International Journal of Forecasting 33 (2017) 848–863

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Volatility measures and Value-at-Risk

Dennis Bams[a,b], Gildas Blanchard[a], Thorsten Lehnert[c,*]

[a] *Limburg Institute of Financial Economics (LIFE), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands*
[b] *Open University Heerlen, P.O. Box 2960, 6401 DL Heerlen, The Netherlands*
[c] *Luxembourg School of Finance, University of Luxembourg, 4, rue Albert Borschette, L-1246 Luxembourg, Luxembourg*

ARTICLE INFO

Keywords:
Value-at-Risk
Option implied volatility
Volatility risk premium
Time-series
GARCH models

ABSTRACT

We evaluate and compare the abilities of the implied volatility and historical volatility models to provide accurate Value-at-Risk forecasts. Our empirical tests on the S&P 500, Dow Jones Industrial Average and Nasdaq 100 indices over long time series of more than 20 years of daily data indicate that an implied volatility based Value-at-Risk cannot beat, and tends to be outperformed by, a simple GJR-GARCH based Value-at-Risk. This finding is robust to the use of the likelihood ratio, the dynamic quantile test or a statistical loss function for evaluating the Value-at-Risk performance.

The poor performance of the option based Value-at-Risk is due to the volatility risk premium embedded in implied volatilities. We apply both non-parametric and parametric adjustments to correct for the negative price of the volatility risk. However, although this adjustment is effective in reducing the bias, it still does not allow the implied volatility to outperform the historical volatility models.

These results are in contrast to the volatility forecasting literature, which favors implied volatilities over the historical volatility model. We show that forecasting the volatility and forecasting a quantile of the return distribution are two different objectives. While the implied volatility is useful for the earlier objective function, it is not for the latter, due to the non-linear and regime changing dynamics of the volatility risk premium.

1. Introduction

Numerous empirical works have demonstrated the superiority of option-implied volatility (IV) over historical volatility models for predicting the future return volatility. We contribute to this literature by evaluating and comparing the merits of IV and historical volatility models for Value-at-Risk (VaR) forecasting. Volatility forecasting and VaR forecasting are two different objectives. Hence, this paper extends the comparison between IV and time series information into another field. Furthermore, we show that the best volatility forecast is not the best VaR forecast. The results of our multiple and complementary back-testing procedures show that the IV-based VaR cannot outperform a standard historical volatility VaR model.

The non-linear and regime changing dynamics of the volatility risk premium embedded in option prices explains the disappointing performance of IV for VaR forecasting. The risk neutral expected volatility implied from options differs from the physical volatility. Both IV-based volatility forecasts and VaR forecasts need to be corrected to address this difference. However, in the case of volatility prediction, a simple correction is sufficient to transform IV into the best forecast. On the other hand, the simple corrections applied in this paper do not allow the IV-based VaR to outperform the historical volatility model based VaR. The quantile prediction exercise concentrates on forecasting the occurrence of tail events. The complex dependence structure between the volatility risk premium and extreme

* Corresponding author.
E-mail address: thorsten.lehnert@uni.lu (T. Lehnert).

http://dx.doi.org/10.1016/j.ijforecast.2017.04.004

returns affects the quantile forecasting power of IV adjusted measures.

Despite earlier studies indicating the poor information content of IV (Canina & Figlewski, 1993), the great majority of volatility prediction literature has concluded that IV is either superior to historical volatility model forecasts (Blair et al., 2010; Christensen & Prabhala, 1998; Fleming, 1998), or at least complements it (Beckers, 1981; Day & Lewis, 1992). This conclusion has been reached based on a variety of markets, including: equity indices (Corrado & Miller, 2005; Yu et al., 2010), individual equities (Taylor et al., 2010), currencies (Charoenwong et al., 2009) and commodities (Szakmary et al., 2003). Only intraday time series volatility information appears to compete with IV information (Pong et al., 2004; Taylor & Xu, 1997).

Many of the studies in this field justify their investigations in the name of risk management (Frijns et al., 2010; Martens & Zein, 2004). However, the earlier literature concentrated on volatility forecasting rather than on its risk management applications. This paper evaluates whether the IV's superior volatility forecasting power translates into a superior VaR prediction relative to historical volatility models. More recently, a stream of studies has investigated the benefits of IV in the context of VaR forecasting (for a comprehensive survey, see Nieto & Ruiz, 2016). Despite this growing body of evidence, no consensus has been reached. Giot (2005) and Jeon and Taylor (2013) show that, for equity indices, a combination of IV and time series information results in superior VaR predictions. For currency markets, the findings of Chong (2004) and Christoffersen and Mazzotta (2005) suggest that IV models provide poor forecasts of the tail of the returns distribution.

We identify the volatility risk premium as a key component, which could explain the discrepancies in the conclusions reached. The relative size of the volatility risk premium across markets, such as equity and currency, is a distinctive factor. Furthermore, the various methodological choices provide implicit adjustments for the volatility risk premium. For instance, incorporating IV into the conditional volatility (quantile) equation, or aggregating IV with other forecasts, provides a source of non-transparent adjustment. Therefore, the IV-based VaR forecast performance is very sensitive to the methodological settings adopted in various investigations.

The empirical design adopted here differs from those of more recent studies that have been dedicated to the use of IV in a VaR forecasting exercise. First, we can account for and isolate the volatility risk premium affecting the IV-based forecasts formally. This approach provides transparent information and quantifies the impact of the volatility risk premium. It also allows us to distinguish clearly the relative benefits of standalone IV-based forecasts. Second, we perform a formal evaluation of the relative performances of IV models and historical volatility models. Previous studies on the topic have relied on traditional back-testing tests, which evaluate the individual forecasting performances but cannot establish the statistical superiority between two competing models.

IV-based forecasts, although efficient, have been documented to be biased (Lamoureux & Lastrapes, 1993; Szakmary et al., 2003; Yu et al., 2010). Chernov (2007) formally identifies the origin of this bias, namely the volatility risk premium (VRP). Since the price of volatility risk is negative, the risk neutral expectation of the volatility implied from options is higher than the physical expectation of that is volatility relevant for the VaR calculation. This feature has a negative effect on the predictive power of IV-based models (Tsiaras, 2009). Prokopczuk and Wese Simen (2014) were the first to formally acknowledge and account for the variance risk premium in an assessment of the forecasting performance of IV. They demonstrate empirically that a simple adjustment or correction of implied volatilities for the variance risk premium improves the forecasting performance relative to standard IV models both in and out-of-sample.

We account for the volatility risk premium in the same spirit as Prokopczuk and Wese Simen (2014), but apply the procedure in the VaR prediction context instead. Moreover, we focus on three major equity indices with publically available implied volatility indices, rather than the commodity market, which requires a proprietary option dataset. The behaviours of the IV and the volatility risk premium differ across markets and assets (Martin et al., 2009), meaning that it is relevant to evaluate the performances of adjusted IV forecasts on more mainstream markets, such as major equity indices.

We assess the information content, for the purpose of VaR measurement, of historical volatility models, IV and IV-adjusted for the volatility risk premium over an extended time period. We adjust for the volatility risk premium both non-parametrically and parametrically, and compare the results obtained. We measure the VaR for three major stock indices, S&P 500, DJIA and Nasdaq 100, using three implied volatility indices SPX (1990–2013), VXD (1997–2013) and VXN (2001–2013). We study the 95% and 99% confidence level VaR, as well as the one day out-of-sample and one month out-of-sample VaR. The one month VaR is included in this study in order to match the maturity of the implied volatility indices, even though it is not used in practice.

We evaluate the VaR models both individually and relative to each other. Individual performances are assessed using traditional backtesting tests. We use both the LR test of Kupiec (1995), which measures whether the number of exceptions is consistent with the confidence level, and the dynamic quantile test of Engle and Manganelli (2004). Relative forecasting performances are established using a VaR specific loss function. The loss function's statistical difference is obtained through a bootstrapping approach, as per Chen and Gerlach (2013), and a formal conditional predictability test proposed by Giacomini and White (2006).

The results from the volatility forecasting literature cannot be transposed to a quantile forecast application. Although IV information generally outperforms time series information for volatility prediction, our results suggest that such is not the case for VaR forecasting: IV and IV-adjusted VaR do not outperform GJR-GARCH VaR. In fact, under certain circumstances, the latter historical volatility model is even found to outperform all of the other models. The GJR-GARCH passes the LR and dynamic quantile tests, at the 5% significance level, more successfully than the other VaR models, including the IV-based