*International Journal of Business and Economics*, 2002, Vol. 1, No. 1, 17-24

# Intra-Day Features of Realized Volatility: Evidence from an Emerging Market

**Burc Kayahan and Thanasis Stengos***
*Department of Economics, University of Guelph, Canada*

**Burak Saltoğlu**
*Department of Economics, Marmara University, Turkey*

**Abstract**

In this paper we investigate the intra-day properties of a recently proposed realized volatility concept using Istanbul Stock Exchange (ISE) 5-minute data returns for the period 1997 to 2000. Using GARCH as a benchmark, we confirm recent findings in the literature that realized volatility provides a better fit than the normal GARCH model.

*Key words*: intra-day volatility; realized volatility; Istanbul Stock Exchange

*JEL classification*: C15; C22; G15

## 1. Introduction

Investigating the volatility patterns with high frequency financial data has become a popular area over the last decade. This particular interest stems from the fact that intra-day volatility dynamics has many implications for return predictability and risk management. As Andersen and Bollerslev (1997) conjectured, economists did not deal with intra-day return dynamics because of the inadequacy of standard time series models in dealing with high frequency data. Following the ARCH methodology of Engle (1982), researchers tried to model the dynamics of intra-day return volatility. One strand of research involves the interrelation between returns in geographically separated financial markets that trade sequentially [see Engle et al. (1990) and Hamao et al. (1990)]. Another line of research in this context is the investigation of the lead and lag relations between two or more markets that trade simultaneously [see Baillie and Bollerslev (1991) and Chan et al. (1991)]. Andersen and Bollerslev (1997) and Andersen, Bollerslev, Diebold, and Labys (2001) (hereafter ABDL) investigated the effects of intra-day seasonality. An excellent review paper by Goodhart and O'Hara (1997) and a recent textbook by Dacarogna et al. (2001) highlight

Received October 23, 2001, accepted January 28, 2002.

*Correspondence to: Department of Economics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada. Email: tstengos@uoguelph.ca.

the issues and problems that arise in high frequency finance. More recently, Andersen et al. (2000), ABDL (2001), and Andersen, Bollerslev, Diebold, and Ebens (2001) (hereafter ABDE), introduced a new concept of volatility, namely, *realized volatility* (RV). It has certain advantages over traditional generalized autoregressive conditional heteroskedasticity (GARCH) models and it has been shown to perform better both in sample and out of sample.

Despite these recent advances, there are no empirical studies that use high frequency financial data from an emerging market. Since these markets are in general very volatile, the findings from these countries can help the scientific community to better examine the performance of newly developed econometric models. The main goal of this paper is to fill this gap by investigating the time series properties of intra-day Istanbul Stock Exchange (ISE) returns. In particular, we investigate the performance of the RV model using ISE intra-day return series and compare it with the traditional GARCH framework. This study is important for two reasons. First, we are able to offer an additional empirical assessment of the theoretical implications of the RV framework. Second, ISE has always appeared to be one of the most volatile stock markets even among other emerging markets. Therefore, investigating such a volatile stock market will have implications about the performance of other emerging markets.

The data used in this study covers 5-minute transaction prices of ISE-100 composite index between 30/12/1997 and 06/03/2000 comprising 25,273 observations. The data set is quite unique and has not been used elsewhere. The RV concept has been explored and the comparison of the benchmark GARCH estimates has been made via standardizing returns at various frequencies and checking the normality of these standardized returns. This approach has also been employed by Andersen (2000). Confirming previous findings, the RV model is shown to perform better than the normal GARCH model and it seems to better explain the intra-day volatility dynamics of the ISE.

The paper proceeds as follows. Section 2 provides some theoretical background of the RV concept. The estimation procedures of GARCH and RV are presented in Section 3, whereas in Section 4 we present and discuss the empirical results. Finally, we conclude.

## 2. Realized and Integrated Volatility for Univariate Diffusion Processes

The literature of empirical measurement of volatility depends on the arguments that are introduced by Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (1998). The theory of integrated and realized volatility has been further discussed in ABDL (2001) and ABDE (2001). In particular, the continuous-time logarithmic price process $p_t$ can be defined as: $dp_t = \sigma_t dW_t,\ t \geq 0$ where $W_t$ denotes standard Brownian Motion and $\sigma_t$ is a strictly stationary process. Then, the discretely sampled returns with $m$ observations per period can be computed with the following formula:

$$r_{(m),t} = p_t - p_{t-1/m} = \int_0^{1/m} \sigma_{t-1/(m+\tau)} dW_{t-1/m+\tau}, \qquad t = 1/m,\ 2/m, \tag{1}$$

where $dW_{t-1/m+\tau}$ denotes the Wiener process and $t$ is the time subscript. By definition the expected returns are equal to zero for all return horizons, and the unit time interval is normalized by setting $m$ equal to 1 in order to represent one day. Furthermore, $\sigma_t$ and $W_t$ are assumed to be independent, which would then lead to the fact that the variance for $h$-period returns for $h > 0$ and $r_{(1/h),t+h}$ conditional on sample path $\{\sigma_{t+\tau}\}_{\tau=0}^{h}$ is given by the definition below:

$$\sigma_{t,h}^2 \equiv \int_0^{1/m} \sigma_{t+\tau}{}^2 d\tau \ . \tag{2}$$

This description of integrated volatility is then concluded to provide a natural definition of price volatility or volatility in a continuous time setting. In other words, the volatility for *h* periods is identical to the integral of past volatility of higher frequencies. Unfortunately, the integrated volatility is unobservable and therefore needs to be estimated. The main approach that has been used in this paper as well as in the paper of Andersen et al. (2000) is to take a sum of high frequency intra-day squared returns in order to compute the daily RV estimates. It can be shown that

$$p\lim_{m\to\infty} \sum_{j=1,\dots,mh} r_{(m),t+j/m}^2 = \sigma_{t,h}^2 \ . \tag{3}$$

In words, RV is a consistent estimator of integrated volatility. Hence, the summation of sufficiently high numbers of high frequency discrete-time intra-day squared returns can be used to estimate the integrated volatility over any horizon accurately and adequately. One of the main reasons for the adoption of the RV concept is that it is free of measurement error as long as $m \to \infty$, unlike the parametric estimates that are likely to suffer from specification and measurement errors. Andersen (2000) argues that the daily squared returns display an extremely noisy behavior and therefore the estimation of realized volatility via high frequency data should provide more accurate and robust measures of volatility. To this end, Andersen (2000) computes the daily GARCH series and then compares the goodness of fit of the alternative estimates of the realized volatility. We also adopted a similar approach in this paper.

## 3. Comparing the Performance of Realized Volatility and GARCH Models

### 3.1 Estimation of GARCH Figures

Following Engle (1982) and Bollerslev (1986), the GARCH model has been a

**Table 1. GARCH Results of MA (12) Model for 5-Minute Returns**

| | ORDERS OF GARCH TERMS | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ARCH Orders | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 98356.76 | 98361.48 | 98366.27 | 98366.53 | 98363.31 | 98425.32 | 98635.83 | 98251.25 | 98917.96 |
| 2 | 97916.76 | 97919.76 | 98279.95 | 98033.23 | 98189.15 | 98651.82 | 98722.57 | 98942.46 | 98975.46 |
| 3 | 97980.98 | 97857.05 | 98002.48 | 98154.98 | 98048.83 | 98480.99 | 98842.45 | 98903.34 | 98952.80 |
| 4 | 97918.86 | 97886.82 | 97742.53 | 98081.68 | 98127.75 | 98836.06 | 98954.23 | 98867.38 | 98967.58 |
| 5 | 97870.93 | 98261.76 | 97948.28 | 97940.78 | 98091.41 | 98052.62 | 98919.05 | 98081.42 | 98137.33 |
| 6 | 98052.81 | 98584.07 | 97811.11 | 98112.32 | 97850.64 | 97811.97 | 97819.24 | 97871.04 | 98031.94 |
| 7 | 97752.92 | 97832.92 | 97881.14 | 97904.70 | 97756.60 | 97771.96 | 97745.52 | 97766.00 | 97791.42 |
| 8 | 97797.18 | 97806.24 | 97739.87 | 97754.34 | 97766.83 | 97662.81 | 97802.99 | 97817.77 | 97838.46 |
| 9 | 97742.50 | 97757.38 | 97769.59 | 97779.21 | 97744.14 | 97689.74 | 97822.13 | 97839.82 | 97859.71 |

standard tool of analysis in empirical time series. In this paper, we use an MA (12) specification to compute the GARCH estimates that capture the correlation structure of the returns. In order to find the best fit we create a nine-by-nine matrix of the log likelihood values that belong to the various combinations of GARCH and ARCH orders and then pick the combination that produces the highest value of the log likelihood. The respective best-fit specification is then used as the basis for the estimation of conditional variance values. For daily returns the best GARCH specification is found to be (2,9) as shown in Table 1.

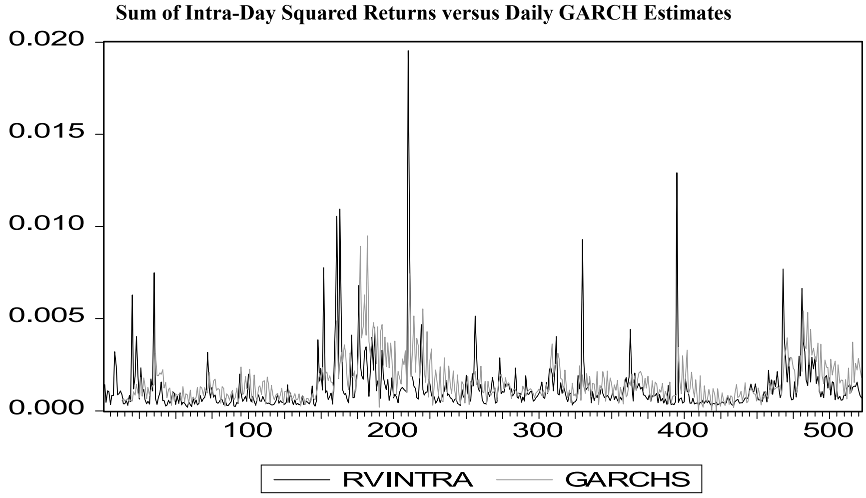
### 3.2 Estimation of Realized Volatility

Andersen (2000) argues that daily squared returns are very noisy and therefore the estimation of RV from high frequency data should provide more accurate and robust measures of volatility. To this end, he computes the daily GARCH series and then compares the goodness of fit of the two alternative estimates of RV. To compare the performance of RV we adopt a similar approach with the only difference that we choose the order of the specification that produces the highest log likelihood value in computing the conditional variance series, instead of relying on a simple (1,1) model. Hence, RV is computed both via summation of high frequency intra-day returns and by squared daily returns. Then both of these estimators are regressed on the best GARCH fit, namely (2,9).

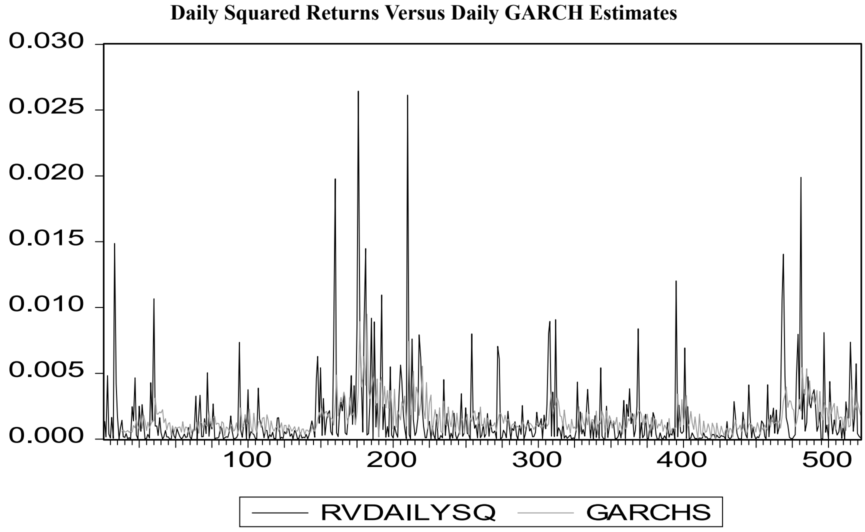
Figure 1 shows the relative fits. As may be seen, GARCH shows a better fit when the sum of intra-day squared returns series is used as the dependent variable rather than daily squared returns. The 5-minute returns are chosen as the high frequency level for the estimation of RV with intra-day data as compared with the daily squared returns. Hence, GARCH estimates are observed to capture the volatility patterns in our sample better when RV is estimated with high frequency data. Moreover, the $R^2$ measure associated with regression of sum of intra-day squared returns on GARCH estimates was computed to be 0.21, whereas the respective $R^2$ measure for daily squared returns was 0.08. The difference between these two is not as high as that found by Andersen and Bollerslev (1997) (they were 0.479 and 0.047 respectively). Overall though, the sum of intra-day squared returns appears to offer better estimates for RV than the daily squared returns.

**Figure 1. Comparison of Realized Volatility Estimates**

**Sum of Intra-Day Squared Returns versus Daily GARCH Estimates**




**Daily Squared Returns Versus Daily GARCH Estimates**



### 3.3 Comparison of Realized Volatility and GARCH Estimates

In order to compare the efficiency of GARCH and RV estimates we will standardize the returns and then analyze the distributions of these standardized returns as in Andersen et al. (2000). This is a very simple, yet informative method to conduct the comparison between the two models.

In the absence of any short-run predictability in the mean, the univariate return series can be defined as: $R_t = \sigma_t \varepsilon_t$ where $\varepsilon_t$ is independently and identically distributed with a zero mean and unit variance and $\sigma_t$ represents the time-$t$ conditional standard deviation. Then, $\sigma$ -standardized returns can be obtained by a simple modification of the definition, such as:

$$\varepsilon_t = \frac{r_t}{\sigma_t}.$$

In reality, $\sigma_t$ is not known of course and needs to be estimated. First we will use GARCH estimates of the volatility in order to compute the standard deviation, which would be used in turn to standardize the returns. Then, in order to make a comparison with the results of Andersen et al. (2000), the returns need to be standardized by the RV estimate of $\sigma_t$ as well. However, since the RV estimator is calculated for daily frequencies, the GARCH estimates will be computed from daily returns.

## 4. Empirical Findings

We assess the performance of these two sets of estimates by comparing the distributional properties of the standardized returns. On the one hand, Table 2 clearly indicates that the RV model appears to generate the standardized return series which are close to being normal. On the other hand, the GARCH model produces standardized returns that are nowhere near normality. There is a wide range of values, a very significant Jarque-Bera test statistic, negative skewness and fat tails. Despite the fact that the GARCH standardized distribution has approximately zero mean and unit variance, it still has negative skewness and much more importantly excess kurtosis.

This result is not surprising since it is known in the literature that stock returns tend to follow non-normal unconditional sampling distributions, particularly in the form of excess kurtosis. As described in Bollerslev et al. (1992), the conditional normality assumption in ARCH generates some degree of unconditional excess kurtosis, however generally less than enough to fully account for the fat-tailed properties of the data. Andersen et al. (2000) have also found similar results in their paper. They conclude that standardization by $\sigma_t$ using GARCH is insufficient to eliminate the excess kurtosis, whereas standardization with the $\sigma_t$ using RV is able to accomplish that goal.

**Table 2. Descriptive Statistics for Standardized Daily Stock Returns of ISE**

| | $\sigma_t$ (GARCH) Standardized | $\sigma_t$ (RV) Standardized |
|---|---|---|
| Mean | -0.009366 | 0.055964 |
| Median | -0.037995 | -0.056891 |
| Maximum | 3.047982 | 2.682061 |
| Minimum | -9.620199 | -2.792690 |
| Std. Dev. | 1.108361 | 1.064142 |
| Skewness | -1.331161 | 0.091569 |
| Kurtosis | 13.481700 | 2.279461 |
| Jarque-Bera | 2485.273000 | 12.021560 |
| Probability | 0 | 0.002452 |
| Observations | 510 | 522 |

## 5. Conclusion

In this study we assess the relative performance of RV using intra-day returns from an emerging market. This estimate appeared to be more successful than that of a conventional GARCH model. This result is in agreement with the recent findings from the high frequency finance literature that relies primarily on data from established markets. One interesting comparison along these lines can be made by comparing realized and implied volatilities from option pricing following Engle et al. (1994) and Christensen and Prabhala (1998). However, since an official options market does not exist in Turkey, this line of research is left for the future.

## References

Andersen, T. G., (2000), "Some Reflections on Analysis of High-Frequency Data," *Journal of Business and Economic Statistics*, 18(2), 146-153.

Andersen, T. G. and T. Bollerslev, (1997), "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115-158.

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, F. X. Diebold, and P. Labys, (2000), "Exchange Returns Standardized by Realized Volatility are (Nearly) Gaussian," *NBER Working Paper*, 7488.

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, (2001), "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42-45.

Baillie, T. and T. Bollerslev, (1991), "Intraday and Inter Market Volatility in Foreign

Exchange Rates," *Review of Economic Studies*, 58, 565-585.

Barndorf-Nielsen, O. E. and N. Shephard, (1998), "Aggregation and Model Construction for Volatility Models," *Manuscript*. Oxford: Department of Mathematical Sciences, University of Aarhus and Nuffield College.

Bollerslev, T., (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31(3), 307-327.

Bollerslev T., R. Y. Chou, and K. F. Kroner, (1992), "ARCH Modelling in Finance," *Journal of Econometrics*, 52, 5-59.

Chan, K., K. C. Chan, and G. A. Karolyi, (1991), "Intraday Volatility in the Stock Index and Stock Index Futures Market," *Review of Financial Studies*, 4, 657-684.

Christensen, B. J. and N. R. Prabhala, (1998), "The Relation between Implied and Realized Volatility," *Journal of Financial Economics*, 50(2), 125-150.

Dacarogna, M. M., R. Gencay, U. A. Muller, R. B. Olsen, and O. V. Pictet, (2001), *An Introduction to High Frequency Finance*. Toronto: Academic Press.

Engle, R. F., (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.

Engle, R. F., T. Ito, and W. L. Lin, (1990), "Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market," *Econometrica*, 58, 525-542.

Engle, R. F., J. Noh, and A. Kane, (1994), "Forecasting Volatility and Option Prices of the S&P 500 Index," *Journal of Derivatives*, 17-305.

Goodart, C. A. E. and M. O'Hara, (1997), "High Frequency Data in Financial Markets: Issues and Applications," *Journal of Empirical Finance*, 4, 73-114.

Hamao, Y., R. W. Masulis, and V. Ng, (1990), "Correlations in Price Changes and Volatility across International Stock Markets," *Review of Financial Studies*, 3, 281-308.