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.

© 2015 Elsevier B.V. All rights reserved.

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/© 2015 Elsevier B.V. All rights reserved.

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 $y_t = \sigma_t \varepsilon_t, \varepsilon_t \overset{iid}{\sim} N(0,1)$ 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."

متن کامل مقاله
دریافت فوری

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات