Int. Journal of Economics and Management 13 (1): 153-164 (2019)

IJEM
International Journal of Economics and Management
Journal homepage: http://www.ijem.upm.edu.my

Dynamic Average-Forecast Value-at-Risk by Using High Frequency IPC Mexican Index

CHIN WEN CHEONG[a*], LIU CHENGZHI[b], JING SHENGZHE[b] AND YE ZHIQING[b]

[a]*Department of Mathematics, Xiamen University, Malaysia*
[b]*School of Economics and Management, Xiamen University Malaysia*

ABSTRACT

This study investigates the dynamic volatility movements and market risk of the high-frequency Mexican IPC (Indice de Precios y Cotizaciones) index. Based on the heterogeneous market hypothesis framework, the high frequency 5-min interval data have been utilised to examine the return and volatility of IPC index. By using high-frequency realised volatility and bi-power volatility estimators in the heterogeneous autoregressive model, the IPC Mexican market was found to be in concordance with the investment structure suggested by the heterogeneous market hypothesis. Besides various volatility estimators, the heterogeneous autoregressive model was improved with the enhancement of autoregressive conditional heteroscedasticity effect to capture the volatility of the realised volatility. To obtain a better forecast, the combination forecasts were applied by using various averaging methods and the forecast evaluations were examined by using various forecast loss functions. Finally, the forecasted results were utilised in determining the Mexican IPC stock market risk via the value-at-risk based on normal and heavy-tailed distributions.

JEL Classification: G14, G17, C52, C58
Keywords: Heterogeneous market hypothesis; high frequency volatility; value at risk

Article history:
Received: 12 October 2018
Accepted: 28 April 2019

[*] Corresponding author: Email: wcchin@xmu.edu.my

INTRODUCTION

Over the last couple of decades, the efficient market hypothesis (Fama, 1998; Malkiel, 2003) in terms of market information, has been rigorously investigated by using the financial market data, which included closed-daily and high-frequency data. Ideally, this hypothesis suggested that all historical market information is revealed by the market price. In other words, none of the market participants can excel in financial investments by using the market timing strategy as well as optimal asset selection. Based on the traditional efficient market hypothesis (EMH), new proposed hypotheses were introduced to complement the EMH. These included heterogeneous market hypothesis (HMH) and adaptive market hypothesis (AMH). The relevant studies for HMH were commonly conducted by using high-frequency data that normally included 1-min or 5-min interval data. The accessibility of high-frequency trading data due to the data storage technology has encouraged the intensive high-frequency data analysis in various financial markets.

For this specific study, the Mexican Indice de Precios y Cotizaciones (IPC) index was selected, which acted as an important indicator to reflect the general and comprehensive performances of the Mexican Stock Exchange (BMV). In addition, as the largest stock exchange in Mexico and the fifth stock exchange in America, BMV plays an irreplaceable role in the American financial markets. To the authors' knowledge and a thorough search of high-frequency Mexican IPC literature yielded only limited related articles. These included the study by Buncic and Gisler (2016) that used the heterogeneous autoregressive high-frequency model to examine the forecasts of realised volatility between the U.S. equity market and other selected foreign equity markets, including the Mexican IPC index. In another study by Buncic and Gisler (2017), they investigated the importance of jumps and the leverage effect on the high-frequency realised volatility in seventeen international equity markets, including the Mexican stock market. Nava et al. (2016) selected the Mexican stock market as one of the four datasets to measure the influence of different timescales on the high-frequency dynamics by decomposing these financial time series into simple oscillations associated with distinct time scales. Rodríguez (2017) analysed the Mexican stock market by using a state-space model that combined long memory and level shifts by decomposing the underlying data process into a mixture model and long memory dynamics. It is worth to note that for all the aforementioned studies, except for Rodríguez (2017), the analysis was not mainly focused on the Mexican stock market. Given the importance of the Mexican stock market as part of Latin America and also the global emerging market, it is hoped that the analysis can be added to the literature of high-frequency analysis of Mexican stock markets. This study can definitely provide better understanding of the underlying volatility movements to academicians or practitioners who are involved in the portfolio analysis and risk management of Mexican stock market.

In the analysis, two high-frequency volatility estimators were used, namely the realised volatility (RV) and bipower variation volatility (BV), to re-examine the HMH in the Mexican stock market. By using the Heterogeneous Autoregressive Model (Corsi, 2009) with the enhancement of asymmetric ARCH feature, the Mexican IPC index was modelled and estimated by using the 5-min data. As compared to RV, the BV has the advantage of jump-robust property. For comparison purposes, the ARCH models with student's t distribution were also used in the study forecast evaluations. After evaluating the best forecast model for volatility, the performances for the individual and average combined forecasts were further examined, which will be used further to determine the market risk. Volatility usually connects with determining the market risk for investment decision. Understanding the statistical behaviour of volatility movement is beneficial to stock market forecasting and portfolio strategy underlying the Mexican IPC index. For the application in finance, the value-at-risk is determined based on the estimation results.

The remaining of this study is arranged as follows: Section 2 explains the formation of high-frequency RV and BV and also introduces all methodologies, including HAR(RV)-TARCH model, HAR(BV)-TARCH model, ARCH-t model, and TARCH-t model. Section 3 discusses the empirical study in terms of estimation, model diagnoses, forecast evaluation, the application in finance by using value-at-risk approach, and combination of different model forecasts to increase the accuracy of results. Finally, Section 4 summarises and concludes this research.

REVIEW OF LITERATURE

Over the years, new definitions have been recommended to complement the classical efficient market hypothesis (EMH). These include the adaptive market hypothesis (Lo, 2004; Lo, 2005) and the heterogeneous market hypothesis (HMH). With new nonlinear features, the heterogeneous market hypothesis (Muller et al., 1993; Dacorogna, 2001) has proposed non-homogeneous market participants who interpreted market information based on the trading preferences and opportunities. This hypothesis emphasised on high-frequency data analysis, which included the FOREX and stock markets in the empirical studies by Muller et al. (1993) and Dacorogna (1998). This hypothesis also suggested that the financial market consisted of market participants with various investment strategy durations. The results of combining various investment time horizons have generated the 'seemingly' like long-range dependence volatility (Chin and Lee, 2018; Chin et al., 2017) in some empirical financial market studies.

For HMH analysis, high-frequency data were required to examine the presence of heterogeneous markets. These data were collected every minute from the daily trading activities through information technology facilities. Earlier studies conducted by Andersen and Bollerslev (1998) and Baillie et al. (2007) had proven that forecasts associated with the high-frequency data indicated better forecast evaluation performances and this has intensively expanded the interest of research over high-frequency data. Among the pioneer studies of high-frequency data analysis, Andersen and Bollerslev (1998) estimated the latent volatility as realised volatility (RV) by accumulating the sum of return products within a day. Corsi (2009) on the other hand accumulated the realised volatility of daily, weekly, and monthly data in order to capture the heterogeneity in the FOREX market. However, RV had an estimation issue when facing the abrupt jumps in the financial market (Barndorff-Nielson & Shephard, 2004). To handle this issue, the bipower variation by using the product of two consecutive absolute returns was introduced by Barndorff-Nielsen and Shephard (2004), which helped to lead to a more consistent estimation.

Besides the solid supports from the financial hypothesis in the model specifications, the methodologies of the forecast were also important to obtain a good forecast in volatility. To improve the forecast accuracy, the combination of multiple forecasts for the forecasting period was often used to overcome the shortcoming of individual forecast models. Combining the multiple individual forecasts can somehow capture the robustness of each particular individual forecasting model and therefore reduce the forecast error. This idea was raised by Bates and Granger (1969) who proposed that combining different forecasts had substantially improved the accuracy of the forecast results. In a study by Newbold and Granger (1974) for 80 monthly and 26 quarterly time series, respectively, they concluded that most of the time the combination of forecasts was superior to the individual forecast models. In another empirical study by Armstrong (2001) found that the ex-ante errors for the equally weighted average forecast in 30 empirical comparisons were reduced with an average of about 12.5% and ranged from 3% to 24%. Other empirical studies by Stock and Watson and Aiolfi (2001) have also shown that the forecasting model that used forecast combination methods showed the best performance of forecast results.

The selected Mexican IPC index acts as the largest stock exchange in Mexico and the fifth stock exchange in America, BMV plays an irreplaceable role in the financial market. Recently, Horenstien and Snir (2017), Herrera et al. (2015), and Torre et al. (2016) conducted an empirical study in regard to the portfolio planning in this area; besides, Choudhry (1996) and Aggarwal et al. (1999) completed relative research that focused on the AR-GARCH models. However, practical studies about this topic are limited, especially for high-frequency data of the IPC index. Therefore, to explore the emerging financial market and find some meaningful stylised facts had inspired to research around high-frequency data by modelling and forecasting the time series of return and volatility, calculating the value-at-risk, and eventually combining multiple forecasts for the forecasting period to improve the accuracy of forecasted results. Besides the high-frequency data analysis, the investigation of co-movement of stock prices of Mexican's IPC and the ASEAN markets (Yeoh, et al., 2015; Jiang, et al., 2017; Chen, 2018) is also an interesting topic in order to examine the correlations between these markets. However this topic will not be included in this specify study.

RESEARCH METHODOLOGY

The high-frequency heterogeneous autoregressive (HAR) volatility model was based on the heterogeneous market hypothesis concepts. The HAR model consists of various volatility components corresponding to time horizons, such as volatility series lag of a day, a week, and a month, respectively, which are denoted by $\sigma_{t-1}^{2,day}$, $\sigma_{t-1}^{2,week}$ and $\sigma_{t-1}^{2,month}$ in this specific study. $\sigma_t^{2,day}$ is an integrated volatility, which is the integral of the instantaneous variance over the high-frequency one-day interval (5 min was taken as an interval in this study). The selection of data frequency (Andersen et al., 2011) has a direct impact on the market microstructure noise, which will cause bias to the estimation of realised volatility. Market microstructure noise is one of the potential drawbacks when using high- frequency data (Arwartani et al., 2009) and will provide additional variation that is not related to the latent volatility. The possible sources of these variations are bid-ask bounce effects and discreteness effect of prices recording. Some early studies of high-frequency realised volatility estimators have used 5-min frequency for exchange rate volatility (Andersen et al., 2001) and stock markets (Anderson et al., 2001b). Bandi and Russel (2006) on the other had suggested the optimal frequency to be 0.4 min to 13.8 min in their stock market analysis. Until now, there is still no single optimal frequency (Potter et al., 2008) that has been recommended in realised volatility estimations. In one study by Liu et al. (2015), they found that the 5-min realised volatility outperformed 400 different estimators on equities, equity indices, exchange rates, and interest rates. Therefore, it has been a common practice (Shin and Hwang, 2015) to implement a data frequency between 5 min to 30 min for the high frequency data analysis.

The selected volatility representations are realised volatility (RV), $\sigma_t^{2,day} = \sum_{i=1}^{N} r_{t,i}^{2}$, and bipower variance (BV), $\sigma_t^{2,day} = c\sum_{j=1}^{N-1} |r_{t,j}r_{t,j+1}|$, where, c is a positive constant. Based on Barndorff-Nielsen and Shephard (2002), without the presence of abrupt jumps, the RV is a consistent estimator for integrated volatility. Although RV's variance may be reduced by the high sampling frequency, at the same time, it may increase its possibility of estimation bias issue. Therefore, Barndorff-Nielsen and Shephard (2004) recommended a jump-robust bipower variation (BV) volatility estimator to overcome this estimation bias issue.

In this study, the HAR model was used with the improvement of asymmetric autoregressive conditional heteroskedastic (ARCH) impact. The specifications for HAR(RV)-TGARCH and HAR(BV)-TGARCH models are formulated as follows:

$$\ln(\sigma_{RV,t}^{2,day}) = \theta_{RV} + \theta_{RV,d}\ln(\sigma_{RV,t-1}^{2,day}) + \theta_{RV,w}\ln(\sigma_{RV,t-1}^{2,week}) + \theta_{RV,m}\ln(\sigma_{RV,t-1}^{2,month}) + \varepsilon_{RV,t}$$

$$\ln(\sigma_{BV,t}^{2,day}) = \theta_{BV} + \theta_{BV,d}\ln(\sigma R_{BV,t-1}^{2,day}) + \theta_{BV,w}\ln(\sigma_{BV,t-1}^{2,week}) + \theta_{BV,m}\ln(\sigma_{BV,t-1}^{2,month}) + \varepsilon_{BV,t}$$

where $\varepsilon_{.,t}$ follows a TGARCH model in the realised volatility (Corsi et al., 2008) and each of the HAR volatility components can be computed by using the equations $\sigma_t^{2,week} = \frac{1}{5}\sum_{t=1}^{5} ln(\sigma_t^{2,day})$ and $\sigma_t^{2,month} = \frac{1}{22}\sum_{t=1}^{22} ln(\sigma_t^{2,day})$. Because of the non-normality of financial time series, a student-t with the density function is used as follows:

$$f_{student-t}(a_n|v) = \frac{\Gamma\left(\frac{v+1}{2}\right)1}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi(v-2)}}(1+\frac{a_n^2}{v-2})^{-\frac{v+1}{2}}, v > 2$$

Where, Γ (·) is a gamma function with v as its degree of freedom. Through EVIEWS, the estimations were conducted by using the parametric maximum likelihood approach. To check the appropriateness of the models, the Ljung–Box portmanteau tests was used to examine the standardised residual for the return equation and squared standardised residuals for the variance equation with the null hypothesis of no serial correlation in both series. Models that passed the diagnostic check will be selected based on the information criteria, such as Akaike, Schwarz, and Hanna-Quinn information criteria. The one-day ahead forecasts were set to h out-of-sample forecast with h = 1, 2. . . 116. Three loss functions were selected in forecast evaluations, namely the root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

Average forecast methodology

In this study, several competitive forecasts were combined into a single forecast by forecast averaging methods to improve the forecast results. A few commonly used averaging weighting (ω) schemes were used, such as the simple-mean (SM), simple-median (SMed), and least squares (LS) approaches. The k-step-head forecasts will be rewritten as an individual model forecasts given by:

$$V_{t+k}^{Average} = \omega_1 V_{t+k}^1 + \omega_2 V_{t+k}^2 + \cdots + \omega_u V_{t+k}^u.$$

The first averaging method SM calculated the arithmetic mean of the forecasts at each observation in the forecast sample, while the second method SMed took the median of the forecasts at every observation in the forecast sample. The implicit (0, 1) weights were time-varying as each forecast method may be the median for some observations but not others. And finally, the least squares weighting method was calculated by regressing the forecasts against the actual values and then by using the coefficients from the regression as weights through the equation:

$$V_{t+k}^{Average} = \omega_0 + \omega_1 V_{t+k}^1 + \omega_2 V_{t+k}^2 + \cdots + \omega_u V_{t+k}^u + \epsilon_{t+k}$$

RESULT AND DISCUSSION

This study has selected the Mexican Stock Exchange, which was ranked the second largest in Latin American stocks. The IPC index indicated the BMV overall performance. It was made up of a balanced weighted selection of shares that were representative of all shares listed on the stock exchange from various sectors across the economy. The price index plot is illustrated in Figure 1. In this study, the in-sample data started from January 2010 and ended in December 2015 (1,479 days. Meanwhile for the forecast evaluations, the data from July 2015 until December 2015 (116 days) were reserved.

Table 1 Descriptive statistics

STATISTICS	log(RV)	log(BV)	RETURN
Mean	-4.335147	-4.456058	0.0001
Median	-4.393404	-4.498928	0.000154
Maximum	-2.268931	-2.630372	0.018098
Minimum	-5.158297	-5.225779	-0.025994
Standard Deviation	0.352961	0.300923	0.004071
Skewness	0.848931	0.712774	-0.328899
(t test)	(13.32848*)	(11.190778*)	(-5.16381904*)
Kurtosis	4.360555	4.168003	5.764403
(t test)	(34.231051*)	(32.719487*)	(45.251481*)
Jarque-Bera	291.7234*	209.3043*	497.5985*

Notes: * indicated significance at 5% level

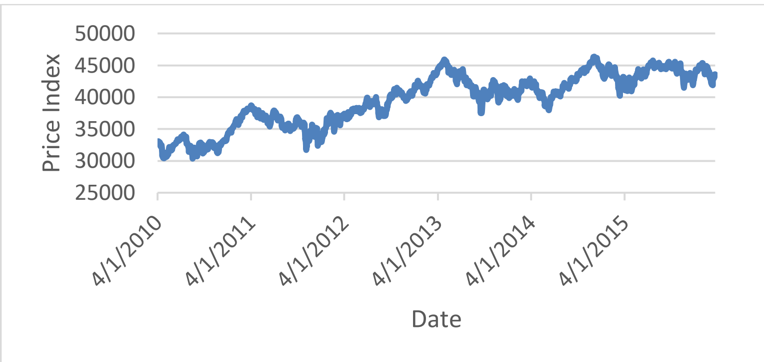


Figure 1 Price index of Mexican IPC

Table 1 shows the overall descriptive statistics for return series and two high-frequency volatility, namely the realised volatility and bi-power realised volatility. For the statistics of skewness and excess kurtosis, all log(RV), log(BV) and return were statistically against the normality based on the t-tests and Jarque-Bera normality tests. Therefore, the model should include the non-Gaussianity assumption.

Estimations

Table 2 The Estimations for HAR models

Estimation	HAR(RV)-TGARCH	HAR(BV)-TGARCH
θ_0	-2.191892*(0.322642)	-1.465731*(0.272320)
θ_d	0.166021*(0.032944)	0.198229*(0.032599)
θ_w	0.217005*(0.052594)	0.344321*(0.053454)
θ_m	0.414492*(0.052145)	0.323539*(0.049912)
a_0	0.063078*(0.027016)	0.033806*(0.014207)
a_1	0.088090*(0.033196)	0.098822*(0.036090)
β_1	0.796498*(0.080372)	0.809389*(0.070277)
δ	-0.083043*(0.042323)	-0.086109*(0.042538)
v	7.425573*(1.30771)	10.09088*(1.634002)
Model selection		
AIC	1.948531	1.459819
SIC	1.983460	1.494748
HIC	1.961617	1.472905
Diagnostic		
$\gamma_t, LB(12)$	20.953(0.451)	13.866(0.309)
$\gamma_t^2, LB(12)$	4.3794(0.976)	4.3747(0.976)

Notes: 1. The parentheses values represent standard errors and p-values for estimation and diagnostic checking respectively.
2. * denotes the 5% level of significance
3. γ_t represents the standardized residuals.

Table 2 reports on the maximum likelihood estimations with different time horizons of additive volatility cascade, namely the daily, weekly, and monthly volatilities under the student-t distributed error assumption for the two HAR models, HAR (RV)-TARCH and HAR (BV)-TARCH model, respectively.

In the model estimation for HAR models, every coefficient of heterogeneous components could be seen, namely the lags for daily component (θ_d), weekly component (θ_w) and monthly component (θ_m), all had sufficient evidences to show that they were statistically different from 0% to 5% level of significance. For the RV model, the lag of monthly volatility component gave the greatest impact to the recent daily volatility. This was followed by the lag of weekly volatility and the weakest by the lag of daily volatility components. On the other hand, the BV models showed an approximately similar impact for the lag of monthly and weekly volatility components to the recent volatility. These findings suggested that the impact of prior volatility of different time horizons were all statistically significant and concluded that it was in concordance with the concept of heterogeneous market hypothesis, whereby the Mexican stock markets consisted of heterogeneous market participants with different preferences in investment durations; In this context, the short-term (daily), middle-term (weekly), and long-term (monthly) investments. As a conclusion, the Mexican IPC index supported the presence of the heterogeneous market hypothesis where different trading horizons of market participants have different interpretations in regard to the inflow market information.

Table 3 The Maximum Likelihood Estimations for GARCH models

Estimation	GARCH-t	TGARCH-t
$\emptyset_0$	0.000253* (9.46E-05)	0.000122* (9.23E-05)
$\emptyset_1$	0.018656* (0.028703)	0.026525* (4.81E-08)
a_0	3.13E-07* (1.30E-07)	2.07E-07* (0.011262)
a_1	0.078473* (0.016059)	-0.047658* (0.019453)
β_1	0.903745* (0.019508)	0.958617* (0.008280)
δ		0.146144*(0.019453)
v		9.801465*(2.826014)
Model selection		
AIC	-8.350309	-8.382402
SIC	-8.327037	-8.355252
HIC	-8.341591	-8.372231

Table 3 Cont.

Diagnostic		
$\gamma_t, LBQ(12)$	9.2875(0.678)	8.2851 (0.762)
$\gamma_t^2, LBQ(12)$	15.609 (0.210)	12.994 (0.369)

Notes:
1. The parentheses values represent standard errors and p-values for estimation and diagnostic checking respectively.
2. * denotes the 5% level of significance
3. γ_t represents the standardized residuals.

As for the HAR-TGARCH component, there was sufficient evidence which showed that the coefficient of ARCH effect and GARCH effect were all statistically different from 0% to 5% level of significance for both BV and RV, respectively. In addition, the leverage effect (δ) was also statistically significant at 5% level of significance. This implied that the volatility of the realised volatility (Corsi, 2008) was a time-varying process. Therefore, it was necessary to include the ARCH effect in the HAR models. In the model diagnostic, under the null hypothesis of serially uncorrelated series, there was insufficient evidence for serial correlations for standardised and squared standardised residuals of both models.

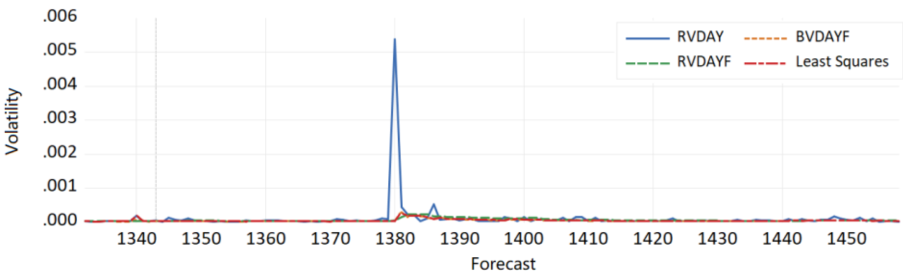
The TGARCH(1, 1)-t model indicated the smallest value, followed by GARCH, HAR-BV-TGARCH(1, 1)-t, and HAR-RV-TGARCH(1, 1)-t models. As compared to RV, the BV performed slightly better in the HAR modelling since the BV had the ability to capture the abrupt jump in the volatility. Concisely, in the estimation performance, the jump-robust realised volatilities (BV) outperformed the standard realised volatilities.

Forecast evaluations

Table 4 Dynamic Forecast Evaluations

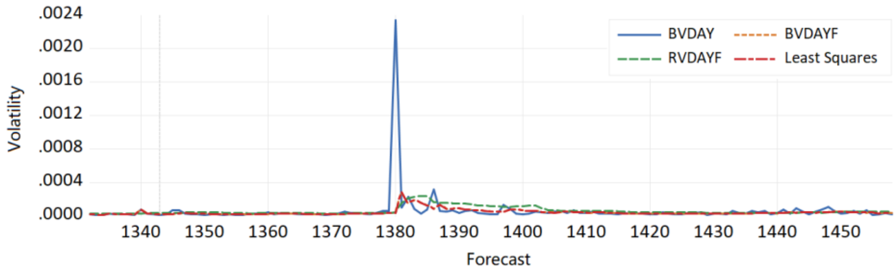
Actual: RV	Forecast evaluation		
Forecast method	MAE	RMSE	MAPE
Individual: HAR(RV)-TGARCH	0.00008369	0.00050023 *,**	72.57115488
HAR(BV)-TGARCH	0.00008068 *,**	0.00050119	51.06501412 *
GARCH-t	0.00009570	0.00050587	62.04323336
TARCH-t	0.00009761	0.00050619	66.30466374
Average: Simple mean	0.00008310	0.00050254	46.63066000
Simple median	0.00008439	0.00050305	46.40041266 **
Least-squares	0.00008369	0.00050023 **	72.57113360
Actual: BV	**Forecast evaluation**		
Forecast method	MAE	RMSE	MAPE
Individual: HAR(RV)-TGARCH	0.00005110	0.00021915	95.14545878
HAR(BV)-TGARCH	0.00004161 *	0.00021773 *,**	56.18827691
GARCH-t	0.00004781	0.00022120	52.51741209 *
TARCH-t	0.00004969	0.00022159	57.40503898
Average: Simple mean	0.00003892 **	0.00021801	42.05743195
Simple median	0.00003900	0.00021835	39.40507510 **
Least-squares	0.00004161	0.00021773 **	56.18828106

Notes: * indicates the smallest (best) value for individual forecast only. ** indicates the smallest (best) value for individual and average forecasts.



Note: The spike in the position of 1380 (date:25-Aug-2015) indicated the drastic change of price indices with the increase of approximate 2.5% (difference by 1000 points) from the previous day (24-Aug-2015).

Figure 2 Forecast Comparison using RV as the actual volatility



Note: The spike in the position of 1380 (date:25-Aug-2015) indicated the drastic change of price indices with the increase of approximate 2.5% (difference by 1000 points) from the previous day (24-Aug-2015).

Figure 3 Forecast Comparison using BV as the actual volatility

Table 4 reports on the dynamic forecast evaluations, which consisted of 116 days from July 2015 to December 2015 for MAE, RMSE, and MAPE for the four models. Figures 2 and 3 illustrate the forecast evaluations when RV and BV were used as the actual volatility values. By using the dynamic forecast approach, the estimated parameters would be used for the next one-day-ahead forecast. For individual forecast performance by using RV and BV as actual volatility series, it was found that in general, the individual HAR(BV)-TGARCH type models performed better than the GARCH-type models. The HAR(BV) performed better than HAR(RV) because RV series had a higher intensity of noisiness, whereas the BV series had the feature to smoothen or eliminate the possible jumps for the consecutive observations. It was also worth to note that except for MAPE evaluation under the BV as actual volatility, the GARCH-t model outperformed other models. This result indicated that the best-estimated models (HAR-BV) did not always guarantee the best forecast results. For the overall forecast evaluations among the individual forecasts and average forecasts, it was found that almost all smallest forecast loss functions were dominant by the average forecasts, such as least squares and simple median methods. These outcomes suggested that the average forecasts gathered all advantages of each individual forecast and provided the best forecast by averaging them. In other words, it was worthy to implement the average forecasts to obtain a more accurate forecast result.

The value-at-risk determination

Table 5 Value-at-Risk Determination based on actual RV (Dynamic forecast)

	HAR(RV) -TGARCH	HAR(BV) -TGARCH	TGARCH	Simple mean	Simple median	Least-squares
$\widehat{\sigma_t^2}(1)$	0.00003490	0.00002320	0.00001250	0.00002035	0.00001789	0.00003491
VaR						
5% quantile	-0.00956950	-0.00778527	-0.00569012	-0.00728579	-0.00682484	-0.00957034
5% VaR	-9569.4991	-7785.2712	-5690.1161	-7285.7887	-6824.8422	-9570.3370
VaR						
1% quantile	-0.01477851	-0.01203231	-0.00880755	-0.01126353	-0.01055407	-0.01477980
1% VaR	-14778.5077	-12032.3112	-8807.5509	-11263.5322	-10554.0659	-14779.7973

Note: Value-at-risk calculates with $1 million of capital

Table 6 Value-at-Risk Determination based on actual BV (Dynamic forecast)

	HAR(RV) -TGARCH	HAR(BV) -TGARCH	TGARCH	Simple mean	Simple median	Least-squares
$\widehat{\sigma_t^2}(1)$	0.00003490	0.00002320	0.00001250	0.00002035	0.00001789	0.00002324
VaR						
5% quantile	-0.00956950	-0.00778527	-0.00569012	-0.00728579	-0.00682484	-0.00779196
5% VaR	-9569.4991	-7785.2711	-5690.1161	-7285.7887	-6824.8422	-7791.9619
VaR						
1% quantile	-0.01477851	-0.01203231	-0.00880755	-0.01126353	-0.01055407	-0.01204261
1% VaR	-14778.5077	-12032.3112	-8807.5509	-11263.5322	-10554.0659	-12042.6092

Note: Value-at-risk calculates with $1 million of capital

For the application in finance, the market risk for the Mexican IPC was computed by using the value-at-risk approach. Three student-t models, namely the HAR(RV)-TGARCH, HAR(BV)-TGARCH, and TGARCH, were used for comparison. In this specific evaluation, the one-day-ahead forecast was only calculated and the student-t distributed return was obtained by the AR-TGARCH model. The degree of freedom for the conditional return was estimated as 9.801465 with 5% level of significance. For student-t models, the long position for IPC market α % quantile one-day-ahead VaR was defined as:

$$VaR(1) = Capital \times \left(\hat{r}(1) - \frac{t_{df}}{\sqrt{\frac{df}{df-2}}} \hat{\sigma}(1) \right)$$

Where, $Capital$ is the amount of money invested, df is the degree of freedom of a standardised student-t model, and $\hat{r}$ and $\hat{\sigma}$ are the forecasted return and volatility. The forecasted volatility was obtained from either GARCH-t, HAR(RV)-TGARCH or HAR(BV)-TGARCH models. Assuming that a capital of $1million were invested to hold a long financial position of the IPC stock market. Overall, the 5% and 1% VaRs are reported in Tables 5 and 6. Based on the dynamic forecast evaluations in Table 4, it was found that the individual HAR(BV)-TGARCH model had the best evaluation results among the three models. The 5% and 1% VaRs of one-day ahead HAR(BV) with student-t distributed return of 9.80147 degrees of freedom are:

$$HAR(BV) - VaR_{(1)}^{\alpha=0.05} = 1million \times \left(0.00009 - \frac{1.83311}{\sqrt{\frac{9.80147}{9.80147-2}}} \sqrt{0.0000232} \right) = -\$7785$$

$$HAR(BV) - VaR_{(1)}^{\alpha=0.01} = 1million \times \left(0.00009 - \frac{2.82144}{\sqrt{\frac{9.80147}{9.80147-2}}} \sqrt{0.0000232} \right) = -\$12032$$

The negative value indicated a loss, which is positioned at the left tail of the return distribution. In short, under a 5% level of significance, the potential loss for the next day was $7,785. Similarly, in the VaR under a 1% level of significance, the loss was $12,032. From Table 5, among the RV and BV models, it seems that the BV reported a lower VaR because it had the feature to handle the possible jumps in the volatility. Both the high-frequency models indicated higher VaR as compared to the daily T-GARCH model. This implied that the daily T-GARCH model had a tendency to underestimate the market risk and provide incorrect information to investors. Consequently, this will allow the investor to encounter a higher risk of investment.

According to the tables of critical values, it was also worth to mention that for low significance level, such as α=0.05, the 5% quantile for the student-t distribution was 1.635, which was slightly smaller than the standardised normal distribution with the value of 1.645. On the other hand, for higher significance level, say α=0.01, the 1% quantile for the student-t distribution is 2.517, larger than the 1% quantile normal distribution, which is 2.236. In short, the fat-tail effect by using a standardised student-t distribution will only provide large VaRs for higher significance level, such as 0.01 in this specific study.

Next, three combination forecasts based on individual models were used for computing the market risk for the Mexican IPC. Consider the least-squares combination method as an example. The 5% and 1% VaRs of one-day ahead least-squares average forecast with the student- t distribution return of 9.80147 degrees of freedom are:

$$(Least - squares) - VaR_{(1)}^{\alpha=0.05} = 1million \times \left(0.00009 - \frac{1.83311}{\sqrt{\frac{9.80147}{9.80147-2}}} \sqrt{0.00003491} \right) = -\$9570$$

$$(Least - squares) - VaR_{(1)}^{\alpha=0.01} = 1million \times \left(0.00009 - \frac{2.82144}{\sqrt{\frac{9.80147}{9.80147-2}}} \sqrt{0.00003491} \right) = -\$14780$$

In short, under the 5% and 1% levels of significance, the potential loss for the next day was $9,570 and $14,780, respectively. According to the dynamic forecast evaluations in Table 4, it was found that in most times the combination forecasts presented better forecast outcomes as compared to the individual models. Therefore, the forecasts that came from these methods may be more reliable for investors.

CONCLUSION

This study used a modified heterogeneous autoregressive model with various high-frequency realised volatility to re-examine the heterogeneous market hypothesis for the Mexican stock market. The empirical discoveries showed that the jump-robust volatility outperformed the standard realised volatility and ARCH-type volatility in the forecast evaluations. For better forecast outcomes, the combination forecasts from three models were used and the forecasts were utilised in determining value-at-risk. In conclusion, the study enhances the literature on market information efficiency analysis, especially in the empirical case study of high frequency heterogeneous market hypothesis. The empirical results offer an alternative way to forecast and determine market risk, particularly in the analysis of investment portfolio strategy and risk.

REFERENCES

Aiolfi, M & Timmermann, A 2004, Structural Breaks and the Performance of Forecast Combinations. [Online] Available at: http://citeseerx.ist.psu.edu/

Andersen, T, Bollerslev, T, Diebold, F & Ebens, H 2001a, ‘The distribution of realized stock return volatility’, *Journal of Financial Economics*, vol. 61, no. 1, pp. 43-76.

Andersen, T. Bollerslev, T, Diebold, F & Labys, P 2001b, ‘The distribution of realized exchange rate volatility’, *Journal of the American Statistical Association*, pp. 42-55.

Andersen, TG & Bollerslev, T 1998, ‘Answering the skeptics: Yes, standard volatility models do provide accurate forecasts’, *International Economic Review*, vol. 39, no. 4, pp. 885-905.

Andersen, TG, Bollerslev, T, Diebold, FX & Labys, P 2003, ‘Modeling and Forecasting Realized Volatility’, *Econometrica*, vol. 7, no. 2, pp. 529-626.

Armstrong, JS 2001, Combining forecasts. [Online], Available at: http://repository.upenn.edu/marketing_papers/34.

Arwartani, B, Corradi, V & Distaso, W 2009, ‘Assesing Market Microstructure Effects via Realized Volatility Measures with an Application to the Dow Jones Industrial Averages Stocks’, *Journal of Business & Economics Statistics*, vol. 27, pp. 251-265.

Baillie, RT, Han, Y, Myers, RJ & Song, J 2007, ‘Long memory models for daily and high frequency commodity futures returns’, *Journal of Futures Markets*, vol. 27, no. 7, pp. 643-668.

Bandi, FM & Russel, JR 2006, ‘Separating Microstructure Noice from Volatility’, *Journal of Financial Economics*, vol. 79, pp. 655-692.

Barndorff-Nielsen, OE & Shephard, N 2004, ‘Power and Bipower Variation with Stochastic Volatility and Jumps’, *Journal of Financial Econometrics*, vol. 2, no. 1, pp. 1-37.

Barndorff-Nielsen, OE & Shephard, N 2002, ‘Estimating Quadratic Variation Using Realised Volatility’, *Journal of Applied Econometrics*, vol. 17, no. 5, pp. 457-477.

Bates, JM & Granger, CWJ 1969, ‘The Combination of Forecasts’, *Operational Research Quarterly*, vol. 20, no. 4, pp. 451-468.

Bollerslev, T 1986, ‘Generalized Autoregressive Conditional Heteroskedasticity’, *Journal of Econometrics*, vol. 31, no. 3, pp. 307-327.

Buncic, D & Gisler, KIM 2016, ‘Global equity market volatility spillovers: A broader role for the United States’, *International Journal of Forecasting*, vol.32(4), pp.1317-1339.

Buncic, D & Gisler, KIM 2017, ‘The role of jumps and leverage in forecasting volatility in international equity markets’, *Journal of International Money and Finance*, vol. 79, pp. 1-19.

Chen, P 2018, ‘Understanding international stock market comovements: A comparison of developed and emerging markets’, *International Review of Economics & Finance*, vol. 56, pp. 451-464.

Chin, WC & Lee, MC 2018, ‘S&P500 volatility analysis using high-frequency multipower variation volatility proxies’, *Empirical Economics*, vol. 54, no. 3, pp. 1297-1318.

Chin, WC, Lee, MC & Tan, PP 2017, ‘Heterogenous market hypothesis evaluation using multipower variation volatility’, *Communication in Statistics- Simulation and Computation*, vol. 46, no. 8, pp. 6574-6587.

Choudhry, T 1996, ‘Stock market volatility and the crash of 1987: Evidence from six emerging markets’, *Journal of International Money and Finance*, vol. 15, no. 6, pp. 969-981.

Corsi, F 2009, ‘A simple approximate long memory model of realized volatility’, *Journal of Financial Economics*, vol. 7, no. 2, pp. 174-196.

Corsi, F, Mittnik, S, Pigorsch, C & Pigorsch, U 2008, ‘The Volatility of Realized Volatility’, *Econometric Reviews*, vol. 27, no. 1-3, pp. 46-78.

Dacorogna, M, Mller, U, Olsen, R & Pictet, O 2001, ‘Defining efficiency in heterogeneous markets’, *Quantitative Finance*, vol. 1, no. 2, pp. 198-201.

Dacorogna, MM, Müller, UA, Pictet, OV & Olsen, RB 1998, Modelling Short-Term Volatility with GARCH and Harch Models. [Online]

Daniel Buncic, KIMG 2016, ‘Global equity market volatility spillovers: A broader role for the United States’, *International Journal of Forecasting*, vol. 32, no.4, pp. 1317-1339.

Engle, RF 1982, ‘Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation’ *Econometrica*, vol. 50, no. 4, pp. 987-1007.

Fama, E 1998, ‘Market efficiency, Long-term returns, and behavioral finance’, *Journal of Financial Economics*, vol. 49, no. 3, pp. 283-306.

Glosten, LR, Jagannathan, R & Runkle, DE 1993, ‘On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks’, *The Journal of Finance*, vol. 48, no. 5, pp. 1779-1801.

Herrera, FL, Salgado, RJS & Akec, SC 2015, ‘Volatility dependence structure between the Mexican Stock Exchange and the World Capital Market’, *Investigación Económica*, vol. 74, no. 293, pp. 69-97.

Horenstein, AR & Snir, A 2017, ‘Portfolio choice in Mexico. Journal of Behavioral and Experimental Finance’, vol 16, pp. 1-13.

Jiang, YH, He, N, Joe, YM 2017, ‘Co-movement of ASEAN stock markets: New evidence from wavelet and VMD-based copula tests’, *Economic Modelling*, vol. 64, pp. 384-398.

Liu, L, Patton, AJ & Sheppard, K 2015, ‘Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes’, *Journal of Econometrics*, vol. 187, no. 1, pp. 293-311.

Lo, AW 2004, ‘The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective’, *The Journal of Portfolio Management*, vol. 30, no. 5, pp. 15-29.

Lo, AW 2005, ‘Reconciling Efficient Markets with Behavioral Finance: The Adaptive Markets Hypothesis’, *Journal of Investment Consulting*, vol. 7, no. 2, pp. 21-44.

Malkiel, BG 2003, ‘The Efficient Market Hypothesis and Its Critics’, *Journal of Economic Perspectives*, vol. 17, no. 1, pp. 59-82.

Muller, UA, Dacorogna MM, Dave, RD, Pictet, OV, Olsen, RB & Ward, JR, 1993, ‘Fractals and intrinsic time - a challenge to econometricians’, Luxembourg, *XXXIXth International AEA Conference on Real Time Econometrics*.

Nava, N, Matteo, TD & Aste, T 2016, ‘Time-dependent scaling patterns in high frequency financial data’, *The European Physical Journal Special Topics*, vol. 225, no. 10, p. 1997-2016.

Newbold, P & Granger, CWJ 1974, ‘Experience with forecasting univariate time series and the combination of forecasts’, *Journal of the Royal Statistical Society*, vol. 137, no. 2, pp. 131-165.

Potter, MD, Martens, M & Dijk, D 2008, ‘Predicting the Daily Covariance Matrix for S&P 100 Stocks Using Intraday Data - But Which Frequency to Use?’, *Econometric Reviews*, vol. 27, pp. 199-229.

Reena Aggarwal, CI & Leal, R 1999, ‘Volatility in emerging stock markets’, *The Journal of Financial and Quantitative Analysis*, vol, 34, no. 1, pp. 33-35.

Rodríguez, G, 2017, ‘Modeling Latin-American stock and Forex markets volatility: Empirical application of a model with random level shifts and genuine long memory’, *The North American Journal of Economics and Finance*, vol 42, pp. 393-420.

Shin, DW & Hwang, E 2015, ‘A Langrangian Multiplier Test for Market Microstructure Noice with Applications to Sampling Interval Determination for Realized Volatility’, *Economics Letters*, vol. 129, pp. 95-99.

Stock, JH & Watson, MW 2001, ‘Vector Autoregressions’, *Journal of Economic Perspectives*, vol. 15, no. 4, pp. 101-115.

Timmermann, A 2006, Chapter 4 Forecast Combinations. Handbook of Economic Forecasting, vol. 1, pp. 135-196.

Torre, ODl, Galeana, E & Aguilasocho, D 2016, 'The use of the sustainable investment against the broad market one. A first test in the Mexican stock market', *European Research on Management and Business Economics*, vol. 22, no. 3, pp. 117-123.

Yeoh, BK, Zainudin, A & Hooy, CW 2015, 'Modeling Time-Varying Stock Market Integration in ASEAN: A Kalman Filter Approach', *International Journal of Economics and Management*, vol. 9, no. 2, pp.232-247.

ACKNOWLEDGMENT

This work was supported by Xiamen University Malaysia Research Fund (XMUMRF/2019-C3/IMAT/0006).