Forecasting Daily Stock Volatility Using GARCH-CJ Type Models with Continuous and Jump Variation

Issam BOUSALAM*[a], Moustapha HAMZAOUI[b], and Otman ZOUHAYR[c]

[a,b]Department of Economics, Abdelmalek Essaädi University of Tangier - Morocco
[c]Albanki-Alternative Banking Institute - France

January 20, 2016

Abstract

In this paper we decompose the realized volatility of the GARCH-RV model into continuous sample path variation and discontinuous jump variation to provide a practical and robust framework for non-parametrically measuring the jump component in asset return volatility. By using 5-minute high-frequency data of MASI Index in Morocco for the period (January 15, 2010 - January 29, 2016), we estimate parameters of the constructed GARCH and EGARCH-type models (namely, GARCH, GARCH-RV, GARCH-CJ, EGARCH, EGARCH-RV, and EGARCH-CJ) and evaluate their predictive power to forecast future volatility. The results show that the realized volatility and the continuous sample path variation have certain predictive power for future volatility while the discontinuous jump variation contains relatively less information for forecasting volatility. More interestingly, the findings show that the GARCH-CJ-type models have stronger predictive power for future volatility than the other two types of models. These results have a major contribution in financial practices such as financial derivatives pricing, capital asset pricing, and risk measures.

JEL-Classification: C22, F37, F47, G17.

Keywords: GARCH-CJ; Jumps variation; Realized volatility; MASI Index; Morocco.

*For correspondence and data request (`boussalam.issam@gmail.com`)

1 Introduction

A common finding in much of the empirical finance literature is that asset returns volatility exhibits "clustering" and "persistence" features. This is why Engle (1982) proposed the AutoRegressive Conditional heteroskedasticity (ARCH) model which was generalized later by Bollerslev (1986) to take into account bigger regression order and proposed the GARCH model. Nelson (1991) found that the asset volatility is "asymmetric" relatively to bad and good news on the market, then he modified the GARCH model and built an exponential GARCH model (EGARCH). These models (GARCH and EGARCH) were found to be more powerful in predicting future volatility (Andersen and Bollerslev, 1998).

Despite the fact that GARCH style models have been continuously proved to be stronger for predicting asset returns volatility, seeking to improve the accuracy of future volatility prediction is an endless process and constitutes the premise of quantitative financial analysis. This is because measuring and predicting accurately the asset returns volatility has too much practical uses in financial asset pricing, financial derivative pricing, and financial risk management.

In order to enhance the accuracy of volatility forecasting, Koopman et al. (2005) introduced the realized volatility (RV) as an exogenous variable into the volatility equation of GARCH model. They built a GARCH-RV model and found that the it has stronger predictive power than the traditional GARCH model. The same results were found by Fuertes et al. (2009) and Frijns et al. (2011).

But in realistic financial markets, the process of asset volatility is not completely continuous but contains some jump components. In fact, Andersen et al. (2007) and Huang et al. (2013) studied the HAR-type RV model and found that model built with continuous sample path variation and discontinuous jump variation that decomposed from RV has stronger power than the undecomposed HAR-RV model in measuring and predicting the asset volatility.

Based on these findings, we estimate that it makes sense to split the exogenous variable RV introduced in GARCH-RV model into a continuous sample path variation and discontinuous jumps variation in order to further enhance the predictive power of GARCH-RV model. Similarly, in this paper we will also extend the EGARCH model to an EGARCH-RV model and an EGARCH-CJ model. Next, we estimate parameters of the above mentioned models and evaluate their forecasting power for the future volatility to identify which volatility model has stronger power for the asset volatility measurement and prediction. This by using the 5-minute high-frequency data of the broad based Moroccan All Shares Index for a 5 years period ranging from January 15, 2010 to January 29, 2016.

The remaining of this paper is as follows, Section 2. discusses the construction of the GARCH-CJ-type models, Section 3. presents the empirical results of parameters estimation and predictive power evaluation, and Section 4. serves to conclude.

2 Model Specification

2.1 GARCH-CJ Model building

2.1.1 GARCH-RV Model Construction

Stock return volatility cannot be directly observable but can be measured in the asset return series. Financial literature shows that return volatility is "clustering" and "persistent" over time. Engle (1982) proposed the AutoRegressive Conditional heteroskedasticity (ARCH) model that captures the clustering feature and Bollerslev (1986) generalized it to take into account bigger regression order and proposed the GARCH model. Scholars generally use the GARCH(1,1) model described by:

$$\begin{gathered} r_t = \ln\left(\frac{I_t}{I_{t-1}}\right) = E(r_t|\Psi_{t-1}) + \epsilon_t \\ \epsilon_t = \sigma_t \cdot z_t \quad , \quad z_t \sim \psi(0,1,v) \\ \sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \end{gathered} \tag{1}$$

I_t is the price of the index at time t and Ψ_{t-1} contains all information up to day $t-1$. ϵ_t are the random innovations (surprises) with $E(\epsilon_t) = 0$ and they are split into a white noise disturbance z_t and a time-dependent standard deviation σ_t characterizing the typical size of the error terms. $\psi(.)$ marks a conditional density function and v denotes a vector of parameters needed to specify the probability distribution of z_t. σ_t is the volatility and ω,α and β are parameters to be estimated.

Seeking to improve the explanatory and the predictive power of the traditional GARCH model, Koopman et al. (2005) incorporated the Realized Volatility (RV) as an exogenous variable into the volatility model GARCH(1,1) and built the GARCH-RV model expressed as follows:

$$\begin{gathered} r_t = E(r_t|\Psi_{t-1}) + \epsilon_t \quad , \quad \epsilon_t = \sigma_t \cdot z_t \\ \sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \lambda RV_{t-1} \end{gathered} \tag{2}$$

λ is a parameter to be estimated as for ω,α and β, and RV_{t-1} is the realized volatility at time $t-1$. Martens (2002) and Koopman et al. (2005) emphasized the importance of using high-frequency intraday returns to the measuring and forecasting of volatility and expressed the realized volatility as a function of overnight return variance.

$$RV_t = \sum_{i=1}^{N} r_{t,i}^2 + r_{t,n}^2 = \sum_{j=1}^{M} r_{t,j}^2 \quad , \quad M = N+1 \tag{3}$$

By assuming N equally divided parts of a trading day, $r_{t,1}$ represents the log-return for the first period (part) of the day where $r_{t,1} = \ln(I_{t,1}/I_{t,0})$ and $I_{t,0}$ is the opening price at Day t, $r_{t,2}$ is the log-return for the second period; ..., and $r_{t,N}$ expresses the N^{th} return at Day t. Finally, $r_{t,n} = r_{t,M} = \ln(I_{t,1}/I_{t-1,c})$ where $I_{t-1,c}$ is the closing price in Day $t-1$.

2.1.2 GARCH-CJ Model Construction

There is empirical evidence that stock markets exhibit fractal features and financial asset price volatility is not continuous but rather generated by a jump process. The nonlinear properties of the stock market volatility is almost due to big information shocks and investors' irrational behaviors. Therein, in order to improve the predictive power of the GARCH-RV model, Andersen et al. (2007) decomposed the realized volatility (RV) in model (2) into a continuous sample path variation denoted C_j and a discontinuous jump variation J_t.

Alternatively, Barndorff-Nielsen and Shephard (2006) introduced the Realized Bipower Variation (RBV) with more robustness properties described by:

$$RBV_t^{[r,s]} = \left\{ \left(\frac{h}{M}\right)^{1-(r+s)/2} \right\} \sum_{j=1}^{M-1} \left|r_{j,t}\right|^r \left|r_{j+1,t}\right|^s, \qquad r, s \geq 0. \tag{4}$$

Where r and s are constants[1], h is a fix time interval and M is the sample frequency within interval h. Barndorff-Nielsen and Shephard (2006) demonstrated that when a stochastic volatility and an infrequent jumps process exist, then the difference between RV and RBV estimates the quadratic variation of the jump component Jt when $M \to \infty$.

$$RV_t - RBV_t \xrightarrow{M\to\infty} J_t. \tag{5}$$

Given a limited sample size, the jumps variation Jt calculated in (5) may not be always positive and to overcome this issue, we treat J_t in the following way:

$$J_t = \text{Max}[RV_t - RBV_t, 0]. \tag{6}$$

When calculating the discontinuous jumps variation J_t a problem of accuracy occurs for an intraday data sampled at unequal frequency. This is why Barndorff-Nielsen and Shephard (2006) introduced a Z_t statistic to test for J_t. Z_t is described by:

$$Z_t = \frac{(RV_t - RBV_t)RV_t^{-1}}{\sqrt{\text{Max}(1, RTQ_t/RBV_t^2)(1/M)((\pi/2) + \pi - 5)}} \longrightarrow \mathcal{N}(0,1). \tag{7}$$

Where

$$RTQ_t = M\mu_{4/3}^{-3}\left(\frac{M}{M-4}\right)\sum_{j=4}^{M} \left|r_{t,j-4}\right|^{4/3}\left|r_{t,j-2}\right|^{4/3}\left|r_{t,j}\right|^{4/3}, \tag{8}$$

$$\left(\mu_{4/3} = \text{E}\left(|Z_t|^{4/3}\right) = 2^{2/3}\Gamma\left(\frac{7}{6}\right)\Gamma\left(\frac{1}{2}\right)^{-1}\right).$$

RTQ_t is the Realized Tripower Quarticity which is an asymptotically unbiased estimator of integrated quarticity in the absence of microstructure noise.

The calculation of RBV_t relies mainly on the sampling frequency of intraday data which might result in some convergence issues when the sampling frequency is sufficiently high. This is due to

[1] Usually $r = s = 1$ is given so that $RBV_t^{[1,1]} = \sum_{j=1}^{M-1} |r_{j,t}|^r |r_{j+1,t}|^s$

several factors and one of these is the market microstructure. Andersen et al. (2012) introduced the Median Realized Volatility ($MedRV_t$) as a robust estimator for J_t instead of the biased RV_t. The alternative $MedRV_t$ uses two-sided truncation, picking the median of three adjacent absolute returns and is expressed by (9). Similarly, RTQ_t used for Z_t calculation in (7) is replaced by $MedRTQ_t$ described hereafter by (10).

$$MedRV_t = \frac{\pi}{6-4\sqrt{3}+\pi}\left(\frac{M}{M-2}\right) \times \sum_{i=2}^{M-1} Med(|r_{t,i-1}|,|r_{t,i}|,|r_{t,i+1}|)^2 \tag{9}$$

$$MedRTQ_t = \frac{3\pi M}{9\pi+72-52\sqrt{3}}\left(\frac{M}{M-2}\right) \times \sum_{i=2}^{M-1} Med(|r_{t,i-1}|,|r_{t,i}|,|r_{t,i+1}|)^4. \tag{10}$$

By replacing RV_t and RTQ_t in (7) with $MedRV_t$ and $MedRTQ_t$ respectively, we calculate the Z_t statistic and get the estimator for both discontinuous jump variation J_t and continuous sample path variation C_j at $1-\alpha$ significance level. In this paper, based on previous research, we choose a confidence level α of 99%. J_t and C_t are then defined as:

$$J_t = I(Z_t > \phi_\alpha)(RV_t - MedRV_t), \tag{11}$$

$$C_t = I(Z_t \leq \phi_\alpha)RV_t + I(Z_t > \phi_\alpha)MedRV_t. \tag{12}$$

Finally, according to the above RV_t decomposition into C_t and J_t, the GARCH-RV model in (2) becomes the GARCH-CJ model expressed as follows:

$$\begin{aligned} r_t &= E(r_t|\Psi_{t-1}) + \epsilon_t \quad , \quad \epsilon_t = \sigma_t \cdot z_t \\ \sigma_t^2 &= \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \lambda C_{t-1} + \gamma J_{t-1} \end{aligned} \tag{13}$$

2.2 EGARCH-CJ Model specification

In response to the weakness of traditional GARCH model to capture all the leptokurtosis of the error terms and to handle the asymmetric responses of volatility, Nelson (1991) constructed the Exponential GARCH (EGARCH) model on the basis of the baseline GARCH model. Most commonly, researchers use the EGARCH(1,1) model described by:

$$\begin{aligned} r_t &= E(r_t|\Psi_{t-1}) + \epsilon_t \quad , \quad \epsilon_t = \sigma_t \cdot z_t \\ \ln\sigma_t^2 &= \omega + \alpha\big(|z_{t-1}| - E[|z_{t-1}|]\big) + \beta\ln(\sigma_{t-1}^2) + \theta z_{t-1}. \end{aligned} \tag{14}$$

Following the method discussed in Section 2.1.1, we get the EGARCH-RV model by introducing the log of the one-period-lagged realized volatility (RV_{t-1}). Thus, equation (14) becomes:

$$\ln\sigma_t^2 = \omega + \alpha\big(|z_{t-1}| - E[|z_{t-1}|]\big) + \beta\ln(\sigma_{t-1}^2) + \theta z_{t-1} + \lambda\ln(RV_{t-1}). \tag{15}$$

We split RV_{t-1} into C_{t-1} and J_{t-1}, we take their logarithms and replace them in (15), thus we obtain the EGARCH-CJ model described as follows:

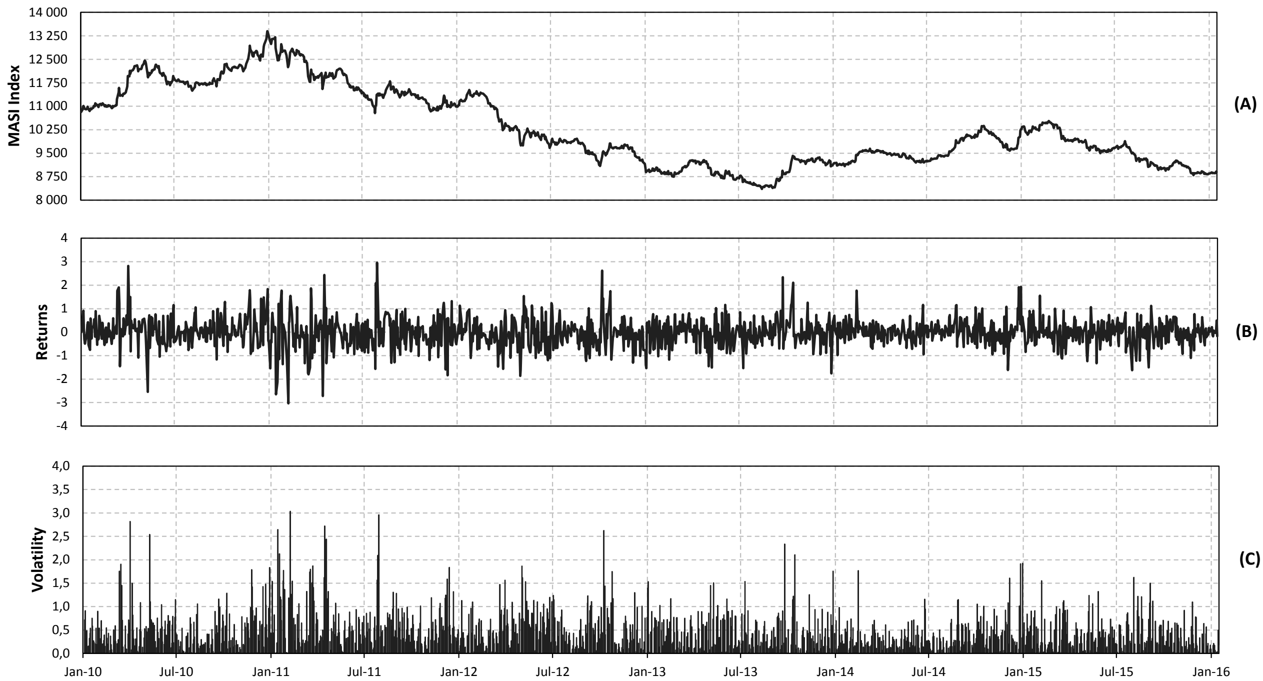
$$r_t = E(r_t|\Psi_{t-1}) + \epsilon_t \qquad , \qquad \epsilon_t = \sigma_t \cdot z_t$$
$$\ln \sigma_t^2 = \omega + \alpha\big(|z_{t-1}| - E[|z_{t-1}|]\big) + \beta \ln(\sigma_{t-1}^2) + \theta z_{t-1} + \lambda \ln(C_{t-1}) + \gamma \ln(J_{t-1} + 1). \tag{16}$$

3 Empirical Results and Comparative Analysis of Models' Predictive Power

3.1 Data and Empirical Properties

3.1.1 Sample Statistics

Our data set is the Moroccan All Shares Index (MASI), recorded at 5 minutes (5-min) intervals during the sample period of January 15, 2010 to January 29, 2016. Data is acquired from Bloomberg®. The Casablanca Stock Exchange opens at 9:30 (GMT) and the first record of the MASI index for that day is registered at 9:31. The market closes at 15:30 (GMT) and the last record of the day is registered at 15:31. Therefore, considering a 5-min intervals during one trading day and by using the moving average interpolation for missed data we obtain 144 daily index records. Overall, our sample period consists of 1,506 days. We eliminated weekends and holidays during which the market was closed.



(A) MASI Index level (B) 5-min returns (log difference, in percent) (C) 5-min volatility (absolute return, in percent). Sample period is January 15, 2010 - January 29, 2016 (216,864 5-mins, 1,506 days). Data source: Bloomberg®

Figure 1: Moroccan All Shares Index (MASI) at 5-min intervals

Table 1 below presents descriptive statistics of all variables needed to estimate the GARCH-type models described before, *i.e.* intraday returns $r_{t,i}$, Realized Volatility RV_t, continuous sample path variation C_t and discontinuous jump variation J_t, and their respective logarithms: $\ln(RV_t)$, $\ln(C_t)$ and $\ln(J_t + 1)$.

Table 1: Summary Statistics of Study's Variables

	Mean	Std. dev.	Skewness	Kurtosis	Jarque-Bera	ADF t-statistic
$r_{t,i}$	-0.0129	0.6003	0.0815	5.655	444.09^{***}	-34.064^{***}
RV_t	1.1639	1.5647	-4.9915	37.873	10615.02^{***}	-18.125^{***}
C_t	0.8523	1.1063	-5.1326	60.345	19221.69^{***}	-11.934^{***}
J_t	0.3116	1.1170	-9.6719	91.238	42360.17^{***}	-22.872^{***}
$\ln(RV_t)$	0.3594	0.3681	-0.5231	3.152	121.65^{***}	-5.166^{***}
$\ln(C_t)$	0.2967	0.2390	-0.3266	2.791	95.95^{***}	-5.710^{***}
$\ln(J_t+1)$	0.1199	0.2476	-3.4885	13.478	1592.14^{***}	-21.246^{***}

(***) denotes significance at 1% level of significance.

We can clearly observe from Table 1 that returns $r_{t,i}$ and realized volatility RV_t are not normally distributed. These are fat-tailed which implies that volatility in Moroccan stock market is high. Furthermore, the ADF t-statistics are all significant at 99% level of confidence, we can easily reject the null hypothesis of unit root existence in the series. This allows us to use the variables for further models analysis and estimation of parameters.

3.1.2 Estimation of Models' parameters and Comparison

The method of estimation adopted in this paper is maximum likelihood, and parameters of the six competing models (GARCH, GARCH-RV, GARCH-CJ, EGARCH, EGARCH-RV and EGARCH-CJ) were estimated under two assumptions for errors distribution, *i.e.* the normal distribution and Student-t distribution. Goodness of fit is compared using the log-likelihood, Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC).

From Table 2 below, by comparing log-likelihood and information criterion AIC and SIC, we can see that the EGARCH-type models (*i.e. EGARCH, EGARCH-RV and EGARCH-CJ*) outperform the GARCH-type models (*i.e. GARCH, GARCH-RV and GARCH-CJ*) in terms of goodness of fit of the data. This means that volatility on the stock market has an asymmetric response relatively to bad news and good news. Furthermore, both of GARCH-type models and EGARCH-type models fit better the data when residuals are assumed to be following a Student-t distribution.

Table 2: Log-likelihood, AIC and SIC for GARCH-type Models and EGARCH-type Models

	Gaussian distribution			Student-t distribution			
	LL	AIC	SIC	LL	AIC	SIC	d.f.
GARCH(1,1)	-1288.76	1.715	1.726	-1245.90	1.659	1.674	5.935^{***}
GARCH-RV	-1262.13	1.729	1.732	-1236.18	1.693	1.711	6.344^{***}
GARCH-CJ	-1238.56	1.753	1.754	-1225.44	1.745	1.737	7.119^{***}
EGARCH(1,1)	-1288.49	1.716	1.730	-1245.54	1.660	1.678	5.926^{***}
EGARCH-RV	-1259.28	1.732	1.733	-1223.66	1.695	1.701	6.845^{***}
EGARCH-CJ	-1254.75	1.754	1.762	1219.25	1.711	1.712	6.731^{***}

Note — LL is the log-likelihood score. d.f. are degrees of freedom of t-distribution and are all significant at 1% level of significance (***). LL, AIC and SIC were calculated using 5-min returns of the MASI Index for the period covering January 15, 2010 to January 29, 2016.

Tables 3 bellow shows that coefficients (λ) of newly added exogenous variables RV_{t-1} and

$\ln(RV_{t-1})$ are all significantly positive at 1% or 5% level of significance. This indicates that volatility in Moroccan stock market exhibits pronounced persistence and last period volatility may affect current period volatility ; this result is consistent with (Koopman et al., 2005). As for the newly GARCH-CJ and EGARCH-CJ models, the coefficients (λ) for C_t are significantly positive at 10% significance level, and the coefficients (γ) for J_t are non significant only when the residual errors in the GARCH-CJ model are assumed to follow a Student-t distribution, otherwise significant.

These estimation results indicate that, in the Moroccan stock market, the lagged continuous sample path variation contains relatively more information for predicting the current volatility, while the lagged discontinuous jump variation contains relatively less information for forecasting. This finding leads us to test for which models has more predictive power for future volatility. Also, the leverage effect is negative (negative estimates for θ), meaning that the volatility in the stock market is more influenced by bad news than good news.

Table 3: Estimates of Parameters for GARCH-type Models and EGARCH-type Models

	Normally distributed residuals			Student-t distributed residuals		
	GARCH	GARCH-RV	GARCH-CJ	GARCH	GARCH-RV	GARCH-CJ
ω	0.0883^{***}	0.0786^{**}	0.0735^{**}	0.0747^{***}	0.1892^{**}	0.1956^{**}
α	0.2355^{***}	-0.2968^{***}	-0.3341^{***}	0.2441^{***}	-0.4219^{***}	-0.3955^{***}
β	0.5273^{***}	0.3541^{**}	0.3917^{***}	0.5655^{***}	0.4123^{***}	0.3963^{**}
λ		0.1254^{**}	0.1349^{**}		0.1784^{**}	0.1996^{*}
γ			0.0533^{*}			0.0378
d.f.				5.935^{***}	6.344^{***}	7.119^{***}
	EGARCH	EGARCH-RV	EGARCH-CJ	EGARCH	EGARCH-RV	EGARCH-CJ
ω	-0.5098^{***}	0.2514^{***}	0.2763^{***}	-0.4873^{***}	0.3649^{***}	0.3821^{***}
α	0.3849^{***}	-0.4159^{***}	-0.4236^{***}	0.3906^{***}	-0.6144^{***}	0.6232^{***}
β	0.8031^{***}	0.7810^{**}	0.7749^{***}	0.8260^{**}	0.6971^{**}	0.6892^{*}
θ	-0.0130	-0.0985	-0.0948	-0.0268	-0.0828	-0.0847^{*}
λ		0.1365^{*}	0.1289^{**}		0.1437^{*}	0.1510^{*}
γ			0.0348^{*}			0.0458^{*}
d.f.				5.926^{***}	6.845^{***}	6.731^{***}

Note — d.f. are degrees of freedom of t-distribution and are all significant at 1% level of significance. Models' parameters are estimated using 5-min returns of the MASI Index for the period covering January 15, 2010 to January 29, 2016. ***, **, and * denote significance at the 1%, 5%, and 10% significance level respectively.

3.2 Forecasting Methodology and Evaluation Criteria

3.2.1 In-Sample Forecasting

In this paper, we use a loss-function to determine whether the GARCH-CJ-type models have better predictive power than GARCH and GARCH-RV-type models. We compare predictive power of these volatility models using four measures, namely, Mean Absolute Error (MAE), Heteroskedasticity-adjusted Mean Absolute Error (HMAE), Root Mean Squared Error (RMSE), and Heteroskedasticity-adjusted Root Mean Squared Error (HRMSE). In general, the smaller are these four statistics, the better is the predictive power of the volatility models. Statistics of MAE, HMAE, RMSE and HRMSE are calculated using formulae in (17). This paper follows the works of Koopman et al. (2005) and Corsi (2009) who used the realized volatility RV as a substitute for the volatility in Day t.

$$
\begin{aligned}
\text{MAE} &= \frac{1}{n}\sum_{i=1}^{n}\left|\sigma_t^2 - \widehat{\sigma}_t^2\right|, \\
\text{HMAE} &= \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\sigma_t^2 - \widehat{\sigma}_t^2}{\sigma_t^2}\right|, \\
\text{RMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\sigma_t^2 - \widehat{\sigma}_t^2)^2}, \\
\text{HRMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\sigma_t^2 - \widehat{\sigma}_t^2}{\sigma_t^2}\right]^2}.
\end{aligned} \tag{17}
$$

Where n denotes the size of the predictive sample, σ_t^2 is the real volatility substituted by RV_t, and $\widehat{\sigma}_t^2$ is the predicted volatility.

Values of in-sample predictive power indexes for the GARCH-type models and EGARCH-type models are listed in Table 4 below.

Table 4: In-Sample Forecast Evaluation

	Errors following normal distribution				Errors following t distribution			
	MAE	HMAE	RMSE	HRMSE	MAE	HMAE	RMSE	HRMSE
GARCH(1,1)	3.5981	0.9837	7.1927	1.3671	3.5647	0.9846	7.2239	1.5410
GARCH-RV	3.5467	0.9216	6.8913	0.9180	3.4988	0.9517	7.2603	1.6131
GARCH-CJ	3.2830	0.8217	6.9516	0.8692	3.4207	0.8946	7.2554	1.6128
EGARCH(1,1)	3.5218	0.9610	7.0220	1.2593	3.5158	0.9126	6.9373	1.5416
EGARCH-RV	3.4894	0.9154	6.8556	1.1346	3.4791	0.8978	6.6210	1.1246
EGARCH-CJ	3.4412	0.8999	6.8373	1.1299	3.4697	0.8615	6.5431	1.1222

Our full sample consists of 216,864 observations (5-min returns) corresponding to 1,506 days from January 15, 2010 to January 29, 2016. GARCH and EGARCH-type models are estimated over the first 195,264 observations of the full sample, i.e. over the period January 15, 2010 to June 15, 2015

Table 4 shows that all values for GARCH-CJ-type models are smaller than that of both GARCH-RV and GARCH type models consecutively. This leads us to conclude that in in-sample volatility forecasting, the GARCH-CJ-type models perform better than their counterparts and have more predictive power. However, when comparing forecasting power of volatility models given normal and student-t distribution for residuals, the findings are mixed and inconclusive regarding which error distribution assumption contributes better to boost the predictive power of the models. See that for the same given model of the six competing models, the four measures when assuming normal distribution for errors are not all smaller (alternatively, higher) than those for a student-t assumption for errors distribution, and judging the predictive power of models relies on which measure is used to for the comparison.

3.2.2 Out-Of-Sample Forecasting

Compared to the in-sample prediction of the models, the results of out-of-sample forecasting are more interesting since they have more practical value. As for the in-sample predictive power evaluation, we divided the full sample of 5-min returns (216,864 observations, January 15, 2010 - January 29, 2016) into two parts. The first part for models parameters estimation covers the period from January 15, 2010 to June 15, 2015, and the second part used for prediction covers the remaining 150 days till January 29, 2016. We still use the same loss function to evaluate the predictive power as for in-sample forecasting.

Table 5 below presents the values for out-of-sample forecasting measures. As in in-sample predictive power evaluation, it is found that GARCH-CJ type-models perform better than GARCH-RV and GARCH type-models for predicting future volatility. Also, the EGARCH type-models has smaller measures values than GARCH type-models which supposes that the former have more predictive power. More interestingly, the assumption for normal distribution of errors allows the GARCH type-models to predict better future volatility. This result is not the same for EGARCH type-models where predictive power measures are not scattered similarly as for the GARCH type-models, and the predictive power judgment depends also here on the measure used for evaluation.

Table 5: In-Sample Forecast Evaluation

	Errors following normal distribution				Errors following t distribution			
	MAE	HMAE	RMSE	HRMSE	MAE	HMAE	RMSE	HRMSE
GARCH(1,1)	0.977	0.965	1.210	1.062	0.972	0.958	1.194	1.048
GARCH-RV	0.971	0.946	1.209	0.988	0.945	0.936	1.183	0.991
GARCH-CJ	0.965	0.937	1.095	0.967	0.923	0.913	1.001	0.984
EGARCH(1,1)	1.002	0.966	1.164	1.059	0.973	0.951	1.189	1.005
EGARCH-RV	0.979	0.954	1.137	0.976	0.939	0.926	1.102	0.975
EGARCH-CJ	0.958	0.929	0.996	0.946	0.914	0.890	0.978	0.972

Our full sample consists of 216,864 observations (5-min returns) corresponding to 1,506 days from January 15, 2010 to January 29, 2016. GARCH and EGARCH type-models are estimated over the first 195,264 observations of the full sample, i.e. over the period January 15, 2010 to June 15, 2015

Based on discussions in sections 3.2.1 and 3.2.2, we conclude that among all the competing models, on top of their best fitting for intraday returns volatility, the GARCH-CJ-type models perform better when forecasting future volatility. Thus, introducing the realized volatility into GARCH model and splitting it into continuous sample path variation (C_t) and discontinuous jumps variation (J_t) enhances the model's explanatory and predictive powers.

4 Concluding remarks

In this paper, we constructed a GARCH-CJ type model with continuous sample path variation and discontinuous jump variation based on the GARCH-RV model introduced by Koopman et al. (2005). In order to test the model's validity, we performed an empirical study using 5-min high-frequency data of the broad based Moroccan All Shares Index (MASI Index) for the period covering January 15, 2010 to January 29, 2016. Then we estimated the parameters of the six competing models, namely, GARCH, GARCH-RV, GARCH-CJ, EGARCH, EGARCH-RV and EGARCH-CJ. We also evaluated each model's predictive power using a loss function by calculating four measures (MAE, HMAE, RMSE, and HRMSE) in both cases of in-sample and out-of-sample forecasting.

The estimation results show that EGARCH-type models fit better the data meaning that the volatility in the Moroccan stock market has asymmetric responses with regard to good news and bad news. Indeed, the leverage effect estimates are negative which means that volatility on the Moroccan stock market is more sensitive to bad news than good news. Also, the distribution of the MASI's returns is found to be leptokurtic indicating that volatility is high in the Moroccan stock market. A result that is consistent with other findings of studies on emerging financial markets. Further conclusions are drawn from the estimation results as follows:

(1) The GARCH-type models and EGARCH-type models fit better the data when a Student-t distribution is assumed for residuals ;

(2) Volatility in the Moroccan stock market exhibits pronounced persistence considering the significant positive estimates for introduced realized volatility (RV) ;

(3) The lagged continuous sample path variation contains relatively more information for predicting the current volatility than the lagged discontinuous jump variation ;

(4) According to predictive power of the models, the GARCH-CJ are found to be better than GARCH and GARCH-RV-type models for forecasting future volatility. This result was found when performing both in-sample and out-of-sample forecasting.

These findings mean that it makes sense to split the realized volatility in the GARCH-RV model into a continuous sample path and discontinuous jumps variations to enhance the models explanatory and predictive power of daily volatility in financial practices such as financial derivatives pricing, capital asset pricing, and risk measures.

Despite the fact that the constructed GARCH-CJ-type models have shown better performance for predicting stock index volatility, it is still necessary to improve the accuracy of measuring and predicting volatility with further improvements for the GARCH-CJ model in our forthcoming research by introducing more significant exogenous variables that impact volatility.

References

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, pages 885–905.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, 89(4):701–720.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93.

Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of financial Econometrics*, 4(1):1–30.

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*, page nbp001.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Frijns, B., Lehnert, T., and Zwinkels, R. C. (2011). Modeling structural changes in the volatility process. *Journal of Empirical Finance*, 18(3):522–532.

Fuertes, A.-M., Izzeldin, M., and Kalotychou, E. (2009). On forecasting daily stock volatility: The role of intraday information and market conditions. *International Journal of Forecasting*, 25(2):259–281.

Huang, C., Gong, X., Chen, X., and Wen, F. (2013). Measuring and forecasting volatility in chinese stock market using har-cj-m model. In *Abstract and Applied Analysis*, volume 2013. Hindawi Publishing Corporation.

Koopman, S. J., Jungbacker, B., and Hol, E. (2005). Forecasting daily variability of the s&p 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 12(3):445–475.

Martens, M. (2002). Measuring and forecasting s&p 500 index-futures volatility using high-frequency data. *Journal of Futures Markets*, 22(6):497–518.

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