Bu makaleye atıfta bulunmak için/To cite this article:
EROĞLU, B.A İKİZLERLİ, D. YENER, H. (2021). Reexamination of the BIST 100 Stock Price Volatility with Heterogeneous Autoregressive Realized Volatility Models. Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 25 (2) , 457-476.

Reexamination of the BIST 100 Stock Price Volatility with Heterogeneous Autoregressive Realized Volatility Models

Burak Alparslan EROĞLU [(*)]
Deniz İKİZLERLİ [(**)]
Haluk YENER [(***)]

Abstract: *In this study, we employ the heterogeneous autoregressive model framework on the (half) daily returns of the BIST 100 index between the years 2016 and 2019. This framework helps us understand the short, medium, and long-term patterns of the volatility dynamics for the return series. Notably, we analyze how leverage effect and jumps in the return series affect the realized volatility of the BIST 100 index. For the analysis, we employ sixteen models, and the results from these models show that there is a leverage effect, albeit small. The effect of jumps is significant and is present either in the short-term or long-term, depending on the type of model utilized for the analysis. We also detect a diurnal effect at the session level, implying that the realized volatility of the BIST 100 index is lower in the morning sessions.*

Keywords: *Realized volatility, heterogenous autoregressive model, the BIST 100 index*

BIST 100 Endeksi Hisse Senedi Fiyat Volatilitesinin Heterojen Otoregresif Gerçekleşen Volatilite Modeliyle Yeniden İncelenmesi

Öz: *Bu çalışmada, BIST 100 Endeksi'nin 2016-2019 yılları arasındaki üç yıllık dönem için günlük getirilerinin volatilitesi, getirilerin volatilite dinamikleri üzerindeki kısa, orta ve uzun vadeli etkilerini anlamak için heterojen otoregregresive modelleme kullanılarak analiz edilmektedir. Özellikle, kaldıraç etkisinin ve getirideki sıçramaların BIST 100 Endeksi'nin volatilite dinamiklerini nasıl etkilediği araştırılmaktadır. Analiz için, on altı farklı modeli veri setine uygulunmaktadır ve bu modellerin sonuçları küçük de olsa bir kaldıraç etkisi olduğunu göstermektedir. Sonuçlara göre sıçrama etkisi, analiz için kullanılan model türüne bağlı olarak kısa veya uzun vadede istatistiki olarak anlamlı çıkmaktadır. Ayrıca, BIST 100 Endeksinin sabah seansında gerçekleşen oynaklığın daha düşük olduğunu işaret edilerek, seans seviyesinde mevsimsel bir etki olduğu da gösterilmektedir.*

Anahtar Kelimeler: *Gerçekleşen volatilite, Heterojen ototregresif model, BIST 100 endeksi*

Makale Geliş Tarihi: *15.11.2020*

Makale Kabul Tarihi: *14.05.2021*

[*)] Dr. Öğr. Üyesi, İstanbul Bilgi Üniversitesi İşletme Fakültesi Ekonomi Bölümü (e-posta: burak.eroglu@bilgi.edu.tr) ORCID ID.https://orcid.org/0000-0001-6814-747X

[**)] Dr. Öğr. Üyesi, İstanbul Bilgi Üniversitesi İşletme Fakültesi İşletme Bölümü (e-posta: deniz.ikizlerli@bilgi.edu.tr) ORCID ID.https://orcid.org/0000-0002-3443-8073

[***)] Dr. Öğr. Üyesi, İstanbul Bilgi Üniversitesi İşletme Fakültesi İşletme Bölümü (e-posta: haluk.yener@bilgi.edu.tr) ORCID ID.https://orcid.org/0000-0003-2654-5810

I. Introduction

An essential aspect of financial time series is the serial autocorrelation of their squared returns, hence the predictability of their volatility (Engle, 1982; Bollerslev, 1986). With the predictability comes the widespread application of the times series modeling in the critical fields of finance. Particular applications include the volatility modeling in the option pricing formula of Black and Scholes (1973), variance-covariance matrix modeling in the optimal portfolio selection specification of Merton (1970), and risk management (see, for example, Christoffersen and Diebold, 1998). In this paper, an application of volatility modeling is undertaken to the intra-day return series of the Borsa Istanbul 100 (BIST 100) index. The application involves the heterogeneous autoregressive model of realized volatility (HAR-RV) method of Corsi (2009). As described in the sequel, the method offers an alternative type of modeling to the seminal volatility modeling of Engle (1982) and Bollerslev (1986).

Financial time series presents a series of well-known stylized facts such as volatility clustering (Engle, 1982; Bollerslev, 1986) and fat-tailed (asymmetric in cases) return distributions (Mandelbrot, 1963). As additional facts, we also observe asymmetric responses of volatility to shocks (Black, 1976) and long memory property (Baillie et al., 1996). The former, also known as the "leverage effect," is well captured in the offsprings of generalized autoregressive conditional heteroscedasticity (GARCH) models as shown in the exponential-GARCH (EGARCH) model of Nelson (1991) and threshold-GARCH (T-GARCH) model of Glosten et al. (1993). For the long memory property, Baillie et al. (1996) propose a fractionally integrated generalized autoregressive conditional heteroscedasticity (FIGARCH) model to account for the lack of long memory property in GARCH models.

In time series models, autocorrelations may very well be statistically significant for large time lags. In effect, as mentioned in Karanasos et al. (2004) (see also the references therein), autocorrelations of squared returns of high-frequency data tend to decline very slowly. This type of autocorrelation behavior suggests the presence of long memory in the volatility. The primary outcome of long memory is that shocks to the volatility take a long time to die out. In other words, the market does not immediately react to information flow. Instead, it responds slowly over time. A second aspect is the leverage effect, as mentioned above. In his seminal work, Black (1976) showed that volatility tends to react more to a negative shock than a positive shock of the same size. One final aspect is the discontinuous (and discrete) paths of the high-frequency data as returns are found to display frequent jumps at the high-frequency level (see, for example; Andersen et al., 2007; Lee and Mykland, 2008; Bajgrowicz et al., 2015).

With the availability of high-frequency data, the volatility modeling literature witnessed a significant step. Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2002) propose a new approach called "realized" volatility that takes advantage of the information in high-frequency returns. The squared of intraday returns sampled at very short intervals are accumulated to construct daily realized volatility. This construction allows volatility to be treated as an "observed" instead of a latent variable. When the volatility is observable, we can employ much simpler econometric models than the models required when the volatility is latent. Under the theory of quadratic variation, the realized volatility is a consistent estimator and an error-free volatility measure for the actual volatility in many studies (see, for example, Andersen and Bollerslev, 1998; and Barndorff-Nielsen and Shephard, 2002).

Later, using the realized volatility approach, Corsi (2009) proposes the heterogeneous autoregressive model of realized volatility (HAR-RV), which is a predictive regression. The model includes lagged daily, weekly, and monthly RV measures in explaining the future RV. The study suggests that simple HAR-RV can capture some 'stylized facts' in financial volatility such as long memory, fat tail distribution, and multiscaling. The results in Corsi (2009) also show that the HAR-RV model is better than the GARCH type models and the ARFIMA-RV model in forecasting performance. Afterward, to further improve forecasting performance, Corsi and Reno (2009, 2012) propose an extension by including the leverage effect. Moreover, the authors also consider the low-frequency averages of such leverage effects (Corsi and Reno, 2009, 2012).

In this study, the HAR-RV model of Corsi (2009) and Corsi and Reno (2009, 2012) is applied to analyze intraday returns of the BIST 100 index. As done in the studies mentioned above and as suggested by Lee and Mykland (2008) and Andersen et al. (2007), return series are separated into their jump and continuous components since they have different dynamics. Furthermore, high-frequency returns are considered at the trading session-level rather than the daily level because there are two trading sessions in Borsa Istanbul; one is the morning session, and the other is the afternoon session.

As mentioned previously, the basic HAR model includes three partial components: short-term traders with daily or higher trading frequency, medium-term traders with weekly trading frequency, and long-term traders with monthly or lower trading frequency. We also add one more frequency into the frequency list, namely session-based or half-daily frequency. Thus, we consider four frequencies for the volatility measures in our model. We adopt an extension of the HAR-RV model, which also takes the leverage effects and jump dynamics into consideration when modeling the dynamics of realized volatility.

We consider the intraday volatility analysis of the BIST 100 index because it is one of the major markets in the Middle East and North Africa (MENA) region (Lagoarde-Segot and Lucey, 2007). Volatility is much higher in the emerging equity markets than it is in the developed markets (Bekaert and Harvey, 1997), and long-run economic growth is, in turn, impeded by the high volatility in emerging markets (Levine and Zervos, 1998). Thus, understanding the characteristics of volatility in emerging stock markets is vital for policymakers and investors. Analyzing the BIST 100 index volatility is also essential because a significant proportion of investors are non-domestic, and the index is significantly affected by international news and flows (Inci and Ozenbas, 2017).

The high-frequency analysis of the BIST 100 index in the finance literature is not exhaustive. To the best of our knowledge, Bildik (2001) is the first study that thoroughly analyzes the high-frequency return dynamics of the Istanbul Stock Exchange (ISE; earlier name for BIST). The study shows that the intra-day returns in ISE follow a W-shaped pattern, while volatility follows an L-shaped pattern between 1996 and 1999. That is, volatility is higher in the morning session, and it is lower in the afternoon session. The same pattern for ISE'e intra-day returns is also confirmed by Temizel (2008) for the data that extend between 1998 and 2003. However, Temizel (2008) also shows that the volatility also follows a W-shape pattern rather than an L-shaped pattern shown in Bildik (2001).

In any case, we may observe the intra-day variation in volatility. In effect, as Inci and Ozenbas (2017) show, by considering ISE'e high-frequency data set between 1998 and 2014, there is a volatility smirk during the morning session is ISE and volatility smile in the afternoon session. In essence, as Bildik (2001), there is a relation between the intra-day patterns of volatility and volume. One can also see this relation from the results of Koksal (2012), showing that spreads in ISE follow an L-shape pattern, while the number of trades and volume follow a U-shape pattern for a short-term period in between May and July of the year 2008.

While the findings of the studies mentioned above are generally in line with those of the finance literature, they usually do not take the asymmetric and long memory properties into account. The main contribution of this study is to consider both properties above while also checking the intra-day variation of the BIST 100 index volatility. Particularly, we conduct the analysis under the jump dynamic. The findings from the data that extends between 2016-04-01 and 2019-01-31 are as follows: (i) the jumps significantly affect the intra-day BIST 100 index volatility both in the short- and long-term depending on the type model employed for the analysis; (ii) the BIST 100 index volatility is lower in the morning session and higher in the afternoon session; (iii) there is a certain level of asymmetry to shocks in the BIST 100 index volatility; (iv) the models that account for jumps and leverage effect along with diurnal patterns are better predictors of the BIST 100 index volatility.

The rest of the paper is constructed as follows. The methodology and data are presented in Section 2. For the methodology, four HAR-RV type models are considered. The first model is the simplest model without leverage and jump effects. The second model is the one with the leverage effects with no jumps. In the third model, only the jump effect is considered. Finally, in the fourth model, all stylized facts of volatility modeling are considered. That is, the final model includes both the jump and leverage effects. Furthermore, considering the dummy variables to understand the trading session and the day of the week effect, the number of models considered for the study increases to sixteen. After presenting the data and the summary of statistics in Subsection 2.3, the results are presented in Section 3. In Section 4, a summary of the key findings is provided.

II. Methodology and Data

This section introduces the methodology and the dataset utilized in the current study. We first describe the econometric tools, including the construction of different volatility measures and the estimation of the heterogeneous autoregressive realized volatility (HAR-RV) models; then, we present the dataset, which consists of the BIST 100 index price.

The BIST 100 index is a market value-weighted index consisting of 100 national market firms. The index excludes investment trusts and is the most important index of the Borsa Istanbul. Therefore, its analysis offers an opportunity to understand the dynamics of the Turkish equity markets. Moreover, as Turkey is among the major emerging markets, our analysis also offers a benchmark for understanding the high-frequency dynamics of the equity markets in emerging markets. We proceed with the description of the methodology employed for the analysis.

A. Volatility Measures

Following Corsi and Reno (2009), we denote the logarithm of the index price with P_t and consider for the log-price evolution a jump-diffusion model of the form:

$$dP_t = \mu_t dt + \sigma_t dW_t + c_t dN_t, \qquad \textbf{(1)}$$

where μ_t is the predictable drift component, σ_t is the instantaneous volatility of the continuous part, W_t is the standard Brownian motion, N_t is a counting process with time-varying intensity λ_t, and c_t is *i.i.d.* process determining the jump magnitude. Corsi and Reno (2009) demonstrates that the quadratic variation of the price process within an interval (say, one day or one trading seance) can be represented as,

$$\tilde{\sigma}_t = \int_t^{t+1} \sigma_s^2 ds + \sum_{t \le \tau_j \le t+1} c_{\tau_j}^2, \qquad \textbf{(2)}$$

where, τ_j denotes the jump times within a time interval. Because we model the price evolution as a stochastic process by employing a jump-diffusion model, as shown above, we use the notion of the quadratic variation to model the “realized volatility” of the returns. To estimate this object, we provide additional notations that will be utilized in the sequel.

Throughout the paper, we denote the trading session (time) index with *t*. We assume that there are two trading sessions in a day, namely the morning and afternoon sessions, as is the case with the Borsa Istanbul. We consider the price data in seconds. To this end, we assume that each trading session is separated into n equal intervals, meaning a total of n seconds. We consider such a split as we have a regular dataset, implying that there is one trade per second. In other words, there is always a trade and only one trade per each n=1, 2,… in a trading session. Then, we let $\{P_{t,i}\}_{i=1}^{n}$ for all $t = 1,2, \ldots$ denote the high-frequency evolution of the price series in session t, where i = 1, …, n is for the high-frequency price observations during a session. We define the index return measured in seconds as $r_{t,i} = [P_{t,i} - P_{t,i-1}] \times 100$, since $P_{t,i}$ denotes the logarithm of the index price. Our objective is to provide a measure of the "realized volatility" within a trading session. This measure is an estimator for the quadratic variation per trading session and is formulated as

$$RV_t = \sum_{i=1}^{n} r_{t,i}^2. \quad \textbf{(3)}$$

As shown by Andersen et al. (2003), the above is a consistent estimator of $\tilde{\sigma}_t$ as $n \to \infty$. Since we consider high-frequency data, we have considerable value for n implying consistency in our estimation of volatility.

Corsi (2009) shows that the two volatility measures mentioned above are linked with a simple relation. To construct the relation, we let $g(\cdot)$ be a monotonic function, such as $g(x) = x$ and $g(x) = log(x)$. We can write $g(RV_t) = g(\tilde{\sigma}_t^2) + \epsilon_t$, where ϵ_t is a random and independent measurement error term. Using this relation, Corsi (2009) and Corsi and Reno (2009) construct the estimable volatility models; we refer to these seminal articles for understanding the construction of models that we use for the analysis of the high-frequency BIST 100 index returns.

Before introducing the realized volatility models, we first define the low-frequency aggregates of the volatility measure as,

$$g(RV_t)^{(k)} \coloneqq \frac{1}{k}\big(g(RV_t) + g(RV_{t-1}) + \cdots + g(RV_{t-k+1})\big). \quad \textbf{(4)}$$

Most often, we use letters instead of a number for k in our notation. For instance, when $k = 2$, we write the daily aggregated volatility measure as $g(RV_t)^{(d)} = (1/2)\big(g(RV_t) + g(RV_{t-1})\big)$, where d stands for "daily." In other words, when $k = 2$, we compute the sample average of the last two trading sessions. Similarly, we denote the weekly ($k = 10$) and monthly ($k = 44$) volatility with letters *w* and *m*, respectively. In sum, as evident from the above formulation, we consider a moving average of session-based realized volatility values in accordance with the period considered for the analysis.

B.Heterogeneous Autoregressive Realized Volatility Models

After introducing the critical volatility measures, we present the novel approach of Corsi (2009) to uncover the long-memory properties of these measures. This model is called the Heterogeneous Autoregressive Realized Volatility (HAR-RV) model. In this modeling framework, Corsi (2009) examines the relationship between the high-frequency (daily) realized volatility and the set of regressors that includes the past high frequency and aggregate low frequency (weekly and monthly) volatility measures.

In this study, we follow Corsi's (2009) construction with a minor modification. We add one more frequency into the frequency list, namely the session-based (or half-daily) frequency. In this respect, we consider our modification under four models, which we now explain. The first one is constructed on a baseline scenario that directly follows Corsi (2009). We then improve the baseline scenario by considering the leverage effect explaining the abnormal, asymmetric movements of the volatility series. Besides, we employ two extra models that we employ to understand the effect of jumps first without the leverage effect.

C.The Basic HAR-RV Model

The basic HAR-RV model that we utilize in this study consists of the following regression model:

$$g(RV_{t+1}) = \beta_0 + \beta_1 g(RV_t) + \beta_2 g(RV_t)^{(d)} + \beta_3 g(RV_t)^{(w)} + \beta_4 g(RV_t)^{(m)} + \epsilon_t. \quad \textbf{(5)}$$

Thanks to its linear form, we can estimate the unknown population parameters in Equation (5) using the ordinary least squares (OLS) method. One can further employ the Newey and West (1987) covariance for the standard errors of the parameter estimates to remove any further problematic issues in the estimation. In the above, our modification consists of the treatment of the realized volatility per session, as captured by $g(RV_t)$, as well as the daily, weekly, and monthly realized volatilities. As we also see, the first model does not capture the stylized facts of returns series, including the asymmetric movements of returns with respect to shocks and jumps in their high-frequency evolution. To account for these, we proceed to the second model, where we describe the HAR-RV Model with the leverage effect.

D.The HAR-RV Model with Leverage Effects (LHAR-RV)

The basic model introduced in the previous section lacks some crucial features of the volatility dynamics. To improve upon the basic model, Corsi and Reno (2009, 2012) propose an extension by including the so-called "leverage effect," which indicates that volatility tends to react more to a negative shock than to a positive shock of the same size (see Black, 1976). Moreover, the authors also consider the low-frequency averages of such leverage effects (Corsi and Reno, 2009, 2012).

Let $1\{.\}$ be an indicator function. We then define $r_t^+ = r_t \times 1\{r_t > 0\}$, $r_t^- = r_t \times 1\{r_t \leq 0\}$ as the positive and negative parts of the return series at time *t*, respectively. Similarly, we can create the low-frequency aggregates of these objects as,

$$\begin{aligned} r_t^{+(k)} &= \frac{1}{k}(r_t + r_{t-1} + \cdots + r_{t-k+1}) \times 1\{r_t + r_{t-1} + \cdots + r_{t-k+1} > 0\}; \\ r_t^{-(k)} &= \frac{1}{k}(r_t + r_{t-1} + \cdots + r_{t-k+1}) \times 1\{r_t + r_{t-1} + \cdots + r_{t-k+1} \leq 0\}, \end{aligned} \quad \textbf{(6)}$$

respectively for the positive and negative returns. We utilize the above two objects to control the long-memory effects of the positive or negative part of the return series. Accordingly, we can analyze how daily, weekly, or monthly positive or negative parts of the returns influence the volatility dynamics. Embedding these objects into the basic HAR-RV model, we obtain the following regression equation,

$$\begin{aligned} g(RV_{t+1}) = {} & \beta_0 + \beta_1 g(RV_t) + \beta_2 g(RV_t)^{(d)} + \beta_3 g(RV_t)^{(w)} + \beta_4 g(RV_t)^{(m)} \\ & + \beta_5 r_t^{+} + \beta_6 r_t^{-} + \beta_7 r_t^{+(d)} + \beta_8 r_t^{-(d)} + \beta_9 r_t^{+(w)} + \beta_{10} r_t^{-(w)} \\ & + \beta_{11} r_t^{+(m)} + \beta_{12} r_t^{-(m)} + \epsilon_t. \end{aligned} \tag{7}$$

We can also estimate this regression equation using the OLS procedure with the Newey-West (1987) serial correlation and Heteroscedasticity robust covariance estimates.

E.HAR-RV model with Jumps (HAR-CJ)

In Equations (5) and (7), we ignore the presence of the jump component in the data generation mechanism. However, in the recent literature, jump dynamics have a growing importance in the analysis of volatility and return series. To accommodate the jump patterns in the current framework, we apply the tools developed by Corsi and Reno (2009, 2012). In their setup, Corsi and Reno (2009, 2012) separate the jump and continuous part of the volatility. We will not give the full description of this procedure but provide a sketch of the underlying mechanism[1].

To this end, we first employ a jump detection mechanism on the intraday price series. In this regard, we can use the jump tests devised by Lee and Mykland (2008), Jiang and Oomen (2008), Corsi et al. (2010), and Barndorff-Nielsen and Shephard (2006). If we detect a jump at *t*, then we estimate the quadratic jump variation $J_t = \sum_{t \le \tau_j \le t+1} c_{\tau_j}^2$ by employing the methods suggested by Corsi and Reno (2009, 2012). We denote the estimate of the quadratic jump variation as $\hat{J}_t$. Subtracting this component from the realized volatility, we obtain an estimate for the continuous part of the quadratic variation. That is, the continuous part of $C_t = \int_t^{t+1} \sigma_s^2 ds$ is estimated by $\widehat{C}_t = RV_t - \widehat{J}_t$. Now, using these two objects, we build a new regression equation as,

$$\begin{aligned} g(RV_{t+1}) = {} & \beta_0 + \beta_1 g(C_t) + \beta_2 g(C_t)^{(d)} + \beta_3 g(C_t)^{(w)} + \beta_4 g(C_t)^{(m)} \\ & + \beta_1 \tilde{g}(J_t) + \beta_2 \tilde{g}(J_t)^{(d)} + \beta_3 \tilde{g}(J_t)^{(w)} + \beta_4 \tilde{g}(J_t)^{(m)} + \epsilon_t, \end{aligned} \tag{8}$$

where in some cases, we need to modify the function $g(\cdot)$ slightly for the jump part, thus we denote the transformation as $\tilde{g}(\cdot)$. This model investigates the jump and continuous parts separately. In this regard, this separation may give more insights into the volatility dynamics.

[1] The details of the procedure can be found in Corsi and Reno (2009, 2012).

F.HAR-RV Model with Jumps and Leverage Effect (LHAR-CJ)

Our final model is a combination of the previous two models. In this model, we consider both the leverage effects and jump dynamics. The following equation is the broadest model for the volatility dynamics:

$$
\begin{aligned}
g(RV_{t+1}) = \beta_0 &+ \beta_1 g(C_t) + \beta_2 g(C_t)^{(d)} + \beta_3 g(C_t)^{(w)} + \beta_4 g(C_t)^{(m)} \\
&+\beta_5 \tilde{g}(J_t) + \beta_6 \tilde{g}(J_t)^{(d)} + \beta_7 \tilde{g}(J_t)^{(w)} + \beta_8 \tilde{g}(J_t)^{(m)} \\
&+\beta_9 r_t^{+} + \beta_{10} r_t^{-} + \beta_{11} r_t^{+(d)} + \beta_{12} r_t^{-(d)} + \beta_{13} r_t^{+(w)} + \beta_{14} r_t^{-(w)} \\
&+\beta_{15} r_t^{+(m)} + \beta_{16} r_t^{-(m)} + \epsilon_t.
\end{aligned} \tag{9}
$$

Finally, we also include some deterministic regressors, such as session and day dummies. These inclusions are essential for understanding the heterogeneity in volatility dynamics based on trading sessions or the day of the week. For the sake of not complicating the notation, we omit the demonstration of the model that contains the inclusion of dummy variables. We note, however, that the addition of dummy variables is relatively straightforward.

G. Data

Our dataset consists of the high-frequency BIST 100 index price data between 2016-04-01 and 2019-01-31. We retrieve the data from the BIST Datastore. As we discussed earlier, there are two trading sessions in the BIST. The first session is between 10:00-13:00, and the second one is between 14:00-17:00[2]. The separation of daily trading into two sessions also allows us to analyze the heterogeneity due to the session differences.

In our sample, we only consider the business days, excluding national holidays and weekends. Accordingly, we have observations based on $T = 1420$ trading sessions. We calculate the realized volatility RV_t by using 10800 secondly data within each trading session. So, in total, we have 1420 x 10800, implying close to 15.4 million observations. We provide the summary statistics of the data utilized for our analysis here below. Note that the data is provided based on the returns from the minutely data rather than the secondly data as we estimate the statistics related to jumps (i.e., number of jumps) from returns within a minute interval.

Table 1. Return Statistics for Minutely BIST100 Returns (dlog(Price)*100)

	Overall	2016	2017	2018	2019
Mean	-0.00012	-0.00013	0.00000	-0.00027	0.00043
Std. Dev.	0.01941	0.01706	0.01537	0.02409	0.01999
Median	0.00031	0.00020	0.00033	0.00033	0.00109

[2] Even though the price data is available between 13:00-14:00, the index price exhibits almost stable pattern during this period.

Kurtosis	32.03097	16.32362	17.11742	32.17472	5.68634
Skewness	-0.81413	-0.12264	-0.44415	-1.03542	-0.31216
Min	-0.61333	-0.24140	-0.34890	-0.61333	-0.12153
Max	0.40839	0.36062	0.36814	0.40839	0.11615
Sample Size	256376	67093	91132	90209	7942
JB-Stat	9031359.48	496429.19	759777.64	3215396.29	2517.02
p-value(JB)	0.001	0.001	0.001	0.001	0.001

Note: Std. Dev. is the standard deviation of the returns. JB-Stat is the Jarque-Bera normality test statistics. Below this statistic is the p-value(JB) denoting the p-value of the Jarque-Bera test statistic.

From the above table, we see that the mean and median values of the minutely return series are quite close to zero (hence similar values for both statistics), suggesting that the distribution of the returns is somewhat symmetric. This situation can also be seen from the small negative skewness values. Nevertheless, from the large kurtosis values, we observe that the returns series possess fat-tails, implying the existence of the large outliers in the high-frequency data. In effect, we detect very high minutely return levels from the maximum and minimum values in the data. For example, from the overall values, we see that, on a minute basis, the prices may drop around 0.61% at worst, while they may increase around 0.40% at best. When annualized, we may see how large the effect of these values is. In essence, the Jarque-Bera statistics confirm the highly non-normal distribution of the returns as well. In sum, from the summary statistics, we observe that the minutely distribution o BIST is highly non-normal with fat tails (see Mandelbrot, 1963).

Given our short initial analysis above, understanding the variability of returns through the analysis of leverage and jump effects is quite natural, as evident from the distribution of data (i.e., high kurtosis and negative skewness). Moreover, we seek to understand whether there is a seasonal effect between the trading session and the days of the week. To account for the former, in some of our models, we use the session dummy $\{D_t\}_{t=1}^{T}$, which takes value one during the morning trading session, and zero otherwise. To account for the latter, we utilize the weekday dummies. We simply have five dummy variables to check whether there is a difference between the days of the week. These dummy variables are denoted as $\{D_{j,t}\}_{t=1}^{T}$ for $j \in \{mon, tue, wed, thu, fri\}$. In our estimation exercises, we omit the Monday dummy to avoid the dummy variable trap (multicollinearity).

III.Results

In this section, we present the estimation results for the sixteen volatility models derived from Equations (5)-(9). These models consist of two types of specifications for $g(\cdot)$, namely linear (identity) and logarithmic specifications. That is, we set $g(x) = x$ or $g(x) = log(x)$ respectively based on the aforementioned specifications. Furthermore, as discussed in Section 2.2.3, for the jump component, when we use $g(x) = log(x)$, we set $\tilde{g}(x) = log(x + 1)$ (see Corsi and Reno, 2009). As a result, we have eight models to estimate; four models with $g(x) = x$ and another four models with $g(x) = log(x)$. In addition to these models, we also consider models with and without dummy interventions. Therefore, with dummy specifications and our original eight models, we have another extra eight models leading to sixteen models.

We first exhibit the full sample estimates for these models and discuss our findings. Next, we employ an out-of-sample forecast exercise to understand which model has the best forecasting performance.

A. Full Sample Estimation Results

We present the full sample OLS estimation results in Tables 2-5. In these tables, we report the OLS coefficient estimates, their t-statistics, p-values of the t-statistics, R-squared (R^2), adjusted R-squared ($\bar{R}^2$) and the F-statistic for the overall significance of the model.

Table 2 exhibits the estimation results for the models without jump dynamics and dummy variables. That is, the table includes the results from the variations of the models in Equations (5) and (7). These variations are done, as we discussed right above. In the HAR-RV and LHAR-RV models, the dependent variable is RV_{t+1}, and in the logHAR-RV and logLHAR-RV models, the dependent variable is $log(RV_{t+1})$. In all the cases, R^2 and $\bar{R}^2$ vary around 0.45 and 0.5. This result indicates that these models explain almost half of the variation in the realized volatility (log-realized volatility for log models). Moreover, all the models are overall significant, according to the F-test. Thus, the models explain the realized volatility (RV) measures. Besides, the intercept terms are all significant as there must be some level of base volatility. In effect, we note at the outset that the intercept terms are significant in all sixteen models we employ for the analysis.

Continuing our analysis with the results in Table 2, we observe that the level models (i.e., the HAR-RV and LHAR-RV) possess long-memory properties since the coefficients of daily, weekly, and monthly volatility are significant at the 5% significance level except for the daily RV in the HAR-RV model (it is significant at 10%). In both log models, the monthly volatility is not significant; thus, log RV exhibits shorter memory than the level counterpart. The other crucial finding is that the leverage effect is weak for the BIST 100 return volatility, because most of the leverage variables, as can be observed from the results of the LHAR-RV and logLHAR-RV models, are insignificant except in few cases. These include weekly positive leverage in the LHAR-RV model at the 5% significance level and negative weekly leverage at the 10% significance level. The coefficient values suggest that average weekly positive shocks on average, in the LHAR-RV model, tend to increase the future volatility, while the average weekly adverse shocks, in the logLHAR-RV model, tends to decrease the future volatility. However, the latter relation is relatively weak. Thus, we may see that the leverage effect is at best moderate and happens in the medium term (i.e., on a weekly basis).

Table 2. The OLS Estimation Results for the Models without Jump Dynamics and Dummy Variable Interventions

	HAR-RV			logHAR-RV			LHAR-RV			logLHAR-RV		
	Coef.	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Intercept	**0.05**	2.62	0.01	**-0.16**	-6.10	0.00	**0.05**	2.18	0.03	**-0.25**	-6.28	0.00
$\boldsymbol{log(RV_t)}$	——	——	——	**0.15**	3.20	0.00	——	——	——	**0.13**	2.78	0.01
$\boldsymbol{log(RV_t)^{(d)}}$	——	——	——	**0.37**	5.65	0.00	——	——	——	**0.37**	5.52	0.00
$\boldsymbol{log(RV_t)^{(w)}}$	——	——	——	**0.28**	5.72	0.00	——	——	——	**0.25**	4.74	0.00
$\boldsymbol{log(RV_t)^{(m)}}$	——	——	——	0.05	0.93	0.35	——	——	——	0.06	1.06	0.29
$\boldsymbol{r_t^+}$	——	——	——	——	——	——	0.00	-0.17	0.86	0.00	0.11	0.91
$\boldsymbol{r_t^{+(d)}}$	——	——	——	——	——	——	-0.02	-0.47	0.64	-0.01	-0.21	0.84
$\boldsymbol{r_t^{+(w)}}$	——	——	——	——	——	——	**0.15**	2.19	0.03	0.18	1.55	0.12
$\boldsymbol{r_t^{+(m)}}$	——	——	——	——	——	——	-0.16	-1.19	0.23	-0.14	-0.54	0.59
$\boldsymbol{r_t^-}$	——	——	——	——	——	——	-0.03	-0.64	0.52	-0.03	-0.83	0.41
$\boldsymbol{r_t^{-(d)}}$	——	——	——	——	——	——	-0.02	-0.58	0.57	-0.03	-0.72	0.47
$\boldsymbol{r_t^{-(w)}}$	——	——	——	——	——	——	-0.09	-1.53	0.13	*-0.11*	-1.67	0.09
$\boldsymbol{r_t^{-(m)}}$	——	——	——	——	——	——	0.01	0.05	0.96	-0.23	-1.43	0.15
$\boldsymbol{RV_t}$	**0.33**	2.92	0.00	——	——	——	**0.29**	3.21	0.00	——	——	——
$\boldsymbol{RV_t^{(d)}}$	*0.16*	1.71	0.09	——	——	——	**0.17**	1.94	0.05	——	——	——
$\boldsymbol{RV_t^{(w)}}$	**0.24**	3.25	0.00	——	——	——	**0.18**	2.50	0.01	——	——	——
$\boldsymbol{RV_t^{(m)}}$	**0.17**	3.37	0.00	——	——	——	**0.21**	4.09	0.00	——	——	——
	$\boldsymbol{R^2}$	$\boldsymbol{\bar{R}^2}$	$\boldsymbol{F}$	$\boldsymbol{R^2}$	$\boldsymbol{\bar{R}^2}$	$\boldsymbol{F}$	$\boldsymbol{R^2}$	$\boldsymbol{\bar{R}^2}$	$\boldsymbol{F}$	$\boldsymbol{R^2}$	$\boldsymbol{\bar{R}^2}$	$\boldsymbol{F}$
	0.4448	0.44	283.4	0.496	0.494	348.3	0.453	0.448	96.9	0.501	0.496	117.7

Note: Coeff. stands for the coefficient estimates, t-stat is the t-statistics with Newey-West standard errors, p-val is the p-value of the t-statistic, $\boldsymbol{R^2}$ is the R-squared (coefficient of determination) of the model, $\boldsymbol{\bar{R}^2}$ is the adjusted R-squared of the model, and $\boldsymbol{F}$ is the F-statistic fort he overall significance of the model. **Bold** font indicates significance at the 5% significance level, and *italic* font indicates significance at 10%.

In Table 3, we observe the results for the model with jump dynamics (see Equations (8) and (9)). Here as well, we omit the dummy variables. One consequential issue arises in the level models. While the continuous part of the volatility does not significantly affect the RV in any frequency, the jump component only influences the RV in the highest frequency. That is, the RV is affected by the nearest jump value. This result is not present in the log models. While the jump components significantly affect the RV at session and month levels, and this effect is weakened at day and week levels, the continuous part is significant at every frequency level. We observe that the leverage effect is somewhat significant again both in the level and log models. We see that the weakly average of positive returns is significant at a 5% level, and its coefficient is positive. Therefore, an increase in returns in the medium term (i.e., medium-term shock) positively affects the RV. The same is also valid under the log model; however, the significance level is weaker than 10%. We again observe that the weakly average negative returns are significant at a 10% significance value in the level and log models. Moreover, the monthly average of negative returns is significant at a 10% level as well. The coefficients of all these values are negative. Because the coefficients are weakly significant, the economic effect of these variables on the realized volatility is minor. In other words, medium and long-term shocks have a weak effect on the dynamics of the realized volatility under the log model. Finally, the R^2 value of all models is similar to what we observed in the analysis of the previous model.

Table 3. The OLS Estimation Results for the Models with Jump Dynamics and without Dummy Interventions

	HAR-CJ			logHAR-CJ			LHAR-CJ			logLHAR-CJ		
	Coef.	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Intercept	**0.07**	4.19	0.00	**-0.74**	-3.65	0.00	0.06	3.42	0.00	**-1.01**	-4.79	0.00
$\boldsymbol{log(C_t)}$	——	——	——	**-0.12**	-2.30	0.02	——	——	——	**-0.13**	-2.54	0.01
$\boldsymbol{log(C_t)^{(d)}}$	——	——	——	**0.17**	2.83	0.00	——	——	——	**0.17**	2.77	0.01
$\boldsymbol{log(C_t)^{(w)}}$	——	——	——	**0.25**	3.17	0.00	——	——	——	**0.25**	3.29	0.00
$\boldsymbol{log(C_t)^{(m)}}$	——	——	——	**-0.18**	-2.35	0.02	——	——	——	**-0.22**	-2.98	0.00
$\boldsymbol{log(1+J_t)}$	——	——	——	**0.32**	5.05	0.00	——	——	——	**0.31**	5.05	0.00
$\boldsymbol{log(1+J_t)^{(d)}}$	——	——	——	*0.41*	1.69	0.09	——	——	——	*0.40*	1.68	0.09
$\boldsymbol{log(1+J_t)^{(w)}}$	——	——	——	-0.03	-0.09	0.92	——	——	——	-0.19	-0.71	0.48
$\boldsymbol{log(1+J_t)^{(m)}}$	——	——	——	**1.24**	3.70	0.00	——	——	——	**1.52**	4.53	0.00
$\boldsymbol{r_t^+}$	——	——	——	——	——	——	-0.01	-0.27	0.79	0.01	0.42	0.67
$\boldsymbol{r_t^{+(d)}}$	——	——	——	——	——	——	-0.02	-0.59	0.56	0.00	0.00	1.00
$\boldsymbol{r_t^{+(w)}}$	——	——	——	——	——	——	**0.14**	2.12	0.03	*0.21*	1.83	0.07
$\boldsymbol{r_t^{+(m)}}$	——	——	——	——	——	——	-0.15	-1.18	0.24	-0.22	-0.78	0.44
$\boldsymbol{r_t^-}$	——	——	——	——	——	——	-0.03	-0.80	0.42	-0.04	-0.92	0.36
$\boldsymbol{r_t^{-(d)}}$	——	——	——	——	——	——	-0.02	-0.38	0.70	-0.03	-0.61	0.54
$\boldsymbol{r_t^{-(w)}}$	——	——	——	——	——	——	*-0.09*	-1.67	0.09	*-0.12*	-1.67	0.10
$\boldsymbol{r_t^{-(m)}}$	——	——	——	——	——	——	0.04	0.27	0.78	*-0.34*	-1.89	0.06
$\boldsymbol{C_t}$	0.28	0.35	0.73	——	——	——	0.13	0.18	0.85	——	——	——
$\boldsymbol{C_t^{(d)}}$	0.57	0.88	0.38	——	——	——	0.58	0.87	0.38	——	——	——
$\boldsymbol{C_t^{(w)}}$	0.39	0.35	0.73	——	——	——	0.01	0.00	1.00	——	——	——
$\boldsymbol{C_t^{(m)}}$	0.24	0.28	0.78	——	——	——	0.58	0.53	0.59	——	——	——
$\boldsymbol{J_t}$	**0.34**	2.70	0.01	——	——	——	**0.33**	2.59	0.01	——	——	——
$\boldsymbol{J_t^{(d)}}$	0.08	0.55	0.58	——	——	——	0.09	0.66	0.51	——	——	——
$\boldsymbol{J_t^{(w)}}$	0.20	1.00	0.32	——	——	——	0.21	0.98	0.33	——	——	——
$\boldsymbol{J_t^{(m)}}$	0.13	0.66	0.51	——	——	——	0.12	0.52	0.61	——	——	——
	$\boldsymbol{R^2}$	$\boldsymbol{\bar{R}^2}$	$\boldsymbol{F}$	$\boldsymbol{R^2}$	$\boldsymbol{\bar{R}^2}$	$\boldsymbol{F}$	$\boldsymbol{R^2}$	$\boldsymbol{\bar{R}^2}$	$\boldsymbol{F}$	$\boldsymbol{R^2}$	$\boldsymbol{\bar{R}^2}$	$\boldsymbol{F}$
	0.447	0.444	142.6	0.500	0.498	176.6	0.454	0.448	72.8	0.506	0.501	89.9

To increase the predictability of the volatility models and examine the effect of trading days and trading sessions, we include the dummy variables in the regression equations. The results associated with this construction are given in Tables 4 and 5. In Table 4, we consider the dynamics without jump components, and in Table 5, we modify the regression equation with jump dynamics.

The inclusion of the dummy variables does not alter the previous significant results drastically. That is, the non-dummy variables that were found to be significant in the analysis of the results in Table 2 are again significant here as well. Insignificant variables remain the same. From the analysis of the dummy variables, we observe that only the session dummy is significant with a negative sign in all models. This issue indicates that the realized volatility is lower in the morning session. This finding is in contrast to that in Bildik (2001), asserting that volatility (of Istanbul Stock Exchange) follows an L-shaped path. According to Bildik (2001), the volatility of the intra-daily returns is higher in the morning session and diminishes towards the closing between the years1996 and

1999. From our findings, it seems the dynamics of RV in between the sessions changed through the years.

Table 4. The OLS Estimation Results for the Models without Jump Dynamics and with Dummy Variable Interventions

	HAR-RV			logHAR-RV			LHAR-RV			logLHAR-RV		
	Coef.	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Intercept	**0.08**	3.71	0.00	**-0.08**	-2.50	0.01	**0.07**	2.97	0.00	**-0.17**	-3.73	0.00
$log(RV_t)$	----	----	----	**0.23**	5.21	0.00	----	----	----	**0.22**	4.72	0.00
$log(RV_t)^{(d)}$	----	----	----	**0.29**	4.60	0.00	----	----	----	**0.29**	4.52	0.00
$log(RV_t)^{(w)}$	----	----	----	**0.28**	5.64	0.00	----	----	----	**0.24**	4.69	0.00
$log(RV_t)^{(m)}$	----	----	----	0.05	0.93	0.35	----	----	----	0.06	1.10	0.27
r_t^+	----	----	----	----	----	----	0.00	-0.03	0.97	0.01	0.28	0.78
$r_t^{+(d)}$	----	----	----	----	----	----	-0.02	-0.52	0.61	-0.02	-0.43	0.67
$r_t^{+(w)}$	----	----	----	----	----	----	**0.15**	2.32	0.02	*0.19*	1.78	0.08
$r_t^{+(m)}$	----	----	----	----	----	----	-0.16	-1.20	0.23	-0.14	-0.54	0.59
r_t^-	----	----	----	----	----	----	-0.03	-0.63	0.53	-0.03	-0.71	0.48
$r_t^{-(d)}$	----	----	----	----	----	----	-0.02	-0.45	0.65	-0.02	-0.57	0.57
$r_t^{-(w)}$	----	----	----	----	----	----	-0.10	-1.61	0.11	*-0.13*	-1.76	0.08
$r_t^{-(m)}$	----	----	----	----	----	----	0.01	0.07	0.94	-0.23	-1.45	0.15
RV_t	**0.37**	3.31	0.00	----	----	----	**0.34**	3.84	0.00	----	----	----
$RV_t^{(d)}$	0.11	1.22	0.22	----	----	----	0.13	1.45	0.15	----	----	----
$RV_t^{(w)}$	**0.24**	2.90	0.00	----	----	----	**0.18**	2.38	0.02	----	----	----
$RV_t^{(m)}$	**0.17**	3.22	0.00	----	----	----	**0.21**	3.14	0.00	----	----	----
$D_{t,s}$	**-0.07**	-5.48	0.00	**-0.15**	-7.12	0.00	**-0.07**	-5.73	0.00	**-0.15**	-7.26	0.00
$D_{t,tue}$	0.01	0.70	0.48	0.02	0.60	0.55	0.01	0.66	0.51	0.02	0.65	0.52
$D_{t,wed}$	0.02	1.12	0.26	0.02	0.77	0.44	0.02	0.96	0.34	0.02	0.77	0.44
$D_{t,thu}$	-0.01	-0.44	0.66	-0.03	-1.09	0.27	-0.01	-0.64	0.52	-0.03	-1.16	0.25
$D_{t,fri}$	0.02	0.91	0.36	0.00	-0.15	0.88	0.01	0.71	0.48	-0.01	-0.30	0.77
	R^2	$\bar{R}^2$	F	R^2	$\bar{R}^2$	F	R^2	$\bar{R}^2$	F	R^2	$\bar{R}^2$	F
	0.459	0.456	133.1	0.516	0.513	166.9	0.467	0.460	72.1	0.521	0.515	89.6

In Table 5, we consider the jump component of the volatility measure. The findings show similar patterns as in Table 3; there are minor changes in the log models, and the level models do not exhibit changes in significance properties and sign of the coefficients. In both log models, the session-based effect of the continuous part of the realized volatility becomes insignificant. The daily effect of the same variables becomes significant only at the 10% significance level. Furthermore, as in Table 4, we see that only the session dummy is significant among all other dummy variables. Given that the sign of its coefficient is negative, the interpretation remains the same.

Table 5. The OLS Estimation Results for the Models with Jump Dynamics and Dummy Interventions

	HAR-CJ			logHAR-CJ			LHAR-CJ			logLHAR-CJ		
	Coef.	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Intercept	**0.10**	5.23	0.00	**-0.66**	-3.21	0.00	**0.09**	4.37	0.00	**-0.93**	-4.11	0.00
$log(C_t)$	----	----	----	-0.06	-1.26	0.21	----	----	----	-0.07	-1.47	0.14
$log(C_t)^{(d)}$	----	----	----	*0.11*	1.87	0.06	----	----	----	*0.11*	1.79	0.07
$log(C_t)^{(w)}$	----	----	----	**0.25**	3.13	0.00	----	----	----	**0.25**	3.31	0.00
$log(C_t)^{(m)}$	----	----	----	**-0.17**	-2.15	0.03	----	----	----	**-0.22**	-2.64	0.01
$log(1+J_t)$	----	----	----	**0.33**	5.39	0.00	----	----	----	**0.32**	5.29	0.00
$log(1+J_t)^{(d)}$	----	----	----	0.38	1.64	0.10	----	----	----	0.37	1.62	0.11

$log(1+J_t)^{(w)}$	——	——	——	-0.01	-0.04	0.97	——	——	——	-0.19	-0.72	0.47
$log(1+J_t)^{(m)}$	——	——	——	**1.23**	3.87	0.00	——	——	——	**1.51**	4.65	0.00
r_t^+	——	——	——	——	——	——	0.00	-0.18	0.86	0.02	0.52	0.60
$r_t^{+(d)}$	——	——	——	——	——	——	-0.02	-0.60	0.55	-0.01	-0.20	0.84
$r_t^{+(w)}$	——	——	——	——	——	——	**0.14**	2.29	0.02	**0.22**	1.93	0.05
$r_t^{+(m)}$	——	——	——	——	——	——	-0.15	-1.20	0.23	-0.22	-0.76	0.44
r_t^-	——	——	——	——	——	——	-0.03	-0.69	0.49	-0.03	-0.81	0.42
$r_t^{-(d)}$	——	——	——	——	——	——	-0.01	-0.30	0.76	-0.02	-0.55	0.58
$r_t^{-(w)}$	——	——	——	——	——	——	*-0.10*	-1.67	0.09	*-0.13*	-1.85	0.06
$r_t^{-(m)}$	——	——	——	——	——	——	0.04	0.28	0.78	*-0.34*	-1.85	0.06
C_t	0.52	0.63	0.53	——	——	——	0.38	0.49	0.62	——	——	——
$C_t^{(d)}$	0.32	0.48	0.63	——	——	——	0.36	0.51	0.61	——	——	——
$C_t^{(w)}$	0.41	0.35	0.72	——	——	——	0.01	0.01	1.00	——	——	——
$C_t^{(m)}$	0.21	0.24	0.81	——	——	——	0.56	0.53	0.60	——	——	——
J_t	**0.35**	2.74	0.01	——	——	——	**0.34**	2.74	0.01	——	——	——
$J_t^{(d)}$	0.06	0.41	0.68	——	——	——	0.08	0.50	0.61	——	——	——
$J_t^{(w)}$	0.20	1.01	0.31	——	——	——	0.21	1.00	0.32	——	——	——
$J_t^{(m)}$	0.13	0.67	0.50	——	——	——	0.12	0.54	0.59	——	——	——
$D_{t,s}$	**-0.07**	-4.80	0.00	**-0.15**	-7.05	0.00	-0.07	-5.27	0.00	**-0.15**	-7.04	0.00
$D_{t,tue}$	0.01	0.63	0.53	0.03	0.96	0.33	0.01	0.61	0.54	0.03	1.11	0.27
$D_{t,wed}$	0.02	1.01	0.31	0.03	1.32	0.19	0.01	0.91	0.36	0.04	1.48	0.14
$D_{t,thu}$	-0.01	-0.53	0.60	-0.02	-0.60	0.55	-0.01	-0.71	0.48	-0.02	-0.56	0.57
$D_{t,fri}$	0.02	0.91	0.36	0.00	-0.01	0.99	0.01	0.71	0.48	0.00	-0.11	0.92
	R^2	$\bar{R}^2$	F	R^2	$\bar{R}^2$	F	R^2	$\bar{R}^2$	F	R^2	$\bar{R}^2$	F
	0.462	0.457	92.7	0.520	0.515	117.1	0.468	0.460	58.5	0.526	0.518	73.7

Finally, in Table 6, we compare the in-sample fit of the alternative models by using the Modified Akaike Information Criteria (AICc) and Adjusted R-squared adapted to the log transformation. One immediate property for the AICc is that the log models have positive, and the level models have negative AICc values. This difference can be natural since these two types of models cannot be directly comparable. However, we can use AICc to compare the models with the same dependent variables. In this regard, the best level model is the LHAR-RV with dummy variables, and the best log model is the LHAR-CJ model with dummy variables. Furthermore, when we check the adjusted R-squared values, the overall best models are the LHAR-CJ and LHAR-RV models with dummy variables. However, it is worth noting that other level models are also performing well.

Table 6. Modified AIC of the Estimated models

	AICc		Adjusted R2	
	w/o Dummy	w Dummy	w/o Dummy	w Dummy
HAR-RV	-92.231	-119.235	0.443	0.456
logHAR-RV	1312.543	1265.810	0.340	0.354
LHAR-RV	-96.358	**-122.525**	0.448	**0.460**
logLHAR-RV	1314.521	1267.954	0.359	0.370
HAR-CJ	-90.103	-117.302	0.444	0.457
logHAR-CJ	1308.545	1262.452	0.380	0.389
LHAR-CJ	-90.930	-117.007	0.448	**0.460**
logLHAR-CJ	1308.110	**1261.963**	0.396	0.402

B. Out-of-sample Forcast Comparison

In this section, we compare the forecasting performance of the realized volatility models. For comparison, we split the data set into training (in-sample) and testing (out-of-sample) periods. In the testing period, we generate static forecasts of the realized volatility. We choose one financial year as the testing period (252 days). The remaining data is used for the estimation. After obtaining the forecasts, we compute the root mean squared error (RMSE), mean absolute error (MAE), mean percentage error (MPE) and mean absolute percentage error (MAPE). These statistics are suggested by Hyndman and Koehlercan (2006), and they can be represented as,

$$
\begin{aligned}
RMSE &= \sqrt{\frac{1}{T_f}\sum_{t=T-T_f+1}^{T}\left(RV_{t+1|t} - RV_{t+1}\right)^2}\ ; \\
MAE &= \frac{1}{T_f}\sum_{t=T-T_f+1}^{T}\left|RV_{t+1|t} - RV_{t+1}\right|\ ; \\
MPE &= \frac{1}{T_f}\sum_{t=T-T_f+1}^{T}\left[100 \times \left(RV_{t+1|t} - RV_{t+1}\right)/RV_{t+1}\right]\ ; \\
MPE &= \frac{1}{T_f}\sum_{t=T-T_f+1}^{T}\left|100 \times \left(RV_{t+1|t} - RV_{t+1}\right)/RV_{t+1}\right|,
\end{aligned}
\qquad \textbf{(10)}
$$

where $T_f = 252$ is the sample size of the out-of-sample period, $RV_{t+1|t}$ is the one step ahead forecast of RV_{t+1}.

According to RMSE, the best forecast model is the LHAR-RV with dummy variables. On the one hand, the MPE criterion selects the best model as the logLHAR-CJ without dummy variables. On the other hand, both MAE and MAPE, which are based on absolute deviations, point the LHAR-CJ with dummy variables. Overall, we see that the results with dummy variables exhibit lower errors, hence better forecasts. Even though there is no absolute winner of the forecast comparison, we see that the LHAR-CJ with dummy variables is selected twice as the winning model. Therefore, we can conclude that in modeling the realized volatility of the BIST 100 index, taking the leverage, jump, and structural change (through dummy variables) effects into account provides better performance.

Table 2. Forecasting Performance of Different Realized Volatility Measure

		RMSE	MAE	MPE	MAPE
w/o dummy	HAR-RV	0.3339	0.2069	-7.9471	28.9014
	logHAR-RV	0.3461	0.2044	-8.1164	28.3019
	LHAR-RV	0.3320	0.2049	-7.4983	28.5678
	logLHAR-RV	0.3386	0.2011	-7.8248	27.9219
	HAR-CJ	0.3400	0.2040	-5.3950	27.8016
	logHAR-CJ	0.3448	0.2142	**-11.0335**	30.4127
	LHAR-CJ	0.3482	0.2009	0.8377	25.9602
	logLHAR-CJ	0.3438	0.2102	-6.5078	28.8005
w dummy	HAR-RV	0.3302	0.2030	-7.7598	28.2545
	logHAR-RV	0.3420	0.2002	-7.6177	27.4715
	LHAR-RV	**0.3283**	0.2012	-7.3299	27.9404
	logLHAR-RV	0.3351	0.1964	-7.3265	27.0111
	HAR-CJ	0.3359	0.2000	-5.2305	27.1505
	logHAR-CJ	0.3417	0.2095	-10.4933	29.5216
	LHAR-CJ	0.3440	**0.1963**	0.9527	**25.2573**
	logLHAR-CJ	0.3417	0.2045	-5.9114	27.7687

IV.Conclusion

In this paper, we consider the modeling of the realized volatility of the BIST 100 index. For the analysis, we consider four models that take the leverage and jump effects of returns into account. The models we consider are a slightly modified version of those employed in Corsi (2009) and Corsi and Reno (2009, 2012). The first one we utilize is a simple model without the consideration of leverage effect and jumps. Then, we construct a second model by embedding the leverage effect to capture the asymmetry in the dynamics of volatility. In the third model, we consider jumps only, and in the final model, we consider both leverage and jumps.

After the construction of the models, we augment the number of models to be used for the analysis based on the specification of the function $g(\cdot)$ and the consideration of the dummy variables. In this respect, we consider sixteen model specifications. Our results show that the models explain the variation in the dynamics of the realized volatility quite well; based on R-squared values, almost half of the variation is explained. We also see that the level models (i.e., when $g(\mathrm{x}) = x$) have long-memory property, while the log models (i.e., when $g(\mathrm{x}) = \log(x)$) have shorter memory property relative to its level counterpart. Besides, the leverage effect in the BIST 100 volatility, although present, is not very strong. Thus, the response to shocks is not entirely asymmetric. We also observe that jumps significantly affect the BIST 100 realized volatility both at the short-term and long-term levels depending on the specification of the model. Finally, we see that only the session-level dummy variable is significant. The result shows that the volatility of the BIST 100 is lower in the morning session. This finding contrasts with the L-shape assertion of Bildik (2001).

For the model selection exercise based on the AICc values, the LHAR-CJ and LHAR-RV models with dummy variables are overall the best models. In effect, from the out-of-sample forecast exercise, we see that the models with dummy variables usually perform better than those without the dummy variables. Particularly, the forecast exercise shows that the LHAR-CJ model is selected the most based on the forecast evaluation measures employed in the study. Thus, we conclude that the level based HAR models with leverage and jump effects perform the best in predicting the realized volatility of the BIST 100 index.

References

Andersen, T. G., & Bollerslev, T. (1998). Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. The Journal of Finance, 53(1), 219-265.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2010). Parametric and nonparametric volatility measurement. In Handbook of Financial Econometrics: Tools and Techniques (pp. 67-137). North-Holland.

Fischer, B. (1976). Studies of stock price volatility changes. In Proceedings of the Business and Economic Statistics Section (pp. 177-181).

Bajgrowicz, P., Scaillet, O., & Treccani, A. (2016). Jumps in high-frequency data: Spurious detections, dynamics, and news. Management Science, 62(8), 2198-2217.

Bekaert, G., & Harvey, C. R. (1997). Emerging equity market volatility. Journal of Financial Economics, 43(1), 29-77.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 74(1), 3-30.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Estimating quadratic variation using realized variance. Journal of Applied econometrics, 17(5), 457-477.

Barndorff-Nielsen O. E., Shephard N., (2006). Econometrics of Testing for Jumps in Financial Economics Using Bi-power Variation, Journal of Financial Econometrics, 1(4), 1-30.

Bildik, R. (2001). Intra-day seasonalities on stock returns: evidence from the Turkish Stock Market. Emerging Markets Review, 2(4), 387-417.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81(3), 637-654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics, 7(2), 174-196.

Corsi, F., and Reno, R. (2009). HAR volatility modeling with heterogeneous leverage and jumps. Working paper, SSRN.

Corsi, F., & Reno, R. (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. Journal of Business & Economic Statistics, 30(3), 368-380.

Corsi, F., Pirino, D., & Renò, R. (2009). Volatility forecasting: The jumps do matter. Working paper.

Corsi, F., Pirino, D., & Reno, R. (2010). Threshold bi-power variation and the impact of jumps on volatility forecasting. Journal of Econometrics, 159(2), 276-288.

Christoffersen, P. F., & Diebold, F. X. (2000). How relevant is volatility forecasting for financial risk management?. Review of Economics and Statistics, 82(1), 12-22.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the Econometric Society, 987-1007.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance, 48(5), 1779-1801.

Hyndman, R.J. and Koehler, A.B. (2006) "Another look at measures of forecast accuracy". International Journal of Forecasting, 22(4), 679-688.

Inci, A. C., & Ozenbas, D. (2017). Intraday volatility and the implementation of a closing call auction at Borsa Istanbul. Emerging Markets Review, 33, 79-89.

Jiang, G. J., & Oomen, R. C. (2008). Testing for jumps when asset prices are observed with noise–a "swap variance" approach. Journal of Econometrics, 144(2), 352-370.

Karanasos, M., Psaradakis, Z., & Sola, M. (2004). On the Autocorrelation Properties of Long-Memory GARCH Processes. Journal of Time Series Analysis, 25(2), 265-282.

Köksal, B. (2012). An analysis of intraday patterns and liquidity on the Istanbul stock exchange.

Lagoarde-Segot, T., & Lucey, B. M. (2007). Capital market integration in the Middle East and North Africa. Emerging Markets Finance and Trade, 43(3), 34-57.

Lee, S. S., & Mykland, P. A. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. The Review of Financial Studies, 21(6), 2535-2563.

Levine, R., & Zervos, S. (1998). Stock markets, banks, and economic growth. American Economic Review, 88(3), 537-558.

Mandelbrot, B. (1963). The variation of certain speculative prices. The Journal of Business, 45(4), 542-543.

Merton, R. C. (1975). Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory 3(4), 621-661.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the Econometric Society 59(2), 347-370.

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation. Econometrica, 55(3), 703-708.

Temizel, F. (2008). İstanbul Menkul Kıymetler Borsasında Hisse Senedi Fiyatlarında Gün İçi Yapılar. Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 22(1), 475-495.