Marczak, Martyna; Beissinger, Thomas

Conference Paper
Bidirectional Relationship between Investor Sentiment and Excess Returns: New Evidence from the Wavelet Perspective


Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association


This Version is available at:
http://hdl.handle.net/10419/145836

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Bidirectional Relationship between Investor Sentiment and Excess Returns: New Evidence from the Wavelet Perspective

Martyna Marczak* Thomas Beissinger†
University of Hohenheim, Germany University of Hohenheim and IZA, Germany

Abstract

We propose to use the wavelet concept of the phase angle to determine the lead–lag relationship between investor sentiment and excess returns that are related to the bubble component of stock prices. The wavelet phase angle allows for decoupling short– and long–run relations and is additionally capable of identifying time–varying comovement patterns. Based on the monthly S&P500 index and two alternative monthly US sentiment indicators we find that in the short run (until 3 months) sentiment is leading returns whereas for periods above 3 months the opposite can be observed. Moreover, the initially strong positive relationship becomes less pronounced with increasing time horizon, thereby indicating that the over– or undervaluation in the short run is gradually corrected in the long run.

JEL Classification: G11, G14, C22, C32

Keywords: wavelet phase angle, wavelet analysis, sentiment indicator, excess returns, speculative bubble, stock market

*Corresponding author: University of Hohenheim, Department of Economics, Schloss, Museumsfluegel, D-70593 Stuttgart, Germany, e-mail: marczak@uni-hohenheim.de
†University of Hohenheim, Department of Economics, Schloss, Museumsfluegel, D-70593 Stuttgart, Germany, e-mail: beissinger@uni-hohenheim.de
1 Introduction

The last decade saw a huge increase in the number of studies dealing with the impact of investor sentiment on stock prices. Many studies simply consider a linear regression of future stock returns on an indicator of investor sentiment, see Bathia and Bredin (2013) and the references therein. However, such an approach implies a unidirectional causality running from sentiment to stock returns. To check for a potential influence of returns on sentiment other studies estimate a VAR model and/or perform Granger causality tests, see, for example, Brown and Cliff (2004) and Kim and Kim (2014); for a non-linear causality framework see Dergiades (2012). However, so far there is scant evidence on whether the lead–lag relationship between sentiment and stock returns may change over time or exhibit specific patterns over the business cycle, the exceptions being the studies by Li (2015) and Lutz (2015).

Complementary to the existing literature, we suggest to resort to wavelet analysis, and more specifically, to the wavelet concept of the phase angle, for a more detailed picture on the lead–lag relationship between sentiment and returns. To the best of our knowledge, this is the first paper employing wavelet analysis to this research question. Wavelet analysis distinguishes between different horizons at which the comovements are measured and thus allows to derive conclusions about the short–run and long–run relationship between stock returns and sentiment. The distinctive feature of wavelet analysis is that it allows the relationship between different periodic components of two time series to be variable in time. This makes it possible to capture changes in behavioral patterns or to uncover asymmetric effects of investor sentiment in different periods, like stock market expansions and contractions. We demonstrate the usefulness of the wavelet phase angle by applying it to S&P500 returns and two measures of US investor sentiment for the period from 1970.M1 to 2014.M9.

Our study is in line with two basic premises from behavioral finance. First, investors are subject to sentiment, i.e. stock market expectations about future earnings and investment risk that are unrelated to fundamentals. Second, there are limits to arbitrage, because betting against investors guided by sentiment is costly and risky (Baker and Wurgler, 2007). From these premises follows that stock returns not only reflect changes in fundamental values but also consist of an “excess returns” component, i.e. a “bubble premium” that is related to changes in investor sentiment. Therefore, it makes sense to use excess returns instead of total returns in our analysis of sentiment effects on the stock market.\footnote{The focus on excess returns is in line with Brown and Cliff (2005), who also relate investor sentiment to some indicator of stock market mispricing.} In this paper, excess returns are defined as the deviations of total returns...
from their fundamental part derived from the well-known static Gordon model (Gordon, 1962). Investor sentiment is by its very nature unobservable and difficult to measure. We extract two indicators for investor sentiment from a set of 9 “direct” sentiment proxies and technical indicators that have been suggested in, e.g., Brown and Cliff (2004) and Baker and Wurgler (2007) using two alternative approaches, principal component analysis and a simple factor model.

2 Excess Returns and Investor Sentiment

We first set out the procedures to obtain the excess return component based on the S&P500 index and two sentiment indexes. The generated data are given on a monthly frequency in the time span 1970.M1 – 2014.M9.

To calculate excess returns that are caused by deviations of stock prices from their fundamental values, the stock price index \( P_t \) must be decomposed into the fundamental price \( P_t^f \) and the bubble component \( P_t^b \). The fundamental price is related to the future stream of dividends and is determined in this paper using the well-known static Gordon model (Gordon, 1962), according to which the fundamental price of an asset is given by:

\[
P_t^f = \frac{1 + g_t^e}{r_t^e - g_t^e} Y_t,
\]

where \( Y_t \) denotes dividends, \( g_t^e \) is the expected growth rate of dividends, and \( r_t^e \) is the expected rate of return. We compute \( g_t^e \) as the 10-year moving average of dividend growth rates. To obtain \( r_t^e \) we refer to a simple CAPM, according to which

\[
r_t^e = \bar{r}_t + \beta R P_t,
\]

where \( \bar{r}_t \) is the risk-free rate of return approximated in this paper by the Moody’s 30–year BAA corporate bond yield. \( R P_t \) is the market risk premium calculated here by the 10–year moving averages of the difference \( (r_t^m - \bar{r}_t) \), with \( r_t^m \) being the market rate of return. Assuming that the S&P500 covers the market portfolio, \( \beta \) is equal to one and \( r_t^m \) corresponds to the actual return \( r_t = (P_t + Y_t - P_{t-1})/P_{t-1} \). All variables are expressed in real terms by deflating nominal values with the consumer price index (CPI).\(^2\) Once \( P_t^f \)

and $P_t^b = P_t - P_t^f$ are obtained, total returns can be decomposed into two parts:

$$r_t = \frac{P_{t-1}^f}{P_{t-1}} r_t^f + \frac{P_{t-1}^b}{P_{t-1}} r_t^b$$

The second component will be referred to as excess returns and will be used in the subsequent wavelet analysis.

In the literature, various approaches have been proposed to quantify investor sentiment. Some studies employ data on “direct” sentiment measures based on investor surveys like the American Association of Individual Investors (AAII) survey or the Investor Intelligence (II) survey; see, e.g., Brown and Cliff (2004). Other studies proxy investor sentiment by, among others, a consumer confidence index (e.g. Lemmon and Portnaiaguina, 2006), various measures reflecting investor mood (e.g. Hirshleifer and Shumway, 2003; Edmans et al., 2007; Tetlock, 2007), and stock market related measures like market liquidity (Baker and Stein, 2004) and closed–end fund discount (Neal and Wheatley, 1998).

A well–known and often used composite sentiment index is provided by Baker and Wurgler (2006).\(^3\) Unfortunately, this index has not been updated after 2010. In order to take account of recent developments in financial markets, we therefore had to construct our own composite sentiment index. For that purpose, we combine “direct” sentiment proxies based on surveys with technical indicators. As for the former, we use the bull–bear spread (BBS) computed with the data from the II survey, and the consumer confidence index (CCI) provided by the Conference Board. Technical indicators can be classified into different categories. The first one represents market breadth and the corresponding variable is the so–called Arms index (ARMS):

$$\text{ARMS} = \frac{\text{ADV}/\text{ADV VOL}}{\text{DECL}/\text{DECL VOL}},$$

where ADV and DECL give the number of advancing and declining issues on the NYSE, respectively, whereas ADV VOL and DECL VOL refer to the cumulative number of issues from the group advancing and declining issues within a given time period. The variables capturing trading activity are the percentage changes in NYSE short interest and in NYSE real margin debt. The next indicator describes market volatility and is given by the ratio of implied volatility VIX (CBOE Volatility Index for S&P500) and realized volatility (RV). The latter is computed with the extreme–value method proposed by Parkinson (1992). Finally, the remaining three indicators are mutual fund flows (MFF) provided by the

\(^3\)This index can be downloaded from [http://people.stern.nyu.edu/jwurgler/](http://people.stern.nyu.edu/jwurgler/).
Investment Company Institute, IPO number and IPO first–day returns. A justification for using these variables in the construction of a composite sentiment index can, for example, be found in Brown and Cliff (2004) and Baker and Wurgler (2006, 2007). The final dataset consists of 9 sentiment series and is characterized by a ragged–edge structure as not all series are available in the entire time span.4

Based on these sentiment series we construct composite sentiment indexes using two alternative approaches: principal components analysis and a simple factor model. These approaches have been commonly used in the construction of sentiment measures; see, e.g., Brown and Cliff (2004) and Baker and Wurgler (2007). Prior to index extraction all data have been standardized.

From the principal component analysis we obtain a sentiment indicator, denoted SENTPC, as the first principal component of a restricted dataset including BBS, CCI, ARMS, percentage change in NYSE real margin debt, IPO number and IPO first–day returns. The remaining 3 sentiment proxies not observable in the entire time span are excluded in the construction of SENTPC.

An alternative sentiment indicator, denoted SENTFM, is derived as the common factor component, \( z_t \), in the following factor model framework:

\[
\begin{align*}
y_t &= \mu + \theta z_t + u_t, \quad u_t \sim NID(0, \Sigma_u) \\
z_{t+1} &= \phi z_t + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma^2)
\end{align*}
\]

where \( y_t \) denotes the vector of 9 sentiment proxies, \( \mu \) is the vector of intercepts and \( u_t \) is the vector of idiosyncratic components with diagonal covariance matrix \( \Sigma_u \). The common factor component follows an AR(1) process, and its contribution to the observed series is expressed by the vector of factor loadings \( \theta \). It is assumed that \( \varepsilon_t \) and \( u_t \) are mutually uncorrelated. The model parameters are estimated by maximum likelihood, and \( z_t \) is extracted by the application of the Kalman filter and smoother. These algorithms are capable of handling missing values and ragged–edge data, and thus allow for using the complete set of 9 sentiment proxies.

It can be argued that sentiment is to some extent also driven by rational factors and can thus incorporate a fundamental part. To remove this part, we regress SENTPC and

---

SENTFM, respectively, on three monthly macroeconomic variables capturing business cycle effects: the growth rate of the industrial production index (IPI), the unemployment rate and the Purchasing Managers Index (PMI). The adjusted versions of SENTPC and SENTFM are nearly coincident with the original ones.

Figure 1 depicts both sentiment indexes along with excess returns. It is evident that both SENTPC and SENTFM quite reasonably reproduce bullish and bearish phases on the stock market. However, they differ from each other with regard to the extent of the oscillations.

Figure 1: Sentiment indexes obtained with principal component analysis (SENTPC) and a factor model (SENTFM), respectively depicted with excess returns

3 Wavelet Phase Angle

To uncover the lead–lag relationship between sentiment and returns, we propose to use the concept of the wavelet phase angle. An advantage of this concept compared to its frequency–domain counterpart is that it carries information about the relationship of the considered variables both in time and frequency. This is because wavelet functions are local in the time and frequency domain so that the resulting wavelet transform of a time series gives its two–dimensional representation. In contrast, sine and cosine functions used in the Fourier transform provide a one–dimensional representation of a series only.

\(^5\)Data on the IPI, the unemployment rate and the PMI have been downloaded from [http://research.stlouisfed.org/fred2/](http://research.stlouisfed.org/fred2/)
The wavelet phase angle between two series $y_t$ and $x_t$ is defined as:

$$
\phi_{xy}(\tau, s) = \arctan \left[ \frac{\Im(W_{xy}(\tau, s))}{\Re(W_{xy}(\tau, s))} \right],
$$

(2)

where $\tau$ and $s$ are time and scale parameter, respectively. Scale $s$ is inversely related to the angular frequency $\omega$ and their functional relation depends on the type of wavelet function. In the case of the Morlet wavelet chosen in this paper it holds that $s = 2\pi/\omega$. In eq. (2), $W_{xy}(\tau, s)$ denotes the wavelet cross–spectrum given by $W_x(\tau, s) W_y^*(\tau, s)$, where $W_j(\cdot)$, $j = x, y$, is the continuous wavelet transform of $j$, and $^*$ labels the complex conjugate. $\Im(\cdot)$ and $\Re(\cdot)$ denote the imaginary and real part, respectively. For details concerning the properties of wavelet functions as well as computational aspects the reader is referred to, e.g., Aguiar-Conraria and Soares (2014) and Marczak and Gómez (2015).

The phase angle $\phi_{xy,\psi}(\tau, s)$ is due to the properties of arctangent a multivalued function whose values are given by the respective principal value $\pm n\pi$, where $n = 0, 1, 2...$, and the principal value lies in $(-\pi/2, \pi/2)$. For interpretation purposes, it is though useful to limit values of the phase angle to the interval $[-\pi, \pi]$. A rationale for this restriction is provided by Marczak and Beissinger (2013). Note that $\phi_{xy,\psi}(\tau, s) \equiv \pm \pi/2$ for $\Re(W_{xy,\psi}(\tau, s)) = 0$ and $\Im(W_{xy,\psi}(\tau, s)) \gtrless 0$. Values of the phase angle inform about the in–phase or anti–phase relation between the components of $x_t$ and $y_t$. If $\phi_{xy,\psi}(\tau, s) \in (-\pi/2, \pi/2)$, the respective components are positively related to each other (in–phase movement), whereas in the case of $\phi_{xy,\psi}(\tau, s) \in [-\pi, -\pi/2) \lor (\pi/2, \pi]$ a negative relationship (anti–phase movement) between them is established. If, for given $\tau$ and $s$, it holds that $0 < \phi_{xy,\psi}(\tau, s) < \pi/2$ or $-\pi < \phi_{xy,\psi}(\tau, s) < -\pi/2$, $y_t$ is said to lag $x_t$ at $(\tau, s)$. Values satisfying $\pi/2 < \phi_{xy,\psi}(\tau, s) < \pi$ or $-\pi/2 < \phi_{xy,\psi}(\tau, s) < 0$ imply leading behavior of $y_t$ over $x_t$ at $(\tau, s)$. If $\phi_{xy,\psi}(\tau, s) = 0$, both series are said to be in phase for given $(\tau, s)$.

To reduce the complexity in the interpretation of phase angle values, it is useful to derive the tendency in the relationship between two series in the time and scale dimension. For that purpose, we average phase angle values separately over time and scale by employing the concept of a mean suited for data measured on a circular scale; see, e.g., Zar (1999).

---

6 The computation of $\phi_{xy}(\tau, s)$ is carried out in Matlab using the ASToolbox by Aguiar-Conraria and Soares (2011).
4 Results

Figure 2 depicts the estimated mean phase angle values with their corresponding 95% confidence bounds in the case of SENTPC and SENTFM, respectively. In the right panels of Figure 2, the horizontal axis represents periods computed according to the formula \( p = \frac{2\pi}{\omega} \) which in the case of the chosen Morlet wavelet reduces to \( p = s \). The depicted range of periods between 2 and 36 months is also used to obtain the mean phase angle values in Subfigures 2a and 2c. The lower bound is restricted by the Nyquist frequency whereas the upper bound is set to 3 years so as to capture the long-run relationship between returns and sentiment.\(^7\)

It can be seen that the results are similar for both sentiment indices. In the entire time interval the mean phase angle takes on values between 0 and \( \frac{\pi}{2} \) suggesting that sentiment is positively related to returns and is lagging behind. Even though this pattern seems to be stable over time, until the mid-1970’s and around 2000 the mean phase angle tends towards zero, meaning that the lagging behavior of sentiment is less pronounced in these time intervals. The result that sentiment seems to lag stock returns is in line with corresponding time-domain results by Brown and Cliff (2004) for the US and Kim and Kim (2014) for twenty countries based on Granger causality tests.

However, the overall picture can mask effects attributed to different horizons at which the comovements are measured. The phase angle values averaged over time allow for disentangling the information about the short– and long–run relationship between sentiment and excess returns. In the short run – up to 3 months – sentiment is leading returns, as indicated by values between \(-\pi/2\) and 0. Positive values observed for periods above 4 months suggest that at longer time horizons returns are leading sentiment. Since this pattern dominates across all periods between 2 and 36 months, sentiment is lagging behind in Subfigures 2a and 2c. It is also worth noting that with increasing time horizon the phase angle gradually increases towards the marking line at the value \( \pi/2 \). This means that the relationship between returns and sentiment becomes less positive at longer time horizons, and after around 3 years the phase angle is close to the area that indicates negative relationship between sentiment and excess returns, with sentiment leading. This observation suggests that after the initial market overvaluation (undervaluation) induced by a high (low) sentiment level the market price slowly adjusts towards its intrinsic value. The hypothesis of the reversal of short–term returns in the long–run has has been confirmed by a range of studies; see, e.g., De Bondt and Thaler (1985), Poterba and Summers

\(^7\)The boundary at 3 years represents a compromise between interpretability and accuracy of results. On the one hand, increasing the boundary could contaminate findings with information of long-run lead–lag relation which can hardly exist. On the other hand, a too low upper bound reduces the number of phase angle values involved in the calculation of the mean values.
Figure 2: Mean phase angle between excess returns and two sentiment indicators, SENTPC and SENTFM, respectively; red dots: point estimates, black lines: corresponding 95% confidence bounds

(1988), Cutler et al. (1991), and Brown and Cliff (2005). Our findings additionally show that during this adjustment process it is the returns that drive the sentiment and not vice versa.

5 Conclusions

In this paper we reassess the relationship between stock returns and investor sentiment. Even though this research question has been examined in a large number of studies using time-domain methods, this paper contributes to the literature by proposing the wavelet concept of the phase angle. This method offers two advantages: first, it uncovers the
lead–lag relationship between the interesting variables over different time horizons, and second, it allows for a time–varying pattern in the relationship.

We compute the wavelet phase angle between excess S&P500 returns, i.e. returns obtained from the bubble component of stock prices, and two US sentiment indicators from 1970.M1 to 2014.M9. The analysis yields several important results. First, in the short run (until 3 months) sentiment is leading excess returns and the relationship between them is positive. Second, between 3 and 36 months the lead–lag behavior is reversed but the relationship stays positive. Third, the positive relationship becomes less pronounced with longer time horizons, which can be interpreted as a gradual correction of the initial mispricing and is consistent with the hypothesis of the short–term returns reversal in the long run. Fourth, though the wavelet phase angle would uncover any time–varying comovement patterns, the relationship between investor sentiment and stock returns seems to be quite stable over time.

Our results suggest that the lead–lag relationship between investor sentiment and stock returns is best uncovered using methods, like wavelet analysis, that allow for disentangling short– and long–term effects. Moreover, the results also suggest that those studies that are ignoring the feedback mechanism from stock returns to sentiment may only tell one part of the story.
References


