Computational Statistics and Data Analysis 56 (2012) 217–230

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

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

Modelling and forecasting noisy realized volatility[☆]

Manabu Asai[a], Michael McAleer[b,c,d], Marcelo C. Medeiros[e,*]

[a] Faculty of Economics, Soka University, Japan
[b] Econometric Institute, Erasmus School of Economics, Erasmus University, Rotterdam, The Netherlands
[c] Tinbergen Institute, The Netherlands
[d] Institute of Economic Research, Kyoto University, Japan
[e] Department of Economics, Pontifical Catholic University of Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 14 June 2010
Received in revised form 20 April 2011
Accepted 19 June 2011
Available online 6 July 2011

Keywords:
Realized volatility
Diffusion
Financial econometrics
Measurement errors
Forecasting
Model evaluation
Goodness-of-fit

ABSTRACT

Several methods have recently been proposed in the ultra-high frequency financial literature to remove the effects of microstructure noise and to obtain consistent estimates of the integrated volatility (IV) as a measure of ex post daily volatility. Even bias-corrected and consistent realized volatility (RV) estimates of IV can contain residual microstructure noise and other measurement errors. Such noise is called "realized volatility error". As such errors are ignored, we need to take account of them in estimating and forecasting IV. This paper investigates through Monte Carlo simulations the effects of RV errors on estimating and forecasting IV with RV data. It is found that: (i) neglecting RV errors can lead to serious bias in estimators; (ii) the effects of RV errors on one-step-ahead forecasts are minor when consistent estimators are used and when the number of intraday observations is large; (iii) even the partially corrected R^2 recently proposed in the literature should be fully corrected for evaluating forecasts. This paper proposes a full correction of R^2. An empirical example for S&P 500 data is used to demonstrate the techniques developed in this paper.

© 2011 Published by Elsevier B.V.

1. Introduction

Given the rapid growth in financial markets and the continual development of new and more complex financial instruments, there is an ever-growing need for theoretical and empirical knowledge of volatility in financial time series.

There is, however, an inherent problem in using models where the volatility measure plays a central role. The conditional variance is latent, and hence is not directly observable. It can be estimated, among other approaches, by the (Generalized) Autoregressive Conditional Heteroskedasticity, or (G)ARCH, family of models proposed by Engle (1982) and Bollerslev (1986) or stochastic volatility (SV) models (see, for example, Taylor, 1986). McAleer (2005) provides an exposition of a wide range of volatility models, and Asai et al. (2006) provides a review of the growing literature on multivariate SV models.

More recently, Andersen and Bollerslev (1998) have showed that ex post daily volatility is best measured by aggregating 288 squared five-minute returns. The five-minute frequency was suggested as a trade-off between accuracy, which is theoretically optimized using the highest possible frequency, and microstructure noise that can arise through the bid-ask bounce, asynchronous trading, infrequent trading, and price discreteness, among other factors.

☆ The authors are grateful to Marcel Scharth for efficient research assistance, and to an associate editor and two anonymous referees for helpful comments and suggestions. For financial support, the first author acknowledges the Japan Society for the Promotion of Science and the Australian Academy of Science, the second author acknowledges the Australian Research Council, the National Science Council, Taiwan, and the Japan Society for the Promotion of Science, and the third author wishes to acknowledge CNPq, Brazil.

* Corresponding author. Tel.: +55 21 3527 1078; fax: +55 21 3527 1084.
E-mail address: mcm@econ.puc-rio.br (M.C. Medeiros).

0167-9473/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.csda.2011.06.024

Ignoring the remaining measurement error, the ex post volatility essentially becomes "observable", and hence can be modelled directly, rather than being treated as a latent variable. Based on the theoretical results of Barndorff-Nielsen and Shephard (2002), Andersen et al. (2003) and Meddahi (2002), several recent studies have documented the properties of realized volatility constructed from high frequency data. However, it is well known that neglecting microstructure noise in calculating realized volatility (RV) can lead to biased and inconsistent estimates of integrated volatility (IV) as a true measure of daily volatility.

Several methods have recently been proposed in the ultra-high frequency financial literature to remove the effects of microstructure noise and to obtain consistent estimates of the IV (see Barndorff-Nielsen et al. (2008), Hansen et al., 2008, and Zhang et al. 2005). For an extensive review of the realized volatility literature, see McAleer and Medeiros (2008a) and Bandi and Russell (2007).

Nevertheless, even bias-corrected and consistent realized volatility estimates of the IV can contain residual microstructure noise and other measurement errors that should not be ignored. Furthermore, the consistency of the above-mentioned estimators is derived under some (strong) assumptions about the microstructure noise. Whenever some of these assumptions are not met in practice, the estimators turn to be inconsistent. Finally, if the number of intraday observations is small (due to illiquidity effects or data availability), the remaining measurement error may not be negligible. Barndorff-Nielsen and Shephard (2002) refer to such remaining noise as the "realized volatility (RV) errors". They suggested a method to estimate the continuous-time SV model, in which volatility follows a non-Gaussian Ornstein–Uhlenbeck (OU) process (see also Corradi and Distaso (2006) for a discussion of measurement errors and realized volatility).

The contribution of this paper is two-fold. First, we extend Barndorff-Nielsen and Shephard's (2002) approach and estimate three different models of IV. The common features between Barndorff-Nielsen and Shephard (2002) and the current paper is the use of state space representation to remove such RV errors. This paper deals with discrete-time SV models, in which the logarithm of IV follows a K-component model, a long memory model (ARFIMA), or a heterogeneous autoregressive (HAR) model. Our K-component model corresponds to the continuous-time SV model of Chernov et al. (2003). Monte Carlo simulation experiments are presented to investigate the effects of the RV errors on the estimators and forecasts of these three models. Second, we show that, in the presence of RV errors, the R^2 correction proposed by Andersen et al. (2005) is only a partial correction. We provide a corrected R^2 measure in Mincer–Zarnowitz regressions when the dependent variable is a noisy RV measure.

An empirical example is used to show that neglecting the RV error can lead to serious bias in estimating IV, and that the new method can eliminate the effects of the errors. Finally, the fully corrected R^2 proposed in this paper is needed in most cases.

The plan of the remainder of the paper is as follows. Section 2 discusses the effects of RV error on estimating and forecasting IV. Section 3 presents the results of Monte Carlo simulation experiments regarding the effects of RV error, using the K-component, long memory and HAR models. Section 4 proposes a new method to fully correct R^2 in the presence of RV error. The results of an empirical example are analysed in Section 5. Some concluding remarks are given in Section 6.

2. Realized volatility and the significance of measurement errors

Suppose that, along day t, the logarithmic prices of a given asset follow a continuous-time diffusion process:

$$\mathrm{d}p(t+\tau) = \mu(t+\tau)\mathrm{d}\tau + \sigma(t+\tau)\mathrm{d}W(t+\tau), \quad 0 \le \tau \le 1,\ t = 1, 2, \ldots,$$

where $p(t+\tau)$ is the logarithmic price at time $t+\tau$, $\mu(t+\tau)$ is the drift component, $\sigma(t+\tau)$ is the instantaneous volatility (or standard deviation), and $W(t+\tau)$ is a standard Brownian motion. In addition, suppose that $\sigma(t+\tau)$ is orthogonal to $W(t+\tau)$, such that there is no leverage effect. This assumption is standard in the realized volatility literature.

Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2002) showed that daily returns, defined as $r_t = p(t) - p(t-1)$, are Gaussian conditionally on $\Im_t \equiv \Im\{\mu(t+\tau-1), \sigma(t+\tau-1)\}_{\tau=0}^{\tau=1}$, the σ-algebra (information set) generated by the sample paths of $\mu(t+\tau-1)$ and $\sigma(t+\tau-1)$, $0 \le \tau \le 1$, such that

$$r_t|\Im_t \sim N\left(\int_0^1 \mu(t+\tau-1)\mathrm{d}\tau, \int_0^1 \sigma^2(t+\tau-1)\mathrm{d}\tau\right).$$

The term $\mathrm{IV}_t^2 = \int_0^1 \sigma^2(t+\tau-1)\mathrm{d}\tau$ is known as the *integrated variance*, which is a measure of the day-t ex post volatility. The integrated variance is typically the object of interest as a measure of the true daily volatility.

In general, $\sigma(t+\tau)$, or a function of $\sigma(t+\tau)$ such as $\sigma^2(t+\tau)$ or $\ln\sigma^2(t+\tau)$, is assumed to follow a continuous-time diffusion process (see Ghysels et al. (1996) for example). Integrating on τ, the Brownian motion of the diffusion process becomes a Gaussian variable, such that the integrated variance is a random variable. In this sense, IV_t^2 plays the same role as the stochastic variance in the class of "Stochastic Volatility (SV)" models. From this viewpoint, the connections among the integrated variance, stochastic variance, and conditional variance are clear. As shown by Nelson (1990), conditional variance models are approximations to continuous-time SV models. In the conditional variance model, the current variance is determined by past information sets, indicating that the approximation can be improved. Usually, continuous-time SV models are approximated by the Euler–Maruyama method, and the resulting models are called "discrete-time" SV models.

Download English Version:

https://daneshyari.com/en/article/416513

Download Persian Version:

https://daneshyari.com/article/416513

Daneshyari.com