. Intuitions, emotional states, and biases significantly shape how people think and act. Measuring mood or sentiment is challenging, but surveys and data collection methods, such as confidence indices and consensus forecasts, offer some solutions. Recently, the availability of web data, including search engine queries and social media activity, has provided high-frequency sentiment measures. For example, the Italian National Statistical Institute's Social Mood on Economy Index (SMEI) uses Twitter data to assess economic sentiment in Italy. The relationship between SMEI and financial market activity, specifically the FTSE MIB index and its volatility, is examined using a trivariate Vector Autoregressive model, taking into account the impact of the COVID-19 pandemic.

Keywords: VAR, Granger Causality, sentiment analysis, financial market, forecasting

JEL Codes: C1, C32, C53, G4

1 Non-technical summary

Both positive and negative emotions can impact judgments and economic decisions. Although measuring mood or sentiment is challenging, tools like surveys and confidence indices help. In fact online data, such as search engine queries and social media activity, provide more frequent sentiment measurements. For instance, the Italian National Statistical Institute created the Social Mood on Economy Index (SMEI) using Twitter data to gauge economic sentiment in Italy. This study examines the relationship between the SMEI and financial market activity, particularly the FTSE MIB index, using a model that considers also possible effects of the COVID-19 pandemic.

Two main data sources are used: the SMEI, which measures daily Italian economic sentiment through tweets, and the FTSE MIB, which represents the performance of 40 major Italian stocks and includes volatility data. The analysis uses Vector Auto-regressive Models (VAR) to study the relationships between the SMEI, FTSE MIB returns, and volatility from February 10, 2016, to March 8, 2020. Granger causality tests are then conducted to determine if past values of one variable can predict current values of another, revealing potential bidirectional influences.

"*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 on a mathematical expectation, 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 a result of animal spirits – of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities.*"

John Maynard Keynes, General Theory Of Employment, Interest And Money.

2 Introduction

John Maynard Keynes introduced as early as 1936 the idea that emotions and instincts (the "*animal spirits*") rather than mere rational analysis, play a crucial role in decision making, particularly among investors.

Moods, be they negative or positive, affect judgment and decision-making, even when prompted by unrelated events. This applies to economic and financial decisions, as well: Kahneman (2011) has become the reference of choice on how intuitions, emotional status, and biases shape judgments and affect thinking, behavior, and decisions. Therefore, information is not the only factor at play, and rationality is not always the main engine behind actions.

If the theoretical reasoning is clear, the way mood or sentiment are measured is somewhat of a challenge; in the attempt to isolate an aggregate result, several suggestions are presented in the literature. One option is to conduct surveys and collect data about how economic agents judge the evolution of the general economic conditions or their own. This is the case, for example, of the various confidence indices (both business and consumer) released monthly by most central statistical offices; or also of the various consensus forecast exercises conducted polling several research institutes delivering median forecasts, the spread of opinions, and the changes relative to the previous release of the exercise (cf. Gallo et al., 2002, on the way participants dynamically influence, and possibly bias, each other).

More recently, the widespread availability of information on the web has spurred a host of indicators derived from search engine searches: D'Amuri and Marcucci (2017) is an example using Google searches to build a leading indicator on unemployment (a way to project sentiment about job security). Moreover, the diffusion of social forums has fostered a thriving line of research based on textual analysis of the content of opinions shared, and reactions to economic or market news. This is conducive to examining both the level (akin to confidence) and the change (mood swings), and it has the overwhelming advantage of being available at the daily level, which is most relevant when analyzing financial data.

Twitter (now X) seems to be the natural outlet for this expression of sentiments, helped by a large number of single messages, the possibility of replying to one another, and of classifying the content by the use of "tags". Angelico et al. (2022) document the timeliness and accuracy of deriving a measure of inflation expectations from a massive amount of "tweets" released in Italy (initially, millions that boil down to several hundreds of thousands after processing and cleaning). In general, the ultra-high frequency nature of this textual data offers a very rich pool of information to be extracted (with the awareness that all sorts of manipulation are possible on that forum in directing opinions).

Starting from October 2018, the Italian National Statistical Institute (ISTAT) is publishing a high-frequency index computed in real-time from Italian Twitter's public stream data, the

"Social Mood on Economy Index" (SMEI), providing a daily measure of the sentiment about the Italian economy. The index is calculated on an average of 26,000 tweets per day. This experimental statistic is updated quarterly with the time series starting in February 2016.

Financial market activity is interpreted as an expression of beliefs and sentiments in producing equilibrium prices and returns from trading; by the same token, market volatility (i.e. the variability of returns) can be seen as inversely related to the consensus on how information reaching the market points to the evolution of the market itself. Typically, a downturn in the market is a reaction to bad news and is characterized by high volatility. Since the early 1990s, a market-based measure of volatility extracted from the implied volatilities of put and call options (at the money - 30 days to expiration) on a market index came to be known as the "fear and greed index" (the VIX is derived from the S&P500, Whaley, 1993, but other option-based volatility indices are available).

Financial investment, being driven by profit incentives, is an interesting field in which it is possible to analyze the properties of the SMEI, that is, its capability to represent a relevant factor interacting with financial variables about the stock market activity in Italy. The latter is represented by two variables: the returns on an aggregate index, the Milano Indice di Borsa (FTSE MIB – the benchmark stock market index for the Borsa Italiana, made up of the 40 most-traded stocks), and its volatility (represented by a range-based volatility measure, Garman and Klass, 1980).

After a discussion of the existing literature (Section 3), we discuss the features of the variables used (Section 4), documenting, in particular, the content of available volatility measures and the SMEI (a relatively novel index). The relationship between the SMEI and the market behavior represented by the FTSE-MIB is discussed in Section 5. We suggest (Section 6) a simple trivariate Vector Autoregressive model between three variables (market returns on the index, the volatility and the SMEI) to investigate which variables are in-sample relevant to increase the forecasting capability (a simple Granger-causality test on augmenting the benchmark univariate AR model). We split the analysis between in-sample (Subsection 6.1), and out-of-sample, where we perform a Diebold-Mariano test (Diebold and Mariano, 1995) to assess when the VAR has a superior performance than the AR and for which variables. Some consideration of the impact of the COVID-19 pandemic on these relationships is in order since our in-sample period ends with the wide outbreak of the virus in March 2020.

The main findings can be summarized as follows. The results indicate that before COVID-19, market volatility was the only variable significantly influenced by past values of other variables. However, during the pandemic, the relationships shifted: past volatility influenced both SMEI and returns, past returns impacted volatility, and SMEI also began to affect volatility, though less significantly. This suggests that the pandemic significantly altered the dynamics among these variables. The study uses then the Diebold-Mariano test to compare the predictive abilities of a univariate autoregression (AR) model and a VAR model in an out-of-sample context. By conducting rolling regressions and generating forecasts, the results show that the only variable for which the VAR is predictively superior to the AR model is the range-based volatility, indicating that both lagged SMEI and returns are valuable information for forecasting market activity turbulence. Moving forward, the ISTAT SMEI index, tracking social mood from short messages on social platforms like X, is a valuable tool for understanding market dynamics, especially during unexpected events. The paper suggests also further research into how sentiment indices relate to market returns and volatility and highlights the potential use of other market activity measures, like a VIX-type volatility index, to enhance analysis. Additionally, the unique availability of SMEI data on weekends and holidays raises questions about its impact on market activity at the start of the trading week, which may be worth being addressed.

3 Earlier contributions and issues

John Maynard Keynes, as early as 1936, introduced the idea that emotions and instincts, the famed "animal spirits" , may play a crucial role in decision-making, especially among investors. More recently, studies have explored the potential connections between public sentiment indices and economic and financial variables.

A stream of papers indeed found that sentiment and stock market dynamics can be highly causal related. For instance, Brown and Cliff (2004) examine the relationship between "investor sentiment" and stock market returns. They first build a sentiment measure starting from survey data on investor sentiment (like bullish investor expectations of above-average returns) and using Kalman filter and principal component analysis as means of extracting composite unobserved sentiment measures. They then explore the bi-directional relation between these investor sentiment measures and the near-term stock returns in a vector autoregression (VAR) framework. They find that changes in the composite measures of investor sentiment are highly correlated with contemporaneous market returns, but this correlation does not directly reveal the causal relation between sentiment and the market. Then the VAR analysis reveals that market returns clearly cause future changes in sentiment. However, very little evidence suggests sentiment causes subsequent market returns.

Similar results are displayed in Wang et al. (2006), who test whether sentiment is useful for volatility forecasting purposes. In fact, they find that most of the sentiment measures they use are caused by returns and volatility rather than vice versa. In addition, they find that lagged returns cause volatility. Finally, all sentiment variables have extremely limited forecasting power once returns are included as a forecasting variable. Tetlock (2007) explores instead how media content influences investor sentiment and, consequently, stock market movements by using daily content from a popular Wall Street Journal column. He found that high media pessimism predicts downward pressure on market prices followed by a reversion to fundamentals, and unusually high or low pessimism predicts high market trading volume.

In the same vein, Gilbert and Karahalios (2010) show how estimating emotions from weblogs provides novel information about future stock market prices. From a dataset of over 20 million *LiveJournal* posts, they construct a metric of anxiety, worry and fear called the Anxiety Index. Using then a Granger causality framework, they find that increases in expressions of anxiety predict downward pressure on the S&P 500 index. These findings are then confirmed via Monte Carlo simulations and show how the mood of millions in a large online community, even one that primarily discusses daily life, can anticipate changes in a seemingly unrelated system. Zhang et al. (2011) explore how sentiment and activity on Twitter can be leveraged to understand investor behavior and predict stock market indicators from Dow Jones, NASDAQ, and S&P 500. Starting from a randomized sample of tweets they measured daily "collective hope and fear" and analyzed then the correlation between these indices and the stock market indicators. The analysis shows that Twitter sentiment is significantly correlated with stock market movements. Positive sentiment on Twitter is often associated with rising stock prices, while negative sentiment correlates with declining prices. The volume of Twitter activity is also found to be a useful predictor, with higher tweet volumes indicating increased market attention and potential volatility. Also Bollen et al. (2011) start from Twitter to investigate whether public mood states derived from feeds are correlated to the value of the Dow Jones Industrial Average (DJIA) over time. They analyzed over 9.8 million tweets from 2.7 million users over six months to assess the sentiment of each tweet as either positive or negative. Then, a Granger causality analysis and a Self-Organizing Fuzzy Neural Network are used to investigate the hypothesis that public mood states are predictive of changes in the DJIA closing values. The authors find that the inclusion of certain public mood dimensions indeed improves the accuracy of standard stock market prediction models. Twitter data are used also by Rao and Srivastava (2012) to investigate how sentiment analysis can be employed to predict stock market movements. The study, based on more than 4 million tweets between June 2010 to July 2011, finds that there is a significant correlation between Twitter sentiment and discussions and stock market movements. Positive sentiment is often associated with rising stock prices, while negative sentiment correlates with falling prices. The volume of tweets is also found to be a useful predictor, with higher volumes indicating greater market attention and potential volatility. Da et al. (2015) instead avail of daily Internet search volume from millions of households to investigate the relationship between investor market-level and asset prices, particularly how fear-based sentiment impacts market dynamics. By aggregating the volume of queries related to household concerns they construct a Financial and Economic Attitudes Revealed by Search (FEARS) index as a new measure of investor sentiment. The FEARS index was then found to have significant predictive power regarding future market returns. In fact, between 2004 and 2011, they found that FEARS (i) predict short-term return reversals, (ii) predict temporary increases in volatility, and (iii) predict mutual fund flows out of equity funds and into bond funds. A different approach is applied by Aggarwal and Mohanty (2018) who make use of principal component analysis (PCA) to build sentiment index as a proxy for Indian stock market sentiments over a time frame from April 1996 to January 2017. Three types of variables enter the calculation of the index: indirect market measures (mostly indicators like for instance price to earning ratios, dividend yields or price to book ratios) and Indian and US macro variables. The index is then used to estimate via OLS regressions the impact of Indian investor sentiments on contemporaneous stock returns of Bombay Stock Exchange, National Stock Exchange and various sectoral indices. The study finds that there is a significant positive correlation between the sentiment index and stock index returns. Chen et al. (2019) investigate whether sentiment analysis of social media posts could be used to predict the direction of stock price movements. The authors apply seven different techniques of data mining to predict stock price movement of Shanghai Composite Index for the period April 2016 to May 2018. The findings suggest that sentiment analysis of social media posts could provide valuable insights into the potential direction of stock price movements; for instance sentiment derived from *Eastmoney*, a social media platform for the Chinese financial community, further enhances model performances.

Nyman et al. (2021) investigates the influence of news and narratives in financial systems, especially in utilizing big data for evaluating systemic risk. The paper examines the application of textual analysis techniques and big data analytics to extract valuable insights from news and narrative sources, assisting in identifying and assessing systemic risks within financial systems. Their results highlight how our measures of sentiment and narrative consensus correlate well with, and in some cases even appear to 'cause', certain economic and financial variables. Also Huang et al. (2019), use computational text analysis technique to construct sentiment indices for 20 countries from 1980 to 2019. The authors then assess whether these sentiment indices trigger early warning indicators (EWIs) ahead of financial crises. For each sentiment index, an EWS is triggered each time there is a spike, i.e. when the index value is above 2 standard deviations from a backward-looking average of 24 months. They find that, for each country in our sample, at least one of the indicators would have successfully anticipated most crises in a window of 24 months. As regards techniques to analyse text data, Loughran and McDonald (2011) have investigated how textual analysis techniques are utilized to interpret financial reports. The authors particularly focus on a large sample of 10-K filings submitted to the Securities and Exchange Commission (SEC) from 1994 to 2008. Relevant to our study, they also find significant relations between the sentiment measures they build and economic and financial variables; for instance, they found that some measures are significant in regressions estimating abnormal shares trading volumes.

Turning to the statistical properties of the SMEI index, Righi et al. (2020) analyze the relationships of this metric with some daily and monthly macroeconomic indicators coming from traditional and non-traditional sources. They use several non-traditional sources to produce time series to relate to the SME, such as the daily number of COVID-19 deaths and new positive cases reported by the Civil Protection Department or macroeconomic indicators coming from Target2 and BI-COMP series (on POS and ATM transactions), but also the Bank of Italy electronic card transaction and e-commerce transaction monthly series and the consumer confidence indicators. They found that the monthly average of the daily series of the level of the SME index shows a low contemporaneous correlation and a weak predictive power of the SME index for the traditional monthly indicators. On the other hand, they observed a positive correlation between the SMEI and the BI-COMP POS daily transaction series.

Authors	Period	Markets-included	Sentiment Index	Method	Results
Gilbert and Karahalios (2010)	2008	S&P 500	”Anxiety Index” (an aggregate measure of anxiety and worry) estimated starting from a dataset of over 20 million posts made on the social networking service LiveJournal. (daily frequency)	Granger causality and Monte Carlo simulations	The paper statistically shows that a broad index of mood from an online community has novel predictive information about the stock market.
Rao and Srivastava (2012)	June 2010 - July 2011	DJIA and NASDAQ-100	Twitter feeds (daily frequency)	Granger causality Expert Model Mining System (EMMS)	High correlation (up to 0.88 for returns) between stock prices and Twitter sentiments
Bollen et al. (2011)	February 2008- December 2008	DJIA	Twitter feeds (daily frequency)	Granger causality and Self-Organizing Fuzzy Neural Network	The accuracy of DJIA predictions can be significantly improved by the inclusion of specific public mood dimensions
Zhang et al. (2011)	March 2009 - September 2009	DJIA, NASDAQ-100 and S&P 500	Twitter feeds (daily frequency)	Correlation analysis	Emotional tweet percentage significantly negatively correlated with Dow Jones, NASDAQ and S&P 500, but displayed significant positive correlation to VIX.
Loughran and McDonald (2011)	1994–2008	NASDAQ	Sentiment measures built via textual analysis starting from 10-K report forms (monthly frequency)	Logit regression	The authors find a significant relation between the sentiment measures they build and the economic and financial variables; for instance, they found that some measures are significant in regressions estimating abnormal shares trading volumes
Nyman et al. (2021)			BoE daily markets reports; broker research reports; Thomson-Reuters News archive (daily frequency).	A combination of natural language processing (NLP) and machine learning algorithms.	The study finds that certain narratives, when amplified through news media, can exacerbate systemic risk by influencing market participants’ expectations and behaviors. Shifts in sentiment and the emergence of negative narratives often precede market downturns, suggesting that these textual indicators can serve as early warning signals for systemic risk.
Tetlock (2007)	January 1984 - September 1999	DIJA	Wall Street Journal’s (WSJ’s) “Abreast of the Market” column on U.S. stock market returns (daily frequency).	VAR Estimates	High values of media pessimism induce downward pressure on market prices; unusually high or low values of pessimism lead to temporarily high market trading volume.
Aggarwal and Mohanty (2018)	April 1996 - January 2017	Bombay Stock Exchange, National Stock Exchange, various Indian sectoral indices	PCA to build sentiment index, then OLS (monthly frequency).	A composite sentiment index for Indian stock market, built with principal component analysis, starting from indirect market measures (such as price to earnings ratio, price to book ratio, dividend yield ...) and macro variables for Indian and US markets.	The study finds that there is a significant positive correlation between the sentiment index and stock index returns
Da et al. (2015)	2004-2011	S&P 500, NASDAQ-100, Russell 2000	A measure of investor sentiment (Attitudes Revealed by Search (FEARS)) built by aggregating daily Internet searches from millions of households (daily frequency)	Regression Analysis	The investor sentiment constructed by the authors was found to have significant predictive power regarding future market returns
Huang et al. (2019)	1980 to 2019		News articles from the Financial Times (daily frequency).		
Chen et al. (2019)	April 2017 - May 2018	Shanghai Composite Index.	Sentiment derived from Eastmoney, a social media platform for the financial community in China (daily frequency).	various machine learning models, such as logistic regression, support vector machines (SVM), and neural networks	The study finds that incorporating social media sentiment significantly improves the accuracy of stock price movement predictions. Positive sentiment in social media posts is generally associated with upward stock price movements, while negative sentiment correlates with downward movements. The volume of social media activity also plays a role in prediction accuracy, with higher activity levels often indicating more reliable sentiment signals.
Brown and Cliff (2004)	March 1965 - December 1998	S&P 500, Russell 2000	Two survey-based investor sentiment measures. Indirect sentiment measures such as advance-decline ratio, short interest, and closed-end fund discounts (weekly frequency).	Principal Component, Kalman filter, VAR	The paper finds strong evidence of co-movement between aggregate sentiment measures and the market but little evidence of short-run predictability in returns (using a variety of methods).
Wang et al. (2006)	February 1990 - December 2001	S&P 100	S&P 100 (OEX) put–call trading volume ratio (PCV); OEX put–callopen interest ratio (PCO); NYSE ARMS index. Two survey-based sentiment ratios provided by investment information providers. (daily frequency)	Granger causality / regressions	The authors test whether sentiment is useful for volatility forecasting purposes. In fact they find that most of sentiment measures they use are caused by returns and volatility rather than vice versa. In addition they find that lagged returns cause volatility. Finally, all sentiment variables have extremely limited forecasting power once returns are included as a forecasting variable.

4 A look at the variables involved

This study hinges on two high-frequency data sources. The first data source is the "Social Mood on Economy Index" (SMEI), an experimental index first released by ISTAT (the Italian National Institute of Statistics) in October 2018 with daily values starting on to February 10, 2016. The index provides daily measures of the Italian sentiment on the economy. These measures are derived from samples of Italian public tweets captured in real-time. The production of the index involves the collection and processing of tweets containing at least one word belonging to a specific set of filtered keywords, which has been designed by subject-matter experts. On average, this procedure processes about 26,000 tweets per day.

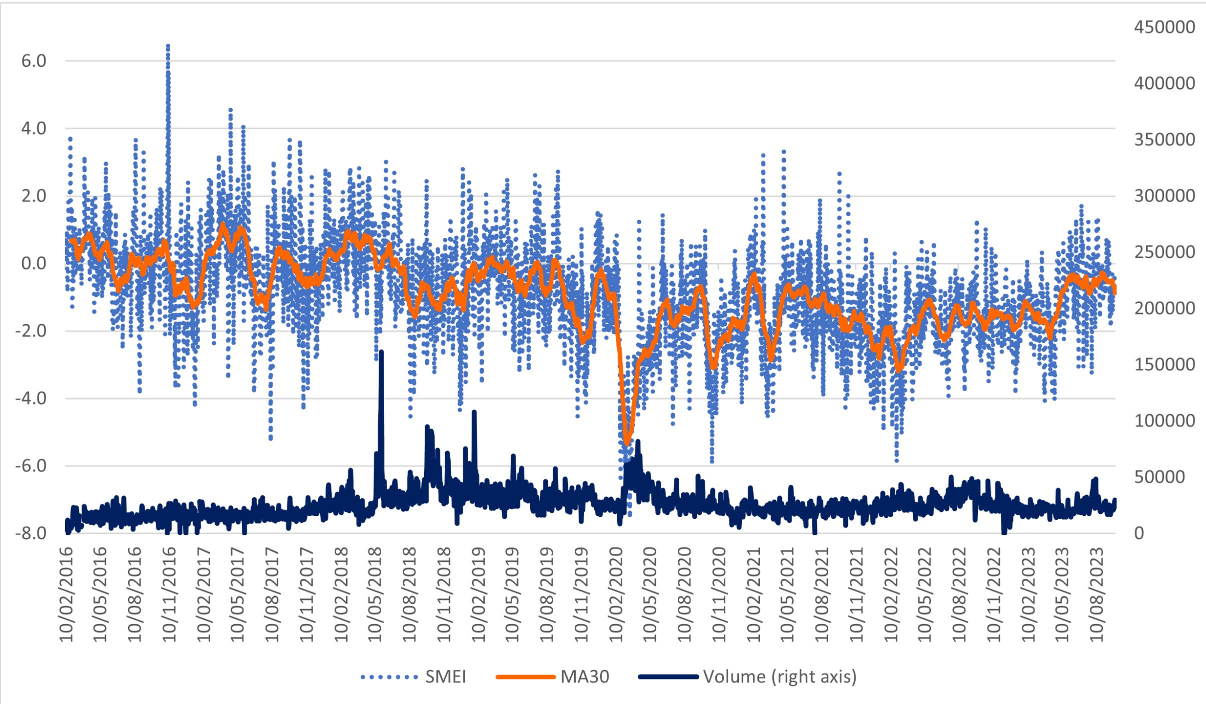

Figure 1: Social Mood on Economy Index: Daily values and 7-day and 30-day Moving Averages. Sample: February 10, 2016–September 30, 2023.

Figure 1 displays the values of the SMEI index for the whole time span, together with the corresponding 30-days (orange line) moving average, while the time evolution of the number of tweets that have been collected and analysed to compute the daily index, i.e. the "volume" of tweets underlying the daily values of the index can be found in dark blue. The higher the value of the index, the "better" and positive is the sentiment of the day.

It is clear from Figure 1 that the daily sentiment swings rather wildly, while themonthly moving average has a more stable pattern. A substantial slump in sentiment occurs between November 2016 and January 2017, followed by a sudden rise in the index. Oscillations in both directions followed until the absolute minimum of the indicator coinciding with the first lockdown occurred in March 2020. Other minimums are then found in October 2020, when the second lockdown was announced, and in February 2022, following the Russian aggression to Ukraine and the ensuing banking and economic sanctions. Starting from the raw daily data of the SMEI a trend is extracted, once two seasonal components are removed.[1]

Looking at the time volume of Tweets we observe clear peaks. The absolute maximum occurred on 29th May 2018, when the BTP-BUND spread exceeded 300 bps. Other peaks happened on 31 January 2019 (Italy fell into technical recession in 18Q4) and 10th April 2020 (when the Eurogroup decided on the economic policy response to the COVID-19 crisis). Watching more closely at the COVID-19 period, the volume of tweets showed a marked increase at the start of March 2020 and the volume doubled through April 2020. After the end of Spring 2020 volumes fell back to a level similar to the pre-pandemic ones. We are well aware that measuring sentiment swings with a Twitter-based indicator like SMEI has some limitations for the ensuing analysis. For instance, such an indicator does not produce a representative sample, neither of the whole Italian population nor of the FTSE MIB investors. We can anyway proxy also in the light of similar studies listed in the above section.

Figure 2 displays then the SMEI trend from the outbreak of COVID-19 to end-March 2022. The impact of the various phase of the pandemic and of the stringency of the containment measures is clearly visible in the chart.

The second set of data we used are the daily closing prices of the FTSE MIB. The FTSE MIB is the primary benchmark Index for the Italian equity markets. The FTSE MIB Index measures the performance of 40 Italian equities that captures approximately 80% of the domestic market capitalization .

Beside the closing prices this paper also uses volatility indices for the FTSE MIB index. The *FTSE Implied Volatility Index* (FTSE IVI) is a series of end-of-day mean volatility, derived from the at-the-money put and call implied volatilities on the FTSE MIB index options [for details on the calculation see `https://www.lseg.com/content/dam/ftse-russell/en_us/documents/ground-rules/ftse-implied-volatility-index-series-ground-rules.pdf`]. Indices for 30,

[1] The seasonal adjustment of the daily time series of the SMEI is described at `www.istat.it/it/files//2018/07/methodological-note-social-mood.pdf`

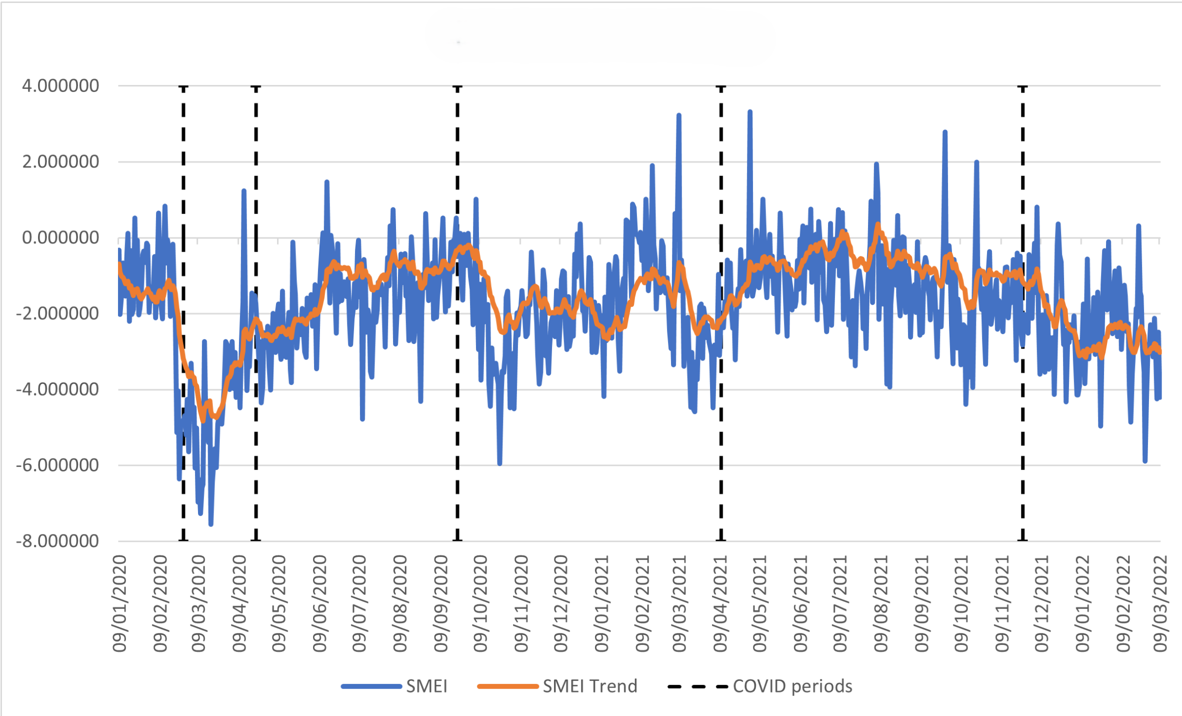

Figure 2: Evolution of the SMEI during the COVID-19 period.

60, 90 and 180 day implied volatility estimates are available. The FTSE IVI is a forward-looking indicator that provides market participants with information and risk management tools and also acts as an indicator of market sentiment and volatility. The FTSE (30-day) IVI is displayed in Figure 3 for the whole sample period of our study, together with the Garman-Klass (GK) volatility estimator defined as in equation (2).

Two remarkable volatility peaks happen on 18 March 2020 and 07 March 2022, in conjunction with the outbreak and escalation of the COVID-19 pandemic and of the Russian aggression against Ukraine respectively.

Next to the FTSE (30 days) IVI, two alternative volatility measures can be calculated for the same Italian market index, using four commonly available intradaily prices (Open, High, Low, and Close). The first is the Parkinson volatility estimator (Parkinson, 1980), based on the highest price, H_t, recorded on day t, and the lowest L_t:

$$V_{P,t} = \frac{1}{4log(2)} \left(logH_t - logL_t\right), \tag{1}$$

The second is the Garman-Klass volatility estimator (Garman and Klass, 1980) that uses all four prices recorded during day t, incorporating the opening (O_t) and the closing price (C_t) in

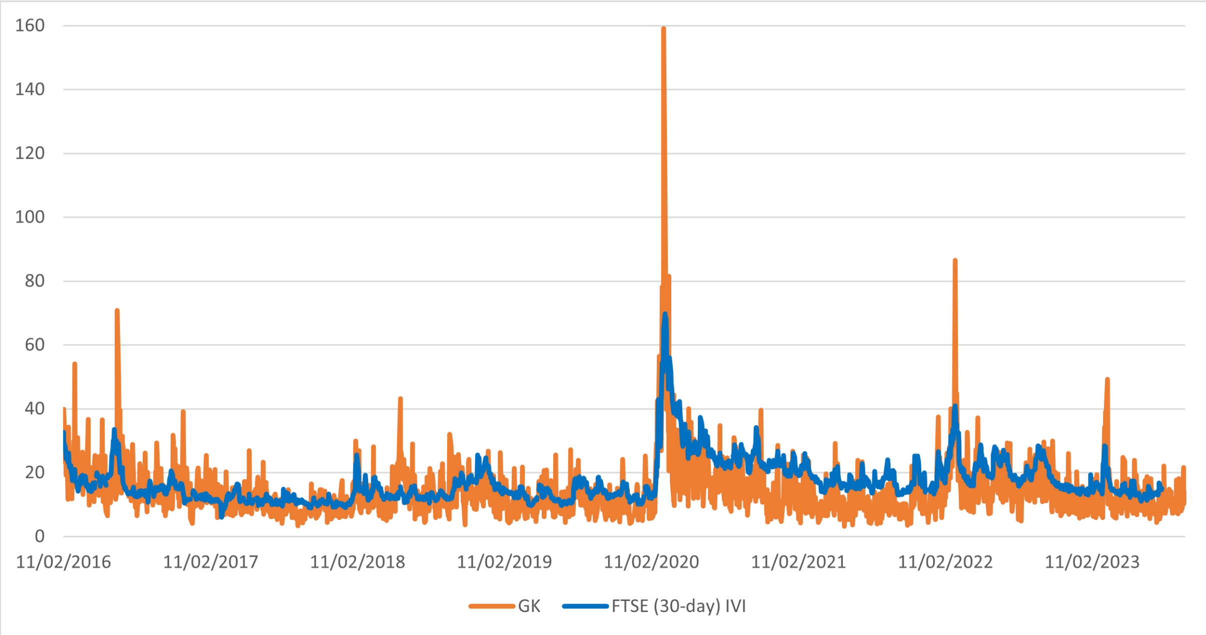

Figure 3: FTSE (30-day) IVI and GK estimator.

the estimation:

$$V_{GK,t} = 0.5\left(ln\frac{H_t}{L_t}\right)^2 - (2ln2-1)\left(ln\frac{C_t}{O_t}\right) \tag{2}$$

Both are end-of-day measures of the volatility at day t and they share much information with the FTSE (30-day) IVI: we summarize their features graphically in Figure 4. $V_{P,t}$ is systematically never higher than $V_{GK,t}$ with a sort of lower bound factor of approximately 1/2.

Two features must be noticed. First, the pattern of both indicators looks very similar to the one of the FTSE (30-day) IVI. Especially the Garman-Klass has a high positive linear correlation with the FTSE (30-day) IVI and the 45-degree line points to a strongest linear relationship. For the two indicators, Garman-Klass and Parkison, the correlation coefficient with the FTSE (30-day) IVI is 0.65 and 0.62 respectively. This will allow us to use these indicators as proxies of the FTSE (30-day) IVI since the data are not easily accessible. Second, the two volatility indices are highly correlated: as a matter of fact, the correlation coefficient between the Parkinson and the Garman-Klass volatility estimators is **.93** over our whole sample. Since they provide very similar results, the models we present in the next chapter are based on the Garman-Klass volatility in annualized percentage terms.

Table 1 displayes the descriptive statistics of the main variables used in this paper. First to observe that the number of observation of the FTSE (30-days) IVI is lower with respect to

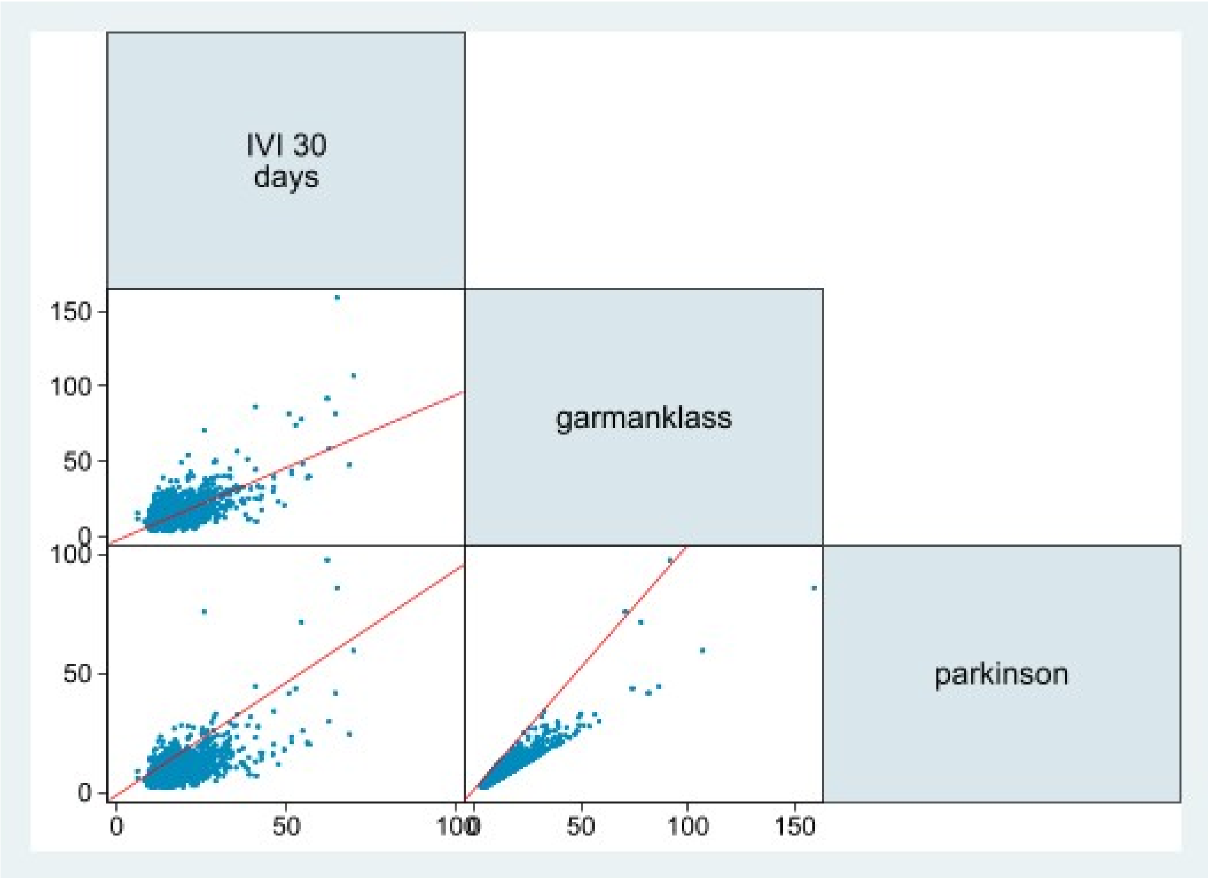

Figure 4: Bivariate scatter plots of the three volatility measures of the FTSE MIB (45-degree line in red).

Table 1: Descriptive statistics of the variables involved

	Obs	Mean	sd	Skewness	min	max
SMEI	1945	-0.876	2.525	-0.174	-7.550	6.532
Trend SMEI	1945	-0.563	1.149	-0.617	-4.827	1.974
MIB (thousand)	1945	21.455	2.819	0.042	14.894	28.162
Volume MIB (mill.)	1945	453.893	187.828	0.355	1.01	999.88
FTSE IVI	1872	17.24	6.92	2.573	6.02	69.73
V_P MIB	1945	8.677	6.029	5.381	2.096	97.51
V_{GK} MIB	1945	14.265	9.111	4.822	3.14	159.16
MIB return	1944	0.428	22.471	-1.949	-18.54	8.55

the other variables due to availability of the data. In the observed period between 2016 and 2023 the average sentiment index is negative (-0.876) with a relative high standard deviation, meaning that the sentiment widely varies during the observed period. Similar observations hold for the trend series of the SMEI, but reducing the standard deviation with respect of the original series and smoothing the picks. As regards the MIB, it is a positive series by definition. In the period relevant to this paper its minimum is 14.894 thousand reached in March 12, 2020 and the maximum 28.162 thousand reached on January 5, 2022. The volume, which represents the daily exchange of buying and selling operations, it is also a positive series and with a very high standard deviation. The volatility, as in its IVI form, of the Italian stock Index has average value 17.24, with a maximum value of 69.73 reached on March 16, 2020. The volatility of the MIB computed using the Garman-Klass formula or the Parkinson formula are close to each other for the mean and values and close enough to the FTSE IVI. Of course, the Garman-Klass version provides us with more information taking into consideration not only high and low value of the MIB index in the day but also its value at the market opening and closing. This is reflected in the higher standard deviation (9.111) of the Garman-Klass version concerning the simplified Parkinson one (6.029). As said, the two estimators have a very high correlation coefficient of 0.93. From the Skewness we see that none of the series is symmetric, only the MIB has very close to 0 Skewness, but not being a unimodal series this does not give us a lot of information. To conclude the table, we have included the return of the MIB as:

$$MIB\_return_t = ln(MIB)_t - ln(MIB)_{t-1} \tag{3}$$

The return of the FTSE-MIB is on average on the whole period 0.42% , varying from a minimum negative return of -18% to a maximum of 8.5% reached on March 24, 2020. We used the return of the MIB as one of the main variables to gauge the impact of the SMEI.

5 Some empirical evidence on the relationship between the SMEI and the FTSE-MIB

Figure 5 brings together the daily observations of the GK volatility and the correspondent values of the SMEI and its trend. We see very clear how negative peaks in the SMEI correspond almost always to positive peaks of the GK volatility, those highlighting a possible relation between the two components that may possibly affect each other. At first glance we could interpret this as a lower sentiment in the Italian population when there is more uncertainty of the Italian equity markets, mainly represented by the MIB.

Table 2 below and Figure 5 show that for the whole period (February 2016 to September 2023) a negative correlation between the FTSE-MIB closing prices and both the daily values

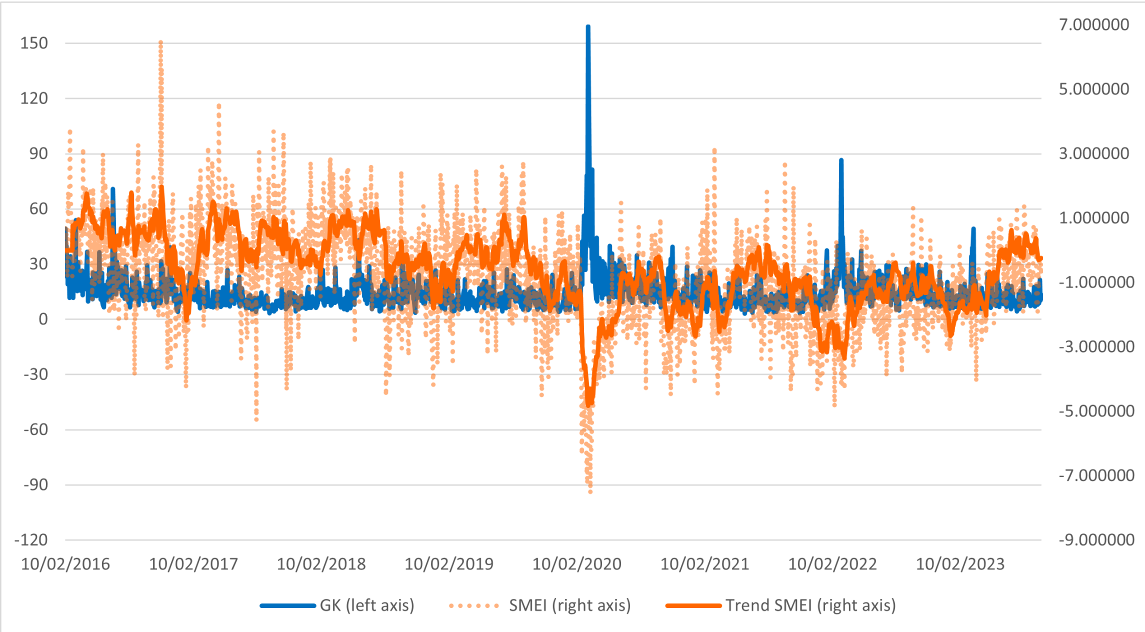

Figure 5: GK and SMEI: sample from 10/02/2016 to 30/09/2023.

Table 2: Correlation between the SMEI and the FTSE-MIB.

	Whole period	Pre-COVID-19	COVID-19	Post-COVID-19
SMEI	-0.1098	0.0155	0.3139	-0.1660
MA-30 SMEI	-0.2531	-0.1696	0.4900	-0.6779

and the trend of the SMEI. If we breakdown the timeline in pre-, during- and post- COVID-19 we see the correlations change. For the pre-COVID-19 there is a slightly positive correlation between the SMEI Index and the FTSE-MIB values, while using the monthly MA of the SMEI the correlation is again negative with the FTSE-MIB values. During COVID-19 a large positive correlation between the SMEI and the FTSE-MIB values appears, both using the SMEI series and the monthly MA. Post-COVID-19 the correlation turned negative again and with a stronger magnitude, especially when using the MA-30 days of the SMEI Index, signalling a clear divergence between economic sentiment and stock valuations.

However, the fact that two series are correlated does not necessarily implies that changes in one series “cause” changes in the other series. For this reason, we then perform a Granger causality test to understand possible predictive relationships between the SMEI series and the return of the MIB. In the next chapter we deepen the models and the results obtained.

6 The impact of mood on volatility

6.1 Granger causality in-sample

The basis of the analysis are Vector Auto-regressive Models (VAR) with the SMEI, the MIB return and the Garman-Klass estimator for the MIB volatility, using lag L=1,...,5 lags and restricted to the period from 10 February 2016 (start date of our dataset) to 8 March 2020. Then, we run the Granger causality test with null hypothesis:

H0: Lagged values of X do not cause Y

Granger causality in a VAR model implies a correlation between the past values of one variable and the current values of other variables. In some cases, both variables X and Y are found to be influenced by the other's lagged values leading to a bidirectional Granger causality.

Table 3: Granger causality restricted to the pre COVID-19. Sample: 10/02/2016 - 08/03/2020.

Dep.Variable	**Excluded**	**Chi-sq**	**df**	**Prob > Chi-sq**
SMEI	MIB return	5.25	5	0.386
SMEI	V_{GK}	7.14	5	0.211
SMEI	ALL	12.71	10	0.240
MIB return	SMEI	2.09	5	0.835
MIB return	V_{GK}	3.64	5	0.601
MIB return	ALL	5.70	10	0.840
V_{GK}	SMEI	1.75	5	0.882
V_{GK}	MIB return	24.62	5	0.000
V_{GK}	ALL	26.76	10	0.003

From Table 3, we clearly see how the null hypothesis is rejected with a high level of confidence when the dependent variable used is the Garman-Klass volatility of the MIB. In particular, the lagged return of the MIB (as expected) and both the lagged return of the MIB and SMEI when put together have a strong influence on the current Garman-Klass volatility. Those, clearly showing a strong correlation of both SMEI and the the return of the MIB with the volatility. Looking at the other two dependent variables, we cannot reject the null hypothesis of no Granger causality in the remaining displayed cases. We could probably say that the SMEI receives a certain influence by factors as the return of the MIB and its volatility but there are probably also other major external factors contributing to its behaviour. As for the return of the MIB this is not Granger caused by any of the two lagged variables, as the null hypothesis cannot be rejected.

Table 4: Granger causality restricted to the COVID-19 and post COVID-19 period. Sample: 09/03/2020 - 30/09/2023.

Dep.Variable	**Excluded**	**Chi-sq**	**df**	**Prob > Chi-sq**
SMEI	MIB return	7.08	5	0.214
SMEI	V_{GK}	24.54	5	0
SMEI	ALL	29.39	10	0.001
MIB return	SMEI	9.19	5	0.102
MIB return	V_{GK}	20.87	5	0.001
MIB return	ALL	24.74	10	0.006
V_{GK}	SMEI	21.85	5	0.001
V_{GK}	MIB return	126.97	5	0
V_{GK}	ALL	149.98	10	0

In Table 4, the results of the Granger causality test on the same variables and models are reported, this time related to the COVID-19 and post-COVID-19 periods, namely, from 9 March, 2020 to 30 September, 2023. In this shorter period of time characterised by a unique emergency period with major impact also on the economic landscape and on the life of the Italian population, we see results that differ from the table analysed before. The null hypothesis of no Granger causality can be rejected in most of the cases, interesting to observe how the joint impact of the lagged SMEI and GK volatility Granger causes the return of the MIB. Also for the SMEI the null hypothesis has to be rejected at 1% confidence level, leading to the fact that during this period the joint effect of the lagged volatility and return of the MIB Granger cause the social and economic mood of the Italian population. Concerning the Garman-Klass volatility we observe consistent results with the pre-COVID-19 period with respect to the Granger causality joint effect of the other two lagged variables and from the return of the MIB itself. Moreover the lagged SMEI seems to have also an effect as the null hypothesis is rejected in this case.

The Granger causality tests displayed highlight the existing relation between the three variables: SMEI, return of the MIB and the GK volatility of the MIB. The main result we focus on is the explainability and predictability of the volatility of the MIB using the lagged SMEI and the return of the MIB. In the next section we explore the out of sample models.

6.2 Out of sample - Diebold-Mariano forecast comparison test

To complete our analysis, we performed an extension of the Granger causality test to ascertain whether extra information is valuable in an out-of-sample framework. The question is then whether the forecasts for a variable produced by a simple AR with five lags, that is, using its past can be significantly outperformed by the forecasts obtained using the corresponding equation of a VAR model that includes additional variables (with the same number of lags). The comparison is done employing the Diebold-Mariano (DM) test (Diebold and Mariano, 1995), where the null hypothesis is one of equal performance of the two sets of forecasts according to a simple loss function, say, the Mean Square Error (MSE) or the Mean Absolute Error (MAE). The difference between either loss for the two sets, suitably standardized, is the DM test statistic with a limit Gaussian distribution. In our case, we test the null hypothesis against a one-sided alternative where the VAR outperforms the AR.

Table 5: Diebold Mariano (DM) test statistics for the one-step ahead forecast comparison between the VAR and the AR for the SMEI Index, the Garman and Klass market volatility, and the returns of the FTSE-MIB.

Series	**MSE**	**MAE**
SMEI		
DM-stat	-2.252	-2.188
p-value	0.988	0.986
GK Volatility		
DM-stat	**1.944**	**1.779**
p-value	0.026	0.038
MIB Returns		
DM-stat	-1.657	-3.021
p-value	0.951	0.998

Note: A positive value indicates a better performance of the VAR model; the one-sided p-values are calculated accordingly. Sample: 09/03/2020 - 30/09/2023.

The forecasts are generated recursively, by fixing the initial parameter estimation period between Feb. 10, 2016, and Mar. 8, 2020 (corresponding to the same 1087 observations used in-sample before), and producing one-step ahead results for 66 periods (approximately, three months) using the historical values for the lagged variables in the models. The fixed window is then moved forward by 66 periods, keeping 1087 observations for estimation (until Jun. 27, 2023), and 66 for projection (until Sep. 29, 2023). The number of overall forecast values is thus 792 for each set of models, from which the forecast errors are computed.

The results are presented in Table 5 by variable, with the DM test statistic value by MSE and MAE loss functions, accompanied by the one-sided p-values calculated as to detect a better performance of the extended information VAR model. The evidence somewhat complements the outcome of the in-sample analysis: out of the three variables, only GK volatility benefits from the extended information set; in our setup, the two models can be considered equivalent for the other two variables, with the interpretation that only the own past should be considered as relevant.

Projecting the behavior of volatility from the VAR can be appreciated graphically between the end of May 2020 and the end of September 2023, as in Figure 6. Except for the burst of volatility on the occasion of the Russian aggression in Ukraine in February 2022, the profile of the one-step ahead forecast follows the actual values rather closely.

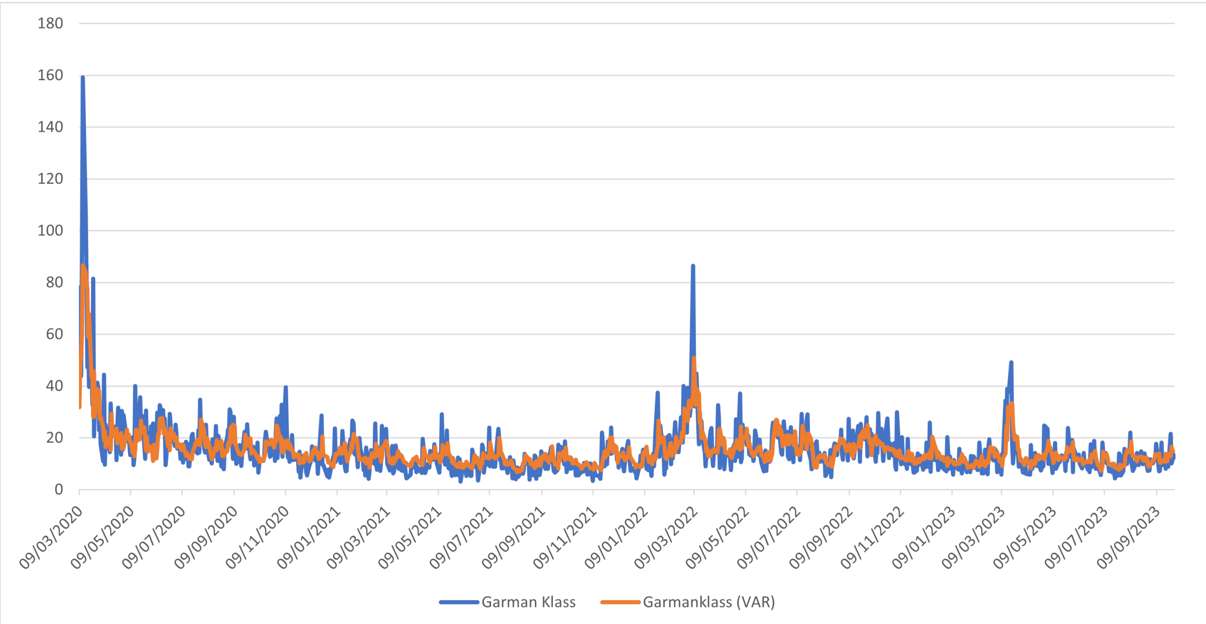

Figure 6: GK volatility versus its one-step ahead VAR one-step ahead forecasts generated in chunks of 66 days before re-estimating.

7 Concluding remarks

The availability of a plethora of users' participation in social forums, such as X (formerly known as Twitter) poses the serious challenge of validating the informational content of what is being expressed in each message. Several attempts are present in the literature, aimed at intercepting relevant words and synthesizing them into indices that can be monitored through time to follow what sentiment is prevailing in one economic environment.

An example of such exercises is given by the one performed by the Italian National Institute of Statistics (ISTAT; cf. Righi et al., 2020), with the Social Mood on Economy Index (SMEI) produced as a daily (weekends included) indicator of sentiment in the Italian context.

In this paper, we investigated some properties of the SMEI in its relationship with market performance of the Milan Stock Exchange, as represented by the FTSE-MIB index, here considered as the time series of both daily returns (first differences of log-prices at close) and daily volatility (expressed as an easily calculable range based on the open, high, low and close prices within the day Garman and Klass, 1980). The research question is one where we consider these two time series together with the SMEI within a stationary VAR model, exploring what relationship, if any can be established, employing an in-sample Granger causality test. We deemed it necessary to break the overall sample (spanning between Feb. 10, 2016, and Sep. 30, 2023) into two sub-samples, given the insurgence of the COVID-19 pandemic in March 2020. Such an epochal event has spurred a series of emergency measures that have disrupted for a long time regular economic and social activities, contributing to a different perception of uncertainty and risk.

The results show that for the pre COVID-19 period, the only variable significantly being affected by lagged values of other variables is the volatility singularly for the returns (presumably due to the so-called leverage effect by which negative returns increase market volatility), so strongly so, that the joint test for both variables (i.e. considering SMEI as well) turns out to be significant. By contrast, when the second sub-sample is considered, we notice that, for single variable tests, lagged volatility Granger-causes both SMEI and returns, lagged returns affect volatility, and lagged SMEI this time affects volatility (only marginally significant – p-value of around 10 % – for returns). The picture given, therefore, is one in which the pandemic turns out to significantly change the dynamic relationships among the variables considered in-sample.

The question can also be addressed dynamically in an out-of-sample context, whereby we resort to a different test, the well-known Diebold-Mariano test of superior predictive ability, holding a univariate autoregression as the benchmark. The framework we built is one in which we resort to rolling regressions, holding an estimation sample to a window of 1087 observations, producing 66 one-step ahead forecasts with both the univariate and the VAR models. In this case, the output shows that the only variable for which the VAR is predictively superior to the AR model is the range-based volatility, indicating that both lagged SMEI and returns are valuable information for forecasting market activity turbulence.

Moving forward, we think that the index built by ISTAT is a welcome addition to the panorama of signal extraction procedures from massive amounts of short messages exchanged over an important social forum such as X. To the best of our knowledge, such an index is subject to revisions and improvements, but the bottom line is that measuring social mood gives an important contribution to explaining market dynamics, especially at times of unexpected and devastating events. As a further indication, it may be advisable to provide evidence of the dynamic relationship of any synthetic sentiment index with returns and volatility as a way to assess the leading properties of sentiment onto market activity.

Other measures of market activity could be used, such as a VIX-type volatility index built from the implied volatilities of near-to-expiration put and call options written on the index. Such information is not freely available but could complement the analysis of the interaction in the market dynamics.

As mentioned, an interesting feature of the SMEI is its availability during the weekend and holidays, prompting the curiosity of whether the ”social mood” accumulated during market closures could generate a different impact on the outcomes of the first day of the trading week. This would require the generation of a pseudo-time series for the returns and the volatility over a seven-day week and proper care be exerted in detecting the direction of the impact.

References

Aggarwal, D. and Mohanty, P. (2018). Do Indian stock market sentiments impact contemporaneous returns? *South Asian Journal of Business Studies*, 7:332–346.

Angelico, C., Marcucci, J., Miccoli, M., and Quarta, F. (2022). Can we measure inflation expectations using twitter? *Journal of Econometrics*, 228(2):259–277.

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

Brown, G. W. and Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, 11(1):1–27.

Chen, S., Gao, T., He, Y., and Jin, Y. (2019). Predicting the stock price movement by social media analysis. *Journal of Data Analysis and Information Processing.*

Da, Z., Engelberg, J., and Gao, P. (2015). The sum of all fears investor sentiment and asset prices. *Review of Financial Studies*, 28:1–32.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

D'Amuri, F. and Marcucci, J. (2017). The predictive power of google searches in forecasting us unemployment. *International Journal of Forecasting*, 33(4):801–816.

Gallo, G. M., Granger, C. W., and Jeon, Y. (2002). Copycats and common swings: The impact of the use of forecasts in information sets. *IMF Staff Papers*, pages 4–21.

Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *The Journal of Business*, 53:67–78.

Gilbert, E. and Karahalios, K. (2010). Widespread worry and the stock market. *Proceedings of the International AAAI Conference on Web and Social Media*, 4(1):58–65.

Huang, C., Simpson, S., Ulybina, D., , and Roitman, A. (2019). News-based sentiment indicators. Technical report.

Kahneman, D. (2011). *Thinking, Fast and Slow.* Farrar, Straus and Giroux, New York.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65.

Nyman, R., Kapadia, S., and Tuckett, D. (2021). News and narratives in financial systems: Exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, 127:104119.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *The Journal of Business*, 53:61–65.

Rao, T. and Srivastava, S. (2012). Analyzing stock market movements using Twitter sentiment analysis. pages 119–123.

Righi, A., Iannaccone, R., Zardetto, D., and Zurlo, D. (2020). The Italian Social Mood on Economy Index and some Covid-19 crisis implications. *Banca d'Italia and Federal Reserve Board Join Conference*, pages 1481–1512.

Tetlock, P. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62:1139–1168.

Wang, Y.-H., Keswani, A., and Taylor, S. J. (2006). The relationships between sentiment, returns and volatility. *International Journal of Forecasting*, 22(1):109–123.

Whaley, R. E. (1993). Derivatives on market volatility: Hedging tools long overdue. *Journal of Derivatives*, 1:71–84.

Zhang, X., Fuehres, H., and Gloor, P. (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.

Annex

VAR trivariate models.

	Coefficient	Std. err.	z	P > $\|z\|$	[95% conf.	interval]
Index						
Index						
L1.	0.499	0.023	21.97	0	0.455	0.54
L2.	0.072	0.025	2.85	0.004	0.022	0.122
L3.	0.097	0.025	3.84	0	0.048	0.147
L4.	0.011	0.025	0.42	0.672	-0.039	0.061
L5.	0.097	0.023	4.28	0	0.053	0.142
MIB return						
L1.	4.264	1.891	2.26	0.024	0.558	7.970
L2.	-0.346	1.943	-0.18	0.859	-4.155	3.462
L3.	2.484	1.944	1.28	0.201	-1.326	6.294
L4.	-1.274	1.920	-0.66	0.507	-5.037	2.489
L5.	-1.681	1.833	-0.92	0.359	-5.274	1.912
V_{GK}						
L1.	6.449	6.552	0.98	0.325	-6.392	19.291
L2.	-5.598	6.723	-0.83	0.405	-18.776	7.579
L3.	-3.312	6.675	-0.5	0.62	-16.396	9.771
L4.	-10.251	6.657	-1.54	0.124	-23.297	2.796
L5.	-3.620	6.317	-0.57	0.567	-16.001	8.760
cons	-0.052	0.058	-0.89	0.374	-0.166	0.063
MIB return						
Index						
L1.	0.0003	0.0003	1	0.32	-0.0003	0.0009
L2.	-0.0002	0.0003	-0.55	0.58	-0.0008	0.0004
L3.	0.0001	0.0003	0.35	0.725	-0.0005	0.0007
L4.	0.00003	0.0003	0.1	0.919	-0.0006	0.0007
L5.	0.0002	0.0003	0.82	0.412	-0.0003	0.0008
MIB return						
L1.	-0.035	0.024	-1.47	0.143	-0.082	0.012
L2.	0.068	0.025	2.74	0.006	0.019	0.116
L3.	0.014	0.025	0.58	0.559	-0.034	0.063
L4.	-0.023	0.024	-0.93	0.351	-0.071	0.025
L5.	0.042	0.023	1.8	0.072	-0.004	0.088
V_{GK}						
L1.	0.082	0.083	0.98	0.327	-0.082	0.245
L2.	0.014	0.085	0.16	0.871	-0.153	0.182
L3.	-0.145	0.085	-1.71	0.087	-0.312	0.021
L4.	0.184	0.085	2.17	0.03	0.018	0.349
L5.	-0.014	0.080	-0.17	0.865	-0.17	0.144
cons	-0.0004	0.0007	-0.55	0.582	-0.0019	0.001

	Coefficient	Std. err.	z	$P > \|z\|$	[95% conf.	interval]
$\mathbf{V}_{GK}$						
Index						
L1.	-5.3E-05	8.3E-05	-0.64	0.525	-0.0002	0.0001
L2.	-4.2E-05	9.3E-05	-0.45	0.649	-0.0002	0.0001
L3.	-7.3E-05	9.3E-05	-0.79	0.429	-0.0003	0.0001
L4.	-0.00002	9.3E-05	-0.22	0.83	-0.0002	0.0002
L5.	-8.2E-05	8.3E-05	-0.99	0.322	-0.0003	8.1E-05
MIB return						
L1.	-0.072	0.007	-10.42	0	-0.086	-0.059
L2.	-0.039	0.007	-5.46	0	-0.053	-0.025
L3.	-0.030	0.007	-4.17	0	-0.044	-0.016
L4.	-0.013	0.007	-1.84	0.065	-0.027	0.001
L5.	-0.009	0.007	-1.41	0.159	-0.023	0.004
V_{GK}						
L1.	0.266	0.024	11.11	0	0.219	0.313
L2.	0.119	0.024	4.84	0	0.071	0.167
L3.	0.158	0.024	6.45	0	0.109	0.205
L4.	0.119	0.024	4.88	0	0.071	0.167
L5.	0.064	0.023	2.76	0.006	0.018	0.109
cons	0.002	0.001	10.56	0	0.002	0.003

Acknowledgements

The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the European Central Bank nor of the Corte dei conti.

Stefano Borgioli
European Central Bank, Frankfurt am Main, Germany; email: stefano.borgioli@ecb.europa.eu

Giampiero M. Gallo
Corte dei Ccnti, Florence, Italy; CRENoS -, University of Cagliari, Cagliari, Italy; email: giampiero.gallo@nyu.edu

Chiara Ongari
European Central Bank, Frankfurt am Main, Germany; email: chiara.ongari@ecb.europa.eu

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-6897-3 ISSN 1725-2806 doi:10.2866/3237842 QB-01-24-024-EN-N