CAUSALITIES BETWEEN SENTIMENT INDICATORS AND STOCK MARKET RETURNS UNDER DIFFERENT MARKET SCENARIOS

Her-Jiun Sheu, National Chi Nan University
Yang-Cheng Lu, Ming Chuan University
Yu-Chen Wei, National Chiao Tung University and Ming Chuan University

ABSTRACT

This paper investigates the causal relationships between sentiment and returns under different market scenarios. In contrast to previous studies that subjectively identify the bullish and bearish markets, we apply a threshold model to detect the extreme level of investors' sentiment econometrically. The empirical results show that most of the sentiment measures exhibit a feedback relationship with returns while ignoring different market states. However, sentiment could be a leading indicator if the higher or lower levels of sentiments were to be distinguished. Among them, the bullish/bearish indicator of ARMS, which is named after its creator, Richard Arms (1989), is a leading indicator if the market is more bearish (in the higher regime). Otherwise, the leading effect of the derivatives market sentiment indicators (the put-call trading volume and option volatility index) is discovered if the market is more bullish (in the lower regime). Our empirical findings further confirm the noise trader explanation that the causal direction would run from investors' sentiment to market behavior.

JEL: C32; G10

KEYWORDS: Investor sentiment, Stock market returns, Granger causality, threshold model

INTRODUCTION

The behavioral models of securities markets regard investors as being of two types: rational arbitrageurs who are sentiment-free and irrational traders who are prone to exogenous sentiment. In considering that investors may either overreact or under-react to extreme levels of sentiment indicators, we examine whether the sentiment indicators are classified according to multiple regimes by using the multivariate threshold model. Since previous studies have usually defined the extreme level subjectively, this paper analyzes the different states more objectively. The causality relationships between stock market returns and sentiment indicators are more significant when the different states are distinguished. The empirical results lead us to conclude that sentiment in both the stock and derivative markets gives rise to distinct lead-lag relationships with returns.

While there is an abundant literature on implied volatility indexes in developed markets, little research has been conducted in the context of emerging markets. Index options involving the Taiwan Stock Exchange Capitalization Weighted Stock Index (*TAIEX* index options, abbreviated as TXO) were first traded on December 24, 2001. The *TAIEX* covers all of the listed stocks on the Taiwan Stock Exchange (TWSE) excluding preferred stocks, full-delivery stocks and newly-listed stocks, which are listed for less than one calendar month. The statistical data published in the annual report of the Futures Industry Association (FIA) in 2003 show that the trading volume of the TXO grew significantly faster in 2003 than in 2002. The FIA is the only association that is representative of all organizations having an interest in the futures market. The FIA has more than 180 corporate members, and reaches thousands of industry participants. Further information may be found on the website http://www.futuresindustry.org/. The FIA annual report in 2006 further indicates that *TAIEX* options is ranked sixteenth in the world in terms of trading volume meaning that it is among the top 20 derivatives contracts. In addition, the trading volume of *TAIEX* options still exhibits a high growth rate. These motivate us to investigate the possible causalities between market sentiment and stock returns in Taiwan.

The analysis is conducted on a daily basis and the sentiment indicators used in this study include the TXO put-call trading volume ratio (TPCV), the TXO put-call open interest ratio (TPCO), the option market volatility index (TVIX) and the ARMS index. Our major focus of concern is on whether the causal relationship between sentiment and returns differs when investors' sentiment is at an extreme level identified optimistically by the threshold model. Our major findings suggest that there is nonlinearity in the sentiment indicators. The causality between sentiment and returns leads to different results when the sentiment index is at an extremely high or low level, or else reflects a typical regime. In the ordinary market scenario, there is low negative correlation as well as bi-directional causality. When the market overacts, the sentiment indicators Granger cause the returns. Among them, the ARMS index Granger causes the stock returns in the median and higher regimes, while the sentiment indicators in the derivatives market Granger cause the returns in the median and lower regimes. Our empirical findings further confirm the noise trader explanation that the causal direction runs from sentiment to market behavior.

The remainder of this paper is organized as follows. Section 2 briefly discusses the relevant literature. Section 3 outlines the measurement and summary statistics of the data. Section 4 summarizes the empirical design of the paper. Section 5 reports the empirical results, confirming that the nonlinear model better captures the dynamic causal relationship between the sentiment index and the stock market index. Section 6 concludes the paper.

LITERATURE REVIEW

Early papers (Friedman, 1953; Fama, 1965) argued that noise traders are unimportant in the financial price formation process because trades made by rational arbitrageurs drive prices close to their fundamental values. However, the market anomalies, for example, the under-reaction and overreaction of stock prices, challenge the efficient markets theory. De Long, Shleifer, Summers and Waldmann (DSSW (1990) hereafter) modeled the influence of noise trading on equilibrium prices and motivated empirical attempts to substantiate the proposition that 'noise traders' risks influence price formation'. Lee, Jiang and Indro (2002) tested the impact of noise trader risk on the formation of conditional volatility and expected returns. Their empirical results show that sentiment is a systematic risk that is priced. Baker and Wurgler (2006) also indicated that investor sentiment affects the cross-section of stock returns. They found that when beginning-of-period proxies for sentiment are low, subsequent returns are relatively high for small stocks, young stocks, high volatility stocks, unprofitable stocks, non-dividend-paying stocks, extreme growth stocks and distressed stocks. If sentiment indicators are risk factors in the time series of returns, they will have the ability to predict the future returns on portfolios, even after appropriately adjusting for other risk factors. These findings all support the need for research on the relationship between stock market returns and indicators of investor sentiment.

The causal relationships between sentiment indicators and stock market returns are mixed in previous studies. Clarke and Statman (1998) found that the sentiment of newsletter writers, whether bullish or bearish, does not forecast future returns, but that past returns and the volatility of those returns do affect sentiment. Causality would thus run from sentiment to market behavior if the noise trader explanation were to be accepted. However, Brown and Cliff (2004) and Solt and Statman (1988) documented that returns cause sentiment rather than the other way round. Brown and Cliff (2004) used a large number of sentiment indicators to investigate the relationship between sentiment and equity returns and found that returns cause sentiment rather than the opposite being the case. Brown (1999) supported the DSSW theory that irrational investors acting in concert and giving a noisy signal can influence asset prices and generate additional volatility. His tests used volatility instead of returns and his results indicated that deviations from the average level of sentiment are associated with increases in fund volatility only during trading hours. Wang, Keswani and Taylor (2006) further tested the relationships between sentiment, returns and volatility. They also found strong and consistent evidence that sentiment measures, both in levels and first differences, are Granger-caused by returns. Banerjee, Doran and Peterson (2007) found that future returns are significantly related to both volatility index (VIX) levels and innovations for most portfolios, where the VIX is treated as a proxy variable for sentiment. While the causality test results presented above do not provide evidence of a consistent relationship between noise traders' sentiments and subsequent price movements, it might be possible that a relationship exists, but only in some special market scenarios.

The frame dependence theory, proposed by Shefrin (2000) in behavioral finance, argues that investors' decisions are sensitive to different market scenarios. This motivates us to investigate whether there are dynamic causal relationships between sentiments and returns. Besides considering both positive and negative market scenarios, we infer that investors may exhibit dissimilar behaviors depending on the level of sentiment, and therefore different dynamic relationships may exist between stock market returns and sentiment indicators. Giot (2005) found that for very high (low) levels of the VIX, future returns are always positive (negative). His findings suggested that extremely high levels of the VIX might signal attractive buying opportunities. Banerjee et al. (2007) examined the relationship between returns and the VIX, the proxy variable for sentiment, for different levels of market performance and relatively high or low levels of volatility. Banerjee et al. (2007) defined those returns above and those below the sample median as constituting a 'bull market' and a 'bear market', respectively. Volatilities above the median level of the VIX are said to be in a 'high volatility' period and those below the median in a 'low volatility period'. They provided two analyses, one of the 'bull and bear market' and the other of 'high and low volatility'. Their findings suggested that the market states based on directional movements (positive and negative returns) or volatility levels (above or below the average) do not make a difference. On the contrary, we believe that the results will be misunderstood if the separation of the different market states is defined subjectively.

To sum up, we apply the threshold model to examine the threshold effect of the sentiment indicators. Higher and lower regimes of sentiment indicators will be detected objectively. Therefore, the causality relationship needs to be tested for different market scenarios.

DATA

The daily sentiment indicators used consist of the TXO put-call trading volume ratio (TPCV), the TXO put-call open interest ratio (TPCO), the TXO volatility index (TVIX) and the TAIEX ARMS index. To do this, we use data that are fully quoted on the Taiwan Futures Exchange (TAIFEX) and the Taiwan Stock Exchange (TSE). The study period extends from 2003 to 2006, encompassing 993 trading days.

Investor Fear Gauge – Option Volatility Index

Options market-based implied volatility can reflect the expectations with respect to price changes in the future, and it can be treated as an indicator of sentiment. Olsen (1998) indicated that the volatility index has been viewed as a 'sentiment indicator' in the recent behavioral finance literature and can be regarded as a market indicator of rises and falls in the underlying index. Whaley (2000) and research conducted by the Chicago Board Options Exchange (CBOE) have indicated that the greater the fear, the higher the VIX level is. Therefore, the volatility index is commonly referred to as the 'investor fear gauge'. Baker and Wurgler (2007) also treated option-implied volatility as one of the sentiment measures in investigating the investor sentiment approach. Therefore, we adopt the Taiwan stock market volatility index (TVIX) as one of the sentiment proxy variables in the Taiwan options market.

In 1993, the CBOE introduced the Volatility Index (VIX) based on the S&P 100 index options that can be defined as the magnitude of price variation for the following 30 days. The new version of the Volatility Index published in 2003 is based on S&P 500 index options prices. In March 2004, the CBOE futures exchange (CFE) introduced volatility futures, and volatility options were launched in February 2006. The underlying index is just the VIX published in 2003. The volatility index has become a tradable derivative. Since the CBOE published the new volatility index in 2003, we construct the TVIX by adjusting the last revision of the CBOE volatility index. The construction of the CBOE's new volatility index incorporates information from the skewness of volatility by using a wider range of strike prices including the out-of-the-money call and put option contracts rather than just the at-the-money series. The new volatility index is more precise and robust than the original version. However, the fundamental features of the volatility index between the old and new versions remain the same. For details of the index's construction, the interested reader may refer to the white book published by the CBOE in 2003, http://www.cboe.com/micro/vix/vixwhite.pdf. In the construction of the Taiwan stock market VIX, the interest rate has been adjusted accordingly. The risk-free rate is calculated from the monthly average one-year deposit rates at the Bank of Taiwan, Taiwan Cooperative Bank, First Bank, Hua Nan Bank and Chang Hwa Bank. The CBOE's volatility index (VIX) uses put and call options in the two nearest-term expiration months in order to bracket a 30-day calendar period. With 8 days left to expiration, CBOE's VIX 'rolls' to the second and third contract months in order to minimize pricing anomalies that might occur close to expiration. However, the nearest-term expiration contract usually has high trading volume and the next nearest-term contract usually has low trading volume in the Taiwan options market even if the nearest-term contract is traded on the last trading day. In considering the market structure of liquidity and trading volume for the second and third contract months, we have revised the rollover rule from 8 days to 1 day prior to expiration in constructing the volatility index in Taiwan.

Put-Call Trading Volume and Open Interest Ratios

The put-call trading volume ratio equals the total trading volume of puts divided by the total trading volume of calls (TPCV). Like the TVIX, market participants view the TPCV as a fear indicator, with higher levels reflecting bearish sentiment. When market participants are bearish, they buy put options to hedge their equity positions or to speculate bearishly. By contrast, a low level of TPCV is associated with a lower demand for puts, which reflects bullish sentiment.

The put-call open interest ratios can be calculated using the open interest of options instead of trading volume (TPCO). When the total option interest increases, most of it comes from higher investor demand for TXO puts. Thus the TPCO tends to be higher on days when the total open interest is high.

ARMS Index

The ARMS index is named after its creator, Richard Arms (1989), and is an indicator of bullish or bearish sentiment. The ARMS index on day t is equal to the number of advancing issues scaled by the trading volume (shares) of advancing issues divided by the number of declining issues scaled by the trading volume (shares) of declining issues. It is measured as:

$$ARMS_t = \frac{\#Adv_t / AdvVol_t}{\#Dec_t / DecVol_t} = \frac{DecVol_t / \#Dec_t}{AdvVol_t / \#Adv_t} \qquad (1)$$

where $\#Adv_t$, $\#Dec_t$, $AdvVol_t$, and $DecVol_t$, respectively, denote the number of advancing issues, the number of declining issues, the trading volume of advancing issues, and the trading volume of declining issues.

ARMS can be interpreted as the ratio of the number of advances to declines standardized by their respective volumes. If the index is greater than one, more trading is taking place in declining issues, while if it is less than one, the average volume of advancing stocks outpaces the average volume of declining stocks. Its creator, Richard Arms, argued that if the average volume of declining stocks far outweighs the average volume of rising stocks, then the market is oversold and this should be treated as a bullish sign. Likewise, he argued that if the average volume of rising stocks far outweighs the average volume of falling stocks, then the market is overbought and this should be treated as a bearish sign.

Summary Statistics of the Data

Table 1 contains summary statistics of all the variables discussed in the study. The returns display excess kurtosis, negative skewness and almost no serial correlation. The contemporaneous relationships among many measures of investor sentiment and market returns depicted in Table 2 are shown to be strong. Figure 1 shows the daily evolution of the TAIEX and returns from 2003 to 2006. Figure 2 is the daily evolution of the sentiment indices from 2003 to 2006.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Autocorrelation			
					ρ_1	ρ_2	ρ_3	ρ_4
TAIEX	6,030.7580	732.2869	-0.4379	3.2624	0.9850	0.9700	0.9550	0.9410
R	0.0006	0.0120	-0.3855	6.3835	0.0390	-0.0110	0.0250	-0.0420
TVIX	20.7318	5.4899	0.9942	3.9072	0.9710	0.9530	0.9390	0.9230
TPCV	0.7835	0.1669	0.8043	4.3116	0.4640	0.3470	0.2820	0.2280
TPCO	0.9307	0.2597	1.1246	5.2412	0.9410	0.8720	0.8010	0.7370
ARMS	0.7168	0.3820	9.0595	175.3529	0.1190	0.0690	0.0010	-0.0120
ΔTVIX	-0.0029	1.2995	1.2845	16.4393	-0.2030	-0.0490	0.0360	-0.0510
ΔTPCV	0.0004	0.1729	-0.0767	4.3869	-0.3920	-0.0550	-0.0050	-0.0220
ΔTPCO	0.0004	0.0885	-3.0162	35.3451	0.0870	0.0250	-0.0670	-0.0420
ΔARMS	-0.0010	0.5087	-0.9781	91.5070	-0.4700	0.0110	-0.0320	0.0110

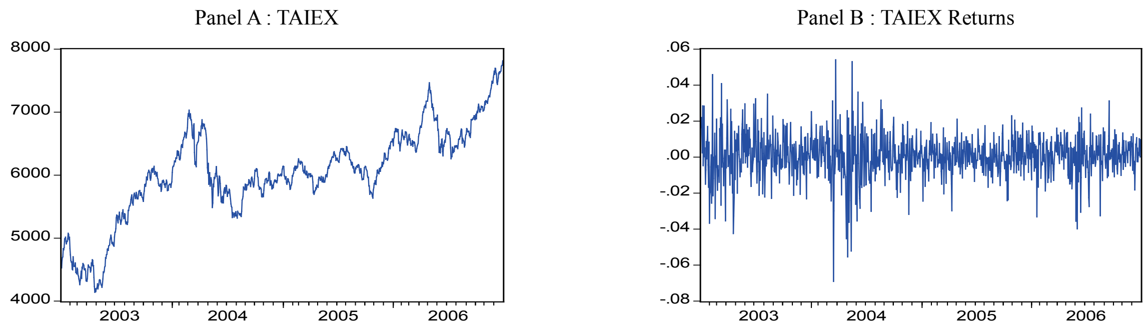
This table presents the summary statistics for the return on the Taiwan stock exchange capitalization weighted stock index (TAIEX) and various sentiment measures, namely, the Taiwan volatility index (TVIX), the put-call volume ratio (TPCV), the put-call open interest ratio (TPCO) and the ARMS ratio. The period covers 1/2/2003 to 12/29/2006.

Table 2: Contemporaneous Correlations

	R	TVIX	TPCV	TPCO	ARMS	ΔTVIX	ΔTPCV	ΔTPCO	ΔARMS
TAIEX	0.0442	-0.4035***	0.0559*	0.2024***	-0.0282	0.0016	0.0007	-0.006	0.0035
R		-0.0885***	-0.2773***	0.1509***	-0.3542***	-0.2537***	-0.2622***	0.3703***	-0.2635***
TVIX			-0.0759**	-0.1974***	0.1028***	0.1197***	-0.0006	-0.0406	-0.0117
TPCV				0.0345	0.1476***	0.0667**	0.5179***	-0.146***	0.021
TPCO					-0.162***	0.0395	-0.0036	0.1704***	0.0222
ARMS						0.0316	0.0404	-0.1609***	0.6638***
ΔTVIX							0.0674**	-0.0431	-0.0285
ΔTPCV								-0.1509***	0.0486
ΔTPCO									-0.0596*

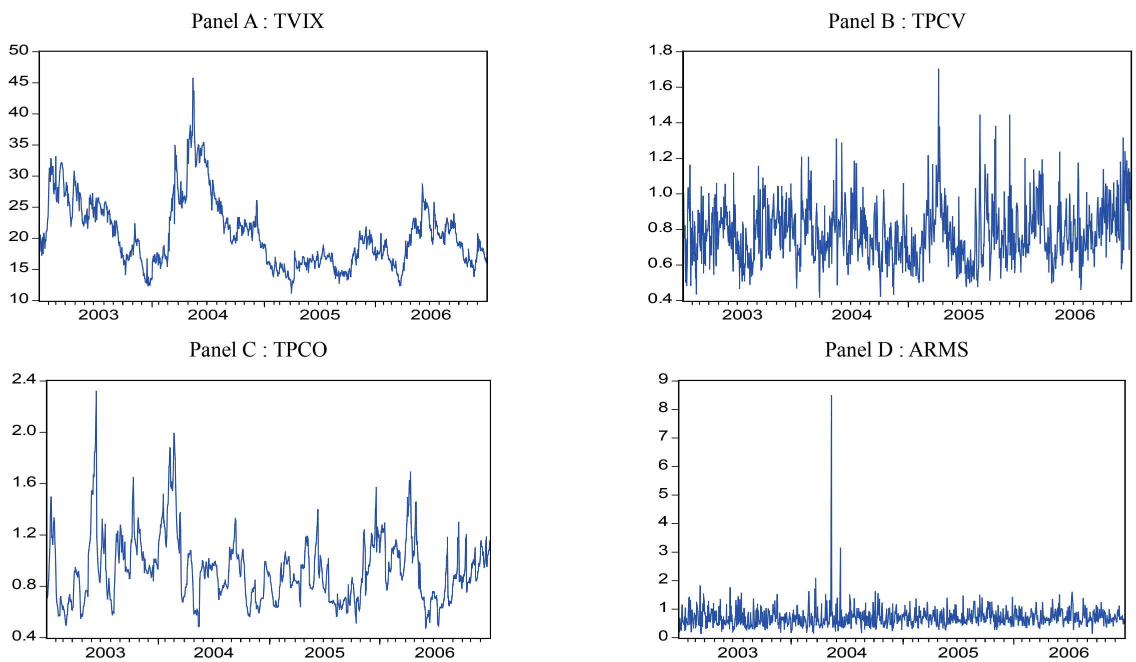
*The pairwise correlations are for selected variables used in the analysis. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.*

Figure 1: Daily Evolution of the TAIEX and TAIEX Returns from 2003 to 2006



This figure shows the daily evolution of the TAIEX and TAIEX returns from 2003 to 2006. TAIEX represents the Taiwan stock exchange capitalization weighted stock index. TAIEX returns are calculated as the logarithmic difference in the daily TAIEX, i.e., $R_t = lnS_t - lnS_{t-1}$, *where* R_t *represents the TAIEX market returns on day t, and* S_t *and* S_{t-1} *are the daily closing prices of the TAIEX on day t and t-1, respectively.*

Figure 2: Daily Evolution of the Sentiment Indices from 2003 to 2006



This figure shows the daily investor sentiments during 2003 to 2006. The Taiwan volatility index (TVIX) is calculated using daily data quoted on the Taiwan Futures Exchange (TAIFEX) and the Taiwan Stock Exchange (TWSE). The method used to construct the TVIX refers to the essence of the last revision of the volatility index of the CBOE and the interest rate, and the rollover rule is revised accordingly. The ARMS, put-call trading volume ratio (TPCV) and put-call open interest ratio (TPCO) are calculated using daily data quoted on the TWSE and TAIFEX.

RESEARCH DESIGN

Causality Tests

We test for Granger causality between sentiment and returns by estimating bivariate VAR models (Granger, 1969, 1988). The Granger causality tests examine whether the lags of one variable enter the equation to determine the dependent variables, assuming that the two series (sentiment index and stock market return) are covariance stationary and the error items are i.i.d. white noise errors.

We estimate the models using both levels and changes in sentiment measures since it is not easy to determine which specification should reveal the primary effects of sentiment. For example, suppose investor sentiment decreases from very bullish to bullish. One might anticipate a positive return due to the still bullish sentiment, but on the other hand, since sentiment has decreased, it is also possible for someone to expect a reduction in the return. The general model we use here can be expressed as follows:

$$
\begin{aligned}
R_t &= C_1 + \sum_{p=1}^{L} \alpha_{1p} R_{t-p} + \sum_{p=1}^{L} \beta_{1p} Senti_{t-p} + \varepsilon_{1t} \ , \\
Senti_t &= C_2 + \sum_{p=1}^{L} \rho_{2p} R_{t-p} + \sum_{p=1}^{L} \beta_{2p} Senti_{t-p} + \varepsilon_{2t} \ ,
\end{aligned}
\tag{2}
$$

where R_t denotes the stock market returns and $Senti_t$ represents the sentiment levels or the sentiment changes. The sentiment indices include TVIX, TPCV, TPCO and ARMS. In the bivariate Granger causality tests, the returns do not Granger cause the sentiment measures if the lagged values R_{t-p} do not enter the $Senti_t$ equation. Similarly, the returns do not Granger cause the sentiment measures if all the ρ_{2p} equal zero as a group based on a standard F-test. Meanwhile, the sentiment measures do not Granger cause the returns if all the β_{1p} equal zero.

Causality Relationship under Different Market Scenarios

We examine the causality relationship under the positive and negative market return scenario. The model may alternatively be written as:

$$
\begin{cases}
R_t = C_{11} + \sum_{p=1}^{L} \rho_{11p} R_{t-p} + \sum_{p=1}^{L} \beta_{11p} Senti_{t-p} + \varepsilon_{11t} \\
Senti_t = C_{12} + \sum_{p=1}^{L} \rho_{12p} R_{t-p} + \sum_{p=1}^{L} \beta_{12p} Senti_{t-p} + \varepsilon_{12t}
\end{cases}
, \text{ if } R_t \geq 0
\tag{3-1}
$$

$$
\begin{cases}
R_t = C_{21} + \sum_{p=1}^{L} \rho_{21p} R_{t-p} + \sum_{p=1}^{L} \beta_{21p} Senti_{t-p} + \varepsilon_{21t} \\
Senti_t = C_{22} + \sum_{p=1}^{L} \rho_{22p} R_{t-p} + \sum_{p=1}^{L} \beta_{22p} Senti_{t-p} + \varepsilon_{22t}
\end{cases}
, \text{ if } R_t < 0
\tag{3-2}
$$

where $R_t \geqq 0$ represents the positive return scenario and $R_t < 0$ is the negative return scenario. The threshold variable of the return is also substituted as a sentiment variable. There are three scenarios examined in the following study, the extremely high sentiment (top 20%), the extremely low sentiment (bottom 20%) and the typical sentiment group (median 60%).

The Oversold and Overbought Scenarios Identified by the Threshold Model

A two-regime version of the threshold autoregressive (TAR) model developed by Tong (1983) is expressed as follows:

$$y_t = I_t\left[\beta_{10} + \sum_{i=1}^{p}\beta_{1i}y_{t\text{-}i}\right] + \left(1\text{-}I_t\right)\left[\beta_{20} + \sum_{i=1}^{p}\beta_{2i}y_{t\text{-}i}\right] + \varepsilon_i,\ I_t = \begin{cases}1, & if\ y_{t\text{-}1} \geq \gamma \\ 0, & if\ y_{t\text{-}1} < \gamma\end{cases} \quad (4)$$

where y_t is the series of interest, β_{1i} and β_{2i} are the coefficients to be estimated, i=1…p, p is the order of the TAR model, γ is the value of the threshold, and I_t is the Heaviside indicator function. One problem with Tong (1983)'s model is that the threshold may not be known. When γ is unknown, Chan (1993) shows how to obtain a super-consistent estimate of the threshold parameter. The general form of Chan's model can be described as:

$$y_t = \begin{cases}\beta_{10} + \beta_{11}y_{t\text{-}1} + \cdots + \beta_{1p}y_{t\text{-}p} + c_1e_t,\ if\ y_{t\text{-}d} < \gamma \\ \beta_{20} + \beta_{21}y_{t\text{-}1} + \cdots + \beta_{2p}y_{t\text{-}p} + c_2e_t,\ if\ y_{t\text{-}d} \geq \gamma\end{cases} \quad (5)$$

For a TAR model, the procedure is to order the observations from the smallest to the largest such that $y_1 < y_2 < y_3 \ldots < y_T$. For each value of y_i, let $\gamma = y_i$, and let the Heaviside indicator be set according to this potential threshold in order to estimate a TAR model. The regression equation with the smallest residual sum of squares contains a consistent estimate of the threshold. Chan (1993) indicates that each data point within the band has the potential to be the threshold. However, it may be inefficient to examine the threshold effect of each value. Therefore, we adopt the grid search method whereby n sample points within the estimation period are selected to test the threshold effect and we set n equal to 100. In order to classify the oversold and overbought regimes, we apply the threshold test twice in the above and below average levels of each sentiment indicator. The highest and lowest 10 percent of the values are excluded from the search to ensure an adequate number of observations on each side of the threshold.

EMPIRICAL RESULTS AND ANALYSIS

Granger-Causality Tests under Different Market Scenarios

The lag lengths of the TAIEX returns and sentiment indices are determined before the causality test is performed. The numbers of lagged terms in the VAR models are decided parsimoniously by the Akaike information criterion (AIC) and the Schwarz criterion (SC). Table 3 presents the general causality tests. The results show that there is a feedback relationship between returns and sentiment, in both levels and first differences, and including TVIX and ARMS. As for the other two derivatives market sentiment indicators, TPCV and TPCO, these have no leading effect.

Table 3: General Causality Tests between Returns and Sentiment

Sentiment	Hypothesis			
	H_{01}	H_{02}	H_{03}	H_{04}
TVIX	2.7533 (0.0642)*	3.9627 (0.0193)**	4.3919 (0.0364)**	6.4175 (0.0115)**
TPCV	0.0196 (0.9806)	0.9918 (0.3713)	0.1901 (0.9031)	6.4853 (0.0002)***
TPCO	0.4045 (0.5249)	51.7436 (<0.0001)***	3.0538 (0.0809)*	30.2449 (<0.0001)***
ARMS	4.8131 (0.0083)***	19.369 (<0.0001)***	2.5839 (0.0173)**	9.0376 (<0.0001)***

*The numbers of lagged terms in the VAR models are decided parsimoniously by the Akaike information criterion (AIC) and the Schwarz criterion (SC). H_{01}: Granger-noncausality from sentiment to returns, i.e., sentiment does not cause returns. H_{02}: Granger-noncausality from returns to sentiment, i.e., returns do not cause sentiment. H_{03}: Granger-noncausality from changes in sentiment to returns, i.e., changes in sentiment do not cause returns. H_{04}: Granger-noncausality from returns to changes in sentiment, i.e., returns do not cause changes in sentiment. Values in the table and the parentheses are F test statistics and p-values, respectively. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.*

The positive and negative market return scenarios indicate whether the market returns are greater than zero or not. The results of these Granger-causality tests are presented in Table 4. The TVIX Granger causes returns when the return is greater than zero. However, the sentiment indicators are Granger-caused by returns while the return is smaller than zero. In short, TVIX could be a leading indicator while the market returns are positive.

Table 4: Causality Tests between Returns and Sentiment – Considering the Positive and Negative Market Return Scenarios

Sentiment	Hypothesis			
	H_{01}	H_{02}	H_{03}	H_{04}
Panel A :Positive Return				
TVIX	33.5609 (<0.0001)***	0.4766 (0.6212)	3.8761 (0.0495)**	0.0607 (0.8056)
TPCV	0.4318 (0.6496)	2.1598 (0.1164)	1.0806 (0.3568)	5.9478 (0.0005)***
TPCO	4.9796 (0.0261)**	15.1619 (0.0001)***	0.782 (0.377)	7.5925 (0.0061)***
ARMS	13.4788 (<0.0001)***	9.4277 (0.0001)***	4.7919 (0.0001)***	5.8443 (<0.0001)***
Panel B Negative Return				
TVIX	23.7999 (<0.0001)***	4.9421 (0.0075)***	0.0029 (0.9569)	9.9774 (0.0017)***
TPCV	1.2442 (0.2891)	0.2446 (0.7831)	0.8122 (0.4876)	2.5514 (0.0551)*
TPCO	2.2443 (0.1348)	61.2698 (<0.0001)***	0.118 (0.7314)	42.7464 (<0.0001)***
ARMS	0.5613 (0.5708)	11.2781 (<0.0001)***	1.3231 (0.2451)	4.7347 (0.0001)***

*This table presents the causality tests between returns and sentiment considering the positive and negative market return scenarios. The numbers of lagged terms in the VAR models are decided parsimoniously by the Akaike information criterion (AIC) and the Schwarz criterion (SC). H_{01}: Granger-noncausality from sentiment to returns, i.e., sentiment does not cause returns. H_{02}: Granger-noncausality from returns to sentiment, i.e., returns do not cause sentiment. H_{03}: Granger-noncausality from changes in sentiment to returns, i.e., changes in sentiment do not cause returns. H_{04}: Granger-noncausality from returns to changes in sentiment, i.e., returns do not cause changes in sentiment. Values in the table and the parentheses are F test statistics and p-values, respectively. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.*

The other situations with which we are concerned in this study are whether the sentiment is grouped in the top 20% or the bottom 20%. Most of the results, which are presented in Table 5, show that there is no distinct causal relationship between sentiment and returns although the TVIX and TPCV Granger cause returns while in the bottom 20%. Considering that the critical values of the overreaction scenarios are determined subjectively, the feedback relationship may be mixed.

Table 5 Causality Tests between Returns and Sentiment – Sentiments Grouped at the Top, Median and Bottom Levels

Sentiment	Hypothesis			
	H_{01}	H_{02}	H_{03}	H_{04}
Panel A: Top 20% of the Sentiment				
TVIX	3.8314 (0.0233)**	8.6299 (0.0003)***	0.0076 (0.9308)	56.8343 (<0.0001)***
TPCV	1.943 (0.1461)	0.8829 (0.4152)	1.4031 (0.2432)	1.9424 (0.1242)
TPCO	0.4743 (0.4918)	2.7216 (0.1006)	0.0001 (0.9938)	0.2948 (0.5877)
ARMS	1.4385 (0.2398)	3.3526 (0.037)**	1.1535 (0.3333)	2.6497 (0.0173)**
Panel B: Median of the Sentiment				
TVIX	2.686 (0.069)*	8.3332 (0.0003)***	0.0053 (0.9417)	4.2173 (0.0405)**
TPCV	1.4351 (0.2389)	1.0379 (0.3548)	1.5621 (0.1975)	1.7258 (0.1605)
TPCO	28.2164 (<0.0001)***	25.3552 (<0.0001)***	0.701 (0.4028)	27.3828 (<0.0001)***
ARMS	13.3482 (<0.0001)***	3.2147 (0.0409)**	7.4548 (<0.0001)***	0.6305 (0.7059)
Panel C: Bottom 20% of the sentiment				
TVIX	0.3011 (0.7404)	3.3127 (0.0385)**	3.6214 (0.0585)*	0.7986 (0.3726)
TPCV	10.5245 (<0.0001)***	1.7001 (0.1854)	4.4979 (0.0045)***	1.8991 (0.1312)
TPCO	20.0555 (<0.0001)***	16.3919 (0.0001)***	0.7241 (0.3959)	0.7776 (0.379)
ARMS	0.3037 (0.7384)	1.7525 (0.1761)	2.1734 (0.0474)**	5.4883 (<0.0001)***

*This table presents causality tests between returns and sentiment considering the sentiments grouped at the top, median and bottom levels. The numbers of lagged terms in the VAR models are decided parsimoniously by the Akaike information criterion (AIC) and the Schwarz criterion (SC). H_{01}: Granger-noncausality from sentiment to returns, i.e., sentiment does not cause returns. H_{02}: Granger-noncausality from returns to sentiment, i.e., returns do not cause sentiment. H_{03}: Granger-noncausality from changes in sentiment to returns, i.e., changes in sentiment do not cause returns. H_{04}: Granger-noncausality from returns to changes in sentiment, i.e., returns do not cause changes in sentiment. Values in the table and the parentheses are F test statistics and p-values, respectively. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.*

Finally, there is the causality test between the returns and sentiment indicators in the extreme levels of investor sentiment that are determined by the threshold model. The threshold tests for each sentiment indicator are presented in Table 6 and the percentages for each regime classified by threshold model are shown in Table 7. The threshold tests show that the higher regime of TVIX and the lower regime of ARMS are not significant. Besides, the other sentiment indicators give rise to significant critical values of the higher and lower regimes that can represent the oversold and overbought situations.

Table 6 Threshold Test

	Upper regime			Lower regime		
Sentiment	Threshold Value	F test statistic	p-value	Threshold Value	F test statistic	p-value
TVIX	22.3673	0.3000	(0.5459)	17.9989	3.4807	(0.0293)**
TPCV	0.9612	8.3180	(0.0003)***	0.7377	6.3779	(0.0018)***
TPCO	1.1807	7.0289	(0.0009)***	0.7633	10.7162	(<0.0001)***
ARMS	1.0648	5.0219	(0.0038)***	0.5045	1.7117	(0.12)
ΔTVIX	0.9236	8.5034	(0.0002)***	-1.2803	10.3877	(<0.0001)***
ΔTPCV	0.1876	11.7296	(<0.0001)***	-0.1243	4.8070	(0.0084)***
ΔTPCO	0.0345	4.1566	(0.0159)**	-0.0223	7.7502	(0.0005)***
ΔARMS	0.3477	103.4980	(<0.0001)***	-0.0896	87.9864	(<0.0001)***

*This table presents the threshold tests. The upper regime is the regime above the average level of the sentiment indicators. The lower regime is the regime below the average level of the sentiment indicators. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.*

Table 7 Percentage of Each Regime Classified by Threshold Model

	Higher Regime	Typical Regime	Lower Regime
TVIX			39%
TPCV	14%	42%	44%
TPCO	14%	57%	29%
ARMS	10%		
ΔTVIX	15%	74%	10%
ΔTPCV	13%	68%	19%
ΔTPCO	24%	50%	26%
ΔARMS	15%	46%	39%

This table presents the percentages of different regimes classified by the threshold model. The higher regime is the regime above the higher threshold which is above the average level of the sentiment indicators. The lower regime is the regime below the lower threshold which is below the average level of the sentiment indicators. The typical regime is the regime between the higher and lower thresholds of the sentiment indicators. The blank of the higher regime and typical regime of TVIX indicates that the threshold test is not significant in the upper regime of the TVIX level. The blank of the lower regime and typical regime of ARMS indicates that the threshold test is not significant in the lower regime of the ARMS level.

The results of the causality relationship in the oversold and overbought situations are shown in Table 8. We can find that the market sentiment indicator, ARMS, leads returns while in the upper regime. Both the equity market and derivatives market sentiment indicators, ARMS and TPCV, Granger cause returns in the median regime. In the lower regime, only the sentiment indicators in the derivatives market, TVIX and TPCV, Granger cause returns. From these findings, we can conclude that the equity or derivatives markets sentiment indicators perform differently in terms of the lead-lag relationship between returns while the sentiments are in the higher, median or lower regimes. Our study suggests that investors can adjust their portfolios by analyzing the sentiment indicators for different scenarios.

CONCLUSIONS

In this paper, we have examined the causal relationship between investors' sentiment and stock market returns. The difference between this paper and the previous literature is that we identify the extreme level of sentiment econometrically by using the threshold model. Our analysis is conducted in three steps by using equity market data. We first construct the sentiment indicators in the equity and derivatives markets including the ARMS index, option volatility index, put-call trading volume ratio and put-call open interest ratio. We then examine the threshold of the sentiment indicators to test whether the sentiment could be classified into oversold, overbought and ordinary regimes. Finally, we investigate the relationships and causal directions for the different market scenarios.

Table 8 Causality Tests between Returns and Sentiment - Application of the Multivariate Threshold Model

Sentiment	Hypothesis			
	H_{01}	H_{02}	H_{03}	H_{04}
Panel A: Upper regime (above the higher threshold)				
TVIX			0.2056 (0.6509)	43.7114 (<0.0001)***
TPCV	0.7388 (0.4796)	1.1357 (0.3243)	0.5386 (0.6567)	0.7615 (0.5178)
TPCO	0.0732 (0.7871)	0.5083 (0.4771)	0.0589 (0.8084)	0.0922 (0.7616)
ARMS	3.4356 (0.0364)**	0.8965 (0.4115)	1.5796 (0.1574)	3.0123 (0.0085)***
Panel B : Typical regime (between the two thresholds)				
TVIX			0.065 (0.7988)	12.5205 (0.0004)***
TPCV	5.7821 (0.0033)***	0.1239 (0.8835)	2.3032 (0.0759)*	3.1818 (0.0235)**
TPCO	29.5417 (<0.0001)***	39.5976 (<0.0001)***	1.4719 (0.2256)	16.3007 (0.0001)***
ARMS			5.2997 (<0.0001)***	0.5891 (0.7391)
Panel C : Lower regime (below the lower threshold)				
TVIX	4.4883 (0.0118)**	5.8011 (0.0033)***	3.8007 (0.0541)*	0.5805 (0.4479)
TPCV	7.0184 (0.001)***	1.3517 (0.2599)	4.4569 (0.0048)***	1.6189 (0.1865)
TPCO	10.7606 (0.0012)***	18.2885 (<0.0001)***	0.8613 (0.3542)	0.4873 (0.4858)
ARMS			3.2284 (0.0042)***	4.4251 (0.0002)***

*This table presents the causality tests between returns and sentiment by applying the multivariate threshold model. The numbers of lagged terms in the VAR models are decided parsimoniously by the Akaike information criterion (AIC) and the Schwarz criterion (SC). H_{01}: Granger-noncausality from sentiment to returns, i.e., sentiment does not cause returns. H_{02}: Granger-noncausality from returns to sentiment, i.e., returns do not cause sentiment. H_{03}: Granger-noncausality from changes in sentiment to returns, i.e., changes in sentiment do not cause returns. H_{04}: Granger-noncausality from returns to changes in sentiment, i.e., returns do not cause changes in sentiment. The blank spaces for the causality tests in the higher regime of the TVIX, the typical regime of TVIX and ARMS, and the lower regime of ARMS indicate that the threshold test is not significant in that scenario. Therefore, the causality tests are not examined in these scenarios. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.*

The empirical results show that the causal relationships between the sentiment indicators and returns are mixed if the market scenario is not classified according to investors' sentiments. The TVIX Granger causes returns in the scenario that returns are greater than zero. Although previous studies (Simon and Wiggins, 2001; Giot, 2005) define the top 20% and bottom 20% as the extreme levels of sentiment, the causality information is still mixed. The linearity test of sentiment shows that the threshold effect is significant except in the higher regime of TVIX and the lower regime of ARMS in levels. When the threshold level is decided objectively, we find that ARMS Granger causes returns in the upper regimes. The sentiment indicators in the derivatives market including TPCV and TVIX Granger cause returns in the typical and lower levels. ARMS (TPCV and TVIX) could be the leading indicator if the market is more bearish (bullish). In conclusion, ARMS (sentiments in the derivatives market) will lose the leading effect in the overbought (oversold) scenario.

We find that the causality relationship is confused if the market scenarios are not taken into account. A leading characteristic of the sentiment indicators would be captured if the extreme scenarios were to be identified. Our empirical findings confirm the noise trader explanation that the causality would run from sentiment to market behavior. The results also support the view that accurate models of prices and expected returns need to assign a prominent role to investor sentiment.

This study is limited to the assumptions of the overreaction regime identified by the upper or lower thresholds of the sentiment indicators. Other econometric methodology, for example the smooth transition autoregressive (STAR) model that is viewed as a generalization of a nonlinear model, could be applied in further research to capture the transition process from bullish regimes to bearish regimes or vice versa. Besides, the information content of the investors' overreaction could be applied to the trading strategy or other portfolio management for further research.

REFERENCES

Arms, R. W. (1989). *The Arms Index (TRIN): An introduction to the volume analysis of stock and bond markets*. McGraw-Hill Companies.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61, 1645-1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21, 129-151.

Banerjee, P. S., Doran, J. S., & Peterson, D. R. (2007). Implied volatility and future portfolio returns. *Journal of Banking & Finance*, 31, 3183-3199.

Brown, G. W. (1999). Volatility, sentiment, and noise traders. *Financial Analysts Journal*, 55, 82-90.

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

Clarke, R. G., & Statman, M. (1998). Bullish or bearish. *Financial Analysts Journal*, 54, 63-72.

Chan, K. S. (1993). Consistentency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Annals of Statistics*, 21, 520-533.

De Long, J. B., Shleifer, A., Summers, L. G., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98, 703-738.

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

Friedman, M. (1953). The case for flexible exchange rate. *Essays in Positive Economics*, Chicago: Univ. Chicago Press.

Giot, P. (2005). Relationships between implied volatility indexes and stock index returns. *Journal of Portfolio Management*, 31, 92-100.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37, 424–38.

Granger, C. W. J. (1988). Some recent developments in a concept of causality. *Journal of Econometrics*, 39, 199-211.

Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking & Finance*, 26, 2277-2299.

Olsen, R. A. (1998). Behavioral finance and its implications for stock-price volatility. *Financial Analysts Journal*, 54, 10-18.

Shefrin, H. (2000). *Beyond Greed and Fear*, Harvard Business School Press, Boston.

Simon, D. P., & Wiggins, R. A. (2001). S&P futures returns and contrary sentiment indicators. *The Journal of Futures Markets*, 21, 447-462.

Solt, M. E., & Statman, M. (1988). How useful is the sentiment index? *Financial Analysts Journal*, 44, 45-55.

Tong, H. (1983). *Threshold models in non-linear time series analysis*. Springer-Verlag: New York.

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

Whaley, R. E. (2000). The investor fear gauge. *Journal of Portfolio Management*, 26, 12-17.

ACKNOWLEDGEMENTS

This study was funded by the project of the National Science Council in Taiwan (Grant number: NSC 96-2416-H-009 -020 -MY3). The authors wish to thank the anonymous reviewers for their excellent comments, resulting in a significant improvement in the quality of this paper.

BIOGRAPHY

Dr. Her-Jiun Sheu is a Professor in the Department of Banking and Finance and President at National Chi Nan University in Taiwan. He can be contacted at National Chi Nan University, No.1, University Rd., Puli, Nantou County, 54561, Taiwan, ROC. Email: hjsheu@ncnu.edu.tw

Dr. Yang-Cheng Lu is an Associate Professor in the Finance Department at Ming Chuan University in Taiwan. He can be contacted at Ming Chuan University, No.250, Sec. 5, Jhongshan N. Rd., Shihlin District, Taipei City 111, Taiwan, ROC. Email: yclu@mcu.edu.tw

Yu-Chen Wei is a Ph.D candidate in the Management Science Department at National Chiao Tung University and Lecturer in the Finance Department at Ming Chuan University in Taiwan. She can be contacted at National Chiao Tung University, No. 1001, University Rd., Hsinchu, 300, Taiwan, ROC. Email: midclaire@gmail.com