Intraday and Interday Time-Zone Volatility Forecasting

Petko S. Kalev*

Department of Accounting and Finance

Monash University

23 October 2006

Abstract

The paper develops a global volatility estimator and forecasts market uncertainty. The study utilizes tick-by-tick data of equity indices, globally traded stocks and commodity futures from different time zones. We aim to produce an efficient and robust risk estimator to sudden changes of the price movements and price discontinuities. The global volatility estimator takes into account the history of price dynamics from all three main time zones. The results of the study - projected global and local one-day ahead, one-week ahead and up to one-month ahead volatilities - are crucial for making optimal investment allocations and informed decisions.

JEL classification: C14, G15, and C53

* Corresponding author: Petko Kalev, Department of Accounting and Finance, Faculty of Business and Economics, P.O.Box 197, Caulfield East, VIC 3145 Australia. Tel: (+613) 9903 2431 Fax: (+613) 9903 2422 Email: Petko.Kalev@BusEco.Monash.edu.au.

I. Introduction

The paper proposes a uniform measuring, modelling and forecasting of the global market volatility by utilizing the highest possible frequency data (continuously observed prices) of securities from all three time zones: Asia-Pacific; Western Europe and the East Coast of the USA. We argue that by using tick-by-tick (intraday) data from financial markets within all three time-zones (around the globe), a volatility estimator, built on daily aggregation of the intraday variability, will produce superior out-of-sample forecasts to both the observed global volatility estimate and to the volatility estimates from any of the individual time zones.

Volatility is central to risk management, yet it is unobservable. Various approaches have been adopted to measure, model and forecast volatility (see Engle (1982), Bollerslev (1986), among others). Back in 1980, Merton argued that arbitrarily accurate estimates of volatility can be obtained through progressively higher sampling of market returns. Nowadays, modelling high frequency financial data plays an important role in analysing both price discovery and managing risk. It offers an optimal basis for asset pricing, risk and portfolio management, and derivatives valuation.

The study builds on a rapidly growing literature of realized volatility (RA thereafter), see among others Andersen and Bollerslev, 1998; Anderson et al., 2001, 2002, 2003; Barndorff-Nielsen and Shepard, 2002a, 2002b, 2003a, 2003b; Bollen and Inder, 2002. The RA is a simple and easy to estimate non-parametric measure. The RA estimator is computed as the sum of intraday sampled squared relative prices, where these returns are sampled usually at very short intervals and most of the time, the interval is set to five minutes. We are interested in financial markets other than the currency market, where quotes are observed 24 hours seven days per week. Our aim is to develop a global volatility estimator and global volatility forecasting framework with applications to equity and futures commodities markets. For each time zone, firstly we consider the main stock market indices, then we concentrate on stocks that are traded globally (known as global stocks), and finally, we model the intraday price volatility of crude oil futures and wheat futures traded around the clock. Tick-by-tick data, transaction prices together with the best bid and ask price and traded volume are provided by Reuters.

To forecast the price volatility of homogenous securities, traded in different time zones, we extend the MIxed DAta Sampling (MIDAS) regression framework recently developed by Ghysels, Santa-Clara and Valkanov (2005, 2006) (GSV thereafter). We evaluate the performance of our out-of-sample projections, based on intraday (hourly) periods, and also interday – that is: for one-day ahead; one-week ahead; two-week ahead; three-week ahead and; one-month ahead horizons (where one month is assumed to be equal to 22 working days). We also investigate whether the forecasts, generated by our model and corrected for price discontinuities (volatility jumps), outperform alternative volatility forecasts, in particular, daily historical volatility; implied volatility forecasts derived from option prices; and GARCH-J (GARCH which is augmented for jumps) type models. In addition, we compare the relative forecasting performance of the realized volatility as defined in Andersen and Bollerslev (1998) and Anderson's et al. with that of the VARHAC realized volatility of Bollen and Inder (2002). The latter estimator makes use of all transaction data and hence one can argue it is more efficient than the sum of the squared logarithmic returns (realized volatility).

The current paper contributes to the literature as follows.
First, we consider modelling and forecasting of the realized volatility of futures (options) instruments from major financial derivates market representatives in the three time zones. The three biggest financial centres in the world are Asia-Pacific (Tokyo, Singapore, Sydney); Western Europe (London & Frankfurt) and; United States – East Coast (Chicago & New York).[1] This adds to the vast literature of globalization of capital markets and improves the understanding of the asset prices

[1] To the best of our knowledge, volatility modelling of securities from these time zones: Asia-Pacific; Western Europe and the East Coast of the USA, has not been previously a subject of a robust, efficient and uniform investigation.

volatility structures' and correlations' and as well as the volatility transmission in a global trading environment.

Second, our project will produce a more robust and efficient estimator of price volatility for the one-day ahead; one-week ahead; two-weeks ahead; three-weeks ahead and; one-month ahead horizons. Risk managers and derivatives traders will have more reliable volatility estimates to start their trading day when a particular financial time zone opens for trading. Next, consistent with Aït-Sahalia, Mykland and Zhang (2005), we develop an optimal sampling frequency framework for realized volatility. The latter approach contrasts to the majority of studies (Andersen and Bollerslev, 1998; Anderson et al., 2001, 2002, 2003; Barndoerff-Nielsen and Shepard, 2002a, 2002b, 2003a, 2003b), which work only with equally spaced five minute frequencies. We also consider the realized range (high-low range) estimator, which is found to be superior to the realized variance in the presence of bid-ask bounce and infrequent trading (Martens and van Dijk, 2006). In addition, we will employ and modify the VARHAC realized volatility estimator of Bollen and Inder (2002), which is based on tick-by-tick data (all transacted prices), and compare the global VARHAC realized volatility estimator and the realized volatility estimator based on equally spaced *m*-minute frequencies. Finally, we contribute to the literature by extending the estimation and evaluation framework of multivariate forecasting models for logarithmic realized volatilities. In order to build a global (time-zone) volatility estimator, we utilize the intraday information and extend the MIDAS framework to forecast future uncertainty for either any of the geographical time zones or for the global market.

The rest of the paper is organized as follows. Section II introduces the data and methodology employed in this study. Section III **will** summarize the results. Section IV **will** conclude.

II Data and Methodology

A. Data

The data for this project are provided by the Securities Industry Research Centre of Asia-Pacific (SIRCA). The data sample consists of intraday transaction and quote data (tick-by-tick transaction data and midpoint of the best bid-ask quotes). SIRCA has been collecting tick-by-tick transaction data from the Reuters screen since January 1996. We note that historical data on global stocks are available since November 1998 when the first global stock DaimlerChrysler AG was officially registered.

Data are collected for the following financial instruments, time zones and regional markets:

1). **Equity Indices - Spot and Options on Futures Data**: Zone One: Nikkei 225 (Tokyo); All Ordinaries (Australia); Hang Seng (Hong Kong) and KOSPI 200 (South Korea); Zone Two: DAX (Germany); FT-SE 100 (UK); CAC 40 (France) and; Zone Three: S&P 500 Composite Index (USA); TSE 300 Composite Index (Canada)
2). **Global Shares - Spot and Options on Futures Data**: DaimlerChrsyler AG, Nokia and UBS, all three are traded on the European stock markets, USA and Tokyo;
3). **Crude Oil Futures and Options on Futures Data**: Zone One: SIMEX (Singapore); Zone Two: IPE (London) and; Zone Three: NYMEX (New York) and;
4). **Wheat Futures and Options on Futures Data**: Zone One, China Zhengzhou Commodity Exchange; Zone Two: LIFFE and MATIF and; Zone Three: CBOT, Minneapolis Grain Exchange and Mid American Commodity Exchange.

B. Theoretical Foundation

We define the logarithmic price of an asset with p(t). A continuous time jump diffusion process, which is a starting point in asset pricing literature, can be expressed in stochastic differential equation form as follows:

$$d\,p(t) = \mu(t)\,d(t) + \sigma(t)\,dW(t) + \kappa(t)\,dq(t), \qquad 0 \le t \le T. \qquad (1)$$

where $\mu(t)$ denotes the continuous and locally bounded variation process, $\sigma(t)$ is the stochastic volatility (assumed to be strictly positive), $W(t)$ is the standard Brownian motion, $\kappa(t)$ denotes the size of the corresponding jumps, while $dq(t)$ is a jump counting process which takes value of one at the time of a jump and zero otherwise.

In continuous time the cumulative price process is given by r(t) = p(t) – p(0). The quadratic variation or notional variance process is given as follows:

$$[r,r]_t = \int_0^t \sigma^2(s)ds + \sum_{0<s\leq t} \kappa^2(s) . \qquad (2)$$

We are interested in modelling and forecasting volatility while utilizing intraday price information from securities traded on more than one market and more than one (two) time zone. Initially, we define the necessary volatility proxies and measures for a single market with m-equally spaced trading intervals.

Next, we extend the methodology of volatility modelling and forecasting considering all securities traded in a particular time zone, and finally, considering all assets traded globally. When a particular financial market operates, the price series is sampled discretely over m equally spaced intraday intervals. The (1/m)-period intraday return series is denoted with $r_{t,1/m}$, where $r_{t,1/m} = p(t) - p(t - 1/m)$.[2] Following Andersen and Bollerslev (1998), the daily realized volatility is the summation of all (1/m)-period squared returns

$$RV_{t+1,1}(1/m) = \sum_{j=1}^{m} r^2_{t+j/m,1/m}. \qquad (3)$$

In the proposed project, we also consider an alternative RV estimator. Bollen and Inder (2002) introduce the VARHAC volatility estimator ($RV_{BI,t}$). This estimator makes use of all transaction data, and hence, one can argue it is more efficient than the sum of the squared logarithmic returns. The realized volatility - either RV_t or $RV_{BI,t}$ - is, however, consistent with the integrated volatility only in the absence of discontinuities in the stochastic process (Barndorff-Nielsen and Shephard, 2002a, b). The realized bi-power variation (BV_t) measure (Barndorff-Nielsen and Shephard, 2004), defined as summation of the adjacent absolute log-returns, is as follows

$$BV_{t+1,1}(1/m) = \pi/2 \times \sum_{j=2}^{m} | r_{t+(j)/m,1/m} | \, | r_{t+(j-1)/m,1/m} |, \qquad (4)$$

[2] For notation simplicity we normalize the daily time interval to unity.

The BV_t is a consistent estimator of the integrated volatility, and hence, can be used to estimate the empirical discontinuities and jumps (Barndorff-Nielsen and Shephard, 2004). After imposing a non-negative truncation, the empirical jumps are measured as

$$J_{t+1,1}(1/m) = \max [\, RV_{t+1}(1/m) - BV_{t+1}(1/m), 0 \,]. \tag{5}$$

C. Building the Global Volatility Estimator

The first stage of building a global volatility estimator consists of calculating the optimal length of sampling of (1/m)-period intraday returns. In order to overcome the problem of market microstructure noise and arrive to an optimal interval length of the intraday sampled return series, we adapt the approach developed by Aït-Sahalia et al. (2005). We restrict 1996, the first year of data available for any of the securities in our sample, to be a learning (pre-model estimation) period. We calculate the optimal sampling intraday interval length for any security and time zone: $1/m_1$, $1/m_2$ and $1/m_3$, accordingly.

The second stage is to calculate the realized volatility, realized bi-power variation, implied volatility and other standard volatility measures. We estimate several models of volatility for the time period of 1000 trading days (approximately, for the period 1997-2001). We obtain daily volatility estimates from any individual time zone $RV_{t+1,1}(1/m_j)$ (where j=1, 2 & 3). The global volatility estimator is therefore defined as follows

$$GRV_{t+1,1}(1/max) = \sum_{j=1}^{3} RV^{*}_{t+1,j}(1/max), \tag{6}$$

where $RV^{*}_{t+1,j}(1/max)$ stands for the non-overlapping realized volatilities from all three time zones and *max* is the maximum sampling length of three time zone intervals. The three time zones are: time zone 1 (TZ1) - Asia-Pacific; time zone 2 (TZ2) - Western Europe and time zone 3 (TZ3) - the East Coast of the USA. The global volatility estimator is calculated according to the clock motion from open-to-close return series. When there is an overlap, like the one between TZ2 and TZ3, preference will be given to the zone where the more liquid market operates. The liquidity is measured by the trading turnover of securities, as measured in 1996.

However, if we are interested in a local j-zone volatility estimator, then the full price history of securities traded on the financial market associated with the j-th zone will be considered. Recall that, the realized volatility (the sum of the squared logarithmic returns) is computed from return series for any asset from any time zone. We also consider the realized range (high-low range) based on 15 (30) minutes intervals as suggested by Martens and van Dijk (2006).

D. Methods of Volatility Modelling and Forecasting

We adapt the MIDAS regression by Ghysels, Santa-Clara and Valkanov (2005, 2006) as our main forecasting tool (model). By definition, mixed regression (MIDAS) models specify conditional expectations as a distributed lag of regressors

$$Y_{t+H,t} = \mu_H + \Phi_H \, f\,\{b_H(k,\theta), X_{t-k+H}\} + \varepsilon_{t+H}, \qquad (7)$$

where $Y_{t+H,t}$ is the measure of the actual volatility observed at daily, one-week, two-weeks, three-weeks and one-month periods, f{.} is a function of lag polynomials with lag coefficients (weights) $b_H(k,\theta)$ and independent variable(s) X_{t-k+H}. The MIDAS regression framework models price volatility in a uniform and robust manner (Ghysels et al. 2005a). This framework provides extra flexibility – the volatility measure on the left hand side, $Y_{t+H,t}$, might be sampled on lower frequencies than the variables from the right hand side, which are sampled at various higher frequencies. For any time horizon, H, we consider as a dependent variable the global realized volatility, GRV_t, and/or the realized j-th zone volatility, $GRV_{t+1,1}(1/m_j)$. The proposed regressors include: the previous day (period) GRV_{t-1}; the squared return series; the absolute return series; and the realized power (sum of high frequency absolute returns). We also consider the daily volatility estimator, ($RV_{BI,t}$) of Bollen and Inder (2002), being independent variable. Consistent with Ghysels et al. (2004b), we estimate the MIDAS regression by the method of quasi-maximum likelihood.

Alternatively, we model the global volatility by utilizing the GARCH(1,1) model. We augment the variance equation of the GARCH model to account for jumps (and/or discontinuities), where the extended GARCH is defined as

$$r_t = \mu + \varepsilon_t \text{ , where } \varepsilon_t \mid \Omega_{t-1} \sim \text{iid } (0,\sigma^2_t) \qquad (8)$$

$$\sigma^2_t = \omega + \alpha\, \varepsilon^2_{t-1} + \beta\, \sigma^2_{t-1} + \lambda\, J_{t-1}\ , \tag{9}$$

where J_{t-1} the jump time series of a security traded on a single time zone (preferably the most recently operating market) or the global jump proxy $GJ_{t-1}(1/max)$, which is the global equivalent of the global volatility estimator GRV $_{t+1}(1/max)$. Consistent with Martens and Zein (2004), we augment the variance equation of the GARCH(1,1) model with the implied volatility, $IV_{t,T}$, computed from call (put) options on the futures prices.

$$\sigma^2_t = \omega + \alpha\, \varepsilon^2_{t-1} + \beta\, \sigma^2_{t-1} + \lambda\, IV_{t-1}\ . \tag{10}$$

When fitting econometrics models, such the MIDAS or GARCH models, to return series, and particularly, those built on high frequency data, it is often observed that model residuals still tend to be heavy tailed. Researchers have imposed distributional restrictions on the error term, for instance, assuming that the residuals are Student's *t* or Generalized Error Distributed (Nelson, 1991). We adapt the Mittnik et al. (2002) approach and model both the time series mixed model (equation, 7) and the GARCH(1,1) model under stable Paretian distributions (McCulloch, 1996).

E. Models Evaluation

To evaluate the out-of-sample time series volatility forecasts, we follow the popular approach of correlating actual data (historical realized volatilities) to the forecasts from our models. We consider the following OSL of realized volatility from day t+1 to t+H, $RV_{t+1,H}$, on a constant and the implied volatility, $IV_{t+1,H}$, and time series forecast, $TS_{t+1,H}$,

$$RV_{t+1,H} = \beta_0 + \beta_1\, IV_{t+1,H} + \beta_2\, TS_{t+1,H} + \varepsilon_{t+1,H}\ , \tag{11}$$

where the out-of-sample forecasting horizon, H, takes values 1, 5, 10, 15 and 22 (the forecast are for one-day ahead, one-week ahead, two-, three-weeks ahead and one-month). In addition, we evaluate the performance of the out-of-sample forecasted volatility for the period of 2002 to 2005, comparing the Root Mean Squared Error (RMSE), which is estimated in a consistent heteroskedasticity manner (Martens and Zein, 2004).

REFERENCES

Aït-Sahalia, Y., P. A. Mykland and L. Zhang (2005). How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise, Review of Financial Studies, 18, 351-416

Andersen, T.G. and T. Bollerslev (1998). Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, International Economic Review, 39, 885-905.

Andersen, T.G., T. Bollerslev and F.X. Diebold (2004). Parametric and Non-Parametric Volatility Measurement, in Handbook of Financial Econometrics (L.P Hansen and Y. Aït-Sahalia, eds.). Elsevier Science, New York.

Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001). The Distribution of Realized Stock Return Volatility, Journal of Financial Economics, 61, 43-76.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003). Modeling and Forecasting Realized Volatility, Econometrica, 71, 579-625.

Barndorff-Nielsen, O.E. and N. Shephard (2002a). Estimating Quadratic Variation Using Realized Variance, Journal of Applied Econometrics, 17, 457-477.

Barndorff-Nielsen, O.E. and N. Shephard (2002b). Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models, Journal of Royal Statistical Society, Series B, 64, 253-280.

Barndorff-Nielsen, O.E. and N. Shephard (2003a). Realised Power Variation and Stochastic Volatility Models, Bernoulli, 9, 243-265.

Barndorff-Nielsen, O.E. and N. Shephard (2003b). How Accurate is the Asymptotic Approximation to the Distribution of Realised Volatility, in Identification and Inference for Econometric Models. A Festschrift for Tom Rothenberg, (edited by Don Andrews, Jim Powell, Paul Ruud and Jim Stock), Econometric Society Monograph Series. Cambridge: Cambridge University Press.

Barndorff -Nielsen, O.E. and N. Shephard (2004). Power and Bipower Variation with Stochastic Volatility and Jumps, Journal of Financial Econometrics, 2, 1-37.

Bollen, B. and B. Inder (2002). Estimating Daily Volatility in Financial Markets Utilizing Intraday Data, Journal of Empirical Finance, 9, 551-562.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.

Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, Econometrica, 45, 987-1007.

Ghysels, E., P. Santa-Clara and R. Valkanov (2005). There is a Risk-Return Tradeoff After All, Journal of Financial Economics, 76(3), 509-548.

Ghysels, E., P. Santa-Clara and R. Valkanov (2006). Predicting Volatility: Getting the Most out of Return Data Sampled at Different Frequencies, Journal of Econometrics, 131(1-2), 59-95.

Maldelbrot, B. (1963). The Variation of Certain Speculative Prices, Journal of Business, 36, 394-419.

Martens, M. and D. van Dijk (2006). Measuring Volatility with the Realized Range, Journal of Econometrics, forthcoming.

Martens, M. and J. Zein (2004). Predicting Financial Volatility: High-frequency Time-Series Forecast vis-à-vis Implied Volatility, Journal of Futures Markets, 24(11), 1005-1029.

McCulloch, J. H. (1996), Financial Applications of Stable Distribution, In: Maddala, G., and C. Rao (Eds.), Handbook of Statistics, Vol. 14, Amsterdam: Elsevier.

Melton, R. C. (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation, Journal of Financial Economics, 8, 323-361.

Mittnik, S., M. S. Paolella, and S. T. Rachev (2002). Stationarity of stable power-GARCH processes, Journal of Econometrics, 106, 97-107.

Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns; A New Approach, Econometrica, 59, 347-344.