Physica A 460 (2016) 201–209

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Daily happiness and stock returns: Some international evidence

Wei Zhang [a,b], Xiao Li [a,c], Dehua Shen [a,c,d,*], Andrea Teglio [d]

[a] College of Management and Economics, Tianjin University, Tianjin, China
[b] Key Laboratory of Computation and Analytics of Complex Management Systems, Tianjin, China
[c] China Center for Social Computing and Analytics, Tianjin University, Tianjin, China
[d] Department of Economics, Universitat Jaume I, Castellón, Spain

HIGHLIGHTS

- Happiness sentiment is positively related to index return.
- Happiness sentiment can provide additional explanatory power for index return.
- Happiness sentiment can granger-cause the changes in index return.
- Happiness sentiment influences the index return and intraday volatility.
- Happiness sentiment is related to macroeconomic events.

ARTICLE INFO

Article history:
Received 17 January 2016
Received in revised form 30 March 2016
Available online 14 May 2016

Keywords:
Happiness sentiment
Social media
Range-based volatility
Kendall correlation coefficient
Granger causality

ABSTRACT

In this paper, we examine the relations between the daily happiness sentiment extracted from Twitter and the stock market performance in 11 international stock markets. By partitioning this happiness sentiment into quintiles from the least to the happiest days, we first show that the contemporary correlation coefficients between happiness sentiment and index return in the 4 and most-happiness subgroups are higher than that in least, 2 and 3-happiness subgroups. Secondly, the happiness sentiment can provide additional explanatory power for index return in the most-happiness subgroup. Thirdly, the daily happiness can granger-cause the changes in index return for the majority of stock markets. Fourthly, we find that the index return and the range-based volatility of the most-happiness subgroup are larger than those of other subgroups. These results highlight the important role of social media in stock market.

1. Introduction

Traditional financial theory is based on the assumptions of investors' rationality and homogeneity that leaves no role for irrational behavior in asset pricing [1]. However, the behavioral economics has shifted the academic focus to investigate the relations between investor sentiment and asset prices. Meanwhile, market participants play an active role in generating "fashion" in stock market, which leads to investor's maniac behavior.[1] All the facts recognize the role of sentiment in explaining and predicting the stock returns.

* Corresponding author at: College of Management and Economics, Tianjin University, Tianjin, China.
E-mail address: shen@uji.es (D. Shen).

[1] A notable example is the "DotCom Mania" described by Ref. [2].

http://dx.doi.org/10.1016/j.physa.2016.05.026
0378-4371/

Table 1
Statistical property of raw daily happiness and daily happiness index. This table reports the statistical property of raw daily happiness and the deseasonalized daily happiness index. The raw daily happiness is the original index downloaded from the website (http://hedonometer.org/index.html) and the daily happiness index is residual of the regression model of raw daily happiness on weekday dummies and holiday dummies.

Index	N	Mean	Median	Std.	Kurtosis	Max	Min
Raw daily happiness	2451	6.0092	6.0050	0.0428	6.0469	6.3560	5.8820
Daily happiness index	1751	0.0000	−0.0038	0.0338	0.4023	0.1946	−0.1317

The existence of noise traders and psychological biases are viewed as the assumptions by most theoretical models on investigating the impact of investor sentiment on stock prices. De Long et al. [3] firstly observe that irrational noise traders could not be offset by limited arbitrageurs; rather, with erroneous beliefs and diverse sentiment, they could actually affect the stock prices and earn higher expected returns. Barberis et al. [4] develop a parsimonious model of investor sentiment to partially explain the underreaction and overreaction. Daniel et al. [5] propose a model populated with overconfidence and self-attribution bias and show that both of them can cause the changes in autocorrelations of returns and excess volatility. Hong et al. [6] also develop a model with boundedly rational investors (news watchers and momentum traders) and show their explanatory power for the short-run continuation and long-run reversal of the prices. All these theoretical predictions provide a general framework for analyzing and interpreting the empirical results.

In order to investigate the implications of the above-mentioned theoretical predictions, scholars endeavor to construct various proxies for investor sentiment in their empirical work. In general, there are mainly three categories of proxies to measure investor sentiment according to the source of sentiment from which the proxy is extracted. The first category refers to the stock market-based proxies. This category includes the fluctuations in discounts of closed-end funds [7], the net mutual fund redemptions [8], bid–ask spreads and turnover [9], the combinations of the closed-end fund discount, NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues and the dividend premium [10,11] as well as the portfolio allocations to equity versus cash and fixed-income securities [12]. As for the second category, empiricists use the survey-based proxies such as the Investor Intelligence by American Association of Individual Investors [13–17], Consumer Confidence by the Conference Board (CBIND) and Consumer Confidence by University of Michigan Survey Research Center [18], UBS/Gallup survey [19] and animusX-Investors sentiment [20]. The third category is the proxies extracted from the news and social media content. This category includes sentiment extracted from the Raging Bull [21], Yahoo! Finance [22–24], Wall Street Journal [25], Twitter [26], Google Insights [27], Facebook [28] and Google Trends [29]. These studies usually construct a time series proxy for investor sentiment and examine the contemporaneous and lead-lag relations between the constructed sentiment proxy and market variables (short-term returns, abnormal trading volume and excess volatility) as well as give mixed findings on the predictability and correlation coefficient.

In this paper, we also attempt to examine the relations between the investor sentiment released by social media and the stock market activities. Naturally, it is also connected to plenty of studies on the dynamics between online behaviors and stock market movements [30–44]. The main contribution of our paper compared with others is to exploit a novel sentiment proxy, i.e., the daily happiness index extracted from Twitter, across 11 international stock markets, perform the contemporaneous correlation regression model with the consideration of trading volume as well as provide impact of happiness on index return and range-based volatility with cross-sectional analysis.

The remainder of this paper is organized as follows. Section 2 describes the daily happiness index and capital data. Section 3 performs the contemporaneous correlation, the Granger causality and the cross-sectional analysis between the daily happiness and major stock indexes. Section 4 sets forth the robustness test and Section 5 concludes.

2. Data description

We obtain raw daily happiness from the website (http://hedonometer.org/index.html). The raw daily happiness scores are derived from the natural language processing technique based on a random sampling of about 10% (50 million) of all messages posted in Twitter. To quantify the happiness of the language, the Amazon's Mechanical Turk service is used to score the level of happiness of selected words appeared on Twitter, e.g., laughter, joy, successful, winning, excellent, rainbow. To avoid the spurious happiness events like "Thanksgiving" and "Christmas Eve" and the potential seasonality (the happiness rises during the end of the week), we regress the raw daily happiness on weekday dummies and holiday dummies and keep the residuals as the daily happiness index employed in this paper. The deseasonalized time series allows us to analyze the trend of the daily happiness independently of the predictable seasonal patterns and reveals the intrinsically changes in daily happiness. Table 1 reports the statistical property of the raw daily happiness and the daily happiness index employed in this paper. Fig. 1 illustrates the raw daily happiness and Fig. 2 illustrates the daily happiness index, which spans from September 10th, 2008 to May 27th, 2015. Totally, there are 2451 and 1751 calendar days, respectively. In particular, both KPSS test and Augmented Dickey–Fuller test confirm the stationary property of the daily happiness index. For the empirical analysis in the following sections, the daily happiness index corresponds to the trading days of each stock market.

Our analysis is conducted on 10 developed stock markets in the world, i.e., USA, Canada, France, Germany, UK, Hong Kong, South Korea, Japan, Australia and New Zealand. Almost all of these countries and region have large proportion of Twitter users. The capital market data are from YAHOO! Finance (http://finance.yahoo.com) including the opening price of