Journal of Economic Behavior & Organization 117 (2015) 62–81

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

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

Forecasting volatility with empirical similarity and Google Trends☆

Alain Hamid, Moritz Heiden*

Department of Statistics, University of Augsburg, Universitaetstr. 16, D-86159 Augsburg, Germany

ARTICLE INFO

Article history:
Received 22 July 2014
Received in revised form 16 April 2015
Accepted 5 June 2015
Available online 19 June 2015

Keywords:
Empirical similarity
Google trends
Investor attention
Volatility
Forecasting

ABSTRACT

This paper proposes an empirical similarity approach to forecast weekly volatility by using search engine data as a measure of investors attention to the stock market index. Our model is assumption free with respect to the underlying process of investors attention and significantly outperforms conventional time-series models in an out-of-sample forecasting framework. We find that especially in high-volatility market phases prediction accuracy increases together with investor attention. The practical implications for risk management are highlighted in a Value-at-Risk forecasting exercise, where our model produces significantly more accurate forecasts while requiring less capital due to fewer overpredictions.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

When examining how volatility in stock markets arises, one of the most interesting questions includes the role played by investors behavior, especially the interest they take in the market as well as their uncertainty about its future outcome. Recent research focusses on retail investors attention on the stock market, as studies show that retail investors trades do indeed move markets on short (weekly) horizons (see e.g. Barber et al., 2009) and investors disagreement causes higher trading volume (e.g. Li and Li, 2014). Traditionally, interest in the market is measured by indirect proxies like volume, turnover and news. While volume might be the natural candidate to link investor attention and volatility, several studies, e.g. Brooks (1998) and Donaldson and Kamstra (2005) demonstrate that it does not improve the accuracy of volatility forecasts. News as an alternative measure are mostly irregular and may underly a considerable publication lag. Recent publications use internet message postings (Kim and Kim, 2014), Facebook users sentiment data (Siganos et al., 2014) or search frequencies (Vozlyublennaia, 2014) to assess the influence of retail investors attention on the stock market. Among these studies, Da et al. (2011), Vlastakis and Markellos (2012) and Andrei and Hasler (2013), suggest that Google search volume is a driver of future volatility.

In this paper our scope of work is twofold. Since previous studies, prominently Vlastakis and Markellos (2012) and Vozlyublennaia (2014), have focused on analyzing the in-sample properties of the relationship between investor attention and volatility by using Google search volume, we take the discussion further and concentrate on the predictability in an

☆ The authors gratefully acknowledge the support of the Research Center Global Business Management at Augsburg University, as well as financial support from the German Research Foundation(DFG)-Project "Wishart Processes in Statistics and Econometrics: Theory and Applications". We further thank two anonymous reviewers for their questions and comments. Both authors contributed equally to this work.

* Corresponding author. Tel.: +49 8215984027.
E-mail addresses: alain.hamid@wiwi.uni-augsburg.de (A. Hamid), moritz.heiden@wiwi.uni-augsburg.de (M. Heiden).

http://dx.doi.org/10.1016/j.jebo.2015.06.005
0167-2681/© 2015 Elsevier B.V. All rights reserved.

out-of-sample forecasting framework. We argue, that if investor attention is not only correlated with, but indeed a cause of volatility, this should enable us to form superior predictions from a model that makes use of the dynamics of both time-series. As theoretical investigations and in-sample comparisons in previous research (e.g. Vlastakis and Markellos, 2012; Andrei and Hasler, 2013; Vozlyublennaia, 2014) have found that volatility and investor attention show strong correlation on short horizons, we focus on the weekly horizon, which is the shortest one possible, as Google has restricted the access to (nonstandardized) daily data. We statistically test the ability of Google data to forecast volatility compared to relevant benchmark models using the Model Confidence Set approach of Hansen et al. (2011). We do not find that simply adding a Google component in a regression context results in significant better out-of-sample forecasts compared to conventional models of volatility, even if a cross-correlation and in-sample analysis implies a strong dependence between investor attention and volatility.

This leads us to the question, whether the dynamics of investor attention can be correctly depicted by an additive component in conventional time-series models for volatility. As an alternative, we suggest including Google data in the framework of empirical similarity introduced by Gilboa et al. (2006), which has already been applied in studying behavioral phenomena in portfolio theory by Golosnoy and Okhrin (2008). An adaption of the model, making it suitable for forecasting volatility was proposed in Golosnoy et al. (2014). Lieberman (2012) suggests a simple extension to an AR(1) model, where the autoregressive parameter is determined by empirical similarity. We follow his approach and augment an AR(1) model by a time-varying coefficient determined by the empirical similarity between last periods Google data and volatility. The unique assumption behind the model is, that volatility increases and decreases with investor attention, depending on the previous level of volatility. This approach allows us to study the relationship between investor attention and volatility while being more flexible in the dynamics of the process, as it allows for stationary, non stationary and explosive behavior. Thereby, the model provides a simple, yet flexible framework for forecasting volatility. Comparing and testing predictive accuracy, we find that particularly in crisis phases, our model significantly improves upon standard time-series models with and without additional Google components. This is consistent with the theory of Andrei and Hasler (2013), who state that in "panic states" where volatility is high, investors pay more attention to the market. In an economic application forecasting weekly Value-at-Risk (VaR), we show that more accurate volatility forecasts also lead to improved VaR forecasts. Since VaR exceedances tend to cluster in crisis periods (see e.g. Candelon et al., 2011), our model is beneficial for risk management as forecasting accuracy translates in more precise VaR forecasts as well as less overall capital requirements.

2. Theory and prior literature

Retail investors behavior and their impact on the stock market are well documented in the agent-based literature, e.g. Lux and Marchesi (1999) and Alfarano and Lux (2007), where uninformed investors (noise traders) act as an additional source of volatility. Barber et al. (2009) study the trading behavior of individual investors and confirm that their buying and selling leads to over-, respectively underpricing of the assets on a weekly horizon. A recent study by Li and Li (2014) on household investors suggests, that not all of their trading behavior is unsophisticated or random. In their sample of 30 years of survey data, they find that dispersion of believes about the economic outlook among investors are positively related to stock market trading. However, as the authors point out, they do not test or necessarily assume rationality behind the trades.

Search engine data as a measure of investor attention follows a similar path. The use of e.g. Google to find information on a certain stock does not imply nor deny a rationale behavior, but seems to be strongly linked to stock market participation (see e.g. Preis et al., 2010). Being the most commonly used search engine for collecting information on the internet, accounting for 77.46% [1] of all desktop user search queries worldwide in 2013, Google search frequency data is a regular (daily) and contemporary data source. As Da et al. (2011) point out, Google is likely to be representative of the general internet search behavior, but searching a term is rather a measure for retail investors than professional investors attention.

Google data has already been applied in forecasting flu (Ginsberg et al., 2009), economic indicators (Choi and Varian, 2012) and private consumption (Vosen and Schmidt, 2011). Koop and Onorante (2013) use Google data in a dynamic model selection approach for macroeconomic nowcasting, stressing the fact that including the Google variables in a regression framework might not always be optimal because of the nonlinear dynamics of the attention process. Da et al. (2011) construct a Google search volume Index, showing that it captures the attention of retail investors, while being different from existing proxies for investor attention. They find that the search volume is highly time-varying, rises in periods of high volatility and retail investors are likely to create additional noise in the market. In a study of 30 NYSE and NASDAQ stocks, Vlastakis and Markellos (2012) show, that information demand is positively related to volatility in a GARCH framework. Motivated by these findings Andrei and Hasler (2013) set up a dynamic equilibrium model where variance increases quadratically with investors attention and uncertainty. Especially when they control for lagged volatility, attention is an even more powerful driver of future volatility. This suggests that volatility increases as attention increases, not the other way round. The working mechanism of attention in their model is twofold. Higher attention in bad times accelerates the incorporation of news into prices and increases volatility. Due to higher attention and elevated information demand, investors are more informed, resulting in lower uncertainty, which finally decreases the volatility. Despite these findings on the relationship between volatility and investor attention, the before mentioned articles do not test out-of-sample predictability. Partly, this may be

[1] As reported by http://www.netmarketshare.com.

Download English Version:

https://daneshyari.com/en/article/7243038

Download Persian Version:

https://daneshyari.com/article/7243038

Daneshyari.com