13 Measuring and Modeling Risk Using High-Frequency Data

Wolfgang Härdle, Nikolaus Hautsch and Uta Pigorsch

13.1 Introduction

Volatility modelling is the key to the theory and practice of pricing financial products. Asset allocation and portfolio as well as risk management depend heavily on a correct modelling of the underlying(s). This insight has spurred extensive research in financial econometrics and mathematical finance. Stochastic volatility models with separate dynamic structure for the volatility process have been in the focus of the mathematical finance literature, see Heston (1993) and Bates (2000), while parametric GARCH-type models for the returns of the underlying(s) have been intensively analyzed in financial econometrics.

The validity of these models in practice though depends upon specific distributional properties or the knowledge of the exact (parametric) form of the volatility dynamics. Moreover, the evaluation of the predictive ability of volatility models is quite important in empirical applications. However, the latent character of the volatility poses a problem. To what measure should the volatility forecasts be compared to? Conventionally, the forecasts of daily volatility models, such as GARCH-type or stochastic volatility models, have been evaluated with respect to absolute or squared daily returns. In view of the excellent in-sample performance of these models, the forecasting performance, however, seems to be disappointing.

The availability of ultra-high-frequency data opens the door for a refined measurement of volatility and model evaluation. An often used and very flexible model for logarithmic prices of speculative assets is the (continuous-time) stochastic volatility model:

\[ dY_t = (\mu + \beta \sigma_t)dt + \sigma_t dW_t, \quad (13.1) \]
where $\sigma^2_t$ is the instantaneous (spot) variance, $\mu$ denotes the drift, $\beta$ is the risk premium, and $W_t$ defines the standard Wiener process. The object of interest is the amount of variation accumulated in a time interval $\Delta$ (e.g., a day, week, month etc.). If $n = 1, 2, \ldots$ denotes a counter for the time intervals of interest, then the term

$$\sigma^2_n = \int_{(n-1)\Delta}^{n\Delta} \sigma^2_t dt$$

is called the actual volatility, see Barndorff-Nielsen and Shephard (2002). The actual volatility is the quantity that reflects the market risk structure (scaled in $\Delta$) and is the key element in pricing and portfolio allocation. Actual volatility (measured in scale $\Delta$) is of course related to the integrated volatility:

$$V(t) = \int_0^t \sigma^2_s ds.$$  

(13.3)

It is worth noting that there is a small notational confusion here: the mathematical finance literature would denote $\sigma_t$ as “volatility” and $\sigma^2_t$ as “variance”, see Nelson and Foster (1994), for example.

An important result is that $V(t)$ can be estimated from $Y_t$ via the quadratic variation:

$$[Y_t]_M = \sum (Y_{t_j} - Y_{t_{j-1}})^2,$$

(13.4)

where $t_0 = 0 < t_1 < \cdots < t_M = t$ is a sequence of partition points and $\sup_j |t_{j+1} - t_j| \to 0$. Andersen and Bollerslev (1998) have shown that

$$[Y_t]_M \overset{p}{\to} V(t), \ M \to \infty.$$  

(13.5)

This observation leads us to consider in an interval $\Delta$ with $M$ observations

$$RV_n = \sum_{j=1}^{M} (Y_{t_j} - Y_{t_{j-1}})^2$$

(13.6)

with $t_j = \Delta\{(n - 1) + j/M\}$. Note that $RV_n$ is a consistent estimator of $\sigma^2_n$ and is called realized volatility. Barndorff-Nielsen and Shephard (2002) point out that $RV_n - \sigma^2_n$ is approximately mixed Gaussian and provide the asymptotic law of

$$\sqrt{M}(RV_n - \sigma^2_n).$$  

(13.7)

The realized volatility turns out to be very useful in the assessment of the validity of volatility models. For instance, reconciling evidence in favor of the forecast accuracy of GARCH-type models is observed when using realized