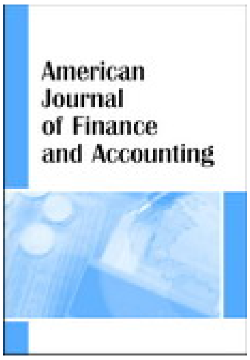
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



American J. of Finance and Accounting

ISSN online: 1752-7775 - ISSN print: 1752-7767
https://www.inderscience.com/ajfa

Investor sentiment metrics and stock market returns: a study of the causality relationship using VAR models

Dorsaf Ben Aissia, Nizar Neffati

DOI: 10.1504/AJFA.2023.10057975

Article History:

Received:	28 July 2022
Last revised:	29 July 2022
Accepted:	25 June 2023
Published online:	07 November 2023

Investor sentiment metrics and stock market returns: a study of the causality relationship using VAR models

Dorsaf Ben Aissia*

ISCAE,
University of Manouba, Tunisia
Email: dorsaf_bi@yahoo.fr
*Corresponding author

Nizar Neffati

Higher Institute of Computer Science and Management,
University of Kairouan, Tunisia
Email: neffati.nizar2016@gmail.com

Abstract: This paper examines the causality relationship between investor sentiment metrics and stock market returns. It considers survey, market, and composite sentiment indexes. It also introduces a dummy variable detecting the effect of economic crisis and decomposes sentiment into rational and irrational components. It uses VAR models and Granger tests, estimates Impulse Reaction Functions (IRFs) of the non-expected movement in investor sentiment, and proposes a forecast error variance decomposition (FEVD) approach to emphasise the importance of these movements on variables of the VAR models. Based on US data (S&P 500, Dow Jones, and NASDAQ indexes) from July 1965 to December 2021, we find a negative and significant relationship between investor sentiment and stock returns. This relationship is primarily explained by the irrational component of sentiment. In addition, we find a bi-directional Granger causality between stock returns and investor sentiment. Still, the IRFs and the FEVD study confirm the superiority of the survey indexes over the market indexes.

Keywords: investor sentiment metrics; stock market returns; causality relationship; VAR models; Granger tests; impulse response function; forecast error variance decomposition; FEVD.

JEL codes: G40; G41.

Reference to this paper should be made as follows: Ben Aissia, D. and Neffati, N. (2023) 'Investor sentiment metrics and stock market returns: a study of the causality relationship using VAR models', *American J. Finance and Accounting*, Vol. 7, No. 2, pp.90–124.

Biographical notes: Dorsaf Ben Aissia is a Professor of Finance. She joined the Higher Institute of Accounting and Business Administration in 2011. She holds a PhD from the Université of Tunis, awarded in 2008 and an Accreditation to Supervise Research (HDR) from the same university, awarded in 2016. His research interests are the stock market return and investor behaviour. He has published many academic papers (+15 papers) in international journals.

Nizar Neffati is an Assistant Professor in Finance. He holds a PhD in Finance from the University of Mannouba. He joined the FSEG of Sousse in 2000 and the Higher Institute of Computer Science and Management, Department of Finance and Accounting, in 2004 (University of Kairouan). His research interests are in capital market and behavioural finance.

1 Introduction

The last two decades have witnessed the burgeoning of finance literature that confirms the relevance of investor sentiment in shaping stock market returns[1] (e.g., Brown and Cliff, 2005; Chung et al., 2012; Beer et al., 2013; Ding et al., 2019). However, the unobservable nature of investor sentiment is the source of one major empirical problem namely the causality linkage between investor sentiment and stock returns (e.g., Chu et al., 2016; Cagli et al., 2020).

A stream of relevant research examines the causality between stock returns and investor sentiment using linear and nonlinear causality tests. Specifically, Chu et al. (2016) test for nonlinear causal relationships based on the method of Péguin-Feissolle et al. (2013) and find a strong bi-directional nonlinear causality between stock returns and investor sentiment. Still, Li et al. (2017) conduct a quantile Granger non-causality test and find that the causal relationship between investor sentiment and stock returns strengthens when a tail quantile interval is considered. Also, Cagli et al. (2020) employ a novel Granger causality test developed by Shi et al. (2018) to detect and date the changes in causal relationships. Their findings indicate that considering nonlinearities for the sample period could change the causal relationship between investor sentiment and the market return.

Another major problem in investor sentiment literature is to identify the best measure of investor sentiment. Relevant papers review investor sentiment metrics and focus on the data used to build these measures (market, survey, text, and media sentiment measures). For instance, DeVault et al. (2019) test the hypothesis that different variables capture investor sentiment-induced mispricing even if they are unrelated. They document a puzzling correlation between sentiment metrics. Zhou (2018) reviews different measures of investor sentiment and finds evidence that they explain returns on stocks that are difficult to value and costly to arbitrage. Moreover, he discusses the thorny issue of aggregating investor sentiment measures over various sources and time horizons to model the evolution of investor sentiment. In this paper, we examine the causality linkage between investor sentiment and stock market returns on the US market, using different indexes based on market and survey data.

The contribution of the paper to the existing literature is fourfold. First, while the related literature employs few indexes, we use different investor sentiment measures (individual and institutional) and propose new composite indexes that aggregate market and survey data using a principal component analysis (PCA). Second, we test the effect of the economic crisis on the sentiment-return relationship. Indeed, Canbaş and Kandır (2009) propose a model considering dummy variables controlling for a financial crisis. One of the explanations for considering an economic crisis is that investors and the market, in general, react differently depending on the economic condition (expansion/recession). Third, we study the decomposition of investor sentiment into rational and irrational components. The idea behind this decomposition is that we hypothesise that it is the irrational component of investor sentiment that affects the stock market returns (e.g., Qiu and Welch, 2006; Verma and Soydemir, 2006; Schmeling, 2009). Fourth, we complement our estimation of the VAR models and the Granger non-causality test by using the impulse response functions (IRFs) and the forecast error variance decomposition (FEVD). Indeed, the IRFs approach allows us to detect the effect of unexpected movements in investor sentiment (shocks) on stock performance while the decomposition of the variance of the forecast error emphasises the importance of these shocks on the variation of the VAR models' variables.

We find a negative and significant relationship between investor sentiment and stock returns. This relationship is more significant for survey indexes than for market indexes. Also, we show that rational component variables are not significant while irrational component variables are negative and significant. In addition, we find a bi-directional Granger causality between stock returns and investor sentiment. Still, our study of the impulse response functions confirms the superiority of the survey indexes over the market indexes and that it is the irrational component of investor sentiment that affects stock market returns. Finally, the FEVD shows that, for big firms, returns are mainly due to the sentiment of institutional investors rather than that of individual investors. Especially, the irrational component of the institutional investors' index contributes largely to variations in returns over a 12-month horizon.

The remainder of our paper is organised as follows. Section 2 reviews the literature and develops the hypotheses. Section 3 describes the data. Section 4 details the methodology. Section 5 details and discusses the empirical results. Section 6 concludes.

2 Literature overview and hypotheses development

2.1 The sentiment-returns causality linkage

The pioneer theoretical model work of De Long et al. (1990) of the sentiment-return relationship has been tested extensively in the recent empirical literature (i.e., Baker and Wurgler, 2007; Verma and Verma, 2008; Schmeling, 2009; Bathia and Bredin, 2013). Researches highlight two essential divergences concerning the nature of the market return and the investor sentiment linkage:

1 the direction and the degree of significance of the relationship between sentiment and return

2 the expected sign of this negative (respectively positive) causal relationship in terms of market timing.

Studying the effect of investor sentiment on stocks returns, the first stream of works uses different measures of sentiment, periods, frequencies, and markets and reports a negative impact of the sentiment on subsequent returns: Fisher and Statman (2003) and Brown and Cliff (2004) for small stocks; Brown and Cliff (2005), Baker and Wurgler (2007), Verma and Verma (2008), Schmeling (2009) and Bathia and Bredin (2013) for the irrational sentiment; and Swaminathan (1996), Fisher and Statman (2000), Simon and Wiggins (2001) and Wang (2001) for the large hedgers.

A second line of the research reports the absence of a significant impact of sentiment on subsequent returns: Solt and Statman (1988), Clarke and Statman (1998) and Wang (2001) for sentiment individual investors; Kling and Gao (2008) for China market, Atukeren et al. (2013) for Spain market; and Spyrou (2012) for Turkey market. Finally, relevant research finds that shifts in sentiment have a positive impact on subsequent returns, i.e., Neal and Wheatley (1998) and Wang (2001) for the sentiment of large speculators; Brown and Cliff (2004) for the sentiment of institutional investors; Beaumont et al. (2008) and Verma and Verma (2008) for rational sentiment; and Lux (2011) for the German market.

Studying the causality linkage, related works analyse the effect of performance on sentiment. Wang et al. (2006) highlighted a cause-and-effect relationship between performance and sentiment using the put/call ratio and the advances-declines ratio as measures. They conclude that sentiment shifts are caused by Granger returns and not vice versa. This result is consistent with that advanced by Brown and Cliff (2004). More recently, Barber et al. (2009) show that in the long run, there was a negative relationship between stock returns and investor sentiment. Still, Lao et al. (2018) find a positive effect of yield shocks on sentiment and a negative impact of sentiment factor shocks on market returns.

To summarise, empirical findings on the existence, nature, and sign of sentiment-return causality linkage are inconclusive and conclusions on the issue remain mixed. In light of the above arguments, we propose the following hypothesis.

H_1 Asset returns predict investor sentiment and the relationship between them is negative.

2.2 The economic crisis effect

Related sentiment studies examine whether the effect of investor sentiment on stock returns differs depending on the economic state (i.e., Lutz, 2016; Bouteska, 2020). Lutz (2016) shows that the effect of economic condition on sentiment are asymmetric and concludes that during the contraction period (going from peak to trough), a high sentiment predicts low future returns, while the relationship between investor sentiment and future returns is positive but relatively weak during the period of expansion of sentiment (passage from the trough to the peak). Still, Canbaş and Kandır (2009) propose a model considering two dummy variables controlling for the 1999 Marmara earthquake in Turkey and the 2001 financial crisis. One of the explanations for considering an economic crisis is that investors and the market, in general, react differently depending on the economic condition (expansion/recession). Hence, we hypothesise that:

H_2 Economic crisis predicts the relationship between investor sentiment and market returns.

2.3 The rational and irrational components of sentiment

Consistent with previous studies (i.e., Brown and Cliff, 2005; Shleifer and Summers, 1990), some measures of investor sentiment may be likely to contain elements of rational and irrational investor sentiment. According to Verma and Soydemir (2006), investors' optimism (respectively pessimism) may be a rational reflection of expectations for the coming period, an irrational enthusiasm, or a combination of both. Therefore, it is important to control all the information that the sentiment can contain on rational factors. Furthermore, Qiu and Welch (2006) show that some important macroeconomic pieces of information are correlated with current economic conditions and capture the rational sentiment. Still, Qiu and Welch (2006), Baker and Wurgler (2006), Lemmon and Portniaguina (2006), Verma and Soydemir (2006) and Schmeling (2009) emphasise that it is an irrational component of investor sentiment that affects the stock market returns. Based on these arguments, we draw the following hypothesis.

H_3 Market returns predict the irrational component of sentiment.

3 Data and preliminary analysis

This study focuses on the US market. It analyses aggregate returns and different sentiment metrics. Table 1 summarises the different data used.

Table 1 Description of variables used in this study

Code	*Definition and source*	*Source*
	Stock markets returns	
S&P500	The S&P500 index for medium companies	Datastream
DJX	The Dow Jones Industrial Average (DJX) index for large companies	Above
NQX	The NASDAQ index for small companies.	Above
	Investor sentiment indexes	
AAII	The American Association of Individual Investors index (%)	Above
$AAII_{Bull}$	The ratio of the bullish percentage to the bearish percentage of AAII index	Above
$AAII_{Bear}$	The ratio of the bearish percentage to the bullish percentage	Above
II	The Investors Intelligence index (%)	Above
II_{Bull}	The ratio of the bullish percentage to the bearish percentage.	Above
II_{Bear}	The bearish percentage to the bullish percentage.	Above
VIX	The implied volatility index.	Above
PCR	The pull-call ratio.	Above
$SENT_1$	The author's index was calculated using PCA approach.	The author's calculation is based on Datastream data.
$SENT_2$	The author's index was calculated using PCA	The author's calculation is based on data

Table 1 Description of variables used in this study (continued)

Code	*Definition and source*	*Source*
	Fundamental variables (Fund)	
Growth	Economic growth is the monthly changes in the industrial production index.	Welch and Goyal (2008) data
*TB*30	Short-term interest rate is the yield on the one-month US treasury bill.	Above
PRE	The economic risk premium is the term structure of interest rates (difference in monthly yields on three-month and one-month treasury bills).	Above
TMS	Future economic expectations: is the term spread (yields spread on the 10-year US treasury bond and three-month treasury bill).	Above
BC	Business conditions: is the default spread (difference in yields on Baa and Aaa corporate bonds)	Above
Div	Is the dividend yield for the value weighted Center for Research in Security Prices (CRSP) index over the past 12 months.	Above
INFL	Inflation is the monthly changes in the consumer price index.	Above
RM-RF	Is the market return premium over the risk-free rate.	Above
SMB	Is the average return on the three small portfolios minus the average return on the three big portfolios.	Kenneth French's website
HML	Is the average return on the two value portfolios minus the average return on the two growth portfolios.	Above
UMD	Is the average return of a high prior return portfolio over a low prior return portfolio.	Above

The first group of data contains equally-weighted price series data on the S&P500, the Dow Jones Industrial Average index, and the NASDAQ to approximate the overall performance of the US stock market. Following Da et al. (2011), Vozlyublennaia (2014) and Mbanga et al. (2019), we use the Dow Jones Industrial Average (DJX) index to represent large companies, the S&P500 (S&P500) index for medium-sized firms, and the NASDAQ (NQX) index for small firms. We estimate the monthly compounded returns of these price series for the period July 1965 to December 2021. The returns are calculated as the first difference of the natural logarithm of the index.

$$R_{i,t} = \mathrm{Ln}\left(\frac{P_t}{P_{t-1}}\right) \tag{1}$$

where $R_{i,t}$ denotes the return to stock index *i* in month *t*.

The second group of data includes three kinds of metrics: surveys, market, and composite indexes. The study uses survey data based on the American Association of Individual Investors (AAII) and Investors' Intelligence (II), which capture the sentiment of individual and institutional investors; these data are obtained from the DataStream database and are monthly for AAII and II. It also uses the implied volatility index (VIX) and the put/call ratio as market data. These variables are daily and weekly from the Chicago Board Options Exchange (CBOE). In our paper, we use the principal component approach (PCA)[2] to extract two composite indexes corresponding to the survey composite index and market composite index, namely, SENT1 and SENT 2. The results are reported in panel B of Table 2 which details the composition of each index.

- $SENT_1 = AAII_{Bull}, AAII_{Bear}, II_{Bull}$ and $AAII_{Bear}$
- $SENT_2 = VIX$ and PCR.

The third group is relative to the fundamental variables. The rationale underlying the use of these data is obtaining irrational components of sentiments. We follow Verma and Verma (2008) and retain eight firm characteristics namely, economic growth, short-term interest rates, economic risk premia, future economic expectations, business conditions, dividend yield, inflation, excess returns on market portfolio (RMF), SMB and HML factors and momentum factor. These variables are obtained from Welch and Goyal's (2008) data[3]. We test the stationarity of our data to avoid potential spurious regressions. We use the augmented Dickey-Fuller (ADF), the Phillip-Perron (PP), and Dickey-Fuller-generalised least squares (DF-GLS) (1996) unit-root tests to determine the appropriate order of integration for each variable.

Table 2 Investor sentiment metrics and stocks returns data

	Panel A		
	Variable	*Start date*	*End date*
Stock market returns	S&P500 EW	1965:07	2021:12
	DJX	1965:07	2021:12
	NQX	1965:07	2021:12
Survey measures	AAII	1965:02	2021:12
	II	1965:07	2021:12
Market measures	VIX	1990:01	2021:12
	Put/Call total	1995:09	2021:12
	Panel B		
Index	*Description*	*Method*	*Period*
$SENT_1$	$AAII_{Bull}$	PCA	1965:07–2021:12
	$AAII_{Bear}$		1965:07–2021:12
	II_{Bull}		1965:06–2021:12
	II_{Bear}		1965:06–2021:12
$SENT_2$	VIX	PCA	1990:01–2021:12
	PCR		1995:09–2021:12

Note: This table reports the time-period of the monthly data of the different investor sentiment metrics and the stock market returns. The variable definition is listed in the Appendix. Panel A displays stock market returns, survey measures, and market measures. Panel B reports the composition and the period of the two composite indexes we construct through the aggregation approach (PCA) namely: $SENT_1$ and $SENT_2$.

4 Methodology

This section describes the methodology used to assess the causality relationship between sentiment and returns, on the US stock market. We first estimate this relationship using a VAR model. Second, we study the causality linkage between them based on the Granger non-causality tests. Finally, we analyse, using the impulse response function (IRF) generated from the VAR model, the impact of the non-expected movement in investor sentiment on stocks returns and analysis the decomposition of the variance of the forecast error of these movements.

4.1 VAR models and Granger causality tests VAR models

4.1.1 VAR models

Related relevant works (i.e., Verma and Verma, 2008; Brown and Cliff, 2004) underline that market returns and investor sentiment are most likely to interact instantly or with some lags. This approach is called the vector autoregressive (VAR) model and allows us a better understanding of the relationship between the two variables.

To analyse these interactions among variables, we adopt the VAR methodology of Brown and Cliff (2004) and Canbaş and Kandır (2009). According to these authors, the introduction of VAR models makes it possible to analyse the interdependencies between several variables, in particular between sentiment and market performance. They argue that VAR models are more suitable for describing the dynamic behaviour of economic and financial time series data and for forecasting (i.e., Sims, 1980). Following Brown and Cliff (2004), we use the general model of *VAR*(*p*) given by:

$$Y_t = \mu + \sum\nolimits_{i=1}^{p} \varphi_i Y_{t-i} + \varepsilon_t \tag{2}$$

where Y_t is a vector of all vectors y of every variable in the system, φ_i is $n \times n$ matrix of parameters of the lag period, and ε_t is an $n \times 1$ vector of identically and independently distributed errors. Using this model, we analyse to what extent the shift in one variable affects other variables and by how much the shock to one variable impacts other variables.

As in Canbaş and Kandır (2009), we introduce in the VAR model a binary variable (*DUMMY*) to control for the economic crisis effect. Model (2) can be further converted to the following:

$$RM_t = a_0 + \alpha_i RM_{t-i} + \sum\nolimits_{i=1}^{n} \beta_i Sentiment_{t-i} + \lambda DUMMY_1 + \varepsilon_{1t}$$

$$Sentiment_t = b_0 + \varphi_i RM_{t-i} + \sum\nolimits_{i=1}^{n} \delta_i Sentiment_{t-i} + \eta DUMMY_1 + \varepsilon_{2t} \tag{3}$$

Equations (3) can be rewritten in a VAR simple matrix form as:

$$\begin{pmatrix} R_M \\ SENT_t \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \end{pmatrix} + \sum\nolimits_{i=1}^{p} \begin{pmatrix} \alpha_{11}(i) & \alpha_{12}(i) \\ \alpha_{21}(i) & \alpha_{22}(i) \end{pmatrix} + \begin{pmatrix} \lambda \\ \eta \end{pmatrix} * DUMMY_1 + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{1t} \end{pmatrix} \tag{4}$$

where Y_t = [*SENT*, *RM*] is a vector of endogenous variables, *SENT* is the proxy of the investor sentiment (*AAII*, *II*, $SENT_1$, and $SENT_2$), *RM* is the market return calculated based on a composite index (S&P500, Dow Jones Industrial Average index and le NASDAQ), *DUMMY* is a control variable for the various financial crises affecting the US market, n is the order of the VAR process and ε_{1t}, ε_{2t} are two error terms.

In our paper, our sentiment indexes are analysed, first, according to their nature. Indeed, we examine individual sentiment measure (*AAII*), in opposition to institutional investors measure (*II*) and direct sentiment measurement ($SENT_1$) to indirect sentiment measure ($SENT_2$). The idea behind this is to detect if the fluctuations in returns are due to a shock of sentiment relating to the type of investor: individual or institutional, or to the nature of the index used: market measures versus survey measures. Our two vectors of endogenous variables are defined as follows:

$$Y_{1t} = \left[AAII, II, S\&P500_{ret}, DJX_{ret}, NQX_{ret}\right]'$$

$$Y_{2t} = \left[SENT_1, SENT_2, S\&P500_{ret}, DJX_{ret}, NQX_{ret}\right]'$$

Second, we consider in our model a decomposition of investor sentiment into rational and irrational components as in Verma and Soydemir (2006), Verma and Verma (2008) and Sayim and Rahman (2015). We then model the rational component of investor sentiment as a set of fundamentals representing investor rational expectations based on economic risk factors. We run the following regression:

$$Sent_{1t} = \gamma_0 + \sum_{i=1}^{n} \gamma_j FUND_{jt} + \xi_t \tag{5}$$

Where $Sent_{1t}$ is the investor sentiment at time t, γ_0 is a constant, γ_j is the parameter to be estimated, *FUND* are fundamentals and ξ_t is the error term. We consider as fundamentals the variables:

1 economic growth (i.e., Fama, 1970; Schwert, 1990)

2 short-term interest rates (i.e., Campbell, 1991)

3 economic risk premia (i.e., Campbell, 1987; Ferson and Harvey, 1991)

4 future economic expectations variables (i.e., Fama, 1990)

5 business conditions (i.e, Fama and French, 1989; Keim and Stambaugh, 1986)

6 dividend yield (i.e., Campbell and Shiller, 1988a, 1988b; Fama and French, 1988; Hodrick, 1992)

7 inflation (i.e, Fama and Schwert, 1977; Sharpe, 2002)

8 excess returns on market portfolio (i.e, Lintner, 1965; Sharpe, 1964)

9 SMB factor (small minus big) and HML factor (high minus low) (i.e., Fama and French, 1993)

10 momentum factor (UMD) (i.e., Jegadeesh and Titman, 1993).

ξ_t is the error term representing the irrational component of sentiment also called noise (i.e., Hirshleifer, 2001; Brown and Cliff, 2005; Verma and Verma, 2008).

Table 3 Effect of fundamentals on investor sentiment measures

Decomposition of sentiment measures into rational and irrational components: OLS

We estimate the following regression:

$$Sent_{1t} = \gamma_0 + \sum_{i=1}^{n} \gamma_j FUND_{jt} + \xi_t$$

Variables are US investor sentiments $Sent_{1t}$, $FUND_{jt}$ is the set of fundamental factors indicating rational investor expectations based on several risk variables which are commonly accepted and used to value asset prices in the literature: US economic growth (*Growth*), economic risk premium ($PRE = T90 - T30$), future economic expectations variables ($TMS = B10 - T30$), US business conditions ($BC = BAA - AAA$), dividend yield (*Div*), inflation (*INF*), short-term interest rates (*TB*30), excess return on the market portfolio (RM_{excess}), Fama and French (2015) and momentum effect Jegadeesh and Titman (1993) (*SMB*, *HML*, *RMW*, *CMA*, *UMD*). OLS regression models. *, **, and *** denote significance level at the 10%, 5%, and 1%.

	(1)	*(2)*	*(3)*	*(4)*
	AAII	*II*	$SENT_1$	$SENT_2$
Growth	0.013	0.040***	0.124	0.161*
	(1.072)	(3.971)	(1.547)	(1.930)
INF	–5.360***	–0.235	21.393	25.525*
	(–2.760)	(–0.156)	(1.634)	(1.867)
BC	–9.188***	–1.512	–90.986***	30.262**
	(–4.686)	(–1.099)	(–6.883)	(2.062)
RM_{excess}	0.658***	0.290**	–0.797	–5.615***
	(4.028)	(2.359)	(–0.724)	(–4.157)
TB30	1.048	3.460**	11.215	–35.437**
	(0.550)	(2.300)	(0.873)	(–2.420)
TMS	0.199	0.529	–22.867***	–42.505***
	(0.228)	(1.074)	(–3.891)	(–4.479)
PRE	–0.457	–1.515***	–26.092***	–40.572***
	(–0.732)	(–5.056)	(–6.197)	(–5.549)
Div	–0.002*	0.000	0.007	0.004
	(–1.664)	(0.582)	(0.921)	(0.430)
SMB	0.519**	0.421**	3.108**	–3.508**
	(2.271)	(2.382)	(2.017)	(–2.009)
HML	0.437	0.811***	3.962*	3.347
	(1.397)	(3.268)	(1.879)	(1.456)
RMW	–0.002	0.000	–0.015	–0.022
	(–0.818)	(0.004)	(–0.712)	(–0.916)

Notes: All reported absolute t-values in parentheses are based on robust standard errors adjusted for heteroskedasticity.
***, **, * indicate statistical significance at the 1%, 5% and 10% levels.

Table 3 Effect of fundamentals on investor sentiment measures (continued)

	(1) *AAII*	*(2)* *II*	*(3)* *SENT1*	*(4)* *SENT2*
CMA	0.003	–0.005	–0.009	–0.065**
	(0.607)	(–1.448)	(–0.309)	(–2.086)
UMD	–0.284*	0.178	0.016	1.274
	(–1.958)	(1.490)	(0.016)	(1.222)
cons	0.699***	0.647***	1.884***	1.422**
	(10.826)	(22.318)	(4.326)	(2.231)
R-squared	0.223	0.249	0.424	0.461
F-statistic	7.76	14.55	19.95	16.67
Prob (F-stat)	0.000	0.000	0.000	0.000
DW	1.342	0.682	0.777	1.374

Notes: All reported absolute t-values in parentheses are based on robust standard errors adjusted for heteroskedasticity.
***, **, * indicate statistical significance at the 1%, 5% and 10% levels.

Table 3 represents regression results for equation (6) and details used variables. Still, the irrational component affects returns as follows:

$$R_t = \alpha_0 + \alpha_1 \widehat{Sent}_{1t} + \alpha_2 \xi_t + \rho_t \tag{6}$$

where α_0 is a constant, α_1 and α_2 are parameters to be estimated and ρ_t is the random error term. In this model, the parameters α_1 capture the impact of rational investor sentiment, while parameters α_2 capture the impact of irrational investor sentiment. We define:

$$Y_{3t} = \left[AAII_{rat}, AAII_{irrat}, II_{rat}, II_{irrat}, S\&P500_{ret}, DJX_{ret}, NQX_{ret}\right]'$$

$$Y_{4t} = \left[SENT1_{rat}, SENT1_{irrat}, SENT2_{rat}, SENT2_{irrat}, S\&P500_{ret}, DJX_{ret}, NQX_{ret}\right]'$$

Besides, we use respectively the Akaike information criteria (AIC) of Akaike (1974) and the Schwarz Bayesian information criterion (SBIC) of Schwarz (1978) to determine the order of the VAR model. We retain a delay of two months for Y_{1t} and Y_{3t} and of three months for Y_{2t} and Y_{4t}.

In addition, several tests are used to assess the quality of the multivariate estimates [i.e., the test of the Lagrange multiplier of autocorrelation of the residuals (LM test), the normality test of the residuals (Jarque-Bera) and the stability test of the estimated VAR].

Table 4 Effect of fundamentals on investor sentiment measures

Decomposition of sentiment measures into rational and irrational components: Cochrane-Orcutt model

We estimate the following regression:

$$Sent_{1t} = \gamma_0 + \sum_{i=1}^{n} \gamma_j FUND_{jt} + \xi_t$$

Variables are US investor sentiments $Sent_{1t}$, $FUND_{jt}$ is the set of fundamental factors indicating rational investor expectations based on several risk variables which are commonly accepted and used to value asset prices in the literature: US economic growth (*Growth*), economic risk premium ($PRE = T90 - T30$), future economic expectations variables ($TMS = B10 - T30$), US business conditions ($BC = BAA - AAA$), dividend yield (*Div*), inflation (INF), short-term interest rates (*TB*30), excess return on the market portfolio (RM_{excess}), Fama and French (2015) and momentum effect Jegadeesh and Titman (1993) (*SMB*, *HML*, *RMW*, *CMA*, *UMD*). OLS regression models. *, **, and *** denote significance level at the 10%, 5%, and 1%.

	(1)	*(2)*	*(3)*	*(4)*
	AAII	*II*	$SENT_1$	$SENT_2$
Growth	0.003	–0.004	–0.062	0.154**
	(0.011)	(0.006)	(0.052)	(0.077)
INF	–4.417**	1.015	–0.838	24.025*
	(2.028)	(1.129)	(10.893)	(14.533)
BC	–10.128***	–2.943	–121.178***	32.776*
	(2.500)	(2.223)	(21.074)	(18.190)
RM_{excess}	0.628***	–0.114*	–2.793***	–5.482***
	(0.146)	(0.069)	(0.715)	(1.224)
TMS	0.196	–0.008	–18.473**	–27.608**
	(1.156)	(0.816)	(9.078)	(11.071)
PRE	–0.631	–2.049***	–24.450***	–28.137***
	(0.823)	(0.576)	(6.500)	(8.399)
Div	–0.002	–0.001	0.011	0.021*
	(0.002)	(0.002)	(0.013)	(0.012)
SMB	0.498**	0.415***	1.237	–4.079**
	(0.205)	(0.101)	(1.001)	(1.577)
HML	0.517*	0.170	1.379	3.114
	(0.282)	(0.144)	(1.376)	(2.118)
RMW	0.000	0.003*	–0.005	–0.029
	(0.003)	(0.002)	(0.014)	(0.023)

Notes: All reported absolute t-values in parentheses are based on robust standard errors adjusted for heteroskedasticity.
***, **, * indicate statistical significance at the 1%, 5%, and 10% levels.

Table 4 Effect of fundamentals on investor sentiment measures (continued)

	(1) AAII	(2) II	(3) $SENT_1$	(4) $SENT_2$
CMA	0.000	0.002	0.003	–0.048*
	(0.004)	(0.002)	(0.019)	(0.028)
UMD	–0.394***	–0.034	–0.596	1.859**
	(0.126)	(0.066)	(0.607)	(0.924)
C	0.719***	0.735***	2.044***	0.306
	(0.083)	(0.061)	(0.655)	(0.716)
R-squared	0.182	0.096	0.213	0.341
F-statistic	6.510	5.04	7.96	10.93
Prob (F-stat)	0.000	0.000	0.000	0.000
DW	2.048	2.131	2.028	2.097

Notes: All reported absolute t-values in parentheses are based on robust standard errors adjusted for heteroskedasticity.
***, **, * indicate statistical significance at the 1%, 5%, and 10% levels.

4.1.2 Granger non-causality tests

One of the methods proposed in the recent related literature (i.e., Shi et al., 2018; Cagli et al., 2020), to analyse the linear causality between two variables, is that of Granger (1969, 1988). The implementation of the Granger causality test requires the prior estimation of the VAR models in their reduced forms given by:

$$RM_t = a_0 + \alpha_i RM_{t-i} + \sum\nolimits_{i=1}^{m} \beta_i Sentiment_{t-i} + \lambda DUMMY_1 + \varepsilon_{1t} \tag{7}$$

$$Sentiment_t = b_0 + \varphi_i RM_{t-i} + \sum\nolimits_{i=1}^{m} \delta_i Sentiment_{t-i} + \eta DUMMY + \varepsilon_{2t} \tag{8}$$

Granger's causality test uses the Wald statistics following the χ^2 distribution under the joint hypothesis that a variable Y does not cause X in the sense of Granger. Since the null hypothesis of Granger non-causality and the alternative hypothesis of Granger causality are:

H_0 Absence of causality in the sense of Granger (*Sentiment ~ RM*), if:
$\beta_1 = \beta_2 = \ldots = \beta_m = 0$.

H_1 Causality in the sense of Granger (*Sentiment* $\rightarrow$ *RM*), if:
$\beta_1 \neq \beta_2 \neq \ldots \neq \beta 1_m \neq 0$.

We test the null hypothesis that market returns do not Granger cause investor sentiment in equation (7) and that investor sentiment does not Granger cause market returns in equation (8).

4.2 Impulse response functions (IRFs)

Our study uses different measures of investor sentiment and indicates the resulting relationship between the expected variations of this factor on market returns. Still, the VAR models ignore the impact of non-expected movements on investor sentiment and stock market returns (i.e., Statman et al., 2006; Verma and Soydemir, 2009) which could create a misspecification problem.

Therefore, this study uses impulse response functions (IRFs) generated from the VAR model to examine the effect of unexpected movements in investor sentiment on stock performance. The shock response functions analyse the impact of an exogenous shock on the contemporary and future values of the variables of the VAR model. It detects the impact of a punctual shock of unexpected changes in investor sentiment on the current and future values of market returns. Still, IRFs represent the behaviour of series in response to shocks while holding the effects of other variables constant.

Since IRFs functions are nonlinear in the estimated parameters, it is necessary to have an idea of its precision, via its variance, to construct confidence intervals or confidence bands. The latter is constructed around the mean response by applying Monte Carlo methods (e.g., Litterman, 1986). The confidence interval calculated from the bootstrapping procedure is 95%. Our papers use the generalised impulse response of Pesaran and Shin (1998), in which a set of orthogonalised shocks does not depend on the order of the variables retained in the VAR.

4.3 The forecast error variance decomposition

The study of the impulse response functions tells us how a shock caused to one variable prop- agates into the system while affecting other variables (the sign of the effect: positive or negative). However, one of its limitations is its inability to determine the magnitude of this shock. The decomposition of the forecast error variance (FEVD: *forecast error variance decomposition*) makes it possible to provide elements of solutions to these limits. Indeed, if the analysis of the IRFs measures the nature of the shock on variables (positive or negative effect), the decomposition of the variance of the forecast error emphasises the importance of innovations (shocks) on the variation of the variables.

Therefore, we complete our study with an analysis of the decomposition of the variance of the forecast error. The aim is to calculate the contribution of innovations (shocks) to the variance of the error. To do this, we calculate the contribution of innovation j to the variance of the forecast error of variable x at horizon h: (x_j, $T + h$). Given that, the variance of the forecast error at horizon h with orthogonal shocks is written:

$$\delta_h = x_{iT+h} - \frac{x_{iT+h}}{\Psi_T} = \sum_{s=0}^{h-1} \theta_{i1s} \varepsilon_{1,T+h-s} \tag{9}$$

For a given variable x_i, the forecast error is defined as follows:

$$\delta_h = x_{iT+h} - x_{iT+h} / \Psi_T = \sum\nolimits_{s=0}^{h-1} \theta_{i1s} \varepsilon_{1,T+h-s} + \ldots + \ldots \sum\nolimits_{s=0}^{h-1} \theta_{iks} \varepsilon_{1,T-h-s} \tag{10}$$

Thus, the variance of the forecast error at horizon h is written as follows:

$$V(\delta_{ih}) = \sigma^2_{\varepsilon 1t} \sum_{s=0}^{h-1} \theta_{i1s} + \ldots + \sigma^2_{\varepsilon 1t} \sum_{s=0}^{h-1} \theta_{iks} \tag{11}$$

The contribution of innovation j to the variance of the forecast error of variable i at the horizon h is:

$$VD_{ij}(h) = \frac{\sigma^2_{\varepsilon 1t} \sum_{s=0}^{h-1} \theta_{iks}}{V(\delta_{ih})} \tag{12}$$

Therefore, for a *VAR* model with k variables, there are k^2 values of $VD_{ij}(h)$.

In this paper, we calculate the contribution of the shocks in AAII, II, SENT1, and SENT2 to the variations in the returns of S&P500ret, DJXret, and NQXret.

5 Empirical results and discussion

5.1 VAR model results

Tables 5-8 report the regression results of market returns on sentiment indexes following equation (4). Table 5 presents the results for the variable (Y_{1t}). We find that coefficient estimates associated with institutional investors sentiment (II) are negative and significant for the returns of large, medium-sized, and small firms respectively S&P500ret, DJXret, and NQXret, with a one-month lag. Still, we find that the measure reflecting the sentiment of individual investors (AAII) is negatively correlated with stock returns for a lag of two months. It is not significant for a lag of one month. These results confirm that market investor sentiment is negatively affected by recent market performance. In light of these results, we accept Hypothesis H_1 which stipulates that asset returns predict investor sentiment and that the relationship between them is negative.

However, Table 6 displays opposite results for the composite indicators used in our analysis (i.e., $SENT_1$ and $SENT_2$). Indeed, Table 6 reports that associated coefficients are negative and significant for the market index $SENT_1$ (aggregation of AAII and II), but positive and insignificant for the survey index $SENT_2$. This finding is also consistent with the conclusion of Ben Aissia and Neffati (2022) who highlight that market-aggregated measures outperform survey indexes when estimating the effect of investor sentiment on stock returns. Tables 5 and 6 display also results for the dummy variable proxy of economic crisis. Results show that the associated coefficient estimates are negative and significant for both Y_{1t} and Y_{2t}. This finding stipulates that a financial crisis matters when the valuation of the relationship between stock market returns and investor sentiment (i.e., Canbaş and Kandır, 2009; Lutz, 2016; Bouteska, 2020). We then accept Hypothesis H_2 which emphasises that economic crisis predicts the relationship between investor sentiment and market returns.

Table 5 Vector auto-regression estimates (Y_{1t})

Independent variable	*Lag*	*Dependent variable*				
		AAII	*II*	*S&P500ret*	*DJXret*	*NQXret*
AAII	1	0.266***	0.0381	–0.00187	–0.000984	–0.00391
		(0.0568)	(0.0270)	(0.00276)	(0.00274)	(0.00417)
	2	0.125**	–0.0113	–0.00455*	–0.00536**	–0.00868**
		(0.0566)	(0.0269)	(0.00275)	(0.00272)	(0.00415)
II	1	–0.153	0.567***	–0.0122**	–0.0133**	–0.0156*
		(0.113)	(0.0538)	(0.00550)	(0.00545)	(0.00830)
	2	0.184*	0.239***	0.00640	0.00701	0.00874
		(0.106)	(0.0502)	(0.00513)	(0.00509)	(0.00775)
S&P500ret	1	–2.271	5.527**	0.292	0.234	0.272
		(4.764)	(2.262)	(0.231)	(0.229)	(0.349)
	2	2.325	–2.459	0.323	0.442*	0.281
		(4.820)	(2.289)	(0.234)	(0.232)	(0.353)
DJXret	1	–0.951	1.646	–0.331*	–0.287	–0.380
		(3.857)	(1.831)	(0.187)	(0.186)	(0.283)
	2	–3.387	2.120	–0.256	–0.299	–0.181
		(3.874)	(1.840)	(0.188)	(0.186)	(0.284)
NQXret	1	5.086***	0.403	0.0606	0.0571	0.139
		(1.464)	(0.695)	(0.0711)	(0.0705)	(0.107)
	2	1.361	0.163	–0.0536	–0.100	–0.0236
		(1.476)	(0.701)	(0.0717)	(0.0710)	(0.108)
DUMMY		–0.350**	0.0562	–0.0309***	–0.0319***	–0.0307**
		(0.176)	(0.0837)	(0.00856)	(0.00848)	(0.0129)
Constant		0.0223	–0.0553**	0.00862***	0.00888***	0.00985**
		(0.0528)	(0.0251)	(0.00256)	(0.00254)	(0.00387)

Notes: This table reports coefficient estimates of the model VAR model (Y_{1t}). The list of variable definitions and data sources is provided in the Appendix. All reported standard errors (SE) in parentheses are based on robust standard errors adjusted for heteroskedasticity and clustered at the firm level. *, ** and *** refer to significance at the 10%, 5%, and 1% levels, respectively.

Tables 7 and 8 report results for the decomposition of investor sentiment into rational and irrational components (i.e., Y_{3t} and Y_{4t}). We find that market returns are explained primarily by irrational sentiment. Indeed, results show that coefficients associated with the rational component (i.e., $AAII_{rat}$, II_{rat}, $SENT1_{rat}$, and $SENT2_{rat}$) are not significant while those associated with an irrational component are negative and significant for $SENT1_{irrat}$ and II_{irrat}. These findings are similar to those reported by Verma and Verma (2008) and Verma and Soydemir (2009). In light of these results, we accept Hypothesis H_3 which stipulates that it is the irrational component of the investor sentiment which predicts stock returns.

Table 6 Vector autoregression estimates (Y_{2t})

Independent variable	*Lag*	*Dependent variable*				
		SENT1	*SENT2*	*S&P500ret*	*DJXret*	*NQXret*
SENT1	1	0.439***	0.134	–0.0193***	–0.0202***	–0.0254***
		(0.0648)	(0.103)	(0.00579)	(0.00565)	(0.00925)
	2	0.185***	0.138	0.00105	0.00305	0.000820
		(0.0718)	(0.114)	(0.00641)	(0.00626)	(0.0102)
	3	0.0781	–0.0315	0.00706	0.00614	0.00528
		(0.0591)	(0.0940)	(0.00528)	(0.00516)	(0.00845)
SENT2	1	0.0052	0.244***	–0.000150	0.000196	0.00156
		(0.0412)	(0.0655)	(0.00368)	(0.00359)	(0.00588)
	2	–0.0118	0.225***	0.00273	0.00322	0.00486
		(0.0413)	(0.0656)	(0.00369)	(0.00360)	(0.00589)
	3	0.0519	0.300***	0.00267	0.00130	0.00414
		(0.0411)	(0.0654)	(0.00368)	(0.00359)	(0.00588)
S&P500ret	1	8.012***	–3.731	0.450*	0.324	0.472
		(3.037)	(4.829)	(0.271)	(0.265)	(0.434)
	2	–3.081	7.268	0.239	0.389	0.0475
		(3.080)	(4.897)	(0.275)	(0.269)	(0.440)
	3	–2.773	–3.762	0.706**	0.580**	1.334***
		(3.118)	(4.957)	(0.279)	(0.272)	(0.445)
DJXret	1	1.171	2.560	–0.477**	–0.381*	–0.593*
		(2.462)	(3.914)	(0.220)	(0.215)	(0.352)
	2	3.733	–4.669	–0.124	–0.198	0.0885
		(2.446)	(3.888)	(0.218)	(0.213)	(0.349)
	3	2.584	0.603	–0.447**	–0.366*	–0.723**
		(2.457)	(3.906)	(0.219)	(0.214)	(0.351)
NQXret	1	–0.270	0.247	0.0203	0.0369	0.0840
		(0.892)	(1.418)	(0.0797)	(0.0778)	(0.127)
	2	0.688	–1.770	–0.0446	–0.0939	0.00201
		(0.893)	(1.419)	(0.0798)	(0.0778)	(0.127)
	3	0.0929	1.373	–0.102	–0.101	–0.294**
		(0.896)	(1.424)	(0.0800)	(0.0781)	(0.128)
DUMMY		–0.0836	0.215	–0.0311***	–0.0314***	–0.0273
		(0.117)	(0.185)	(0.0104)	(0.0102)	(0.0167)
Constant		0.0263	–0.0720	0.0108***	0.0110***	0.0132**
		(0.0378)	(0.0601)	(0.00338)	(0.00330)	(0.00540)

Notes: This table reports coefficient estimates of the model VAR model (Y 2t). The list of variable definitions and data sources is provided in Appendix A. All reported standard errors (SE) in parentheses are based on robust standard errors adjusted for heteroskedasticity and clustered at the firm level. *, ** and *** refer to significance at the 10%, 5%, and 1% levels, respectively.

Table 7 Vector autoregression estimates (Y_{3t})

Independent variable	*Lag*	*Dependent variable*						
		AAIIrat	*AAIInoise*	*IIrat*	*IIrat*	*S&P500ret*	*DJXret*	*NQXret*
$AAII_{rat}$	1	0.616***	–0.206	–0.0393	–0.0848	0.0893	0.0863	0.0442
		(0.0807)	(0.183)	(0.0308)	(0.0990)	(0.0767)	(0.0760)	(0.116)
	2	0.0478	0.297	–0.0169	0.0399	–0.0925	–0.0803	–0.149
		(0.0804)	(0.182)	(0.0307)	(0.0986)	(0.0765)	(0.0757)	(0.116)
$AAII_{noise}$	1	–0.0159	0.259***	–0.00571	0.0621**	–0.0244	–0.0172	–0.0341
		(0.0241)	(0.0546)	(0.00918)	(0.0295)	(0.0229)	(0.0227)	(0.0346)
	2	–0.0115	0.107**	0.0162*	–0.00554	–0.0334	–0.0409*	–0.0635*
		(0.0238)	(0.0540)	(0.00908)	(0.0292)	(0.0226)	(0.0224)	(0.0342)
II_{rat}	1	–0.274*	–0.0789	0.587***	0.621***	–0.0763	–0.0609	–0.198
		(0.157)	(0.356)	(0.0599)	(0.193)	(0.149)	(0.148)	(0.226)
	2	0.283*	0.0758	0.382***	–0.562***	0.0391	0.0196	0.186
		(0.157)	(0.357)	(0.0601)	(0.193)	(0.150)	(0.148)	(0.227)
II_{noise}	1	–0.0923**	–0.0270	0.0191	0.482***	–0.0918**	–0.100**	–0.127*
		(0.0455)	(0.103)	(0.0174)	(0.0558)	(0.0433)	(0.0429)	(0.0655)
	2	0.0457	0.110	–0.00727	0.240***	0.0485	0.0529	0.0623
		(0.0411)	(0.0933)	(0.0157)	(0.0504)	(0.0391)	(0.0387)	(0.0592)

Notes: This table reports coefficient estimates of the model VAR model (Y_{3t}). The list of variable definitions and data sources is provided in the Appendix. All reported standard errors (SE) in parentheses are based on robust standard errors adjusted for heteroskedasticity and clustered at the firm level. *, **, and *** refer to significance at the 10%, 5%, and 1% levels, respectively.

Table 7 Vector autoregression estimates (Y_{3t}) (continued)

Independent variable	*Lag*	*Dependent variable*						
		AAIIrat	*AAIInoise*	*IIrat*	*IIrat*	*S&P500ret*	*DJXret*	*NQXret*
S&P500ret	1	–0.0587	–0.441	0.179*	0.931***	0.256	0.209	0.228
		(0.255)	(0.578)	(0.0972)	(0.313)	(0.242)	(0.240)	(0.367)
	2	0.651**	–0.419	0.0365	–0.429	0.330	0.434*	0.366
		(0.258)	(0.585)	(0.0985)	(0.317)	(0.245)	(0.243)	(0.371)
DJXret	1	–0.162	0.0286	0.0934	0.0566	–0.362*	–0.319*	–0.395
		(0.197)	(0.448)	(0.0754)	(0.242)	(0.188)	(0.186)	(0.284)
	2	–0.648***	0.0932	0.0222	0.278	–0.265	–0.311*	–0.183
		(0.198)	(0.450)	(0.0756)	(0.243)	(0.189)	(0.187)	(0.285)
NQXret	1	0.00222	0.694***	–0.0277	0.00547	0.0544	0.0509	0.136
		(0.0759)	(0.172)	(0.0290)	(0.0931)	(0.0722)	(0.0715)	(0.109)
	2	–0.0799	0.279	–0.0268	0.0322	–0.0388	–0.0843	–0.0285
		(0.0767)	(0.174)	(0.0293)	(0.0941)	(0.0729)	(0.0722)	(0.110)
DUMMY		–0.0365***	0.0145	0.00714*	–0.0116	–0.0286***	–0.0292***	–0.0314**
		(0.00973)	(0.0221)	(0.00371)	(0.0119)	(0.00926)	(0.00917)	(0.0140)
Constant		0.186***	–0.0506	0.0490***	–0.0245	0.0327	0.0303	0.0747
		(0.0449)	(0.102)	(0.0171)	(0.0551)	(0.0427)	(0.0423)	(0.0647)

Notes: This table reports coefficient estimates of the model VAR model (Y_{3t}). The list of variable definitions and data sources is provided in the Appendix. All reported standard errors (SE) in parentheses are based on robust standard errors adjusted for heteroskedasticity and clustered at the firm level. *, **, and *** refer to significance at the 10%, 5%, and 1% levels, respectively.

Table 8 Vector autoregression estimates (Y_{4t})

Independent variable	*Lag*	*Dependent variable*						
		SENT1rat	*SENT1noise*	*SENT2rat*	*SENT2rat*	*S&P500ret*	*DJXret*	*NQXret*
$SENT1_{rat}$	1	0.849***	−0.442**	0.193	0.128	−0.00131	−0.00645	0.00833
		(0.0698)	(0.186)	(0.122)	(0.257)	(0.0162)	(0.0158)	(0.0258)
	2	−0.0574	0.603**	0.0205	−0.217	−0.0155	−0.00511	−0.0538
		(0.0909)	(0.243)	(0.158)	(0.334)	(0.0211)	(0.0206)	(0.0337)
	3	0.0798	−0.264	−0.0801	0.188	0.00380	−0.000838	0.0204
		(0.0654)	(0.174)	(0.114)	(0.241)	(0.0152)	(0.0148)	(0.0242)
$SENT1_{noise}$	1	0.0565**	0.362***	0.175***	−0.0886	−0.0193***	−0.0195***	−0.0272***
		(0.0259)	(0.0691)	(0.0451)	(0.0953)	(0.00602)	(0.00588)	(0.00959)
	2	−0.0186	0.160**	−0.00495	0.117	0.00202	0.00393	0.00328
		(0.0280)	(0.0747)	(0.0487)	(0.103)	(0.00651)	(0.00635)	(0.0104)
	3	0.0584**	0.0154	−0.142***	0.0683	0.00795	0.00730	0.00687
		(0.0238)	(0.0635)	(0.0415)	(0.0876)	(0.00554)	(0.00541)	(0.00882)
$SENT2_{rat}$	1	0.0467	0.115	0.337***	−0.550***	−0.00239	−0.000102	0.0128
		(0.0570)	(0.152)	(0.0993)	(0.210)	(0.0133)	(0.0129)	(0.0211)
	2	−0.0797	−0.0776	0.376***	0.193	0.00598	0.00244	0.00521
		(0.0566)	(0.151)	(0.0985)	(0.208)	(0.0132)	(0.0128)	(0.0209)
	3	0.107*	−0.0243	0.191*	0.462**	0.000584	−0.000560	−0.00481
		(0.0566)	(0.151)	(0.0986)	(0.208)	(0.0132)	(0.0129)	(0.0210)
$SENT2_{noise}$	1	−0.0256	−0.0109	0.000779	0.191***	0.00106	0.00154	0.00281
		(0.0171)	(0.0456)	(0.0298)	(0.0629)	(0.00398)	(0.00388)	(0.00633)
	2	−0.0129	−0.00140	0.0286	0.109*	0.00320	0.00453	0.00422
		(0.0171)	(0.0456)	(0.0298)	(0.0629)	(0.00398)	(0.00388)	(0.00634)
	3	−0.0102	0.0345	−0.0155	0.241***	0.00353	0.00263	0.00403
		(0.0170)	(0.0453)	(0.0296)	(0.0625)	(0.00395)	(0.00386)	(0.00629)

Notes: This table reports coefficient estimates of the model VAR model (Y_{4t}). The list of variable definitions and data sources is provided in the Appendix. All reported standard errors (SE) in parentheses are based on robust standard errors adjusted for heteroskedasticity and clustered at the firm level. *, **, and *** refer to significance at the 10%, 5%, and 1% levels, respectively.

Table 8 Vector autoregression estimates (Y_{4t}) (continued)

Independent variable	*Lag*	*Dependent variable*						
		SENT1rat	*SENT1noise*	*SENT2rat*	*SENT2rat*	*S&P500ret*	*DJXret*	*NQXret*
S&P500ret	1	4.723***	2.467	−0.836	−3.520	0.500*	0.365	0.566
		(1.176)	(3.135)	(2.047)	(4.324)	(0.273)	(0.267)	(0.435)
	2	1.402	−4.183	−0.987	6.966	0.158	0.341	−0.174
		(1.234)	(3.290)	(2.148)	(4.538)	(0.287)	(0.280)	(0.457)
	3	1.261	−5.845*	−4.516**	0.421	0.672**	0.544*	1.359***
		(1.243)	(3.316)	(2.165)	(4.573)	(0.289)	(0.282)	(0.460)
DJXret	1	–0.273	2.663	1.031	1.068	−0.490**	−0.393*	−0.548
		(0.975)	(2.599)	(1.697)	(3.585)	(0.227)	(0.221)	(0.361)
	2	–1.475	5.579**	1.107	−4.536	−0.111	−0.206	0.152
		(0.973)	(2.595)	(1.694)	(3.579)	(0.226)	(0.221)	(0.360)
	3	0.686	2.573	3.830**	−0.0864	−0.428*	−0.370*	−0.695*
		(0.964)	(2.571)	(1.678)	(3.545)	(0.224)	(0.219)	(0.357)
NQXret	1	–0.742**	0.812	0.404	−0.954	0.005	04	0.0274
		(0.356)	(0.949)	(0.620)	(1.309)	(0.0828)	(0.0808)	(0.132)
	2	–0.614*	0.988	0.581	−1.687	−0.0277	−0.0877	0.0271
		(0.359)	(0.958)	(0.626)	(1.322)	(0.0835)	(0.0815)	(0.133)
	3	–0.626*	1.114	0.860	0.816	−0.0985	−0.0949	−0.308**
		(0.359)	(0.957)	(0.625)	(1.320)	(0.0834)	(0.0815)	(0.133)
DUMMY		0.0264	−0.0976	0.129	−0.0132	−0.0305***	−0.0302***	−0.0275
		(0.0462)	(0.123)	(0.0805)	(0.170)	(0.0107)	(0.0105)	(0.0171)
Constant		–0.007	39	0.0395	−0.0427	0.0354	0.0105***	0.0103***
		(0.0154)	(0.0410)	(0.0267)	(0.0565)	(0.00357)	(0.00349)	(0.00569)

Notes: This table reports coefficient estimates of the model VAR model (Y_{4t}). The list of variable definitions and data sources is provided in the Appendix. All reported standard errors (SE) in parentheses are based on robust standard errors adjusted for heteroskedasticity and clustered at the firm level. *, **, and *** refer to significance at the 10%, 5%, and 1% levels, respectively.

However, the various estimates of the VAR models used as well as the coefficients obtained do not allow us to decide concerning the sense of the sentiment-performance relationship. Consequently, we deploy a linear causality test in the sense of Granger to determine the direction of a possible cause-and-effect relationship between sentiment and market return.

Table 9 WALD causality tests

Null hypothesis (H0)	*Wald test: Y_{1t}*		*Null hypothesis (H0)*	*Wald test: Y_{2t}*	
	χ^2-statistics	*p-values*		*χ^2-statistics*	*p-values*
II~AAII	3.808	0.149	SENT2~SENT1	2.449	0.485
S&P500ret~AAII	0.821	0.663	S&P500ret~SENT1	7.810**	0.050
DJAret~AAII	1.481	0.477	DJAret~SENT1	5.366	0.147
NQXret~AAII	12.165***	0.002	NQXret~SENT1	0.546	0.909
AAII~II	2.746	0.253	SENT1~SENT2	7.201	0.066
S&P500ret~II	5.940	0.051	S&P500ret~SENT2	2.515	0.473
DJAret~II	3.199	0.202	DJAret~SENT2	1.717	0.633
NQXret~II	0.468	0.791	NQXret~SENT2	2.148	0.542
AAII~S&P500ret	2.391	0.303	SENT1~S&P500ret	10.671**	0.014
II~S&P500ret	5.391	0.068	SENT2~S&P500ret	1.456	0.693
DJAret~S&P500ret	4.442	0.109	DJAret~S&P500ret	7.161	0.067
NQXret~S&P500ret	1.604	0.448	NQXret~S&P500ret	2.188	0.534
AAII~DJAret	3.603	0.165	SENT1~DJAret	11.508***	0.009
II~DJAret	5.963	0.051	SENT2~DJAret	1.237	0.744
S&P500ret~DJAret	5.700	0.058	S&P500ret~DJAret	9.320**	0.025
NQXret~DJAret	2.902	0.226	NQXret~DJAret	3.603	0.308
AAII~NQXret	5.708	0.058	SENT1~NQXret	7.743	0.052
II~NQXret	4.162	0.125	SENT2~NQXret	3.017	0.389
S&P500ret~NQXret	1.784	0.410	S&P500ret~NQXret	9.256**	0.026
DJAret~NQXret	1.879	0.391	DJAret~NQXret	5.446	0.142

Notes: This table reports the χ^2-statistics of the Wald causality tests. The list of variable definitions and data sources is provided in the Appendix. All reported p-values in parentheses are based on robust standard errors adjusted for heteroskedasticity. *, ** and *** refer to significance at the 10%, 5%, and 1% levels, respectively.

5.1 Granger tests results

To study the possibility of causality in the Granger sense between the variables (sentiment-stocks returns), we apply a Wald χ^2 test. Table 9 displays the results of the causality test for the various models retained. The implementation of the Granger non-causality test is performed under SATA 15.1 via the VAR-Granger procedure. Results show that for the AAII sentiment index, the associated probabilities are greater than 0.05. Thus, the null hypothesis that sentiment does not Granger cause returns cannot be rejected. Still, we cannot reject the null hypothesis that returns (i.e., S&P500ret, DJXret) do not cause AAII except for small firms (NQXret). Indeed, We find a p-value of 0.002 associated with the relation AAII-NQXret Therefore, for AAII, returns explain the individual investor sentiment for small firms whereas investor sentiment does not influence stock returns at all.

Still, Table 8 reports p-values greater than 0.05 for the institutional investors in both directions. These results stipulate an absence of a possible causality linkage between the sentiment of institutional investors and returns. These results confirm those of Brown and Cliff (2004) and Kling and Gao (2008), who find similar results for the USA and China. Nevertheless, our results contradict those of other studies (i.e., Fisher and Statman, 2000; Baker and Wurgler, 2006 and Canbaş and Kandır, 2009).

To better argue these findings, we need to examine the composite indexes results. We predict to have better results since related works (DeVault et al., 2019; Ben Aissia and Neffati, 2022) find that composite indexes outperform individual investor sentiment measures in explaining stock returns. Results show a bi-directional causality for $SENT_1$. Indeed, we find a p-value of 0.009 for mid-sized firms (i.e., DJAret) and a p-value of 0.026 for small firms (i.e., NQXret). However, the opposite causality is not verified. Still, no relationship exists between $SENT_2$ and performance. This means that only $SENT_1$ variable contains useful information to predict performance. To conclude, we find that the causality linkage between investor sentiment and stock returns is verified in both senses. If market returns predict individual investor measures for small stocks, the market composite index explains returns for small and medium-sized stocks.

5.2 IRFs results

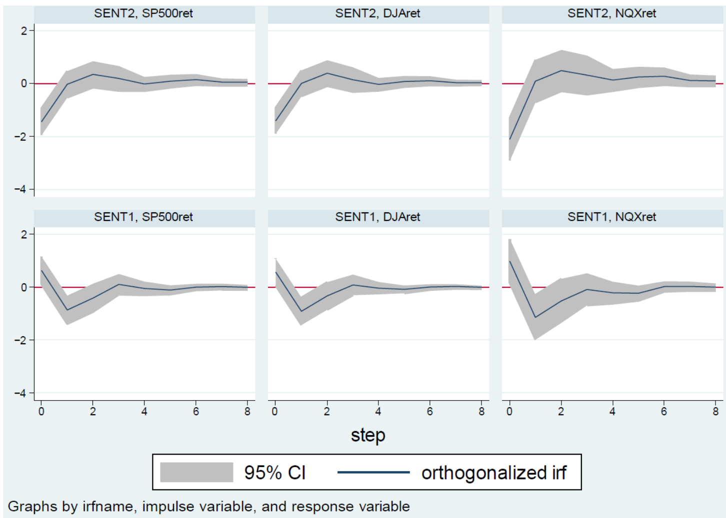
Figure 1 shows the response of market returns following a shock affecting one of the sentiment factors over a horizon of 12 months. Indeed, we represent the response of each variable to a shock of 1%. We distinguish between the shocks affecting the indirect measures of sentiment ($SENT_1$ and $SENT_2$) and the direct measures (AAII and II). We also decompose each of these indicators into rational and irrational components.

5.2.1 Effect of a shock affecting SENT1, SENT2, AAII and II

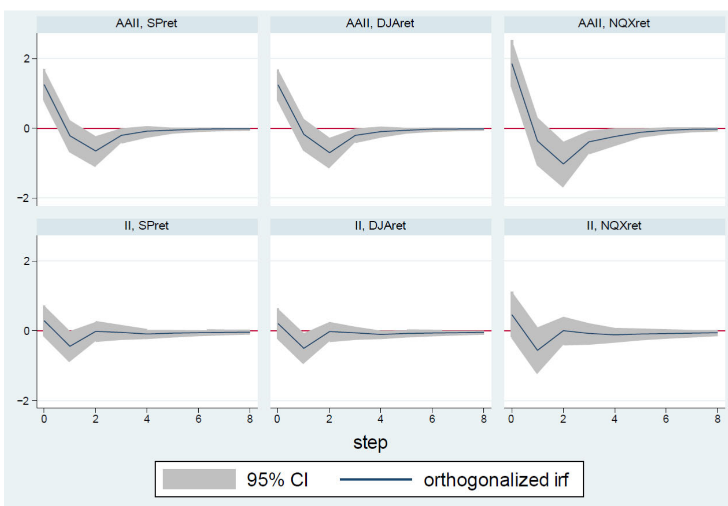
As expected, we find that a positive sentiment shock in $SENT_1$ leads to a decrease in aggregate returns (S&P500ret, DJAret, and NQXret) during the first month. The response of the different sentiment measures is nearly identical [Figure 1(a)]. Still, the impact of the shock is more pronounced for small firms (NQXret).

However, Figure 1(a) shows that a shock in $SENT_2$ has a positive and significant effect during the first month and not significant during the following months. The adjustment to a steady state is quite fast for large and medium-sized firms (S&P500ret and DJAret). Figure 1(b) reports the response of market index returns to a shock affecting individual (AAII) and institutional (II) sentiment measures. The results are similar to those of the aggregate index SENT1 with a more pronounced effect for AAII sentiment. Our conclusions are then similar to those reported by Verma and Verma (2008), who find that the sentiment variable is a contrarian factor of short-term stock performance. Our conclusions confirm our previous finding of the superiority of the survey indexes to better predict the causality link between investor sentiment and stock market returns.

Figure 1 Return S&P500ret, DJXret and NQXret response to US sentiment shock, (a) return response to US investor sentiment shock SENT1 and SENT2 (b) return response to US investor sentiment shock AAII and II, (c) return response to US investor sentiment shock $\text{AAII}_{\text{rational}}$ (d) return response to US investor sentiment shock $\text{AAII}_{\text{irrational}}$ (e) return response to US investor sentiment shock $\text{II}_{\text{rational}}$ (f) return response to US investor sentiment shock $\text{II}_{\text{irrational}}$ (g) return response to US investor sentiment shock $\text{SENT1}_{\text{rational}}$ (h) Return response to US investor sentiment shock $\text{SENT1}_{\text{irrational}}$ (i) return response to US investor sentiment shock $\text{SENT2}_{\text{rational}}$ (j) return response to US investor sentiment shock $\text{SENT2}_{\text{irrational}}$ (see online version for colours)



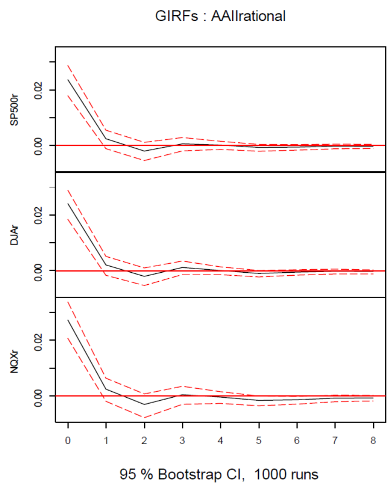
(a)



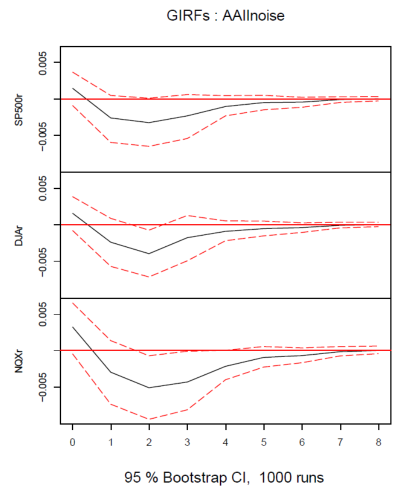
(b)

Notes: US returns impulse response functions to US investor sentiment with two standard error bands. The dashed lines on each graph represent the upper and lower 95% confidence band. When the upper and lower bands carry the same sign, the response becomes statistically significant. On each graph, percentage returns are plotted on the vertical axis, and time on the horizontal axis.

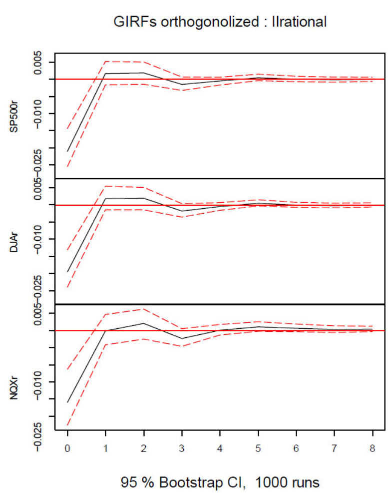
Figure 1 Return S&P500ret, DJXret and NQXret response to US sentiment shock, (a) return response to US investor sentiment shock SENT1 and SENT2 (b) return response to US investor sentiment shock AAII and II, (c) return response to US investor sentiment shock $\text{AAII}_{\text{rational}}$ (d) return response to US investor sentiment shock $\text{AAII}_{\text{irrational}}$ (e) return response to US investor sentiment shock $\text{II}_{\text{rational}}$ (f) return response to US investor sentiment shock $\text{II}_{\text{irrational}}$ (g) return response to US investor sentiment shock $\text{SENT1}_{\text{rational}}$ (h) Return response to US investor sentiment shock $\text{SENT1}_{\text{irrational}}$ (i) return response to US investor sentiment shock $\text{SENT2}_{\text{rational}}$ (j) return response to US investor sentiment shock $\text{SENT2}_{\text{irrational}}$ (continued) (see online version for colours)



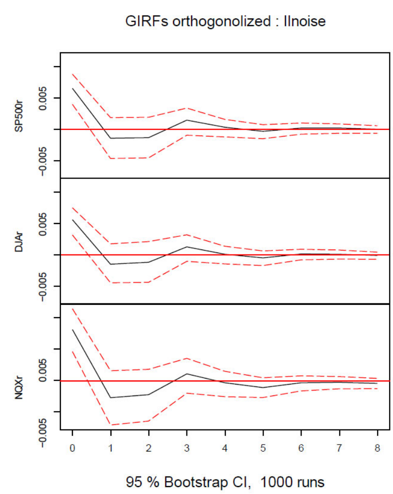
(c)



(d)



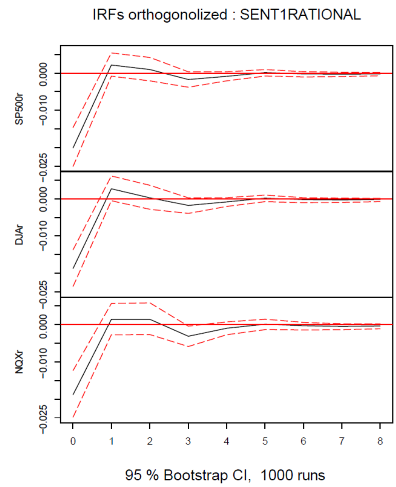
(e)



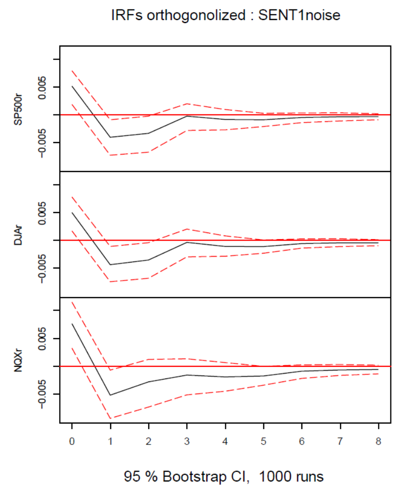
(f)

Notes: US returns impulse response functions to US investor sentiment with two standard error bands. The dashed lines on each graph represent the upper and lower 95% confidence band. When the upper and lower bands carry the same sign, the response becomes statistically significant. On each graph, percentage returns are plotted on the vertical axis, and time on the horizontal axis.

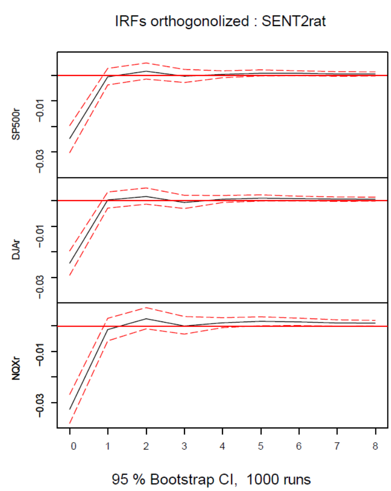
Figure 1 Return S&P500ret, DJXret and NQXret response to US sentiment shock, (a) return response to US investor sentiment shock SENT1 and SENT2 (b) return response to US investor sentiment shock AAII and II, (c) return response to US investor sentiment shock $AAII_{rational}$ (d) return response to US investor sentiment shock $AAII_{irrational}$ (e) return response to US investor sentiment shock $II_{rational}$ (f) return response to US investor sentiment shock $II_{irrational}$ (g) return response to US investor sentiment shock $SENT1_{rational}$ (h) Return response to US investor sentiment shock $SENT1_{irrational}$ (i) return response to US investor sentiment shock $SENT2_{rational}$ (j) return response to US investor sentiment shock $SENT2_{irrational}$ (continued) (see online version for colours)



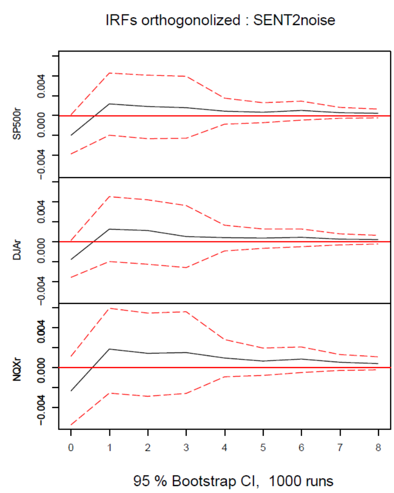
(g)



(h)



(i)



(j)

Notes: US returns impulse response functions to US investor sentiment with two standard error bands. The dashed lines on each graph represent the upper and lower 95% confidence band. When the upper and lower bands carry the same sign, the response becomes statistically significant. On each graph, percentage returns are plotted on the vertical axis, and time on the horizontal axis.

Table 10 Forecast error variance decomposition for *VAR* model (Y_{1t} and Y_{2t})

Horizon	*Model Y_{1t}: Choc AAII*					*Model Y_{1t}: Choc II*				
	AAII	*II*	*SP500r*	*DJAr*	*NQXr*	*AAII*	*II*	*SP500r*	*DJAr*	*NQXr*
1	1.00	0.00	0.00	0.00	0.00	0.02	0.98	0.00	0.00	0.00
2	0.96	0.00	0.02	0.00	0.02	0.05	0.78	0.16	0.00	0.00
3	0.94	0.00	0.02	0.01	0.03	0.06	0.75	0.18	0.00	0.01
4	0.94	0.00	0.02	0.01	0.03	0.05	0.75	0.18	0.00	0.01
5	0.93	0.00	0.02	0.01	0.03	0.05	0.75	0.19	0.00	0.00
6	0.93	0.00	0.02	0.01	0.03	0.05	0.74	0.20	0.00	0.00
7	0.93	0.00	0.02	0.01	0.03	0.05	0.74	0.20	0.00	0.00
8	0.93	0.00	0.02	0.01	0.03	0.05	0.74	0.20	0.00	0.00
9	0.93	0.00	0.02	0.01	0.03	0.05	0.74	0.20	0.00	0.00
10	0.93	0.00	0.02	0.01	0.03	0.05	0.74	0.20	0.00	0.00
11	0.93	0.00	0.02	0.01	0.03	0.05	0.74	0.21	0.00	0.00
12	0.93	0.00	0.02	0.01	0.03	0.05	0.74	0.21	0.00	0.00
Horizon	*Model Y_{2t}: Choc SENT1*					*Model Y_{2t}: Choc SENT2*				
	SENT1	*SENT2*	*SP500r*	*DJAr*	*NQXr*	*SENT1*	*SENT2*	*SP500r*	*DJAr*	*NQXr*
1	1.00	0.00	0.00	0.00	0.00	0.01	0.99	0.00	0.00	0.00
2	0.82	0.02	0.15	0.00	0.01	0.01	0.99	0.00	0.00	0.00
3	0.78	0.03	0.18	0.00	0.02	0.01	0.98	0.00	0.00	0.01
4	0.78	0.02	0.18	0.00	0.02	0.02	0.97	0.00	0.00	0.00
5	0.77	0.02	0.19	0.00	0.02	0.02	0.97	0.00	0.00	0.01
6	0.76	0.02	0.20	0.00	0.02	0.02	0.97	0.00	0.00	0.01
7	0.76	0.02	0.20	0.00	0.02	0.02	0.97	0.00	0.00	0.01
8	0.76	0.02	0.21	0.00	0.02	0.02	0.97	0.00	0.00	0.01
9	0.76	0.02	0.21	0.00	0.02	0.03	0.96	0.00	0.00	0.01
10	0.76	0.02	0.21	0.00	0.02	0.03	0.96	0.00	0.00	0.01
11	0.75	0.02	0.21	0.00	0.02	0.03	0.96	0.00	0.00	0.01
12	0.75	0.02	0.21	0.00	0.02	0.03	0.96	0.00	0.00	0.01

Table 11 Forecast error variance decomposition for *VAR* model (Y_{3t})

Horizon	*Model Y_{3t}: Choc $AAII_{rat}$*							*Model Y_{3t}: Choc $AAII_{noise}$*						
	$AAII_{rat}$	AA_{noise}	$II_{rational}$	II_{noise}	*SP500r*	*DJAr*	*NQXr*	$AAII_{rat}$	$AAII_{noise}$	$II_{rational}$	II_{noise}	*SP500r*	*DJAr*	*NQXr*
1	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.28	0.00	0.72	0.00	0.00	0.00	0.00
2	0.00	0.96	0.00	0.00	0.01	0.00	0.03	0.35	0.00	0.61	0.00	0.02	0.01	0.01
3	0.00	0.94	0.00	0.00	0.01	0.00	0.04	0.34	0.00	0.60	0.00	0.02	0.02	0.01
4	0.00	0.93	0.00	0.00	0.01	0.00	0.05	0.34	0.00	0.61	0.00	0.02	0.02	0.01
5	0.01	0.93	0.00	0.00	0.01	0.00	0.05	0.33	0.00	0.61	0.00	0.02	0.02	0.01
6	0.01	0.93	0.00	0.00	0.01	0.00	0.05	0.33	0.00	0.61	0.00	0.03	0.02	0.02
7	0.01	0.93	0.00	0.00	0.01	0.00	0.05	0.32	0.00	0.61	0.00	0.03	0.02	0.02
8	0.01	0.93	0.00	0.00	0.01	0.00	0.05	0.32	0.00	0.61	0.00	0.03	0.02	0.02
9	0.01	0.92	0.00	0.00	0.01	0.00	0.05	0.31	0.00	0.62	0.00	0.03	0.02	0.02
10	0.01	0.92	0.00	0.00	0.01	0.00	0.05	0.31	0.00	0.62	0.00	0.03	0.02	0.02
11	0.01	0.92	0.00	0.00	0.01	0.00	0.05	0.31	0.00	0.62	0.00	0.03	0.02	0.02
12	0.01	0.92	0.00	0.00	0.01	0.00	0.05	0.30	0.00	0.62	0.00	0.04	0.02	0.02
Horizon	*Model Y_{3t}: Choc II_{rat}*							*Model Y3t : Choc II_{noise}*						
	$AAII_{rat}$	$AAII_{noise}$	$II_{rational}$	II_{noise}	*SP500r*	*DJAr*	*NQXr*	$AAII_{rat}$	$AAII_{noise}$	$II_{rational}$	II_{noise}	*SP500r*	*DJAr*	*NQXr*
1	0.28	0.00	0.72	0.00	0.00	0.00	0.00	0.01	0.02	0.04	0.93	0.00	0.00	0.00
2	0.35	0.00	0.61	0.00	0.02	0.01	0.01	0.09	0.03	0.07	0.74	0.07	0.00	0.00
3	0.34	0.00	0.60	0.00	0.02	0.02	0.01	0.12	0.02	0.07	0.70	0.09	0.00	0.01
4	0.34	0.00	0.61	0.00	0.02	0.02	0.01	0.12	0.02	0.07	0.69	0.09	0.00	0.01
5	0.33	0.00	0.61	0.00	0.02	0.02	0.01	0.12	0.02	0.07	0.68	0.10	0.00	0.01
6	0.33	0.00	0.61	0.00	0.03	0.02	0.02	0.12	0.02	0.07	0.68	0.10	0.00	0.01
7	0.32	0.00	0.61	0.00	0.03	0.02	0.02	0.12	0.02	0.07	0.68	0.11	0.00	0.01
8	0.32	0.00	0.61	0.00	0.03	0.02	0.02	0.12	0.02	0.07	0.67	0.11	0.00	0.01
9	0.31	0.00	0.62	0.00	0.03	0.02	0.02	0.12	0.02	0.07	0.67	0.11	0.00	0.01
10	0.31	0.00	0.62	0.00	0.03	0.02	0.02	0.12	0.02	0.07	0.67	0.11	0.00	0.01
11	0.31	0.00	0.62	0.00	0.03	0.02	0.02	0.12	0.02	0.07	0.67	0.11	0.00	0.01
12	0.30	0.00	0.62	0.00	0.04	0.02	0.02	0.12	0.02	0.07	0.67	0.11	0.00	0.01

Table 12 Forecast error variance decomposition for *VAR* model (Y_{4t})

Horizon	Model Y_{4t}: Choc $SENT1_{rat}$							Model Y_{4t}: Choc $SENT1_{noise}$						
	$SENT1_{rat}$	$SENT1_{noise}$	$SENT2_{rational}$	$SENT2_{noise}$	$SP500r$	$DJAr$	$NQXr$	$SENT1_{rat}$	$SENT1_{noise}$	$SENT2_{rational}$	$SENT2_{noise}$	$SP500r$	$DJAr$	$NQXr$
1	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.97	0.00	0.00	0.00	0.00	0.00
2	0.77	0.01	0.09	0.01	0.13	0.00	0.00	0.08	0.85	0.03	0.00	0.03	0.00	0.01
3	0.65	0.01	0.12	0.01	0.20	0.00	0.00	0.08	0.83	0.03	0.00	0.04	0.00	0.02
4	0.58	0.01	0.12	0.01	0.26	0.00	0.00	0.07	0.83	0.03	0.00	0.04	0.00	0.02
5	0.53	0.02	0.12	0.02	0.31	0.00	0.00	0.08	0.83	0.03	0.00	0.04	0.00	0.02
6	0.48	0.02	0.12	0.02	0.35	0.00	0.00	0.08	0.83	0.03	0.00	0.04	0.00	0.02
7	0.45	0.02	0.12	0.02	0.38	0.00	0.00	0.08	0.82	0.03	0.00	0.04	0.00	0.02
8	0.43	0.03	0.11	0.02	0.41	0.00	0.00	0.08	0.82	0.03	0.01	0.04	0.00	0.02
9	0.41	0.03	0.11	0.02	0.43	0.00	0.00	0.08	0.82	0.03	0.01	0.04	0.00	0.02
10	0.39	0.03	0.10	0.02	0.45	0.00	0.00	0.08	0.82	0.03	0.01	0.04	0.00	0.02
11	0.37	0.03	0.09	0.02	0.47	0.00	0.00	0.08	0.82	0.03	0.01	0.04	0.00	0.02
12	0.36	0.03	0.09	0.02	0.49	0.00	0.00	0.08	0.82	0.03	0.01	0.04	0.00	0.02
Horizon	Model Y_{4t}: Choc $SENT2_{rat}$							Model Y4t: Choc $SENT2_{noise}$						
	$SENT1_{rat}$	$SENT1_{noise}$	$SENT2_{rational}$	$SENT2_{noise}$	$SP500r$	$DJAr$	$NQXr$	$SENT1_{rat}$	$SENT1_{noise}$	$SENT2_{rational}$	$SENT2_{noise}$	$SP500r$	$DJAr$	$NQXr$
1	0.09	0.01	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00
2	0.08	0.01	0.89	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.97	0.00	0.00	0.00
3	0.08	0.02	0.83	0.00	0.05	0.01	0.01	0.00	0.01	0.01	0.96	0.00	0.00	0.01
4	0.07	0.02	0.79	0.00	0.09	0.01	0.01	0.00	0.01	0.01	0.96	0.00	0.00	0.01
5	0.06	0.02	0.77	0.00	0.12	0.01	0.01	0.00	0.01	0.01	0.96	0.00	0.00	0.01
6	0.06	0.02	0.75	0.00	0.15	0.01	0.01	0.00	0.01	0.01	0.96	0.00	0.00	0.01
7	0.06	0.02	0.73	0.00	0.17	0.01	0.01	0.00	0.01	0.01	0.96	0.00	0.00	0.01
8	0.05	0.02	0.71	0.00	0.19	0.01	0.01	0.00	0.01	0.01	0.96	0.00	0.00	0.01
9	0.05	0.02	0.70	0.00	0.21	0.01	0.01	0.00	0.02	0.01	0.96	0.00	0.00	0.01
10	0.05	0.01	0.68	0.00	0.23	0.01	0.01	0.00	0.02	0.01	0.96	0.00	0.00	0.01
11	0.05	0.01	0.67	0.00	0.24	0.01	0.01	0.00	0.02	0.01	0.96	0.00	0.00	0.01
12	0.05	0.01	0.66	0.00	0.26	0.01	0.01	0.00	0.02	0.01	0.96	0.00	0.00	0.01

5.2.2 Effect of a shock affecting the rational and irrational component

Figures 1(c), 1(d), 1(e) and 1(f) report the impulse responses of S&P500, DJA, and NQX indexes return following an increase of 1% in rational and irrational components of individual investor sentiment (AAII) and (II).

We find that the effect of the irrational component is negative and significant during the first month (and zero during the other periods), for the AAII and II sentiment indexes. However, the response to the rational component is positive and significant during the first two months (and insignificant during the remaining months), for the AAII index.

Results reported in Figures 1(g) and 1(i) for aggregate indices $SENT_1$ and $SENT_2$ are similar to our previous results. Indeed, only the irrational component for the $SENT_1$ is negative and significant, as it is predicted in related relevant research. We conclude that our study of the impulse response functions confirms that it is the irrational component of investor sentiment that affects stock market returns.

5.3 FEVD results

Tables 10–12 report the results of the FEVD analysis. We find the S&P500ret is mainly determined by the shocks of the sentiment index (II) and the shock of the index SENT1 (respectively 18% and 20%). For the returns of the DJA and NQX indexes, the contribution of the shock of the two measures is about 3%.

Still, the decomposition of each INDEX into rational and irrational components shows that the contribution of the rational component of the AAIIrat index is not significant for Big and mid-size firms (i.e., S&P500 and DJA) and does not exceed 5% for the NQX. However, the irrational component of AAII explains respectively 4% and 2% for the returns of the S&P500 and NQX.

For the institutional investors (II) index, the forecast error variance decomposition analysis shows that the shock of the irrational component (II_{irr}) dominates the fluctuations of SP500ret, with a share of 11%.

However, the results of the aggregate sentiment indices ($SENT_1$ and $SENT_2$) are opposite to our previous results. Indeed, the fluctuations of large companies (i.e., S&P500) are essentially explained by the rational component of sentiment. It is around 49% for $SENT1_{rat}$ and 26% for $SENT2_{rat}$.

6 Conclusions

This paper examines the causality linkage between investor sentiment and stock market returns on the US market, using different indexes based on market and survey data. Moreover, we examine the effect of the economic crisis on this relationship and decompose sentiment into rational and irrational components. We use VAR models and Granger tests, estimate the impulse response functions (IRFs) of the non-expected movement in investor sentiment and propose a forecast error variance decomposition (FEVD) approach to emphasize the importance of these movements on variables of the VAR models. Using a sample of US data (S&P 500, Dow Jones, and NASDAQ indexes) over the period July 1965 to December 2019, our results reveal a negative and significant relationship between investor sentiment and stock returns. This relationship is more significant for survey indexes than for market indexes. Also, we show that rational component variables are not significant while irrational component variables are negative and significant. In addition, we find a bi-directional Granger causality between stock returns and investor sentiment. Still, our study of the impulse response functions confirms the superiority of the survey indexes over the market indexes and that it is the irrational component of investor sentiment that affects stock market returns. Finally, we find using our FEVD analysis that the fluctuation in S&P500 returns is mainly due to the sentiment of institutional investors rather than that of individual investors. The irrational component of institutional investors' index contributes largely to variations in returns over a 12-month horizon.

References

Akaike, H. (1974) 'A new look at the statistical model identification', *IEEE Trans. Automat. Contr.*, Vol. AC-19, pp.716–723.

Atukeren, E., Korkmaz, T. and C¸evik, E.İ. (2013) 'Spillovers between business confidence and stock returns in Greece, Italy, Portugal, and Spain', *International Journal of Finance & Economics*, Vol. 18, No. 3, pp.205–215.

Baker, M. and Wurgler, J. (2006) 'Investor sentiment and the cross-section of stock returns', *The Journal of Finance*, Vol. 61, No. 4, pp.1645–1680.

Baker, M. and Wurgler, J. (2007) 'Investor sentiment in the stock market', *Journal of Economic Perspectives*, Vol. 21, No. 2, pp.129–152.

Baker, M., Wurgler, J. and Yuan, Y. (2012) 'Global, local, and contagious investor sentiment', *Journal of Financial Economics*, Vol. 104, No. 2, pp.272–287.

Barber, B.M., Odean, T. and Zhu, N. (2009) 'Systematic noise', *Journal of Financial Markets*, Vol. 12, No. 4, pp.547–569.

Bathia, D. and Bredin, D. (2013) 'An examination of investor sentiment effect on G7 stock market returns', *The European Journal of Finance*, Vol. 19, No. 9, pp.909–937.

Beaumont, R., van Daele, M., Frijns, B., Lehnert, T. and Muller, A. (2008) 'Investor sentiment, mutual fund flows and its impact on returns and volatility', *Managerial Finance*, Vol. 34, No. 11, pp.772–785.

Beer, F., Zouaoui, M. et al. (2013) 'Measuring stock market investor sentiment', *Journal of Applied Business Research (JABR)*, Vol. 29, No. 1, pp.51–68.

Ben Aissia, D. and Neffati, N. (2022) 'Investor sentiment, bull and bear markets and stock returns', *International Journal of Banking, Accounting, and Finance*, Vol. 13, No. 1, pp.115–144.

Bouteska, A. (2020) 'Some evidence from a principal component approach to measure a new investor sentiment index in the Tunisian stock market', *Managerial Finance*, Vol. 49, No. 3, pp.1129–1145.

Brown, G.W. and Cliff, M.T. (2004) 'Investor sentiment and the near-term stock market', *Journal of Empirical Finance*, Vol. 11, No. 1, pp.1–27.

Brown, G.W. and Cliff, M.T. (2005) 'Investor sentiment and asset valuation', *The Journal of Business*, Vol. 78, No. 2, pp.405– 440.

Cagli, E.C., Ergün, Z.C. and Durukan, M.B. (2020) 'The causal linkages between investor sentiment and excess returns on Borsa Istanbul', *Borsa Istanbul Review*, Vol. 20, No. 3, pp.214–223.

Campbell, J.Y. (1987) 'Stock returns and the term structure', *Journal of Financial Economics*, Vol. 18, No. 2, pp.373–399.

Campbell, J.Y. (1991) 'A variance decomposition for stock returns', *The Economic Journal*, Vol. 101, No. 405, pp.157–179.

Campbell, J.Y. and Shiller, R.J. (1988a) 'Stock prices, earnings, and expected dividends', *The Journal of Finance*, Vol. 43, No. 3, pp.661–676.

Campbell, J.Y. and Shiller, R.J. (1988b) 'The dividend-price ratio and expectations of future dividends and discount factors', *The Review of Financial Studies*, Vol. 1, No. 3, pp.195–228.

Canbaş, S. and Kandır, S.Y. (2009) 'Investor sentiment and stock returns: evidence from Turkey', *Emerging Markets Finance and Trade*, Vol. 45, No. 4, pp.36–52.

Chu, X., Wu, C. and Qiu, J. (2016) 'A nonlinear Granger causality test between stock returns and investor sentiment for Chinese stock market: a wavelet-based approach', *Applied Economics*, Vol. 48, No. 21, pp.1915–1924.

Chung, S-L., Hung, C-H. and Yeh, C-Y. (2012) 'When does investor sentiment predict stock returns?', *Journal of Empirical Finance*, Vol. 19, No. 2, pp.217–240.

Clarke, R.G. and Statman, M. (1998) 'Bullish or bearish?', *Financial Analysts Journal*, Vol. 54, No. 3, pp.63–72.

Da, Z., Engelberg, J. and Gao, P. (2011) 'In search of attention', *The Journal of Finance*, Vol. 66, No. 5, pp.1461–1499.

De Long, J.B., Shleifer, A., Summers, L.H. and Waldmann, R.J. (1990) 'Noise trader risk in financial markets', *Journal of Political Economy*, Vol. 98, No. 4, pp.703–738.

DeVault, L., Sias, R. and Starks, L. (2019) 'Sentiment metrics and investor demand', *The Journal of Finance*, Vol. 74, No. 2, pp.985–1024.

Ding, W., Mazouz, K. and Wang, Q. (2019) 'Investor sentiment and the cross-section of stock returns: new theory and evidence', *Review of Quantitative Finance and Accounting*, Vol. 53, No. 2, pp.493–525.

Fama, E. (1970) 'Efficient capital markets: a review of theory and empirical work', *Journal of Finance*, Vol. 25, No. 2, pp.383–417.

Fama, E.F. (1990) 'Term-structure forecasts of interest rates, inflation and real returns', *Journal of Monetary Economics*, Vol. 25, No. 1, pp.59–76.

Fama, E.F. and French, K.R. (1988) 'Dividend yields and expected stock returns', *Journal of Financial Economics*, Vol. 22, No. 1, pp.3–25.

Fama, E.F. and French, K.R. (1989) 'Business conditions and expected returns on stocks and bonds', *Journal of Financial Economics*, Vol. 25, No. 1, pp.23–49.

Fama, E.F. and French, K.R. (1993) 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics*, Vol. 33, No. 1, pp.3–56.

Fama, E.F. and French, K.R. (2015) 'A five-factor asset pricing model', *Journal of Financial Economics*, No. 116, pp.1–22.

Fama, E.F. and Schwert, G.W. (1977) 'Asset returns and inflation', *Journal of Financial Economics*, Vol. 5, No. 2, pp.115–146.

Ferson, W.E. and Harvey, C.R. (1991) 'The variation of economic risk premiums', *Journal of Political Economy*, Vol. 99, No. 2, pp.385–415.

Fisher, K.L. and Statman, M. (2000) 'Investor sentiment and stock returns', *Financial Analysts Journal*, Vol. 56, No. 2, pp.16–23.

Fisher, K.L. and Statman, M. (2003) 'Consumer confidence and stock returns', *The Journal of Portfolio Management*, Vol. 30, No. 1, pp.115–127.

Gong, X., Zhang, W., Wang, J. and Wang, C. (2022) 'Investor sentiment and stock volatility: new evidence', *International Review of Financial Analysis*, Vol. 80, p.102028, https://doi.org/10.1016/j.irfa.2022.102028.

Granger, C.W.J. (1969) 'Investigating causal relations by econometric models and cross-spectral methods', *Econometrica*, Vol. 37, No. 3, p.424.

Granger, C.W.J. (1988) 'Causality, cointegration, and control', *Journal of Economic Dynamics and Control*, Vol. 12, Nos. 2–3, pp.551–559.

Hirshleifer, D. (2001) 'Investor psychology and asset pricing', *The Journal of Finance*, Vol. 56, No. 4, pp.1533–1597.

Hodrick, R.J. (1992) 'Dividend yields and expected stock returns: alternative procedures for inference and measurement', *The Review of Financial Studies*, Vol. 5, No. 3, pp.357–386.

Hudson, Y. and Green, C.J. (2015) 'Is investor sentiment contagious? International sentiment and UK equity returns', *Journal of Behavioral and Experimental Finance*, Vol. 5, pp.46–59, doi:10.2139/ssrn.2293258.

Jegadeesh, N. and Titman, S. (1993) 'Returns to buying winners and selling losers: implications for stock market efficiency', *The Journal of Finance*, Vol. 48, No. 1, pp.65–91.

Keim, D.B. and Stambaugh, R.F. (1986) 'Predicting returns in the stock and bond markets', *Journal of Financial Economics*, Vol. 17, No. 2, pp.357–390.

Kling, G. and Gao, L. (2008) 'Chinese institutional investors' sentiment', *Journal of International Financial Markets, Institutions, and Money*, Vol. 18, No. 4, pp.374–387.

Lao, J., Nie, H. and Jiang, Y. (2018) 'Revisiting the investor sentiment-stock returns relationship: a multi-scale perspective using wavelets', *Physica A: Statistical Mechanics and its Applications*, Vol. 499, pp.420–427.

Lemmon, M. and Portniaguina, E. (2006) 'Consumer confidence and asset prices: some empirical evidence', *The Review of Financial Studies*, Vol. 19, No. 4, pp.1499–1529.

Li, H., Guo, Y. and Park, S.Y. (2017) 'Asymmetric relationship between investors' sentiment and stock returns: evidence from a quantile non-causality test', *International Review of Finance*, Vol. 17, No. 4, pp.617–626.

Lintner, J. (1965) 'Security prices, risk, and maximal gains from diversification', *The Journal of Finance*, Vol. 20, No. 4, pp.587–615.

Litterman, R.B. (1986) 'Forecasting with Bayesian vector autoregressions – five years of experience', *Journal of Business & Economic Statistics*, Vol. 4, No. 1, pp.25–38.

Lutz, C. (2016) 'The asymmetric effects of investor sentiment', *Macroeconomic Dynamics*, Vol. 20, No. 6, pp.1477–1503.

Lux, T. (2011) 'Sentiment dynamics and stock returns the case of the German stock market', *Empirical Economics*, Vol. 41, No. 3, pp.663–679.

Mbanga, C., Darrat, A. and Park, J. (2019) 'Investor sentiment and aggregate stock returns: the role of investor attention', *Review of Quantitative Finance and Accounting*, pp.1–32.

Neal, R. and Wheatley, S.M. (1998) 'Do measures of investor sentiment predict returns?', *Journal of Financial and Quantitative Analysis*, Vol. 33, No. 4, pp.523–547.

Péguin-Feissolle, A., Strikholm, B. and Teräsvirta, T. (2013) 'Testing the Granger non-causality hypothesis in stationary nonlinear models of unknown functional form', *Communications in Statistics-Simulation and Computation*, Vol. 42, No. 5, pp.1063–1087.

Pesaran, H.H. and Shin, Y. (1998) 'Generalized impulse response analysis in linear multivariate models', *Economics Letters*, Vol. 58, No. 1, pp.17–29.

Qadan, M. and Aharon, D.Y. (2019) 'Can investor sentiment predict the size premium?', *International Review of Financial Analysis*, Vol. 63, pp.10–26.

Qiu, L. and Welch, I. (2006) *Investor Sentiment Measures*, No. 589641, National Bureau of Economic Research.

Sayim, M. and Rahman, H. (2015) 'An examination of US institutional and individual investor sentiment effect on the Turkish stock market', *Global Finance Journal*, Vol. 26, pp.1–17.

Schmeling, M. (2009) 'Investor sentiment and stock return: some international evidence', *Journal of Empirical Finance*, Vol. 16, No. 3, pp.394–408.

Schwarz, G. (1978) 'Estimating the dimension of a model', *Annals of Statistics*, Vol. 6, pp.461–464.

Schwert, G.W. (1990) 'Stock returns and real activity: a century of evidence', *The Journal of Finance*, Vol. 45, No. 4, pp.1237–1257.

Sharpe, S.A. (2002) 'Reexamining stock valuation and inflation: the implications of analysts' earnings forecasts', *Review of Economics and Statistics*, Vol. 84, No. 4, pp.632–648.

Sharpe, W.F. (1964) 'Capital asset prices: a theory of market equilibrium under conditions of risk', *The Journal of Finance*, Vol. 19, No. 3, pp.425–442.

Shi, S., Phillips, P.C. and Hurn, S. (2018) 'Change detection and the causal impact of the yield curve', *Journal of Time Series Analysis*, Vol. 39, No. 6, pp.966–987.

Shleifer, A. and Summers, L.H. (1990) 'The noise trader approach to finance', *Journal of Economic Perspectives*, Vol. 4, No. 2, pp.19–33.

Simon, D.P. and Wiggins, R.A. (2001) 'S&P futures returns and contrary sentiment indicators', *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, Vol. 21, No. 5, pp.447–462.

Sims, C.A. (1980) 'Macroeconomics and reality', *Econometrica*, Vol. 48, pp.1–48.

Solt, M.E. and Statman, M. (1988) 'How useful is the sentiment index?', *Financial Analysts Journal*, Vol. 44, No. 5, pp.45–55.

Spyrou, S. (2012) 'Sentiment changes, stock returns and volatility: evidence from NYSE, AMEX and NASDAQ stocks', *Applied Financial Economics*, Vol. 22, No. 19, pp.1631–1646.

Statman, M., Thorley, S. and Vorkink, K. (2006) 'Investor overconfidence and trading volume', *Review of Financial Studies*, Vol. 19, No. 4, pp.1531–1565.

Swaminathan, B. (1996) 'Time-varying expected small firm returns and closed-end fund discounts', *The Review of Financial Studies*, Vol. 9, No. 3, pp.845–887.

Verma, R. and Soydemir, G. (2006) 'The impact of US individual and institutional investor sentiment on foreign stock markets', *The Journal of Behavioral Finance*, Vol. 7, No. 3, pp.128–144.

Verma, R. and Soydemir, G. (2009) 'The impact of individual and institutional investor sentiment on the market price of risk', *The Quarterly Review of Economics and Finance*, Vol. 49, No. 3, pp.1129–1145

Verma, R. and Verma, P. (2008) 'Are survey forecasts of individual and institutional investor sentiments rational?', *International Review of Financial Analysis*, Vol. 17, No. 5, pp.1139–1155.

Vozlyublennaia, N. (2014) 'Investor attention, index performance, and return predictability', *Journal of Banking & Finance*, Vol. 41, pp.17–35.

Wang, F.A. (2001) 'Overconfidence, investor sentiment, and evolution', *Journal of Financial Intermediation*, Vol. 10, No. 2, pp.138–170.

Wang, W., Su, C. and Duxbury, D. (2022) 'The conditional impact of investor sentiment in global stock markets: a two-channel examination', *Journal of Banking & Finance*, Vol. 138, p.106458.

Wang, Y-H., Keswani, A. and Taylor, S.J. (2006) 'The relationships between sentiment, returns and volatility', *International Journal of Forecasting*, Vol. 22, No. 1, pp.109–123.

Welch, I. and Goyal, A. (2008) 'A comprehensive look at the empirical performance of equity premium prediction', *The Review of Financial Studies*, Vol. 21, No. 4, pp.1455–1508.

Zhou, G. (2018) 'Measuring investor sentiment', *Annual Review of Financial Economics*, Vol. 10, pp.239–259.

Notes

1 It is worth noting that sentiment is found as a major factor not only in equities but also in explaining part of the idiosyncratic volatility (Wang et al., 2022), commodity prices (Qadan and Aharon, 2019), and the looking-forward volatility (Gong et al., 2022).

2 Following Brown and Cliff (2004), Hudson and Green (2015), Baker and Wurgler (2006), and Baker et al. (2012), we use the principal component analysis (PCA) to transform individual variables into different principal components. The first component must explain the most variation and each following component accounts for the highest variance possible.

3 These data are available on the following link: https://sites.google.com/view/agoyal145/.

Appendix

Acronyms, variable definitions and data sources

Approaches	*Acronyms*
PCA	Principal component analysis
VAR	Vector autoregressive
IRF	Impulse response function
FEVD	Forecast error variance decomposition