Information demand and stock return predictability

Dimitris K. Chronopoulos [a,*], Fotios I. Papadimitriou [b,c], Nikolaos Vlastakis [d]

[a] School of Management, University of St Andrews, St Andrews, KY16 9RJ Scotland, UK
[b] Southampton Business School, University of Southampton, Southampton SO17 1BJ, UK
[c] Business School, University of Aberdeen, Aberdeen AB24 3FX, UK
[d] Essex Business School, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK

ARTICLE INFO

ABSTRACT

Recent theoretical work suggests that signs of asset returns are predictable given that their volatilities are. This paper investigates this conjecture using information demand, approximated by the daily internet search volume index (SVI) from Google. Our results reveal that incorporating the SVI variable in various GARCH family models significantly improves volatility forecasts. Moreover, we demonstrate that the sign of stock returns is predictable contrary to the levels, where predictability has proven elusive in the US context. Finally, we provide novel evidence on the economic value of sign predictability and show that investors can form profitable investment strategies using the SVI.

© 2017 Published by Elsevier Ltd.

1. Introduction

The question of whether stock market returns contain a predictable component has attracted much attention from both academics and market participants. As a result, numerous financial and macroeconomic variables have been employed to address the issue (see, *inter alia*, Rozeff, 1984; Campbell and Shiller, 1988; Fama and French, 1988; Hodrick, 1992; Lamont, 1998; Pontiff and Schall, 1998; Lettau and Ludvigson, 2001; Boudoukh et al., 2007; Kellard et al., 2010; Andriosopoulos et al., 2014; Jordan et al., 2014) while a plethora of approaches and testing procedures have emerged in the literature making the overall assessment an extremely difficult task (e.g., Goetzmann and Jorion, 1993; Nelson and Kim, 1993; Ferson et al., 2003; Inoue and Kilian, 2004; Amihud and Hurvich, 2004; Lewellen, 2004).[1]

However, although the empirical evidence on return predictability is mixed when we consider stock returns in the levels, there is a smaller body of literature which suggests that it is more likely to predict the signs of stock returns instead (see Breen et al., 1989; Pesaran and Timmermann, 2000, 1995; Christoffersen et al., 2007; Nyberg, 2011; Chevapatrakul, 2013). An important development in this area is the recent theoretical work by Christoffersen and Diebold (2006) which suggests that the sign of stock returns is predictable given that their volatilities are.[2] The intuition behind this relationship is that

---

* Corresponding author.
 E-mail addresses: dc45@st-andrews.ac.uk (D.K. Chronopoulos), F.Papadimitriou@soton.ac.uk (F.I. Papadimitriou), nvlast@essex.ac.uk (N. Vlastakis).
 [1] For a comprehensive and up-to-date overview of stock return predictability, see Rapach and Zhou (2013).
 [2] A voluminous literature on volatility predictability exists which is not reviewed here. For the importance of this topic and for detailed evidence, see Schwert and Seguin (1990), Pagan and Schwert (1990), Hsieh (1991), Brooks (1998), and Poon and Granger (2003, 2005).

any changes in volatility will alter the probability of observing negative or positive returns. More specifically, a rise in volatility increases the probability of a negative return, given that the expected returns are positive. Christoffersen and Diebold (2006) point out that sign predictability may exist even if there is no mean predictability, while it is also independent of the shape of the return distribution. Such sign predictability can be particularly useful for creating profitable investment strategies.

Building on the work of Christoffersen and Diebold (2006), which provides a more concrete framework for sign predictability, we empirically investigate whether improved volatility forecasts, via a measure for information demand, can indeed lead to better sign forecasts of stock returns. Given that stock return predictability in the levels has been proven difficult to detect in the US context (see, Bossaerts and Hillion, 1999; Goyal and Welch, 2003, 2008; Kostakis et al., 2015), it is of particular interest to investigate whether the sign of US stock returns can be predicted instead. Additionally, we investigate whether such sign return predictions can be exploited to form profitable trading strategies from the perspective of investors. This is important for two reasons. First, statistical significance does not necessarily translate into economic gains for investors (Leitch and Tanner, 1991; Andriosopoulos et al., 2014) and second, apart from the theoretical contribution of Christoffersen and Diebold (2006), there is no empirical evidence regarding the economic value of sign predictability based on their model.

To achieve our goal, we turn to the link between return volatility and information flow. As suggested by Ross (1989), the return volatility is directly related to the rate of flow of information to the market. Hence, information about an event occurring (i.e. a signal) will affect the volatility of stock returns and more so if the informational content of this event is low (Gerlach, 2005; Li, 2005). In this case, the demand for information will increase up to the point that the prevailing ambiguity about the state of the world is resolved (Moscarini and Smith, 2002). In other words, the demand for information about an event is directly related to the "quality" of a signal, which is in turn related to return volatility. In a recent study, Vlastakis and Markellos (2012) construct a proxy for demand for information based on the internet "Search Volume Index" (SVI henceforth) from Google. Their analysis reveals a positive and significant contemporaneous relationship between the SVI and historical (and implied) measures of volatility in the US market. Vlastakis and Markellos (2012) also argue that in the networked world information discovery is shared between traditional information providers (such as those employed by sophisticated investors, e.g. Reuters) and alternative channels on the internet. It follows that information arrives on average at roughly the same time for both institutional and retail investors. Therefore, SVI captures the reaction of retail investors to the same signals that institutional investors observe and react to. Consequently, we expect that the use of the SVI as a measure for demand for information can lead to better volatility forecasts. To this end, we augment a number of GARCH family models with the daily SVI to construct better volatility forecasts. Subsequently, we obtain sign forecasts using the model of Christoffersen and Diebold (2006) which we compare against forecasts from various other competing models. Moreover, based on the produced sign forecasts we formulate appropriate trading strategies and examine whether a real-time investor would be able to exploit any sign predictability. Specifically, we adopt the framework of Granger and Pesaran (2000) and consider an investor who, at each period, has the option to invest her wealth either in stocks or in bonds. Therefore, in a dynamic setting, the investor maximizes her utility function using a time-varying optimal predictive rule to translate sign probability forecasts into investment decisions. Our empirical application uses daily data on the Standard & Poor's 500 (S&P 500) index and spans the period from January 2, 2004 through December 31, 2016.

Our contribution to the literature is threefold. First, we show that the most prominent candidate to predict the sign of daily stock returns is the model suggested by Christoffersen and Diebold (2006) when extended with the SVI variable. Specifically, we find that it outperforms all considered types of competing models either in their standard form or extended with a number of variables which have been shown to work well in previous studies on asset pricing and asset return sign forecasts. For instance, based on the Brier score which is a commonly used scoring rule to evaluate probability forecasts, it produces on average 2.6% more accurate return sign forecasts compared to a naive model which only considers the proportion of positive past returns. This finding highlights the usefulness of the demand for information measure (SVI) in predicting the sign of US stock returns. This is a novel finding in the return predictability literature which is robust to different sample periods (both normal and volatile) and further confirmed by the Diebold and Mariano (1995) test statistic of equal predictive accuracy.

Second, our study offers evidence which suggests that the model by Christoffersen and Diebold (2006) exhibits a good performance in terms of economic value. Moreover, we show that the inclusion of the SVI variable can enhance this model and lead to even higher economic gains for investors. This is the first study to confirm the efficacy of the aforementioned model for real-time investors and to also suggest a new variable that further improves its performance. For example, we find that an active trading strategy that utilises the SVI variable within the Christoffersen and Diebold (2006) framework produces a higher return and a lower standard deviation, which results in a higher Sharpe ratio compared to a simple buy and hold strategy. Our results in this context are robust to different utility functions, levels of risk aversion and estimates of transaction costs. Furthermore, our conclusions remain unaffected under different realistic scenarios (with or without short-sales) and measures of economic value. Third, we demonstrate that the inclusion of the SVI variable in all considered GARCH family models leads to superior stock market volatility forecasts. This finding extends an emerging literature which establishes a relationship between measures of demand for information and asset return volatility (e.g., Vlastakis and Markellos, 2012; Andrei and Hasler, 2015; Vozlyublennaia, 2014; Da et al., 2015; Dimpfl and Jank, 2016; Goddard et al., 2015).

Overall, this study answers calls for research on formulating strategies and assessing their profitability based on sign forecasts (Christoffersen and Diebold, 2006) while it complements the literature by providing new empirical evidence on how investors' attention and reaction to new information can enhance the predictability of asset return signs.

The paper is organised as follows. Section 2 provides a description of our data and the variables used in our study. Section 3 presents the methodological approach and Section 4 discusses the empirical findings. Finally, Section 5 concludes.

2. Data

This paper employs data at a daily frequency covering the period from January 2, 2004 to December 31, 2016. Our sample size is determined by the data availability of our main predictive variable of interest, namely the demand for information which is discussed below. In more detail, we use daily prices of the S&P 500 index obtained from the Center for Research in Security Prices (CRSP). The monthly Treasury bill rate, which proxies for the risk-free rate, is from Ibbotson and Associates and it is available from Kenneth French's website. As in Pesaran and Timmermann (1995) and Chevapatrakul (2013), the excess returns on the stock market, R_t, are computed as $R_t = ((P_t - P_{t-1})/P_{t-1}) - r_t^f$, where P_t is the closing price of the S&P 500 index and r_t^f is the return from holding a one-month Treasury bill over the period from $t-1$ to t.

To capture the demand for information we follow Vlastakis and Markellos (2012) and use a proxy for information flow based on the number of searches on Google. Specifically, we use the SVI which represents the total searches for any keyword on Google in the form of an index, so that the observation with the highest number of searches in a given sample takes the value of 100.

Our main forecasting variable is the change in the SVI, defined as $\Delta SVI_t = SVI_t - SVI_{t-1}$. Given that our goal is to predict the direction of change of the S&P 500 index, the term used as a search keyword is "s&p 500". The selection of this term was decided through a procedure similar to the one followed by Vlastakis and Markellos (2012) and entails inserting the full name of the stock index and checking all similar terms suggested by Google Trends as well as Wordtracker. We finally opted for the term that was the least ambiguous and had the highest search volume.[3] Google offers data on the SVI from January 2004 onwards, thus leaving us with 3263 daily observations. A description of the process we follow to obtain a consistent daily time-series of SVI is provided in Appendix A.

In order to assess the accuracy of volatility forecasts we need to compare them against the true volatility. However, volatility is a latent variable which cannot be directly observed and hence a proxy for it is required. For this purpose we use realised volatility (RV_t) based on 5 min returns with subsampling obtained from the Realized Library of the Oxford-Man Institute.

Moreover, a number of other variables are relevant to the study. Specifically, we use two different bond yields, namely, the short-term bond yield (r_t^s) and the long-term bond yield (r_t^l). The former is measured by the yield of the 3-month Treasury bill, while the latter is measured by the yield of the 10-year Treasury bond. Finally, we also use a bond yield spread, namely, the slope of the Treasury yield curve ($TERM_t$), which is defined as the difference between the yields on the 10-year and the 3-month Treasury securities. The data on the bond yields are obtained from the St. Louis FED's FRED database.

Table 1 presents descriptive statistics of all variables considered in our study. As reflected by the high levels of skewness and kurtosis, all variables display a significant departure from normality. However, as shown in Christoffersen and Diebold (2006), the theoretical result that volatility forecastability implies sign forecastability holds even if returns do not follow a Gaussian distribution. Additionally, as mentioned in Christoffersen et al. (2007), if the distribution is asymmetric then the sign can be predictable even if the expected return is zero. Therefore, there is no concern regarding whether the empirical results in our paper can be affected by the evidence of non-normality in the return series. Turning to ΔSVI, we find that it is positively skewed with excess probability in the tails, as suggested by the higher moments. Moreover, it has a negative autocorrelation of -0.33. Finally, the stationarity of ΔSVI is assessed using three unit root tests: the Augmented Dickey-Fuller test (ADF, Dickey and Fuller, 1979), the Phillips-Perron test (PP, Phillips and Perron, 1988), and the Kwiatkowski, Phillips, Schmidt, and Shin test (KPSS, Kwiatkowski et al., 1992). Whereas for the ADF and PP tests the null hypothesis is the existence of a unit root in the data, the null hypothesis for the KPSS test is stationarity. The corresponding statistics are, respectively, -30.811, -392.260 and 0.4384 and suggest that the ΔSVI is stationary.

For completeness, Fig.1 shows the graph of the SVI variable. As can be seen from the graph, demand for information abruptly increases in October and November 2008 after Lehman's bankruptcy. Furthermore, demand for information also spikes on August 8, 2011 when US and global stock markets crashed following the credit rating downgrade of the US sovereign debt and on June 24, 2016 when the UK voted in favour of exiting the EU. Finally, it reaches a record level on November 9, 2016 when the new US president was elected. Overall, the graph indicates that SVI is associated with demand for information related to the market.

3. Methodology

3.1. The framework of return sign predictability

Christoffersen and Diebold (2006) demonstrate that the sign of asset returns can be predictable even if the expected returns are not, provided that their volatility is. To see this, let R_t denote a series of excess returns, which follows a distribution that depends only on its first two moments, and let Ω_t be the information set available at time t. Moreover, the con-

---

[3] See also Vlastakis and Markellos (2012) for a more detailed discussion of the keyword selection procedure.

Table 1
Descriptive statistics.

	Mean	Standard deviation	Median	Skewness	Kurtosis	Autocorrelation
R_t	0.0003	0.0120	0.0007	−0.0924	14.9055	−0.10
ΔSVI_t	0.0009	7.4178	0.0000	3.3672	191.4147	−0.33
RV_t	0.0001	0.0003	0.0000	11.0943	208.7226	0.68
r_t^S	0.0001	0.0001	0.0000	1.2398	3.0157	1.00
r_t^L	3.1967	1.0744	3.1400	0.1127	1.6701	0.99
$TERM_t$	1.9495	1.0834	2.0800	−0.5519	2.5956	0.99

This table presents summary statistics on all variables considered in this study. R_t denotes excess returns; ΔSVI_t is the change in information demand; RV_t is the realised volatility; r_t^S is the yield on the 3-month Treasury bill; r_t^L the yield on the 10-year Treasury bond; $TERM_t$ is the difference between r_t^L and r_t^S.
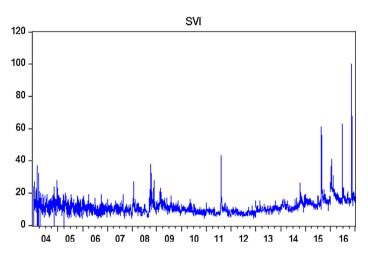

Fig. 1. Time series graph of the information demand variable (SVI). This figure depicts the demand for information variable for the full sample period (i.e. January 2, 2004 to December 31, 2016).

ditional mean and variance of the expected excess returns are denoted, respectively, by $\mu_{t+1|t} = \mathrm{E}(R_{t+1}|\Omega_t)$ and $\sigma_{t+1|t}^2 = Var(R_{t+1}|\Omega_t)$. If $\mu_{t+1|t}$ varies with Ω_t, we say that the return series displays conditional mean dependence and hence it is predictable. Similarly, if $\sigma_{t+1|t}^2$ varies with Ω_t, we can define conditional variance dependence (or volatility predictability). If the probability of a positive return at time t also varies with Ω_t, then the sign of the return series is predictable. In this case, the conditional probability of a positive excess return is given by:

$$
\begin{aligned}
Pr(R_{t+1} > 0|\Omega_t) &= 1 - Pr(R_{t+1} \leqslant 0|\Omega_t) \\
&= 1 - Pr\left(\frac{R_{t+1} - \mu_{t+1|t}}{\sigma_{t+1}} \leqslant -\frac{\mu_{t+1|t}}{\sigma_{t+1}}\right) \\
&= 1 - F\left(-\frac{\mu_{t+1|t}}{\sigma_{t+1}}\right),
\end{aligned}
\tag{1}
$$

where F is the cumulative distribution function of the standardized return $(R_{t+1} - \mu_{t+1|t})/\sigma_{t+1}$. Provided that the conditional volatility is predictable, the sign of the return will also be predictable even if the expected excess return is constant over time but different from zero (i.e. $\mu_{t+1|t} = \mu$ and $\mu \neq 0$). This is because, as Eq. (1) suggests, the probability of a positive return is time varying as volatility moves.[4]

Building on the above framework, and with the aim to provide fresh empirical evidence, in this paper we forecast the probability of observing a positive excess return on the S&P 500 index using forecasts of its expected excess returns as well as its volatility. The methodological approach for obtaining such forecasts is discussed below.

3.2. Volatility forecasting models

As suggested by the work of Christoffersen and Diebold (2006), forecasting the daily stock market return volatility plays a key part in predicting accurately the sign of stock market returns. Therefore, we turn to the GARCH family models that have been shown to perform well in forecasting conditional stock return volatility (Poon and Granger, 2003). These models

---

[4] As mentioned earlier, even if $\mu_{t+1|t} = 0$ the sign of the excess return will continue to be predictable provided that F is asymmetric.

involve the joint estimation of a conditional mean and a conditional variance equation. The return process is assumed to be generated as:

$$R_t = a + bR_{t-1} + \varepsilon_t,$$ (2)

where R_t is the excess return and $\varepsilon_t \sim N(0, \sigma_t^2)$ is the error term. The GARCH (1, 1) process introduced by Bollerslev (1986) and Taylor (1986) is given by:

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2,$$ (3)

where $\alpha, \beta \geqslant 0$; $\omega > 0$ ensures the positiveness of the conditional variance. For the unconditional variance to exist, a necessary and sufficient condition is that $\alpha + \beta < 1$.

As stressed earlier, one of the main objectives of our study is to examine whether investors' demand for information can improve volatility forecasts. Hence, we extend the standard GARCH (1, 1) model to a GARCH-SVI (1, 1) model by incorporating the Google SVI index both in the mean and in the variance equations as follows:

$$R_t = a + bR_{t-1} + c\Delta SVI_{t-1} + \varepsilon_t,$$ (4)

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + \delta\Delta SVI_{t-1},$$ (5)

where ΔSVI is the change in the demand for information.

We also consider popular extensions of the basic GARCH model. One such extension is the GJR due to Glosten et al. (1993) which accommodates the asymmetry in the response of volatility to positive and negative shocks. The conditional variance in the GJR-GARCH model has the following form:

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I_{t-1},$$ (6)

where $I_{t-1} = 1$ if $\varepsilon_{t-1} < 0$ and $I_{t-1} = 0$ if $\varepsilon_{t-1} \geqslant 0$.

Similarly to the GARCH model, we also extend the GJR-GARCH to include the ΔSVI variable as follows:

$$R_t = a + bR_{t-1} + c\Delta SVI_{t-1} + \varepsilon_t,$$ (7)

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I_{t-1} + \delta\Delta SVI_{t-1}.$$ (8)

In addition, we employ the EGARCH model proposed by Nelson (1991) which, unlike the GARCH (1, 1) model, does not require imposing non-negativity constraints. Moreover, similar to the GJR-GARCH, the standard EGARCH model does not enforce a symmetric response of volatility to positive and negative shocks. The EGARCH (1, 1) specification can be expressed as follows:

$$ln(\sigma_t^2) = \omega + \beta ln(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[\frac{|\varepsilon_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right].$$ (9)

As with the other volatility models, we extend the EGARCH (1, 1) specification by including the demand for information proxy variable. The relevant EGARCH-SVI (1, 1) model is then written as:

$$R_t = a + bR_{t-1} + c\Delta SVI_{t-1} + \varepsilon_t,$$ (10)

$$ln(\sigma_t^2) = \omega + \beta ln(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[\frac{|\varepsilon_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right] + \delta\Delta SVI_{t-1}.$$ (11)

To obtain one-step-ahead forecasts of the conditional volatility, we rely on a recursive scheme. Let $T = L + P$ denote the full sample size where L is the number of in-sample observations and P is the number of out-of-sample forecasts. For all models, we initially use the first L observations to get the corresponding set of parameter estimates. The first one-step-ahead forecast error is then constructed by comparing our forecasted volatility against the realised volatility, which serves as a proxy of the latent return volatility. In order to construct the second one-step-ahead forecast error, we use data available through period $L + 1$ and repeat the above procedure. This process continues until all observations are used and we have a sequence of P one-step-ahead forecast errors.

We report the statistics of the out-of-sample forecast errors obtained in different sample periods. In particular we document the mean, standard deviation, and root mean square error (RMSE) of volatility forecast errors resulting from each competing model presented above. The parameters are estimated by Quasi Maximum Likelihood using a Gaussian Likelihood.

3.3. Assessing volatility forecasts

One issue when comparing volatility forecasts is that the model with the smallest forecast error is not necessarily superior to the other competing models. Hence, we need to formally examine whether the estimated RMSEs are significantly different from one another in a statistical sense. For this purpose, we employ the Diebold and Mariano (1995) (DM) statistic which tests for equal predictive accuracy. In particular, Diebold and Mariano (1995) suggest the use of a *t*-statistic for pairwise model comparisons in terms of forecasting performance. Their approach involves taking the loss differential between two competing models scaled by a robust estimator of its variance and comparing it against the standard normal distribution.

When comparing forecasts between non-nested models, such as between the GARCH model against the EGARCH model, the DM statistic has a standard normal distribution (see West, 1996). However, when comparing forecasts from nested models McCracken (2007) shows that the DM statistic has a non-standard limiting distribution and provides asymptotically valid critical values for various combinations of in-sample and out-of-sample proportions. For instance, in our study this case applies when we compare forecasts between a standard form GARCH family model and its extended counterpart (e.g., EGARCH and EGARCH-SVI).

3.4. Return sign forecasting models

To estimate the probability of a positive return on the S&P 500 index we adopt the framework by Christoffersen and Diebold (2006) (C&D) and Christoffersen et al. (2007) and employ a model which allows for the forecasted probability to be conditioned on the predictive variable of interest through volatility forecasts. Based on the relevant model, shown in Eq. (12) below, the $t + 1$ period forecast is generated as:

$$\hat{\pi}_{t+1}^{C\&D} = \widehat{Pr}(R_{t+1|t} > 0) = 1 - \frac{1}{t}\sum_{k=1}^{t} I\left(\frac{R_k - \hat{\mu}_k}{\hat{\sigma}_k} \leqslant \frac{\hat{\mu}_{t+1|t}}{\hat{\sigma}_{t+1|t}}\right), \tag{12}$$

where $I()$ is the indicator function, $\hat{\sigma}_{t+1|t}$ is the one-step-ahead volatility forecast as described above, and $\hat{\mu}_{t+1|t}$ is the one-step-ahead expected return forecast. The $\hat{\mu}_{t+1|t}$ is estimated as an AR(1) process. As explained earlier, in this paper we use the change in the demand for information to enhance the volatility forecasts of the S&P 500 index returns. Therefore, we extend the C&D model by incorporating the ΔSVI variable both in $\hat{\sigma}_{t+1|t}$, as described in Section 3.2, and in $\hat{\mu}_{t+1|t}$.

The probability forecasts from our main model (C&D-SVI) are then compared against those generated by a naive model as well as by a number of alternative models which have been previously used to predict the sign of stock returns. In particular, one-step-ahead probability forecasts from the naive model are generated as follows:

$$\hat{\pi}_{t+1}^{Naive} = \widehat{Pr}(R_{t+1|t} > 0) = \frac{1}{t}\sum_{k=1}^{t} I(R_k > 0). \tag{13}$$

That is, our naive model is based on the empirical distribution function of the return. In other words, we consider the proportion of positive past returns from the beginning of the sample period up to the time the forecast is made. To further assess the robustness of our main model, we also employ some alternative competing models which have been used in the extant literature to construct probability forecasts. Specifically, we consider a dynamic probit model (see, Nyberg, 2011) which uses a binary sign return indicator as the dependent variable and also accounts for potential autocorrelation in the structure of the return signs time series. Based on this model, probability forecasts are constructed as shown in Eq. (14):

$$\hat{\pi}_{t+1}^{DynProbit} = Pr(R_{t+1|t} > 0) = \Phi(\beta_0 + \beta_1 I(R_t > 0)). \tag{14}$$

Moreover, similar to Nyberg (2011) and Chevapatrakul (2013), we consider a number of alternative specifications of the dynamic probit model in Eq. (14) which control for various factors that have been used to predict return signs. Specifically, we employ three models which extend the dynamic probit model by adding, respectively, the short term yield, r_t^S, the long term yield, r_t^L, and the term spread, $TERM_t$, as in Eq. (15) below:

$$\hat{\pi}_{t+1}^{DynProbit-Z} = Pr(R_{t+1|t} > 0) = \Phi(\beta_0 + \beta_1 I(R_t > 0) + \beta_2 Z), \tag{15}$$

where Z can be one of the variables, r_t^S, r_t^L or $TERM_t$, as discussed above. Finally, we also consider a dynamic probit model (DynProbit-All) which incorporates all three variables mentioned above.

To compare the predictive performance of each model we report statistics on the out-of-sample prediction errors obtained in different sample periods. In particular, we document the mean, standard deviation and Brier scores of the probabilistic forecast errors derived from each model. The Brier score is a commonly used scoring rule to evaluate probability forecasts (Brier, 1950). It may be viewed as the mean squared loss function for probabilistic forecasts. The Brier score for a sample of *n* binary forecasts, where the event either occurs or does not occur, is given by $Brier = \frac{1}{n}\sum_{t=1}^{n}(\hat{\pi}_t - o_t)^2$, where $\hat{\pi}_t$ is the predicted probability of the event occurring according to the *t*th forecast, and o_t is equal to one or zero, depending

on whether the event actually occurs or not. To assess whether any identified differences between the competing models are also statistically significant we employ the DM statistic as in the case of comparing volatility forecasts (see Section 3.3).

4. Empirical results

4.1. Volatility forecasting results

Table 2 presents the estimates of the different standard GARCH model specifications employed in our study as well as their extended with the ΔSVI counterparts. These in-sample results indicate that the coefficient of ΔSVI is positive and statistically significant across all different GARCH models.[5] This finding suggests that ΔSVI possesses significant in-sample predictive power for return volatility and enriches standard GARCH models with useful information about market variation. Moreover, the coefficient γ is also found to be significant in both the GJR-GARCH and EGARCH models suggesting the existence of asymmetry in the conditional return distribution. As suggested by Christoffersen et al. (2007) such evidence of asymmetry implies that the sign of returns could be predictable provided that their volatility is. Overall, these results extend an emerging literature which establishes a relationship between measures of demand for information and asset return volatility (e.g., Vlastakis and Markellos, 2012; Andrei and Hasler, 2015; Vozlyublennaia, 2014; Da et al., 2015; Dimpfl and Jank, 2016; Goddard et al., 2015).

Panel A in Table 3 tabulates the forecast error statistics obtained from the volatility models described in Section 3.2 which are employed to produce one-step-ahead out-of-sample daily forecasts of the S&P 500 return volatility. To allow for a more thorough examination of forecasting performance, the computed statistics are also reported for two different sub-sample periods. The first sub-sample spans the period from January 3, 2005 to December 31, 2007 whereas the second one covers the period from January 2, 2008 to December 31, 2016 which includes the recent financial crisis.

Panel A reveals that the inclusion of the ΔSVI variable in all considered volatility models improves their out-of-sample forecasting ability as indicated by the reported RMSEs. This result is robust not only across all model specifications but also in all sample periods under consideration. This is a new finding in the literature which indicates that the change in demand for information conveys additional useful information and enhances daily forecasts of stock return volatility.[6]

On the other hand, if we compare volatility forecasts derived from the GARCH model against those derived from the asymmetric GARCH family models, we observe that the latter exhibit a superior forecasting performance. Specifically, the EGARCH-SVI model exhibits the best forecasting performance, as suggested by the smallest RMSE, followed by the standard EGARCH model. These results are consistent with the study by Awartani and Corradi (2005) who also report that asymmetric GARCH models lead to better volatility forecasts relative to their symmetric counterparts.

The identified RMSE differences discussed above do not necessarily suggest that the competing models produce forecasts which are also different in a statistical sense. Therefore, we need to formally examine whether the inclusion of the ΔSVI variable in our volatility models can indeed lead to significantly better forecasts. To this end, we conduct a test of equal predictive accuracy. As such, Panel B in Table 3 tabulates the computed DM statistics when we compare each volatility model against its extended with the ΔSVI variable counterpart across different periods. The DM statistics suggest that the inclusion of the ΔSVI variable significantly improves the predictive power of all GARCH family models at the 1% level during the full out-of-sample period. This result is robust across all considered periods and establishes that the ΔSVI contains valuable information for predicting stock market volatility.

4.2. Sign predictability results

Given that the EGARCH-SVI and the EGARCH models produce the best volatility forecasts among all competing models under consideration, we now further explore whether we can exploit these forecasts to achieve better sign predictability of daily stock returns as suggested by Christoffersen and Diebold (2006).[7] Using both of these volatility models will also allow us to empirically assess (i) how the C&D model compares against other types of models that are relevant for sign predictability and (ii) whether the extension we propose in this paper with the ΔSVI variable can enhance the sign forecasting power of the C&D model. Panel A in Table 4 shows the forecast error statistics derived from all models described in Section 3.4 which are used for return sign predictability. As in the case of volatility forecasts, these statistics are also reported for two different sub-sample periods (one of which includes the recent financial crisis).

---

[5] As a robustness check, we also estimated the considered GARCH models by additionally including either the VIX^2, the realised volatility or the trading volume. In all cases, ΔSVI remains significant and its coefficient maintains the correct sign. To account for potential multicollinearity problems, we follow Cooper and Priestley (2009) and Andriosopoulos et al. (2014) and we look at the relative performance of ΔSVI and the VIX^2 (Realised Volatility or Trading Volume), when the latter is orthogonalised relative to the former.

[6] Although this paper is theoretically motivated by the link between SVI and stock market volatility, we have also extended the GARCH family models presented in Section 3.2 with the r_t^S, r_t^L, and $TERM_t$ variables. However, these extended models do not generally outperform their standard form counterparts in terms of volatility forecasting power and additionally, none of these models (either standard or extended) outperform the EGARCH-SVI model or the simple EGARCH model. Therefore, based on the theory put forward by Christoffersen and Diebold (2006), we consider only these two volatility models when constructing return sign forecasts.

[7] We have also extended the EGARCH model with either the VIX^2, the realised volatility or the trading volume but none of these alternative model specifications produces a lower RMSE when predicting volatility.

Table 2

In-sample performance of volatility forecasting models.

	ω	α	γ	β	δ
GARCH	0.0001***	0.1054***		0.8707***	
GARCH-SVI	0.0001***	0.0909***		0.8853***	0.0001**
GJR-GARCH	0.0001***	−0.0262***	0.2021***	0.9014***	
GJR-GARCH-SVI	0.0001***	0.1485***	0.0496	0.5894***	0.0001***
EGARCH	−0.3164***	0.1282***	−0.1503***	0.9768***	
EGARCH-SVI	−0.2053***	0.1013***	−0.1187***	0.9865***	0.0449***

All volatility forecasting models are presented in Section 3.2.
* Indicates significance at the 10% level.
** Indicates significance at the 5% level.
*** Indicates significance at the 1% level.

Table 3

Out-of-sample performance of the volatility models.

	GARCH	GARCH-SVI	EGARCH	EGARCH-SVI	GJR-GARCH	GJR-GARCH-SVI
Panel A: Forecast error statistics						
03/Jan/2005–02/Dec/2016						
Mean	0.0293	0.0290	0.0098	0.0092	0.0225	0.0235
Standard deviation	0.2233	0.2226	0.1849	0.1827	0.1983	0.1961
Root mean square error	0.2252	0.2244	0.1851	**0.1829**	0.1995	0.1975
03/Jan/2005–31/Dec/2007						
Mean	0.0037	0.0037	0.0003	-0.0008	0.0039	0.0038
Standard deviation	0.0493	0.0492	0.0479	0.0476	0.0444	0.0437
Root mean square error	0.0494	0.0493	0.0479	0.0476	0.0446	**0.0438**
02/Jan/2008–31/Dec/2016						
Mean	0.0379	0.0375	0.0130	0.0125	0.0287	0.0300
Standard deviation	0.2558	0.2550	0.2116	0.2091	0.2272	0.2247
Root mean square error	0.2585	0.2576	0.2120	**0.2094**	0.2290	0.2266
	GARCH vs GARCH-SVI		EGARCH vs EGARCH-SVI		GJR-GARCH vs GJR-GARCH-SVI	
Panel B: Diebold and Mariano (1995) statistics						
03/Jan/2005–31/Dec/2016	4.53***		2.92***		1.22***	
03/Jan/2005–31/Dec/2007	2.50***		3.41***		2.84***	
02/Jan/2008–31/Dec/2016	5.22***		3.34***		1.38***	

Panel A presents the properties of the volatility prediction errors obtained from the competing GARCH family models presented in Section 3.2. The full out-of-sample period spans January 3, 2005–December 31, 2016. For a more in-depth evaluation, results are also reported for two different sub-sample periods, one of which spans the pre-crisis period and one that includes the recent financial crisis. All models use available data starting from January 2, 2004 to produce one-step-ahead volatility forecasts. Boldface indicates superior performance (i.e. more accurate forecasts). Numbers in Panel A are of order 10^{-3}. Panel B presents the computed Diebold and Mariano (1995) test statistics across different periods. These statistics are employed to test whether the reported RMSE performances between the different GARCH family models and their extended with ΔSVI counterparts reported in Panel A, are statistically different from one another. A positive value indicates that the GARCH family model extended with the ΔSVI performs better than its standard counterpart.
* Indicates significance at the 10% level.
** Indicates significance at the 5% level.
*** Indicates significance at the 1% level.

Based on the produced Brier scores, the results suggest that the model by Christoffersen and Diebold (2006) extended with the ΔSVI variable is the most prominent candidate for predicting the sign of daily stock returns. Specifically, it outperforms all other competing models, either in their standard form or enhanced with a number of variables which have been shown to be relevant for asset return and sign forecasts. This finding is consistent across all periods and is further confirmed by means of the DM statistic (see Panel B of Table 4). It is worth mentioning here that none of the extended dynamic probit models considered in the analysis manages to outperform the baseline dynamic probit model (DynProbit). This finding is broadly in contrast to recent evidence which suggests that variables such as the short term interest rate and the long term interest rate perform well when monthly data are considered (Nyberg, 2011; Chevapatrakul, 2013). Hence, our findings point to the direction that the ΔSVI seems to convey more useful information for predicting the sign of returns at higher frequencies.

To shed more light on our findings, we further explore whether this sign predictability is driven by predictability in the mean or in the variance. To this end, we estimate Eq. (4) and we find that the lagged ΔSVI is not a significant in-sample predictor of stock returns at the level. This result indicates that predictability in the variance drives sign predictability of returns in our sample. This is consistent with the theory of Christoffersen and Diebold (2006) which suggests that sign predictability can exist independently of mean predictability.

Table 4
Out-of-sample performance of the sign forecasting models.

	C&D-SVI	C&D	DynProbit	DynProbit-r^S	DynProbit-r^L	DynProbit-r^{Term}	DynProbit-All	Naive
Panel A: Forecast error statistics								
03/Jan/2005–31/Dec/2016								
Mean	0.4702	0.4771	0.4931	0.4937	0.4945	0.4935	0.4946	0.4958
Standard deviation	0.4414	0.4217	0.0657	0.0652	0.0574	0.0654	0.0670	0.0506
Brier score	**0.6857**	0.6907	0.7022	0.7026	0.7032	0.7025	0.7033	0.7041
03/Jan/2005–31/Dec/2007								
Mean	0.4550	0.4676	0.4894	0.4902	0.4909	0.4899	0.4913	0.4943
Standard deviation	0.4001	0.3182	0.0732	0.0777	0.0676	0.0767	0.0863	0.0566
Brier score	**0.6745**	0.6838	0.6996	0.7002	0.7006	0.6999	0.7009	0.7030
02/Jan/2008–31/Dec/2016								
Mean	0.4752	0.4802	0.4943	0.4948	0.4957	0.4947	0.4957	0.4963
Standard deviation	0.4543	0.4510	0.0630	0.0605	0.0535	0.0611	0.0591	0.0485
Brier score	**0.6894**	0.6930	0.7031	0.7034	0.7040	0.7033	0.7040	0.7045
		C&D vs C&D-SVI	DynProbit vs C&D-SVI	DynProbit-r^S vs C&D-SVI	DynProbit-r^L vs C&D-SVI	DynProbit-r^{Term} vs C&D-SVI	DynProbit-All vs C&D-SVI	Naive vs C&D-SVI
Panel B: Diebold and Mariano (1995) statistics								
03/Jan/2005–31/Dec/2016		2.12[***]	3.17[***]	3.25[***]	3.41[***]	3.32[***]	3.45[***]	3.34[***]
03/Jan/2005–31/Dec/2007		3.43[***]	4.76[***]	4.93[***]	4.86[***]	4.89[***]	5.08[***]	4.80[***]
02/Jan/2008–31/Dec/2016		1.52[***]	2.54[**]	2.60[**]	2.84[***]	2.69[**]	2.81[***]	2.30[**]

In this table, Panel A presents the forecast error statistics from all competing sign forecasting models which are described in Section 3.4. The full out-of-sample period spans January 3, 2005–December 31, 2016. For a more in-depth evaluation, results are also reported for two different sub-sample periods, one of which spans the pre-crisis period and one that includes the recent financial crisis. All models use available data starting from January 2, 2004 to produce one-step-ahead sign forecasts. Boldface indicates superior performance (i.e. more accurate forecasts). Panel B presents the computed Diebold and Mariano (1995) test statistics across different periods. These statistics are employed to test whether the Brier scores obtained from the sign forecasting models are statistically different from the Brier score produced by our preferred model (i.e. C&D-SVI). A positive value indicates that the C&D-SVI model exhibits a superior performance.
[*] Indicates significance at the 10% level.
[**] Indicates significance at the 5% level.
[***] Indicates significance at the 1% level.

4.3. Further analysis of sign predictability: Economic significance

Finding statistical significance in terms of predictive ability is neither a necessary nor a sufficient condition for a profitable investment strategy (Leitch and Tanner, 1991; Diebold and Lopez, 1996). Therefore, in this section we explore whether sign forecasts can help improve investor's decisions and lead to a profitable investment strategy. In particular, we consider an investor with a utility function for wealth w defined as $U(w) = -e^{-Aw}$, where A is the investor's degree of risk aversion.[8] As in Goetzmann et al. (2007), Della Corte et al. (2010), and Andriosopoulos et al. (2014), we assume that the risk aversion coefficient is three.[9] The investor has the option to invest either in the S&P 500 index or in a riskless asset (treasury bills) assuming that sales and purchases of stocks and bonds are subject to transaction costs. The investment decision is then based on the probability of observing a positive return on the index in the next period as generated by each of the predictive models discussed in Section 3.4. To provide a more comprehensive examination of economic value we consider both scenarios where short sales are either allowed or prohibited. Therefore, our paper is relevant to investors operating under different constraints.

4.3.1. The framework for measuring economic significance

In order to assess the economic value of return sign forecasts, we adopt the decision-theoretic framework put forward by Granger and Pesaran (2000). Within this framework, a utility maximizing investor considers the utility of both successful and failed predictions of an upward movement in the stock index before taking action. Let $y_{t+1}^d = I(r_{t+1}^{S\&P500} > 0)$ be an indicator of the realised direction of the return on the S&P 500 index and $\hat{y}_{t+1|t}^d$ be the corresponding directional forecast. Then the investor tries to maximize the following utility function:

$$u_t(\hat{y}_{t+1|t}^d, y_{t+1}^d) = \begin{cases} u_{11,t}, & \text{if } y_{t+1}^d = 1 \text{ and } \hat{y}_{t+1|t}^d = 1 \\ u_{00,t}, & \text{if } y_{t+1}^d = 0 \text{ and } \hat{y}_{t+1|t}^d = 0 \\ u_{01,t}, & \text{if } y_{t+1}^d = 0 \text{ and } \hat{y}_{t+1|t}^d = 1 \\ u_{10,t}, & \text{if } y_{t+1}^d = 1 \text{ and } \hat{y}_{t+1|t}^d = 0 \end{cases}, \tag{16}$$

---

[8] We have also used a logarithmic utility function and our findings are qualitatively similar.
[9] We have also considered investors with $A \in \{1, 2, 4, 5\}$ and our conclusions remain robust to different levels of risk aversion.

where $u_{ij,t}$ is the utility when $y_{t+1}^d = i$ and $\hat{y}_{t+1|t}^d = j$ with $i,j \in \{0,1\}$. We also assume that the utility of a correct prediction is always greater than that of a false prediction. That is $u_{11,t} > u_{10,t}$ and $u_{00,t} > u_{01,t}$. Moreover, let $\hat{\pi}_{t+1}$ denote the forecasted probability of an upward movement for the S&P 500 index formed at the beginning of the period t. Under the above set up the expected utility of taking action is given by $u_{11,t}\hat{\pi}_t + u_{01,t}(1 - \hat{\pi}_t)$, while the expected utility of not taking action is given by $u_{10,t}\hat{\pi}_t + u_{00,t}(1 - \hat{\pi}_t)$. Therefore, the maximizing utility investor will make the prediction $\hat{y}_{t+1|t}^d = 1$ if:

$$u_{11,t}\hat{\pi}_t + u_{01,t}(1 - \hat{\pi}_t) > u_{10,t}\hat{\pi}_t + u_{00,t}(1 - \hat{\pi}_t) \tag{17}$$

or

$$\hat{\pi}_t > \frac{(u_{00,t} - u_{01,t})}{(u_{11,t} - u_{01,t}) + (u_{00,t} - u_{10,t})} \equiv \bar{\pi}_t. \tag{18}$$

In our paper, we consider two alternative scenarios. The first scenario is more restrictive and does not allow for short-sales while under the second scenario short-sales are allowed. In the case of the first scenario we have the following:

$$\omega_t = \begin{cases} 1, & \text{if } \hat{y}_{t+1|t}^d = 1 \\ 0, & \text{if } \hat{y}_{t+1|t}^d = 0 \end{cases} \tag{19}$$

where ω_t represents the portfolio weight attributed to the stock market index. The realised return from the active trading strategy is then given by:

$$R_{p,t+1}^{Active} = \omega_t R_{m,t+1} + (1 - \omega_t)R_{f,t+1}, \tag{20}$$

where $R_{m,t+1}$ is the return on the S&P 500 index and $R_{f,t+1}$ denotes the return on the riskless asset. In this setup the investor may decide to fully invest either in treasury bills that yield $R_{f,t+1}$ or in the S&P 500 index, which yields $R_{m,t+1}^R$ in the event of a market rise and $R_{m,t+1}^F$ in the event of a market fall. Following Granger and Pesaran (2000), historical averages of positive and negative index returns of up to time t are used to predict $R_{m,t+1}^R$ and $R_{m,t+1}^F$, respectively. In our analysis we also consider transaction costs for stocks and bonds which are denoted by τ_m and τ_f respectively. The payoff matrix for this scenario is described in Table 5.

Based on Table 5 and assuming the investor is fully invested in the risk-free asset she will switch to the stock index (i.e. $\hat{y}_{t+1|t}^d = 1$) if:

$$\hat{\pi}_t > \frac{-e^{-A(1+R_{f,t+1})} + e^{-A\xi(1+R_{t+1}^F)}}{e^{-A\xi(1+R_{t+1}^F)} - e^{-A\xi(1+R_{t+1}^R)}} \equiv \bar{\pi}_t^{upper}, \tag{21}$$

where $\xi = (1 - \tau_m)(1 - \tau_f)$. Now, if the investor is fully invested in the stock index she will switch to the risk-free asset (i.e. $\hat{y}_{t+1|t}^d = 0$) only if:

$$\hat{\pi}_t < \frac{-e^{-A(1+R_{t+1}^F)} + e^{-A\xi(1+R_{f,t+1})}}{e^{-A(1+R_{t+1}^R)} - e^{-A(1+R_{t+1}^F)}} \equiv \bar{\pi}_t^{lower}. \tag{22}$$

Finally, she will remain inactive if $\bar{\pi}_t^{upper} < \hat{\pi}_t < \bar{\pi}_t^{lower}$.

Regarding the second scenario of our economic analysis, we follow Campbell and Thompson (2008) and Maio (2013) and we extend the strategy presented above to allow for the possibility of short-selling the stock index. Based on this scenario, the investor can exploit further her predictions for stock market movements and potentially achieve higher economic gains. The corresponding weights for this strategy are now given by:

$$\omega_t = \begin{cases} 1.5, & \text{if } \hat{y}_{t+1|t}^d = 1 \\ -1, & \text{if } \hat{y}_{t+1|t}^d = 0 \end{cases} \tag{23}$$

Therefore, the investor will allocate 150% of her wealth to the stock index in case of a positive return prediction, borrowing 50% at the risk-free rate. In case of a negative return prediction she will sell short the stock index at 100% and invest the proceeds in the risk-free asset.[10] The relevant payoff matrix based on these weights is described in Table 6.

As can be seen in Table 6, the investor will go long on the index if:

$$\hat{\pi}_t > \frac{-e^{-A(1-R_{t+1}^F)} + e^{-A\xi(1+1.5R_{t+1}^F - 0.5R_{f,t+1})}}{e^{-A\xi(1+1.5R_{t+1}^R - 0.5R_{f,t+1})} - e^{-A\xi(1+1.5R_{t+1}^R - 0.5R_{f,t+1})} - e^{-A(1-R_{t+1}^F)} + e^{-A(1-R_{t+1}^R)}} \equiv \bar{\pi}_t^{upper}, \tag{24}$$

---

[10] We have also considered alternative weighting schemes within the range of $[-1, 1.5]$ and our results are qualitatively similar.

Table 5
Payoff matrix (no short-sales).

Action taken at time t.	Market at time t + 1:	
	Rises	Falls
Buy the riskless asset:		
Yes	$\xi(1 + R_{f,t+1})$	$\xi(1 + R_{f,t+1})$
No	$1 + R_{m,t+1}^R$	$1 + R_{m,t+1}^F$
Buy the stock index:		
Yes	$\xi(1 + R_{m,t+1}^R)$	$\xi(1 + R_{m,t+1}^F)$
No	$1 + R_{f,t+1}$	$1 + R_{f,t+1}$

This table displays the payoffs to an investor who has the option to invest either in the stock index or in a riskless asset (Treasury bills) in different states of the world (i.e. market rise or fall), given that short sales are not allowed. For a detailed description see Section 4.3.1. $\xi = (1 - \tau_m)(1 - \tau_f)$ where τ_m and τ_f are the transaction costs for stocks and bonds, respectively. R_f, is the risk-free rate of return, while R_m^R, and R_m^F are the rates of return of the stock index corresponding to a market rise or fall, respectively. Following Granger and Pesaran (2000), historical averages of positive and negative index returns are used to predict R_m^R and R_m^F.

Table 6
Payoff matrix (short-sales allowed).

Action taken at time t.	Market at time t + 1:	
	Rises	Falls
Buy the stock index		
Yes	$\xi(1 + 1.5\, R_{m,t+1}^R - 0.5\, R_{f,t+1})$	$\xi(1 + 1.5\, R_{m,t+1}^F - 0.5\, R_{f,t+1})$
No	$1 - R_{m,t+1}^R$	$1 - R_{m,t+1}^F$
Short-sell the stock index		
Yes	$\xi(1 - R_{m,t+1}^R)$	$\xi(1 - R_{m,t+1}^F)$
No	$1 + 1.5\, R_{m,t+1}^R - 0.5\, R_{f,t+1}$	$1 + 1.5\, R_{m,t+1}^F - 0.5\, R_{f,t+1}$

This table displays the payoffs to an investor who has the option to invest either in the stock index or in a riskless asset (Treasury bills) in different states of the world (i.e. market rise or fall), given that short sales are allowed. For a detailed description see Section 4.3.1. $\xi = (1 - \tau_m)(1 - \tau_f)$ where τ_m and τ_f are the transaction costs for stocks and bonds, respectively. R_f, is the risk-free rate of return, while R_m^R, and R_m^F are the rates of return of the stock index corresponding to a market rise or fall, respectively. Following Granger and Pesaran (2000), historical averages of positive and negative index returns are used to predict R_m^R and R_m^F.

whereas she will sell short the index if:

$$\hat{\pi}_t < \frac{-e^{-A\xi(1+1.5R_{t+1}^F - 0.5R_{f,t+1})} - e^{-A\xi(1-R_{t+1}^F)}}{e^{-A\xi(1-R_{t+1}^F)} - e^{-A\xi(1-R_{t+1}^R)} - e^{-A(1+1.5R_{t+1}^R - 0.5R_{f,t+1})} + e^{-A(1+1.5R_{t+1}^R - 0.5R_{f,t+1})}} \equiv \bar{\pi}_t^{lower}. \tag{25}$$

As before, she will remain inactive if $\bar{\pi}_t^{upper} < \hat{\pi}_t < \bar{\pi}_t^{lower}$.

The optimal predictor that maximizes the expected utility of the investor under either scenario is given by:

$$\hat{y}_{t+1|t}^d = (1 - h_{t+1|t} - \lambda_{t+1|t})\hat{y}_t^d + h_{t+1|t}, \tag{26}$$

where $h_{t+1|t} = I(\hat{\pi}_t > \bar{\pi}_t^{upper})$, and $\lambda_{t+1|t} = I(\hat{\pi}_t < \bar{\pi}_t^{lower})$.

Under both scenarios presented above the return of each active strategy is compared against the return of a passive "buy-and-hold" (B&H) strategy which goes 100% long on the stock index throughout the investment horizon. To further evaluate the economic performance of each active strategy and to be comparable with related studies we initially employ the commonly used Sharpe ratio. We then compare all Sharpe ratios against the corresponding Sharpe ratio obtained from the passive B&H strategy. This is important since different strategies might involve different levels of risk. As a further check, we test whether the Sharpe ratio of any active trading strategy is also statistically different from the Sharpe ratio of the passive B&H strategy. To this end, we employ White's (2000) Reality Check which additionally accounts for data mining concerns that can arise when many predictors are considered. The null hypothesis tested is that the maximum Sharpe ratio of all competing active strategies is less than or equal to the Sharpe ratio of the B&H strategy.

However, it is argued that the Sharpe ratio is associated with certain limitations which can lead to misleading inferences when ranking competing investment strategies (e.g., Goetzmann et al. 2007). Specifically, as it is based on the mean-variance

framework its use is appropriate only when returns follow a normal distribution or when the investor's preferences are described by a quadratic utility function (Zakamouline and Koekebakker, 2009; Homm and Pigorsch, 2012). To alleviate this concern in our paper, we additionally employ the economic performance measure (EPM) due to Homm and Pigorsch (2012). The EPM is obtained by dividing the mean of the portfolio return by its economic index of riskiness, which is a measure proposed by Aumann and Seranno (2008). In contrast to the Sharpe ratio, this alternative measure can account not only for the mean and variance of the strategy's returns but also for their higher order moments and hence, it gives a more realistic representation of what investors consider in practice when evaluating investment opportunities (Golec and Tamarkin, 1998; Harvey and Sidique, 2000).

4.4.1. Empirical evidence on the economic significance

This section addresses the important question of whether an active strategy based on each of the sign predicting models discussed in Section 3.4 can lead to economic gains for investors when compared against the passive B&H strategy which serves as a benchmark. For all strategies under consideration we report average portfolio returns, standard deviations, and Sharpe ratios (all annualized). We also report the EPM measure, the Jarque and Bera (1987) test for normality and the White's Reality Check test. All performance measures are computed net of transaction costs and, in line with the statistical analysis presented in the previous sections, across different periods.

The literature considers a number of ways for estimating transaction costs. For instance, Fama and Blume (1966) approximate transaction costs using the floor trader cost, estimated at 5 basis points (bps) for each one-way trade. This approach is also adopted in some recent studies (e.g., Hsu and Kuan, 2005; Bekiros, 2010). Another main approach involves the estimation of bid-ask spread and commission (S&C) costs (see Stoll and Whaley, 1983; Bhardwaj and Brooks, 1992). Within this context we estimate the S&C costs to be 2.43 bps. For the bid-ask spread we use data from an S&P 500 exchange traded fund while the commission cost is obtained from Anand et al. (2012). Finally, we also consider a method which infers transaction costs from stock price behaviour due to Lesmond et al. (1999). This is a more comprehensive measure to estimate the cost of trading as it implicitly includes not only the spread and commission components but also short sale costs, immediacy costs, and some of the price impact costs. Based on this method we estimate the transaction costs to be 2.4 bps.

Table 7 presents the results of economic value under the first scenario where short-sales are not allowed, with respect to the more comprehensive measure of transaction costs due to Lesmond et al. (1999).[11] However, our findings are also robust to the different measures of transaction costs discussed above.[12]

During the full out-of-sample period (January 1, 2005–December 31, 2016), the passive B&H strategy generates an average return of 4.1%, which combined with a standard deviation of 19.5% leads to a Sharpe ratio of 0.317. The active strategy based on the C&D-SVI model clearly outperforms the passive strategy as it yields a Sharpe ratio of 0.324. This finding, which reflects the good predictive performance of the C&D-SVI model in terms of sign predictability, is robust across all considered periods and is the result of both higher average returns and lower volatility.[13] In other words, this active strategy will switch from the stock index into the risk-free asset whenever negative returns are predicted, thus effectively reducing the downside risk of the portfolio. As a further check, Table 7 presents the results of the White's Reality Check procedure which tests the null hypothesis that the best active strategy in terms of Sharpe ratios is not superior to the B&H strategy. Our findings indicate that the C&D-SVI, which is the best strategy in our case, overall leads to statistically better economic gains relative to the passive strategy. Finally, considering the more comprehensive EPM measure due to Homm and Pigorsch (2012) the relative performance of the active strategies remain unaffected, thus strengthening the robustness of our findings.[14]

Turning to the second and less restrictive scenario, where short-sales are allowed, the corresponding results are tabulated in Table 8. Across all periods, we observe that the best strategy is still the one associated with the C&D-SVI model followed by the one based on the C&D model. Interestingly, in this scenario the remaining active strategies not only underperform the B&H strategy as in the first scenario but they consistently lead to economic losses. This is congruent with the notion that good predictive performance in statistical terms (as documented in Section 4.2) does not guarantee economic gains (see, Leitch and Tanner, 1991). Overall, the results based on these alternative models under both scenarios are broadly in contrast to recent evidence by Nyberg (2011) and Chevapatrakul (2013) who report economic gains, albeit without allowing for short-sales. A plausible explanation for this finding could be the use of higher frequency data in our study which are associated with lower expected returns and potentially higher number of trades that accumulate greater transaction costs. Moreover, as in the first scenario (i.e. where short-sales are not allowed), the White's Reality Check suggests that the strategy based on the C&D-SVI model is significantly better than the B&H strategy in terms of Sharpe ratios. Finally, we note that our results remain robust when we employ the EPM measure.

Overall, our paper is the first study to confirm the usefulness of the C&D model for real-time investors and to also suggest a new variable, ΔSVI, that further enhances its performance. This is demonstrated by considering different realistic scenarios

---

[11] We round up the transaction costs to 3 bps and hence we employ a slightly more conservative value of one-way transaction costs. This means that each round-trip (selling/buying the bond and buying/selling the stock) costs 6 bps.

[12] The transaction costs employed in our paper are generally in agreement with those reported in a number of previous studies (e.g., Gencay, 1998; Bajgrowicz and Scaillet, 2012).

[13] To give a better picture of when any positive profit from this strategy would be eliminated, we also estimate its break-even transaction costs. Specifically, we find that these costs are 6.2 bps and 8.5 bps, respectively, for the first and second scenario considered in our study.

[14] Following Homm and Pigorsch (2012), we set the EPM measure equal to zero if the average excess returns of a portfolio are negative.

Table 7
Out-of-sample economic significance (no short-sales).

	C&D-SVI	C&D	DynProbit	DynProbit-r^S	DynProbit-r^L	DynProbit-r^{Term}	DynProbit-All	Naive	B&H	White's Reality Check
03/Jan/2005–31/Dec/2016										
Return	0.0526	0.0447	0.0481	0.0332	0.0106	0.0177	0.0039	0.0123	0.0413	
Standard deviation	0.1594	0.1538	0.1708	0.1458	0.1035	0.1512	0.1048	0.0011	0.1953	
Jarque-Bera test/100	548.25[***]	731.32[***]	213.21[***]	254.76[***]	59.18[***]	198.78[***]	56.22[***]	7.20[***]	165.82[***]	
Sharpe ratio	0.3241	0.2810	0.2887	0.2120	0.0360	0.1100	−0.0252	0.0044	0.3170	0.093
EPM * 100	0.0838	0.0631	0.0659	0.0357	0.0010	0.0096	0	0	0.0795	
03/Jan/2005–31/Dec/2007										
Returns	0.0694	0.0648	0.0660	0.0339	0.0417	0.0331	0.0392	0.0419	0.0605	
Standard deviation	0.1146	0.1073	0.1238	0.1187	0.0858	0.1206	0.1123	0.0004	0.1238	
Jarque-Bera test/100	1.45[***]	2.38[***]	1.51[***]	2.56[***]	13.87[***]	2.13[***]	4.86[***]	0.55[***]	1.52[***]	
Sharpe ratio	0.2892	0.2600	0.2497	−0.0057	0.0417	−0.0104	0.0328	−0.0388	0.2514	0.028
EPM * 100	0.0660	0.0530	0.0490	0	0	0	0.0008	0	0.0500	
02/Jan/2008–31/Dec/2016										
Return	0.0469	0.0380	0.0421	0.0330	0.0002	0.0125	−0.0079	0.0024	0.0405	
Standard deviation	0.1718	0.1664	0.1838	0.1537	0.1080	0.1601	0.1021	0.0003	0.2139	
Jarque-Bera test/100	388.63[***]	501.31[***]	149.13[***]	210.02[***]	41.44[***]	157.63[***]	59.43[***]	83.47[***]	98.76[***]	
Sharpe ratio	0.3502	0.3043	0.3142	0.2696	0.0347	0.1412	−0.0466	0.0422	0.3482	0.108
EPM * 100	0.0906	0.0674	0.0719	0.0578	0.0009	0.0158	0	0	0.0901	

This table presents the results of economic value when short-sales are not allowed. The full out-of-sample period spans January 3, 2005–December 31, 2016. For a more in-depth evaluation, results are also reported for two different sub-sample periods, one of which spans the pre-crisis period and one that includes the recent financial crisis. All models use available data starting from January 2, 2004. Returns, standard deviations and Sharpe ratios are all annualized. EPM is a more comprehensive performance measure due to Homm and Pigorsch (2012) which accounts not only for the mean and variance of the strategy returns but also for their higher order moments. Finally, we report the *p*-values for the White's (2000) Reality Check, which tests the null hypothesis that the maximum Sharpe ratio of all competing active strategies is less than or equal to the Sharpe ratio of the B&H strategy.
[*] Indicates significance at the 10% level.
[**] Indicates significance at the 5% level.
[***] Indicates significance at the 1% level.

and various performance measures. These results are generally in line with the statistical results presented in the previous sections and further establish the usefulness of the ΔSVI variable in return sign predictability. Our findings complement the studies by Christoffersen and Diebold (2006) and Christoffersen et al. (2007) given the lack of empirical evidence regarding the economic efficacy of their suggested model, which is of particular interest to market participants.

5. Conclusion

In this paper we empirically investigate whether a measure for information demand, can lead to better sign forecasts of stock returns. This is of particular interest given that US stock returns have been proven difficult to predict in the levels. Specifically, we augment a number of competing GARCH family models with a measure of information demand (approximated by SVI from Google) and construct one-step-ahead volatility forecasts, which we use within the Christoffersen and Diebold (2006) framework to predict the sign of stock returns. We then focus on the economic significance of sign predictability and formulate appropriate trading strategies from the perspective of a real-time investor. This is a very important dimension of our study given the lack of empirical evidence within this framework and the fact that improved sign predictability in a statistical sense does not necessarily mean higher economic value for investors. In particular, we adopt the approach by Granger and Pesaran (2000) and consider an investor whose goal is to maximize her utility function using a time-varying optimal rule which translates sign probability forecasts into investment decisions. Our empirical application relies on daily data from the S&P 500 index and covers the period from January 2, 2004 through December 31, 2016.

Turning to our results, we show that extending all considered GARCH family models with the change in demand for information variable (ΔSVI) leads to more accurate return volatility forecasts. We find that this improvement in forecasting power is statistically significant and robust across different sample periods (including the recent financial crisis). This result is in line with a small body of relevant work and highlights the usefulness of the ΔSVI in predicting the volatility of returns during both normal and turbulent periods in the stock market.

Moreover, when we employ the forecasted volatilities to predict the sign of daily stock returns we find that the model suggested by Christoffersen and Diebold (2006) extended with the ΔSVI variable exhibits the best performance among all considered models. Specifically, it outperforms all types of models either in their standard form or extended with a number

Table 8
Out-of-sample economic significance (short-sales allowed).

	C&D-SVI	C&D	DynProbit	DynProbit-r^S	DynProbit-r^L	DynProbit-r^{Term}	DynProbit-All	Naive	B&H	White's Reality Check
03/Jan/2005–31/Dec/2016										
Return	0.1330	0.0989	−0.0711	−0.0711	−0.0711	−0.0711	−0.0711	−0.0711	0.0413	
Standard deviation	0.2562	0.2502	0.1953	0.1953	0.1953	0.1953	0.1953	0.1953	0.1953	
Jarque-Bera test/100	258.06***	314.86***	165.59***	165.59***	165.59***	165.59***	165.59***	165.59***	165.82***	
Sharpe ratio	0.5834	0.4934	−0.3169	−0.3169	−0.3169	−0.3169	−0.3169	−0.3169	0.3170	0.057
EPM * 100	0.2737	0.1959	0	0	0	0	0	0	0.0795	
03/Jan/2005–31/Dec/2007										
Returns	0.1826	0.1402	0.0023	0.0023	0.0023	0.0023	0.0023	0.0023	0.0605	
Standard deviation	0.1685	0.1512	0.1238	0.1238	0.1238	0.1238	0.1238	0.1238	0.1238	
Jarque-Bera test/100	1.80***	4.43***	1.53***	1.53***	1.53***	1.53***	1.53***	1.53***	1.53***	
Sharpe ratio	0.8899	0.7010	−0.2517	−0.2517	−0.2517	−0.2517	−0.2517	−0.2517	0.2514	0.010
EPM * 100	0.6260	0.3890	0	0	0	0	0	0	0.0500	
02/Jan/2008–31/Dec/2016										
Return	0.1155	0.0953	−0.0957	−0.0957	−0.0957	−0.0957	−0.0957	−0.0957	0.0405	
Standard deviation	0.2794	0.2754	0.2139	0.2139	0.2139	0.2139	0.2139	0.2139	0.2139	
Jarque-Bera test/100	162.43***	182.90***	98.65***	98.65***	98.65***	98.65***	98.65***	98.65***	98.76***	
Sharpe ratio	0.5342	0.4693	−0.3372	−0.3372	−0.3372	−0.3372	−0.3372	−0.3372	0.3374	0.054
EPM * 100	0.2296	0.1772	0	0	0	0	0	0	0.0901	

This table presents the results of economic value when short-sales are allowed. The full out-of-sample period spans January 3, 2005–December 31, 2016. For a more in-depth evaluation, results are also reported for two different sub-sample periods, one of which spans the pre-crisis period and one that includes the recent financial crisis. All models use available data starting from January 2, 2004. Returns, standard deviations and Sharpe ratios are all annualized. EPM is a more comprehensive performance measure due to Homm and Pigorsch (2012) which accounts not only for the mean and variance of the strategy returns but also for their higher order moments. Finally, we report the *p*-values for the White's (2000) Reality Check, which tests the null hypothesis that the maximum Sharpe ratio of all competing active strategies is less than or equal to the Sharpe ratio of the B&H strategy.
* Indicates significance at the 10% level.
** Indicates significance at the 5% level.
*** Indicates significance at the 1% level.

of variables motivated by previous studies on asset pricing and asset return sign forecasts. This finding holds across all considered periods and is further confirmed by means of the Diebold and Mariano (1995) statistic.

Finally, we find that there is consistency between the statistical results and the results of economic value and show that the model of Christoffersen and Diebold (2006) extended with the ΔSVI variable offers the highest gains for investors. This confirms its usefulness in the context of return sign predictability. These novel findings hold when we consider different utility functions, levels of risk aversion and estimates of transaction costs. Our findings are also robust to different realistic scenarios (with or without short-sales) and measures of economic value, giving further support to the main conclusions of the paper.

Appendix A

The Google Trends database provides data from January 2004 onwards at weekly frequency. However, if the requested period is up to a quarter, the data is available at daily frequency. There is a caveat that comes with this, though, due to the way SVI is calculated. Specifically, the provided index is always scaled to the highest value within the specified sample. This means that if one obtains daily SVI data for a time period spanning more than a quarter by downloading data one quarter at a time, such data are not comparable and a time-series of SVI cannot be constructed in a consistent manner. Others in the literature (e.g., Da et al., 2015) have addressed this issue by calculating changes (and losing one observation at the start of each quarter in the process), which gets rid of any scaling issues. In our paper we adopt an alternative approach to construct a daily time-series of SVI in a consistent manner. This approach exploits a feature of Google Trends that allows the user to compare SVI for the same term between different time periods. The important property of this feature is that, in order to allow for the comparison, all different subsamples are scaled to the single highest value found among them. Therefore, if one compares between two different quarters, the data Google Trends provides will be daily and under the same scaling. Therefore, it becomes an issue of identifying the quarter that contains the day with the highest SVI (through repeated downloads) and then downloading all other quarters in comparison to that one. This allows the construction of a series that is properly scaled at the levels, which preserves all the information contained in the variable, with the added bonus of not losing any observations within the sample.

The algorithm detailing the steps required for constructing a consistent time-series of SVI using Google Trends is as follows. Taking advantage of the fact that Google Trends allows the simultaneous comparison of up to five different time periods at the same time:

1. Download the first five quarters in the sample.
2. Find the quarter that contains the value of 100. Name this quarter Q_0.
3. Download Q_0 together with four other quarters each time until you have downloaded all quarters contained in the time period of interest.
4. If you reach the final quarter in the period of interest and $100 \in Q_0$, then proceed to step 6.
5. If during any stage in the execution of step 3 $100 \notin Q_0$, then this means that one of the other quarters that were downloaded together with Q_0 now contains 100. The quarter that contains 100 now becomes the new Q_0. Repeat steps 3 and 4.
6. Join all downloaded quarters together in one vector, in chronological order, to obtain a consistent time-series of SVI.

References

Amihud, Y., Hurvich, C.M., 2004. Predictive regressions: a reduced-bias estimation method. J. Fin. Quant. Anal. 39 (4), 813–841.
Anand, A., Irvine, P., Puckett, A., Venkataraman, K., 2012. Performance of institutional trading desks: an analysis of persistence of trading costs. Rev. Fin. Stud. 25, 557–598.
Andrei, D., Hasler, M., 2015. Investor attention and stock market volatility. Rev. Fin. Stud. 28 (1), 33–72.
Andriosopoulos, D., Chronopoulos, D.K., Papadimitriou, F.I., 2014. Can the information content of share repurchases improve the accuracy of equity premium predictions? J. Emp. Fin. 26, 96–111.
Aumann, R.J., Seranno, R., 2008. An economic index of riskiness. J. Polit. Econ. 116, 810–836.
Awartani, B.M.A., Corradi, V., 2005. Predicting the volatility of the S&P 500 stock index via GARCH models: the role of asymmetries. Int. J. Forecast. 21, 167–183.
Bajgrowicz, P., Scaillet, O., 2012. Technical trading revisited: False discoveries, persistence tests, and transaction costs. J. Financ. Econ. 106 (3), 473–491.
Bekiros, S.D., 2010. Heterogeneous trading strategies with adaptive fuzzy Actor-Critic reinforcement learning: a behavioural approach. J. Econ. Dynam. Control 34, 1153–1170.
Bhardwaj, R.K., Brooks, L.D., 1992. The January anomaly: effects of low share price, transaction costs, and bid-ask bias. J. Fin. 47, 553–574.
Bollerslev, T., 1986. Generalised autoregressive heteroscedasticity. J. Economet. 31, 307–327.
Bossaerts, P., Hillion, P., 1999. Implementing statistical criteria to select return forecasting models: what do we learn? Rev. Fin. Stud. 12 (2), 405–428.
Boudoukh, J., Michaely, R., Richardson, M., Roberts, M.R., 2007. On the importance of measuring payout yield: implications for empirical asset pricing. J. Fin. 62 (2), 877–915.
Breen, W., Glosten, L.R., Jagannathan, R., 1989. Economic significance of predictable variations in stock index returns. J. Financ. 44 (5), 1177–1189.
Brier, G.W., 1950. Verification of forecasts expressed in terms of probabilities. Mon. Weather Rev. 78, 1–3.
Brooks, C., 1998. Predicting stock index volatility: can market volume help? J. Forecast. 17 (1), 59–80.
Campbell, J.Y., Shiller, R.J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. Rev. Fin. Stud. 1 (3), 195–228.
Campbell, J.Y., Thompson, S.B., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? Rev. Fin. Stud. 21 (4), 1509–1531.
Chevapatrakul, T., 2013. Return sign forecasts based on conditional risk: evidence from the UK stock market index. J. Bank. Fin. 37, 2342–2353.
Christoffersen, P.F., Diebold, F.X., 2006. Financial asset returns, direction-of-change forecasting, and volatility dynamics. Manage. Sci. 52 (8), 1273–1287.
Christoffersen, P.F., Diebold, F.X., Mariano, R., Tay, A., Tse, Y., 2007. Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: international evidence. J. Fin. Forecast. 1, 3–24.
Cooper, I., Priestley, R., 2009. Time-varying risk premiums and the output gap. Rev. Financ. Stud. 22, 2801–2833.
Da, Z., Engelberg, J., Gao, P., 2015. The sum of all FEARS investor sentiment and asset prices. Rev. Fin. Stud. 28, 1–32.
Della Corte, P., Sarno, L., Valente, G., 2010. A century of equity premium predictability and the consumption-wealth ratio: an international perspective. J. Emp. Fin. 17, 313–331.
Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc. 74, 427–431.
Diebold, F.X., Lopez, J.A., 1996. Forecast Evaluation and Combination. NBER Working Paper No. 192.
Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econ. Statist. 13, 253–263.
Dimpfl, T., Jank, S., 2016. Can internet search queries help to predict stock market volatility? Euro. Fin. Manage. 22 (2), 171–192.
Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns. J. Fin. Econ. 22 (1), 3–25.
Fama, E.F., Blume, M.E., 1966. Filter rules and stock market trading. J. Bus. 39, 226–241.
Ferson, W.E., Sarkissian, S., Simin, T.T., 2003. Spurious regressions in financial economics? J. Fin. 58 (4), 1393–1413.
Gencay, R., 1998. Optimization of technical trading strategies and the profitability in security markets. Econ. Lett. 59, 249–254.
Gerlach, J.R., 2005. Imperfect information and stock market volatility. Fin. Rev. 40, 173–194.
Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess returns on stocks. J. Fin. 48 (5), 1779–1801.
Goddard, J., Kita, A., Wang, Q., 2015. Investor attention and FX market volatility. J. Int. Fin. Mark., Inst. Money 38, 79–96.
Goetzmann, W.N., Jorion, P., 1993. Testing the predictive power of dividend yields. J. Fin. 48 (2), 663–679.
Goetzmann, W.N., Ingersoll, J., Spiegel, M., Welch, I., 2007. Portfolio performance manipulation and manipulation-proof performance measures. Rev. Fin. Stud. 20, 1503–1546.
Golec, J., Tamarkin, M., 1998. Bettors love skewness, not risk, at the horse track. J. Polit. Econ. 106, 205–225.
Goyal, A., Welch, I., 2003. Predicting the equity premium with dividend ratios. Manage. Sci. 49 (5), 639–654.
Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Fin. Stud. 21 (4), 1455–1508.
Granger, C.W.J., Pesaran, M.H., 2000. Economic and statistical measures of forecast accuracy. J. Forecast. 19 (7), 537–560.
Harvey, C.R., Sidique, A., 2000. Conditional skewness in asset pricing tests. J. Fin. 55, 1263–1295.
Hodrick, R.J., 1992. Dividend yields and expected stock returns: alternative procedures for inference and measurement. Rev. Fin. Stud. 5 (3), 357–386.
Homm, U., Pigorsch, C., 2012. Beyond the Sharpe ratio: an application of the Aumann-Seranno index to performance measurement. J. Bank. Fin. 36, 2274–2284.
Hsieh, D.A., 1991. Chaos and nonlinear dynamics: application to financial markets. J. Fin. 46 (5), 1839–1877.
Hsu, P.H., Kuan, C.M., 2005. Re-examining the profitability of technical analysis with data snooping checks. J. Fin. Economet. 3, 606–628.
Inoue, A., Kilian, L., 2004. In-sample or out-of-sample tests of predictability: which one should we use? Econ. Rev. 23, 371–402.
Jarque, C.M., Bera, A.K., 1987. A test of normality of observations and regression residuals. Int. Statist. Rev. 55, 163–172.
Jordan, S.J., Vivian, A.J., Wohar, M.E., 2014. Forecasting returns: new European evidence. J. Emp. Fin. 26, 76–95.

Kostakis, A., Magdalinos, T., Stamatogiannis, M.P., 2015. Robust econometric inference for stock return predictability. Rev. Fin. Stud. 28 (5), 1506–1553.

Kellard, N.M., Nankervis, J.C., Papadimitriou, F.I., 2010. Predicting the equity premium with dividend ratios: reconciling the evidence. J. Emp. Fin. 17 (4), 539–551.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationary against the alternative of a unit root. J. Econometrics 54, 159–178.

Lamont, O., 1998. Earnings and expected returns. J. Fin. 53 (5), 1563–1587.

Leitch, G., Tanner, J.E., 1991. Econometric forecast evaluation: profits versus the conventional error measures. Am. Econ. Rev. 81, 580–590.

Lesmond, D.A., Ogden, J.P., Trzcinka, C.A., 1999. A new estimate of transaction costs. Rev. Fin. Stud. 12 (5), 1113–1141.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth and expected stock returns. J. Fin. 56 (3), 815–849.

Lewellen, J., 2004. Predicting returns with financial ratios. J. Fin. Econ. 74 (2), 209–235.

Li, G., 2005. Information quality, learning, and stock market returns. J. Fin. Quant. Anal. 40, 595–620.

Maio, P., 2013. The "Fed model" and the predictability of stock returns. Rev. Fin. 17, 1489–1533.

McCracken, M.W., 2007. Asymptotics for out-of-sample tests of Granger causality. J. Economet. 140, 719–752.

Moscarini, G., Smith, L., 2002. The law of large demand for information. Econometrica 70, 2351–2366.

Nelson, D.B., 1991. Conditional heteroscedasticity in asset returns: a new approach. Econometrica 59 (2), 347–370.

Nelson, C.R., Kim, M.J., 1993. Predictable stock returns: the role of small sample bias. J. Fin. 48 (2), 641–661.

Nyberg, H., 2011. Forecasting the direction of the US stock market with dynamic binary probit models. Int. J. Forecast. 27, 561–578.

Pagan, A.R., Schwert, G.W., 1990. Alternative models for conditional stock volatility. J. Economet. 45 (1–2), 267–290.

Pesaran, M.H., Timmermann, A., 1995. Predictability of stock returns: robustness and economic significance. J. Financ 50 (4), 1201–1228.

Pesaran, M.H., Timmermann, A., 2000. A recursive modelling approach to predicting UK stock returns. Econ. J. 110, 159–191.

Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335–346.

Pontiff, J., Schall, L.D., 1998. Book-to-market ratios as predictors of market returns. J. Fin. Econ. 49 (2), 141–160.

Poon, S.H., Granger, C.W.J., 2003. Forecasting volatility in financial markets: a review. J. Econ. Lit. 41, 478–539.

Poon, S.H., Granger, C.W.J., 2005. Practical issues in forecasting probability. Fin. Anal. J. 61 (1), 45–56.

Rapach, D.E., Zhou, G., 2013. Forecasting stock returns. In: Elliott, Graham, Timmermann, Allan (Eds.), Handbook of Economic Forecasting, vol. 2. Elsevier, Amsterdam.

Ross, S.A., 1989. Information and volatility: the no-arbitrage martingale approach to timing and resolution irrelevancy. J. Fin. 44, 1–17.

Rozeff, M., 1984. Dividend yields are equity risk premiums. J. Portf. Manage. 11 (1), 68–75.

Schwert, G.W., Seguin, P.J., 1990. Heteroskedasticity in stock returns. J. Fin. 45 (4), 1129–1155.

Stoll, H.R., Whaley, R., 1983. Transaction costs and the small firm effect. J. Fin. Econ. 12, 57–79.

Taylor, S., 1986. Modelling Financial Time Series. Wiley, Chichester.

Vlastakis, N., Markellos, R.N., 2012. Information demand and stock market volatility. J. Bank. Fin. 36 (6), 1808–1821.

Vozlyublennaia, N., 2014. Investor attention, index performance, and return predictability. J. Bank. Fin. 41, 17–35.

West, K.D., 1996. Asymptotic inference about predictive ability. Econometrica 64, 1067–1084.

White, H., 2000. A reality check for data snooping. Econometrica 68 (5), 1097–1126.

Zakamouline, V., Koekebakker, S., 2009. Portfolio performance evaluation with generalized Sharpe ratios: beyond the mean and variance. J. Bank. Fin. 33, 1242–1254.