Economic Modelling 49 (2015) 162–171

Contents lists available at ScienceDirect

Economic Modelling

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

How does Google search affect trader positions and crude oil prices?

Xin Li [a], Jian Ma [b], Shouyang Wang [c], Xun Zhang [c,*]

[a] School of Management, University of Chinese Academy of Sciences, 80 Zhongguancun East Road, Beijing, China
[b] Department of Information Systems, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong
[c] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 55 Zhongguancun East Road, Haidian District, Beijing 100190, China

ARTICLE INFO

Article history:
Accepted 20 April 2015
Available online 15 May 2015

Keywords:
Internet-based data
Google search volume index
Crude oil price
Granger causality
Trader positions

ABSTRACT

Novel data series constructed from Internet-based platforms such as Google have been widely applied to analyze economic and financial indicators and have been demonstrated to be effective in short-term forecasts. However, few studies have demonstrated the role of Google search data in analyzing trader positions and energy price volatility. This paper uses the Google search volume index (GSVI) to measure investor attention, and investigate the relationships among the GSVI, different trader positions, and crude oil prices from January 2004 to June 2014. The empirical results present some new evidences. First, the GSVI measures investor attention from noncommercial and nonreporting traders, rather than commercial traders. Second, the feedback loop between GSVI and crude oil price is verified. Third, the GSVI improves the forecast accuracy of crude oil price in recursive one-week-ahead forecasts. This paper contributes to existing literature by incorporating open source Internet-based data into the analysis and prediction of crude oil prices, as well as other prices in financial markets in the Big Data Era.

1. Introduction

The prediction of crude oil price has attracted considerable attention from investors because crude oil is one of the most important commodities in the global market. However, it is challenging to accurately predict crude oil price because it is deeply affected by many complicated factors, such as economic growth, inventories, interests rates, and U.S. dollar exchange rates etc. (Alquist and Kilian, 2010; Benhmad, 2012; Carfí and Musolino, 2014; Ding et al., 2014; Wang and Chueh, 2013; Yu et al., 2008). Recently, investor attention has become a newly emerging concept that can be considered as an important factor in price fluctuations (Vlastakis and Markellos, 2012). However, it is hard to measure investor attention due to some of its intrinsic features, such as the subjectivity of public concern and difficulty in census data collection. For the past few years, Internet-based data such as the Google Search Volume Index (GSVI) have been widely used to measure investor attention in financial markets, which is relatively objective compared with traditional measures of investor attention (Da et al., 2010; Drake et al., 2012; Vlastakis and Markellos, 2012). Studies have seldom analyzed the roles of the GSVI in influencing trader positions and energy prices. This paper proposes to use the weekly GSVI data to measure investor attention and investigates how the GSVI affects different trader positions and crude oil price volatility.

The penetration of Internet technology has made it possible to capture investor attention via various platforms, such as forums, weblogs, microblogs, and search engines (Agichtein et al., 2008; Asur and Huberman, 2010; Kietzmann et al., 2011). Although a large quantity of data has the potential to be a measure of investor attention, the statistical data from search engines is the most persuasive data for several reasons. First, these data are time series, which are easier to deal with compared to unstructured or semi-structured data in the form of text or images from online platforms. There are many sophisticated analysis techniques that can be applied to manipulate time series data (Box et al., 2008; Brillinger, 1981). Secondly, search is widely recognized as a revealed attention measure, which is new and direct. Da et al. (2011) proved that the GSVI was correlated with but different from existing proxies for investor attention using a sample of Russell 3000 stocks. Actually, search terms such as 'crude oil price' demonstrate that some investors are paying attention. Therefore, search volume data could be a direct measure of investor attention. Google is the most popular search engine for collecting information in the United States. In March 2013, Google accounted for 67.1% of all search queries performed in the United States according to a report issued by comScore Inc.[1] Thus, the statistical data of search behaviors from Google are representative of investor attention. Google Trends[2] is a public tool provided by Google Inc. that shows the search volumes of a particular term from January 2004 at a weekly frequency. Therefore, the search volume data from

* Corresponding author. Tel: +86 1062622016.
E-mail addresses: leexin111@163.com (X. Li), isjian@cityu.edu.hk (J. Ma), sywang@amss.ac.cn (S. Wang), zhangxun@amss.ac.cn (X. Zhang).

[1] comScore Inc., a global leader in measuring the digital world and preferred source of digital business analysis, issues the comScore Explicit Core Search Share Reports every month. The report investigates the use of search engines in total U.S., including home and work locations in March 2013. More detailed information can be obtained at: http://www.comscore.com/Insights/Press_Releases.

[2] The GSVI data series can be downloaded from http://www.google.com/trends/.

http://dx.doi.org/10.1016/j.econmod.2015.04.005
0264-9993/

Google Trends denoted as GSVI is utilized in this paper to directly measure investor attention.

The existing literature has demonstrated that the GSVI can represent economic and financial markets trends (Bank et al., 2011; Da et al., 2011; Saiz and Simonsohn, 2013). Specifically, Guo and Ji (2013) used the GSVI to represent public concerns regarding oil prices, oil demand, financial crises, and the effect of Libyan war on oil markets. Tushar and Saket (2013) applied the GSVI and Twitter sentiment to predict oil, gold, and market indices, and proved that the forecasting performance of the GSVI was superior. To the best of our knowledge, this paper is the first to use the GSVI to investigate the relationship among investor attention, three types of investors (noncommercial, commercial, and non-reporting traders), and crude oil price volatility. The present study leverages the GSVI series to measure investor attention and examines its relationship with traders in the crude oil futures market. Does the GSVI reflect the attention of all investors, or only some of them? How does the investor attention measured by GSVI affect crude oil prices? It is important and innovative to investigate the impacts of online user generated data on crude oil markets from an econometric modeling perspective in the Big Data Era.

In commodity futures markets, the main trading participants are identified as noncommercial, commercial, and non-reporting traders according to their motivations by the Commodity Futures Trading Commission (CFTC)[3]. The commercial traders, including dealers, producers and manufacturers, etc., engage in risk hedging, whereas the non-commercial traders, such as brokers, trades and hedge funds, are mostly speculators(Büyüksahin and Harris, 2011; Sanders et al., 2004; Zhang and Wang, 2009). Non-reporting traders are usually small speculators. The GSVI is generated by Internet users, who include the public, researchers, investors, etc. In financial studies, the GSVI has proved to be an indicator that captures retail investor attention. Furthermore, speculators such as noncommercial and non-reporting traders are more likely to apply the Google search engine for crude oil information searches compared to commercial traders. It is feasible to extrapolate that the GSVI measures some speculators' attention rather than all investors in the crude oil market. Because investors may search for relevant information before they make decisions about buying or selling their trading positions, search data are likely a direct and early signal reflecting their position changes. Therefore, it is expected that the GSVI series drives the changes of some investors' trading positions.

Although there is no consensus on the relationship between speculative trading and oil prices, some researchers have insisted that speculators' trading positions have affected crude oil price dynamics (Cifarelli and Paladino, 2010; Zhang and Wang, 2009). In this research, we agree with their conclusions and will not discuss the relationship between trader positions and price volatility. In view of the aforementioned analysis, high investor attention that is measured by the GSVI can lead to proactive operations of trading positions, which can further affect the volatility of the crude oil markets. In addition, great fluctuations in the crude oil market result in market anomalies, which contain extremely valuable opportunities to speculate. The extreme changes of crude oil prices can draw great attention from market participants, especially the speculators. Therefore, it is possible to assume that there is a feedback loop between investor attention and crude oil prices. Because investor attention influences crude oil prices, it should improve the forecasting accuracy of crude oil prices.

In this empirical study, multiple data sources from the Google Trends, Energy Information Administration (EIA) website and Commitments of Traders (COT) reports by CFTC, are used to examine the relationship among investor attention, trading positions, and crude oil prices from January 2004 to June 2014. First, this research individually tests the causal relationship between the GSVI and trading positions of noncommercial, commercial and non-reporting traders, to verify that the GSVI reflects some investors' attention rather than all investors' attention. Second, we conduct groups of Granger causality tests to depict the potential feedback between investor attention and crude oil prices. Third, we explore the impacts of investor attention on crude oil prices from the forecasting perspective. A recursive out-of-sample forecast is conducted to evaluate the performance of the GSVI. This research contributes to the empirical literature by introducing Internet-based data to analyze investors' trading activities and commodity prices. A new measure of investor attention is constructed by using the GSVI, and then using this new measure to help forecast crude oil prices. It is demonstrated that the GSVI captures the speculators' attention instead of the attention from commercial traders. In addition, the bi-directional causality between investor attention and crude oil price is investigated. Investor attention measured by the GSVI can improve the crude oil price forecasting accuracy.

The remainder of the paper proceeds as follows. The research background is described in Section 2. Data description is presented in Section 3. The empirical results and analysis are illustrated in Section 4. Finally, a discussion and conclusions are presented in Section 5.

2. Research background

Existing research has examined the applications of the GSVI in the analysis and forecast of economic and financial indicators. In the research field of economic forecasting, Askitas and Zimmermann (2009) applied econometric models to forecast the German unemployment using the GSVI under complex and fast changing conditions. Swallow and Labbé (2013) tested whether the GSVI improved the fit and efficiency of nowcasting models for Chile automobile sales. Humphrey (2010) employed the GSVI to predict national and local existing home sales, and found that it improved the forecasting accuracy. Choi and Varian (2012) used the GSVI to forecast many economic indicators, such as automobile sales, unemployment claims, travel destination planning and consumer confidence. Vosen and Schmidt (2012) introduced a new monthly indicator using the GSVI for private consumption in Germany, and found that the new indicator outperformed the survey-based indicators. (Toth and Hajdu, 2012) used the GSVI to nowcast Hungarian household consumption and proved its predictive power in Hungary. Guzman (2011) incorporated the GSVI to predict inflation expectation.

In the field of financial analysis, the GSVI was considered as a favorable indicator of investor attention. Da et al. (2010) first proposed a direct measure of investor attention on the stock market through Google Trends, which was correlated but different from the existing proxies of the investor attention, such as turnover and extreme returns. They found that the new measure captured the retail investors' attention. Their work was the first to illustrate the usefulness of the GSVI in financial applications. Thereafter, Vlastakis and Markellos (2012) employed the GSVI to measure information demand for the largest 30 stocks traded on the New York Stock Exchange, and used the Reuters NewsScope Archive as the information supply. Their research concluded that investors needed more information when their level of risk aversion increased. Joseph et al. (2011) used the GSVI as the proxy for investor sentiment of retail investors, and revealed that GSVI reliably predicted abnormal stock returns and trading volumes. Andrei and Hasler (2011); Kita and Wang (2012) used the GSVI to measure the investor attention on the stocks and analyzed its relationship with volatility in the stock and foreign exchange markets. Gao et al. (2011) studied the association between information acquisition that is measured by the GSVI and daily trading activities. They confirmed a positive volume–search association, which indicated that the GSVI was associated with a 9% increase in the trading volume. Other research including Bank et al. (2011), Drake et al. (2012), Jacobs and Weber (2012) and Zhang et al. (2013) used the GSVI to predict revenue

[3] The commercial traders use futures contracts for hedging as defined in CFTC Regulation 1.3(z). The detailed information could be seen in the website http://www.cftc.gov/MarketReports/CommitmentsofTraders/ExplanatoryNotes/index.htm.