

Stock prices, dividends, earnings, and investor sentiment

Chung Baek¹

Published online: 4 July 2015
© Springer Science+Business Media New York 2015

Abstract We investigate the relationship between fundamental market variables and investor sentiment. Our study focuses on empirical aspects that have not been explored by previous studies. We find that sentiment is co-integrated with earnings and sentiment changes cause earnings changes. Under extreme market events, however, sentiment changes tend to move more closely together with stock returns. We also find that the predictive power of sentiment changes increases with subsequent medium-term earnings changes and sentiment changes are asymmetrically more sensitive to high earnings. Our study provides a new insight to stock market participants.

Keywords Investor sentiment · Co-integration · Causality · Copula · Structural break

JEL Classification G02 · G10

1 Introduction

Many studies have examined the role of investor sentiment in the stock market. Although traditional theories do not consider investor sentiment as one of fundamental factors that affect the stock market, Shefrin and Statman (1994) show that investor sentiment can serve as one of main drivers in stock pricing. As DeBonds and Thaler (1985), Lakonishok et al. (1994) and Barberis et al. (1998) argue, in fact, investor sentiment appears to influence irrational or unexpected valuations unexplained by traditional theories or models.

A number of studies empirically examine how investor sentiment is related to stock returns. Some studies focus on the direct causal relationship between stock returns and sentiment. Jansen and Nahuis (2003) employ survey data conducted under the European

✉ Chung Baek
cbaek@troy.edu

¹ Department of Economics, Finance, and RMI, Sorrell College of Business, Troy University, Troy, AL 36082, USA

Commission and find that stock returns causes consumer confidence. Brown and Cliff (2004) show that stock returns cause sentiment using survey data as direct sentiment measures. Wang et al. (2006) and Spyrou (2012) also report similar results. However, Schmeling (2009) finds bi-directional causality between stock returns and consumer confidence across different countries and Chen and Kuo (2014) shows the similar bi-directional causality using the Conference Board Consumer Confidence Index (CBCCI) as a proxy for consumer confidence.

On the other hand, other studies investigate more general relationships. Using Michigan survey as consumer confidence index and the Wilshire 5000 stock index, Otoo (1999) finds that sentiment increases with stock prices and stock returns play a role as a leading indicator. Fisher and Statman (2003) adopt two separate sentiment measures: American Association of Individual Investors (AAII) survey as an individual sentiment measure and Merrill Lynch survey as an institutional sentiment measure and find that individual sentiment shows a significant relationship and sentiment has some predictive power for small cap stock returns. Baker and Wurgler (2006) find that stock returns are conditional on sentiment and conclude that the asset pricing model should incorporate sentiment.

There are also studies that consider different aspects in the market. Huth et al. (1994) show that the sentiment indices are useful to forecast aggregate spending and business activity using the University of Michigan survey as a consumer sentiment index and the Conference Board's index of consumer confidence. Fisher and Statman (2006) use the bullish sentiment index as an individual sentiment and find that the bullish sentiment index can work as a market timing instrument. Liao et al. (2011) show that investor sentiment can play a critical role in explaining fund manager herding based on the principal component analysis. Chen and Kuo (2014) investigate how investor sentiment affects the Eurodollar option smile and find a significant relationship between interest rate volatility smile and sentiment. As the use of online trading increases, several studies attempt to create the sentiment index from online stock message boards that measure individual sentiment. These studies include Tumarkin and Whitelaw (2001), Antweiler and Frank (2004) and Das and Chen (2007).

Most previous studies provide their results based on a single proxy for sentiment. Baker and Wurgler (2006), however, argue that all single proxies for investor sentiment are imperfect and defective. They construct a sentiment index on the basis of the first principal component of six major sentiment proxies. Since their index represents overall sentiment, it can be considered the composite sentiment index that combines individual and institutional sentiments. For our study, we choose the composite sentiment index developed by Baker and Wurgler (2006).

Although previous studies provide various information about the role of investor sentiment, there still exists room to delve into investor sentiment in the stock market. First, unlike previous studies, our study examines the relationship between investor sentiment and fundamental market variables: stock prices, dividends, and earnings. It seems important to study investor sentiment with dividends and earnings as well as stock prices because they constitute fundamental market ratios and provide critical information to market participants. Second, most previous studies directly investigate the causal relationship without testing the co-integration. There are many studies that investigate the co-integration associated with market variables other than investor sentiment. These studies include Chang and Nieh (2001), Assaf (2006) and Gosh and Clayton (2006). We implement the co-integration test in order to identify the long run equilibrium between market variables and investor sentiment. If two variables are co-integrated, we need to incorporate an error correction term into the causality test. Third, with real data rather than nominal

data we attempt to identify the true relationship between market variables and investor sentiment by making our results comparable over the whole period of our study regardless of inflation. Fourth, we examine how investor sentiment moves with market variables assuming that extreme market events occur. We calculate bivariate probabilities based on two copulas and see how the dependence structure between investor sentiment and market variables affects their relationship under extreme market scenarios. Lastly, we measure the predictive power of investor sentiment. We not only examine the predictive power of investor sentiment over the whole period of our study but also see how the predictive power of investor sentiment alters if there exist structural breaks.

We describe all data in Sect. 2 and show empirical results and discussions from Sect. 3 through Sect. 6. Then, we conclude in Sect. 7.

2 Data

We use S&P 500 stock index data downloadable from Robert Shiller's website. His data provide all real values for market prices, dividends, and earnings. For the sentiment index, we employ the composite sentiment index of Baker and Wurgler (2006) obtained from Jeffrey Wurgler's website. His website provides changes in the composite sentiment index as well as the composite sentiment index level. All stock and sentiment data are monthly data. We select the data from 1965 to 2010 because the latest (monthly) version of the composite sentiment index covers July 1965 to December 2010. Figures 1, 2, 3 and 4 show real prices, real dividends, real earnings and the composite sentiment index, respectively.

We also collect monthly data of some economic variables from the website of the Federal Reserve Bank of St. Louis. These variables include the consumer price index, industrial production index, and 10-year Treasury note. They are used as control variables in two regressions to measure the predictive power of sentiment changes. We truncate data at the beginning or the end of the whole period if necessary for short or long-term computations.

3 Co-integration and causality

First of all, we need to implement the unit root test for all variables. We examine if variables have a unit root in their levels and differences. The typical unit root test requires the following regressions:

Fig. 1 Historical real prices

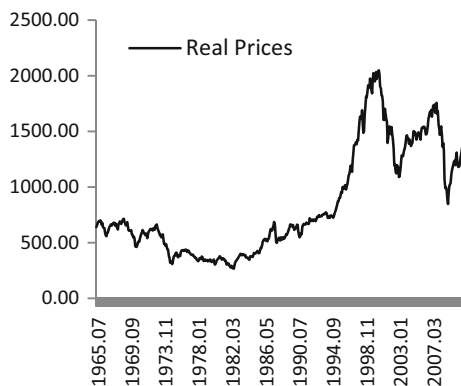


Fig. 2 Historical real dividends

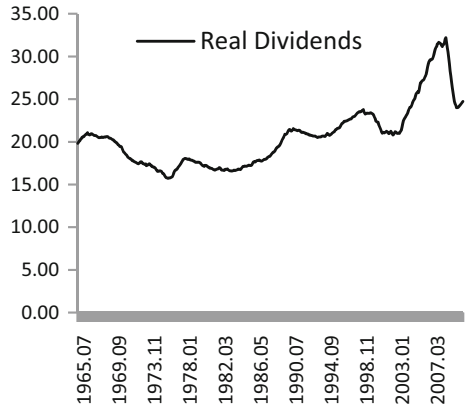


Fig. 3 Historical real earnings

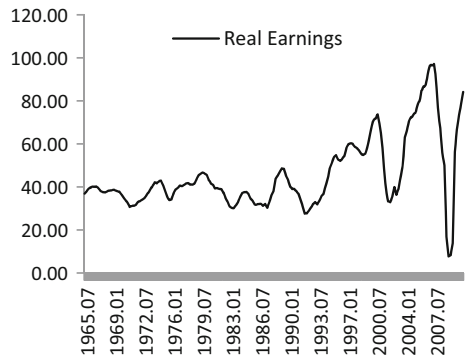
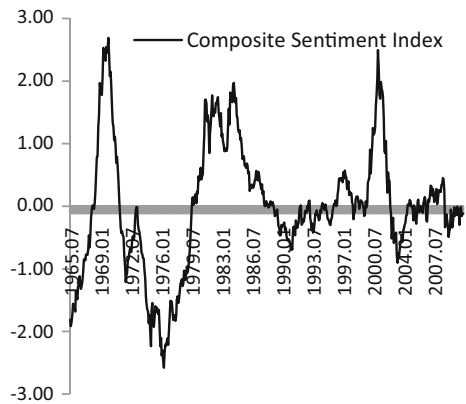


Fig. 4 Historical sentiment index



$$\Delta x_t = \beta_0 + \beta_1 x_{t-1} + \sum_{j=1}^n \gamma_j \Delta x_{t-j} + \varepsilon_t \tag{1}$$

$$\Delta x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 t + \sum_{j=1}^n \gamma_j \Delta x_{t-j} + \varepsilon_t \tag{2}$$

where the null hypothesis is $\beta_1 = 0$ (x_t has a unit root). Equation (1) has no trend whereas Eq. (2) has a trend. Although several rules compete for determining the optimal lag length, there seems to be no perfect rule to choose the number of lags. Throughout our study, we adopt the well-known Akaike information criterion (AIC) that has been traditionally and widely used. Instead, to minimize the optimal lag problem, we show co-integration results with three cases including one lag more and less than the optimal lag length.

In Table 1, we show results for the unit root test using the Augmented Dickey–Fuller (ADF) and the Phillips–Perron (PP) methods. For variables’ levels in Panels A and B, we fail to reject the null hypotheses for all variables at both 1 and 5 % significance levels except only for sentiment with no trend under the ADF method. In Panels C and D, the same unit root test is repeated with variables’ log differences. Since, however, Baker and Wurgler (2006) provide changes in the composite sentiment index and they are also downloadable from Wurgler’s website, we use them for sentiment differences. All null hypotheses with differences are rejected even at the 1 % significance level. As a result, real prices, real dividends, real earnings, and sentiment turn out to be integrated of order one.

Since all variables are I(1), we investigate if there exists the co-integrating relationship between sentiment and market variables. The co-integration test identifies the existence of the long run equilibrium between variables. If two variables are co-integrated, we have to construct the vector error correction model (VECM) for the purpose of conducting the causality test. Then, we can see how two variables move toward their long run equilibrium through the error correction term.

We employ the Johansen test to identify the existence of the co-integration between sentiment and market variables. To get statistically robust results, we test the co-integration with one lag ahead of and behind the optimal lag length as well. In Table 2, R is the

Table 1 Unit root test

	Prices	Dividends	Earnings	Sentiment
<i>Panel A: Level (augmented Dickey–Fuller)</i>				
Constant with no trend	-1.1071	-1.5462	-1.3085	-3.1925*
Constant with trend	-2.4276	-3.2173	-2.6633	-3.1749
<i>Panel B: Level (Phillips–Perron)</i>				
Constant with no trend	-0.7174	-0.1345	-0.8888	-2.3717
Constant with trend	-1.7766	-1.5220	-1.6841	-2.2570
<i>Panel C: Log difference (augmented Dickey–Fuller)</i>				
Constant with no trend	-5.5313**	-4.3371**	-5.1815**	-4.9768**
Constant with trend	-5.5702**	-4.4454**	-5.1840**	-4.9674**
<i>Panel D: Log difference (Phillips–Perron)</i>				
Constant with no trend	-17.7580**	-8.1754**	-8.6571**	-24.8460**
Constant with trend	-17.7740**	-8.2257**	-8.6551**	-24.8530**

We conduct the unit root test using the following regressions:

$$\text{Constant with no trend: } \Delta x_t = \beta_0 + \beta_1 x_{t-1} + \sum_{j=1}^n \gamma_j \Delta x_{t-j} + \varepsilon_t$$

$$\text{Constant with trend: } \Delta x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 t + \sum_{j=1}^n \gamma_j \Delta x_{t-j} + \varepsilon_t$$

where the null hypothesis is $\beta_1 = 0$ (x_t has a unit root). We use both ADF and PP tests for levels and log differences. For sentiment differences, we use changes in the composite sentiment index obtained from Wurgler’s website

*, ** Statistical significance at the 5 and 1 % levels, respectively

Table 2 Co-integration—Johansen test

	Prices and sentiment			Dividends and sentiment			Earnings and sentiment					
	N	R	Trace	Maximum eigenvalue	N	R	Trace	Maximum eigenvalue	N	R	Trace	Maximum eigenvalue
9 Lags	0	1	11.814	10.658	1 Lag	0	8.022	6.619	2 Lags	0	24.096**	16.819*
	1	0	1.156	1.156	1	1	1.403	1.403		1	7.278	7.278
10 Lags	0	1	14.156	12.876	2 Lags	0	8.022	6.619	3 Lags	0	20.948*	13.637
	1	0	1.280	1.280	1	1	1.403	1.403		1	7.311	7.311
11 Lags	0	1	16.334	15.043*	3 Lags	0	8.380	6.863	4 Lags	0	19.210*	11.446
	1	0	1.291	1.291	1	1	1.516	1.516		1	7.765	7.765

R is the number of co-integrating vectors and N is the number of lags included in the co-integration test. To get statistically robust results, we test the co-integration with one lag ahead of and behind the optimal lag length. While the null hypothesis for the trace test is that the number of co-integrating vectors is less than or equal to R, the null hypothesis for the maximum eigenvalue test is that the number of co-integrating vectors is R

*, ** Statistical significance at the 5 and 1 % levels, respectively

number of co-integrating vectors and N is the number of lags included in the test. The null hypothesis for the trace method is that the number of co-integrating vectors is less than or equal to R whereas the null hypothesis for the maximum eigenvalue method is that the number of co-integrating vectors is equal to R .

For real dividends and sentiment, we fail to reject the null hypotheses with all three lags using both trace and maximum eigenvalue methods. This means that there is no co-integrating relationship between real dividends and sentiment. Regarding real prices and sentiment, we fail to reject the null hypotheses except only for the case that the number of lags is 11. We, however, identify the existence of the co-integration between real earnings and sentiment. With the trace method, we reject the null hypotheses with all three lags when R is 0 and fail to reject the null hypotheses when R is 1. The maximum eigenvalue method also shows the same result in the case that the number of lags is 2. As a result, only real earnings and sentiment are co-integrated.

Now, we examine the causal relationship between market variables and sentiment. Since all variables should be stationary for the Granger causality test, we use stock returns, dividend changes, earnings changes, and sentiment changes because they are all stationary as shown in Table 1. According to Engle and Granger (1987), there has to be at least one directional causality if two variables are co-integrated. Also, we should incorporate the error correction term into the causality test in order to correctly identify the causal relationship between co-integrated variables. To do this, we construct the VECM for variables that are co-integrated and the vector autoregressive (VAR) model for variables that are not co-integrated.

$$\Delta x_t = \alpha + \sum_{j=1}^n \beta_j \Delta x_{t-j} + \sum_{j=1}^n \gamma_j \Delta y_{t-j} + \varepsilon_t \tag{3}$$

$$\Delta x_t = \alpha + \alpha_1 \widehat{e}_{t-1} + \sum_{j=1}^n \beta_j \Delta x_{t-j} + \sum_{j=1}^n \gamma_j \Delta y_{t-j} + \varepsilon_t \tag{4}$$

where α_1 is the adjustment coefficient of the error correction term and \widehat{e}_{t-1} is obtained from the co-integrating regression, $x_t = \mu_0 + \mu_1 y_t + \varepsilon_t$. We use the VECM for earnings changes and sentiment changes and the VAR model for other variables. When co-integrated variables deviate from their long run equilibrium, the departure should be corrected through the error correction term in the VECM. We calculate F-values for the null hypothesis that $\gamma_j = 0$.

In Panel A of Table 3, we do not find any significant causal relationship. In particular, the causality between dividend changes and sentiment changes is not statistically significant at all. Since the dividend payout is much more likely to be affected by the firm's dividend policy, this finding is not surprising.

For stock returns and sentiment changes, the causal relationship between them does not appear to be statistically significant. This finding is inconsistent with those of previous studies (see Brown and Cliff 2004; Spyrou 2012) that report one directional causality from stock returns to sentiment changes. We may consider a very weak causality from stock returns to sentiment changes but its F-value (1.38) is significant at most at the 20 % significant level. Since we use real values instead of nominal values, the true causal relationship between real stock returns and sentiment changes seems to be statistically insignificant.

On the other hand, Panel B of Table 3 shows that sentiment changes cause earnings changes rather than vice versa. This finding is new and interesting because it implies that

Table 3 Causality test

Null hypothesis	F-statistic	
<i>Panel A: Causality with the vector autoregressive (VAR) model</i>		
$\Delta S \not\Rightarrow \Delta P$ (Δ Sentiment does not cause Δ Prices)	1.10	
$\Delta P \not\Rightarrow \Delta S$ (Δ Prices does not cause Δ Sentiment)	1.38	
$\Delta S \not\Rightarrow \Delta D$ (Δ Sentiment does not cause Δ Dividends)	0.45	
$\Delta D \not\Rightarrow \Delta S$ (Δ Dividends does not cause Δ Sentiment)	0.00	
Null hypothesis	F-statistic	α_1 (t-statistic)
<i>Panel B: Causality with the vector error correction model (VECM)</i>		
$\Delta S \not\Rightarrow \Delta E$ (Δ Sentiment does not cause Δ Earnings)	2.09*	-0.0003 (-3.12**)
$\Delta E \not\Rightarrow \Delta S$ (Δ Earnings does not cause Δ Sentiment)	1.78	-0.1578 (-3.24**)

We use the following VAR to test causality:

$$\Delta x_t = \alpha + \sum_{j=1}^n \beta_j \Delta x_{t-j} + \sum_{j=1}^n \gamma_j \Delta y_{t-j} + \varepsilon_t$$

The causality between sentiment and earnings is tested using the following VECM:

$$\Delta x_t = \alpha + \alpha_1 \hat{e}_{t-1} + \sum_{j=1}^n \beta_j \Delta x_{t-j} + \sum_{j=1}^n \gamma_j \Delta y_{t-j} + \varepsilon_t$$

where α_1 is the adjustment coefficient of the error correction term and \hat{e}_{t-1} is obtained from the co-integrating regression, $x_t = \mu_0 + \mu_1 y_t + \varepsilon_t$. We calculate F-values for the null hypothesis that $\gamma_j = 0$

*, ** Statistical significance at the 5 and 1 % levels, respectively

we can use sentiment changes as a leading indicator for earnings changes. As Schmeling (2009) mentions, since sentiment is influenced by various economic factors and news media, the fact that sentiment changes lead earnings changes appears to be quite reasonable. The adjustment coefficient of the error correction term is statistically negatively significant for both directions. This means that whenever variables deviate from their long run equilibrium the error will be inversely adjusted toward the long run equilibrium.

4 Copula and extreme movements

In this section, we are interested in investigating how sentiment changes move with stock returns, dividend changes, and earnings changes under extreme market events (e.g., stock market crash or bubble). A copula is very useful to explain multivariate extreme movements. We use two asymmetric Archimedean copulas: Clayton copula and Gumbel copula. Archimedean copulas are popular and widely used for many applications. While the Clayton copula can be used to examine extreme movements under the bearish market scenario the Gumbel copula can be used to see extreme movements under the bullish market scenario. According to Klugman et al. (2008), a multivariate copula is the joint distribution function of Uniform random variables as follows:

$$K(u_1, u_2, \dots, u_p) = \Pr(U_1 \leq u_1, U_2 \leq u_2, \dots, U_p \leq u_p) \tag{5}$$

where $U_i \sim U(1, 0)$. According to Sklar’s theorem, for any joint distribution function F, there exists a unique copula K that satisfies the following equation.

$$F(x_1, x_2, \dots, x_p) = K(F_1(x_1), F_2(x_2), \dots, F_p(x_p)) \tag{6}$$

Therefore, we can build a multivariate joint distribution with marginal distributions and a specifically defined copula. One of different classes in copula is the Archimedean copula that can be easily constructed with a defined generator. The form of Archimedean copula is as follows:

$$K(u_1, u_2, \dots, u_p) = \pi^{-1}(\pi(u_1) + \pi(u_2) + \dots + \pi(u_p)) \tag{7}$$

where $\pi(u)$ is a generator. Since we investigate a copula-based relationship between market variables and sentiment, we use bivariate copulas.

It is common to have dependence measures associated with the copula. The most popular dependence measure is Kendall’s tau (τ) that has a mathematical link with Archimedean copulas. The definition of Kendall’s tau is as follows:

$$\tau(X, Y) = \Pr[(X_1 - X_2)(Y_1 - Y_2) > 0] - \Pr[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{8}$$

where iid bivariate random variables (X_1, Y_1) and (X_2, Y_2) have marginal distribution $F_X(x)$ for X_1 and X_2 and marginal distribution $F_Y(y)$ for Y_1 and Y_2 . The relationship between Kendall’s tau and the copula function K is derived with the following equation:

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 K(u_1, u_2)k(u_1, u_2)du_1du_2 - 1 \tag{9}$$

where $k(u_1, u_2) = \frac{\partial^2 K(u_1, u_2)}{\partial u_1 \partial u_2}$. The Clayton copula has greater dependence weight to the negative tail and can be constructed as follows:

$$K(u_1, u_2) = \text{Max}\left[(u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, 0 \right] \tag{10}$$

$$\tau = \theta / (2 + \theta) \tag{11}$$

where $\pi(u) = -\frac{1}{\theta}(1 - u^{-\theta})$ and θ is the dependence parameter. Equation (11) is the mathematically derived relationship between Kendall’s tau and the single parameter in the Clayton copula. The Gumbel copula has greater dependence weight to the positive tail and can be derived as follows.

$$K(u_1, u_2) = \exp\left\{ -\left[(-\ln u_1)^\lambda + (-\ln u_2)^\lambda \right]^{1/\lambda} \right\} \tag{12}$$

$$\tau = 1 - 1/\lambda \tag{13}$$

where $\pi(u) = (-\ln u)^\lambda$ and λ is the dependence parameter.

In Table 4, we calculate Kendall’s tau for each bivariate case and estimate the single parameter of the copula function. Panel A shows dependence parameter estimates in both Clayton and Gumbel copulas. For the bivariate case of dividend changes and sentiment changes, the dependence parameter of the Gumbel copula is not defined.

In Panel B of Table 4, we calculate probabilities for extreme movements using copulas. Under the bearish market scenario with the Clayton copula, the probability that stock returns and sentiment changes fall within their lowest 10th percentile at the same time is 3.13 %. This is about 2.4 times as high as the probability (1.28 %) that earnings changes and sentiment changes fall within their lowest 10th percentile at the same time. The probability that stock returns and sentiment changes fall within their lowest 1st percentile

Table 4 Copula parameter estimates and extreme probabilities

Bivariate vectors	Dependence parameter (θ) of Clayton copula			Dependence parameter (λ) of Gumbel copula		
<i>Panel A: Dependence parameter</i>						
($\Delta P, \Delta S$)	0.4108			1.2054		
($\Delta D, \Delta S$)	−0.0256			Not defined		
($\Delta E, \Delta S$)	0.0520			1.0260		
Bivariate vectors	Clayton copula			Gumbel copula		
	Lowest 1st P	Lowest 5th P	Lowest 10th P	Highest 1st P	Highest 5th P	Highest 10th P
<i>Panel B: Probabilities for extreme values</i>						
($\Delta P, \Delta S$)	0.00223916	0.01358606	0.03129187	0.00229718	0.01287337	0.02923971
($\Delta D, \Delta S$)	0.00005397	0.00194870	0.00865584	Not defined	Not defined	Not defined
($\Delta E, \Delta S$)	0.00024429	0.00374717	0.01279743	0.00044339	0.00411496	0.01298007

We use two Archimedean copulas: Clayton and Gumbel. While the Clayton copula has greater dependence weight to the negative tail the Gumbel copula has greater dependence weight to the positive tail. The lowest nth P is the probability that both values fall within their lowest nth percentiles at the same time and the highest nth P is the probability that both values fall within their highest nth percentiles at the same time. We calculate Kendall’s tau for each bivariate case and estimate the single parameter of the copula using the relationship between Kendall’s tau and the copula function. Under the bullish market scenario with Gumbel copula, we calculate the highest nth percentile using the survival function of the copula

at the same time is 0.22 % which is about nine times as high as the probability (0.024 %) that earnings changes and sentiment changes fall within their lowest 1st percentile at the same time. Probabilities for dividend changes and sentiment changes are the lowest in all three percentiles. As a result, under the assumption that large losses occur in the stock market, stock returns and sentiment changes are likely to move most closely together.

On the other hand, under the bullish market scenario with the Gumbel copula, we calculate the highest nth percentile using the survival function of the copula.

$$K_S(u_1, u_2) = 1 - u_1 - u_2 - K(u_1, u_2) \tag{14}$$

where $K_S(u_1, u_2)$ is the survival function of the copula, $K(u_1, u_2)$. The probability that stock returns and sentiment changes fall within their highest 1st percentile at the same time is 0.23 % which is about five times as high as the probability (0.044 %) that earnings changes and sentiment changes fall within their highest 1st percentile at the same time. In the same way, the probability that stock returns and sentiment changes fall within their highest 10th percentile at the same time is 2.92 % which is about 2.2 times as high as the probability (1.30 %) that earnings changes and sentiment changes fall within their highest 10th percentile at the same time. When we assume that large gains occur in the stock market, stock returns and sentiment changes are also likely to move most closely together.

As we see in the previous section, only earnings and sentiment are co-integrated and their causal relationship is formed from sentiment changes to earnings changes. When, however, we assume extreme market events that may occur in the stock market, stock returns and sentiment changes turn out to move most closely together. The dependence structure between stock returns and sentiment changes may provide important information under extreme market events.

5 Sentiment and prediction

To investigate the predictive power of sentiment changes, we adopt two approaches. One is to use incremental adjusted R^2 shown by Spyrou (2012) and Schmeling (2009) and the other is to use prediction criterion (PC) developed by Amemiya (1980). Although the recent related studies employ incremental adjusted R^2 s to examine the predictive power of sentiment changes, we also use PC because it explicitly considers a loss function to measure prediction errors. Since dividend changes are not significantly related to sentiment changes in previous sections, we examine the predictive power of sentiment changes only for stock returns and earnings changes. We construct two separate regressions to examine the predictive power of sentiment changes for future stock returns and earnings changes.

$$\Delta y_{t+i} = \alpha_0 + \sum_{j=1}^n \beta_j \Delta c_{jt} + \varepsilon_t \tag{15}$$

$$\Delta y_{t+i} = \alpha_0 + \sum_{j=1}^n \beta_j \Delta c_{jt} + \Delta s_t + \varepsilon_t \tag{16}$$

where Δy_{t+j} is the subsequent i -month stock (real) return or earnings change at time t and Δs_t is the sentiment change at time t . We use some macroeconomic variables as control factors (c_{jt}). These variables are consumer price index (CPI), industrial production index (IPI), and 10-year Treasury note rate (TN). Since the GDP data are available only on a quarterly or annual basis, instead, we include IPI as one of control factors.

As shown in Spyrou (2012) and Schmeling (2009), we check the incremental adjusted R^2 to see if sentiment changes add some explanatory power to forecast future stock returns or earnings changes. We calculate the incremental adjusted R^2 by examining the adjusted R^2 of Eq. (16) relative to that of Eq. (15). We look into the predictive power of sentiment changes to forecast subsequent 1-month stock returns or earnings changes through subsequent 12-month stock returns or earnings changes. Thus, we can compare short-, medium-, and long-term predictive powers of sentiment changes.

Schmeling (2009) shows that the sentiment factor adds some explanatory power to predict future short- and medium-term stock returns by reporting incremental adjusted R^2 s. Spyrou (2012) finds that sentiment changes add a weak explanatory power to forecast stock returns in terms of incremental adjusted R^2 s.

In Table 5, Panel A shows incremental adjusted R^2 s by subtracting adjusted R^2 s of Eq. (15) from those of Eq. (16). When we investigate the predictive power of sentiment changes for subsequent stock returns, it turns out that sentiment changes have the most predictive power for subsequent 1-month stock returns. The incremental adjusted R^2 for subsequent 1-month stock returns is 0.86 %. Sentiment changes also show some predictive power for subsequent 6- and 11-month stock returns but their incremental adjusted R^2 s are relatively small.

For subsequent earnings changes, sentiment changes have the most predictive power for subsequent 6-month earnings changes. Its incremental adjusted R^2 is 1.29 %. Also, sentiment changes have some predictive power for subsequent 7-, 8-, 9-, and 10-month earnings changes. In other words, sentiment changes have some explanatory power for future medium-term earnings changes. Although adjusted R^2 s are small, it seems to be typical as found in previous studies.

On the other hand, Panel B shows decremental PCs between Eqs. (15) and (16). Since the PC should be minimized, the positive decremental PC represents the added explanatory

Table 5 Predictive power of sentiment changes—stock returns and earnings changes

Dependent variable (stock return or earnings changes)	Stock return and sentiment change			Earnings changes and sentiment change		
	Adj R ² of Eq. (15)	Adj R ² of Eq. (16)	Incremental R ²	Adj R ² of Eq. (15)	Adj R ² of Eq. (16)	Incremental R ²
<i>Panel A. Adjusted R²</i>						
Subsequent 1-month	0.0458	0.0544	0.0086	0.0886	0.0883	-0.0003
Subsequent 2-month	-0.0017	-0.0034	-0.0017	0.0907	0.0908	0.0001
Subsequent 3-month	0.0024	0.0007	-0.0017	0.1100	0.1085	-0.0015
Subsequent 4-month	0.0069	0.0067	-0.0002	0.0389	0.0374	-0.0015
Subsequent 5-month	0.0127	0.0110	-0.0017	0.0254	0.0237	-0.0017
Subsequent 6-month	0.0032	0.0088	0.0056	0.0299	0.0428	0.0129
Subsequent 7-month	0.0011	0.0014	0.0003	0.0192	0.0244	0.0052
Subsequent 8-month	0.0158	0.0141	-0.0017	0.0321	0.0370	0.0049
Subsequent 9-month	0.0142	0.0134	-0.0008	0.0390	0.0447	0.0057
Subsequent 10-month	0.0159	0.0151	-0.0008	0.0547	0.0585	0.0038
Subsequent 11-month	0.0169	0.0186	0.0017	0.0989	0.0975	-0.0014
Subsequent 12-month	-0.0040	-0.0047	-0.0007	0.0759	0.0745	-0.0014
<i>Panel B. Prediction criterion (PC)</i>						
Dependent variable (stock return or earnings changes)	Stock return and sentiment change			Earnings changes and sentiment change		
	PC ($\times 10^{-2}$) of Eq. (15)	PC ($\times 10^{-2}$) of Eq. (16)	Decremental PC ($\times 10^{-2}$)	PC ($\times 10^{-2}$) of Eq. (15)	PC ($\times 10^{-2}$) of Eq. (16)	Decremental PC ($\times 10^{-2}$)
Subsequent 1-month	0.1375	0.1369	0.0006	0.3001	0.3003	-0.0002
Subsequent 2-month	0.1433	0.1438	-0.0005	0.2974	0.2976	-0.0002
Subsequent 3-month	0.1434	0.1439	-0.0005	0.2891	0.2900	-0.0009
Subsequent 4-month	0.1426	0.1430	-0.0004	0.3114	0.3124	-0.001
Subsequent 5-month	0.1424	0.1429	-0.0005	0.3162	0.3174	-0.0012
Subsequent 6-month	0.1446	0.1443	0.0003	0.3167	0.3139	0.0028

Table 5 continued

Dependent variable (stock return or earnings changes)	Stock return and sentiment change		Earnings changes and sentiment change	
	PC ($\times 10^{-2}$) of Eq. (15)	PC ($\times 10^{-2}$) of Eq. (16)	PC ($\times 10^{-2}$) of Eq. (15)	PC ($\times 10^{-2}$) of Eq. (16)
Subsequent 7-month	0.1447	0.1451	0.3228	0.3224
Subsequent 8-month	0.1438	0.1442	0.3212	0.3211
Subsequent 9-month	0.1440	0.1442	0.3204	0.3201
Subsequent 10-month	0.1429	0.1434	0.3144	0.3146
Subsequent 11-month	0.1433	0.1435	0.3119	0.3123
Subsequent 12-month	0.1447	0.1450	0.3167	0.3173

According to Amemiya (1980), prediction criterion (PC) is calculated as follows:

$$PC = \hat{\sigma}_1^2 \left(1 + \frac{K_1}{T} \right)$$

$$\hat{\sigma}_1^2 = (T - K_1)^{-1} y' M_1 y$$

$$M_1 = I - X_1 (X_1' X_1)^{-1} X_1'$$

where K_1 , is the number of regressors; T , is the sample size; y , is the vector of the dependent variable; and X_1 , is the vector of explanatory variables

power of sentiment changes. As we see in the 4th and 7th columns, the results shown in Panel B are very similar to those of Panel A. Particularly, we confirm that sentiment changes have the most predictive power for medium-term earnings changes.

6 Impact of structural breaks on the predictive power of sentiment changes

Since earnings changes and sentiment changes are most closely related on the basis of the results obtained from previous sections, we attempt to identify structural breaks in real earnings and examine whether the predictive power of sentiment changes increases or decreases in sub-periods divided by structural breaks of real earnings. Structural breaks often play an important role when we investigate time series variables because permanent structural changes in the long run may influence a specific relationship between variables.

To detect structural breaks in real earnings, we employ well-known methods: Cumulative Sum Chart and Chow test. The cumulative sum chart shows detailed movements in a time series and thus, we can identify candidates for structural breaks. Then, with the Chow test, we test whether or not those breaks really cause structural changes. We calculate the cumulative sum of real earnings using the following equation:

$$SUM_t = SUM_{t-1} + (E_t - \bar{E}) \quad \text{for } t = 1, 2, \dots, n \tag{17}$$

where E_t is earnings at time t and \bar{E} is the historical average of earnings. The cumulative sum ends at zero starting from a zero value of SUM_0 .

As we see in Fig. 5, it is clear that there exists one single candidate for structural break in the time series of real earnings. It is October 1994 in which the cumulative sum is minimized. In other words, we observe a clear decreasing trend before October 1994 and a

Fig. 5 Historical cumulative sum of real earnings. We calculate the cumulative sum using the following equation: $SUM_t = SUM_{t-1} + (E_t - \bar{E})$ for $t = 1, 2, \dots, n$ where E_t is earnings at time t and \bar{E} is the historical average of earnings. The cumulative sum ends at zero starting from a zero value of SUM_0

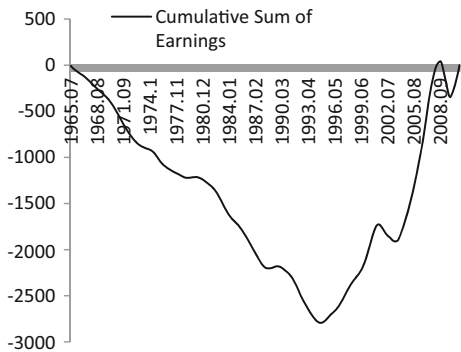


Table 6 Structural break in real earnings

Break point	First period	Second period	N1	N2	Chow test statistic
10/1994	7/1965–10/1994	11/1994–12/2010	352	194	66.93**

We use the Chow test to confirm that two sub-periods are structurally different

$N1$ number of the first period data, $N2$ number of the second period data

** Statistical significance at the 1 % significance level

Table 7 Predictive power of sentiment changes—high and low earnings periods

Dependent variable (earnings changes)	Low earnings period (7/1965–10/1994)			High earnings period (11/1994–12/2010)		
	Adj R ² of Eq. (15)	Adj R ² of Eq. (16)	Incremental R ²	Adj R ² of Eq. (15)	Adj R ² of Eq. (16)	Incremental R ²
<i>Panel A: Adjusted R²</i>						
Subsequent 1-month	0.1346	0.1343	-0.0003	0.2039	0.1998	-0.0041
Subsequent 2-month	0.1459	0.1458	-0.0001	0.1935	0.1990	0.0055
Subsequent 3-month	0.1234	0.1248	0.0014	0.2299	0.2255	-0.0044
Subsequent 4-month	0.1282	0.1279	-0.0003	0.0667	0.0615	-0.0052
Subsequent 5-month	0.1486	0.1463	-0.0023	0.0216	0.0169	-0.0047
Subsequent 6-month	0.1514	0.1491	-0.0023	0.0373	0.0778	0.0405
Subsequent 7-month	0.1649	0.1627	-0.0022	0.0453	0.0661	0.0208
Subsequent 8-month	0.1552	0.1541	-0.0011	0.0759	0.0954	0.0195
Subsequent 9-month	0.1588	0.1592	0.0004	0.0915	0.1127	0.0212
Subsequent 10-month	0.1386	0.1386	0.0000	0.1446	0.1612	0.0166
Subsequent 11-month	0.1158	0.1136	-0.0022	0.2525	0.2490	-0.0035
Subsequent 12-month	0.0933	0.0950	0.0017	0.1794	0.1748	-0.0046
<i>Panel B: Prediction criterion (PC)</i>						
Dependent variable (earnings changes)	Low earnings period (7/1965–10/1994)			High earnings period (11/1994–12/2010)		
	PC ($\times 10^{-3}$) of Eq. (15)	PC ($\times 10^{-3}$) of Eq. (16)	Decremental PC ($\times 10^{-3}$)	PC ($\times 10^{-2}$) of Eq. (15)	PC ($\times 10^{-2}$) of Eq. (16)	Decremental PC ($\times 10^{-2}$)
Subsequent 1-month	0.1883	0.1889	-0.0006	0.7615	0.7676	-0.0061
Subsequent 2-month	0.1889	0.1894	-0.0005	0.7588	0.7641	-0.0053
Subsequent 3-month	0.1949	0.1950	-0.0001	0.7126	0.7202	-0.0076
Subsequent 4-month	0.1943	0.1948	-0.0005	0.8586	0.8676	-0.0090
Subsequent 5-month	0.1910	0.1920	-0.0010	0.8980	0.9076	-0.0096
Subsequent 6-month	0.1904	0.1914	-0.0010	0.8935	0.8692	0.0243

Table 7 continued

Dependent variable (earnings changes)	Low earnings period (7/1965–10/1994)		High earnings period (11/1994–12/2010)		
	PC ($\times 10^{-3}$) of Eq. (15)	PC ($\times 10^{-3}$) of Eq. (16)	PC ($\times 10^{-2}$) of Eq. (15)	PC ($\times 10^{-2}$) of Eq. (16)	Incremental PC ($\times 10^{-2}$)
Subsequent 7-month	0.1891	0.1902	0.8963	0.8895	0.0068
Subsequent 8-month	0.1932	0.1941	0.8795	0.8745	0.0050
Subsequent 9-month	0.1962	0.1970	0.8670	0.8611	0.0059
Subsequent 10-month	0.2033	0.2042	0.8108	0.8071	0.0037
Subsequent 11-month	0.2117	0.2129	0.7648	0.7729	-0.0081
Subsequent 12-month	0.2175	0.2182	0.8213	0.8293	-0.0080

clear increasing trend after October 1994. Since we focus on structural changes over the long run, we ignore some small fluctuations over the short run.

To statistically confirm if real earnings have a structural break in October 1994, we use the Chow test. In Table 6, the test statistic, 66.93, is significant at the 1 % level. This means that two sub-periods (before and after October 1994) are structurally significantly different. While the first sub-period is considered the low earnings period because earnings are decreasing relative to the historical average, the second sub-period is considered the high earnings period because of its increasing trend. Although it is not the purpose of our study to explain why earnings have structurally changed since the break date, it may be attributed to significant changes in corporate world (e.g., the fourth wave in mergers and acquisitions and remarkable progresses in IT industry) that have been initiated since the mid-1980s.

Using Eqs. (15) and (16), we re-calculate incremental adjusted R^2 s and decremental PCs in both low and high earnings periods. In Table 7, we find that the predictive power of sentiment changes considerably increases for subsequent medium-term (6-, 7-, 8-, 9-, and 10-month) earnings changes in the high earnings period. Those incremental adjusted R^2 s range from 1.66 to 4.05 %. These results appear to be quite significant when compared with the result that incremental adjusted R^2 s shown in Spyrou (2012) range from 0.2 to 2.41 % and those shown in Schmeling (2009) range from 1 to 4 %. The decremental PCs in the high earnings period range from 0.37 basis points to 2.43 basis points for subsequent medium-term earnings changes. This is consistent with the results shown in Panel A. As a matter of fact, positive incremental adjusted R^2 s and decremental PCs in the high earnings period are overwhelmingly concentrated in subsequent medium-term earnings changes.

Consequently, adjusted R^2 s and PCs for medium-term earnings changes significantly positively change in the high earnings period after the structural change. Thus, sentiment changes seem to be asymmetrically more sensitive to high earnings. If we are to use sentiment changes as a leading indicator for earnings changes, it may be critical to recognize high or low earnings.

7 Conclusion

Most previous studies employ various single proxies for investor sentiment and focus on investigating the relationship between stock returns and sentiment changes. To overcome the imperfection of single sentiment proxies, we use the composite sentiment index developed by Baker and Wurgler (2006). Our study is different from previous studies in several aspects.

First, we examine how sentiment is related to dividends and earnings as well as stock prices. These market variables are important because they form fundamental market ratios and provide critical information to market participants. In addition, we use real data rather than nominal data to draw the true relationship between market variables and sentiment. Second, we examine not only the causal relationship but also the co-integrating relationship. If two variables are co-integrated, their causality should be tested with an error correction term. Third, we show how sentiment changes move together with stock returns, dividend changes, and earnings changes under extreme market scenarios. This provides information about the relative dependence structure associated with extreme values. Fourth, we investigate the predictive power of sentiment changes for stock returns and earnings changes and see if sentiment changes asymmetrically respond with structural changes.

We discover that sentiment is co-integrated with earnings rather than stock prices or dividends and the causal relationship is formed from sentiment changes to earnings changes. Thus, sentiment changes appear to lead earnings changes. Sentiment changes, however, tend to move more closely together with stock returns under extreme market scenarios. This means that the relationship between stock returns and sentiment changes can be better explained by the dependence structure of their extreme movements.

The predictive power of sentiment changes is larger for future earnings changes than future stock returns and tends to be concentrated in future medium-term earnings changes. Also, when the entire period is divided into the low earnings period and the high earnings period by the break date identified in real earnings time series, the predictive power of sentiment changes decreases in the low earnings period and increases in the high earnings period. Our study provides a new insight to stock market participants.

References

- Amemiya T (1980) Selection of regressors. *Int Econ Rev* 21:331–354
- Antweiler W, Frank M (2004) Is all that talk just noise? The information content of internet stock message boards. *J Finance* 59:1259–1295
- Assaf A (2006) Canadian REITs and stock prices: fractional cointegration and long memory. *Rev Pac Basin Financ Mark Polic* 9:441–462
- Baker M, Wurgler J (2006) Investor sentiment and the cross-section of stock returns. *J Finance* 61:1645–1680
- Barberis N, Schleifer A, Vishny R (1998) A model of investor sentiment. *J Financ Econ* 49:307–343
- Brown G, Cliff M (2004) Investor sentiment and the near-term stock market. *J Empir Financ* 11:1–27
- Chang T, Nieh C (2001) International transmission of stock price movements among Taiwan and its trading partners: Hong Kong, Japan and the United States. *Rev Pac Basin Financ Mark Polic* 4:379–402
- Chen C, Kuo I (2014) Investor sentiment and interest rate volatility smile: evidence from Eurodollar options markets. *Rev Quant Financ Acc* 43:367–391
- Das S, Chen M (2007) Yahoo! For Amazon: sentiment extraction from small talk on the web. *Manag Sci* 53:1375–1388
- Debondt W, Thaler R (1985) Does the stock market overreact? *J Finance* 40:793–805
- Engle R, Granger C (1987) Co-integration and error correction: representation, estimation, and testing. *Econometrica* 55:251–276
- Fisher K, Statman M (2003) Consumer confidence and stock returns. *J Portf Manag* Fall 30:115–127
- Fisher K, Statman M (2006) Market timing in regressions and reality. *J Financ Res* 29:293–304
- Gosh A, Clayton R (2006) Debt and equity market reaction to employment reports. *Rev Pac Basin Financ Mark Polic* 9:431–440
- Huth W, Eppright D, Taube P (1994) The indexes of consumer sentiment and confidence: leading or misleading guides to future buyer behavior. *J Bus Res* 29:199–206
- Jansen W, Nahuis N (2003) The stock market and consumer confidence: European evidence. *Econ Lett* 79:89–98
- Klugman S, Panjer H, Willmot G (2008) *Loss models from data to decisions*. Wiley, Hoboken
- Lakonishok J, Schleifer A, Vishny R (1994) Contrarian investment, extrapolation, and risk. *J Finance* 49:1541–1578
- Liao T, Huang C, Wu C (2011) Do fund managers herd to counter investor sentiment? *J Bus Res* 64:207–212
- Otto M (1999) *Consumer sentiment and the stock market*. Board of Governors of the Federal Reserve System, Washington, DC, pp 1–16
- Schmeling M (2009) Investor sentiment and stock returns: some international evidence. *J Empir Financ* 16:394–408
- Shefrin H, Statman M (1994) Behavioral capital asset pricing theory. *J Financ Quant Anal* 29:323–349
- Shiller R <http://www.econ.yale.edu/~shiller/data.htm>
- Spyrou S (2012) Sentiment changes, stock returns and volatility: evidence from NYSE, AMEX and NASDAQ stocks. *Appl Financ Econ* 22:1631–1646

-
- Tumarkin R, Whitelaw R (2001) News or noise? Internet message board activity and stock prices. *Finance Anal J* 57:41–51
- Wang Y, Keswani A, Taylor S (2006) The relationships between sentiment, returns, and volatility. *Int J Forecast* 22:109–123
- Wurgler J <http://people.stern.nyu.edu/jwurgler>