Journal of Banking & Finance 58 (2015) 1–14

Contents lists available at ScienceDirect

Journal of Banking & Finance

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

Stock market volatility: Identifying major drivers and the nature of their impact ☆

Stefan Mittnik [a,*], Nikolay Robinzonov [a], Martin Spindler [a,b]

[a] Department of Statistics and Center for Quantitative Risk Analysis, Ludwig Maximilians University Munich, Akademiestr. 1/I, 80799 Munich, Germany
[b] Max Planck Society, Munich, Germany

ARTICLE INFO

Article history:
Received 19 February 2014
Accepted 3 April 2015
Available online 28 April 2015

JEL classification:
C55
C58
G17
E00

Keywords:
Componentwise boosting
Financial market risk
Forecasting
GARCH
Exponential GARCH
Variable selection

ABSTRACT

Financial-market risk, commonly measured in terms of asset-return volatility, plays a fundamental role in investment decisions, risk management and regulation. In this paper, we investigate a new modeling strategy that helps to better understand the forces that drive market risk. We use componentwise gradient boosting techniques to identify financial and macroeconomic factors influencing volatility and to assess the specific nature of their influence. Componentwise boosting is capable of producing parsimonious models from a, possibly, large number of predictors and—in contrast to other related techniques—allows a straightforward interpretation of the parameter estimates.

Considering a wide range of potential risk drivers, we apply boosting to derive monthly volatility predictions for the equity market represented by S&P 500 index. Comparisons with commonly-used GARCH and EGARCH benchmark models show that our approach substantially improves out-of-sample volatility forecasts for short- and longer-run horizons. The results indicate that risk drivers affect future volatility in a nonlinear fashion.

1. Introduction

The importance of understanding and reliably modeling financial risk has—again—become evident during the market turbulences in recent years. Accurate volatility predictions for asset prices are crucial when projecting risk measures, such as Value-at-Risk (VaR) or Expected Shortfall, that are commonly used in risk assessment, the design of risk-mitigation strategies, and for regulatory purposes. Although there has been a long tradition in attempting to predict asset prices (cf. Goyal and Welch, 2003; Welch and Goyal, 2008; Cochrane and Piazzesi, 2005; Lustig et al., 2011), the intense interest in volatility modeling began only after the seminal works of Engle (1982) and Bollerslev (1986), and has since become an extensively researched area in the field of financial econometrics.

Despite this tremendous interest, the vast majority of studies on predicting financial-market risk have been confined to conditioning only on past return histories as conditional information.[1] Only relatively few studies have analyzed to what extent the information contained in other financial or macroeconomic variables helps to improve volatility predictions. Employing autoregressive models, Schwert (1989) analyzes the relation of stock volatility and macroeconomic factors, such as GDP fluctuations, economic activity and financial leverage. Engle et al. (2013) use inflation and industrial production in a mixed-frequency GARCH framework to predict the volatility of U.S. stock returns. They show that incorporating economic fundamentals into volatility models pays off in terms of long-horizon forecasting and that macroeconomic fundamentals play a significant role even at short horizons. Flannery and Protopapadakis (2002) analyze the impact of real macroeconomic

☆ A previous version of the working paper was circulated under the title "Boosting the Anatomy of Volatility." Part of the research was conducted while Martin Spindler was visiting the Department of Economics at MIT, Cambridge, USA, with financial support from the German Research Foundation (DFG). He thanks for the hospitality and fruitful research environment.

* Corresponding author at: Ludwig Maximilians University, Department of Statistics, Chair of Financial Econometrics, Akademiestr. 1/I, 80799 Munich, Germany. Tel.: +49 89 2180 3224; fax: +49 89 2180 5044.

E-mail address: finmetrics@stat.uni-muenchen.de (S. Mittnik).

[1] A comparison of alternative VaR forecasting strategies that follow this line is given in Kuester et al. (2006).

http://dx.doi.org/10.1016/j.jbankfin.2015.04.003
0378-4266/

variables on aggregate equity returns; and Engle and Rangel (2008) find that macroeconomic variables help predicting the low-frequency component of volatility. Paye (2012) and, especially, Christiansen et al. (2012) consider extended sets of macroeconomic factors and a broader range of asset classes. Both use conventional linear approaches to model log-transformed realized volatility and include lagged volatility as well as financial and macroeconomic factors as predictors. Christoffersen and Diebold (2000) analyze the predictability of volatility for different markets on a daily basis. Their conclusion is that when the horizon of interest is longer than ten or twenty days, depending on the asset class, then volatility is effectively not predictable. Another interesting line of research focuses on implied volatility, (Canina and Figlewski, 1993; Christensen and Prabhala, 1998; Jiang and Tian, 2005; Prokopczuk and Wese Simen, 2014). While this approach is perfectly appropriate for forecasting purposes, it does not directly allow an analysis of the influence of macroeconomic factors on financial-market volatility.

In view of the limited number of studies and their varying approaches, there is little or no consensus concerning the usefulness of financial and macroeconomic variables for volatility prediction. And it is this issue which we address in this paper. To gain deeper insights into the nature of volatility processes, we employ so-called boosting techniques. As will be demonstrated, given a large set of potential risk drivers, boosting enables us not only to identify the factors that drive or lead[2] market risk, but also to assess the specific nature of their impact. The selection of relevant volatility drivers and the estimation of their particular—potentially nonlinear—influence is accomplished in a data-driven fashion, requiring only minimal subjective decisions concerning model specification.

Although boosting has been shown to be a useful approach in many statistical applications, it has been more or less ignored in empirical economics and finance. Among the very few exceptions are Bai and Ng (2009), who use it for predictor selection in factor-augmented autoregressions, and Audrino and Bühlmann (2009), who apply it to modeling stock-index volatility. In this paper, we demonstrate the usefulness of boosting techniques for modeling financial market risk. The approach we adopt differs from the initial approach of Audrino and Bühlmann (2009) in several aspects—three of which we regard as particularly relevant. First, we go beyond the usual GARCH specification by allowing a large number of exogenous risk drivers to affect volatility, in order to improve our understanding of the nature of volatility processes. Second, we employ a predictor-selection strategy that largely avoids subjective specification decisions. Moreover, instead of the componentwise *knot* selection in bivariate-spline estimation adopted in Audrino and Bühlmann (2009), we employ componentwise *predictor* selection, giving rise to a better interpretability of the estimated model, in order to facilitate the interpretability of the model obtained.

This paper contributes to the existing literature on volatility modeling in several ways. First, we investigate the role of a broad set of potential macroeconomic and financial factors in determining future stock-market volatility. Second, by employing boosting techniques, we gain deeper insight into the nature of the forces driving volatility. Models obtained via boosting techniques can be directly used for forecasting. Alternatively, specifications obtained via boosting—i.e., the selection of risk drivers and the description of the response behavior they induce—can serve as a starting point for more elaborate, possibly, nonlinear model-building procedures. Third, our empirical results strongly suggest that both the use of macroeconomic information and permitting nonlinear relationships help predicting volatility. Conducting forecasting comparisons with commonly employed GARCH and EGARCH benchmarks, we demonstrate that the boosting strategy we adopt clearly outperforms these benchmarks in the short and, especially, in the medium and long run. We show that the source of the short-term improvement is attributable to the factor-selection capabilities of boosting, whereas the medium- and long-term outperformance is due to allowing factors to have nonlinear effects on volatility.

Although not the focus here, our modeling approach can also serve policy and regulatory purposes. The boosting strategy chosen identifies specific regions where factors tend to critically affect market risk. Thus, the approach can help policy makers and regulators to identify critical thresholds at which interventions may be called for and can also help designing financial stabilization mechanisms.

The remainder of the paper is organized as follows. Section 2 details and illustrates the specific boosting algorithm adopted. Section 3 discusses the volatility measure and predictor variables employed, the way multi-step forecasting comparisons are conducted, and the results we obtain. Section 4 concludes.

2. A boosting approach to modeling volatility

Boosting, as put forth in Freund and Schapire (1996), was originally designed to solve binary classification problems. To do so and to achieve any desirable degree of accuracy, it suffices that the classifier (also called base learner) performs only slightly better than random guessing (Kearns and Valiant, 1994; Schapire et al., 1998). Friedman (2001) placed boosting in a regression framework, viewing it as a gradient descent technique. Boosting is especially suitable in applications where there is a large number of—possibly "similar"—predictors, as it curbs multicollinearity problems by shrinking their influence towards zero.

Componentwise boosting combines model estimation and model selection in a unified, iterative framework and has a number of advantages: (i) It selects relevant predictors for the variable of interest and ignores redundant ones. (ii) It easily handles high-dimensional situations where the number of covariates can even exceed the number of observations, a situation where classical approaches, such as (nonlinear) regression analysis and maximum likelihood estimation, typically fail. Moreover, these latter approaches are only applicable *after* the model has been fully specified. (iii) It captures nonlinear dependencies. (iv) In contrast to other flexible prediction methods (such as random forests), componentwise boosting generates results that can be interpreted straightforwardly. (v) Boosting has very good properties concerning prediction, comparable to Lasso. For the linear model, consistency of L_2-boosting in prediction norm was shown in Bühlmann (2006).

Before we start with a more detailed explanation of boosting, let us remark on the difference between boosting and factor modeling and the problem of statistical significance. Linear factors models are usually applied for dimension reduction in large data sets and each factor represents a linear combination of variables. This makes a direct, variable-specific interpretation of factor models more difficult. In contrast, boosting identifies individual variables that influence the dependent variable, not combinations of potential drivers. As of yet, a drawback of boosting concerns significance testing. Sofar, there are no results for inference. This is still subject of ongoing research. As far as prediction is concerned, the focus here, superior performance has, however, been demonstrated.

Volatility modeling via gradient boosting was first considered in Audrino and Bühlmann (2003), who adopted a GARCH-type framework, assuming a stationary return process of the form $y_t = \sigma_t \varepsilon_t, \varepsilon_t \stackrel{iid}{\sim} N(0,1)$ and a rather general dependence of σ_t on

[2] Throughout the paper we use terms like "driver," "factor" and "leading indicator" interchangeably implying only the possibility of Granger causation or "usefulness for prediction."

past returns. Their approach aims purely at prediction, as the resulting model has limited interpretability. A similar model, with neural networks as base learners, was proposed by Matías et al. (2010).

In the analysis below, we use so-called componentwise gradient boosting (see Bühlmann and Yu, 2003; Bühlmann and Hothorn, 2007), which is designed to simultaneously select relevant predictors and to capture the specific nature of their impact. Next, we briefly summarize and motivate our strategy to volatility modeling. More details of the method and a small simulation study illustrating the approach are given in the appendix.

The modeling framework we choose builds on the exponential ARCH specification of Nelson (1991), but is augmented to include—in a rather flexible way—a large number of risk drivers that potentially affect volatility. The total number of predictors can be very large and may, in principle, even exceed the sample size. The specific form of our model is given by

$$\begin{aligned} y_t &= \exp(\eta_t/2)\varepsilon_t \\ \eta_t &= \eta(\mathbf{z}_t) = \beta_0 + f_{\text{time}}(t) + f_{\text{yr}}(n_t) + f_{\text{month}}(m_t) \\ &\quad + \sum_{j=1}^{s} f_j(y_{t-j}) + \sum_{k=1}^{q}\sum_{j=1}^{p} f_{k,j}(x_{k,t-j}), \end{aligned} \tag{1}$$

where $y_t = \log(P_t/P_{t-1})$ denotes the logarithmic return, P_t is the asset prices at time t, and $\varepsilon_t \overset{iid}{\sim} N(0,1)$. The r-dimensional vector $\mathbf{z}_t = (1, t, n_t, m_t, y_{t-1}, \ldots, y_{t-s}, x_{1,t-1}, \ldots, x_{1,t-p}, \ldots, x_{q,t-1}, \ldots, x_{1,t-p})^\top$, with $r = s + qp + 4$, contains the predictor realizations available at or prior to time $t-1$. To keep the exposition simple, we assume, without loss of generality, that η_t has zero mean and, thus, omit β_0.

We specify all $f_{\cdot}(\cdot)$ functions in (1) as regression trees. That is, for any component $z_{i,t}$, $i = 1, \ldots, r$, in $\mathbf{z}_t$ (in the following simply denoted by z), $f_{\cdot}(z)$ is given by

$$f_{\cdot}(z) = \sum_{j=1}^{J_z} \gamma_{zj} \mathrm{I}_{R_{zj}}(z),$$

where $\mathrm{I}_{R_{zj}}(z)$ denotes the indicator function, i.e., $\mathrm{I}_{R_{zj}}(z) = 1$, if $z \in R_{zj} \subset \mathbb{R}$, and $\mathrm{I}_{R_{zj}}(z) = 0$, otherwise; the $R_{zj}, j = 1, 2, \ldots, J_z, J_z \in \mathbb{N}$, denote disjoint regions (or regimes) partitioning the domain of z; and γ_{zj} denotes the corresponding constants representing the impact of z on η_t in that particular region. The number of regimes, J_z, the boundaries of the regions, R_{zj}, and the γ_{zj}-values are not specified in advance, but rather determined by the boosting algorithm. Functions $f_{\text{time}}(\cdot)$, $f_{\text{yr}}(\cdot)$ and $f_{\text{month}}(\cdot)$ capture possible deterministic trend and seasonal components in volatility; $f_j(y_{t-j})$, $j = 1, \ldots, s$, capture the influence of past returns; and $f_{k,j}(x_{k,t-j})$, $j = 1, \ldots, q$, are functions of lagged predictors. The selection process typically excludes some, if not many, of the r predictors from (1), implying that only a subset of the initial r predictors are relevant for explaining volatility. In other words, we may intentionally specify a broad set of predictor candidates, as it tends to get rigorously pruned by the boosting algorithm.

Regression trees can capture complex forms of dependence by recursively partitioning the predictor domain into regions with similar response behavior and assigning a constant response value to each regime.[3] In contrast to the "smooth" specifications of classical nonparametric regression models, regression trees can handle abrupt changes which makes them an attractive choice when modeling financial-market volatility. Moreover, regression trees have the advantage that, in autoregressive dynamic settings, the question of explosive behavior does not arise, as the response is described by (sets of) constants rather than multiplicative autoregressive coefficients. To decide on the partitioning of the domains, we maximize the absolute value of the standardized difference between the means of the two adjacent groups among all possible split positions (see also Hothorn et al., 2006).

Our setup allows for the presence of complex volatility responses that go beyond asymmetry—a key feature, for example, of exponential GARCH models that allow volatility responses to negative shocks to differ from those to positive ones—and permits multiple regimes that are unknown in advance and determined in a data-driven way.[4] The decision, which of the drivers under consideration are relevant, is also part of the data-driven specification process and requires no prior information. As a measure of market risk we use *realized* variance rather than the conventional variance modeled in a GARCH framework, since unobservable covariates, such as lagged variance or error terms, make the selection process nontrivial if not impossible.

We estimate (1) via componentwise gradient boosting,[5] which derives the final model in a highly flexible way by sequentially combining a series of individual predictor components. It, thus, provides a joint procedure for model specification and estimation. Our estimation minimizes the expectation of some (with respect to η differentiable) loss function, L, and solves

$$\hat{\eta}_t = \arg\min_{\eta} \frac{1}{T}\sum_{t=1}^{T} L(y_t, \eta(\mathbf{z}_t; \boldsymbol{\beta})), \tag{2}$$

where $\boldsymbol{\beta}$ denotes the unknown parameter vector to be estimated in a parametric setting. The solution to (2) is derived by reducing the empirical loss in successive steps. The final $\boldsymbol{\beta}$-estimate is given simply by the sum of the estimates obtained in each step and has—in contrast to alternative flexible approaches like bagging or random forests—a direct interpretation.[6]

To estimate the desired characteristic of the conditional distribution, the loss function, L, needs to be appropriately specified. Under the assumption $y_t|\mathbf{z}_t \sim N(0, e^{\eta_t})$, the negative conditional log-likelihood loss function and the negative gradient are, respectively,

$$L_t = \frac{1}{2}\left[\eta_t + \frac{y_t^2}{e^{\eta_t}}\right] \quad \text{and} \quad g_t = -\frac{\partial L_t}{\partial \eta_t} = \frac{1}{2}\left[\frac{y_t^2}{e^{\eta_t}} - 1\right]. \tag{3}$$

Instead of fitting all components of vector $\mathbf{z}_t$ simultaneously, they are fitted individually using the specified base-learner function. At each boosting step, only one component is included, namely the one which correlates most strongly with the negative gradient. Such a step can be viewed as a partial "sub-solution" to the global optimization problem. As base learner we use regression trees with two nodes. Doing so, the algorithm simply splits a predictors sample range optimally into two disjoint regions and assigns constant volatility response values to each region. This seems to be a rather crude way of approximating complex volatility responses. However, by iterating this procedure sufficiently many times, we can—as illustrated in Appendix B—capture rather elaborate response patterns. The estimates obtained during an iteration do not fully enter but rather in terms of a (small) fraction. This form of shrinkage helps to dampen the "greediness" of the gradient technique, which may otherwise be prone to neglecting correlated predictor candidates, and to cure the typical instability of forward selection methods (Breiman, 1996).

[3] For a detailed discussion on regression trees, see Breiman et al. (1984). Below, we use so-called conditional inference trees (Hothorn et al., 2006).

[4] See the results in Section 3.4 and Appendix C for further details.

[5] See Hothorn et al. (2013) for a software implementation.

[6] To avoid overfitting, we start with controlling the bias-variance tradeoff by using a low-variance/high-bias specification. In subsequent steps, the bias will be gradually reduced, with the variance increasing at a slower rate (Bühlmann and Yu, 2003).

Table 1
Description of the financial, macroeconomic and lagged volatility predictor variables employed.

Variable	Abbrev.	Source	Description
A. Equity Market Variables and Risk Factors			
Dividend Price Ratio (Log)	Shiller	D-P	Dividends over the past year (12-month moving sum) relative to current market prices (in logs)
Earnings Price Ratio (Log)	E-P	Shiller	Earnings over the past year (12-month moving sum) relative to current market prices (in logs)
US Market Excess Return	MKT	Fama French	Fama–French's market factor: U.S. stock market return minus one-month T-Bill rate
Size Factor	SMB	Fama French	Fama–French's SMB factor: Return on small stocks minus return on big stocks
Value Factor	HMLFX	Fama French	Fama–French's HML factor: Return on value stocks minus return on growth stock
Short Term Reversal Factor	STR	Fama French	Fama–French's short-term reversal factor: Return on stocks with low prior one-month return minus return on stock with high prior return
S&P 500 Turnover	TURN	CRSP	Turnover for the S&P 500
S&P 500 Return	mreturns	Datastream	Monthly log returns of the S&P 500
CBOE Market Volatility Index	VIX	CBOE	Measure of the implied volatility of S&P 500 index options
Log realized variance	LRVar	Datastream	Log realized variance defined in Eq. (4)
Change of LRVar	LRVar.c	Datastream	Change of the log realized variance
B. Interest Rates, Spreads and Bond Market Factors			
T-Bill Rate (Level)	T-B	Goyal Welch	Three-month T-Bill rate
Rel. T-Bill Rate	RTB	Goyal Welch	T-Bill rate minus its 12 month moving average
Long Term Bond Return	LTR	Goyal Welch	Rate of return on long term government bonds
Rel. Bond Rate	RBR	Goyal Welch	Long-term bond yield minus its 12 month moving average
Term Spread	T-S	Goyal Welch	Difference of long-term bond yield and three-month T-Bill rate
Cochrane Piazzesi Factor	C-P	Cochrane Piazzesi	Measure of bond risk premia; recursively estimated based on Fama-Bliss data
C. FX Variables and Risk Factors			
Return on Dollar Risk Factor	DOL	Lustig et al. (2011)	FX risk premium measure; average premium for bearing FX risk
Average Forward Discount	AFD	Lustig et al. (2010)	Aggregate predictor of FX returns calculated from forward rates and spot rates
D. Liquidity and Risk Variables			
Default spread	DEF	Goyal-Welch	Measure of default risk: BAA minus AAA corporate bond yields
FX average bid-ask spread	BAS	Menkhoff et al. (2011)	Bid-ask spreads as measure of illiquidity in foreign exchange markets
Pastor-Stambaugh liquidity factor	PS	Pastor Stambaugh	Measure of stock market liquidity based on price reversals
TED spread	TED	Datastream	Measure of illiquidity: LIBOR minus T-Bill rate
E. Macroeconomic Variables			
Inflation Rate, Monthly	INFM	Datastream	Monthly (log) growth rate of the U.S. consumer price index
Inflation Rate, YoY	INFA	Datastream	Year-over year (log) growth rate of the U.S. consumer price index
Industrial Production Growth, Monthly	IPM	Datastream	Monthly (log) growth rate of U.S. industrial production
Industrial Production Growth, YoY	IPGA	Datastream	Year-over year (log) growth rate of U.S. industrial production
Housing Starts	H-S	Datastream	Monthly change in housing started
M1 Growth, Monthly	M1M	Datastream	Monthly (log) growth rate of U.S. M1
M1 Growth, YoY	M1A	Datastream	Year-over-year (log) growth rate of U.S. M1
Orders, Monthly	ORDM	Datastream	New orders, consumer goods and materials; monthly growth rate
Orders, YoY	ORDA	Datastream	New orders, consumer goods and materials; year-to-year growth rate
Return CRB Spot	CRB	Datastream	Commodity price spot index; annual log difference
Capacity Utilization	CAP	Datastream	Level to which the productive capacity is used
Employment Growth	EMPL	Datastream	Change in the employed population
Consumer Sentiment	SENT	Datastream	Monthly change in University of Michigan consumer sentiment
Consumer Confidence	CONF	Datastream	Monthly change in consumer confidence index
Diffusion Index	DIFF	Datastream	Philadelphia Fed Business Outlook Survey Diffusion Index
Chicago PM Business Barometer	PMBB	Datastream	Leading indicator of economic health; survey of purchasing managers
ISM PMI	PMI	Datastream	Monthly change in purchasing manager index

As an alternative to two-node regression trees, other specifications, such as linear or spline functions, could be chosen as base learners. For the application at hand, our simple regression-tree specification turned out to be the better choice. This seems largely due to the fact that, in addition to being able to capture nonlinear response patterns, it can best cope with the abrupt and asymmetric volatility responses we observe. Tree specifications are less prone to outliers than linear models and, in contrast to higher-order spines, behave nicely at the borders.

In summary, the modeling strategy we adopt is rather flexible and has the advantage of providing us with interpretable parameter estimates, so that it should help us to gain insights into the role of particular risk drivers and to better understand the nature of volatility processes. To what extent this translates into better risk predictions will be investigated next.

3. Boosting stock-market volatility

To examine the usefulness of boosting for modeling and predicting equity market volatility, we take the S&P 500 stock index as a representative candidate and entertain a range of financial and macroeconomic factors as potential volatility drivers. Next, after describing the data and detailing the boosting specifications, we present the empirical results in two parts. First, we compare the predictive performance of the boosting approach with that of the benchmark candidates and, then, we take a closer look at the causes for the improvements in forecasting accuracy. Finally, we briefly discuss the nature of the impact the driving factors have on stock market volatility.

3.1. Data and model specification

Our monthly S&P 500 index data cover the period December 1989 to December 2010, amounting to 253 months in total. As potential volatility drivers, we consider the 40 financial and macroeconomic factors summarized in Table 1. These factors can be divided into five categories:

(A) *Equity Market Variables and Risk Factors:* This set comprises well-known equity factors, such as dividend price ratio, earnings price ratio and Fama–French factors. Moreover, we include returns of the MSCI world stock market index, the implied volatility index (VIX) derived from S&P 500 index options traded on the Chicago Board Options Exchange, and the turnover of the S&P 500 which might reflect traders' uncertainty about future market valuations.

(B) *Interest Rates, Spreads and Bond Market Factors:* This category comprises of interest rates and spreads employed by Welch and Goyal (2008), namely, the T-Bill rate, relative T-Bill rate, long term bond return and term spread. Moreover, the Cochrane and Piazzesi (2005) bond factor is included.

(C) *FX Variables and Risk Factors:* This set contains the return on Dollar risk factor and average forward discount. For both we refer to Lustig et al. (2011).

(D) *Liquidity and Risk Variables:* As liquidity measures for different markets, we use the default spread, TED spread, FX average bid-ask spread (Menkhoff et al., 2011) and the Pastor and Stambaugh, 2003 liquidity factor.

(E) *Macroeconomic Variables:* This is the largest, group of factors, containing the inflation rate, industrial production, housing starts, M1 growth, orders, return CRB spot, consumer confidence and others.

We include the first and second lag of all 40 factors as potential predictors. We include two lags of log realized variance and changes in log realized variance to capture temporal and state dependence in volatility, and also allow for seasonal components. This gives us altogether $r = 84$ predictors.

As volatility cannot be observed directly, we follow French et al. (1987) and Schwert (1989)[7] and use monthly *log realized variance*, calculated from daily returns, as proxy for market volatility,[8] i.e.,

$$\mathrm{LRVar}_t = \log \sum_{\tau=1}^{M_t} r_{t,\tau}^2, \quad t = 1, \ldots, T, \tag{4}$$

where $r_{t,\tau}$ denotes the τth daily return in month t; and M_t the number of trading days in month t. Fig. 1 shows the log realized-variance series for the equity market in the chosen period.

3.2. Predictive performance

The predictive performance is examined via rolling-window forecasting for the period June 2002 to September 2010. Starting with a history of 153 months, we move the fixed-length window forward month by month, re-estimate, and generate a sequence of one-step-ahead forecasts for 100 months. Applying a direct forecasting approach,[9] we also produce multi-period forecasts for horizons of up to six months by adapting (1) accordingly, i.e.,

$$\begin{aligned} y_{t+h} &= \exp(\eta_{t+h}/2)\varepsilon_{t+h}, \quad h = 1, \ldots, 6, \\ \eta_{t+h} &= \beta_0 + f_{\mathrm{time}}(t+h) + f_{\mathrm{yr}}(n_{t+h}) + f_{\mathrm{month}}(m_{t+h}) \\ &\quad + \sum_{j=0}^{s-1} f_j(y_{t-j}) + \sum_{k=1}^{q}\sum_{j=0}^{p-1} f_{k,j}(x_{k,t-j}). \end{aligned} \tag{5}$$

For two reasons, we adopt a direct forecasting approach rather than a recursive one, i.e., deriving chains of six one-step-ahead predictions. First, whereas recursive predictions with the GARCH benchmark models is the obvious choice, the use of exogenous variables in our boosting approach would either require to also predict those variables in some recursive manner or to use the observed values that were realized after the time period the prediction is made. The former is highly impractical as it requires predictive models for 38 variables; and the latter is impossible in real-time forecasting applications. The second, and here more important reason is that we are interested in examining how the impact of risk drivers change as the forecasting horizon grows. In other words: We want to understand which drivers matter in the short and in the long term, and how does the nature of their impact change.

To assess the predictive performance, we compare multi-step, out-of-sample boosting forecasts to their (recursive) counterparts derived from a GARCH(1,1) and an Exponential GARCH(1,1) (Nelson, 1991) benchmarks.[10] Clearly, there are many potential alternatives that could serve as a benchmark.[11] However, Hansen and Lunde (2005)—addressing the question: "Does anything beat a GARCH(1,1)?"—conclude that there are essentially no or only very little benefits from using more elaborate models, so that the GARCH(1,1) model can be regarded as a challenging benchmark and the EGARCH model is a natural competitor in the context of our model specification.

By allowing exogenous variables to enter our volatility model, the comparison with the standard benchmark models may not seem fair. However, the estimation of GARCH models given a large set of potential explanatory variables and a short data history, as is the case here, is not obvious—especially, when the predictors affect the conditional variance in a complex nonlinear fashion. The ability to meaningfully select relevant drivers and to specify the nature of their influence is the strength of componentwise boosting.[12]

We evaluate the forecasting performance in terms of the mean squared prediction error, i.e., the mean of the squared differences between the realized volatility and the h-step-ahead forecast given by (5).[13]

The results, reported in Table 2, show that model (1) obtained via componentwise boosting clearly outperforms the GARCH and EGARCH benchmark at all horizons considered. As the predictions of all three models are virtually unbiased, the lower MSEs for the boosting model result from the fact that it produces less extreme prediction errors than the benchmarks.

To assess the statistical significance of the forecasting improvements, we apply the Diebold–Mariano test (Diebold and Mariano, 1995) in the modified version of Harvey et al. (1997). The null hypothesis of the test presumes that benchmark forecasts are more accurate than those of the proposed model, so that rejection of the null favors our approach. The p-values of the modified Diebold–Mariano test are reported in Table 3. It turns out that boosting forecasts are significantly better for all medium- and long-term horizon and very competitive in short-term. We also conducted the Giacomini–White test (Giacomini and White, 2006). The results, reported in Table 3, confirm those of the Diebold–Mariano test.

---

[7] Note that Schwert (1989) also investigates the influence of the volatility of macroeconomic variables on stock market volatility, but finds only weak evidence. We, therefore, do not consider macroeconomic volatility as drivers.

[8] For an in-depth review of the realized-volatility concept, we refer to Andersen et al. (2006).

[9] For direct forecasting via boosting in a nonlinear time series context, see Robinzonov et al. (2012).

[10] We compute multi-step (E) GARCH forecasts recursively. By fitting GARCH models, using data sampled at each frequency compatible with horizons $h = 1, \ldots, 6$, we also computed nonrecursive h-step forecasts. These, however, turned out to be rather poor and are not reported here.

[11] Christiansen et al. (2012) use an autoregressive model for realized volatility as benchmark.

[12] Various advanced procedures for variable selection exist, such as Least Angle Regression (LARS, Efron et al., 2004) or the Least Absolute Shrinkage and Selection Operator (LASSO, Tibshirani, 1996). They are, however, geared towards modeling conditional means rather than conditional variances, which is the focus here. For discussions of the selection properties of boosting over LARS- and LASSO-type variable-selection methods see, for example, Bai and Ng (2009) and Mayr et al. (2012).

[13] The h-step expected squared prediction error is given by $\mathrm{ERR}_{t+h} = (\mathrm{RV}_{t+h} - \hat{\eta}_{t+h})^2, h = 1, \ldots, 6$, where $t = 154, \ldots, 253$, i.e., the last one hundred observations covering the period August 2002 to December 2010, and is measured by the average of the observed squared errors. Employing other loss functions, such as the mean absolute error (MAE), gave similar results and left the ranking of the models unchanged.

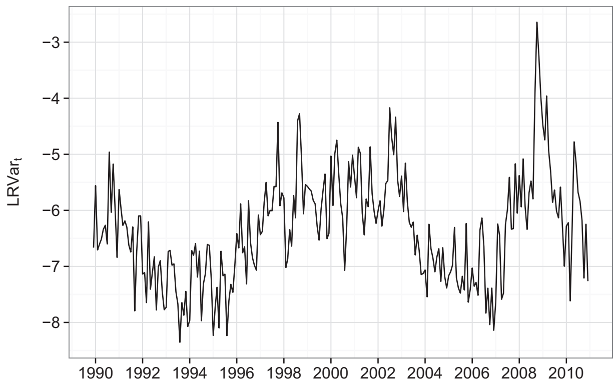


Fig. 1. Time series of monthly realized variance (in logarithms), as defined by (4), of the S&P 500.

Thus, allowing for exogenous factors and nonlinear dependencies helps to improve volatility forecasting for all horizons.

3.3. Where does the improvement come from?

As we have seen in the previous section, the model obtained by boosting delivers superior forecasting results. As the model includes macroeconomic explanatory variables and allows them to affect volatility in a nonlinear fashion, an obvious question is: Where does the predictive improvement mainly come from? Is it simply the inclusion of the additional variables, or does the nonlinear specification improve predictive performance?

As it turns out, the gain in forecasting accuracy is largely attributable to both the explanatory power of the drivers allowing them to affect volatility in a nonlinear fashion. However, the role of the two phenomena changes as the forecasting horizon increases. To disentangle these two sources of improvement, we adjust model (1) and use linear instead of constant base learners by specifying

$$\begin{aligned} y_t &= \exp(\eta_t/2)\varepsilon_t \\ \eta_t &= \beta_0 + \sum_{j=1}^{s}\beta_j y_{t-j} + \sum_{k=1}^{q}\sum_{j=1}^{p}\gamma_{k,j}x_{k,t-j} =: \eta(\mathbf{z}_t), \end{aligned} \tag{6}$$

where the deterministic (seasonal) components are omitted.

For the following discussion, we label the linear specification (6) by "L1" and the tree-based model (1) by "T1." A comparison of L1 and T1 reveals to what extent the gain in accuracy can be attributed to the nonlinear structure. Furthermore, we specify a linear model, labeled "L0" that is nested in L1 and which includes only past returns but omits all macroeconomic factors. A comparison between L0 and L1 reveals to what extent gains can be attributed to the drivers in the linear setting. Analogously, we entertain a fourth, tree-based model, labeled "T0", which includes only lagged returns and, thus, is nested in T1. Altogether, we consider the four models summarized in Table 4.

We subject the four models to the same forecasting exercise described in Section 3.1. The results are summarized in terms of boxplots of the 100 squared prediction errors for each of the six forecasting horizons shown in Fig. 2. The boxplots indicate that models making use of exogenous information, i.e., specifications L1 and T1, drastically increase forecasting accuracy. Thus, a substantial part of the gain in accuracy can be attributed to the predictive power of the selected macroeconomic drivers. The linear specification with exogenous drivers, L1, delivers the best one- and two-step predictions. Beyond that, the nonlinear tree-specification, T1, dominates all other specifications. This observation is in line with Maheu and McCurdy (2002) and Sensier and van Dijk (2004), who argue that changes in volatility are more appropriately captured when allowing for instantaneous breaks and large, abrupt changes rather than gradual adjustments—features for which regression trees are particularly well suited. Only for short-term forecasts does the linear specification outperform the coarser regression-tree setup. For longer-term forecasts (here, after step four), however, the predictive performance of the linear specification drops to that of the benchmark models.

Table 2
Mean squared errors of boosting and benchmark models.

Horizon	GARCH	EGARCH	Boosting
1	0.588	0.753	0.527
2	0.927	0.903	0.712
3	1.025	1.107	0.770
4	1.122	1.526	0.742
5	1.029	1.331	0.664
6	1.353	1.432	0.927

Table 3
Results for the modified Diebold–Mariano test (DM test) and Giacomini–White test (GW test) for the benchmark models GARCH and EGARCH, with $*$, $**$ and $***$ indicating 10, 5, and 1% significance, respectively.

Horizon	GARCH		EGARCH	
	DM test	GW test	DM test	GW test
1	0.275	0.137	0.075*	0.132
2	0.060*	0.219	0.107	0.347
3	0.009***	0.027**	0.035**	0.056*
4	0.001***	0.010***	0.001***	0.002***
5	0.002***	0.006***	0.010***	0.012***
6	0.006***	0.014**	0.038**	0.053*

3.4. What are the influential drivers?

From a theoretical and applied finance point of view and also from a policy-making perspective, it is of interest to identify the risk drivers in equity markets and to assess the specific manner in which they affect volatility. Such knowledge could, for example, be used for developing early-warning mechanisms for market turbulences or designing market-stabilization strategies. For this purpose, the proposed modeling strategy is much better suited than the usual GARCH framework with its black-box nature. Boosting enables us to extract information contained in risk-driving factors that can be helpful for regulation and policy making.

In this section, we present the general volatility response patterns for all horizons considered and, taking a somewhat longer-term perspective, illustratively discuss the results for the six-month horizon in more detail. The results for the other horizons can be interpreted in a similar way. Appendix C shows the plots of the full set of relevant volatility drivers for horizons one through six.

A first observation is that for all horizons only a fairly small number of macroeconomic and financial variables are selected as drivers of market volatility. Their number varies between six and nine. The VIX is identified as an important predictor for realized volatility and is preferred by the selection mechanism across all horizons.[14] This suggests that the options market provides information that leads realized volatility. The VIX signals changes in both positive and negative directions. In principle, two VIX regimes can be identified: values below about 15–18 indicate a decrease whereas values above that range an increase in future S&P500 volatility. The fact that this phenomenon holds for horizons beyond two months

[14] The fact that, in a GARCH setting, option-implied volatility is a helpful volatility predictor beyond that what past (squared) returns offer has been shown before (cf. Claessen and Mittnik (2002) and references therein).

Table 4
Overview of models specified to identify the sources of predictive performance.

Label	Specification
L0	Linear with lagged returns
T0	Tree with lagged returns
L1	Linear with lagged returns + exogenous variables
T1	Tree with lagged returns + exogenous variables

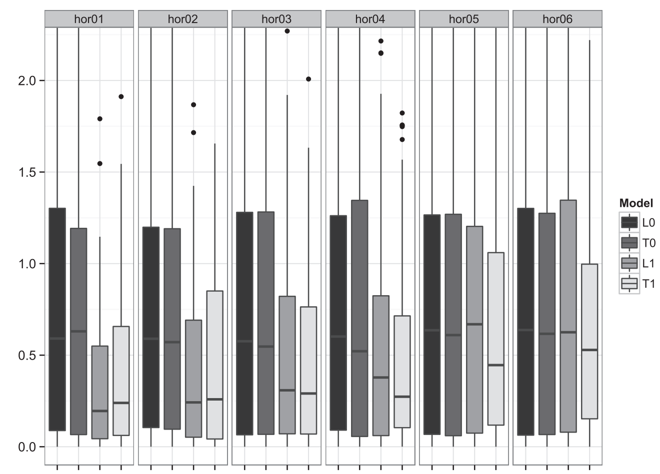


Fig. 2. Forecasting comparison for horizons one through six months between the four models L0, L1, T0 and T1 summarized in Table 4. L* and T* denote linear and tree-based models, respectively, and *0 (*1) denote models that exclude (include) all exogenous drivers.

reflects the persistency of the effects the VIX has. Also, log realized volatility (LRVar) itself, defined in (4), is a strong indicator, which signals changes in volatility for up to five months in advance.

In view of the full set of results shown in Appendix C, we can conclude that VIX and LRVar belong to the very few variables capable of forecasting a *decrease* in realized volatility. The others are TED spread and new orders of consumer goods (monthly and annually). All other macro variables appear to be only useful for predicting volatility *increases*.

In the following, we discuss the role of the drivers for the realized volatility of the S&P500 in more detail, focusing on the longest, i.e., the six-period-ahead horizon. We present the chosen factors, discuss their role and graphically focus on the impact of the most relevant drivers in Fig. 3. In that figure, the ticks along the horizontal axes show the data that were observed for the respective driver during the sample period. We identify the following leading variables: VIX, TED spread, orders (year over year), HML factor, and CRB returns. The built-in variable selection process excluded all other drivers considered. Fig. 3 shows the first and second lags of two important factors. As is to be expected, not all the lags of the relevant predictors relate to future realized volatility. For example, TED spread (Fig. 3, bottom panel), orders and HML factor enter only with their first lag, whereas CRB returns displays a longer-lived impact and enters with its second lag. For VIX both lags are selected and shown in Fig. 3 (upper panel).

VIX values (first lag) below 17 indicate a drop in the volatility and values above this threshold an increase. Specifically, values below 17 suggest a decrease in volatility by 0.40 on the log scale, which corresponds to a decrease of 18% in the volatility (or standard deviation),[15] while values above this threshold signal a volatility increase by 0.25 on log scale (or 13% increase in volatility). For the second lag (Fig. 3, upper panel) we observe three regimes, but the effects are less pronounced compared to the first lag. Values below 17, again, signal a decrease (−0.15 on log scale or −7%) and higher values a slight increase in volatility. Clearly, the VIX, which is the options-implied measure of S&P 500 volatility, is identified as a main predictor beyond lagged log realized variance. Overall, given the persistence of volatility, the observation that past volatility is an important predictor for future volatility is less of a surprise.

The TED spread can be interpreted as a measure of illiquidity. Values below 0.01 have a slight negative effect on next period's volatility. TED-spread values between 0.011 and 0.014 tend to increase volatility by approximately 17% and values above 0.014 by 28%. High levels of illiquidity generate uncertainty and nervousness among market participants and, thus, drive up volatility. The TED spread has been widely recognized as an indicator of perceived credit risk (cf. Brunnermeier (2009) and Mittnik and Semmler (2015) and references therein). Our findings suggest that increased TED spread also spills over and induces risk into equity markets.

Another finding is the regime-dependence of volatility with respect to new orders of consumer goods and materials. A moderate or strong drop in orders increases the future log realized variance slightly, whereas small drops or increases have no influence. Higher orders of consumer goods signal a positive development for producing companies, taking risk out of financial markets, whereas strong decreases induce uncertainty and jumps in stock-market volatility. Orders are known to be a reliable leading indicator for the economic growth. An increase in orders boosts companies' future earnings, lowers debt-equity ratios and, thus, market risk.

4. Conclusions

We have used boosting techniques to assess the influence of a wide range of potential financial and macroeconomic risk drivers for the S&P 500 index. The specific approach chosen relies regression trees as fundamental building blocks and allows us to identify influential volatility drivers together with the particular form of their impact.

Our empirical results give insight into the "anatomy" of volatility by identifying relevant drivers and by estimating for each driver thresholds that partition its domain into areas with similar impact on volatility. By doing so, nonlinear dependencies can be identified in a parsimonious fashion. We do, indeed, find highly nonlinear influences of financial drivers on volatility. This extends the existing research, which has primarily concentrated on linear volatility dynamics. Our results show that allowing for both macroeconomic information and the presence of nonlinear effects helps to assess future behavior of market volatility and to improve predictive performance.

One- through six-month-ahead out-of-sample forecasting applications to monthly log realized variance suggest that our boosting approach performs very favorable. For all the six forecasting horizons considered, the commonly-used GARCH and EGARCH benchmarks are clearly outperformed. What makes the approach appealing is the straightforward and systematic incorporation of exogenous risk drivers. Short-term forecasts also benefit when the exogenous drivers enter the model in a linear fashion. In the medium and long term, however, we gain accuracy by allowing volatility to react asymmetrically and with jumps. These findings confirm those of Engle et al. (2013), who report that the inclusion of macroeconomic variables improves long-run predictability of U.S. stock-return volatility, and are also compatible with those of West and Cho (1995) and Christoffersen and Diebold (2000), who, using information sets that do not contain macroeconomic variables, find that the quality of volatility forecasts decays quickly as the forecasting horizon grows. The empirical results presented

[15] With volatility being specified as $\sigma_t = e^{\eta_t/2}$, changes in η_t have a multiplicative effect on σ_t. If η_t decreases by 0.4, then, σ_t decreases by about 18% since $e^{-0.4/2} \approx 0.82$.

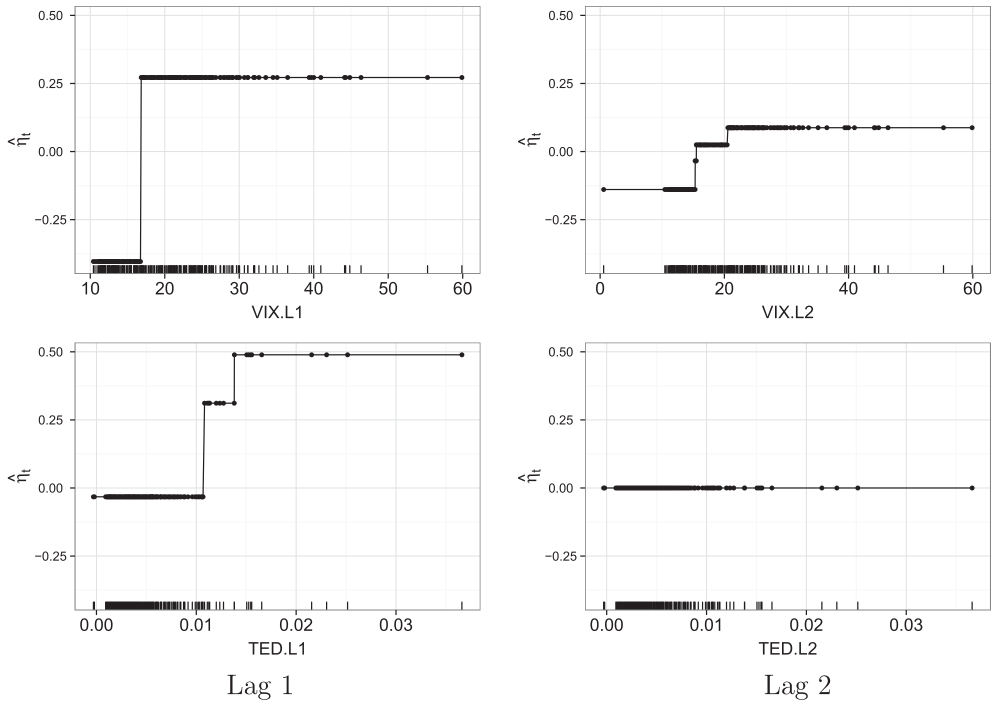


Fig. 3. The two most relevant predictors for the six-month-ahead S&P 500 volatility. Each row shows the estimated impact of the first and second lag of the volatility index (VIX) and TED spread, respectively.

here suggest that there is only a small set of—rather plausible—factors which primarily drive future S&P 500 volatility: volatility itself, captured in terms of the implied volatility index (VIX) and log realized variance (LRVar), the TED spread, and new orders of consumer goods and materials.

The application demonstrates that boosting is well suited for a unified framework for predictor selection and estimation in the context of volatility modeling. This is especially the case in the presence of many potential (and possibly highly dependent) risk drivers. An advantage of the approach is that it can cope with situations where we have—relative to the sample size—a large set of potential predictors. Apart from being useful in terms of variable selection and forecasting, models obtained via boosting can also provide a promising starting point for specifying nonlinear, parametric volatility models.

Acknowledgments

We would like to thank James Hamilton, Torsten Hothorn and Neil Shephard for invaluable discussions and feedback. We are grateful to the participants of the 18th International Conference Computing in Economics and Finance in Prague (2012), the SMU-ESSEC Symposium on Empirical Finance and Financial Econometrics in Singapore (2012), the Annual Meeting of the German Statistical Society in Vienna (2012), the 1st Vienna Workshop on High Dimensional Time Series in Macroeconomics and Finance 2013 and the Boston College/Boston University Econometrics Workshop 2013 for helpful comments. We are particularly thankful to two anonymous referees whose comments and suggestions helped to greatly improve the paper.

Appendix A. Implementation of componentwise boosting algorithm

In our implementation, we follow Friedman (2001) and shrink the coefficient towards zero. Shrinkage helps to dampen the "greediness" of the gradient technique, which may otherwise be prone to neglecting correlated predictor candidates, and "cures" the typical instability of forward selection methods (Breiman, 1996). The "appropriate" shrinkage, set by the shrinkage parameter ν, is empirically determined and can safely vary such $\nu \in [0.01, 0.3]$. The specific choice for ν has little effect on the final estimates, but rather affects the computational time. The updating in the m-th iteration is then given by $\hat{\eta}^{[m]} = \hat{\eta}^{[m-1]} + \nu \hat{f}^{[m]}_{\hat{s}_m}$, where $\hat{\eta}^{[m]} = (\hat{\eta}^{[m]}(\mathbf{z}_1), \ldots, \hat{\eta}^{[m]}(\mathbf{z}_T))^\top$ and $\hat{f}^{[m]}_{\hat{s}_m} = (\hat{f}^{[m]}_{\hat{s}_m}(z_{\hat{s}_m,1}), \ldots, \hat{f}^{[m]}_{\hat{s}_m}(z_{\hat{s}_m,T}))^\top$ are vectors of length T and $\hat{s}_m \in \{1, 2, \ldots, r\}$. Fitting the base learner in a given iteration modifies the gradient evaluation so that, with each iteration, covariates and gradients become increasingly orthogonal.

Without stopping, boosting with stumps will inevitably overfit, making the model useless for prediction. Hence, an appropriate stopping rule is essential. Note that the conditional observations $y_i|\mathbf{z}_i$ and $y_j|\mathbf{z}_j$, for $i \neq j, i, j \in \{1, \ldots, T\}$ are, by assumption, independent. We, therefore, determine the optimal number of boosting steps via bootstrapping by uniformly sampling with probability $1/T$ and with replacement from the observed data. Doing so, each sample makes use of roughly 64% of the original data for training, with the remaining, unselected data points used for evaluation. We repeat this twenty-five times for a large number of boosting steps and choose the step number that produces the lowest average out-of-sample loss.

To summarize, the boosting algorithm for volatility forecasting consists of the following steps:

1. Initialize function estimate $\hat{\eta}^{[0]}_t = \log\left(\frac{1}{T-1}\sum_{t=1}^T (y_t - \bar{y})^2\right)$, $\bar{y} = \frac{1}{T}\sum_{t=1}^T y_t, t = 1, \ldots, T$.
2. For all $z_{i,t} \in \mathbf{z}_t$, specify regression trees, $f_i(z_{i,t}) = \sum_{j=1}^J \gamma_{ij} I_{R_{ij}}(z_{i,t}), i = 1, \ldots, r$, using stumps, i.e., $J = 2$. Set $m = 0$.
3. Increase m by one.

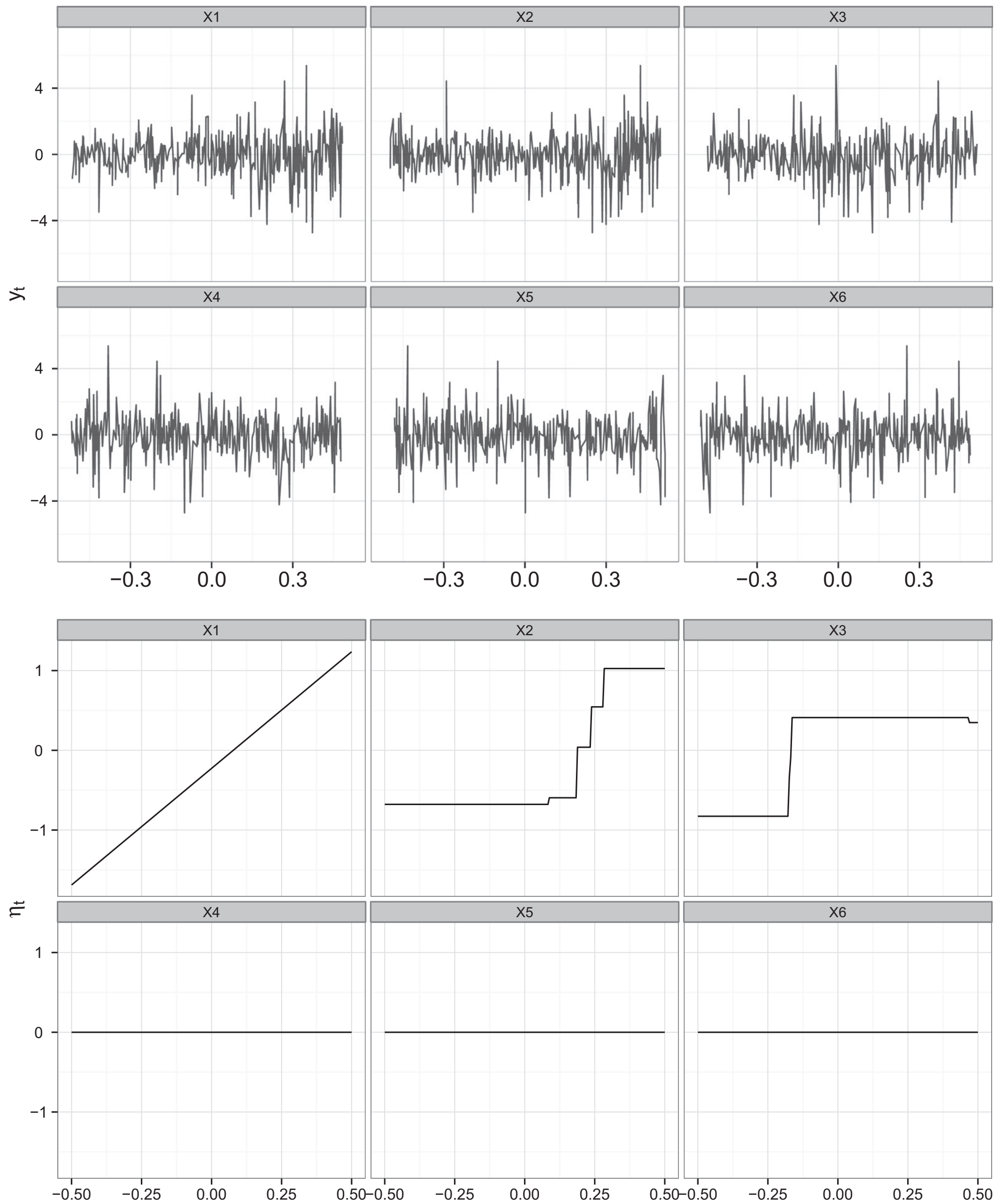


Fig. 4. Partial return components (upper half) simulated from (7), indicating how drivers X_1 through X_6 affect returns, and estimated partial log-volatility (lower half).

4.
   (a) Compute the negative gradient in (3) and evaluate $\hat{\eta}^{[m-1]}(\mathbf{z}_t), t = 1, \ldots, T$.
   (b) Estimate the negative gradient, using the stumps specified in Step 2. This yields r vectors, where each vector is an estimate of the gradient.
   (c) Select the base learner, $\hat{f}^{[m]}_{\hat{s}_m}, \hat{s}_m \in \{1, 2, \ldots, r\}$, that correlates most with the gradient according to the residual-sum-of-squares criterion.
   (d) Update the current estimate by setting $\hat{\eta}^{[m]} = \hat{\eta}^{[m-1]} + \nu \hat{f}^{[m]}_{\hat{s}_m}$, where ν is the shrinkage factor or the step size.
5. Repeat Steps 3 and 4 until the stopping condition applies.

Appendix B. Illustrating how boosting works

Before applying our boosting approach to forecasting volatility, we briefly demonstrate the principle ideas of the proposed method by conducting an illustrative simulation study. To do so, let the data generating process be given by

$$
\begin{aligned}
y_t &= \exp(\eta_t/2)\varepsilon_t \\
\eta_t &= 0.1 + 2x_{1,t} + 2\mathrm{I}_{[0.1,0.5]}(x_{2,t})x_{2,t} - 0.6\mathrm{I}_{[-0.5,-0.2]}(x_{3,t}) \\
&\quad + 0x_{4,t} + 0x_{5,t} + 0x_{6,t},
\end{aligned} \tag{7}
$$

with $\varepsilon_t \overset{iid}{\sim} N(0,1)$ and $x_{i,t}$ being the t-th observation of $X_i \sim U[-0.5, 0.5]$, $i = 1, \ldots, 6, t = 1, \ldots, T, T = 400$. Note that only the first three covariates affect volatility—the first linearly, the second linearly only for $X_2 \in [0.1, 0.5]$, and the third as a step function. The last three covariates, X_4 through X_6, do not contribute, and are included to check for robustness against false detection. We choose linear base learners for all but the second and third predictors, which are fitted with a regression-tree base learner, so that the second equation in (7) is fitted as

$$
\eta_t = \beta_0 + \beta_1 x_{1,t} + \sum_{j=1}^{J_2} \gamma_{2j} \mathrm{I}_{R_{2j}}(x_{2,t}) + \sum_{j=1}^{J_3} \gamma_{3j} \mathrm{I}_{R_{3j}}(x_{3,t}) + \beta_4 x_{4,t} + \beta_5 x_{5,t} + \beta_6 x_{6,t}, \tag{8}
$$

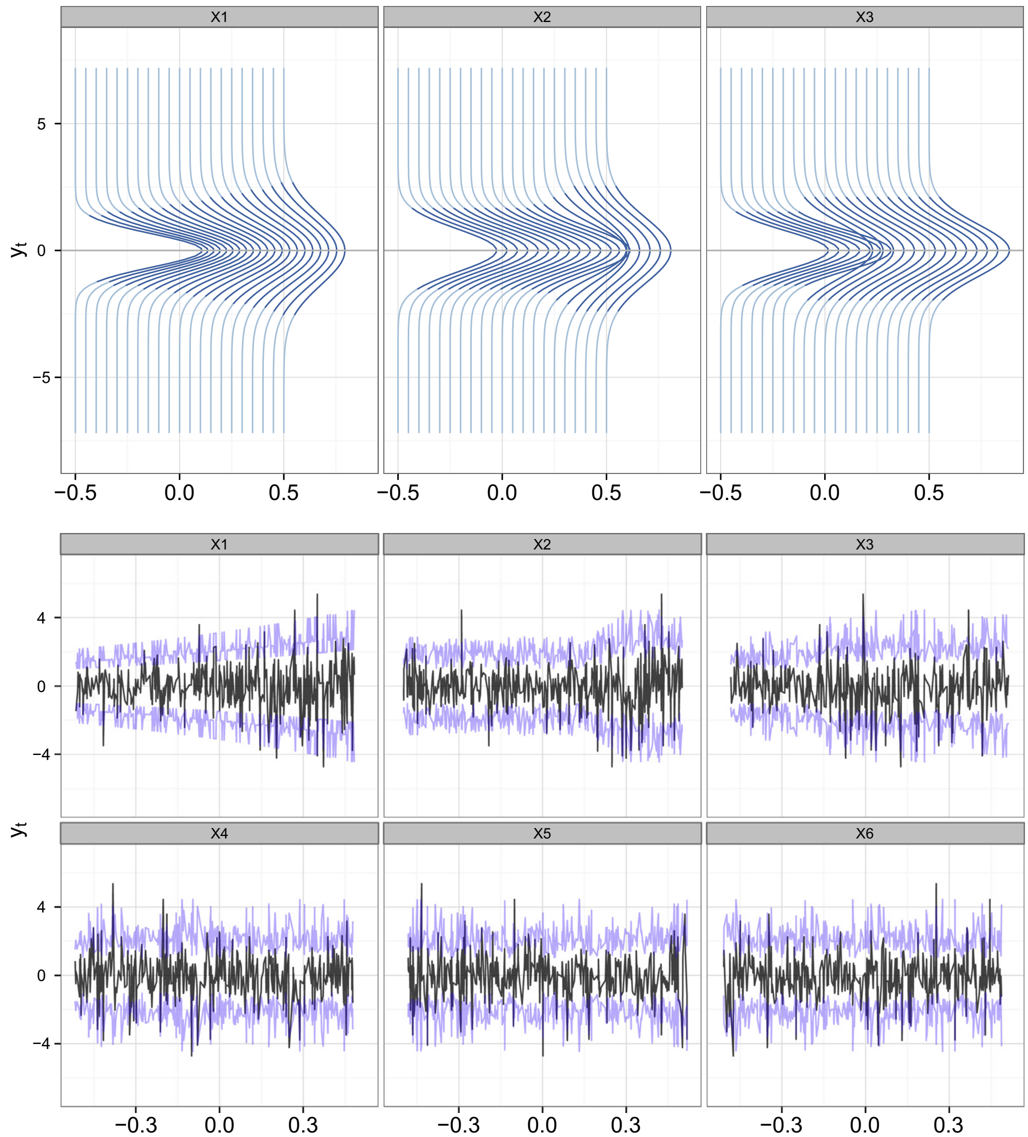


Fig. 5. Partial conditional density estimates (upper half) associated with X_1 through X_3 in (7). Dark segments indicate estimated 95% interquantile ranges, the lighter ones show the estimated tails. Simulated return components (lower half, darker lines) associated with these partial conditional densities; the lighter lines represent 95% interquantile ranges.

where R_{2j} and R_{3j} represent the estimated partitions of the domains of X_2 and X_3.

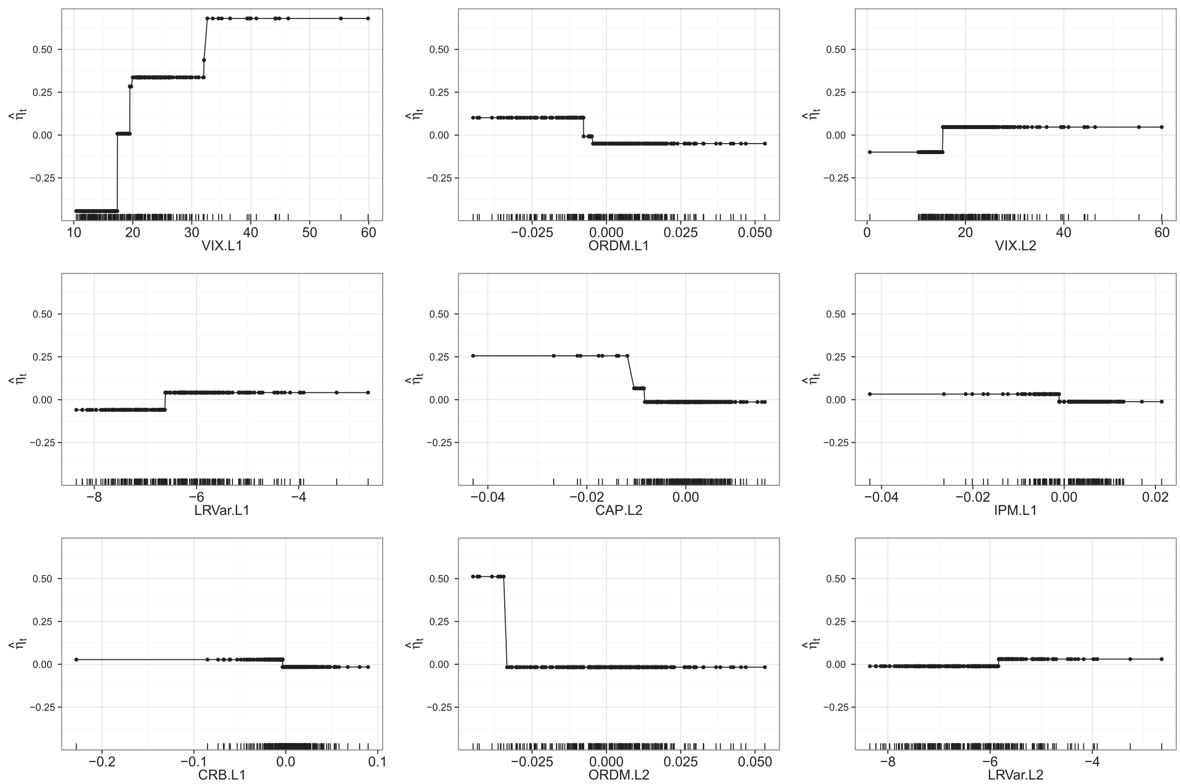
Ideally, the algorithm will recover the β and γ parameter values specified in (8). This means that X_4, X_5 and X_6 should not be selected at all, i.e., $\beta_4 = \beta_5 = \beta_6 = 0$, and that the domain of X_3 should be partitioned such that only $X_3 \in [-0.5, -0.2]$ affects volatility. Regarding X_2, although having linear impact for $X_2 \in [0.1, 0.5]$ and none otherwise, we intentionally chose an "incorrect" base learner by specifying a step function, in order to see whether the influence can still be adequately approximated.

Fig. 4 shows simulated, driver-specific return components (upper half) and the estimated partial impacts on volatility η_t (lower half). The influence of the underlying volatility drivers appears to be captured reasonably well. The parameter estimate $\hat{\beta}_1 = 1.463$ is low due to parameter regularization via early stopping. This is typical for shrinkage methods, where the parameter estimates usually have smaller magnitudes than unregularized solutions and the bias vanishes as the sample size increases. The advantage of early stopping is that, indeed, no redundant predictors are selected, i.e., $\hat{\beta}_4 = \hat{\beta}_5 = \hat{\beta}_6 = 0$. Furthermore, X_3 has the largest jumps near the right boundary of the interval $[-0.5, -0.2]$, and the linear impact for $X_2 \in [0.2, 0.5]$ is also captured, despite the moderate sample size chosen.

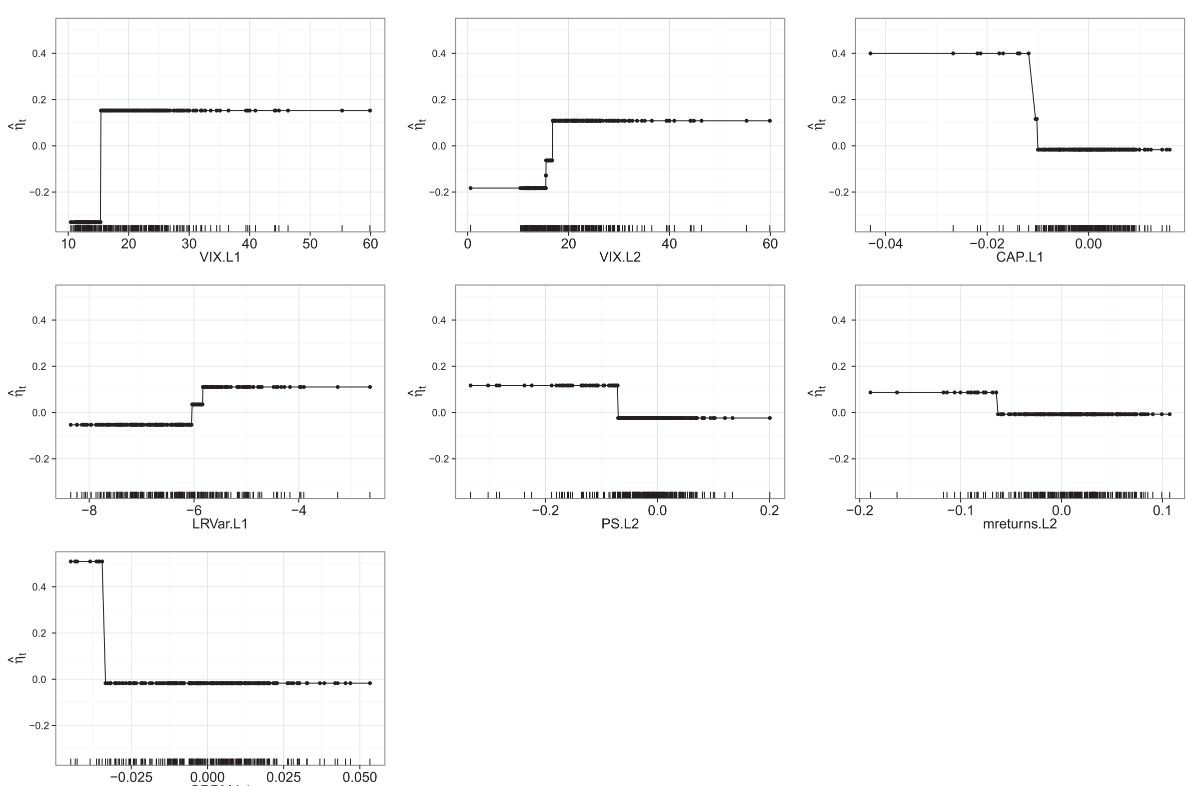
The results shown in Fig. 4 are typical in the sense that the variations in hundreds of repetitions were small. Translating the log scale in Fig. 4 back to standard deviations gives the estimate of the conditional density. Fig. 5 (upper half) shows the estimated partial densities associated with X_1, X_2 and X_3, with the central 95% interquantile ranges represented by the darker segments, and, in the lower half, simulated return components associated with these conditional densities. Visual inspection reveals that variations in volatility are closely captured, a finding that is supported by the fact that the estimates produce a coverage rate of 95.75% for the 95% interquantile range. The partial contribution of each driver is readily obtained in an interpretable way: an increase in X_1 causes the variance to grow proportionally; X_2 has an increased impact on the variance for $X_2 \in [0.1, 0.5]$; the variance contribution markedly decreases for $X_3 \in [-0.5, -0.2]$; and, with $\hat{\beta}_4 = \hat{\beta}_5 = \hat{\beta}_6 = 0$, the estimated conditional density of y_t remains, indeed, invariant with respect to X_4, X_5 or X_6.

Appendix C. Full list of volatility drivers

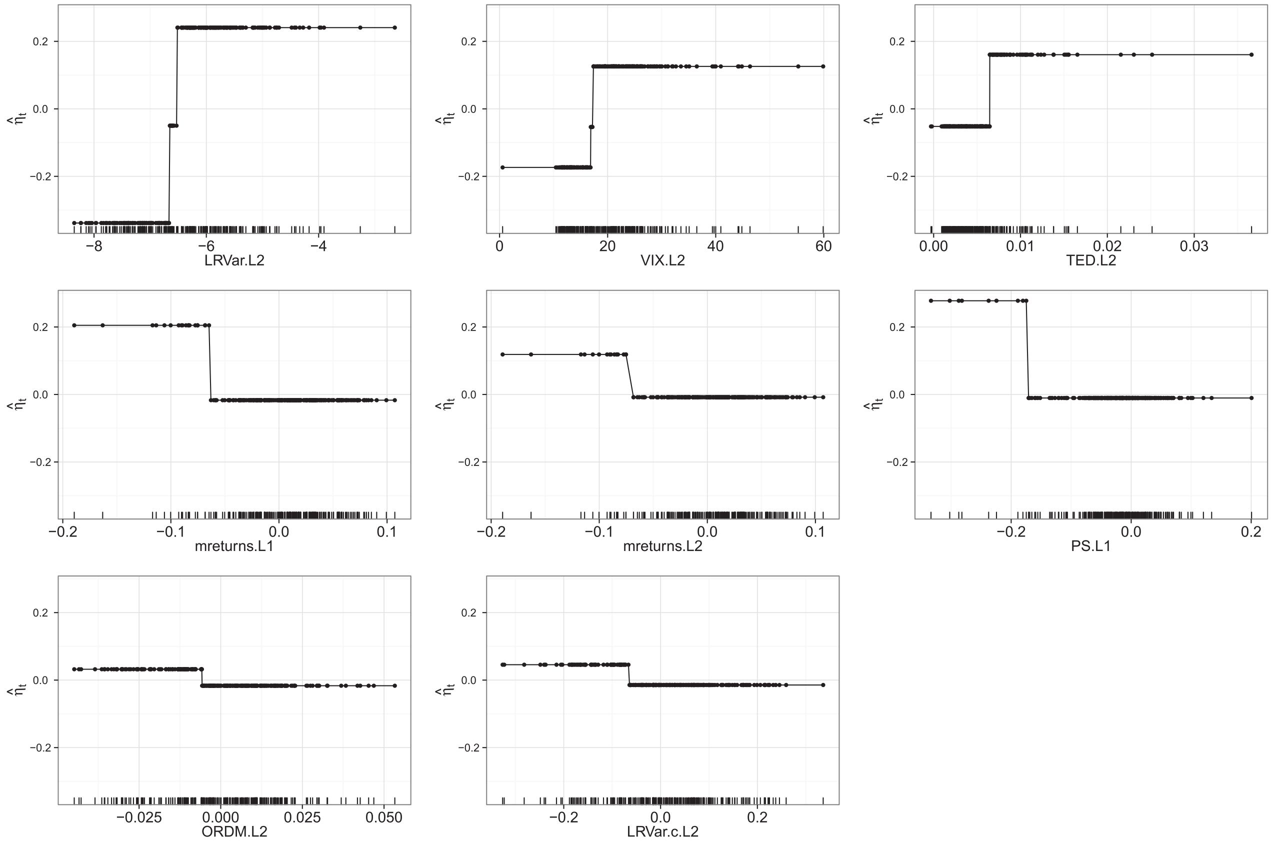
Horizon 1.



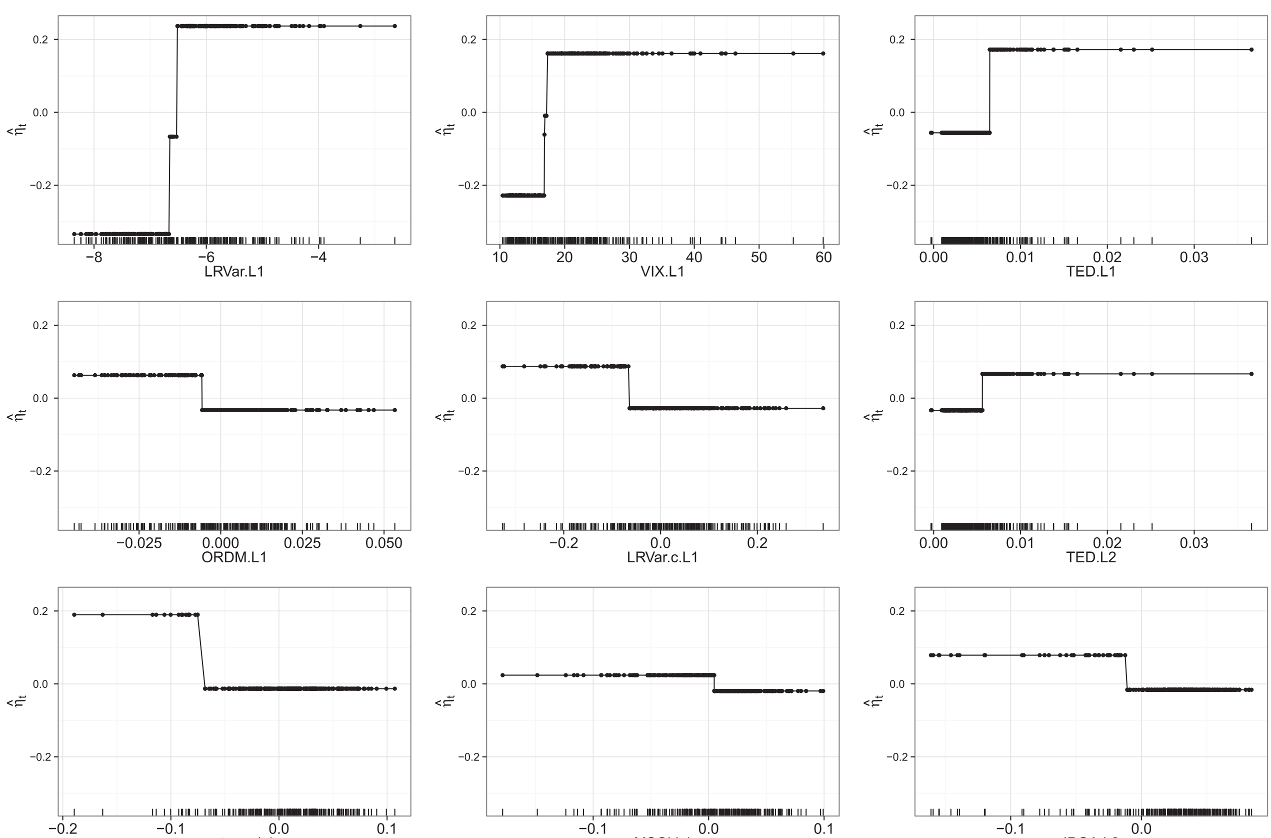
Horizon 2.



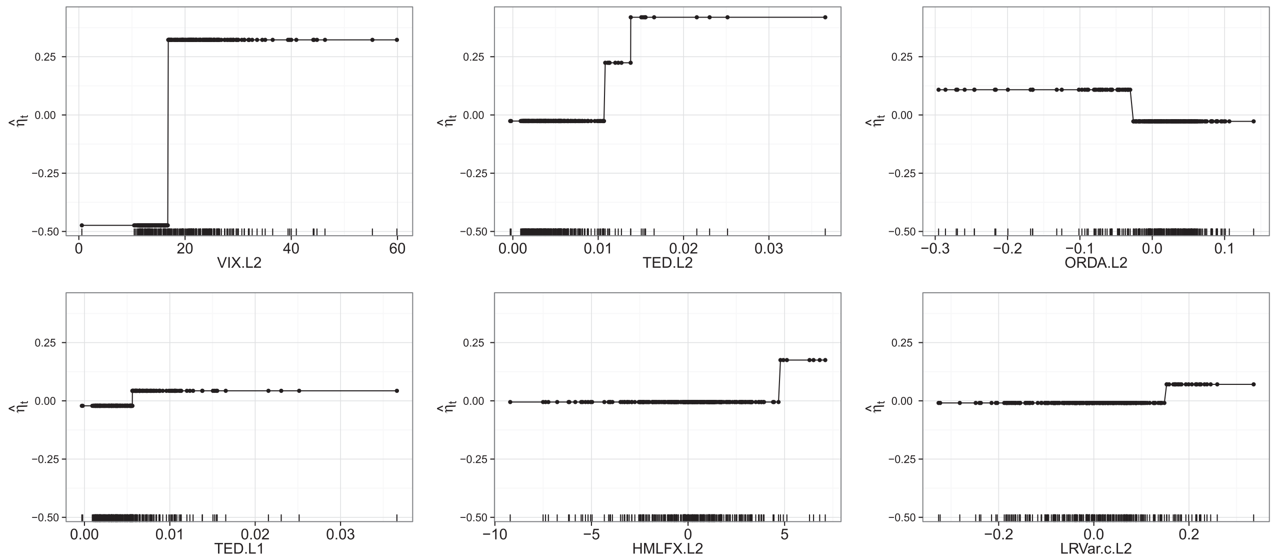
Horizon 3.



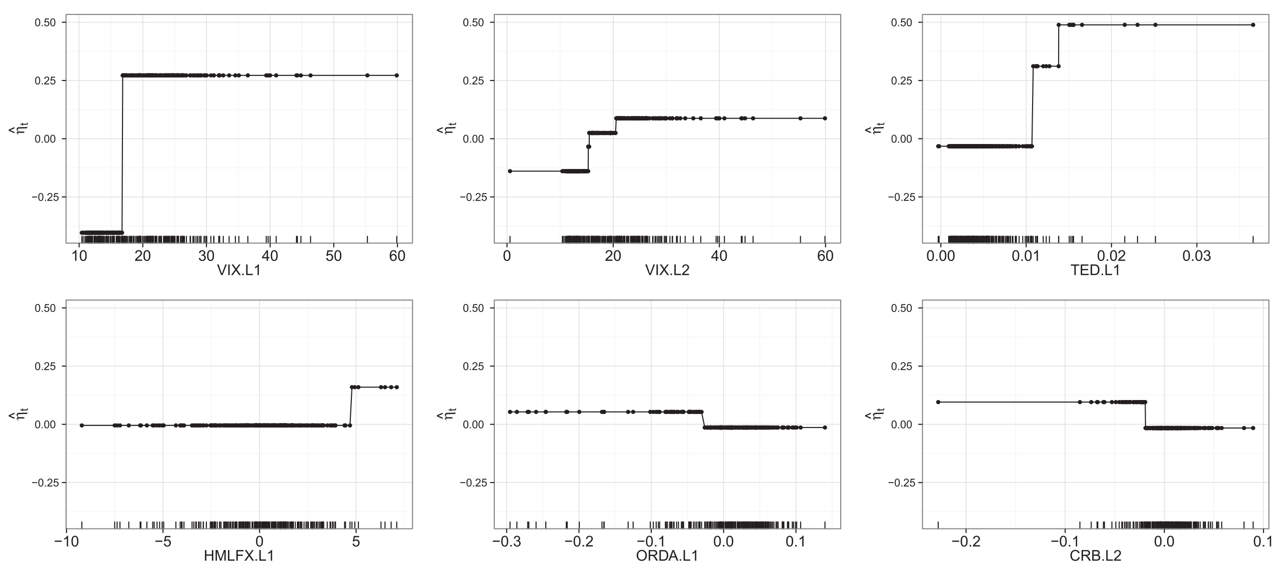
Horizon 4.



Horizon 5.



Horizon 6.



The list of all drivers and their abbreviations is shown in Table 1. The extensions *.L1 and *.L2 refer to the first and the second lag, respectively.

References

Andersen, T.G., Bollerslev, T., Christoffersen, P.F., Diebold, F.X., 2006. Volatility and correlation forecasting. In: Elliott, G., Granger, C., Timmermann, A. (Eds.), Handbook of Economic Forecasting, vol. 1. Elsevier, pp. 777–878, chap. 15.

Audrino, F., Bühlmann, P., 2003. Volatility estimation with functional gradient descent for very high dimensional financial time series. Journal of Computational Finance 6, 65–89.

Audrino, F., Bühlmann, P., 2009. Splines for financial volatility. Journal of the Royal Statistical Society Series B (Statistical Methodology) 71, 655–670.

Bai, J., Ng, S., 2009. Boosting diffusion indices. Journal of Applied Econometrics 24, 607–629.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Breiman, L., 1996. Bagging predictors. Machine Learning 24, 123–140.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Chapman and Hall, New York.

Brunnermeier, M., 2009. Deciphering the liquidity and credit crunch 2007–2008. Journal of Economic Perspectives 23, 77–100.

Bühlmann, P., 2006. Boosting for high-dimensional linear models. The Annals of Statistics 34, 559–583.

Bühlmann, P., Hothorn, T., 2007. Boosting algorithms: regularization, prediction and model fitting. Statistical Science 22, 477–505, with discussion.

Bühlmann, P., Yu, B., 2003. Boosting with the L_2 loss: regression and classification. Journal of the American Statistical Association 98, 324–339.

Canina, L., Figlewski, S., 1993. The informational content of implied volatility. Review of Financial Studies 6, 659–681.

Christensen, B.J., Prabhala, N.R., 1998. The relation between implied and realized volatility. Journal of Financial Economics 50, 125–150.

Christiansen, C., Schmeling, M., Schrimpf, A., 2012. A comprehensive look at financial volatility prediction by economic variables. Journal of Applied Econometrics 27, 956–977.

Christoffersen, P.F., Diebold, F.X., 2000. How relevant is volatility forecasting for financial risk management? The Review of Economics and Statistics 82, 12–22.

Claessen, H., Mittnik, S., 2002. Forecasting stock market volatility and the informational efficiency of the DAX-index options market. European Journal of Finance 8, 302–321.

Cochrane, J.H., Piazzesi, M., 2005. Bond risk premia. American Economic Review 95, 138–160.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13, 253–263.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least angle regression. The Annals of Statistics 32, 407–499 (with discussion, and a rejoinder by the authors).

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.

Engle, R.F., Rangel, J.G., 2008. The Spline-GARCH model for low-frequency volatility and its global macroeconomic causes. Review of Financial Studies 21, 1187–1222.

Engle, R.F., Ghysels, E., Sohn, B., 2013. Stock market volatility and macroeconomic fundamentals. The Review of Economics and Statistics 95, 776–797.

Flannery, M.J., Protopapadakis, A.A., 2002. Macroeconomic factors do influence aggregate stock returns. Review of Financial Studies 15, 751–782.

French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–29.

Freund, Y., Schapire, R., 1996. Experiments with a new boosting algorithm. In: Proceedings of the Thirteenth International Conference on Machine Learning Theory. Morgan Kaufmann, San Francisco, pp. 148–156.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. The Annals of Statistics 29, 1189–1232.

Giacomini, R., White, H., 2006. Tests of conditional predictive ability. Econometrica 74, 1545–1578.

Goyal, A., Welch, I., 2003. Predicting the equity premium with dividend ratios. Management Science 49, 639–654.

Hansen, P.R., Lunde, A., 2005. A forecast comparison of volatility models: does anything beat a GARCH(1,1)? Journal of Applied Econometrics 20, 873–889.

Harvey, D., Leybourne, S., Newbold, P., 1997. Testing the equality of prediction mean squared errors. International Journal of Forecasting 13, 281–291.

Hothorn, T., Hornik, K., Zeileis, A., 2006. Unbiased recursive partitioning: a conditional inference framework. Journal of Computational and Graphical Statistics 15, 651–674.

Hothorn, T., Bühlmann, P., Kneib, T., Schmid, M., Hofner, B., 2013. Model-Based Boosting, R Package Version 2.2-2.

Jiang, G.J., Tian, Y.S., 2005. The model-free implied volatility and its information content. Review of Financial Studies 18, 1305–1342.

Kearns, M., Valiant, L., 1994. Cryptographic limitations on learning Boolean formulae and finite automata. Journal of the Association for Computing Machinery 41, 67–95.

Kuester, K., Mittnik, S., Paolella, M.S., 2006. Value-at-Risk prediction: a comparison of alternative strategies. Journal of Financial Econometrics 4, 53–89.

Lustig, H., Roussanov, N., Verdelhan, A., 2010. Countercyclical Currency Risk Premia, NBER Working Papers 16427, National Bureau of Economic Research, Inc.

Lustig, H., Roussanov, N., Verdelhan, A., 2011. Common risk factors in currency markets. Review of Financial Studies 24, 3731–3777.

Maheu, J.M., McCurdy, T.H., 2002. Nonlinear features of realized FX volatility. The Review of Economics and Statistics 84, 668–681.

Matías, J.M., Febrero-Bande, M., González-Manteiga, W., Reboredo, J.C., 2010. Boosting GARCH and neural networks for the prediction of heteroskedastic time series. Mathematical and Computer Modelling 51, 256–271.

Mayr, A., Hofner, B., Schmid, M., 2012. The importance of knowing when to stop – a sequential stopping rule for component-wise gradient boosting. Methods of Information in Medicine 51, 178–186.

Menkhoff, L., Sarno, L., Schmeling, M., Schrimpf, A., 2011. Carry Trades and Global Foreign Exchange Volatility, Tech. Rep. 8291, C.E.P.R. Discussion Papers.

Mittnik, S., Semmler, W., 2015. Overleveraging, Financial fragility and the banking–macro link: theory and empirical evidence, Macroeconomic Dynamics (forthcoming).

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347–370.

Pastor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Paye, B.S., 2012. 'Déjà Vol': predictive regressions for aggregate stock market volatility using macroeconomic variables. Journal of Financial Economics 106, 527–546.

Prokopczuk, M., Wese Simen, C., 2014. The importance of the volatility risk premium for volatility forecasting. Journal of Banking & Finance 40, 303–320.

Robinzonov, N., Tutz, G., Hothorn, T., 2012. Boosting techniques for nonlinear time series models. AStA Advances in Statistical Analysis 96, 99–122.

Schapire, R., Freund, Y., Bartlett, P., Lee, W., 1998. Boosting the margin: a new explanation for the effectiveness of voting methods. The Annals of Statistics 26, 1651–1686.

Schwert, G.W., 1989. Why does stock market volatility change over time? Journal of Finance 44, 1115–1153.

Sensier, M., van Dijk, D., 2004. Testing for volatility changes in U.S. macroeconomic time series. The Review of Economics and Statistics 86, 833–839.

Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58, 267–288.

Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies 21, 1455–1508.

West, K.D., Cho, D., 1995. The predictive ability of several models of exchange eate volatility. Journal of Econometrics 69, 367–391.