Essays in Financial Econometrics: Conditional Volatility, Realized Volatility and Volatility Spillovers

BURAK KORKUSUZ

Thesis submitted for the degree of Doctor of Philosophy in Finance

Accounting and Finance Division
Stirling Management School
University of Stirling
United Kingdom
January 2023

ABSTRACT

The accurate forecast of stock market volatility is of particular importance for policy makers, investors, and market participants who have certain levels of risk which they can bear. This thesis centres around the conditional volatility, realized volatility, and volatility spillovers in the context of their model extensions. In particular, we examine the behaviour of stock market volatility in a selection of international markets, the ability of different models to provide accurate volatility forecasts, and the nature of the interrelations between markets from the perspective of complex network theory. Focussing on the modelling and forecasting of volatility we compare some well-established conditional volatility models with realized volatility models and further investigate the use of a number of additional parameters in improving the forecast accuracy of the future realized volatility. In this regard, a wide range of additional parameters, from assets to commodities, extreme range estimators to overnight volatility, oil price to gold price, VIX to EPU, bond price to interest rate, are included. Moreover, those are classified as different information channels, namely local, regional, and global. In terms of volatility spillovers, a volatility spillover model is combined with complex network theory in order to construct a volatility network of international financial markets, consisting of nodes and edges. The main contributions of this thesis are four. First, using the thirty different stock market indices and more up-to-date data the realized volatility (HAR-RV) models outperform the conditional volatility (GARCHs) models and, moreover, decomposition of realized volatility into positive and negative realized semi-variances (HAR-PS) improve the forecast accuracy of HAR-RV model. Second, extreme range estimators such as Parkinson and Garman-Klass could contain additional information for forecasting the future realized volatility. Third, the role of global information at improving the forecasts of future realized volatility is more important than that of local and regional information. Lastly, the spillover networks of international financial markets are much denser in crisis periods compared to non-crisis periods and volatility spillovers in COVID-19 Crisis (2020) period are more transitive and intense than Global Financial Crisis (2008) period.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisors Professor David G. McMillan and Dr. Dimos Kambouroudis for their invaluable guidance, support, and expert knowledge throughout the research process. Their mentorship and insightful feedback have been instrumental in shaping the direction and quality of this research work.

I also thank to the examining board members, Professor Jing Chen (external examiner) and Dr. Dionysia Dionysiou (internal examiners) for their valuable time, expertise, and constructive feedback. Their insightful comments and suggestions have strengthened the depth of this thesis.

I would like to acknowledge the support and encouragement from my mother and father, who have been a constant source of love and motivation throughout my academic journey. I am also sincerely grateful to all my friends and colleagues who have provided support, valuable discussions, and encouragement throughout my PhD journey.

Last but not least, I am thankful to the Republic of Türkiye Ministry of National Education for providing me with the necessary resources, infrastructure, and academic environment to pursue my doctoral studies.

List of Contents

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Figures

Chapter 2

Chapter 4

Chapter 5

Glossary

Terms	Definitions
Volatility Clustering	This characteristic of volatility was first stated by Mandelbrot (1963, pp. 418) who explains that "Large changes tend to be followed by large changes…and small changes tend to be followed by small changes". The series of returns are uncorrelated, yet the squared returns display a positive and significant slowly decaying autocorrelation function that accounts for the clustering behaviour.
Leptokurtosis	The distribution of the returns generally exhibits excess kurtosis, meaning that there is more weight in the tails compared to the lognormal distribution. This distribution is modelled as independent and identically distributed process (i.i.d.).
Leverage Effects	The so-called leverage effect is described as the asymmetric response of volatility to positive and negative returns of the same magnitude. More precisely, bad news has a more significant effect than good news.
Long Memory	Stock returns display a weak autocorrelation. According to Ding et al. (1993), they are not independently and identically distributed. The autocorrelation of absolute and squared returns slowly decays that means a sign of long memory. This phenomenon is accounted for by questioning 'how quickly an asset forgets a large shocks in financial volatility'.
Jump	Financial data sometimes exhibit extraordinary variations due to market crash, political turmoil, speculative activity or natural disaster. In practice, jumps are not predictable. However, this stylised fact has a strong positive impact on the future volatility of stock markets.
Realized Volatility	The Realized Variance/Volatility (RV) is one of the most popular volatility measures that can be computed by taking the sum of squared intraday returns. It is important to note that it is tick-by-tick data and 5-minute RV is considered to be a better representative of the true volatility.
Conditional Volatility	Conditional volatility can be defined simply as; "volatility is conditioned on lagged values of itself and of model errors".
Volatility Spillovers	The transmission of information across financial markets is called in the literature as 'volatility spillover effects'. This phenomenon is also referred to as fear connectedness by Diebold and Yilmaz (2014), who are two of the most well-known researchers in this topic.
ARCH/GARCH	According to Engle (1982), if volatility can be correlated over time, then the change in variance can be modelled with the ARCH (Autoregressive Conditional Heteroskedasticity) model. Afterwards, Bollerslev (1986) made a generalization, which does not only catch 'volatility clustering', but also include 'fat tails' in the new model that is the Generalized ARCH (or GARCH) model. There are many different variants of GARCH models for different needs of practitioners and researchers (see –'Glossary to ARCH (GARCH)' (Bollerslev, 2009)– for more information).
HAR-RV	The Heterogenous Autoregressive Model of Realized Variance (Corsi, 2009) is a simple autoregressive type model for modelling and forecasting realized variance. The model consists of three components which are daily, weekly and monthly parameters.

ARFIMA	The long memory autoregressive fractionally integrated moving average (ARFIMA) model is a long memory type model that can be used for modelling and forecasting realized variance. It is an alternative to the HAR-RV model and had been often applied before the HAR-RV was proposed in 2009.
Efficient Market Hypothesis	The efficient market hypothesis (EMH) states that current stock prices reflect all available information and thus impossible to beat the market. In other words, everything in the market is already accurately and fairly priced and no room to make excess returns for investors and market participants, except by chance. There are different forms of this hypothesis such as the weak form, the semi-strong and the strong form. The weak form aims to test whether current stock prices reflect all available information. The semi-strong form is based on event studies, in other words, the announcement effects. The strong form analyses whether some specific groups (insiders) have private information from which to take advantage.
Heterogenous Market Hypothesis	This hypothesis assumes that the main reason of heterogeneity in the financial markets stems from the existence of various types of investors (Muller, Dacorogna, Dave, Olsen, Pictet and von Weizsacker, 1997). Different investors interpret the same information differently depending on their risk appetites.
Network Theory	Network theory is a concept that considers the relationships between different parts of real life complex systems by using nodes and edges.

CHAPTER 1

1. Introduction and Research Background

1.1. Introduction

The volatility forecasting of financial assets has been one of the most important topics in financial econometrics. The term, volatility, refers to the degree of fluctuations in asset prices within a short period of time. Poon and Granger (2003) consider volatility as a *"barometer for the vulnerability of financial markets and the economy"(p. 479)*. As financial volatility is closely related to risk and uncertainty, its sphere of influence is quite wide including, for example, risk management, option pricing, investment analysis, portfolio diversification, and policy-making. Therefore, the accurate forecast of stock market volatility is of significant importance for policy makers, investors, and market participants who have certain levels of risk which they intend not to exceed.

As financial markets pass through calm and crisis periods, volatility is not constant over time. In other words and modern terms, volatility is time-varying, meanwhile it is persistent (serially-correlated). Taking advantage of the persistence feature of volatility, volatility can be predictable to some extent with some econometrics techniques, namely the (Generalized) Autoregressive Conditional Heteroscedasticity ((G)ARCH) model (Engle, 1982; Bollerslev, 1986). Fundamentally, GARCH is a statistical modelling approach in which the volatility tomorrow is described in terms of the volatility today, and the observed returns. In financial econometrics, standard GARCH and its variants have become one of the most widely used time series models for modelling and forecasting volatility. Therefore, the repertoire of GARCH family models is quite wide since academics and practitioners attempt to explain different behaviours and characteristics of volatility (for more information see, Glossary to ARCH (GARCH), Bollerslev, 2009). For instance, EGARCH (Nelson, 1991), GJR-GARCH (Glosten, Jagannathan, and Runkle, 1993), and PGARCH (Ding, Engle, and Granger, 1993) models are suggested to capture the asymmetric behaviour of volatility, whilst GARCH-BEKK (Engle and Kroner, 1995) is one of the multivariate versions of standard GARCH model that can capture the transmission of volatility (volatility spillover effects) across markets.

GARCH type models are usually applied to the daily close-to-close returns, whereas Andersen, Bollerslev, Diebold, and Labys (ABDL, 2003) noted: *"It has become apparent that standard volatility models used for forecasting at the daily level cannot readily accommodate the information in intraday data" (p. 1)*. Since the early 2000s, with the increasing availability of high frequency data (thanks to the developments in data storage technologies) the use of intraday data in volatility research has drawn a lot of attention in academia. The seminal work of Andersen and Bollerslev (1998) points out that high frequency or in other words intraday data contain more information compared to the daily close-to-close data as the data of financial markets include thousands of transaction prices per day. Amongst others, Andersen and Bollerslev (1998), ABDL (2001, 2003), Koopman, Jungbacker and Hol (2005), Engle and Gallo (2006), Sheppard and Sheppard (2010), Celik and Ergin (2014) evidence that intraday data have the potential to better understand the dynamic properties of financial volatility and improve the accuracy of volatility estimation and forecasting. This is because intraday data in fact is the original form of stock market prices. Traders and market practitioners frequently observe intraday price movements when making important trading decisions. Looking at minute-wise data does essentially mean looking at the heart of the process of price formation in financial markets.

The true volatility is unobservable so that a proxy for the true volatility is needed. Daily squared or absolute returns are often employed as a proxy for the actual volatility in the daily return-based methods, especially within GARCH type models. However, the fact that daily squared or absolute returns are known as noisy estimators and thus poor to be substituted for the true volatility. Alternatively, Andersen and Bollerslev (1998) construct a volatility measure which derived from high frequency data. This measure is called as the realized variance that is computed by the sum of squared intraday returns. They demonstrate that the realized variance is a more precise measure of volatility compared to daily close-to-close based squared returns. This is because the realized variance is based on cumulative intraday returns in which the noisy component is shrunk. Preliminary studies (Hol and Koopman, 2002; Martens and Zein, 2004) show that the Autoregressive Fractionally Integrated Moving Average (ARFIMA) is the best-fitted model to the realized variance compared to other alternative models. Afterwards, Corsi (2009) proposes the Heterogenous Autoregressive (HAR) model for modelling the realized variance, which is a simple autoregressive type model but regressed over different time horizons. More recent studies (Barndorff-Nielsen, Kinnebrock and Sheppard, 2010; Corsi and Reno, 2012; Sevi, 2014; Patton and Sheppard, 2015) show that the HAR-RV is a much better model than the others in modelling and forecasting financial volatility. To sum up all the above-mentioned core literature in a few words, easy access to high frequency data since the early 2000s has evolved the studies of volatility modelling and forecasting from the daily-based methods to the intraday-based methods.

In modelling and forecasting realized volatility, the dominant modelling approach is the HAR-RV specification of Corsi (2009) and much of the progress in the aforementioned core literature is limited to the baseline HAR-RV model (without adding any exogenous variables to the baseline model). However, the role of various parameters in improving the forecasting performance of the HAR-RV model is quite important to obtain better volatility forecasts.[1] This is an obvious gap in the literature which this study aims to fill. In this regard, this thesis centres around the conditional volatility, realized volatility, and volatility spillovers in the context of their model extensions. In a nutshell, this thesis examines the behaviour of stock market volatility in a selection of international markets, the ability of extended models to provide accurate volatility forecasts, and the nature of the interrelations between markets but from the perspective of complex network theory.

Broadly speaking, the aims of this thesis are twofold. First, to determine the best-performing model for forecasting the future volatility. To do this, the first three of empirical works analyse the role of various key parameters in improving the in-sample fit and out-of-sample forecasting accuracy of the HAR-RV model. Second, to explore the importance and impact of spillover effects across financial markets. For this purpose the last empirical study does combine a spillover model (GARCH-BEKK) with complex network theory to investigate the volatility spillover network in international financial markets (consisting of nodes and edges).

Some further details of the above-mentioned four empirical exercises are briefly as follows. The first empirical exercise, Chapter 2, compares the conditional volatility (GARCH) models to the realized volatility (HAR-RV and ARFIMA-RV) models. The decomposition of realized volatility into positive and negative realized semivariances (HAR-PS) is also applied to improve the forecast accuracy of HAR-RV model. The second empirical study, Chapter 3, investigates the role of range-based estimators in improving the future realized volatility by

---

[1] These exogenous parameters could be such as some of the well-known global financial indicators (oil prices, interest rates, gold prices, implied volatility indices and economic policy uncertainty indices), cross market information (other stock markets) and other relevant stock markets data (overnight volatility and extreme range estimators).

extending HAR-RV model with some additional (X) variables, which are Parkinson, Garman-Klass, Rogers-Satchell, and Yang-Zhang. The third empirical exercise (Chapter 4) also extends HAR-RV model with various exogenous variables but classifying them according to the different kinds of information channels, namely the local, regional, and global information. With the use of an unusual methodological approach the last empirical research, Chapter 5, analyses the dynamic transmission mechanism of volatility spillovers between some key global financial indicators and the G20 stock markets under the five identified sub-periods. This method is a combination of a bivariate GARCH-BEKK model with complex network theory, which constructs a volatility network of international financial markets using nodes and edges. The sub-periods are determined according to the crisis and non-crisis periods including, for example, Global Financial Crisis (2008) period and COVID-19 (2020) Crisis period.

The main contributions are four. The first study finds the superiority of high frequency based models over daily based models that empirically contribute to the existing literature by conducting a comprehensive exercise with 30 different international stock markets and more up-to-date data. The second empirical work suggests that Parkinson and Garman-Klass estimators could be utilized to produce better forecasts of future realized volatility. Third, the channel of global information are better to improve the future realized volatility forecasts compared to local and regional channels. The final exercise methodologically and empirically contributes to the existing literature. Methodologically, a solution is provided to the difficulty encountered by the bivariate GARCH-BEKK model when dealing with multi-dimensionality issue (i.e. we[2] combine a bivariate GARCH-BEKK model with complex network theory in order to construct a network of international financial markets). Empirically, it is found that the networks of international financial markets are much denser in crisis periods compared to non-crisis periods and financial volatility spreads more rapidly and directly through key financial indicators to the G20 stock markets, especially in crisis periods.

The economic implications of the above-mentioned findings can be helpful in the process of risk and portfolio management. In terms of risk management, accurate volatility forecasting is quite important, especially for policy makers, investors, and market participants who have certain levels of risk which they can bear. This is because all financial actors desire to know today, "what will be the degree of volatility tomorrow?". Thus, the results of the first three of empirical exercises could contain useful information in improving the forecast accuracy of

[2] Throughout the thesis, I prefer to use plural "we" instead of "I" as the published papers from this thesis are co-authored jointly with my supervisory team Professor David McMillan and Dr Dimos Kambouroudis.

stock market volatility and determining the most and least relevant parameters of stock market volatility. For instance, some well-established extreme range estimators (e.g. Parkinson and Garman-Klass) and global information channels (global financial barometers and US market news) as additional variables can help in forecasting the one-day-ahead volatility of stock markets. The gain in forecasting accuracy is believed to be economically significant to minimize risk and maximize return. For example, whilst investors and market participants want to rearrange their stocks or portfolio positions before financial markets become too volatile, policy makers would desire to narrow bid-ask spread in order to restore market liquidity if the future is expected to be more volatile. From the perspective of portfolio management, investors and market participants could use our findings to align their portfolios by reducing their exposure to various risks (i.e. local, regional, and global risks in Chapter 4). To sum up, over the last two decades, the Global Financial Crisis (GFC, 2008) and the COVID-19 Crisis (CVC, 2020) have further highlighted the significance and impact of accurate and efficient volatility forecasting.

The findings of the last empirical research could also provide important information for investors and market participants who wish to diversify their portfolios in international financial markets. This is because considering volatility spillover relations in G20 stock markets is important in being able to manage risks and portfolio diversifications. For example, the existence of global risk factors can be thought as a sign to restrict the possibilities of portfolio diversification, especially during crisis periods when the correlations among investment instruments are high. In the non-crisis periods, more diversified portfolios can be constructed depending on time and market specific information.

1.2. Research Background

This section gives and discusses the key concepts used in this thesis in order to draw a picture of this research's background. The sub-section 1.2.1 gives fundamental information about the well-known conditional volatility models. Afterwards, the Heterogenous Market Hypothesis, the realized volatility models, and the proxies for financial volatility are explained in sub-sections 1.2.2, 1.2.3, and 1.2.4, respectively. These sections contain some essential information for the first three empirical chapters, namely Chapters 2, 3, and 4. The last two sub-sections (1.2.5 and 1.2.6) are related to Chapter 5 and give relevant details and descriptions.

1.2.1. Conditional Volatility (GARCH models)

As financial markets pass through calm and crisis periods, volatility is not constant over time. In other words and modern terms, volatility is time-varying, meanwhile persistent (serially-correlated) it is. Taking advantage of the persistence feature of volatility, volatility can be predictable to some extent with some econometrics techniques, namely the (Generalized) Autoregressive Conditional Heteroscedasticity ((G)ARCH) model (Engle, 1982; Bollerslev, 1986). According to Engle (1982), if volatility can be correlated over time, then the change in variance can be modelled with the ARCH model. In other words, the variance of the current error is conditionally a function of its past values. The ARCH model is in the form of an AR process for the error variance and thus slow to respond to the large shocks so that it is unable to successfully capture the volatility persistence. Afterwards, Bollerslev (1986) suggests a generalized form of the ARCH (known as the so-called GARCH model), assuming an ARMA process for the error variance. The GARCH process does not only catch 'volatility clustering', but also include 'fat tails' in the model equation. Fundamentally, GARCH is a statistical modelling approach in which the volatility tomorrow is described in terms of the volatility today, and the observed returns. However, the GARCH still has some drawbacks, for example, one of the most important one is that the standard GARCH model is symmetric. It does treat positive daily returns the same as negative daily returns. Therefore, many other asymmetric GARCH specifications have been developed in order to explain the asymmetry in financial volatility. For instance, EGARCH (Nelson, 1991), GJR-GARCH (Glosten, Jagannathan, and Runkle, 1993), and PGARCH (Ding, Engle, and Granger, 1993) models are suggested to capture the asymmetric effects of volatility from different perspectives. In financial econometrics, the repertoire of GARCH type models is quite wide since academics and practitioners attempt to explain not only asymmetry, but also different behaviours and characteristics of volatility such as leptokurtosis, long memory, and even volatility spillover effects with the multivariate version of the GARCH model (for more information see, Glossary to ARCH (GARCH), Bollerslev, 2009). For instance, the GARCH-BEKK (Engle and Kroner, 1995) is one of the multivariate versions of the GARCH that captures the transmission of volatility (i.e. volatility spillover effects) across markets. Indeed, standard GARCH and its variants, which are the most widely used time series models, have become the mainstream models for modelling and forecasting volatility.

GARCH family models employ the daily close-to-close returns in order to model volatility. However, Andersen, Bollerslev, Diebold, and Labys (ABDL, 2003) stated: *"It has become apparent that standard volatility models used for forecasting at the daily level cannot readily accommodate the information in intraday data" (p. 1).* Over the last twenty-years, easy access to high frequency data (thanks to the developments in data storage technologies) has opened a new chapter in the literature of volatility research. The seminal work of Andersen and Bollerslev (1998) points out that the data which are higher frequencies than daily-based contain more information as the financial market data include thousands of transaction prices per day. Amongst others, Andersen and Bollerslev (1998), ABDL (2001, 2003), Koopman, Jungbacker and Hol (2005), Engle and Gallo (2006), Sheppard and Sheppard (2010), Celik and Ergin (2014) evidence that intraday data have the potential to better understand the dynamic properties of financial volatility and improve the accuracy of volatility estimation and forecasting. All these researches in the literature result in new methodological approaches in the field of financial econometrics.

1.2.2. Heterogenous Market Hypothesis

When Engle and Lee (1993) suggest to decompose the conditional variance of stock returns into a permanent (long-run) and a transitory (short-run) components in their proposed model, the intuition behind this idea was similar to the Heterogenous Market Hypothesis but these components are modelled without addressing the investment horizons of investors to specific investors groups. However, Muller, Dacorogna, Dave, Olsen, Pictet and von Weizsacker (1997) focus specifically on the aspect of the time horizons of various investors groups in their study other than informed traders and herding behaviours of different speculators (which are some other aspects of the market heterogeneity).

The Heterogenous Market Hypothesis (HMH) is a new concept in the literature of Efficient Market Hypothesis (EMH) rather than being a widely established or universally recognised theory. The HMH in fact claims that market participants are heterogenous or in other words non-homogenous. According to Muller et al. (1997), the main reason of heterogeneity in financial markets stems from the existence of various types of market participants and their different levels of information, risk preferences, trading strategies and beliefs. From this point of view, the HMH attempts to challenge the EMH. This challenge starts with the easy access of high frequency time series data. When researches (Andersen and

Bollerslev, 1998; Martens, 2001; Andersen, Bollerslev, Diebold, and Labys, 2003) evidence a great gain in forecast accuracy thanks to the availability of tick-by-tick data, it is seen that the findings of high frequency data bring counter-evidence to the weak form of market efficiency. This is to say that when the price movements of stock markets are observed at the interval of higher frequency (intraday), the markets look different in comparison with a rational market where stock prices react rapidly to market news and events. The reason is that not all market agents are rational as Shiller (1989) discusses the smartness of investors by stating “most participants in the stock market are not “smart investors” (following the rational expectation model) but rather follow trends and fashions)”.

Different market participants interpret same information in different ways according to their risk preferences. Each investors have different objectives, for example, some investors are completely hedgers whilst some others are completely speculators. Therefore, it is unrealistic to expect that each one of investors react identically to the same news. Even if there are many approaches to categorize heterogenous agents, the most promising way is to consider the expectations of market participants in various time horizons. In practice, those can be grouped in terms of their investment horizons as follows: Investor (over week, month or year), day trader (four hour or daily), intra-day trader (15 min, 30 min or 1 hour), and scalper (1 min or 5 min). From the perspective of the HMH, broadly speaking, market participants could be divided into three different categories depending on their investment horizons such as short-term, middle-term, and long-term investors. Short term (intraday traders) investors usually do their trades only within a given trading day or overnight. Middle term investors carry out their trading activities in a weekly horizon, whilst long-termism is associated with monthly horizon. Short term investors assess the market at a higher frequency (intraday) level and have a shorter memory whereas middle term traders focus on the days up until a week. In a similar vein, long term investors do plan their trading activities based on monthly horizon. Each of these groups of market participants may have their own trading strategies consistent with their trading horizon and have a homogenous appearance within their own groups. In this way, the heterogeneity of investors could be accounted for providing a better understanding of how those different market participants react and perceive to the same news. Based on this idea, Corsi (2009) suggests to divide not only the markets but also the volatility of financial market into different components (as daily, weekly and monthly horizons) in a proposed model in which the next section explain it in more details.

1.2.3. Realized Volatility (HAR-RV models)

The actual volatility is latent and therefore a proxy for the true volatility is needed. For comparison purpose, daily squared or absolute returns are often employed as a proxy to examine the performance of daily return-based methods such as GARCH family models. Here, the issue is that daily squared or absolute returns are noisy estimators and thus poor to be substituted for the true volatility. Alternatively, Andersen and Bollerslev (1998) construct a volatility measure which derived from high frequency data. This measure is called as the realized variance (RV) that is computed by the sum of squared intraday returns. They demonstrate that RV is a more precise measure of volatility compared to daily close-to-close based squared returns. This is because RV is based on cumulative intraday returns in which the noisy component is shrunk. Preliminary studies (Hol and Koopman, 2002; Martens and Zein, 2004) show that the Autoregressive Fractionally Integrated Moving Average (ARFIMA) is the best-fitted model for the RV compared to other alternative models.

The intuition behind Corsi (2009)'s HAR model is that different investors interpret same information differently depending on their risk appetites as it is stated in the HMH section. Based on this argument of the HMH, Corsi (2009) proposes the Heterogenous Autoregressive (HAR) model for RV, which is a simple autoregressive type model but regressed over different time horizons. The HAR-RV model is based on capturing different reactions of different investors through a simple autoregressive process, which is in other words an additive cascade model of different volatility components. The HAR-RV model can simply be estimated by the ordinary least square (OLS) method. The model is a good alternative to the ARFIMA-RV model. The HAR-RV model can also capture long memory characteristics of volatility even though it is not in the class of long memory models. The components of the HAR-RV model separately reflect the information from short to long term trading activities of market participants. The intuition behind the model is that investors could be divided into three different classes depending on their investment horizons such as short-term, middle-term, and long-term investors. In this way, different types of investors are included in the model, enabling to better understand that how those different market participants react and perceive to the same news. Related literature is quite rich for the applications of HAR-RV model. A number of studies (Corsi, 2009; Barndorff-Nielsen, Kinnebrock and Sheppard, 2010; Corsi and Reno, 2012; Sevi, 2014; Patton and Sheppard, 2015) examine the performance of the HAR-RV model and compare with some other alternative models. The superiority of the HAR-RV model is documented over the other alternative models in modelling and forecasting RV. To sum up all the literature in a few words, easy access to high frequency data since the early 2000s have evolved the studies of volatility modelling and forecasting from the daily-based methods to the intraday-based methods and a number of new realized measures are introduced in order to substitute the actual volatility.

1.2.4. Volatility Proxies

Financial price changes are unpredictable that makes the actual volatility time-varying and therefore unobservable. For this reason, we need to employ a target (proxy) volatility measure for the true volatility. Another reason is that a proxy is required to evaluate the accuracy of model fit and forecasts. Therefore, the choice of proxy is critical as it summarizes the true market information.

The current view in the literature is that the use of squared or absolute close-to-close returns are traditional approaches, whereas the realized measures based on high frequency intraday sampling contain more information than the daily based volatility measures and therefore considered to be a better representative of the true volatility (Andersen and Bollerslev, 1998; Andersen et al., 1999). Another alternative might be considered as the intraday high-low range that is introduced against the aforementioned traditional close-to-close approaches. The motivation behind range-based measures is that the close-to-close method takes into account only closing prices of today and yesterday. If two consecutive closing prices are the same or quite similar, the close-to-close measure fails to detect intraday volatility. Therefore, close-to-close volatility is commonly applied by the investors who are concerned with long term investment and passive investing. However, the investors who consider intraday trading and information need more than closing prices. In this regard, first, the high-low range of prices and afterwards OHLC (open, high, low, close) prices have become important in the literature of volatility research and some well-documented range-based volatility estimators are introduced in this area of research (some of the seminal papers, amongst others, Parkinson, 1980; Garman-Klass, 1980; Rogers and Satchell, 1991; Yang and Zhang, 2000).

In terms of the volatility measures which are derived from high frequency data, the superiority of the realized variance as a good proxy for the true volatility is well-documented by some distinguished studies (Barndorff-Nielsen and Shephard, 2002b, and Andersen, Bollerslev, Diebold, and Labys, 2003). Therefore, the realized variance is most frequently-used volatility measure compared to other alternative measures. It is calculated as the sum of squared intraday returns and therefore the choice of frequency interval for intraday data does matter. This issue has been addressed by many scholars. Amongst others, Martens (2001) finds that when the frequency of intraday observations rises, more accurate daily volatility estimation is obtained. However, when the frequency is too high, the efficiency of high frequency data can be distorted due to microstructure noise effect. For this reason, Hol and Koopman (2002) suggest to select the frequency interval between 5 and 30 minutes. One of the most widely-accepted papers for this issue is written by Liu, Patton, and Sheppard (2015) that compares over 400 different realized measures and point out that it is difficult to significantly beat five-minute realized variance so that the general consensus among scholars and practitioners is to use the 5-min realized variance as a target volatility.

1.2.5. Volatility Spillovers (Multivariate GARCH models)

The transmission of fluctuations across markets is known as 'volatility spillover effects' (see, for example, Yu et al., 2015; Rejeb and Arfaoui, 2016; Mensi et al., 2018). This phenomenon is also referred to as fear connectedness by Diebold and Yilmaz (2014). In the last two decades, the Global Financial Crisis (GFC, 2008) and the Covid-19 Crisis (CVC, 2020) indicate the importance and impact of the transmission mechanism of spillover effects across financial markets. This is the fact that the spillover effects across markets is likely to have a profound impact on each economy. However, the transmission of fluctuations varies with the degree of market integration. In the case of such crises, increasing globalisation and financialization of markets allows adverse effects in one market to further intensify existing spillover effects. Consequently, throughout such crises, investors typically sell-off risky assets on fears of financial contagion that results in a further spread of global risk.

Multivariate GARCH models are capable to capture the transmission of fluctuations between markets. The GARCH-BEKK model of Engle and Kroner (1995) is one of the multivariate versions of the GARCH type models. This model is a bi-variate model that can capture volatility spillovers between pairwise markets. In considering multivariate-GARCH models there is a range of alternatives including the CCC (constant conditional correlation; Bollerslev, 1990), DCC (dynamic conditional correlation; Engle, 2002) and the GARCH-VECH (Bollerslev, Engle, and Wooldridge, 1988) models in addition to the GARCH-BEKK. A key advantage of the GARCH-BEKK model is that it does not impose any restriction on the conditional correlation structure between series. In addition, the conditional variances are restricted to ensure they are positive definite, while reducing parameter dimensions. However, it is suggested to use bi-variate form of the GARCH-BEKK model. This is because if a new variable is added into the model, the number of model parameters increases significantly. This issue is worse in terms of the full GARCH-VECH specification as the VECH inherently requires a larger number of parameters than the BEKK to be estimated and this can often lead to non-convergence. The full VECH-GARCH model of Bollerslev et al. (1988) is rarely used in the literature because the number of free parameters increases very fast with the number of variables. In the case of bi-variate VECH-GARCH specification, the model needs to generate 23 model parameters, which is a large number of free parameters for non-linear estimation. The restricted form of VECH or in other words diagonal VECH-GARCH model might be a better alternative to the unrestricted one but the restricted form of VECH does not generate cross-product (spillovers) parameters. This means that the direction of volatility transmission from one market to another cannot be extracted. Therefore, the VECH specification is rarely used and Bollerslev et al. (1988) did not estimate this model in their applications. Similarly, the CCC and DCC GARCH models also do not capture spillover effects from one market to another. Instead, both of the models do extract the magnitude of total spillovers between pairwise markets using econometrics techniques (e.g. see, method of Diebold and Yilmaz, (2009, 2012)). To sum up, except BEKK-GARCH model, the other multivariate GARCH specifications seem not to be suitable for detecting the direction of volatility spillovers.

1.2.6. Complex Network Theory

Complex network theory is a concept that takes into account the relationships among different parts of real complex systems as a network (Hao, An, Zhang, Li and Wei, 2015; An, Zhong, Chen, Li, and Gao, 2014). Network science is defined by the United States National Research Council (Network Science, 2005) as *"the study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena"(p. 28).* According to this definition, it is connected with other principal research disciplines. Network science does provide convenience to quantify and simplify the many parts of real world complex systems. The empirical results of complex systems can be seen in real world networks such as computer networks, technological networks, biological networks, social networks. This means that network science is a multidisciplinary field that helps us to understand the dynamics of real world complex systems by identifying key nodes and structures and uncovering patterns and properties of their interrelations.

A complex network consists of nodes which are connected by edges. The structure of any complex network graph can be represented by the help of an adjacency matrix. In network theory, an adjacency matrix is a square matrix that is employed in order to represent a finite network graph. The elements of an adjacency matrix show whether pairwise nodes are adjacent or not. If two nodes are connected to each other by a directional or bidirectional line, these two nodes are called adjacent. In the special case of a finite network graph, the diagonal elements of the adjacency matrix are zeros. In short, network theory enable us to visualize and analyse the relationships between different nodes in a system. To do this, there are a number of network statistics that indicate the different features of nodes and networks such as degree distribution, shortest path length, network diameter, network density, and clustering coefficient. Those can be defined shortly as follows. The degree distribution of a node is affected by the number of edges the node has and the size of those edges. Shortest path length is defined as the average of the shortest steps between pairwise nodes in a network. Afterwards, a short definition of network diameter is the shortest path between the two most farthest nodes of a network. Network density shows how the number of edges is close to the maximum possible edges in a network and if the network density is equal to unit, the network is called as a complete graph that includes all the possible edges between pairwise nodes. Lastly, clustering coefficient is a measure that shows how all of nodes are well-integrated in a network graph.

Complex networks could have some characteristic features, for example, the small world effect and the superposition phenomenon. The small world effect is first discovered by Watts and Strogatz (1998). It is a phenomenon in the network theory, assuming that no node is independent from the network. In other words, all of nodes are linked to each other either with a direct or undirect tie. Two most widely-used network statistics, which are the average shortest path length and average clustering coefficient are used to find out the small world effect. Another network characteristic present in our networks is the superposition phenomenon, which is basically a physic principle applying all the linear systems such as height in a water wave, intensity of a light wave or pressure in a sound wave. For instance, where two water waves travelling in opposite directions, the size of combined wave is the sum of the both water waves in the intersection point. Similarly, the thickness of an edge between pair nodes is identified by the superposition principle in this work.

1.3. Overview of Thesis

Motivated by the seminal paper of Andersen and Bollerslev (1998), **Chapter 2** mainly focuses on the modelling and forecasting of realized volatility. The one-step-ahead rolling and recursive window techniques are employed for generating the out-of-sample forecasts. With the use of a selection of forecast evaluation criteria and forecast comparison tests, we compare the forecasts obtained from the conditional volatility models (most-widely used GARCHs) to the forecasts of RV-based models (HAR-RV, RSV, and PS and ARFIMA-RV). The results are mostly in line with previous research that HAR-RV model outperforms the others and, moreover, decomposition of realized volatility into positive and negative realized semivariances (HAR-PS) improve the forecast accuracy of HAR-RV model. This chapter empirically contributes to the existing literature by conducting a comprehensive exercise with 30 different international stock markets and more up-to-date data.

Chapter 3 investigates whether extreme range estimators, which is derived from OHLC (open, high, low, close) prices, contain important information for forecasting the future realized volatility in the G7 stock markets. The same methodology is used with the previous chapter but extending the HAR-RV model with an exogenous (X) variable (hereafter HAR-RV-X). In this context, we analyse the additional information content of Parkinson, Garman-Klass, Roger-Satchell, and Yang-Zhang estimators to the future realized volatility. Although the results seem to be inconclusive in the stock markets of group of Seven, Parkinson and Garman-Klass estimators could be utilized to produce better forecasts of future realized volatility. To the best of our knowledge, this is the first study which examines the information content of extreme range information at improving the forecasts of realized volatility.

In **Chapter 4**, the HAR-RV-X model from the previous chapter is used and also further expanded by the Kitchen Sink (KS) strategy. The HAR-RV-KS uses a long list of possible exogenous variables in the model at once. For an inclusive examination, a wide range of exogenous variables from assets to commodities, implied volatility indices to bond rates are involved in this analysis. Moreover, those exogenous variables are classified according to different information channels, namely local, regional, and global. In doing so, we aim to investigate which class of models best helps in forecasting the future realized volatility. In conclusion, whilst the HAR-RV-KS outperforms the HAR-RV and HAR-RV-X (with only one X variable) specifications, the role of global information at improving the forecasts of future realized volatility is more important than the others.

Chapter 5 is relatively different from the previous chapters, considering multivariate GARCH approach with complex network theory to analyse the volatility spillover relations between key global financial indicators and G20 stock markets. Specifically, a bivariate GARCH-BEKK model that captures volatility spillover effects is combined with a complex network theory. Using this synthesis approach, we construct the spillover networks of international financial markets under five identified sample sub-periods including crisis and non-crisis periods. The findings contribute to the literature of volatility spillovers from the network theory perspective as follows. The volatility spillover relations between key global barometers (oil, gold, and bond) and G20 markets vary significantly across five identified sub-periods. Notably, networks are much denser in crisis periods compared to non-crisis periods. In comparing two crisis periods, Global Financial Crisis (2008) and COVID-19 Crisis (2020) periods, the network statistics suggest that volatility spillovers in the latter period are more transitive and intense than the former. This suggests that financial volatility spreads more rapidly and directly through key financial indicators to the G20 stock markets, especially in crisis periods.

Finally, **Chapter 6** summarizes the main findings of each empirical works, respectively. Afterwards, we provide some concluding remarks about obtaining higher forecast accuracy and propagation path of financial volatility in international markets. Lastly, some economic implications of those results are addressed for policy makers, investors, and market participants who have certain levels of risk which they can bear.

CHAPTER 2

A comprehensive exercise:
Forecasting realized volatility

A forecast comparison of conditional volatility models and realized volatility models

ABSTRACT

Easy access to high frequency data since the early 2000s have evolved the studies of volatility forecasting from inter-day based methods to intra-day based methods. Motivated by the seminal paper of Andersen and Bollerslev (1998), this forecasting exercise compares the inter-day (GARCH, EGARCH, PGARCH, and TGARCH) methods to the intra-day (HAR-RV and ARFIMA-RV) methods, including 30 different stock market indices between 2010-2019. One-day-ahead out-of-sample volatility forecasts are generated using both the rolling and recursive windows forecasting methods. The out-of-sample forecast losses are measured by the MSE, MAE, and QLIKE criteria. The conditional Giacomini-White pairwise test is employed to test the forecasting performance of the competing models. The results indicate that the models which employ the realized variance generate more accurate forecasts compared to the models with the conditional variance; the HAR-RV specifications outperform the others. Moreover, the decomposition of realized variance into positive and negative realized semivariances (HAR-RSV and HAR-PS) improves the out-of-sample forecasts. Following the HAR-RV models, the ARFIMA-RV specification is also found to be a superior model against the GARCH models as it uses a more precise measure of the true volatility. However, a few exceptional results are encountered in our sample that is in favour of the EGARCH model. This chapter empirically contributes to the existing literature by conducting a comprehensive exercise with 30 different international stock markets and more up-to-date data.

2.1. Introduction

The fluctuations of asset prices over a short period of time is called as volatility. As Poon and Granger (2003, p. 479) stated, it is a *"barometer for the vulnerability of financial markets and the economy"*. It is associated with risk and uncertainty, therefore, volatility forecasting is of critical importance to the future of the whole economy. Its sphere of influence is quite wide including, for example, portfolio management, option pricing, trading strategies, and monetary policy making.

Although a great number of research papers have been conducted on the predictability of volatility, there is still a lack of consensus in the literature on which one is the best forecasting model. In this context, the data of stock market, exchange rate, and crude oil are the most frequently examined assets. Those have been mostly studied using the GARCH family models since 1980s and GARCH models with daily data have become the mainstream class of models. Towards the turn of the century, the easy access to high frequency data has spurred the volatility forecasting activities in using intraday data. In addition to this, some pioneering researches have introduced realized measures to the literature by taking advantage of high frequency data. Those are followed by the newly developed volatility forecasting models. The Heterogenous Autoregressive model of the realized variance (HAR-RV) was first suggested by Corsi (2009). In recent years, this new model and its variants have become the new mainstream class of models and the recent literature of volatility research has been thriving on it.

Motivated by the seminal paper of Andersen and Bollerslev (1998), this forecasting exercise compares the inter-day (GARCH, EGARCH, PGARCH, and TGARCH) volatility models to the intra-day (HAR-RV, RSV, and PS and ARFIMA-RV) volatility models, including 30 different stock market indices between 2010-2019. In doing so, this chapter empirically contributes to the existing literature by conducting a comprehensive exercise with 30 different international stock markets and more up-to-date data. Majority of the indices consist of the stock markets of developed countries all over the world.[3] We know that volatility is unobservable and therefore we need to substitute a proxy for the true volatility. Andersen and Bollerslev (1998) points out that high frequency-based data contain more information compared to the daily-based data as the data of financial markets include thousands of transaction prices per day. Moreover, Liu, Patton, and Sheppard (2015) compare over 400

[3] Therefore, we are unable to conduct a comparison study between developed and developing countries' stock market indices due to data availability issues.

different realized measures and point out that it is difficult to significantly beat the 5-minute realized variance. For these reasons, we choose to employ the 5-min realized variance as a proxy in this study. One-day-ahead out-of-sample volatility forecasts are generated using both the rolling and recursive windows forecasting methods. The out-of-sample forecast losses are measured using the MSE, MAE, and QLIKE criteria. One of the main findings of this study indicates that the models which employ the realized variance generate more accurate forecasts compared to the models with the conditional variance, which is consistent with the seminal paper of Andersen and Bollerslev (1998).

More detailed findings indicate that the HAR-RV models outperform the other alternative models. Moreover, the decomposition of realized variance into positive and negative realized semivariances (HAR-RSV and HAR-PS) improves the out-of-sample forecasts. Following the HAR-RV models, the ARFIMA-RV specification is also found to be a superior model against the GARCH models as it uses a more precise measure of the true volatility. However, we encounter a few exceptional indices (6 out of 30) in which the EGARCH model outperforms the HAR-RV and ARFIMA-RV models. Furthermore, we use the conditional GW pairwise test in order to investigate whether the forecast errors of the competing models are statistically significant or not. Most of the test results are positive, which confirms the superiority of the HAR-RV specification. However, the test results of these exceptional indices (6 out of 30; e.g. DJI, FTSE, GSPTSE, MXX, OSEAX, SPX), which are in favour of the EGARCH model are insignificant. Therefore, we cannot exactly say that whether the winner model for those exceptional indices are actually superior. This is the fact that the results and findings of every work are specific to market, data frequency, time horizon, and some characteristics of volatility.

This chapter is organised as follows: ***Section 2*** presents the review of related literature. In ***Section 3***, the data and methods used in this study are explained in more detail. Afterwards, ***Sections 4*** gives the empirical results and their evaluations respectively. Finally, the conclusion is presented in ***Section 5***.

2.2. Literature Review

A great number of studies have been conducted to assess the forecasting performance of the GARCH family models and other alternative models. Indeed, the GARCH type models (with daily data) have become the mainstream models for the volatility forecasting. In the 2000s, the availability of high frequency data and the introduction of realized measures changed the direction of the research towards the studies of the realized volatility. Corsi (2009) proposed the Heterogeneous Autoregressive model of the realized variance (HAR-RV). This model and its new variants are now on their way to becoming a new mainstream class of models and the recent literature of volatility forecasting has been thriving on it.

This is the fact that the true volatility is latent and therefore researchers need to provide a proxy of it. There are different ways to do that such as using the data of squared returns and realized measures. The conditional variance of daily returns is captured by the GARCH family models. According to Engle (1982), if volatility can be correlated over time, then the change in variance can be modelled with the ARCH model. Afterwards, Bollerslev (1986) made a generalization, which does not only catch 'volatility clustering', but also include 'fat tails' in the new model that is the GARCH model. The superiority of the GARCH class of models was such that hundreds of other sophisticated GARCH models have been developed. Indeed, the GARCH models with daily data have become the mainstream models for the volatility forecasting.

In the 2000s, the easy access to high frequency data opened a new chapter in the research of financial volatility. A great number of studies indicate that models employing intraday/high frequency data clearly improve the accuracy of volatility forecasts (Andersen and Bollerslev, 1997; Andersen, Bollerslev, Diebold, and Labys, 2001; Martens and Zen, 2004; Koopman, Jungbacker, Hol, 2005; Chortareas, Jiang, Narkervis, 2011; Sevi, 2014). Initially, we need to answer the question 'why do high frequency data improve the forecast accuracy?'. There are a number of reasons. One of the main reasons is that considering the persistency property of volatility, high frequency data can provide a more accurate measure of current volatility because it contains more information for forecasting future volatility. In this sense, it also improves the evaluation of volatility forecasts this is because intraday data is less likely to induce volatility models' inconsistent rankings. Another important reason is that high frequency data enable us to understand the dynamic properties of financial volatility that is very important for modelling and forecasting (Hansen and Lunde, 2010).

Following the availability of high frequency data intraday data are employed to form various volatility measures which are more direct proxies for financial volatility. These volatility measures are called "realized measure or realized variance" in this literature. The superiority of the realized variance as a good proxy for the true volatility is well-documented by some distinguished studies (Barndorff-Nielsen and Shephard, 2002b, and Andersen, Bollerslev, Diebold, and Labys, 2003). Therefore, the realized variance is most frequently-used volatility measure compared to other alternative measures. It is calculated as the sum of squared intraday returns and therefore the choice of frequency interval for intraday data does matter. This issue has been addressed by many scholars. Amongst others, Martens (2001) finds that when the frequency of intraday observations increases, in turn, more accurate daily volatility estimation is obtained. However, when the frequency is too high (in other words the use of ultra-frequency data), the efficiency of high frequency data can be distorted due to microstructure noise effect. For this reason, Hol and Koopman (2002) suggest to select the frequency interval between 5 and 30 minutes. One of the most widely-accepted research articles for this issue is published by Liu, Patton, and Sheppard (2015) that compares over 400 different realized measures and point out that it is difficult to significantly beat five-minute realized variance so that the general consensus among scholars and practitioners is to use the 5-min realized variance as a target volatility. Therefore, this study considers the simple realized variance which is constructed based on the 5-minute intervals of squared returns.

The realized variance approach is introduced by Andersen and Bollerslev (1998). They indicate that the realized variance is a better measure of true volatility compared to daily squared returns because absolute and squared daily returns are noisy estimators for daily volatility. Andersen, Bollerslev, Diebold, and Labys (2003) do not consider the realized variance as only a proxy for the latent volatility but also consider as a variable to be directly modelled. Andersen et al., (2001, 2003) point out that the directly modelled intraday volatility can generate more accurate forecasts compared to the forecasts of GARCH family models. They also indicate that as sampling frequency becomes higher, the realized variance is an increasingly better measure of the true volatility. However, it is important to pay attention that ultra-high frequency data often result in microstructure noise which can make the estimates of some parameters very unstable. Similar evidence is found that high-frequency returns data provide more precise measure of true volatility than daily returns data (Blair, Poon, and Taylor, 2001). Across much research (e.g., Blair et al., 2001; Engle, 2002; Andersen, 2003; Koopman et al., 2005; Bollerslev 2009) the consensus view can be summarised as follows: the realized variance, which is derived from high frequency data contains more information about future volatility compared to the conditional variance based on daily close-to-close prices.

The studies of the realized variance have become one of the most investigated aspects of the volatility forecasting following the paper of Andersen and Bollerslev (1998). They are one of the first researchers, showing that the realized variance is a more accurate measure of volatility compared to the squared returns. Using the data of different financial assets many studies aim to find out the best performing volatility forecasting model. However, the literature has still to reach a consensus. Most of the papers concentrate mainly on the stock markets, yet in the context of single (or several) stocks or market indices. Even though stock market indices become one of the most investigated financial assets, there is still a gap in the literature in terms of the most recent developments in the research of high frequency data such as the introduction of new models and applications of those in the international markets.

There exists a large body of literature, comparing and evaluating the forecasting performance of various volatility models. In the 1990s, papers mainly focus on daily data. In this context, the literature is quite large and a great number of research papers assess the forecasting performance of the GARCH family model and other alternative models. It can be said that the GARCH type models (with daily data) have become a mainstream class of models for the volatility forecasting. Toward the turn of the century the availability of high frequency data has spurred the volatility forecasting activities by employing intraday data. In addition to this, extracting high frequency data research introduced a great numbers of realized measures instead of daily returns. The easy access to high frequency data and the introduction of realized measures (as newly introduced volatility proxies) changed the direction of the research and opened a new area of research. Hence, more recent studies concentrate more on the newly developed models with high frequency data (such as the HAR, MIDAS, HEAVY models) and also seek to extend and supplement this literature. Corsi (2009) proposed Heterogeneous Autoregressive model of the realized variance (HAR-RV). This model and its new variants are now on their way to becoming a new mainstream class of models and the recent literature of volatility forecasting has been thriving on it.

The Autoregressive Conditional Heteroskedasticity (ARCH) model (Engle, 1982) is the model which aims to capture the movements of volatility over time. Afterwards, Bollerslev (1986) proposed the generalised form of the ARCH model that is called the GARCH model. The GARCH model allows practitioners to capture more characteristics of data than the ARCH specification such as volatility clustering and fat tails. However, it still has some drawbacks, for example, one of the most important one is that the standard GARCH model is symmetric. It does treat positive daily returns the same as negative daily returns. Therefore, many other GARCH specifications have been developed in order to explain different volatility behaviours such as leverage effect. Some of extensions of the GARCH models are as follows: the exponential GARCH (EGARCH) model (Nelson, 1991), the threshold GARCH (TGARCH) model (Glosten, Jagannathan, and Runkle, 1993), Asymmetric Power ARCH (PGARCH) model (Ding, Engle, and Granger, 1993).

Pagan and Schwert (1990) conduct research, which is one of the first systematic comparisons of volatility models in the literature. It has been proven by many researchers that the GARCH-type models capture the characteristics of data better than other models. Akgiray (1989) applies a GARCH (1,1) model to monthly US data and find its superiority. Cumby et al. (1993) find that the EGARCH model is superior to the historical volatility models. Similarly, Hansen and Lunde (2005) carry out research comparing sixteen different GARCH specification and find that the GARCH (1,1) is the best performing one among others in datasets of the deutsche mark-dollar exchange rates. However, when performing the out-of-sample forecast for the IBM stocks, the models accommodating the leverage effect outperforms the standard GARCH model. On the other hand, Japanese and Singaporean stock indices are investigated and pointed out that an exponentially weighted moving average (EWMA) model yields better volatility forecasts compared to the ARCH-type models (Tse, 1991; Tse and Tung, 1992). Balaban, Bayar, and Faff (2004) evaluate various simple and GARCH models in 15 different stock markets. They find that the exponential smoothing model provides the best forecasts of volatility, whereas ARCH family models generates the worst forecasts.

Merton (1980) first uses high frequency data for measuring volatility and notes that the sum of squared returns can be used for computing the conditional variance at higher frequencies. The reason is that absolute and squared daily returns are the noisy estimators of volatility. In this context, Andersen and Bollerslev (1998) construct the realized variance as a more accurate true volatility measure in comparison to the daily squared returns. The main function of the realized variance was to be employed as an estimator of the true volatility for evaluating the forecasting performances of the volatility models. In the 2000s, the growing availability of high frequency data has drawn researchers' attention and the potential value of the realized variance used for improving volatility models. It is understood that high frequency data contain more information for the current and also future market volatility. Andersen et al. (2001, 2003) document that when the frequency of data rises, more accurate volatility forecasts are produced. Similar evidence finds that high frequency returns data contains more information than daily returns data (Blair, Poon, and Taylor, 2001; Engle, 2002; Andersen, 2003; Koopman et al., 2005; Bollerslev 2009). McMillan, Speight, and ap Gwilym (2000) compare and evaluate the UK stock market volatility forecasts of simple and GARCH models for daily, weekly, and monthly frequencies. They point out that moving average and GARCH models provide the best forecasts for all frequencies.

Corsi (2009) suggests the Heterogeneous Autoregressive model of the realized variance (HAR-RV) that is based on the Heterogenous Market Hypothesis. The HAR-RV model is in the form of a simple autoregressive model for the realized variance over different time spans, including daily, weekly, and monthly components. In other words, the HAR specification is an additive cascade model of different volatility components. The performance of the HAR-RV model is remarkably good in spite of its simple structure. According to Corsi (2009), the long memory feature of volatility is successfully modelled in a simple and parsimonious way although not belonging to the family of long memory models. It is not essentially a long memory model as the ARFIMA model, but can successfully deal with the persistence feature of volatility.

The superior performance of the HAR-RV is also supported by much research (e.g. Andersen et al. 2011; Patton and Sheppard, 2009; and Bollerslev et al. 2016). Barndorff-Nielsen, Kinnebrock, and Sheppard (2010) introduce positive and negative realized semivariance measures, which are obtained from the signed high frequency intraday returns. Afterwards, the HAR-RV model is expanded by decomposing volatility into jump and continuous components, negative and positive realized semivariances and also taking into account leverage effect (Sevi, 2014). It is also found that the decomposition of the realized variance improves the in-sample forecast performance, whilst failing in the out-of-sample. Patton and Sheppard (2015) find that negative realized semivariance for future volatility is much more important compared to positive realized semivariance. Disentangling the effects of negative and positive realized semivariances or good and bad volatilities successfully improves future volatility forecasts' performance. In this regard, in order to capture the asymmetric effect of the signed returns Patton and Sheppard (2015) suggest an asymmetric HAR model that include positive and negative realized semivariances. Particularly, Patton and Sheppard (2011) discuss the effects of positive and negative realized semivariances on forecasting power at different lags, namely daily, weekly and monthly signed realized semivariances. The decomposition of all the HAR model components (daily, weekly, and monthly) into positive and negative realized semivariances makes the role of the daily component diminished. Also, the increasing number of explanatory variables results in the overfitting issue. Therefore, Patton and Sheppard (2015) update the asymmetric HAR model by decomposing only the daily component, which contains the larger effect on the future volatility compared with the weekly and monthly components. Fang, Jiang, and Luo (2017) point out that the decomposition of the realized variance into the signed semivariances only for the daily component of the HAR model is more important than the model which decompose all the model components. In the case of the decomposition of all the model components, they find that the power of explanatory variables changes from negative semivariances to positive ones. Wu and Hou (2019) use the asymmetric HAR model and also suggest a combination of the asymmetric HAR with time-varying coefficients for the Chinese stock market. Other important realized semivariance related studies account for the good and bad volatilities in terms of the cross section of expected returns (Bollerslev, Li, and Zhao, 2018) and option pricing performance (Feunou and Okou, 2018).

Bucci (2017) conducts a systematic review on forecasting realized volatility that discusses the empirical foundations of different sorts of financial volatility. Forecast accuracy evaluation methods and the MIDAS, GARCH-MIDAS, Realized GARCH, and HEAVY models are analysed. Junior and Pereira (2011) examine MIDAS-RV and HAR-RV models using intraday data in Sao Paulo Stock Market. Whilst there is no difference between the models in the out-of-sample forecasts, the MIDAS-RV model outperforms in the in-sample. Celik and Ergin (2014) support that high frequency based volatility forecasting models outperform the traditional GARCH models. they suggest the HAR-RV-CJ and MIDAS models as the best performing high frequency based volatility models. It is also noted that the MIDAS model has superior performance in crisis periods. Using high frequency data of 100 stocks from different sectors Izzeldin, Kabir Hassan, Pappas, and Tsionas (2019) compare the forecasting performance of HAR and ARFIMA models. They point out that higher frequencies improve the forecasting performance of the models and an outperforming model depends on the frequency of data.

More recent studies examine the predictive ability of the HAR model for different markets and different settings, for example, Izzeldin, Muradoglu, Pappas, and Sivaprasad (2021) investigate the impact of the pandemic on the G7 stock markets at the sectoral level using the Smooth Transition HAR model that allows a switch between different volatility regimes. They find a strong transition evidence to the Covid-19 crisis regime in all stock markets and sectors but this varies according to sectors. Similarly, Ding, Kambouroudis, and McMillan (2023) incorporate the simple Autoregressive (AR) and HAR models with the Smooth Transition and Markov regime-switching approaches, which are non-linear regime switching models, to evaluate the forecasting accuracy of future realized volatility. It is pointed out in their results that the HAR model incorporates regime-switching approaches in particular the Markov regime-switching method with time-varying transition function yields the most accurate forecasting errors. Slim, Tabche, Koubaa, Osman and Karathanasopoulos (2023) focus on the forecasting realized volatility of Bitcoin to investigate the informative role of price duration. They note in their volatility forecasting exercise that price duration yields significant gains when it is applied to the price of Bitcoin.

Caporin (2022) investigates the role of jumps and signed jumps in the context of forecasting future realized volatility using 5-min return data for over 4800 Russell-3000 stock prices between the period of 2003 and 2019. Their findings evidence that the inclusion of jumps and signed jumps mostly does not help to improve the out-of-sample accuracy of the volatility models of his extensive choice, whilst improving the in-sample accuracy of the competing models. One of the most relevant contributions of his work to the literature is that the trade-off between forecasting accuracies and model complexity gives an advantage to the use of the baseline HAR model, which is proposed by Corsi (2009). A similar study (Bu, Hizmeri, Izzeldin, Murphy, and Tsionas (2023)) to Caporin (2022) but contrary to his out-of-sample forecast results that finds statistically significant gains in the in-sample and out-of-sample examines the respective contributions of the various kinds of jump components to total quadratic variation in an ultra-high frequency settings (i.e. using second-wise data). However, they point out the limited contribution of signed jumps in their study using the S&P 500 ETF (SPY) and twenty individual stocks in this index. The contrary views of the abovementioned studies could stem from different settings of the studies such as different frequencies and different markets. For instance, while Caporin (2022) uses 5-min return data, Bu et al. (2023) employ ultra-high frequency data.

To sum up, there is an ongoing debate in the literature on which model best suits the patterns of data in order to understand the strengths and weaknesses of different volatility models. Knowing which volatility model is better is important for risk management strategies, investment decisions, portfolio optimization, and overall investment performance as more accurate models enable us to carry out better estimation of future volatility and more accurate evaluations of market conditions. However, the results and findings are specific to market, data frequency, time horizon, and some characteristics of volatility. This is one of the most important reasons why each contribution in the literature is essential. Motivated by the seminal paper of Andersen and Bollerslev (1998), this study empirically contributes to the existing literature by conducting a comprehensive exercise with 30 different international stock markets and more up-to-date data.

2.3. Data and Methodology

2.3.1. Data Description

The data used in this study is provided by the Oxford-Man Institute of Quantitative Finance Realized Library. 5-min realized (semi) variance and daily returns series are employed for the volatility prediction of 30 different stock market indices.[4] Majority of them consist of the indices of developed countries. We attempt to conduct a comprehensive study by using a great numbers of stock market indices. The full list of index names and abbreviations is given at Table 1.

2.Table 1 The full list of index names and abbreviations

Symbol	Name	Symbol	Name
AEX	Amsterdam Exchange Index	***KS11***	Korea Composite Stock Price Index
AORD	All Ordinaries	***KSE***	Karachi SE 100 Index
BFX	Belgium 20 Index	***MXX***	IPC Mexico Index
BSESN	S&P BSE Sensex	***N225***	Nikkei 225 Index
BVLG	PSI All-Share Index	***NSEI***	NIFTY 50 Index
BVSP	BVSP BOVESPA Index	***OMXC20***	OMX Copenhagen 20 Index
DJI	Dow Jones Industrial Average	***OMXHPI***	OMX Helsinki All Share Index
FCHI	CAC 40 Index	***OMXSPI***	OMX Stockholm All Share Index
FTMIB	FTSE MIB Index	***OSEAX***	Oslo Exchange All Share Index
FTSE	FTSE 100 Index	***RUT***	Russel 2000 Index
GDAXI	DAX Index	***SMSI***	Madrid General Index
GSPTSE	S&P/TSX Composite Index	***SPX***	S&P 500 Index
HSI	HANG SENG Index	***SSEC***	Shanghai Composite Index
IBEX	IBEX 35 Index	***SSMI***	Swiss Stock Market Index
IXIC	Nasdaq 100 Index	***STOXX50E***	EURO STOXX 50 Index

The dataset of this forecasting exercise is the post-2007/2008 global financial crisis period. Every single index consists in the period of 9 years, specifically from January 4, 2010 to October 3, 2019. The number of observations in each index is approximately 2400 trading days. However, total trading days in a year can differ between each countries due to different

[4] The data used in this study cover all the available stock market indices in the Oxford-Man Institute's Quantitative Finance Realized Library. The series are a selection of daily non-parametric estimated volatility.

public holidays and nontrading periods. In this forecasting exercise, the initial sample comprises approximately one year period (330 obs. [2010-2011]), whilst the time interval of out-of-sample volatility forecasts is 8 years (2070 obs. [2011-2019]). We arbitrarily choose the in-sample length as 330 observations considering the length at least one year period to let the regression fit normally and obtain a longer out-of-sample period. This is because the main objective of this work is to evaluate the out-of-sample performance of the models. Lastly, as we discussed the importance of daily and intra-daily information in the section of literature review, this study focuses on daily volatility and generate only one-day ahead out-of-sample forecasts.

2.Table 2 ***Descriptive statistics of realized variance and returns series***

	REALIZED VARIANCE				**RETURNS**				
INDICES	**MEAN**	**ST. DEV.**	**SKEW.**	**EX.KUR.**	**MEAN**	**ST. DEV.**	**SKEW.**	**EX.KUR.**	
AEX	7.23E-05	0.00011	9.1184	135.49	0.00020	0.0104	-0.2617	3.2510	
AORD	4.00E-05	5.37E-05	7.8535	107.25	0.00012	0.0081	-0.3858	1.7869	
BFX	6.66E-05	8.51E-05	8.1125	109.33	0.00013	0.0104	-0.1465	4.1918	
BSESN	6.15E-05	6.24E-05	5.6385	60.804	0.00032	0.0009	-0.1043	1.8976	
BVLG	4.32E-05	4.31E-05	4.4342	33.381	0.00020	0.0103	-0.5226	2.8381	
BVSP	0.00010	0.00010	9.0808	147.76	0.00015	0.0142	-0.1384	1.8879	
DJI	6.85E-05	0.00016	20.937	692.56	0.00037	0.0088	-0.4722	3.9001	
FCHI	9.68E-05	0.00013	6.7683	73.601	0.00012	0.0121	-0.2349	3.9195	
FTMIB	0.00011	0.00015	5.7298	52.499	-4.1E-05	0.0153	-0.3330	4.4745	
FTSE	8.01E-05	0.00016	20.338	656.04	0.00010	0.0092	-0.2287	2.3859	
GDAXI	9.79E-05	0.00013	6.8276	74.264	0.00027	0.0119	-0.2944	2.6917	
GSPTSE	4.47E-05	9.43E-05	18.668	558.51	0.00013	0.0075	-0.3714	2.6558	
HSI	5.82E-05	6.90E-05	8.7771	115.54	7.50E-05	0.0114	-0.2860	2.2446	
IBEX	0.00014	0.00021	11.184	211.10	-0.00012	0.0137	-0.2078	7.5089	
IXIC	6.06E-05	0.00010	10.248	167.65	0.00050	0.0107	-0.4942	3.2251	
KS11	4.48E-05	7.76E-05	14.604	331.30	7.62E-05	0.0093	-0.4493	4.0942	
KSE	5.17E-05	5.62E-05	4.9986	46.235	0.00051	0.0097	-0.4072	2.6439	
MXX	6.33E-05	0.00013	24.929	896.13	0.00010	0.0090	-0.4416	3.4626	
N225	7.19E-05	0.00014	10.093	146.89	0.00029	0.0132	-0.5673	5.4127	
NSEI	6.07E-05	6.82E-05	7.3457	99.860	0.00032	0.0098	-0.1345	1.8773	
OMXC20	8.74E-05	0.00031	29.286	1039.0	0.00043	0.0109	-0.2474	2.7520	
OMXHPI	7.33E-05	0.00046	46.596	2253.9	0.00012	0.0113	-0.3033	3.6854	
OMXSPI	6.31E-05	0.00024	34.113	1410.1	0.00027	0.0110	-0.3570	8.3484	
OSEAX	8.74E-05	0.00015	8.8676	129.06	0.00032	0.0108	-0.2398	2.7012	
RUT	6.62E-05	0.00011	10.280	166.54	0.00034	0.0123	-0.3242	3.9481	
SMSI	0.00013	0.00024	14.539	305.67	-0.00014	0.0136	-0.2924	7.4696	
SPX	6.72E-05	0.00013	11.680	242.69	0.00038	0.0093	-0.4893	4.5094	
SSEC	0.00013	0.00027	7.9734	85.157	-4.6E-05	0.0137	-0.9185	5.9848	
SSMI	5.51E-05	0.00012	21.799	680.43	0.00015	0.0094	-0.6778	6.3145	
STOXX50E	0.00011	0.00018	13.079	315.25	5.05E-05	0.0122	-0.1969	4.4480	

Note: *Those are the main descriptive statistics for the indices used in this study. Figure 1 and 2 in the below illustrate some examples of time series graphs and distributions.*

2.Figure 1 Some example plots of 5-min realized variance, daily return and their distributions

Figure 1 *rv5_spx* *returns_spx*

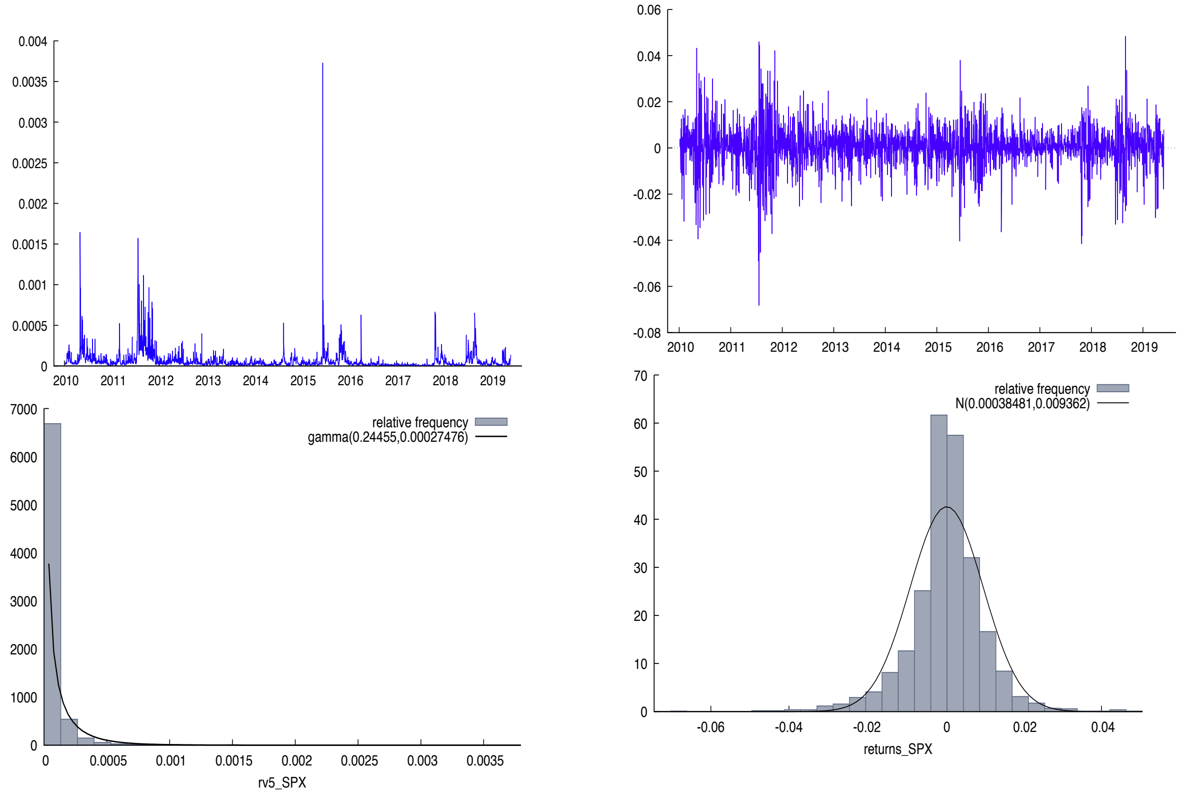
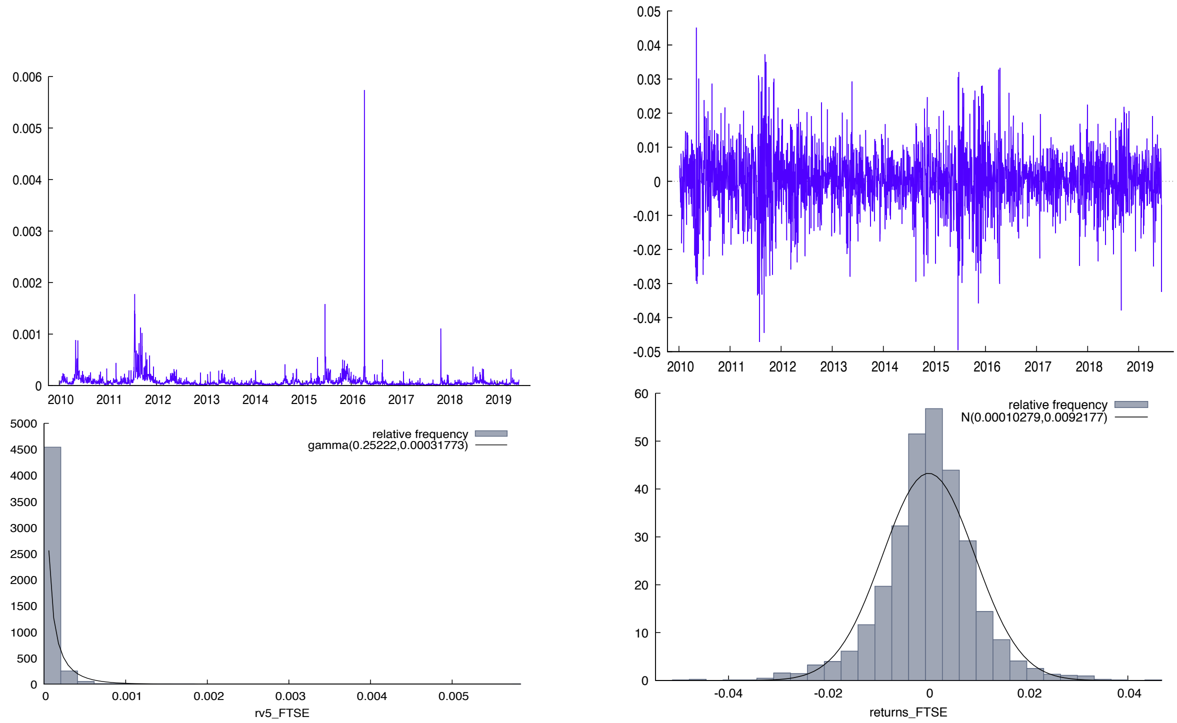


Figure 2 *rv5_ftse* *returns_ftse*



Liu, Patton, and Sheppard (2015) compare over 400 different realized measures and point out that it is difficult to significantly beat five-minutes realized variance. Therefore, we select 5-min realized variance as a proxy for the true volatility. The aim is to estimate this target as close as possible using the two groups of competing models such as conditional volatility (GARCHs) models and realized volatility (HAR-RV) models. It is important here to note that intraday frequencies are aggregated to the daily volatilities (which is called realized volatility). This enables us to compare these two groups of competing models and find out the most accurate genre of models. It is likely to say that if the frequency of data increases, the accuracy of volatility estimation will improve. On the other hand, microstructure noise may have an impact on the efficiency of ultra-high frequency data through measurement errors and price discreteness.

Table 2 presents the first four statistical moments of both series for different indices respectively, namely; mean, standard deviation, skewness, and excess kurtosis. The values of the moments are as commonly seen in the literature. The means of the returns and the realized variance series are close to zero for each index that is consistent. The returns series are moderately negative skewed whereas the series of 5-min realized variance have a high positive skew. On the other hand, the values of the fourth moment indicate the leptokurtic distribution for all the time series. Indeed, it can be pointed out that the series have a non-normal distribution. Figure 1 and 2 illustrate several example graphs for some indices.

2.3.2. Realized Volatility Models

Volatility is latent and therefore we need to provide a proxy of the true volatility. As a proxy, researchers used to employ the daily squared returns until the work of Andersen and Bollerslev (1998). They document that the squared returns is a poor proxy in comparison to the cumulative intraday squared returns. Afterwards, the so-called realized variance which is the sum of the squared intraday returns was defined by Andersen, Bollerslev, Diebold, and Labys (ABDL; 2003). Indeed, the realized variance and the daily squared returns are both unbiased estimates of volatility, yet the realized variance is highly efficient one.

$$RV_t = \sum_{i=1}^{m} r_{t,i}^2 \tag{1}$$

As can be shown in equation (1), the realized variance is calculated as the sum of squared intraday returns, where *m* is the number of intraday observations during day *t*. Theoretically, the higher *m* value is more accurate daily volatility estimation. However, when the number of *m* is too high, the efficiency of high frequency data can be distorted due to microstructure noise effect. ABDL (2003), Martens (2001), and Hol *and* Koopman (2002) suggest the frequency interval between 5 and 30 minutes. In a recent study, comparing over 400 different realized measures Liu, Patton, and Sheppard (2015) point out that it is difficult to significantly beat five-minute realized variance. Therefore, we utilise the 5-min realized variance for the estimations of the HAR and ARFIMA models.

Barndorff-Nielson et al. (2010) decompose the realized variance into positive and negative realized semivariances or good and bad volatilities.

$$RSV_t^+ = \sum_{i=1}^{m} r_{t,i}^2 \, I\,\{r_{t,i} > 0\} \tag{2}$$

$$RSV_t^- = \sum_{i=1}^{m} r_{t,i}^2 \, I\,\{r_{t,i} < 0\} \tag{3}$$

where $I\,\{\cdot\}$ is an indicator function. We should also note that $RV_t = RSV_t^+ + RSV_t^-$.

HAR-RV model

The HAR-RV model is based on the heterogeneous market hypothesis of Muller, Dacorogna, Dave, Olsen, Pictet and von Weizsacker (1997). According to the hypothesis, there are three types of investors that have different risk preferences and different reactions to the same new market information. In addition to the hypothesis, the same researchers develop the Heterogenous Autoregressive Conditional Heteroskedasticity (HARCH) model. Inspired by the HARCH model and its background hypothesis, Corsi (2009) proposes the HAR-RV model that is an additive cascade model of different volatility components. The model is specified as:

$$RV_{t+h}^d = \beta_0 + \beta_d RV_t^d + \beta_w RV_t^w + \beta_m RV_t^m + \varepsilon_{t+h} \tag{4}$$

where RV_t^d is daily realized volatility; RV_t^w refers to weekly realized volatility, and then RV_t^m indicates monthly realized volatility. RV_t^w and RV_t^m can easily be calculated as follows:

$$RV_t^w = \frac{1}{5}(RV_{t-5}^d + RV_{t-4}^d + \cdots + RV_{t-1}^d)$$

$$RV_t^m = \frac{1}{22}(RV_{t-22}^d + RV_{t-21}^d + \cdots + RV_{t-1}^d)$$

The main point of the HAR-RV model is to predict future volatility using three different volatility components; a daily (RV_t^d), a weekly (RV_t^w), and a monthly (RV_t^m) components. The HAR-RV model can simply be estimated by the ordinary least square (OLS) method. The model is such a good alternative to the ARFIMA model. The HAR-RV model can also capture long memory characteristics of volatility even though it is not in the class of long memory models. In practice, the HAR-RV model is found to be such a promising model as the model performance is remarkably good in spite of its simple structure.

Different types of investors have different objectives in financial markets. For instance, some investors are completely hedgers whilst some others are completely speculators. Hence the HAR-RV model is based on capturing different reactions of different investors through the simple autoregressive process. Financial interpretation of the model is that the investors are divided into three different categories. In the model, RV_t^d, RV_t^w, and RV_t^m components represent short-term, middle-term, and long-term investors respectively and indicate the degree of different investors' impact on current realized volatility. In other words, the model coefficients provide an understanding of how these different market participants react and perceive to volatility. Moreover, the HAR-RV model can successfully capture the persistence feature of realized volatility.

HAR-PS and HAR-RSV models

Barndorff-Nielsen, Kinnebrock, and Sheppard (2010) first introduce positive and negative realized semivariance measures, which are obtained from the signed high frequency intraday returns. Patton and Sheppard (2015) decompose only the daily explanatory HAR model component into negative and positive realized semivariances. In this study, we call Patton and Sheppard (2015)'s model as the HAR-PS model. The HAR-PS specification is presented in equation (5):

$$RV_{t+h}^d = \beta_0 + \beta_d^- RSV_t^- + \beta_d^+ RSV_t^+ + \beta_w RV_t^w + \beta_m RV_t^m + \varepsilon_{t+h} \quad (5)$$

Afterwards, it is added one more realized semivariance specification to the model comparison that decomposes not only the daily component, but also separates weekly and monthly components. The model of Patton and Sheppard (2011) is called here as the HAR-RSV and the model is given as follows:

$$RV_{t+h}^{d} = \beta_0 + \beta_d^- RSV_t^- + \beta_d^+ RSV_t^+ + \beta_w^- RSV_t^- + \beta_w^+ RSV_t^+ + \beta_m^- RSV_t^- + \beta_m^+ RSV_t^+ + \varepsilon_{t+h} \quad (6)$$

According to the seminal research of Patton and Sheppard (2011) and Barndorff-Nielsen, Kinnebrock, and Sheppard (2010), the decomposition of realized variance into positive and negative realized semivariances (or good and bad volatilities) adds more information for the prediction of future volatility.

ARFIMA-RV model

The long memory autoregressive fractionally integrated moving average (ARFIMA) model was first developed by Granger and Joyeux (1980). The ARFIMA model is in the class of long memory models and therefore can successfully capture the persistency feature of volatility. Andersen et al. (2003) suggest the univariate ARFIMA model in order to model the realized volatility. An ARFIMA *(p, d, q)* model is presented by:

$$\varphi(L)(1 - L)^{d}(RVt - \mu) = \theta(L)\varepsilon_t \quad (7)$$

where $\varphi(L)$ and $\theta(L)$ are the lag polynomials of the autoregressive *(AR)* and moving average *(MA)* components. ε_t is the error term which is distributed approximately as a Gaussian white noise [$N(0,\sigma_u^2)$]. The fractional differencing parameter is represented by *d* in equation (7). The *AR* and *MA* components explain the short memory properties of volatility and as for the *d*, it accounts for the long memory properties of volatility. The value of *d* is expected between 0 and 0.5 in order to capture long memory property. Andersen et al. (2003) found $d=0.401$. In this context, a general empirical conclusion with ARFIMA model is that this framework outperforms traditional GARCH models which are based on daily returns (Hansen and Lunde, 2010).

2.3.3. Conditional Volatility Models

In this study, we utilise the daily returns to estimate the GARCH family models. The daily returns are calculated by taking the logarithmic difference between today's and yesterday's closing prices.[5] Before passing to the GARCH models, we present the notation of the returns process, (r_t), as follows:

$$r_t = \mu_t + \varepsilon_t \ , \qquad (8)$$

where

$$\varepsilon_t = z_t h_t, \qquad z_t \xrightarrow{i.i.d.} (0,1)$$

The returns process, r_t, consists of two components, namely the conditional mean process, μ_t, and the innovation term, ε_t. The first component could include AR or MA terms and the second component can be decomposed as:

a-) an independent shock (noise term), z_t, with zero mean and constant variance,

b-) the conditional variance, h_t.

GARCH model

The Autoregressive Conditional Heteroskedasticity (ARCH) model was first suggested by Engle (1982). He aimed to capture the conditional variance with the ARCH model. if volatility can be correlated over time, then the change in variance can be modelled using the model. In other words, the variance of the current error is conditionally a function of its past values. The ARCH model is in the form of an AR process for the error variance and thus slow to respond to the large shocks. In other words, it is unable to successfully capture the volatility persistence.[6] Assuming an ARMA process for the error variance, Bollerslev (1986) proposed the Generalised ARCH (GARCH) model. The GARCH model does not only capture 'volatility clustering', but also include 'fat tails'. Therefore, we consider the GARCH and its some variants. The order of *p=q=1* for the GARCH models is sufficient to capture the volatility clustering. In the finance literature, higher order models are rarely estimated. Instead, the parsimonious order (*p=q=1*) is suggested. The specification of the GARCH model is given as follows:

---

[5] The formula of returns calculation is as follows: $r_t = \log{(p_t - p_{t-1})}$.

[6] Therefore, we do not include the ARCH model in our comparison.

$$h_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \, \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \, h_{t-j}^2 \quad (9)$$

The non-negativity constraints must be satisfied with $\alpha_0 > 0$, $\alpha_i \geq 0$ *for i=1,2,…,q and* $\beta \geq 0$ *for j=1,2,…,p*. Essentially, it generates one-step-ahead forecasts for the variance as a weighted average variance, α_0, the previous volatility, $\sum_{i=1}^{q} \alpha_i \, \varepsilon_{t-i}^2$, and the previous forecast variance, $\sum_{j=1}^{p} \beta_j \, h_{t-j}^2$. The model is covariance stationary if $\alpha_i + \beta_j$<1. The sum of the coefficients, $\alpha_i + \beta_j$, explains how persistent is the shocks to volatility. The GARCH model can capture some characteristics of returns such as volatility clustering and leptokurtosis.

In practice, bad news has a more significant effect than good news, which is known as leverage effect. There usually exists a negative correlation between the current stock returns and the future conditional variance. The GARCH model provides a way to model the change in variance, yet it is a symmetric model. It treats negative daily returns the same as positive daily returns. Therefore, the standard GARCH model is incapable of capturing the larger effect of bad news. For this reason, some extensions of the GARCH models have been introduced following Engle and Bollerslev (1982, 1986); such as exponential-GARCH model (Nelson, 1991), the threshold GARCH model (Rabemananjara and Zakoian, 1993), Asymmetric Power ARCH or APGARCH (Ding, Engle, and Granger, 1993) etc, [see –'Glossary to ARCH (GARCH)' (Bollerslev, 2009)– for more variants of GARCH models]. Each of those models attempts to improve the forecast accuracy by adding some idiosyncratic components on the traditional GARCH model.

EGARCH model

Nelson (1991) pointed out that there is a negative significant autocorrelation between returns and volatility. In order to capture the negative asymmetric effects of shocks he suggested the EGARCH as an alternative asymmetric model. The EGARCH is given as follows:

$$\log(h_t^2) = \alpha_0 + \alpha_1 \frac{|\varepsilon_{t-1}|}{h_{t-1}} + \gamma \frac{\varepsilon_{t-1}}{h_{t-1}} + \beta_1 \log(h_{t-1}^2) \quad (10)$$

where the parameter, γ, captures the leverage effect if $\gamma < 0$. A significant α_1 coefficient means that the volatility clustering is captured. This model is effective because the logarithmic form of the model will ensure the non-negativity for the conditional variance even if the model parameters are negative.

PGARCH model

The power GARCH model was developed by Ding, Engle, and Granger (1993), which is based on the standard deviation GARCH model of Taylor (1986) and Schwert (1989). Ding et al. (1993) indicate that both ways – modelling the variance (with the GARCH) and the standard deviation (with the TS-GARCH) – are capable of capturing the autocorrelation pattern of financial data. Here, we consider h_t^δ where the power parameter, δ, is the key point and can be estimated. The specifications of the power GARCH model is:

$$h_t^\delta = \alpha_0 + \alpha_1(|\varepsilon_{t-1}| - \gamma_1\varepsilon_{t-1})^\delta + \beta_1 h_{t-1}^\delta \tag{11}$$

where $\delta > 0$. The model is symmetric if $\gamma_1 = 0$. According to the values of $\delta\ and\ \gamma$, the model can be converted to the some other GARCH models. For example, the PGARCH specifications $\delta = 2\ and\ \gamma = 0$ is actually the standard GARCH model; $\delta = 1\ and\ \gamma = 0$ is the TS-GARCH; and $\delta = 2\ and\ 0 \leq \gamma \leq 1$ is the GJR-GARCH.

TGARCH (GJR) model

Glosten, Jagannathan, and Runkle (1993) proposed the Threshold GARCH model as an alternative to the other asymmetric models. This model includes a dummy variable in order to capture the asymmetric effects. However, if we set the threshold term to zero, the TGARCH model evolves to the standard GARCH model. The Threshold GARCH model specifications can be expressed as follows:

$$h_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \gamma_1\varepsilon_{t-1}^2 I_{t-1} + \beta_1 h_{t-1}^2 \tag{12}$$

In the model, I_{t-1} is the dummy variable that can successfully capture the leverage effect in financial data. The dummy variable, $I, is\ equal\ to$ 1 if $\varepsilon_{t-1} < 0$ and otherwise if $\varepsilon_{t-1} > 0$, it is zero. $\varepsilon_{t-1} > 0$ represents good news and $\varepsilon_{t-1} < 0$ is bad news. The impact of good news is only α_1, whilst bad news is $\alpha_1 + \gamma_1$. Bad news has a greater impact on volatility than good news if $\gamma_1 > 0$.

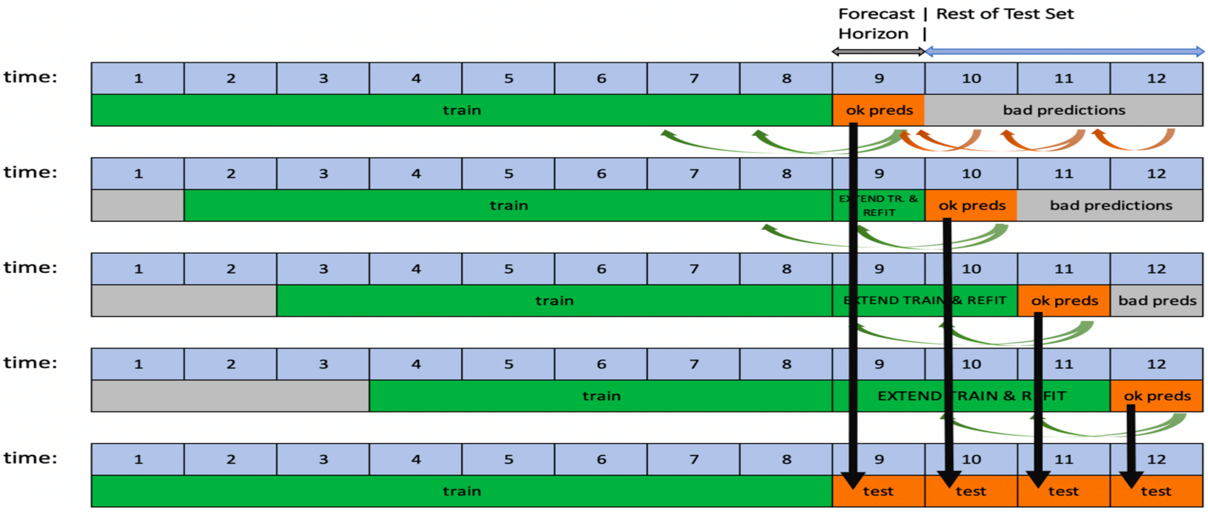
2.3.4. Forecast Evaluation Procedure

We use both the rolling and recursive windows for forecasting. The loss functions; the mean squared error (MSE), the mean absolute error (MAE), and the quasi-Gaussian log-likelihood (QLIKE) are considered in order to compare the models. Lastly, the Giacomini and White (2006) pairwise test is employed to evaluate the forecasting performance of two models.

Rolling and Recursive Windows

Figure 2 illustrates the rolling window's working mechanism.[7] The rolling window is one of the most popular methods in forecasting and therefore this study employs the rolling window technique in order to generate the volatility forecasts of stock markets. Initially, the whole sample needs to be divided into two subgroups such as the initial sample and out-of-sample windows. In the literature, there is no consensus on how to select an appropriate forecasting window. Since the main objective of this work is to evaluate the out-of-sample performance of the models, we arbitrarily choose the initial and out-of-sample windows considering a length that allows the regression fit normally and obtain longer out-of-sample period.

2.Figure 2 Rolling window's working principle



---

[7] Figure 1 is taken from the website: https://docs.h2o.ai/driverless-ai/1-8-lts/docs/userguide/time-series.html

According to Figure 2, the rolling window's working principle does work the way that the estimation sample is then rolled forward by adding one new observation and dropping the most distant observation. In this way, the size of initial sample window used to estimate the models remains at a fixed length. However, the recursive window technique makes use of an expanding windows that does not drop the first observation of the initial sample during the whole forecasting process. In Figure 2, the first four lines demonstrate how the loop command works and then the last line shows the pre-defined training and test samples that are green (time horizon between 1-8) and orange (time horizon between 9-12) respectively. Throughout this thesis, we produce only the one-step-ahead volatility forecasts of the stock markets. The reason is that the forecasts more than one-step-ahead are highly likely to give poorer forecasts due to the lack of information of further prediction. Therefore, they are labelled in the Figure 2 as bad predictions.

In the literature, there is no clear evidence whether the rolling and recursive windows should be employed. For example, Corsi et al. (2008) employ the recursive windows for the realized volatility of the S&P500 index, while Pu, Chen, and Ma (2016) use the rolling scheme for forecasting the realized volatility in the Chinese stock market. Vortelinos (2017) uses both of them for generating the realized volatility forecasts of some US assets and point out that there is no difference between the rolling and recursive methods in terms of forecast accuracy. We should also point out that the initial sample size of both of these methods highly matters and different initial sample lengths could generate different forecast results on the same data. In this thesis, we keep the length of in-sample period a little shorter as 1 to 2 years to obtain longer out-of-sample period but also allow the regression fit feasibly.

Loss functions

Since the main goal of this work is to compare the performance of the competing models, we need to measure the ability of the models using some loss functions. Many different forecasting criteria can be used for comparison purpose. Lopez (2001) points out that it is not clear to decide which measure is the most accurate to which model. On the other hand, Patton (2011) documents the robustness of the QLIKE and MSE criteria. The reason is explained as such: in the case of such a noisy volatility proxy, the QLIKE and MSE provide consistent rankings for volatility models. In this regard, three of the most popular measures are selected, namely the mean squared error (MSE), the mean absolute error (MAE), and the quasi-Gaussian log-likelihood (QLIKE). The loss functions are specified as follows:

$$QLIKE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}[log\widehat{RV_t^2} + \frac{\widehat{RV_t^2}}{RV_t^2}] \tag{13}$$

$$MSE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}[RV_t^2 - \widehat{RV_t^2}]^2 \tag{14}$$

$$MAE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau} \mid RV_t^2 - \widehat{RV_t^2} \mid \tag{15}$$

where RV_t^2 is the proxy of the true volatility and $\widehat{RV_t^2}$ is the volatility forecast. The number of observations is represented by τ. The MAE is based on taking absolute values of errors and equally treats all the losses, while the MSE punishes larger losses. If the absolute value is not taken, it becomes the mean biased error (MBE). The MBE can give useful information, yet the negative and positive errors are likely to cancel out each other. The QLIKE and MSE are frequently used ones in the literature due to being robust to the noisy volatility proxy. Patton and Sheppard (2009) indicate that the QLIKE is powerful in the Diebold-Mariano test, which is quite similar test to the Giacomini and White (GW) test that we use here. Although the MSE, MAE, and QLIKE are the most frequently used criteria, there is still a possibility that such a model with the lower error may not be exactly better than the other model. For this reason, we need to apply the GW test.

Conditional Giacomini-White (GW) test

For instance, we have two different forecasted series, namely X and Y. Assuming that the values of loss functions of X are lower than Y. Can we say that the forecast X has a superior performance compared to the forecast Y? Or is it possible that the difference between the forecasts X and Y is inherently insignificant? In order to test conditional predictive ability Giacomini and White (2006) suggest a pairwise test on equal conditional predictive ability, which examines whether two different forecasting models statistically have the same accuracy or not. The GW test is a better alternative to the so-called test of Diebold and Mariano (1995) and White (1996) (Hereafter DM test). Compare to the DM test, the GW test has several advantages; the GW can capture the effect of estimation uncertainty on relative forecast performance and also it can deal with the forecasts which are based on both nested and non-nested models.

Let $L(y_t; \hat{y_t})$ denote the loss of a model's forecast where $\hat{y_t}$ is the forecasted value and y_t is the proxy of the true value. The difference between the loss from model i and a benchmark model o can be defined as follows:

$$d_{i,t} = L\big(y_t; \widehat{y_{o,t}}\big) - L\big(y_t; \widehat{y_{\iota,t}}\big) \tag{16}$$

The null hypothesis is $H_0: E\big(d_{i,t+\tau} \big| h_t\big) = 0$ and h_t is some information set. Then, the conditional predictive ability (CPA) test statistic can be calculated as a Wald statistic:

$$CPA_t = T(T^{-1} \textstyle\sum_{t=1}^{T-\tau} h_t d_{i,t+\tau})' \, \widehat{\Omega}_T^{-1} \big(T^{-1} \textstyle\sum_{t=1}^{T-\tau} h_t \, d_{i,t+\tau}\big) \sim \chi_1^2 \tag{17}$$

where the symbol, $\widehat{\Omega_T}$ is the Newey and West (1987) HAC estimator for the asymptotic variance of $h_t d_{i,t+\tau}$. Under the squared error metric, the CPA is used to evaluate whether any model outperforms the random walk or not.

2.4. Empirical Results

In this section, we present the in-sample and out-of-sample empirical results of the conditional volatility (GARCH) models and the realized volatility (HAR-RV and ARFIMA-RV) models. Afterwards, the results are tested and evaluated in terms of each stock markets in more detail.

2.4.1. Parameter Estimates and in-sample evaluation

This subsection evaluates the in-sample estimation results of competing models. First, we present the estimation results of the HAR and ARFIMA models together. Afterwards, the GARCH family models' estimation results are given. According to Swanson, Elliott, Ghysels, and Gonzalo (2006), a preferred model should be selected by considering the out-of-sample performance of the models rather than their in-sample fit. It is also worth to note that the in-sample fit of a model does not necessarily mean the out-of-sample accuracy too. Since the main objective of this work is to evaluate the out-of-sample performance of the models, we give a short in-sample evaluation here.

HAR AND ARFIMA model estimation results

The HAR-type models are found to be promising models. In practice, the HAR model exhibits remarkably good performance in spite of its simple structure. The main point of the model is that in terms of the reactions of different market participants to the same information, investors are divided into three different categories; short-term, middle-term, and long-term investors. In other words, the model coefficients provide an understanding of how these different market participants react and perceive to volatility. Table 3 shows some estimation results of competing models for the SPX, FTSE, FCHI, and N225 indices. According to the results, volatility is highly persistent because the sum of the model coefficients ($C + \beta_d + \beta_w + \beta_m$) for each one of the indices is close to unit and highly significant, which means that the volatility of past contains useful information about future volatility. The daily component of the model has relatively larger impact on the future volatility in comparison to the weekly and monthly components, which is consistent with the theory.

Patton and Sheppard (2011) suggest that the decomposition of realized variance into positive and negative realized semivariances (or good and bad volatilities) adds more information for the prediction of future volatility. As it is mentioned in the previous paragraph, the most recent (or daily) realized variance component of the HAR model has larger effect on the future volatility compared with the weekly and monthly components. Therefore, we first decompose only the daily explanatory HAR model component into negative and positive realized semivariances. In this study, we call this model as the HAR-PS model. The results of the HAR-PS model provide strong evidence that the decomposition of the daily realized variance into its signed realized semivariances improves the model fit. This is because the coefficients of negative realized semivariance for the indices are higher and more significant compared to the coefficients on positive realized semivariance. In addition, we add one more realized semivariance specification that decomposes not only the daily component, but also do weekly and monthly components. This model is called as the HAR-RSV. It should be noted that when we decompose all the HAR model components (daily, weekly, and monthly) into positive and negative realized semivariances, the role of the daily component diminishes. Also, the increasing number of explanatory variables results in the overfitting issue by consuming the model's degree of freedom. As can be seen in Table 3, the HAR-PS regression coefficients are more consistent and significant than the HAR-RSV model results. From this point of view, we would expect the HAR-PS model to generate better results.

On the other hand, ARFIMA models do not have a clear economic interpretation, yet they can capture dynamic characteristics of the data such as long memory, asymmetry, and volatility clustering. The fractional differencing parameters, d values, are between 0 and 0.5 for all the indices and highly significant. These values are consistent with the theory. The ARFIMA model is selected among parsimonious models to avoid overfitting issue. In our framework, we adjust the lag of AR and MA components to one (p=q=1). The reason is that parsimonious models generate better results compared to over-parameterised models. In econometric analysis, too many explanatory variables –in this context AR and MA components– consume the degree of freedom in a model, which means that explanatory variables contribute very little to the dependent variable.

Estimation results of the GARCH-type models

In this subsection, we summarise the estimation results of competing GARCH models. In the GARCH models (Table 3, –Panel A, B, C, D, E, and F–), the coefficients of the lagged conditional variance (β_1) and lagged squared residual (α_1) terms are highly significant. The variance intercept terms are quite small. The lagged squared residual terms are usually between 0.1 and 0.2, whilst the lagged conditional variance terms around 0.8. More importantly, the sum of the coefficients in the models is close to one. This convergence towards unity implies that shocks to the conditional variance are quite persistent. For our financial returns data, the results are as expected from the typical GARCH models.

The coefficients of asymmetry (γ) are significant in the EGARCH and TGARCH models, showing the existence of the leverage effect (i.e. the negative shocks tend to have a larger impact on the volatility in comparison to positive shocks). In the PGARCH models, estimated values of the power parameter (δ) are statistically significant and disperse around 1 (e.g. 0.88, 0. 90, 1.01, 0.95). The model uses the conditional standard deviation, as Ding et al. (1993) indicate that both ways – modelling the variance (with the GARCH) and the standard deviation (with the TS-GARCH) – are capable of capturing the pattern of financial data such as leverage effects.

Swanson et al. (2006) emphasise that the best model should be chosen by comparing the out-of-sample performances of competing models rather than their in-sample fit. Moreover, the in-sample fit of a model does not necessarily mean the out-of-sample accuracy of the model. Therefore, we aim to concentrate more on the results of the out-of-sample analysis.

2.Table 3 Full-sample estimation results of the HAR and GARCH models for some indices

Panel A: *HAR model results*

Coef. → ↓ Index	β_0	β_d	β_w	β_m
SPX	*1.23E-05*	*0.329*	*0.212*	*0.273*
	(3.40E-06)	(0.101)	(0.071)	(0.075)
	***	***	***	**
FTSE	*2.02E-05*	*0.090*	*0.328*	*0.326*
	(6.43E-06)	(0.079)	(0.113)	(0.106)
	***		***	***
FCHI	*1.15E-05*	*0.428*	*0.250*	*0.200*
	(3.02E-06)	(0.097)	(0.090)	(0.077)
	***	***	***	***
N225	*1.97E-05*	*0.358*	*0.155*	*0.211*
	(3.92E-06)	(0.096)	(0.069)	(0.059)
	***	***	**	***

Panel B: *ARFIMA model results*

Coef. → ↓ Index	C	d	AR(1)	MA(1)
SPX	*6.92E-05*	*0.302*	*0.642*	*-0.575*
	(5.00E-05)	(0.033)	(0.065)	(0.066)
		***	***	***
FTSE	*8.07E-05*	*0.316*	*-0.187*	*0.034*
	(6.84E-05)	(0.021)	(0.080)	(0.100)
		***	**	
FCHI	*9.65E-05*	*0.378*	*0.579*	*-0.494*
	(9.26E-05)	(0.025)	(0.044)	(0.038)
		***	***	***
N225	*6.90E-05*	*0.267*	*0.091*	*0.029*
	(5.79E-05)	(0.029)	(0.176)	(0.156)
		***		

Panel C: *HAR-RSV model results*

Coef. → ↓ Index	C	β_d^-	β_d^+	β_w^-	β_w^+	β_m^-	β_m^+
SPX	*1.36E-05*	*0.391*	*0.158*	*0.286*	*0.200*	*0.486*	*0.047*
	(3.82E-06)	(0.119)	(0.160)	(0.171)	(0.161)	(0.317)	(0.313)
	***	***		*			
FTSE	*1.58E-05*	*0.085*	*0.163*	*0.252*	*0.471*	*-0.119*	*0.778*
	(4.50E-06)	(0.029)	(0.071)	(0.081)	(0.143)	(0.181)	(0.236)
	***	***	**	***	***		***
FCHI	*1.23E-05*	*0.735*	*-0.011*	*0.839*	*-0.326*	*0.301*	*0.168*
	(3.20E-06)	(0.186)	(0.106)	(0.299)	(0.269)	(0.433)	(0.417)
	***	***		***			
N225	*1.92E-05*	*0.470*	*0.119*	*0.176*	*0.244*	*0.042*	*0.413*
	(3.79E-06)	(0.145)	(0.135)	(0.136)	(0.242)	(0.234)	(0.281)
	***	***					

Panel D: *HAR-PS model results*

Coef. → ↓ Index	C	β_d^-	β_d^+	β_w	β_m
SPX	*1.26E-05*	*0.406*	*0.147*	*0.258*	*0.272*
	(3.38E-06)	(0.118)	(0.159)	(0.089)	(0.081)
	***	***		***	***
FTSE	*1.97E-05*	*0.062*	*0.218*	*0.297*	*0.315*
	(5.41E-06)	(0.103)	(0.152)	(0.128)	(0.100)
	***			**	***
FCHI	1.16E-05	*0.855*	*-0.099*	*0.289*	*0.203*
	(3.13E-06)	(0.208)	(0.105)	(0.097)	(0.083)
	***	***		***	**
N225	*2.02E-05*	*0.459*	*0.132*	*0.198*	*0.218*
	(3.96E-06)	(0.154)	(0.135)	(0.061)	(0.060)
	***	***		***	***

Panel E: *GARCH model results* **Panel F:** *EGARCH model results*

Coef. → ↓ Index	α_0	α_1	β_1	α_0	α_1	γ	β_1
SPX	*3.73E-06* (8.46E-07) ***	*0.171* (0.025) ***	*0.790* (0.024) ***	*-0.683* (0.089) ***	*0.168* (0.023) ***	*-0.217* (0.024) ***	*0.942* (0.008) ***
FTSE	*4.28E-06* (1.27E-06) ***	*0.136* (0.026) ***	*0.813* (0.035) ***	*-0.650* (0.158) ***	*0.165* (0.030) ***	*-0.158* (0.023) ***	*0.945* (0.014) ***
FCHI	*3.85E-06* (1.29E-06) ***	*0.124* (0.026) ***	*0.853* (0.028) ***	*-0.450* (0.111) ***	*0.132* (0.027) ***	*-0.181* (0.026) ***	*0.961* (0.011) ***
N225	*7.16E-06* (2.27E-06) ***	*0.133* (0.025) ***	*0.830* (0.030) ***	*-0.695* (0.140) ***	*0.200* (0.027) ***	*-0.152* (0.025) ***	*0.939* (0.014) ***

Robust standard errors in parentheses.
** %10, **%5, ***%1; the stars are significance levels for p values*

2.4.2. Out-of-sample evaluation

The out-of-sample period is 8 years between 2011-2019, containing approximately 2070 daily observations. A set of GARCH-type models, employing the conditional variance, is compared to the HAR and ARFIMA models, which uses the data of 5-min realized variance. One-day-ahead out-of-sample volatility forecasts are generated using both the rolling and recursive windows forecasting techniques. The out-of-sample forecast losses are measured using the MSE, MAE, and QLIKE criteria and presented in Tables 4, 5, and 6.

We could not find any significant difference between the results of rolling and recursive methods. For this reason, only recursive windows forecasting results are inserted in Tables 4, 5, and 6; the rolling one is in the Appendix. The parallel results of the fixed and expanding windows does not change the way we interpret. In order not to cover the pages with long tables we insert the rolling windows tables in the Appendix Table 1, 2, and 3.

2.Table 4 **QLIKE for recursive windows forecast models**

	GARCH	EGARCH	PGARCH	TGARCH	HAR-RV	HAR-PS	HAR-RSV	ARFIMA-RV
AEX	-8.6969	-8.7343*	-8.7213	-8.7134	**-8.8430****	-8.8420	-8.8407	-8.6583
AORD	-9.1461	-9.1668*	-9.1530	-9.1574	-9.2819	**-9.2853****	-9.2673	-9.0719
BFX	-8.6885	-8.7071*	-8.6997	-8.6786	-8.8527	-8.8528	**-8.8532****	-8.7515
BSESN	-8.7496*	-8.7485	-8.7313	-8.7392	**-8.8623****	-8.8615	-8.8567	-8.7286
BVLG	-8.8701	-8.8942*	-8.8817	-8.8805	**-9.2605****	-9.2395	-9.1763	-9.0899
BVSP	-8.0710	-8.0921*	-8.0805	-8.0754	**-8.3105****	-8.2945	-8.2861	-8.2306
DJI	-8.9694	-8.9953	**-9.0121****	-8.9824	-8.9790	-8.9840*	-8.9828	-8.6638
FCHI	-8.3883	-8.4131*	-8.3987	-8.4065	-8.5450	**-8.5500****	-8.5490	-8.3846
FTMIB	-8.0176	-8.0604*	-8.0430	-8.0400	-8.3441	**-8.3466****	-8.3431	-8.1700
FTSE	-8.6845	-8.6847	-8.6833	-8.6849*	-8.7075	**-8.7084****	-8.6668	-8.5534
GDAXI	-8.3844	-8.3870	-8.3903*	-8.3820	-8.5109	**-8.5140****	-8.5130	-8.3626
GSPTSE	-9.2795	-9.2923*	-9.2830	-9.2848	-9.3334	-9.3357	**-9.3397****	-9.0713
HSI	-8.5916	-8.6143*	-8.5896	-8.5900	**-8.8725****	-8.8724	-8.8666	-8.7170
IBEX	-8.0215	-8.0488*	-8.0353	-8.0308	**-8.1147****	-8.1119	-8.1095	-7.9500
IXIC	-8.8001	-8.8100	-8.8322*	-8.8045	-9.0520	**-9.0558****	-9.0374	-8.9107
KS11	-9.0485*	-9.0410	-9.0353	-9.0190	-9.2606	-9.2596	**-9.2649****	-9.2007
KSE	-8.8503*	-8.8389	-8.8407	-8.8406	-9.0144	-9.0163	**-9.0378****	-8.8708
MXX	-8.7525	-8.7383	-8.7561*	-8.7426	-8.7741	-8.7792	**-8.7819****	-8.2632
N225	-8.4380	-8.4584*	-8.4532	-8.4466	**-8.7708****	-8.7526	-8.7670	-8.6718
NSEI	-8.7568*	-8.7538	-8.7373	-8.7442	**-8.8788****	-8.8780	-8.8744	-8.6978
OMXC20	-8.4406	-8.4485	**-8.4505****	-8.4432	-8.4475*	-8.4293	-8.3141	-8.1909
OMXHPI	-8.6659	-8.6937*	-8.6759	-8.6862	**-8.6986****	-8.6938	-8.6664	-8.6894
OMXSPI	-8.8249	-8.8286	-8.8265	-8.8333*	-8.9579	-8.9805	**-9.0016****	-7.9089
OSEAX	-8.5604	-8.5834*	-8.5817	-8.5734	-8.6303	**-8.6371****	-8.6183	-8.4479
RUT	-8.6053*	-8.5944	-8.6029	-8.5938	-8.9399	**-8.9454****	-8.9416	-8.6094
SMSI	-8.1075	-8.1410*	-8.1257	-8.1217	-8.2246	**-8.2285****	-8.2281	-7.9962
SPX	-8.9798	-8.9978	-9.0184*	-8.9871	-9.0687	**-9.0716****	-9.0634	-8.7008
SSEC	-8.1963*	-8.1574	-8.1450	-8.1816	**-8.3718****	-8.3654	-8.3547	-8.0772
SSMI	-8.9164	-8.9392*	-8.9212	-8.9171	-9.0730	**-9.0746****	-9.0708	-8.8375
STOXX50E	-8.2606	-8.2840*	-8.2768	-8.2740	-8.3647	**-8.3652****	-8.3614	-8.1962

Notes: *** with bold indicates the best performing model, underlined with * is the best performing GARCH model, regardless of the HAR and ARFIMA models.*

2.*Table 5* **MSE for recursive windows forecast models**

	GARCH	EGARCH	PGARCH	TGARCH	HAR-RV	HAR-PS	HAR-RSV	ARFIMA-RV
AEX	1.19E-08	9.94E-09*	1.18E-08	1.31E-08	8.29E-09**	8.40E-09	8.42E-09	9.42E-09
AORD	3.82E-09	3.11E-09*	3.92E-09	3.57E-09	2.34E-09	2.33E-09	2.24E-09**	3.37E-09
BFX	9.73E-09	8.60E-09*	9.72E-09	1.28E-08	4.85E-09**	5.06E-09	5.19E-09	5.47E-09
BSESN	4.75E-09	4.57E-09*	5.71E-09	5.36E-09	2.94E-09	2.93E-09**	2.94E-09	3.76E-09
BVLG	9.24E-09	9.29E-09	8.95E-09*	1.04E-08	1.09E-09**	1.14E-09	1.31E-09	1.87E-09
BVSP	2.39E-08	1.99E-08*	2.75E-08	2.88E-08	9.29E-09**	9.61E-09	9.46E-09	1.06E-08
DJI	2.38E-08	2.13E-08**	2.21E-08	2.44E-08	2.47E-08*	2.54E-08	2.59E-08	2.57E-08
FCHI	1.90E-08	1.63E-08*	2.01E-08	2.39E-08	1.01E-08	9.96E-09**	1.00E-08	1.30E-08
FTMIB	4.75E-08	3.73E-08*	4.36E-08	6.03E-08	1.09E-08	1.08E-08**	1.09E-08	1.76E-08
FTSE	2.42E-08	2.31E-08**	2.37E-08	2.32E-08	2.59E-08*	2.68E-08	2.66E-08	2.67E-08
GDAXI	1.64E-08	1.50E-08*	1.54E-08	2.14E-08	1.31E-08	1.30E-08	1.38E-08	1.27E-08**
GSPTSE	8.28E-09	7.57E-09	8.55E-09	7.54E-09**	8.16E-09	8.35E-09	8.39E-09	8.65E-09
HSI	1.48E-08	1.09E-08*	1.50E-08	1.29E-08	4.17E-09	4.16E-09	4.14E-09**	5.46E-09
IBEX	4.03E-08	3.46E-08*	3.99E-08	5.08E-08	2.80E-08**	3.27E-08	3.38E-08	3.43E-08
IXIC	1.37E-08	9.79E-09*	1.10E-08	1.43E-08	5.53E-09	5.04E-09**	5.05E-09	6.78E-09
KS11	1.00E-08	7.54E-09*	9.20E-09	1.11E-08	4.24E-09	4.85E-09	5.33E-09	4.06E-09**
KSE	6.12E-09*	6.24E-09	6.26E-09	8.27E-09	2.70E-09	2.67E-09**	2.71E-09	3.42E-09
MXX	2.04E-08	1.94E-08**	2.03E-08	2.10E-08	1.98E-08	2.02E-08	2.08E-08	2.07E-08
N225	3.13E-08	2.65E-08*	2.89E-08	3.21E-08	1.36E-08**	1.41E-08	1.38E-08	1.40E-08
NSEI	5.92E-09	5.80E-09*	7.11E-09	6.69E-09	3.88E-09	3.87E-09	3.86E-09**	4.98E-09
OMXC20	3.97E-08	3.95E-08*	3.96E-08	4.13E-08	3.90E-08**	3.93E-08	4.05E-08	4.42E-08
OMXHPI	1.32E-08	1.04E-08*	1.38E-08	1.26E-08	7.85E-09	7.81E-09	7.63E-09**	1.15E-08
OMXSPI	1.36E-08	1.30E-08*	1.57E-08	1.82E-08	8.77E-09	8.25E-09**	8.29E-09	6.77E-08
OSEAX	1.94E-08	1.68E-08**	1.82E-08	1.90E-08	1.76E-08	1.79E-08	1.81E-08	1.95E-08
RUT	2.67E-08	1.70E-08*	1.98E-08	2.67E-08	7.74E-09	7.26E-09**	7.64E-09	1.12E-08
SMSI	6.18E-08	5.50E-08*	5.98E-08	7.23E-08	5.02E-08	6.00E-08	6.21E-08	5.46E-08
SPX	1.45E-08	1.20E-08**	1.28E-08	1.42E-08	1.34E-08	1.36E-08	1.37E-08	1.64E-08
SSEC	5.85E-08*	6.07E-08	6.39E-08	6.04E-08	4.27E-08	4.12E-08**	4.25E-08	6.66E-08
SSMI	1.76E-08	1.98E-08	1.58E-08*	2.27E-08	1.48E-08**	1.69E-08	1.78E-08	1.52E-08
STOXX50E	3.28E-08	2.98E-08*	3.23E-08	3.82E-08	2.81E-08**	3.13E-08	3.13E-08	2.90E-08

Notes: *** with bold indicates the best performing model, underlined with * is the best performing GARCH model, regardless of the HAR and ARFIMA models.*

2.Table 6 **MAE for recursive windows forecast models**

	GARCH	EGARCH	PGARCH	TGARCH	HAR-RV	HAR-PS	HAR-RSV	ARFIMA-RV
AEX	6.19E-05	5.73E-05*	6.13E-05	6.36E-05	3.27E-05	**3.24E-05****	3.26E-05	4.97E-05
AORD	4.04E-05	3.79E-05*	4.08E-05	3.99E-05	2.15E-05	**2.13E-05****	2.17E-05	3.09E-05
BFX	6.10E-05	5.84E-05*	6.15E-05	6.50E-05	2.82E-05	**2.78E-05****	2.79E-05	3.74E-05
BSESN	4.85E-05	4.84E-05*	5.29E-05	5.11E-05	2.62E-05	**2.61E-05****	2.65E-05	4.00E-05
BVLG	7.35E-05	6.80E-05*	7.23E-05	7.39E-05	**1.66E-05****	1.70E-05	1.85E-05	2.78E-05
BVSP	0.000120	0.000111*	0.000121	0.000121	**4.34E-05****	4.39E-05	4.38E-05	5.53E-05
DJI	4.92E-05	4.60E-05*	4.63E-05	5.03E-05	**4.36E-05****	4.38E-05	4.46E-05	6.67E-05
FCHI	8.50E-05	8.20E-05*	8.82E-05	8.97E-05	4.24E-05	**4.19E-05****	4.21E-05	6.42E-05
FTMIB	0.000158	0.000144*	0.000154	0.000158	5.10E-05	**5.02E-05****	5.09E-05	7.92E-05
FTSE	5.10E-05	4.82E-05*	5.05E-05	4.98E-05	4.63E-05	4.63E-05	**4.62E-05****	6.54E-05
GDAXI	7.63E-05*	7.84E-05	7.88E-05	8.46E-05	4.59E-05	**4.53E-05****	4.58E-05	6.31E-05
GSPTSE	3.52E-05	3.43E-05*	3.81E-05	3.44E-05	2.63E-05	2.63E-05	**2.62E-05****	4.31E-05
HSI	8.76E-05	7.95E-05*	8.80E-05	8.48E-05	2.61E-05	**2.60E-05****	2.66E-05	3.47E-05
IBEX	0.000101	9.35E-05*	0.000101	0.000108	**6.46E-05****	6.58E-05	6.65E-05	0.000100
IXIC	7.05E-05	6.83E-05	6.76E-05*	7.29E-05	2.89E-05	**2.79E-05****	2.81E-05	3.97E-05
KS11	5.49E-05	5.34E-05*	5.63E-05	5.99E-05	**2.00E-05****	2.05E-05	2.11E-05	2.54E-05
KSE	5.73E-05*	5.79E-05	5.82E-05	6.17E-05	2.71E-05	2.70E-05	**2.69E-05****	3.65E-05
MXX	5.11E-05	4.99E-05*	5.11E-05	5.18E-05	**3.95E-05****	3.96E-05	4.07E-05	4.99E-05
N225	0.000123	0.000116*	0.000119	0.000122	4.10E-05	4.14E-05	**4.07E-05****	4.88E-05
NSEI	5.22E-05*	5.25E-05	5.77E-05	5.53E-05	2.76E-05	**2.74E-05****	2.78E-05	4.51E-05
OMXC20	7.50E-05	7.36E-05*	7.42E-05	7.69E-05	**5.86E-05****	5.93E-05	6.36E-05	7.64E-05
OMXHPI	8.47E-05	7.85E-05*	8.44E-05	8.15E-05	6.10E-05	6.00E-05	**5.77E-05****	7.05E-05
OMXSPI	7.29E-05*	7.50E-05	7.95E-05	7.89E-05	4.07E-05	3.90E-05	**3.69E-05****	0.000232
OSEAX	6.86E-05	6.40E-05*	6.63E-05	6.76E-05	4.88E-05	**4.86E-05****	4.90E-05	7.09E-05
RUT	9.90E-05	9.87E-05*	9.98E-05	0.000102	3.36E-05	**3.27E-05****	3.31E-05	5.27E-05
SMSI	0.000111	0.000101*	0.000109	0.000115	**6.79E-05****	6.83E-05	6.87E-05	0.000110
SPX	5.23E-05	4.98E-05	4.94E-05*	5.26E-05	3.98E-05	**3.97E-05****	4.04E-05	6.41E-05
SSEC	0.000121*	0.000124	0.000132	0.000122	6.91E-05	**6.84E-05****	7.02E-05	9.09E-05
SSMI	5.30E-05	4.98E-05*	5.24E-05	5.42E-05	**2.60E-05****	2.62E-05	2.70E-05	4.41E-05
STOXX50E	8.69E-05	8.36E-05*	8.81E-05	9.33E-05	5.65E-05	**5.64E-05****	5.65E-05	8.20E-05

Notes: *** with bold indicates the best performing model, underlined with * is the best performing GARCH model, regardless of the HAR and ARFIMA models.*

The GARCH family models, namely GARCH, EGARCH, PGARCH, and TGARCH are compared against the HAR-(RV, RSV, PS) and ARFIMA models. In terms of the loss functions, the results of the Tables are mixed for the above mentioned models. What is obvious in the Tables is that the realized variance is found to be a more precise measure of the true volatility in comparison with the conditional variance. This result is consistent with the seminal work of Andersen and Bollerslev (1998). Considering the performance of competing models, there is no doubt that the HAR models outperform the ARFIMA and GARCH-type models. The success of the HAR specification is documented by the vast majority of the indices. However, there are still some exceptional indices (DJI, FTSE, GSPTSE, MXX, OSEAX, SPX) for the MSE criteria that the EGARCH model outperforms the HAR models. We believe that those exceptions occur by chance or abnormal data structure because the conditional Giacomini and White (GW) test results for those exceptional results are insignificant.

2.*Table 7* ***Summary results for competing models***

Rolling windows	**GARCH**	**EGARCH**	**PGARCH**	**TGARCH**	**HAR-RV**	**HAR-RSV and HAR-PS**		**ARFIMA-RV**
QLIKE Winner *(GARCH winner)*	**0** (1)	**0** (7)	**1** (20)	**0** (2)	**19**	**1**	**9**	**0**
MSE Winner *(GARCH winner)*	**0** (3)	**4** (22)	**0** (1)	**2** (4)	**15**	**3**	**6**	**2**
MAE Winner *(GARCH winner)*	**0** (2)	**1** (23)	**0** (2)	**0** (3)	**12**	**17**	**0**	**0**
Recursive windows								
QLIKE Winner *(GARCH winner)*	**0** (6)	**0** (16)	**2** (6)	**0** (2)	**10**	**6**	**13**	**0**
MSE Winner *(GARCH winner)*	**0** (2)	**5** (25)	**0** (2)	**1** (1)	**10**	**4**	**8**	**2**
MAE Winner *(GARCH winner)*	**0** (5)	**0** (23)	**0** (2)	**0** (0)	**9**	**6**	**15**	**0**
TOTAL WINNER	**0**	**10**	**3**	**3**	**75**	**88**		**4**

First rows (bold numbers) show the numbers of winners for all the indices. Second rows (parenthesis ones) are only for GARCH models.

Table 7 exhibits the summary results of the models for all the indices. Confirming the superiority of the HAR specification, we make another assessment only between the HAR models. The HAR-RV, HAR-RSV, and HAR-PS models are compared to each other and it is found that the decomposition of realized variance into positive and negative realized semivariances improves the out-of-sample forecasts of realized variance. We should remember here that while the HAR-PS model decomposes only the daily component, the HAR-RSV model decomposes the daily, weekly and monthly components. The contribution of the HAR-PS model on the out-of-sample forecasts is larger than the contribution of the HAR-RSV model. The reason is that the decomposition of the daily, weekly and monthly components into the realized semivariances diminishes the role of the daily realized variance which contains the most important information for forecasting future volatility.

Following the HAR type models, the ARFIMA model also outperforms the GARCH family models as it uses such a more precise measure of the true volatility that is 5-min realized variance. This result is confirmed by the MSE and MAE criteria, whereas the QLIKE values are indecisive between the ARFIMA and EGARCH models. In order to make the final decision between them we check the GW conditional predictive ability test results. The results also underpin the ARFIMA's superiority against the GARCH models. In this forecasting exercise, a general empirical conclusion with the ARFIMA model is that it outperforms traditional GARCH models which are based on daily returns. On the other hand, it is inferior to the HAR models, even though both of them use the same volatility measure.

Broadly speaking, it is concluded in this exercise that the GARCH models are inferior to the HAR and ARFIMA models. In this paragraph, we evaluate only the GARCH genre models, regardless of the realized measure models. When we only look at the conditional variance models, the MSE and MAE loss functions draw almost a clear picture. The EGARCH specification has the best performance among its counterparts. However, only one difference reveals here in terms of the rolling and recursive windows techniques. While the rolling windows results of the QLIKE indicate the PGARCH as the best performing model, the recursive QLIKE criteria says that the EGARCH model is the best. The results of the QLIKE criteria seem ambiguous. Nevertheless, we can conclude that the EGARCH model is superior to its counterparts. Thereafter, the results of the loss functions for the other conditional variance models (except EGARCH) does not give useful information to choose the second best and also the worst performing model between the GARCH models.

As it is in the QLIKE and MAE, the realized variance models are expected to be winners with almost a hundred per cent of all the indices. However, 6 out of 30 stock market indices of the MSE results shows that the GARCH family models (especially EGARCH) outperform the realized variance models. The MSE loss function has the most exceptional results (DJI, FTSE, GSPTSE, MXX, OSEAX, SPX) when the realized measures are considered a more precise measure of the true volatility. The number of exceptional results for the QLIKE is two out of thirty (DJI and OMXC20) and the MAE has no exception.

The p-values of the conditional GW test results are reported in Table 8. The null hypothesis is that "both of the models (row and column) statistically have the same accuracy" is tested in terms of squared forecast error. The signs, + and –, show which model outperforms and which model is outperformed. The positive sign (+) indicates the superiority of the column model, whilst the negative one (–) says that the row one outperforms the column model. When we look at the results in terms of the HAR models, most of the results are positive, which confirms the superiority of the HAR specification. However, the test results of some exceptional indices (e.g. DJI, FTSE, GSPTSE, MXX, OSEAX, SPX), which are in favour of the EGARCH model against the HAR models, are insignificant. This means that the column and row models perform equally well, so that we cannot say that whether the row model or column model is superior for those exceptional indices. Indeed, the superiority of the EGARCH model against the HAR for the exceptional indices is likely to be unreliable.

We believe that those exceptions may occur by chance or arise from abnormal data structure because the GW test results says that the column and row models perform equally well for those exceptional results. The exceptional indices mostly consist of the largest stock markets in the world such as the SPX, DJI, GSPTSE, MXX in the continent of America and FTSE, OSEAX, OMXC20 in the Europe. In this context, we need to discuss the largest daily changes of those indices for the post (2007-2008) global financial crisis period. Severe volatility of those markets may cause unpredictable data structure and could make such a best model incapable (and a worst one capable).

According to the GW conditional predictive ability test, it is difficult to say that there is a predictable pattern in the dynamics of those exceptional indices. For instance, $27^{th} April\ 2010$, European sovereign debt crisis resulted in the decline of European stock markets and then spread worldwide. Therefore, the annualized volatility of the European stock markets is higher than the other markets. It lies around 15-25 per cent for the European markets, whilst the other markets are between 10-15 percent in the Appendix Table 4. $6^{th} May\ 2010$ (flash crash), the DJI index dropped almost 1000 points, which is the worst intra-day point loss. $1^{st} August\ 2011$ is another one that the SPX experienced a sharp drop. Those are followed by the stock market crash in China, affecting first Asian stock markets, and 2015-16 stock market sell-off in the US markets. More recent one is the world-wide stock market downturn in 2018. As can be seen in the periods of the financial turbulence, the largest stock markets are first affected ones and therefore is likely to have abnormal data structure. As a consequence, model comparison results would be distorted to some extent.

The decomposition of realized variance into positive and negative realized semivariances adds more information for the prediction of future volatility. This is true and indicated in our results. However, even if the realized semivariance specification (HAR-RSV and HAR-PS) is the winner and improves the out-of-sample forecasts, they could not prevail over the EGARCH model for some exceptional indices (6 out of 30 in MSE; DJI, FTSE, GSPTSE, MXX, OSEAX, SPX). As we discuss it in the previous paragraph, the turbulent times of the aforementioned markets between 2010-2019 could be the reason of abnormal data structure, in turn, this would suggest a good performing model as a bad one and also vice versa.

2.Table 8 **Conditional Giacomini-White test results**

RECURSIVE	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
AEX							
GARCH	0.039 (+)	0.020 (+)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.050 (-)	0.095 (-)	0.084 (-)
ARFIMA-RV	-	-	-	-	-	0.005 (+)	0.002 (+)
HAR-RSV	-	-	-	-	-	-	0.023 (+)
AORD							
GARCH	0.000 (+)	0.000 (-)	0.061 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.082 (-)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.001 (-)	0.000 (+)	0.009 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.001 (+)
HAR-RSV	-	-	-	-	-	-	0.012 (-)
BFX							
GARCH	0.055 (+)	0.012 (+)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.003 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.197 (-)	0.184 (-)	0.181 (-)
ARFIMA-RV	-	-	-	-	-	0.030 (+)	0.046 (+)
HAR-RSV	-	-	-	-	-	-	0.086 (+)
BSESN							
GARCH	0.000 (+)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.097 (+)	0.031 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.183 (+)

RECURSIVE	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
BVLG							
GARCH	0.014 (-)	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.043 (+)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.000 (-)	0.000 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.001 (+)
BVSP							
GARCH	0.000 (+)	0.000 (-)	0.004 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.027 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.129 (-)	0.198 (-)
ARFIMA-RV	-	-	-	-	-	0.005 (+)	0.004 (+)
HAR-RSV	-	-	-	-	-	-	0.096 (+)
DJI							
GARCH	0.032 (+)	0.043 (+)	0.212 (-)	0.076 (-)	0.180 (-)	0.405 (-)	0.340 (-)
EGARCH	-	0.027 (-)	0.097 (-)	0.124 (-)	0.000 (-)	0.043 (-)	0.110 (-)
PGARCH	-	-	0.125 (-)	0.267 (-)	0.009 (-)	0.185 (-)	0.339 (-)
TGARCH	-	-	-	0.062 (-)	0.252 (-)	0.077 (-)	0.062 (-)
HAR-RV	-	-	-	-	0.000 (-)	0.315 (-)	0.436 (-)
ARFIMA-RV	-	-	-	-	-	0.030 (-)	0.002 (+)
HAR-RSV	-	-	-	-	-	-	0.088 (+)
FCHI							
GARCH	0.045 (+)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.035 (-)	0.007 (+)	0.011 (+)
ARFIMA-RV	-	-	-	-	-	0.072 (+)	0.063 (+)
HAR-RSV	-	-	-	-	-	-	0.004 (+)

RECURSIVE	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
FTMIB							
GARCH	0.000 (+)	0.000 (+)	0.026 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.032 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.036 (-)	0.026 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.020 (+)
FTSE							
GARCH	0.011 (+)	0.106 (+)	0.063 (+)	0.443 (-)	0.000 (-)	0.504 (-)	0.532 (-)
EGARCH	-	0.000 (-)	0.000 (-)	0.345 (-)	0.000 (-)	0.124 (-)	0.349 (-)
PGARCH	-	-	0.000 (+)	0.447 (-)	0.001 (-)	0.368 (-)	0.529 (-)
TGARCH	-	-	-	0.356 (-)	0.000 (-)	0.142 (-)	0.364 (-)
HAR-RV	-	-	-	-	0.005 (-)	0.155 (-)	0.521 (-)
ARFIMA-RV	-	-	-	-	-	0.010 (+)	0.003 (-)
HAR-RSV	-	-	-	-	-	-	0.073 (-)
GDAXI							
GARCH	0.085 (+)	0.027 (+)	0.000 (-)	0.005 (+)	0.048 (+)	0.004 (+)	0.006 (+)
EGARCH	-	0.001 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (+)	0.138 (-)	0.067 (+)
ARFIMA-RV	-	-	-	-	-	0.010 (-)	0.000 (-)
HAR-RSV	-	-	-	-	-	-	0.061 (+)
GSPTSE							
GARCH	0.058 (+)	0.000 (-)	0.036 (+)	0.000 (+)	0.224 (-)	0.004 (-)	0.002 (-)
EGARCH	-	0.000 (-)	0.053 (+)	0.025 (-)	0.000 (-)	0.236 (-)	0.190 (-)
PGARCH	-	-	0.007 (+)	0.000 (+)	0.056 (-)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.114 (-)	0.001 (-)	0.354 (-)	0.317 (-)
HAR-RV	-	-	-	-	0.000 (-)	0.529 (-)	0.460 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.092 (+)

RECURSIVE	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
HSI							
GARCH	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.065 (+)	0.005 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.123 (-)
IBEX							
GARCH	0.002 (+)	0.000 (+)	0.002 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.035 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.001 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.205 (-)	0.268 (-)
ARFIMA-RV	-	-	-	-	-	0.020 (+)	0.002 (+)
HAR-RSV	-	-	-	-	-	-	0.174 (+)
IXIC							
GARCH	0.001 (+)	0.004 (+)	0.006 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.007 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.338 (+)	0.168 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.026 (+)
KS11							
GARCH	0.000 (+)	0.000 (+)	0.010 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.001 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.001 (+)	0.434 (-)	0.490 (-)
ARFIMA-RV	-	-	-	-	-	0.443 (-)	0.537 (-)
HAR-RSV	-	-	-	-	-	-	0.554 (+)

RECURSIVE	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
KSE							
GARCH	0.000 (-)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.004 (+)	0.023 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.038 (+)
MXX							
GARCH	0.031 (+)	0.031 (+)	0.002 (-)	0.000 (+)	0.000 (-)	0.002 (-)	0.002 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.014 (-)	0.163 (-)	0.073 (-)	0.068 (-)
PGARCH	-	-	0.012 (-)	0.000 (+)	0.003 (-)	0.001 (-)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.083 (-)	0.495 (-)
ARFIMA-RV	-	-	-	-	-	0.018 (-)	0.003 (+)
HAR-RSV	-	-	-	-	-	-	0.150 (+)
N225							
GARCH	0.000 (+)	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.006 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.100 (-)	0.075 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.023 (+)
NSEI							
GARCH	0.000 (+)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.088 (+)	0.098 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.133 (-)

RECURSIVE	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
OMXC20							
GARCH	0.022 (+)	0.003 (+)	0.001 (-)	0.000 (+)	0.009 (-)	0.000 (-)	0.000 (+)
EGARCH	-	0.008 (-)	0.000 (-)	0.061 (+)	0.001 (-)	0.027 (-)	0.044 (+)
PGARCH	-	-	0.030 (-)	0.000 (+)	0.008 (-)	0.006 (-)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.002 (-)	0.001 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.000 (-)	0.079 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.000 (+)
OMXHPI							
GARCH	0.000 (+)	0.000 (-)	0.001 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.101 (-)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.001 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.018 (+)	0.045 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.000 (-)
OMXSPI							
GARCH	0.078 (+)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.005 (-)	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.001 (+)	0.000 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.017 (+)
OSEAX							
GARCH	0.000 (+)	0.000 (+)	0.096 (+)	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.015 (-)	0.035 (-)	0.071 (-)	0.090 (-)
PGARCH	-	-	0.001 (-)	0.000 (+)	0.046 (-)	0.000 (+)	0.001 (+)
TGARCH	-	-	-	0.000 (+)	0.097 (-)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.024 (-)	0.264 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.181 (+)

RECURSIVE	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
RUT							
GARCH	0.000 (+)	0.004 (+)	0.002 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.001 (-)	0.004 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.020 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.135 (+)	0.126 (+)
ARFIMA-RV	-	-	-	-	-	0.010 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.180 (+)
SMSI							
GARCH	0.000 (+)	0.000 (+)	0.011 (-)	0.000 (+)	0.000 (+)	0.001 (-)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.094 (-)	0.011 (-)
PGARCH	-	-	0.018 (-)	0.000 (+)	0.000 (+)	0.002 (-)	0.000 (-)
TGARCH	-	-	-	0.000 (+)	0.005 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.362 (-)	0.550 (-)
ARFIMA-RV	-	-	-	-	-	0.155 (-)	0.038 (-)
HAR-RSV	-	-	-	-	-	-	0.157 (+)
SPX							
GARCH	0.021 (+)	0.016 (+)	0.110 (+)	0.007 (+)	0.115 (-)	0.022 (+)	0.021 (+)
EGARCH	-	0.020 (-)	0.014 (-)	0.147 (-)	0.000 (-)	0.226 (-)	0.201 (-)
PGARCH	-	-	0.006 (-)	0.024 (-)	0.014 (-)	0.079 (-)	0.042 (-)
TGARCH	-	-	-	0.004 (+)	0.044 (-)	0.002 (+)	0.002 (+)
HAR-RV	-	-	-	-	0.028 (-)	0.092 (-)	0.129 (-)
ARFIMA-RV	-	-	-	-	-	0.050 (+)	0.031 (+)
HAR-RSV	-	-	-	-	-	-	0.001 (+)
SSEC							
GARCH	0.000 (-)	0.000 (-)	0.013 (-)	0.000 (+)	0.000 (-)	0.001 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (-)	0.001 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (-)	0.002 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.001 (-)	0.020 (+)	0.123 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.031 (+)

RECURSIVE	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
SSMI							
GARCH	0.053 (-)	0.100 (+)	0.094 (-)	0.012 (+)	0.026 (+)	0.039 (-)	0.008 (+)
EGARCH	-	0.167 (+)	0.004 (-)	0.194 (+)	0.113 (+)	0.108 (+)	0.107 (+)
PGARCH	-	-	0.078 (-)	0.000 (+)	0.003 (+)	0.100 (-)	0.012 (-)
TGARCH	-	-	-	0.087 (+)	0.080 (+)	0.056 (+)	0.051 (+)
HAR-RV	-	-	-	-	0.084 (-)	0.154 (-)	0.451 (-)
ARFIMA-RV	-	-	-	-	-	0.286 (-)	0.110 (-)
HAR-RSV	-	-	-	-	-	-	0.089 (+)
STOXX50							
GARCH	0.047 (+)	0.020 (+)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (-)	0.000 (-)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.521 (-)	0.594 (-)
ARFIMA-RV	-	-	-	-	-	0.001 (-)	0.000 (-)
HAR-RSV	-	-	-	-	-	-	0.227 (+)

In terms of only the GARCH family models, the asymmetric EGARCH model outperforms the other GARCH models in the vast majority of the indices –except the rolling windows forecasting results of the QLIKE criteria, it indicates the superiority of the PGARCH model–. The reasons behind the success of the EGARCH model could be some important model specifications of it. For instance, the logarithmic form of the model does ensure the non-negativity condition for the conditional variance even if the model parameters are negative. Also, the leverage effect and volatility clustering are captured in the model.

We could not find any significant difference between the rolling and recursive windows forecasting techniques although the ways they work are different from each other. For every single next iteration, the rolling method adds newly observed realized volatility measures or daily returns, and meanwhile it drops the first observation of the initial sample. However, the recursive one makes use of an expanding windows, which the first observation of the initial sample is anchored during the whole forecasting process. Some (Corsi et al., 2008) employ the recursive scheme, whilst some others (Pu, Chen, and Ma, 2016) prefer the rolling one. Vortelinos (2017) uses both of them and found no difference. Clark and McCracken (2008) suggest to combine the rolling and recursive forecasts for improving forecast accuracy. It should be pointed out that the initial sample size of both of these methods highly matters and different initial sample lengths could generate different forecast results on the same data. Indeed, both of the techniques allow to gain a deeper understanding of a model's performance.

Each one of the loss functions has a specific calculation method, causing to produce different results. The QLIKE and MSE are the most popular and frequently used ones in the literature due to being robust to the noisy volatility proxies. In this work, the QLIKE and MAE criteria indicate that the HAR specification is the winner among the competing models in the vast majority of the indices, whereas the MSE has the most exceptional results (6 out of 30 indices, –namely DJI, FTSE, GSPTSE, MXX, OSEAX, SPX– says that the realized variance models are inferior to the conditional variance models). Those exceptions would occur by chance, data structure or perhaps different formulas of the loss functions. In order to investigate whether the forecast errors of two competing models are statistically significant or not, the GW conditional predictive ability test is applied. Using squared forecast error we find equal forecasting performance between the column and row models for the exceptional results. We cannot surely say that the exceptional results of the MSE criteria are reliable.

2.5. Conclusion

This forecasting exercise compares the GARCH, EGARCH, PGARCH, and TGARCH models that use daily data to the models which are derived from high frequency data such as HAR-RV, RSV, and PS and ARFIMA-RV models. In this regard, 30 different stock market indices all over the world between 2010-2019 are included. In doing so, we aim to contribute to the existing literature by carrying out a comprehensive forecasting exercise with 30 different international stock market indices and more up-to-date data.

One-day-ahead out-of-sample volatility forecasts are generated using both the rolling and recursive windows forecasting techniques. The out-of-sample forecast losses are measured using the MSE, MAE, and QLIKE criteria. We also use the conditional GW pairwise test in order to investigate whether the forecast errors of the competing models are statistically significant or not.

One of the main findings of this study indicates that the models which employ the realized variance generate more accurate forecasts compared to the models with the conditional variance, which is consistent with the seminal paper of Andersen and Bollerslev (1998). More specifically, considering the performance of competing models, there is no doubt that the HAR models outperform the ARFIMA and GARCH models. The success of the HAR specification is documented by the vast majority of the indices. Moreover, the decomposition of realized variance into positive and negative realized semivariances (HAR-RSV and HAR-PS) improve the out-of-sample forecasts. However, some exceptional indices (DJI, FTSE, GSPTSE, MXX, OSEAX, SPX) are seen in the results of the MSE and (DJI and OMXC20 in the) QLIKE criteria that the EGARCH model outperforms the HAR models. We believe that those exceptions occur by chance or due to the abnormal data structure because the conditional GW test results report the insignificance of those exceptional results, meaning that both of the competing models perform equally well. In other words, we cannot say that whether the outperforming models for the exceptional indices are actually superior. Following the HAR models, the ARFIMA model also outperforms the GARCH models as it uses a more precise measure of the true volatility, which is 5-min realized variance. It is concluded in this exercise that the GARCH models are inferior to the HAR and ARFIMA models. When we evaluate only the GARCH genre models (regardless of the realized measure models), the EGARCH specification is superior to its counterparts. On the other hand, the worst performing model is unclear. In terms of the results of rolling and recursive methods we could not find any significant difference between them. Nevertheless, the results and findings are specific to market, data frequency, time horizon, and some characteristics of data. This is one of the most important reasons why each contribution in the literature is essential.

In the end, modelling and forecasting financial volatility is always a challenge for academics and practitioners. This chapter emphasizes the significance of high frequency based models for obtaining the most accurate volatility forecasts, in particular HAR-RV model, compared to daily based models. It can be said that the usage of high frequency data compared to daily data at modelling and forecasting financial volatility could enable practitioners to provide more efficient portfolio and risk management strategies by the help of better future volatility forecasts. However, the model selection for the technical and practical processes cannot be generalized even with a large scale of empirical study as the findings could be changing according to different markets, frequencies, time horizons, and specific patterns of volatility. This is one of the most important reasons why each contribution in the literature is essential. Several possible further research ideas based on this chapter would be to investigate considering wider classes of models such as simple and stochastic volatility models (or maybe machine learning and deep learning models, which are quite popular these days) as well as using wider stock market indices. Finally, having documented the superiority of high frequency based volatility models over daily based volatility models, the next two chapters build on the HAR model by adding some additional variables to further explore the model's predictive ability.

CHAPTER 3

Do extreme range estimators improve realized volatility forecasts? Evidence from G7 Stock Markets[8]

Abstract

This chapter investigates whether range estimators contain important information in forecasting future realized volatility. Widely applied range-based estimators are used: Parkinson, Garman- Klass, Roger-Satchell, and Yang-Zhang within a HAR-RV-X framework. Overnight volatility and close-to-close volatility estimators are also included, and the forecasting exercise is applied to G7 stock markets using a rolling window. Using QLIKE, HMSE and MCS forecast criteria, several noteworthy points are reported. The overall findings suggest that while no single model dominates, overnight return volatility achieves the most consistent performance. For example, HAR- RV model forecasts for CAC and DAX indices are improved only by overnight volatility, with some evidence also for SPX. For other indices, forecasts are improved by Parkinson and/or Garman-Klass volatility estimators. Of note, simpler range estimators outperform more complex range estimators. The findings could be important for investors in managing portfolio risk.

[8] A shorter version of this chapter is a published paper in a refereed academic journal. Reference: Korkusuz, B., Kambouroudis, D., & McMillan, D.G. (2023). "Do extreme range estimators improve realized volatility forecasts? Evidence from G7 Stock Markets". *Finance Research Letters, 1544-6123/© 2023 Published by Elsevier Inc*.

3.1. Introduction

The accurate forecast of stock market volatility is of particular importance for investors who have certain levels of risk which they intend not to exceed. In modelling and forecasting stock market volatility, the dominant approach uses the realized volatility (RV, Andersen and Bollerslev, 1998) measure, modelled with the HAR-RV specification of Corsi (2009).[9] The subsequent literature (e.g., Degiannakis and Filis, 2017; Kambouroudis et al., 2021; Peng et al., 2018; Liu et al., 2019; Wang, 2019; Yang and Liu, 2012; Zhang et al., 2020) seeks to improve forecast accuracy by adding exogenous variables in a HAR-RV-X model. These exogenous variables include, for example, cross-market information, implied volatility and EPU (economic policy uncertainty). Further explanatory variables such as the leverage effect, realized semi-variance, jump, and overnight volatility are considered as more directly related to the dependent variable as they capture the stylised facts of volatility. For example, Barndorff-Nielsen et al. (2010) introduce positive and negative realized semi-variance, Corsi and Reno (2012) consider leverage, while Wang et al. (2015) also examine the role of overnight volatility.

An alternative source of information that can improve forecasts is given by OHLC (Open, High, Low, Close) prices. Although RV is the sum of squared intraday returns, it does not necessarily capture high-low range information. Thus, range estimators (e.g., Parkinson, Garman-Klass, Roger-Satchell, and Yang-Zhang) can be introduced to the HAR-RV-X framework. Range estimators (or OHLC estimators) are straightforward to calculate given the availability of open, high, low and close price data. Despite both RV and OHLC being volatility estimators, they capture different frequencies in the data and the information value of OHLC estimators for RV-based forecasts has not been examined. Therefore, this study fills the gap in this literature by investigating whether OHLC estimators that embed range information improve future RV forecasts.

Identifying our work from the existing literature, we can consider, for example, Peng et al. (2018) and Todorova and Husmann (2012). While both focus on the forecasts for a single market, the former incorporates ‘X’ variables into the RV equation and the latter uses the range-based measures as the series to be forecast (i.e., the proxy for volatility). A further paper, Kambouroudis et al. (2021) does include a number of markets in their forecast exercise, but

[9] Andersen et al. (2001) indicate that Realized Variance (RV) is a natural estimator for the integrated variance.

does not include range-based 'X' variables.[10] In recent work that considers range-based forecasts, Petnehazi and Gall (2019) combine with a neural network approach for the Dow Jones Industrial Average, while Wu and Xu (2022) apply the CARR approach to Chinese stock markets. Our study, in essence, combines these approaches by allowing the range-based measures to be the 'X' variables across a selection of markets.[11] In doing so, we believe this chapter addresses a clear gap in the literature.

This chapter is organised as follows: ***Section 2*** gives the literature review. In ***Section 3***, the methods used in this study are introduced. Afterwards, ***Sections 4*** describes the data. In ***Section 5,*** we present the empirical results. Lastly, the conclusion is given in ***Section 6***.

3.2. Literature Review

There are various types of investors in the financial markets. In practice, those can be grouped in terms of their investment horizons as follows: Investor (over week, month or year), day trader (four hour or daily), intra-day trader (15 min, 30 min or 1 hour), and scalper (1 min or 5 min). Even though they are named after differently depending on their investment horizons, they all could be called as "investor or market participants". However, each investors have different objectives, for example, some investors are completely hedgers whilst some others are completely speculators. Therefore, it is unrealistic to expect that each one of investors react identically to the same news. Similarly, but slightly different from practitioners, some researchers divide investors into three different classes depending on their investment horizons such as short-term, middle-term, and long-term investors. In this way, each types of investors could provide a better understanding of how those different market participants react and perceive to the same news. What is explained from the beginning to here is the intuition behind Corsi (2009)'s HAR model. The HAR model is based on capturing different reactions of different investors through a simple autoregressive process, which is in other words an additive cascade model of different volatility components. It can simply be estimated by the ordinary least square (OLS) method and this model is in the form of a simple structure. Related literature

---

[10] Another approach is the CARR (conditional autoregressive range) model of Chou (2005), see Xie (2018) for a recent application.

[11] A further approach, suggested by Christensen and Podolskij (2007) and Martens and van Dijk (2007) also combines the RV and range approaches and does so by calculating the range within intraday intervals before summing to the daily frequency. This requires the availability of intraday interval high and low data as opposed to daily high and low as used here. This would be an interesting avenue to pursue, however, we currently lack availability of the required data.

is quite rich for the applications of HAR model and many studies employ this specification in order to model and forecast the RV.

The most widely used high frequency volatility estimator is known as Realized Volatility (RV) and the RV is best modelled by Corsi (2009)'s HAR-RV specification.[12] Moreover, existing literature (Degiannakis and Filis, 2017; Kambouroudis, McMillan, and Tsakou 2021; Peng et al., 2018; Liu et al., 2019; Wang, 2019; Yang and Liu, 2012) attempt to improve the forecasting accuracy of the HAR-RV model by incorporating a set of exogenous variables in the HAR-RV-X model framework. Those exogenous variables could be obtained from external sources including, for example, cross-market information, Implied Volatility and EPU. Some other explanatory variables such as leverage effect, realized semi-variance, jump, and overnight volatility could also be involved in the model, which are more directly related to dependent variable. Those are usually used to capture stylised facts of volatility. Barndorff-Nielsen, Kinnebrock, and Sheppard (2010) introduce positive and negative realized semivariance measures, which are obtained from the signed high frequency intraday returns. Afterwards, Corsi and Reno (2012), who are the first, add the leverage effect as a component in the HAR-RV model and find that it has a significant impact on the RV for the SPX index. Wang et al. (2015) examine the role of overnight volatility along with the leverage effect for Chinese stock market and point out significant effects of both of them when both are included in the model.

As well as the RV, another volatility estimator is the price range, which is also known as high/low range. The price range is defined as the difference between the highest and lowest prices over an identified time period. Early applications of price range in the area of finance could be seen in the work of Mandelbrot (1971). Afterwards, the Parkinson range-based volatility estimator is introduced in the early 1980s. Parkinson (1980) proposes the scaled high-low price range in order to measure the variability of asset prices. Parkinson (1980) also find the efficiency of Parkinson estimator as roughly 4.9 times higher compared to the efficiency of traditional estimator (that is close-to-close estimator). Extreme range estimators basically assume that the financial prices follow a geometric Brownian motion with two parameters which are the drift (difference between open and close prices) and the volatility. However, Parkinson estimator assumes the zero-drift price process. Following the success of Parkinson

---

[12] Andersen et al. (2001) indicate that Realized Variance (RV) is a natural estimator for the integrated variance.

estimator, some other new range-based estimators have been introduced incorporating with opening and closing prices as well as the scaled high-low price range.

Garman-Klass (1980) attempt to improve the efficiency gain on Parkinson estimator and introduce Garman-Klass (1980) estimator. It is basically a weighted average of the Parkinson volatility estimator and the drift component although they make the same assumption with Parkinson that the price is a zero-drift process. Garman and Klass (1980) claim that their estimator is about 7.4 times more efficient in comparison to the traditional variance based on closing prices. Hereupon, Rogers and Satchell (1991) address the issue of drift more formally contrary to the aforementioned both estimators and suggest their own estimator which is independent of drift process. In other words, their estimator, which is Rogers-Satchell, does not require the assumption of zero-drift and found to be more efficient estimator compared to both of the Parkinson and Garman-Klass estimators. More recently, Yang and Zhang (2000) suggest a new OHLC estimator to overcome the drawbacks of previous volatility estimators. Yang and Zhang (2000) argue that their estimator is the minimum-variance unbiased variance estimator and independent of the drift and opening jumps of the underlying price movements. Many empirical researches (Garman and Klass, 1980; Ball and Torous, 1984; Rogers and Satchell, 1991; Kunitomo, 1992; Yang and Zhang, 2000; Brandt and Diebold, 2006; Brandt and Jones, 2006; Martens and van Dijk, 2007; Chou and Liu, 2011) document that one can employ the high/low range information in order to improve volatility estimation. Moreover, Alizadeh, Brandt, and Diebold (2002) and Brandt and Diebold (2006) indicate that extreme range estimators seem to be robust to microstructure noise, for example, bid-ask bounce. Li and Weinbaum (2001) evidence the efficiency of extreme range estimators for the S&P500 and S&P100 indices. Similar results are also found by Pandey (2002) for the S&P CNX Nifty stock index.

Some researches concentrate on modelling and forecasting of the price-range information. Wang, Hsu, and Liu (2014) evidence that range-based volatility estimators incorporating with the conditional variance of GARCH model does improve the out-of-sample volatility forecasts of Nasdaq100 stock market index. Similarly, Molnar (2016) demonstrates the superior performance of the range-GARCH model over the standard GARCH model using stocks and stock market indices in terms of in-sample and out-of-sample performance. Some others (Miralles-Marcelo, Miralles-Quirós, and Miralles-Quirós, 2013; Raju and Rangaswamy, 2017) also examine the ability of GARCH specification based on the price-range information. Consequently, it is documented that the high-low range information does improve the model estimates and forecasts of the GARCH model. Prior to those studies, Chou (2005) proposes a range based volatility model that uses the price-range information in modelling and forecasting volatility. This is known as the conditional autoregressive range (CARR) model. Chou (2005) and Heng-Chih Chou (2007) employ CARR and GARCH models in order to forecast the out-of-sample volatility of S&P500 index. The findings show that CARR model produces sharper volatility forecasts in comparison with a simple GARCH model. Miralles-Marcelo et al. (2013) examine the CARR model's forecasting performance using the various extreme range estimators of S&P 500 index. They indicate that the Parkinson estimator has the superiority for upward volatilities and trends, whilst the outperformance of the CARR model seems to be in downward trend. However, although the CARR model is a rival to the so-called GARCH model, it has not been paid enough attention due to the popularity of the GARCH model.

The existing literature of extreme range estimators analyse (Molnar, 2012, 2016) and compare the performance of those OHLC-based estimators. Bali and Weinbaum (2005), Li and Hing (2011), Todorova and Husmann (2012), and Jiang et al. (2014) find Garman-Klass (1980) estimator as the optimal range-based estimator. On the other hand, Raju and Rangaswamy (2017) point out the superiority of Yang-Zhang (2000) estimator for in and out-of-sample forecasting performance. Moreover, Petnehazi and Gall (2018) examine the predictability of some extreme range estimators by applying recurrent neural networks on the stock prices of the Dow Jones Industrial Average index and note that changes in the values of extreme range estimators seem to be more predictable compared to those of the close-to-close return based estimators' values. Unlike those empirical studies, Yarovaya et al., (2016) find inconclusive results. All in all, amongst others, Parkinson (1980), Garman-Klass (1980), Rogers-Satchell (1991), and Yang and Zhang (2000) have documented theoretically that range-based volatility estimators are more efficient compared to return-based volatility estimators.

To sum up, close-to-close volatility is the most widely-used traditional method in the literature that considers only the daily closing prices to estimate volatility. However, the current volatility studies extract the volatility (i.e. realized volatility) from high frequency data and intraday high frequency data allow us to estimate daily volatility more precisely. We already point out the importance of intraday information content in the first and second chapters. In addition to inter and intra-day financial data, another kind of data for volatility estimation is daily open, high, and low prices, which are mostly available where closing price is available. In the applications of volatility studies, the fact that the OHLC (open, high, low, close) prices could contain much more valuable information than only closing prices. Accordingly, the daily price range that is the difference between high and low prices is inherently prone to be employed for the volatility estimation. In this regard, some well-established extreme range volatility estimators (e.g. Parkinson, Garman-Klass, Roger-Satchell, and Yang-Zhang), which are derived from the OHLC (open, high, low, close) prices, have been introduced in the literature. Although those estimators have already existed since the early 1980s, academics and practitioners have rarely used them. However, the current literature employing the extreme range estimators began to grow as those range-based estimators can handle the volatility from multiple dimensions.

The OHLC (Open, High, Low, Close) prices could provide another important improvement in helping the RV forecasts. The reason is that high-low range information is not included in the RV even though the RV is the sum of squared intraday returns. This idea brings us to the point that extreme range estimators (e.g. Parkinson, Garman-Klass, Roger-Satchell, and Yang-Zhang) could be utilized for this purpose. Extreme range estimators (or in other words, OHLC estimators) are easy to calculate because of the availability of open, high, low and close prices. Despite the fact that both of RV and OHLC are volatility estimators but in different frequencies, the information value of OHLC estimators for the RV have not been examined yet. Therefore, this study fills the gap in this literature by investigating that question "could OHLC estimators embed this extreme range information into the future RV forecasts?". To the best of our knowledge, this is the first study which examines the information content of extreme range information at improving the forecasts of realized volatility.

3.3. Methodology

3.3.1. HAR-RV-X model

Corsi (2009) proposes the HAR-RV model that is an additive cascade model of different volatility components. The model is specified in Equation 1 as follows:

$$RV_{t+h}^{d} = \beta_0 + \beta_d RV_t^d + \beta_w RV_t^w + \beta_m RV_t^m + \beta_X X_t^d + \varepsilon_{t+h} \quad (1)$$

where RV_t^d is daily volatility component; RV_t^w refers to weekly component, and then RV_t^m indicates monthly component. In the equation, $\beta_X X_t^d$ refers to the exogenous variables (CC, ONV, PK, GK, RS, and YZ) and the RV_t^w and RV_t^m can easily be calculated in Equations 2 and 3 respectively as follows:

$$RV_t^w = \frac{1}{5}(RV_{t-5}^d + RV_{t-4}^d + \cdots + RV_{t-1}^d) \tag{2}$$

$$RV_t^m = \frac{1}{22}(RV_{t-22}^d + RV_{t-21}^d + \cdots + RV_{t-1}^d) \tag{3}$$

The main point of the HAR-RV model is to predict future volatility using three different volatility components; a daily (RV_t^d), a weekly (RV_t^w), and a monthly (RV_t^m) components. The HAR-RV model can simply be estimated by the ordinary least square (OLS) method. The model parameters, RV_t^d, RV_t^w, and RV_t^m theoretically represent short-term, middle-term, and long-term investors respectively and explain the RV via reactions of the different types of investors. In other words, the model coefficients provide an understanding of how these different market participants react and perceive to volatility.

3.3.2. Forecast combination

The forecast combination method is first proposed by Bates and Granger in 1969. It is considered as an attractive technique of forecasting and widely-used in many forecasting researches (Clemen, 1989; Stock and Watson, 2004; Ma, Li, Liu, Zhang, 2017). It simply sums all the individual forecasts and then the sum of individual forecasts is divided by the number of forecasts. In other words, it can be defined as the equal weighted average of the forecasts produced by various models. Therefore, given those forecasts in our work, we have six forecasting models (with the exception of the plain HAR-RV model). The sum of all these forecasts is divided by six in order to generate the combination forecast.

3.3.3. Range Based Volatility Estimators

This section presents the methodological frameworks of range based volatility estimators that are derived from OHLC prices. The estimators we include in this study consist of the close-to-close volatility (4), overnight volatility (5), Parkinson volatility (6), Garman-Klass volatility (7), Roger-Satchell volatility (8), and Yang-Zhang volatility (9).

$$V_{CC,t} = (ln\frac{C_t}{C_{t-1}})^2 \tag{4}$$

Close-to-close volatility estimator or in other words it is known as squared daily returns is one of the most widely used method to measure volatility. Daily squared return is often employed as a volatility proxy for the true volatility. It is simply computed as the above Equation 1, which is the log closing price of today minus the log closing price of yesterday. Close-to-close volatility is commonly applied by the investors who are concerned with long term investment and passive investing. Therefore, long-term investors are usually interested in only closing prices. However, the investors who focus on intraday trading need not only closing prices but also OHLC prices.

$$V_{ONV,t} = (\ln(\frac{O_t}{C_{t-1}}))^2 \tag{5}$$

Equation 5 shows the formula of overnight volatility, which is proposed by Brooks et al. (2000). It requires only today's open and yesterday's close prices to capture accumulated overnight information that could contain important input to improve the persistence of the volatility estimations of indices and so their forecasts (Wang et al., 2015; Kambouroudis, et al., 2021).

$$V_{PK,t} = \frac{1}{4ln2}\left(ln\frac{H_t}{L_t}\right)^2 \tag{6}$$

High-low range can be used to measure the variability of stock market indices. Motivated by high-low range of prices, Parkinson (1980) proposes the scaled high-low price range in order to measure the variability of stock markets, shares or indices. It is found that the efficiency of Parkinson estimator is approximately 4.9 times higher compared to that of the traditional estimators.[13] Extreme range estimators assume that the stock prices follow a geometric Brownian motion with two parameters which are the drift[14] and the volatility. However,

[13] Efficiency of an estimator is defined as follows: $Eff(\widehat{\sigma^2}) \equiv \frac{var(\sigma_s^2)}{var(\widehat{\sigma^2})}$ where σ_s^2 is a simple (traditional) volatility estimator and its efficiency by definition is 1. The denominator, $\widehat{\sigma^2}$, refers to the extreme range estimators.

[14] Drift process means the difference between open and close prices during a trading day.

Parkinson estimator assumes the zero-drift price process. Garman-Klass (1980) in the below Equation (7) suggest a new estimator that incorporates with the drift process.

$$V_{GK,t} = 0.5(ln\frac{H_t}{L_t})^2 - (2\ln 2 - 1)(ln\frac{C_t}{O_t})^2 \tag{7}$$

In doing so, Garman-Klass (1980) attempt to conduct the efficiency gain for range-based estimation contrary to Parkinson estimator. Practically, GK estimator is a weighted average of the Parkinson volatility estimator and the drift (open-to-close squared return) even though they assume same with Parkinson that the price is a zero-drift process. Garman and Klass claim that GK estimator is about 7.4 times more efficient in comparison to the traditional variance based on closing prices. Afterwards, Rogers and Satchell (1991) propose another alternative measure of volatility, which is given by the following equation (8):

$$V_{RS,t} = (ln\frac{H_t}{C_t})(ln\frac{H_t}{O_t}) + (ln\frac{L_t}{C_t})(ln\frac{L_t}{O_t}) \tag{8}$$

Rogers and Satchell (1991) address the issue of drift more formally with their estimator which is independent of drift process. This means that RS estimator does not require the assumption of zero-drift and therefore it is found to be more efficient estimator compared to the estimators in Equations 6 and 7.

$$V_{YZ,t} = (ln\frac{O_t}{C_{t-1}})^2 + k(ln\frac{C_t}{O_t})^2 + (1-k)V_{RS,t} \qquad k = (\frac{0.34}{1.34+\frac{n+1}{n-1}}) \tag{9}$$

Lastly, Yang and Zhang (2000) develop a new volatility estimator for further efficiency gain on previous volatility estimators. They claim about their estimator being the minimum-variance unbiased variance estimator and independent of the drift and opening jumps of the underlying price movements. The related literature already does analyse (Molnar, 2012) and compare the performance of those range-based volatility estimators. GK estimator is found to be as optimal OHLC estimator by Bali and Weinbaum (2005), Li and Hing (2011), Todorova and Husmann (2012), Jiang et al. (2014). Raju and Rangaswamy (2017) suggest the YZ estimator for in and out-of-sample forecasting performance. Furthermore, Yarovaya et al., (2016) find inconclusive results.

3.3.4. Rolling Window Forecasting and Loss Functions

In this study, we first need to decide how to create forecasts, especially the length of historical data to employ for forecasting. In this sense, we produce out-of-sample forecasts using rolling window method as rolling window technique has some advantages compared to recursive window. For instance, when the dynamics and behaviours of the true volatility (chosen proxy) changes over time, the forecasts produced by the models using rolling window method could adapt to this changes faster than those of recursive window method. However, the key in the context of rolling window technique is the choice of window size that can have an impact on forecasting performance of the models, nevertheless, the choice of window size is arbitrary in the literature. However, this is the fact that too little historical data could cause the model being estimated imprecisely and as a result of this, forecasting would not show such a good performance. On the other hand, if the models are estimated using too much historical data, very little out-of-sample data will left for forecasting, which is not enough time span to validate your work empirically. In selecting the window size of rolling method, we follow the works of Ma, Wahab, Liu, and Liu (2018) and Ma, Liu, Huang, and Chen (2017) who also use similar window size. Therefore, the optimal choice of window size could be considered as the 600 observations, which is approximately several years of historical data. In this way, the regression is likely to fit smoothly and a longer out-of-sample period will be obtained. The reason is that the main objective of this work is firstly to evaluate the out-of-sample forecasting performance of the competing models, not to find out an optimal forecasting window size. Moreover, recursive window method is also applied to further robustness of this work.

To evaluate the out-of-sample performance of competing models, we employ two of the most popular loss functions, namely quasi-Gaussian log-likelihood (QLIKE) and heteroskedasticity adjusted mean squared error (HMSE), which are known as robust criterion in the forecasting literature. Equations 10 and 11 give the mathematical formulas of QLIKE and HMSE respectively as follows.

$$QLIKE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}\left[\frac{RV_t}{\widehat{RV}_t} - \log\left(\frac{RV_t}{\widehat{RV}_t}\right) - 1\right] \quad (10)$$

$$HMSE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}\left[1 - \frac{\widehat{RV_t^2}}{RV_t^2}\right]^2 \quad (11)$$

where $\widehat{RV}_t$ denotes the out-of-sample forecasts from the competing models and RV_t is the target volatility for true volatility. τ is the number of out-of-sample forecasting days. The robustness of the QLIKE and MSE is documented in the seminal paper of Patton (2011). Whilst assessing different forecasts for the mean, the MSE criterion could be seen as a natural choice. However, according to Bollerslev et al. (1994) and Bollerslev and Ghysels (1996), the MSE loss function may be less clear in a heteroskedastic environment. For this reason, Bollerslev and Ghysels (1996) suggest the heteroskedasticity-adjusted MSE (HMSE). Given the heteroskedastic feature of financial volatility, the HMSE is the loss function of our choice over the MSE. Additionally, we also use the R-square statistic of Mincer-Zarnowitz (MZ) regression (Mincer and Zarnowitz, 1969). The regression can be expressed as follows.

$$RV_{t+1}^2 = \beta_0 + \beta_1 \widehat{RV_{t+1}^2} + \epsilon_{t+1} \tag{12}$$

The R-square statistic of the MZ regression measures the goodness-of-fit of the forecast series in respect to the target series. In other words, it is revealed that how much of the target volatility is explained by the specific forecast series. One can interpret the R-square statistic of the MZ regression by looking for the highest the R-squares between the competing models' forecasts, which shows the forecast series is good at explaining the target volatility.

3.3.5. Model Confidence Set (MCS) Procedure

Hansen et al. (2011) suggest the well-established MCS procedure to identify the set of superior models by the help of a specific elimination algorithm. At a given level of confidence, the elimination algorithm examines which group of models survive in a set of competing models, as those are defined in terms of a specific loss function without a priori benchmark model. The poorly predictive models are eliminated from the initial set of competing models. There are six different statistics for specifying the set of superior models, the range and semi-quadratic are the most preferred according to suggestion by Hansen et al. (2003). Briefly, the details of the MCS procedure as follows. Let $L_{i,k}$ denote the criterion of model i and $d_{i,j,k} = L_{i,k} - L_{j,k}$ is the differential. The null hypothesis of MCS procedure is $H_{0,M} = E(d_{i,j,k}) = 0, for\ i,j\ \in M$, $M \subset M^0$ and the null is tested against the alternative $H_{1,M} = E(d_{i,j,k}) \neq 0, for\ some\ i,j\ \in M$.

3.4. Data

The data used in this study consists of the stock market indices of G7, group of Seven, namely United States, United Kingdom, Canada, France, Germany, Italy, and Japan. Specifically, daily closing prices, daily range prices (open, high, low, close), and 5-minute realized variance series from June 1, 2009 to October 22, 2021 are used. Those stock market data is obtained from the Oxford-Man Institute's Quantitative Finance Realized Library. The trading days across the G7 stock markets differ from one market to another. Therefore, the data cleaning process have been applied to have a common data sample for the group of Seven. To conduct active trading day synchronization for all the markets, the rows corresponding to the nan-values by any of the markets are omitted.

Table 1a presents the full list of models (and indices) names and abbreviations, whilst Table 1b gives the descriptive statistics of all series used in this study after the date alignment process as noted in the paragraph above, which leads to having approximately 2780 observations for each series. The first 600 observation (2009/06/02 – 2012/01/18) is used for the in-sample estimation and the rest of (2180) observations (2012/01/19 – 2021/10/22) is the out-of-sample to apply the one-step-ahead rolling window forecasting method. The summary statistics are as expected. For example, the means of all the series are close to zero. The negative skewness exists in the table for return series that is often associated with the characteristic of return series. Contrary to return series, range and RV series have positive skewness. The excess kurtosis show that all the series are leptokurtic with higher peak points as well as fatter tails. The Jarque-Bera normality test results indicate that the distribution of each series is non-normal. The Augmented Dickey-Fuller (ADF) test results reject the null hypothesis of a unit root at the 1% significance level, indicating that all the series are stationary. Ljung-Box Q-statistic shows the presence of such correlation up to the fifth order.

Volatility is latent and therefore we need to use a target volatility for the true volatility. There are different approaches to substitute an efficient measure for the actual volatility. The commonly used proxies are extracted from historical data such as the close-to-close (daily squared return) volatility, range based (OHLC) volatilities, and integrated (realized) volatility. Among these proxies, the realized volatility based on high frequency intraday sampling contains more information than the other two methods and therefore considered to be a better representative of the true volatility (Andersen and Bollerslev, 1998; Andersen et al., 1999).

Moreover, Liu, Patton, and Sheppard (2015) compare over 400 different realized measures and suggest that it is difficult to significantly beat five-minutes realized variance. Therefore, this study considers 5-minute realized volatility (RV) of Anderson and Bollerslev (1998) as a target volatility.

3.Table 1a Full list of models (and indices) names and abbreviations

HAR-RV	Heterogeneous Autoregressive Model of RV	**CAC**	France CAC40
HAR-RV-ONV	HAR-RV-Overnight Volatility	**DAX**	Germany DAX30
HAR-RV-CC	HAR-RV-Close-to-Close Volatility	**FTSE**	United Kingdom FTSE100
HAR-RV-PK	HAR-RV-Parkinson	**FTMIB**	Italy FTSE MIB
HAR-RV-GK	HAR-RV-Garman-Klass	**NIKKEI**	Japan NIKKEI 225
HAR-RV-RS	HAR-RV-Rogers-Satchell	**SPTSX**	Canada S&P/TSX
HAR-RV-YZ	HAR-RV-Yang-Zhang	**SPX**	United States S&P500

3.Table 2b Descriptive statistics of G7 data (Return, Range, and 5-min RV)

RETURN	**MEAN**	**STD. DEV.**	**SKEW.**	**EX. KURT.**	**JARQUE-BERA**	**Q(5)**	**ADF**
CAC	0.0002	0.0135	-0.5417	7.8268	7229.12***	19.4829***	-13.49***
GDAXI	0.0003	0.0133	-0.4563	7.9884	7485.72***	16.6816***	-27.25***
FTSE	0.0001	0.0107	-0.8529	8.4088	8524.27***	24.9361***	-13.59***
FTMIB	9.3E-05	0.0165	-1.0084	10.561	13386.5***	12.2867**	-20.67***
NIKKEI	0.0003	0.0138	-0.3026	4.5760	2467.13***	17.0018***	-21.04***
SPTSX	0.0002	0.0096	-1.8155	25.757	78343.5***	41.7380***	-10.93***
SPX	0.0005	0.0110	-0.7790	9.5335	10805.2***	38.5681***	-13.55***
RANGE	**Mean**	**Std. Dev.**	**Skew.**	**Ex. Kurt.**	**Jarque-Bera**	**Q(5)**	**ADF**
CAC	0.0142	0.0089	2.4706	10.741	16192.9***	3733.55***	-8.57***
GDAXI	0.0143	0.0089	2.4779	12.002	19529.4***	3596.39***	-6.56***
FTSE	0.0126	0.0083	3.5583	24.289	74203.1***	3931.79***	-8.16***
FTMIB	0.0180	0.0113	3.5887	28.654	101073***	3382.78***	-8.61***
NIKKEI	0.0116	0.0084	5.1205	57.640	396984***	1871.18***	-8.27***
SPTSX	0.0097	0.0070	3.9901	31.495	122277***	4750.06***	-5.96***
SPX	0.0111	0.0081	2.8837	14.587	28501.5***	4607.59***	-6.05***
5-MIN RV	**Mean**	**Std. Dev.**	**Skew.**	**Ex. Kurt.**	**Jarque-Bera**	**Q(5)**	**ADF**
CAC	0.0001	0.00017	9.9869	166.65	3.26E+06***	4818.41***	-8.10***
GDAXI	0.0001	0.00016	7.7792	91.942	1.007E+06***	5225.84***	-8.77***
FTSE	9.6E-05	0.00023	16.637	392.73	1.79E+07***	2035.20***	-10.3***
FTMIB	0.0001	0.00016	6.5410	66.618	533886***	4674.88***	-6.32***
NIKKEI	7.5E-05	0.00015	11.574	197.14	4.56E+06***	1660.67***	-9.55***
SPTSX	5.5E-05	0.00013	14.468	308.73	1.11E+07***	2497.40***	-10.4***
SPX	7.9E-05	0.00019	10.893	167.64	3.31E+06***	4015.42***	-9.03***

Note: *Asterisk *, **, and *** denote rejections of null hypothesis at 10%, 5%, and 1% significance levels, respectively. The null hypothesis of the third and fourth moments are "Skewness = 0" and "Excess Kurtosis = 3".*

3.5. Results

3.5.1. In-sample evaluation

Table 2 reports the full sample estimation results of the seven HAR-RV-X type models in this study. β_0, β_d, β_w, $and\ \beta_m$ are the model coefficients of plain HAR-RV model, which is explained in the methodology section in more detail. β_X represents the various exogenous variables (extreme range estimators) that is embedded in the HAR-RV model. R^2 is the goodness-of-fit of various HAR-RV-X models for the full sample analysis. This study aims to evaluate the out-of-sample performance of the various forecasting models. Furthermore, the performance in the in-sample for a model does not necessarily have an impact on the out-of-sample performance of the same model. Therefore, this section presents a short in-sample evaluation for the competing models.

First of all, the original parameters of the HAR-RV model such as the daily and weekly coefficients mostly seem to be significant for the G7 markets except the daily coefficients of the FTSE and GSPTSE indices which are not significant. In terms of the monthly component, the coefficients are insignificant, with two notable exceptions which are the FTSEMIB and NIKKEI indices (which are significant). Following that most of the extreme range estimators that are embedded as exogenous variables into the plain HAR-RV model do improve the in-sample accuracy of the HAR-RV model for the group of Seven so that the impact of these exogenous variables on the performance of the HAR-RV model cannot be neglected. This can be better understood by comparing the adjusted R-square value of the baseline model with the adjusted R-squares of the models with an additional variable in Table 2. In short, the HAR-RV-PK can be considered as a promising model in the in-sample evaluation. Amongst the other extreme range estimators, the Parkinson estimator when added to the baseline model (in the case of five indices; CAC, DAX, FTSE, FTMIB and SPX) strengths the model's overall goodness-of-fit much better compared to the other estimators that are used as an additional variable. For other two indices, a moderate improvement can be seen for the SPTSX but the HAR-RV-PK model does not improve the in-sample accuracy for only the NIKKEI.

3.Table 3 full sample estimation results of HAR-RV-X models

Coef. → **↓ Model spec.**	β_0	β_d	β_w	β_m	β_X	$Adj.R^2$
$HAR\text{-}RV_{CAC}$	1.51E-05 (4.48E-06) ***	0.506 (0.091) ***	0.298 (0.061) ***	0.051 (0.049)	–	0.56
$HAR\text{-}RV\text{-}ONV_{CAC}$	1.44E-05 (4.33E-06) ***	0.482 (0.095) ***	0.268 (0.067) ***	0.033 (0.047)	0.118 (0.031) ***	0.59
$HAR\text{-}RV\text{-}CC_{CAC}$	1.49E-05 (4.15E-06) ***.	0.451 (0.090) ***	0.263 (0.079) ***	0.054 (0.046)	0.051 (0.013) ***	0.58
$HAR\text{-}RV\text{-}PK_{CAC}$	1.20E-05 (4.70E-06) **	0.221 (0.084) ***	0.280 (0.072) ***	0.062 (0.054)	0.339 (0.148) **	0.59
$HAR\text{-}RV\text{-}GK_{CAC}$	1.45E-05 (4.56E-06) ***	0.301 (0.142) **	0.303 (0.066) ***	0.060 (0.049)	0.212 (0.173)	0.57
$HAR\text{-}RV\text{-}RS_{CAC}$	1.49e-05 (4.59e-06) ***	0.544 (0.138) ***	0.296 (0.061) ***	0.048 (0.048)	-0.031 (0.078)	0.56
$HAR\text{-}RV\text{-}YZ_{CAC}$	1.55e-05 (4.42e-06) ***	0.408 (0.082) ***	0.299 (0.067) ***	0.049 (0.051)	0.058 (0.050)	0.57
$HAR\text{-}RV_{DAX}$	1.35E-05 (3.57E-06) ***	0.436 (0.092) ***	0.378 (0.104) ***	0.053 (0.062)	–	0.56
$HAR\text{-}RV\text{-}ONV_{DAX}$	1.36E-05 (3.64E-06) ***	0.427 (0.092) ***	0.353 (0.091) ***	0.041 (0.060)	0.066 (0.023) ***	0.58
$HAR\text{-}RV\text{-}CC_{DAX}$	1.37E-05 (3.54E-06) ***	0.399 (0.091) ***	0.353 (0.106) ***	0.052 (0.059)	0.035 (0.008) ***	0.58
$HAR\text{-}RV\text{-}PK_{DAX}$	1.26E-05 (3.62E-06) ***	0.110 (0.094)	0.358 (0.108) ***	0.075 (0.064)	0.337 (0.105) ***	0.60
$HAR\text{-}RV\text{-}GK_{DAX}$	1.48E-05 (4.03E-06) ***	0.107 (0.092)	0.363 (0.100) ***	0.081 (0.063)	0.309 (0.067) ***	0.59
$HAR\text{-}RV\text{-}RS_{DAX}$	1.53E-05 (3.91E-06) ***	0.258 (0.110) **	0.373 (0.104) ***	0.077 (0.064)	0.142 (0.042) ***	0.58
$HAR\text{-}RV\text{-}YZ_{DAX}$	1.42E-05 (3.82E-06) ***	0.369 (0.102) ***	0.375 (0.103) ***	0.051 (0.062)	0.039 (0.038)	0.57

(Continued)

Coef. → ↓ ***Model spec.***	β_0	β_d	β_w	β_m	β_X	$Adj.R^2$
$HAR\text{-}RV_{FTSE}$	*2.22E-05* *(4.81E-06)* ***	*0.066* *(0.060)*	*0.660* *(0.177)* ***	*0.041* *(0.101)*	–	*0.28*
$HAR\text{-}RV\text{-}ONV_{FTSE}$	*2.14E-05* *(4.55E-06)* ***	*0.065* *(0.059)*	*0.646* *(0.177)* ***	*0.045* *(0.099)*	*0.110* *(0.075)*	*0.29*
$HAR\text{-}RV\text{-}CC_{FTSE}$	*2.10E-05* *(4.37E-06)* ***	*-0.013* *(0.075)*	*0.603* *(0.169)* ***	*0.050* *(0.092)*	*0.116* *(0.026)* ***	*0.31*
$HAR\text{-}RV\text{-}PK_{FTSE}$	*2.13E-05* *(4.85E-06)* ***	*-0.147* *(0.085)* *	*0.486* *(0.125)* ***	*0.068* *(0.087)*	*0.430* *(0.111)* ***	*0.32*
$HAR\text{-}RV\text{-}GK_{FTSE}$	*2.07E-05* *(5.04E-06)* ***	*-0.026* *(0.051)*	*0.601* *(0.159)* ***	*0.051* *(0.098)*	*0.204* *(0.058)* ***	*0.29*
$HAR\text{-}RV\text{-}RS_{FTSE}$	*2.22e-05* *(4.81e-06)* ***	*0.070* *(0.047)*	*0.662* *(0.189)* ***	*0.040* *(0.104)*	*-0.007* *(0.060)*	*0.29*
$HAR\text{-}RV\text{-}YZ_{FTSE}$	*2.10e-05* *(4.80e-06)* ***	*0.0002* *(0.050)*	*0.609* *(0.165)* ***	*0.054* *(0.099)*	*0.116* *(0.037)* ***	*0.29*
$HAR\text{-}RV_{FTMIB}$	*1.50E-05* *(3.95E-06)* ***	*0.432* *(0.085)* ***	*0.342* *(0.067)* ***	*0.093* *(0.052)* *	–	*0.53*
$HAR\text{-}RV\text{-}ONV_{FTMIB}$	*1.40E-05* *(3.78E-06)* ***	*0.418* *(0.085)* ***	*0.343* *(0.067)* ***	*0.076* *(0.051)*	*0.057* *(0.020)* ***	*0.54*
$HAR\text{-}RV\text{-}CC_{FTMIB}$	*1.49E-05* *(3.83E-06)* ***	*0.314* *(0.067)* ***	*0.340* *(0.068)* ***	*0.112* *(0.050)* **	*0.043* *(0.007)* ***	*0.58*
$HAR\text{-}RV\text{-}PK_{FTMIB}$	*1.59E-05* *(4.00E-06)* ***	*0.161* *(0.079)* **	*0.349* *(0.078)* ***	*0.143* *(0.050)* ***	*0.145* *(0.027)* ***	*0.58*
$HAR\text{-}RV\text{-}GK_{FTMIB}$	*1.56E-05* *(4.10E-06)* ***	*0.291* *(0.116)* **	*0.344* *(0.067)* ***	*0.114* *(0.053)* **	*0.086* *(0.044)* *	*0.54*
$HAR\text{-}RV\text{-}RS_{FTMIB}$	*1.46E-05* *(4.05E-06)* ***	*0.475* *(0.124)* ***	*0.340* *(0.069)* ***	*0.087* *(0.056)*	*-0.024* *(0.025)*	*0.54*
$HAR\text{-}RV\text{-}YZ_{FTMIB}$	*1.50E-05* *(3.96E-06)* ***	*0.274* *(0.094)* ***	*0.347* *(0.065)* ***	*0.098* *(0.050)* *	*0.074* *(0.028)* ***	*0.55*

(Continued)

Coef. → ↓ *Model spec.*	β_0	β_d	β_w	β_m	β_X	$Adj.R^2$
$HAR\text{-}RV_{NIKKEI}$	2.01E-05 (3.24E-06) ***	0.347 (0.073) ***	0.263 (0.081) ***	0.122 (0.051) **	–	0.28
$HAR\text{-}RV\text{-}ONV_{NIKKEI}$	1.45E-05 (3.40E-06) ***	0.289 (0.059) ***	0.231 (0.094) **	0.138 (0.058) **	0.134 (0.045) ***	0.30
$HAR\text{-}RV\text{-}CC_{NIKKEI}$	1.81E-05 (3.40E-06) ***.	0.233 (0.058) ***	0.242 (0.070) ***	0.123 (0.054) **	0.062 (0.015) ***	0.30
$HAR\text{-}RV\text{-}PK_{NIKKEI}$	2.02E-05 (3.17E-06) ***	0.314 (0.176) *	0.269 (0.091) ***	0.125 (0.053) **	0.023 (0.092)	0.28
$HAR\text{-}RV\text{-}GK_{NIKKEI}$	2.00E-05 (3.13E-06) ***	0.360 (0.142) **	0.261 (0.092) ***	0.121 (0.053) **	-0.008 (0.049)	0.28
$HAR\text{-}RV\text{-}RS_{NIKKEI}$	2.00e-05 (3.12e-06) ***	0.372 (0.128) ***	0.257 (0.091) ***	0.118 (0.053) **	-0.015 (0.029)	0.28
$HAR\text{-}RV\text{-}YZ_{NIKKEI}$	1.88e-05 (3.65e-06) ***	0.264 (0.079) ***	0.249 (0.080) ***	0.142 (0.058) **	0.046 (0.020) **	0.29
$HAR\text{-}RV_{SPTSX}$	1.26E-05 (2.65E-06) ***	0.091 (0.060)	0.637 (0.166) ***	0.034 (0.090)	–	0.31
$HAR\text{-}RV\text{-}ONV_{SPTSX}$	1.49E-05 (2.67E-06) ***	0.042 (0.055)	0.434 (0.110) ***	0.101 (0.071)	0.234 (0.078) ***	0.41
$HAR\text{-}RV\text{-}CC_{SPTSX}$	1.48E-05 (2.67E-06) ***	-0.039 (0.100)	0.591 (0.178) ***	0.075 (0.079)	0.055 (0.010) ***	0.33
$HAR\text{-}RV\text{-}PK_{SPTSX}$	1.35E-05 (2.63E-06) ***	-0.011 (0.057)	0.444 (0.189) **	0.089 (0.087)	0.236 (0.107) **	0.33
$HAR\text{-}RV\text{-}GK_{SPTSX}$	1.31E-05 (2.69E-06) ***	0.051 (0.053)	0.504 (0.199) **	0.068 (0.094)	0.136 (0.090)	0.32
$HAR\text{-}RV\text{-}RS_{SPTSX}$	1.28E-05 (2.68E-06) ***	0.069 (0.050)	0.549 (0.200) ***	0.056 (0.096)	0.084 (0.069)	0.32
$HAR\text{-}RV\text{-}YZ_{SPTSX}$	1.41E-05 (2.73E-06) ***	0.026 (0.046)	0.510 (0.164) ***	0.074 (0.083)	0.077 (0.027) ***	0.33

(Continued)

Coef. → ↓ *Model spec.*	β_0	β_d	β_w	β_m	β_X	$Adj.R^2$
$HAR\text{-}RV_{SPX}$	1.49E-05 (4.12E-06) ***	0.339 (0.079) ***	0.452 (0.136) ***	0.018 (0.053)	–	0.46
$HAR\text{-}RV\text{-}ONV_{SPX}$	1.69E-05 (3.83E-06) ***	0.318 (0.077) ***	0.410 (0.113) ***	0.025 (0.048)	0.095 (0.030) ***	0.47
$HAR\text{-}RV\text{-}CC_{SPX}$	1.63E-05 (3.60E-06) ***.	0.141 (0.104)	0.456 (0.134) ***	-0.011 (0.059)	0.133 (0.027) ***	0.50
$HAR\text{-}RV\text{-}PK_{SPX}$	1.11E-05 (3.95E-06) ***	0.099 (0.077)	0.361 (0.122) ***	0.007 (0.050)	0.452 (0.140) ***	0.49
$HAR\text{-}RV\text{-}GK_{SPX}$	1.38E-05 (4.10E-06) ***	0.253 (0.085) ***	0.403 (0.125) ***	0.017 (0.049)	0.190 (0.082) **	0.47
$HAR\text{-}RV\text{-}RS_{SPX}$	1.45e-05 (4.07e-06) ***	0.305 (0.084) ***	0.433 (0.130) ***	0.019 (0.052)	0.071 (0.038) *	0.46
$HAR\text{-}RV\text{-}YZ_{SPX}$	1.65e-05 (4.09e-06) ***	0.272 (0.074) ***	0.390 (0.122) ***	0.020 (0.048)	0.092 (0.020) ***	0.48

Notes: *$HAR\text{-}RV\text{-}X_{INDEX}$ is the model specification; "HAR-RV" is the plain Heterogenous Autoregressive Model of Realized Volatility. "X" symbolizes the exogenous variables, which are extreme range estimators in our case and "INDEX" indicates the specific stock market in the group of Seven. Robust standard errors in parentheses and the stars indicate the significance levels for p values; * (%10), ** (%5), *** (%1).*

Having evaluated the Parkinson estimator's performance, the other estimators give a mixed results depending on the various indices. If the results are generalized, it could be said that the second and third best additional variables to the HAR-RV model alternate between ONV and CC (either ONV or CC) depending on the different stock market indices in Table 2. Both estimators show the same performance for the indices of DAX and NIKKEI stock markets. However, the HAR-RV-ONV exhibits better goodness-of-fit for SPTSX and CAC, whilst the HAR-RV-CC's adjusted R-squares are significant and higher than the others (except HAR-RV-PK) for FTSE, FTMIB, and SPX. To sum up, except for Parkinson estimator's additional information to the HAR-RV model, the in-sample results are inconclusive. Hereupon, instead of going deeper into the in-sample analysis that does not necessarily have an impact on the out-of-sample performance of the competing models, this work aims to concentrate on the competing models' out-of-sample performance as it is pointed out in the beginning of this section.

3.5.2. Out-of-sample evaluation

This study examines whether extreme range estimators or in other words OHLC estimators contain important information for forecasting the future realized volatility. In this regards, we focus on some well-known estimators in the related literature, including, for example, Close-to-close, Overnight Volatility, Parkinson, Garman-Klass, Roger-Satchell, and Yang-Zhang to examine their potential for generating better forecasting accuracy both in the context of model-based approach (HAR-RV-X model) and model-free approach. Using one-step-ahead rolling window technique, we evaluate the forecast results of the seven competing models to investigate which of these forecasts is closer to the actual volatility that is 5-min realized volatility in our case. To do this, several well established loss functions, namely, QLIKE and HMSE are considered as both of them are pointed out as robust criterion in the literature (Patton, 2011). The MCS procedure that identifies the set of the best models is employed to support further those empirical results. We also repeat this forecasting exercise under recursive window technique in order to underpin our results.[15]

3.5.2.1. Evaluation of results with model-free approach

The model-free approach used in this study considers the extreme range estimators themselves as forecasters for the realized volatility. The first four rows of each market indices in Table 2 are the model-free approach for forecasting the realized volatility, which are Parkinson, Garman-Klass, Rogers-Satchell, and Yang-Zhang estimators, respectively. When we compare the model-free approach to the model-based approach, the model-free approach (considering the estimators themselves as forecasters) significantly outperforms the HAR-RV-X model's forecasts in Table 2. Broadly speaking, the Garman-Klass estimator based on model-free approach is the best-performing forecaster for the majority of indices. The second best-performing forecaster might be considered as the Parkinson estimator, whilst the results of the Rogers-Satchell and Yang-Zhang estimators are a little mixed and mostly inferior to the Parkinson and Rogers-Satchell. An interesting point can be made as follows.[16] While traditional range estimators with simple formulas such as Parkinson and Garman-Klass have potential to

[15] Even if we point out some advantages of rolling window technique, we applied recursive window method for further robustness check.
[16] We already emphasized this point earlier but here it is in terms of the model-free approach.

help in forecasting the realized volatility, the extreme range estimators derived from more complex formulas, namely Roger-Satchell and Yang-Zhang hardly contains new information for forecasting realized volatility in the group of Seven.

3.5.2.2. Evaluation of results with model-based approach

According to Table 2, the realized volatility forecasts of CAC and DAX stock indices are improved by only ONV information while extreme range estimators (Parkinson, Garman-Klass, Roger-Satchell, and Yang-Zhang) do not add any extra information onto the benchmark HAR-RV model. However, the forecast combination of extreme range estimators does work for the CAC and DAX indices. These results draw by the QLIKE and HMSE criteria. Contrary to these results, MZ R-squares show that the forecast of CAC index can be improved by the HAR-RV-GK specification, while the best-performing model of DAX seems to be HAR-RV-PK. When we look at the results of FTSE index, the picture is different from the other indices but all the three loss functions, namely QLIKE, HMSE, and MZ R-square values give the same results as follows. ONV does not improve the forecasting accuracy of the HAR-RV model. This result is consistent with the findings in the work of Kambouroudis et al. (2020) as they point out that FTSE is only exception among other indices which ONV does not improve the forecast performance of the HAR-RV model. On the other hand, unlike CAC and DAX indices, realized volatility forecasts of FTSE index are improved best by the Parkinson range estimator (HAR-RV-PK). However, the inclusion of more sophisticated extreme range based estimators such as Roger-Satchell and Yang-Zhang do not improve the baseline HAR-RV model of FTSE index.

In the case of FTMIB index, the loss functions used in this study produce different results from each other but the common result can be summarized such a way that HAR-RV-ONV is such a promising model for the criterion so that ONV could be considered as an important variable for the FTMIB index. The difference between different criteria is that QLIKE and MZ R-square advise the HAR-RV-CC and HAR-RV-COMB as the best-performing models respectively, whilst the extreme range estimators are outperformed by the benchmark model. On the other hand, the result of HMSE criteria for forecasting FTSEMIB realized volatility select some of the extreme range estimators as the best forecasting models such as HAR-RV-GK, HAR-RV-ONV, and HAR-RV-YZ respectively, which produces empirical counterevidence against the result of QLIKE and MZ R-square criterion.

Afterwards, the results of NIKKEI index indicate that HAR-RV-CC model outperforms the others on the behalf of all the loss functions. However, rest of NIKKEI results' ranking is inconclusive in terms of the choice of the loss functions. For example, QLIKE is in favour of the plain HAR-RV model, meaning that extreme range estimators does not contain valuable information for forecasting NIKKEI realized volatility whereas HMSE shows that only HAR-RV-CC, HAR-RV-COMB and HAR-RV-YZ models improve the forecasting accuracy compared to the benchmark forecasting model. In addition, MZ R-squares for the NIKKEI present similar results with the HMSE values of NIKKEI index.

Following the NIKKEI index, the best-performing forecasting models of SPTSX index are also subject to the choice of loss functions. QLIKE suggests HAR-RV-ONV as the best-performing forecasting model, followed by HAR-RV-COMB and HAR-RV-PK, and HAR-RV-YZ models respectively. On the other hand, the rankings of the best-performing models measured by HMSE are as follows: HAR-RV-PK, HAR-RV-GK, and HAR-RV-COMB respectively. The MZ R-square results of NIKKEI index show that no model improve the baseline model. Here, we need to point out that considering HMSE criteria HAR-RV-ONV adds very little value on the benchmark model, which makes it the worst-performing model compared to the models with extreme range estimators, (while QLIKE selects the same model, HAR-RV-ONV, as the best forecasting model). The overall result for SPTSX is that the benchmark HAR-RV model could be improved best by the Parkinson volatility according to both of the criterion.

Lastly, when the results of SPX index are evaluated by the choice of our loss functions, the findings of best-performing forecasting model on the HMSE and MZ R-square sides contradict with the results of QLIKE. QLIKE is in the favour of HAR-RV-COMB and HAR-RV-GK model. In the same rankings, QLIKE suggests HAR-RV-PK as the second worst-performing forecasting model. However, HMSE and MZ R-square select HAR-RV-PK by far the best model compared to the rest of its counterparts. Likewise, HAR-RV-ONV measured by QLIKE is the fourth best model but when it is measured by HMSE, it seems to be worst performing model.

3.Table 4 *ROLLING WINDOW FORECASTING RESULTS EVALUATED BY QLIKE AND HMSE AND MCS PROCEDURE*

CAC	QLIKE	p-value	Rank	$MZ\text{-}R^2$	CAC	HMSE	p-value	Rank
PK (Model-free)	0.2360	eliminated	–	0.721	PK (Model-free)	0.3508	1.0000	3
GK (Model-free)	0.1947	eliminated	–	**0.793**	GK (Model-free)	**0.2168**	**1.0000**	**1**
RS (Model-free)	0.6907	eliminated	–	0.788	RS (Model-free)	0.2741	1.0000	2
YZ (Model-free)	0.1798	eliminated	–	0.560	YZ (Model-free)	3.6710	eliminated	–
HAR-RV	0.1675	eliminated	–	0.477	HAR-RV	0.7861	0.3078	7
HAR-RV-ONV	**0.1517**	**1.0000**	**1**	0.476	HAR-RV-ONV	**0.7257**	**1.0000**	**4**
HAR-RV-CC	0.1688	eliminated	–	0.466	HAR-RV-CC	0.8068	0.0782	10
HAR-RV-PK	0.1766	eliminated	–	0.482	HAR-RV-PK	0.8008	0.1878	9
HAR-RV-GK	0.1724	eliminated	–	**0.494**	HAR-RV-GK	0.8181	0.1964	8
HAR-RV-RS	0.1700	eliminated	–	0.444	HAR-RV-RS	0.7893	1.0000	6
HAR-RV-YZ	0.1817	eliminated	–	0.493	HAR-RV-YZ	0.8534	0.0378	11
HAR-RV-COMB.	0.1599	1.0000	2	0.487	HAR-RV-COMB.	0.7622	1.0000	5
DAX	QLIKE	p-value	Rank	$MZ\text{-}R^2$	DAX	HMSE	p-value	Rank
PK (Model-free)	0.2141	eliminated	–	0.731	PK (Model-free)	0.3707	1.0000	3
GK (Model-free)	**0.1620**	**1.0000**	**1**	**0.742**	GK (Model-free)	**0.2238**	**1.0000**	**1**
RS (Model-free)	0.4212	eliminated	–	0.715	RS (Model-free)	0.2757	1.0000	2
YZ (Model-free)	0.1963	eliminated	–	0.430	YZ (Model-free)	4.7830	eliminated	–
HAR-RV	0.1756	eliminated	–	0.462	HAR-RV	0.9157	eliminated	–
HAR-RV-ONV	**0.1657**	**1.0000**	**2**	0.464	HAR-RV-ONV	**0.8607**	**1.0000**	**4**
HAR-RV-CC	0.1771	eliminated	–	0.456	HAR-RV-CC	0.9355	eliminated	–
HAR-RV-PK	0.1830	eliminated	–	**0.496**	HAR-RV-PK	0.8990	eliminated	–
HAR-RV-GK	0.1849	eliminated	–	0.469	HAR-RV-GK	0.9134	eliminated	–
HAR-RV-RS	0.1790	eliminated	–	0.436	HAR-RV-RS	0.9216	eliminated	–
HAR-RV-YZ	0.1810	eliminated	–	0.475	HAR-RV-YZ	0.9602	eliminated	–
HAR-RV-COMB.	0.1720	eliminated	–	0.480	HAR-RV-COMB.	0.8967	eliminated	–
FTSE	QLIKE	p-value	Rank	$MZ\text{-}R^2$	FTSE	HMSE	p-value	Rank
PK (Model-free)	**0.2464**	**1.0000**	**1**	0.708	PK (Model-free)	0.4672	1.0000	2
GK (Model-free)	0.3033	eliminated	–	**0.760**	GK (Model-free)	**0.4543**	**1.0000**	**1**
RS (Model-free)	3.9910	eliminated	–	0.585	RS (Model-free)	0.6183	1.0000	3
YZ (Model-free)	0.2728	1.0000	2	0.627	YZ (Model-free)	1.8960	1.0000	4
HAR-RV	0.2851	0.7680	8	0.186	HAR-RV	2.8416	eliminated	–
HAR-RV-ONV	0.3140	eliminated	–	0.184	HAR-RV-ONV	2.8504	eliminated	–
HAR-RV-CC	0.2816	1.0000	4	0.221	HAR-RV-CC	2.7036	eliminated	–
HAR-RV-PK	**0.2806**	**1.0000**	**3**	**0.247**	HAR-RV-PK	**2.5594**	**1.0000**	**5**
HAR-RV-GK	0.2887	1.0000	6	0.195	HAR-RV-GK	2.6977	eliminated	–
HAR-RV-RS	0.2892	0.4778	9	0.140	HAR-RV-RS	2.9010	eliminated	–
HAR-RV-YZ	0.2860	1.0000	7	0.184	HAR-RV-YZ	2.8410	eliminated	–
HAR-RV-COMB.	0.2822	1.0000	5	0.212	HAR-RV-COMB.	2.7220	eliminated	–
FTMIB	QLIKE	p-value	Rank	$MZ\text{-}R^2$	FTMIB	HMSE	p-value	Rank
PK (Model-free)	0.1537	eliminated	–	0.572	PK (Model-free)	2.1890	eliminated	–
GK (Model-free)	**0.1243**	**1.0000**	**1**	**0.711**	GK (Model-free)	1.4530	1.0000	2
RS (Model-free)	0.8281	eliminated	–	0.511	RS (Model-free)	**1.3870**	**1.0000**	**1**
YZ (Model-free)	0.2326	eliminated	–	0.600	YZ (Model-free)	8.7150	eliminated	–
HAR-RV	0.1509	0.5504	6	0.484	HAR-RV	1.6260	0.5534	9
HAR-RV-ONV	0.1492	1.0000	4	0.481	HAR-RV-ONV	1.5580	0.4120	4
HAR-RV-CC	0.1471	1.0000	3	**0.534**	HAR-RV-CC	1.5920	1.0000	7
HAR-RV-PK	0.1518	0.1598	8	0.501	HAR-RV-PK	1.6090	eliminated	–
HAR-RV-GK	0.1519	1.0000	5	0.458	HAR-RV-GK	**1.4940**	**1.0000**	**3**
HAR-RV-RS	0.1583	eliminated	–	0.453	HAR-RV-RS	1.6650	1.0000	8
HAR-RV-YZ	0.1504	0.3996	7	0.461	HAR-RV-YZ	1.5570	0.2766	5
HAR-RV-COMB.	**0.1470**	**1.0000**	**2**	0.503	HAR-RV-COMB.	1.5630	1.0000	6

Note: *Loss functions and MCS procedure results are merged in this table. QLIKE and HMSE values are obtained by one-step-ahead rolling window forecasting method. MZ-R square stands for Mincer-Zarnowitz regression's R squares. Lower the values of QLIKE and HMSE is better while higher the value of MZ R-square is better for the comparison of forecasts. Window size is 600 observations that is used as in-sample estimation. The out-of-sample consists of 2180 observations. P-value and rank results are received from the MCS procedure. Bold numbers show the best-performing models for each indices.*

(Continued)

NIKKEI	QLIKE	p-value	Rank	$MZ\text{-}R^2$	NIKKEI	HMSE	p-value	Rank
PK (Model-free)	0.2079	1.0000	2	0.783	PK (Model-free)	0.2869	1.0000	3
GK (Model-free)	**0.1806**	**1.0000**	**1**	**0.828**	GK (Model-free)	**0.1914**	**1.0000**	**1**
RS (Model-free)	1.0070	eliminated	–	0.743	RS (Model-free)	0.2583	1.0000	2
YZ (Model-free)	0.3027	eliminated	8	0.592	YZ (Model-free)	10.440	eliminated	–
HAR-RV	0.2863	0.9072	5	0.250	HAR-RV	1.6990	0.1100	10
HAR-RV-ONV	0.2903	0.7524	6	0.244	HAR-RV-ONV	1.7850	0.0312	11
HAR-RV-CC	**0.2835**	**1.0000**	**3**	0.259	HAR-RV-CC	**1.6020**	**1.0000**	**4**
HAR-RV-PK	0.5478	eliminated	–	0.234	HAR-RV-PK	1.7560	1.0000	8
HAR-RV-GK	0.4080	eliminated	–	0.250	HAR-RV-GK	1.7240	1.0000	7
HAR-RV-RS	0.3754	eliminated	–	0.254	HAR-RV-RS	1.7870	1.0000	9
HAR-RV-YZ	0.2877	0.6694	7	0.246	HAR-RV-YZ	1.6280	1.0000	6
HAR-RV-COMB.	0.2863	0.9467	4	**0.265**	HAR-RV-COMB.	1.6230	1.0000	5
SPTSX	**QLIKE**	**p-value**	**Rank**	$MZ\text{-}R^2$	**SPTSX**	**HMSE**	**p-value**	**Rank**
PK (Model-free)	0.2150	1.0000	2	**0.631**	PK (Model-free)	0.4000	1.0000	3
GK (Model-free)	0.2167	1.0000	3	0.466	GK (Model-free)	**0.2766**	**1.0000**	**1**
RS (Model-free)	1.3330	eliminated	–	0.378	RS (Model-free)	0.3477	1.0000	2
YZ (Model-free)	**0.1932**	**1.0000**	**1**	0.474	YZ (Model-free)	2.1500	eliminated	–
HAR-RV	0.2532	eliminated	–	0.272	HAR-RV	1.8270	eliminated	–
HAR-RV-ONV	**0.2175**	**1.0000**	**4**	0.262	HAR-RV-ONV	1.7780	eliminated	–
HAR-RV-CC	0.2729	eliminated	–	0.243	HAR-RV-CC	1.7360	eliminated	–
HAR-RV-PK	0.2405	eliminated	–	0.263	HAR-RV-PK	**1.4420**	**1.0000**	**4**
HAR-RV-GK	0.2517	eliminated	–	0.237	HAR-RV-GK	1.5550	eliminated	–
HAR-RV-RS	0.2586	eliminated	–	0.250	HAR-RV-RS	1.7160	eliminated	–
HAR-RV-YZ	0.2480	eliminated	–	0.248	HAR-RV-YZ	1.7200	eliminated	–
HAR-RV-COMB.	0.2304	eliminated	–	**0.273**	HAR-RV-COMB.	1.5670	eliminated	–
SPX	**QLIKE**	**p-value**	**Rank**	$MZ\text{-}R^2$	**SPX**	**HMSE**	**p-value**	**Rank**
PK (Model-free)	0.2114	1.0000	2	**0.754**	PK (Model-free)	0.3105	1.0000	3
GK (Model-free)	0.2332	1.0000	3	0.661	GK (Model-free)	**0.2131**	**1.0000**	**1**
RS (Model-free)	2.7350	eliminated	–	0.560	RS (Model-free)	0.2929	1.0000	2
YZ (Model-free)	**0.2005**	**1.0000**	**1**	0.402	YZ (Model-free)	5.3760	eliminated	–
HAR-RV	0.3012	0.0070	9	0.353	HAR-RV	3.2730	eliminated	–
HAR-RV-ONV	0.2942	1.0000	7	0.253	HAR-RV-ONV	3.4880	eliminated	–
HAR-RV-CC	0.4054	1.0000	11	0.410	HAR-RV-CC	3.1940	eliminated	–
HAR-RV-PK	0.3334	1.0000	10	**0.439**	HAR-RV-PK	**1.9250**	**1.0000**	**4**
HAR-RV-GK	0.2908	1.0000	5	0.311	HAR-RV-GK	2.5470	eliminated	–
HAR-RV-RS	0.2993	0.0016	8	0.251	HAR-RV-RS	3.0710	eliminated	–
HAR-RV-YZ	0.2934	1.0000	6	0.355	HAR-RV-YZ	2.9860	eliminated	–
HAR-RV-COMB.	**0.2833**	**1.0000**	**4**	0.372	HAR-RV-COMB.	2.6100	eliminated	–

Note: *Loss functions and MCS procedure results are merged in this table. QLIKE and HMSE values are obtained by one-step-ahead rolling window forecasting method. MZ-R square stands for Mincer-Zarnowitz regression's R squares. Lower the values of QLIKE and HMSE is better while higher the value of MZ R-square is better for the comparison of forecasts. Window size is 600 observations that is used as in-sample estimation. The out-of-sample consists of 2180 observations. P-value and rank results are received from the MCS procedure. Bold numbers show the best-performing models for each indices.*

Even though the results for the group of Seven in terms of indices and criterion seem to be inconclusive, we can summarize the results as follows. In this sense, the HAR-RV model forecasts of CAC and DAX indices are improved by only overnight volatility, yet the inclusions of extreme range estimators do not add new information on the forecasts of realized volatility. However, the forecast combination of extreme range estimators does work for the CAC and DAX indices. Unlike the CAC and DAX stock market indices, realized volatility forecasts of FTSE index are improved best by the Parkinson range estimator (HAR-RV-PK). However, the inclusion of more sophisticated extreme range based estimators such as Roger-Satchell and Yang-Zhang do not improve the baseline HAR-RV model of FTSE index. In terms of the FTMIB index, HAR-RV-ONV is such a promising model for all the three loss functions so that ONV could be considered as an important variable for the FTMIB stock market. Afterwards, the results of NIKKEI index show that HAR-RV-CC model outperforms the others on the behalf of all the loss functions. In the case of the SPTSX index, the benchmark HAR-RV model could be improved best by the Parkinson volatility according to QLIKE and HMSE criterion (while MZ R-square rejects this result). For the SPX index, the results are mixed in terms of the criterion. For example, HMSE and MZ R-square show that HAR-RV-PK model is the best model by far compared to the rest of its counterparts, whereas QLIKE is in the favour of HAR-RV-COMB and HAR-RV-GK model.

An interesting findings in terms of extreme range estimators is that while traditional range estimators with simple formulas such as Parkinson and/or Garman-Klass have potential to improve the forecasts of realized volatility, the extreme range estimators derived from more complex formulas, namely Rogers-Satchell and Yang-Zhang generally contains no new information for forecasting realized volatility. All in all, when the different characteristics and dynamics of each stock markets are considered, it is highly unlikely to expect one single model is the winner against the others and therefore the rankings of best performing models are changeable from one market to another.

3.6. Conclusion

In this study, we examine whether extreme range estimators, which is derived from OHLC prices, contain important information for the future realized volatility in the G7 stock markets. Including Close-to-close, Overnight Volatility, Parkinson, Garman-Klass, Roger-Satchell, and Yang-Zhang estimators individually as an exogenous variable in the HAR-RV model framework, this work examines each estimator's ability for a better forecasting performance of the future realized volatility. This exercise is conducted using rolling and recursive window methods on the stock markets of group of Seven. The difference between target volatility and volatility forecasts are measured by QLIKE, HMSE and MZ R-square loss functions. Afterwards, those results are tested by the MCS procedure.

The results show several noteworthy points. We can say that the findings of this forecasting exercise are inconclusive on the G7 stock markets. First of all, the model-free approach (considering the estimators themselves as forecasters) significantly outperforms the HAR-RV-X model's forecasts. Second, the HAR-RV model forecasts of CAC and DAX indices are improved by only overnight volatility, whilst the inclusions of extreme range estimators do not add any new information on the forecasts of realized volatility. The other indices (such as FTSE, FTMIB, NIKKEI, SPTSX, and SPX) could be improved by Parkinson and/or Garman-Klass volatility estimators if included into the plain HAR-RV model. Third, while the traditional range estimators derived from simple formulas such as Parkinson and Garman-Klass have potential to improve the forecasts of realized volatility, the extreme range estimators derived from more complex formulas, namely Roger-Satchell and Yang-Zhang do not contain any new information for forecasting realized volatility in the G7 stock markets. Lastly, given the different characteristics and dynamics of each stock markets, it is unrealistic to expect one single model outperforms others in the group of Seven and therefore the rankings of best performing models are changeable from one market to another and also one time period to another. Applying our approach proposed in this chapter to a wider groups of stock market prices and oil and gold prices would be a plausible avenue of future research in order to validate the capacity of seven range-based estimators in forecasting realized volatility within a HAR-RV-X framework.

Overall, this chapter provides a new insight into the way practitioners and academics handle volatility forecasting. For instance, the predictive accuracy of realized volatility forecasts could be improved by the help of the range information. From this aspect, traditional extreme range estimators could be employed in the applications of the realized volatility in order to develop better portfolio and risk management strategies. As well as this, when the access to intraday frequency data is restricted, traditional extreme range estimators derived from the OHLC prices (which are publicly available in many cases) could be used as an alternative approach. This is also quite valuable for academics and practitioners in the case of lacking in tick-by-tick data. In the end, we should also suggest investors, practitioners, and academics about considering global risk factors while forecasting the future realized volatility. In this regard, the next chapter attempts to investigate the role of global information in improving the realized volatility forecasts.

CHAPTER 4

Importance of local, regional, and global information in forecasting realized volatility

ABSTRACT

This study aims to investigate the importance of exogenous volatilities at improving the forecasting accuracy of stock market volatility. In this regard, this volatility forecasting exercise is conducted on the SPX, FTSE, and GDAXI stock markets using the HAR-RV model with a wide range of different exogenous variables (i.e. HAR-RV-X). Additionally and more importantly, to evaluate the forecasting results from a different perspective, we classify our exogenous volatilities according to different information sources, namely local, regional, and global. Using our specific classification method we attempt to find out which class of models best helps with forecasting stock market volatility. In the HAR-RV-X framework, exogenous variables are used in various forms including, for example, each individual exogenous variable separately, forecast combination, and Kitchen-Sink approach. One-day-ahead out-of-sample volatility forecasts are generated using the rolling window mechanism and the QLIKE, HMSE, and HMAE loss functions for the forecast losses are employed. To check robustness of forecasts, we conduct the MCS procedure and different window sizes. The results present several noteworthy points. First, the predictive accuracy of stock market volatility increases where we include the other exogenous volatilities. Second, the integration of various volatilities, namely combination and Kitchen-Sink models, provides stronger forecast performance than the models with a single exogenous variable. Third, given the outperformance of the global information over local and regional information, the results are informative to reveal the dynamics of each markets.

4.1. Introduction

Forecasting financial volatility is a critical task in informing policy makers, investors, and market participants about their future risks and returns and the ways how to optimize their portfolios. Therefore, the accurate volatility forecasts are of interest to all of them. There exists a growing body of literature on generating accurate volatility forecasts. However, forecasting volatility accurately is still a big challenge for academics and practitioners.

The recent literature of volatility forecasting has been building upon the HAR-RV model. A great number of studies in the HAR literature concentrate on the extensions of the HAR model such as the realized semi-variances, jump component, asymmetries and leverage effects (see, among others, Barndorff-Nielsen, Kinnebrock, and Sheppard, 2010; Andersen et al. 2011; Sevi, 2014; Patton and Sheppard, 2015). However, a limited number of studies have been conducted which adds some exogenous variables to the HAR-RV model to improve the model's forecasting accuracy (e.g. Peng et al., 2018; Wang, 2019; Degiannakis and Filis, 2017; Kambouroudis, McMillan, and Tsakou 2021). Some of those studies concentrate on the forecasting oil price realized volatility while a few others examine from the perspective of stock markets. However, this work contributes to this literature from the stock market perspective by forecasting the future realized volatility with the HAR-RV-X specification and seek to reveal the cross-market and cross-asset information flows.

To investigate the impact of various exogenous volatilities at improving the forecasting accuracy of stock market volatility, this study considers a wide range of exogenous volatilities classified as local, regional, and global information sources. The aim of this research is to find out which class of exogenous volatilities best helps with forecasting stock market volatility. In this regard, this volatility forecasting exercise is carried out on the SPX, FTSE, and GDAXI stock market indices using the HAR-RV model with a range of different exogenous variables (i.e. HAR-RV-X) from local, regional, and global information channels. The time span is from 01 July 2009 to 28 May 2020. In the HAR-RV-X framework, exogenous variables are used in various forms including, for example, each individual exogenous variable separately, forecast combination, and the Kitchen-Sink method. One-day-ahead out-of-sample volatility forecasts are produced using the rolling window forecasting method and the QLIKE, HMSE, and HMAE loss functions for computing the forecast losses are employed. For further analysis, the MCS procedure, forecasts' robustness checks, and the cumulative HMSE difference of the baseline and winning models are applied. The results present several noteworthy points. First, the predictive accuracy of stock market volatility increases where we include the other exogenous volatilities. Second, the integration of various volatilities, namely combination and Kitchen-Sink models, improves the forecast performance much better than the models with a single exogenous variable. Lastly, in our samples, general conclusion is that the global information channel for major stock markets exhibits superior performance compared to the channels of local and regional information. Indeed, given the outperformance of the global information over local and regional information, the results could be important to inform market participants about the dynamics of markets.

This study is organised as follows: ***Section 2*** presents the review of related literature. In ***Section 3***, the methods used in this work are introduced in more detail. Afterwards, we give the data description in ***Section 4***. The empirical results and their evaluations are located in ***Sections 5***. Finally, the conclusion is presented in ***Section 6***.

4.2. Literature Review

As the third chapter of this thesis incorporates the baseline HAR-RV model with the extreme range estimators, including a wider group of exogenous variables with the so-called HAR-RV model is a reasonable concept to provide more accurate volatility forecasts. Many different kinds of additional variables could be considered for this purpose such as the VIX, EPU, other stock indices, oil prices, and interest rates. Previous studies, for example, Kambouroudis and McMillan (2016) add the VIX and volume as additional variables to the GARCH model whether these two exogenous variables provide any additional forecast power in the volatility forecasting content or not. They find some evidence that both additional variables improve the volatility forecasts of the US, the UK, and France stock markets. Kanas (2013) also find evidence for the S&P500 index that adding VIX squared in the GARCH equation gives a preferred forecasting results compared to the GARCH model without VIX. Similarly, Kanniainen, Lin, and Yang (2014) for S&P500 and Yang and Liu (2012) for emerging markets prove that exogenous variables, namely the VIX index improve the conditional volatility models' performance. More recently, Wang, Lu, He and Ma (2020) examine which additional predictors (VIX or EPU) that incorporate in the HAR-RV model do better improve the future volatility forecasts for 19 international stock market indices during the period of Covid-19 pandemic. Their results show that the VIX index is more useful for improving the forecasts of future volatility as the VIX is superior for 12 stock markets, whereas the EPU index can improve forecast accuracy only for 5 market indices. As a generalized result, the idea of including various exogenous variables in the literature of realized volatility takes an important place, especially in the HAR-RV model's framework (e.g. Peng et al., 2018; Liu et al., 2019; Wang, 2019).

The idea of employing high frequency data to compute measures of volatility at a lower frequency is first suggested by Merton (1980). After the 2000s, the easy access to high frequency data has spurred the studies of the realized volatility. The realized volatility (RV) is defined as the sum of squared intraday returns and first proposed by Andersen and Bollerslev (1998). It is an alternative measure of daily volatility that used to generate more accurate daily volatility measures. Initially, a number of studies use the ARFIMA model to model and forecast the RV. Afterwards, Corsi (2009) points out the drawbacks of the ARFIMA-RV model, saying it is just a mathematical trick and does not have a clear economic interpretation, and proposed a new model that is the Heterogeneous Autoregressive model of the Realized Variance (HAR-RV). The HAR-RV model is based on Heterogenous Market Hypothesis (HMH) and the recent literature of volatility forecasting has been thriving on it. In the HAR framework, daily realized volatility is modelled using as a function of past daily, weekly, and monthly components. In other words, it is known as an additive cascade model of different volatility components. The intuition behind the HAR-RV model is quite straightforward. It assumes the heterogenous nature of market participants, implying that the components of the HAR-RV model reflect shorter or longer term trading activity of market participants.

A large number of studies in the HAR literature concentrate on the extensions of the HAR model to consider the realized semi-variances, jump component, asymmetries and leverage effects (see, among others, Barndorff-Nielsen, Kinnebrock, and Sheppard, 2010; Andersen et al. 2011; Corsi and Reno, 2012; Sevi, 2014; Patton and Sheppard, 2015). There are also a limited number of studies adding exogenous variables to the HAR-RV model (i.e. HAR-RV-X) to improve the model's forecasting accuracy (e.g. Peng et al., 2018; Wang, 2019; Degiannakis and Filis, 2017). In this context, this area of research concentrates mainly on the realized volatility studies of stock market and oil price. The next paragraph in the HAR-RV-X framework discusses current stock market related works and then the following paragraph is about the volatility forecasting of oil price.

Peng, Chen, Mei, and Diao (2018) investigate whether the G7 countries' stock market indices can improve the forecasting accuracy of Chinese stock market realized volatility or not. They find that the G7 stock markets can contain useful information to forecast the one-day ahead volatility of the Chinese stock market. They also suggest that the best performing model is the Kitchen Sink model that all the G7 stock market indices are used at once in the exogenous part of the HAR-RV model. This study is one of the similar works to ours. As for the differences, their work is carried out from the perspective of Chinese stock market and consider only G7 stock market indices as attached component to the plain HAR-RV model. Liu, Ma, and Zhang (2019) extend the work of Peng et al. (2018) using twenty-seven global stock markets in order to forecast the Chinese stock market realized volatility. They criticise the Kitchen Sink model of Peng et al. (2018) by saying that the model can cause overfitting issue due to pushing all the global stock markets in the same model. Thereupon, they suggest implementing the HAR-RV-X model with time-varying parameter and various forecasting combination strategies to extract global stock information. To compare with the Kitchen Sink model, they produce 27 individual HAR-RV-X models and take a weighted average of the individual forecasts using the mean average, trimmed mean average, median average, and discount mean square prediction error. The findings document that the median combination of time varying HAR-RV model is the best performing for the Chinese stock market. The work of Liu et al. (2019) is also related to ours in the sense that they attempt to extract only global stock information while we consider local, regional, and global. Wang (2019) analyses the linkage between the CBOE VIX index and 13 stock markets of G20 countries. According to the findings, the VIX index as an exogenous component to the HAR-RV model can improve the forecasting performance in the international stock markets and, moreover, the large VIX exhibits a stronger forecast performance on future RV in comparison to the original VIX. Duan, Chen, Zeng, and Liu (2018) examine the impacts of economic policy uncertainty (EPU) and leverage effect on future volatility in the framework of regime switching HAR-RV (MS-HAR-RV). The findings indicate that the MS-HAR-RV model including the EPU and leverage effect with regimes can provide higher forecast accuracy compared to the HAR-RV type and GARCH class models. Mei, Liu, Ma, and Chen (2017) investigate the impacts of realized skewness and kurtosis on stock market realized volatility. In their work, the realized skewness and kurtosis as additional variables incorporate into the HAR-RV model. Out-of-sample results for US and Chinese stock markets indicate that the realized skewness can improve the accuracy of forecasts in the middle and long term forecasting, whereas the realized kurtosis cannot enhance the model's performance. Kambouroudis, McMillan, and Tsakou (2021) examine whether some exogenous variables, including, for example, the implied volatility, leverage effect, overnight returns, and volatility of realized volatility help in forecasting future realized volatility in 10 international stock markets. The empirical findings show that including each these exogenous variables in the HAR-RV model improves the forecasting accuracy of future realized volatility, with the exception of the volatility of realized volatility. It is also pointed out that implied volatility provides stronger forecast than its other counterparts.

Degiannakis and Filis (2017) conduct a relatively similar study with ours, but it is on oil price realized volatility forecasting. They investigate cross-market volatility flows by categorising exogenous volatilities according to four different asset classes such as Stocks, Forex, Macro, and Commodities. It is pointed out that the HAR-RV model with exogenous volatilities exhibit better forecasting performance for oil price volatility compared to the baseline HAR-RV model and the best performing model is suggested as a combination of multiple asset classes' volatilities. Their combination model corresponds to the Kitchen Sink model of this study. Ma, Wahab, and Liu (2018) examine the role of only economic policy uncertainty (EPU) index at improving the forecasting accuracy of crude oil futures' realized volatility. They use only one exogenous variable to the HAR-RV model, but the internal extensions of the HAR model such as the realized semi-variances, jump component, and leverage effects are considered and also some thresholds for the EPU are set. The forecasting accuracy of those models for different horizons are evaluated by both economic and statistics value analysis. According to the results, the EPU achieves higher forecast accuracy than the baseline HAR-RV model and, moreover, the above-threshold EPU exhibits higher forecast performance. Yu (2019) adds the leverage effect and EPU to the HAR-RV model and find that considering both the EPU and leverage effect in the model can substantially help in forecasting the Bitcoin price volatility.

In short, majority of the abovementioned volatility forecasting papers related to the HAR-RV-X model analyse the Chinese stock market and the information flow is from the western international stock markets towards the Chinese stock market. Also, some other related studies focus on the forecasting exercises of oil price realized volatility. Unlike the papers concentrate above, this study seeks to reveal the impacts of the cross-market and cross-asset information flows and our work contributes to the scarce literature of stock market realized volatility using the HAR-RV-X models by extending them in multiple ways. For instance, we classify our exogenous volatilities according to their geographical sources and mainly attempt to inform market participants about the impacts of information flow between the US and EU financial data (i.e. cross-market and cross-asset).

4.3. Methodology

4.3.1. Kitchen-Sink approach

Geographically classified exogenous volatilities such as local, regional, and global information are added to the baseline HAR-RV model as new explanatory variables. In the HAR-RV-X framework, exogenous variables are used in various forms including, for example, each individual exogenous variable separately, forecast combination and Kitchen-Sink approach that includes a sets of exogenous variables at the same time in the same model. The HAR-RV-X model specifications for individual forecasts and Kitchen-Sink forecasts are given in the below Equations (1 and 2) respectively as follows:

$$RV_{t+1}^{d} = \beta_0 + \beta_d RV_t^d + \beta_w RV_t^w + \beta_m RV_t^m + \boldsymbol{\beta_X X_t^d} + \varepsilon_{t+1} \quad (1)$$

$$RV_{t+1}^{d} = \beta_0 + \beta_d RV_t^d + \beta_w RV_t^w + \beta_m RV_t^m + \sum_{\boldsymbol{i=1}}^{\boldsymbol{K}} \boldsymbol{\beta_i X_{i,t}^d} + \varepsilon_{t+1} \quad (2)$$

In Equation 1, the exogenous component, $\beta_X X_t^d$, refers to the i^{th} individual exogenous volatility at day t. We can obtain 13 different individual HAR-RV-X model using this formula. For example, the HAR-RV-GDAXI, HAR-RV-FCHI, HAR-RV-SPX, HAR-RV-VIX, HAR-RV-GOLD, HAR-RV-BOND, HAR-RV-EPU … are obtained from the abovementioned Formula 1. Equation 2 implies the Kitchen-Sink models where $\sum_{i=1}^{K} \beta_i X_{i,t}^d$, represents the multi-exogenous variables. These Kitchen-Sink (KS) models used in this work can be called as follows: the HAR-RV-REGIONAL-KS (GDAXI+FCHI+FTMIB+STOXX), HAR-RV-GLOBAL-KS (SPX+DJI+IXIC+VIX+WTI+GOLD), HAR-RV-LOCAL-KS (BOND+EPU+LIBOR), and HAR-RV-OVERALL-KS (LOCAL+REGIONAL+GLOBAL). The combination method simply takes the average of all the individual forecasts in groups. The forecast combinations are the simple average of all included forecasts, which can be calculated as follows: the sum of individual forecasts is divided to the numbers of individual forecasts. For instance, the formula in the bracket (HAR-RV-BOND + HAR-RV-EPU + HAR-RV-LIBOR / 3) is for the local-combination and the same technique is too applied to derive the local, regional and global forecast combinations.

4.3.2. Rolling window and loss functions

The rolling window is one of the most popular methods in forecasting and therefore this study employs the rolling window technique in order to generate the volatility forecasts of stock markets. Initially, the whole sample needs to be divided into two subgroups such as the initial sample and out-of-sample windows. In the literature, there is no consensus on how to select an appropriate forecasting window. Since the main objective of this work is to evaluate the out-of-sample performance of the models, we arbitrarily choose the initial and out-of-sample windows considering a length that allows the regression fit normally and obtain longer out-of-sample period. The rolling window's working principle does work the way that the estimation sample is then rolled forward by adding one new observation and dropping the most distant observation. In this way, the size of initial sample window used to estimate the models remains at a fixed length. We produce only the one-step-ahead volatility forecasts of the stock markets. The reason is that the forecasts more than one-step-ahead are highly likely to give poorer forecasts due to the lack of information of further point prediction.

To evaluate the out-of-sample accuracy of competing models, we select three of the most popular loss functions in the literature, namely quasi-Gaussian log-likelihood (QLIKE), heteroskedasticity adjusted mean squared error (HMSE), and heteroskedasticity adjusted mean absolute error (HMAE) in line with recent studies (e.g. Zhou, Pan, and Wu, 2019; Ma et al., 2018; Liu et al, 2019).

$$QLIKE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}[log\widehat{RV_t^2} + \frac{\widehat{RV_t^2}}{RV_t^2}] \quad (3)$$

$$\text{H}MSE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau}[1 - \widehat{RV_t^2}/ RV_t^2]^2 \quad (4)$$

$$\text{H}MAE = \frac{1}{\tau}\sum_{t=T+1}^{T+\tau} | 1 - \frac{\widehat{RV_t^2}}{RV_t^2} | \quad (5)$$

where $\widehat{RV_t^2}$ denotes the out-of-sample volatility forecast from competing models and RV_t^2 is a proxy for true market volatility. τ is the number of out-of-sample forecasting days. Each one of the loss functions have a specific calculation method in order to measure the forecast error. According to Patton (2011), these three well-established loss functions can provide consistent rankings for competing volatility models in the case of a noisy volatility proxy.

4.4. Data Description

Daily realized variance series are collected from the Oxford-Man Institute's Quantitative Finance Realized Library.[17] According to the seminal paper of Liu, Patton, and Sheppard (2015), no one measure significantly outperforms the 5-minute realized variance among a set of 400 different volatility estimators. Therefore, we use 5-min realized variance series, which is a widely accepted robust volatility measure. In this study, we employ a wide range of different financial and economic data. For instance, the eight international stock market indices included are: FTSE 100 (UK), GDAXI (Germany), FCHI (France), FTMIB (Italy), STOXX50E (Euro Stox 50), SPX (S&P 500), DJI (Dow Jones Industrial Average), and IXIC (Nasdaq 100). The reason why they are chosen is that we are primarily interested in the information flow between the largest and most active American and European stock markets. Additionally, we use the Federal Reserve Bank of St. Louis (FRED) database in order to obtain the data of the CBOE volatility index (VIX), the Crude Oil Prices (WTI; West Texas Intermediate), the CBOE Gold ETF Volatility Index, the UK and US Government 10-year Treasury Bond Yields, and the London 12-month Interbank offered rates based on POUND (LIBOR-POUND), USD (LIBOR-USD), and EURO (LIBOR-EURO). The volatility index of the GDAXI and Germany Government 10-year Treasury Bond Yield are extracted using the webpage 'investing.com'. The indices of UK and US economic policy uncertainty (EPU) are provided by the webpage of the Economic Policy Uncertainty.[18]

In this study, the time interval of full sample is from 01 July 2009 to 10 April 2020. We are restricted to this time span and also the abovementioned variables due to the scarcity of daily frequency data. The different stock markets have the different trading days. Therefore, we need to align our dataset to the days that all the markets have active trading. To do this, the unmatched days (by any of the series) have to be omitted. In other words, each row of all the series has to match the same date point. After carrying out this data cleaning, we obtain the 2600 observations of the twenty series, which are aligned to the same date points in rows.

[17] https://realized.oxford-man.ox.ac.uk/
[18] https://www.policyuncertainty.com/

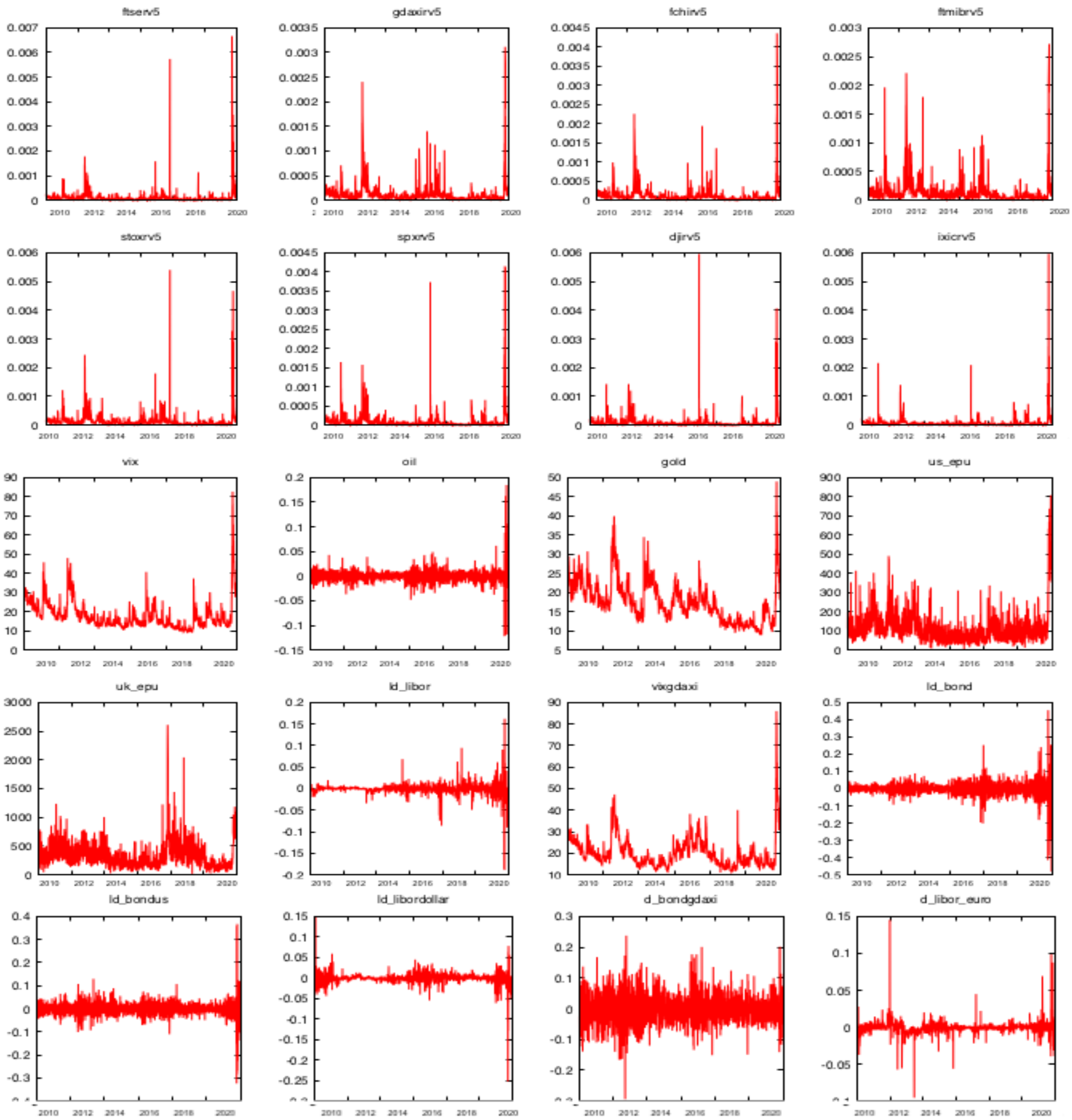
4.Table 1 Descriptive statistics of the series

	Mean	Std. Dev.	Skew.	Ex. Kurt.	Jarque-Bera	Q(5)	ADF
FTSE	9.91E-05	0.00024	14.986***	314.99***	1.09E+07***	2602.15***	-7.975***
GDAXI	0.000110	0.00017	7.7620***	87.446***	863110***	5541.95***	-7.163***
FCHI	0.000112	0.00020	10.123***	150.49***	2.52E+06***	5376.08***	-7.934***
FTMIB	0.000122	0.00017	6.2168***	58.473***	390868***	4748.29***	-6.651***
STOXX50E	0.000131	0.00026	10.683***	156.23***	2.71E+06***	4174.94***	-9.508***
SPX	8.31E-05	0.00021	10.834***	154.84***	2.67E+06***	4770.71***	-9.261***
DJI	8.50E-05	0.00024	12.721***	224.01***	5.55E+06***	3718.72***	-9.684***
IXIC	7.45E-05	0.00019	15.448***	353.84***	1.37E+07***	4368.95***	-9.128***
VIX	17.908	7.3291	2.7481***	13.057***	21950.5***	11127.8***	-5.664***
OIL	-7.52E-06	0.01265	1.4177***	45.893***	231243***	104.958***	-10.28***
GOLD	17.555	5.2736	1.0482***	2.0058***	920.793***	11536.8***	-4.377***
BOND	-0.00109	0.03915	-0.8040***	33.515***	123090***	15.8668***	-9.577***
UKEPU	331.66	206.34	2.4136***	14.328***	25003.1***	4862.01***	-4.916***
LIBOR	-0.00042	0.01175	-1.0463***	85.886***	806962***	255.72***	-8.677***
VIXDAX	20.132	7.1466	2.7106***	13.832***	24122.3***	11337.9***	-5.179***
BONDDAX	-0.00156	0.0438	0.1837***	2.5038***	699.622***	17.0596***	-16.16***
LIBOREU	-0.00066	0.0074	3.4517***	94.700***	984961***	566.049***	-6.561***
USEPU	110.53	78.832	3.0009***	13.736***	24558.5***	6538.77***	-3.510***
BONDUS	-0.00063	0.0292	0.1477***	32.964***	118726***	60.6513***	-8.871***
LIBORUSD	-0.00032	0.0118	-4.3787***	104.06***	1.19E+06***	554.636***	-8.273***

Notes: Asterisk *,**, and *** denote rejections of null hypothesis at 10%, 5%, and 1% significance levels, respectively. The null hypothesis of the third and fourth moments are "Skewness = 0" and "Excess Kurtosis = 3".

Table 1 indicates the descriptive statistics of realized variance series of eight international stock markets, VIX, VIX-GDAXI, Oil Prices, Gold Volatility Index, UK, US and Germany 10-year Treasury Bond Yield, UK and US Economic Policy Uncertainty Index, and London 12-month Interbank offered rates (LIBOR-POUND, USD, and EURO). According to this table, all the series are significantly skewed and leptokurtic at the 99 per cent confidence level, meaning that each series has fat-tail distribution. The Jarque-Bera statistic values show the non-normality of all the series at the 99% confidence level. The Ljung-Box statistic for serial correlation indicates that the null hypothesis of no autocorrelation up to the 5th order are rejected for all the series, suggesting the existence of autocorrelation. Lastly, the Augmented Dickey-Fuller (ADF) test statistic values indicate that all the series are stationary because the null hypothesis of a unit root is rejected for all the series. The line graphs are also plotted in Figure 1.

4.Figure 1 ***Line graphs of all the series***



4.5. Empirical Results

In this work, a wide range of exogenous variables are added to the baseline HAR-RV model to examine whether an exogenous variable(s) improves the predictability of stock market volatility or not. To do this, we categorise our sample exogenous variables in respect to their information sources such as local, regional, and global. Due to time and data limitation, this exercise is conducted only on several stock markets that are the SPX, FTSE, and GDAXI. The classification of additional variables are as follows. Firstly, the BOND, LIBOR, and EPU, which are the financial data of specific markets, are determined to investigate the effects of local information on the future volatility of the stock markets. Secondly, the major European stock markets (FTSE, GDAXI, FCHI, FTMIB, and STOXX50E) are assigned to the regional information and lastly the global information is represented by the SPX, DJI, IXIC, VIX, OIL, and GOLD.

4.5.1. In sample evaluation

Since the main objective of this study is to assess the out-of-sample accuracy of the forecasting models, we prioritise the out-of-sample performance of the models rather than their in-sample fit. It is also worth to point out that high accuracy of model in the in-sample does not necessarily mean that it will exhibit the same performance in the out-of-sample. Therefore, we give a brief in-sample analysis here and then focus more on the out-of-sample performance of the competing models.

Tables 2, 3, and 4 indicate the full sample estimation results of the HAR-RV-X models. For clarity, “HAR-RV-“ is the baseline model and the latter denotes an exogenous variable added to the model (e.g. HAR-RV-SPX). We estimate the HAR-RV-X model using OLS method, which is based on the Newey-West/Bartlett correction allowing for correlation up to the order of 5. Generally speaking, at the 99 per cent significance level, the daily, weekly and exogenous components of the HAR-RV-X models for the SPX and GDAXI are positive and statistically significant. Most of the monthly components for all the three markets are insignificant. The daily components are insignificant only for the FTSE. According to the Tables, an exogenous variable in the models seems to have a positive impact on the volatility of the SPX, FTSE, and GDAXI stock markets. In terms of the models’ goodness-of-fit, the adjusted R-squares seem normal, which are higher than 0.35, 0.50, and 60, respectively for the FTSE, SPX, and GDAXI. One point is that when more than one exogenous variable is added to the HAR-RV-X model (e.g. HAR-RV-GLOBAL and HAR-RV-REGIONAL+GLOBAL+LOCAL), the adjusted R-squares have risen relatively. In other words, the new included explanatory variables relatively contribute to the dependent variable, meanwhile we need to be careful about the numbers of new exogenous variables due to the possibility of the overfitting issue. According to the results, there is a positive sign shows that the majority of exogenous variables add extra information to the baseline models.

4.Table 2

Full sample volatility estimation results of HAR-RV-X models for the S&P 500 stock market

Models	$Constant$	β_1	β_2	β_3	β_X	$Adj. R^2$
HAR-RV (Baseline)	1.48E-05***	0.292***	0.593***	-0.065	-	0.53
REGIONAL INFORMATION						
HAR-RV-GDAXI	8.06E-06**	0.229***	0.548***	-0.076***	0.151***	0.54
HAR-RV-FCHI	9.55E-06*	0.220***	0.533***	-0.074	0.151***	0.54
HAR-RV-FTMIB	2.10E-06	0.222***	0.547***	-0.074	0.188**	0.54
HAR-RV-STOX	9.49E-06**	0.196**	0.548***	0.068	0.131	0.54
HAR-RV-FTSE	1.34E-05***	0.238***	0.583***	-0.076	0.077	0.53
GLOBAL INFORMATION						
HAR-RV-DJI	1.42E-05***	0.732***	0.547***	-0.058	-0.385***	0.54
HAR-RV-IXIC	1.31E-05***	0.087	0.545***	-0.042	0.277	0.54
HAR-RV-VIX	-0.00011***	0.212***	0.521***	-0.301**	9.33E-06***	0.56
HAR-RV-WTI	1.50E-05***	0.282***	0.581***	-0.046	-0.001***	0.54
HAR-RV-GOLD	-1.21E-05	0.288***	0.587***	-0.090	1.70E-06***	0.53
LOCAL INFORMATION						
HAR-RV-BOND	1.49E-05***	0.287***	0.598***	-0.067**	-0.0001	0.53
HAR-RV-USEPU	1.03E-05*	0.291***	0.594***	-0.081	5.24E-08	0.53
HAR-RV-LIBOR	1.44E-05***	0.288***	0.597***	-0.063	-0.0007	0.53

	$Constant$	β_1	β_2	β_3	β_{GDAXI}	β_{FCHI}	β_{FTMIB}	β_{STOXX}	β_{FTSE}	$Adj. R^2$
HAR-RV-REGIONAL	2.02e-06	0.191**	0.547***	-0.074	0.013	-0.064	0.164	0.075	0.013	0.54

	$Constant$	β_1	β_2	β_3	β_{DJI}	β_{IXIC}	β_{VIX}	β_{OIL}	β_{GOLD}	$Adj. R^2$
HAR-RV-GLOBAL	-9.52E-05***	0.270	0.446***	-0.247**	-0.211**	0.233	1.02E-05***	-0.001***	-2.44E-06*	0.57

	$Constant$	β_1	β_2	β_3	β_{BOND}	β_{EPU}	β_{LIBOR}	$Adj. R^2$
HAR-RV-LOCAL	1.13E-05**	0.286***	0.599***	-0.076	-4.84E-05	3.82E-08	-0.0006	0.53

	$Constant$	β_1	β_2	β_3	β_{GDAXI}	β_{FCHI}	β_{FTMIB}	β_{STOXX}	β_{FTSE}	β_{DJI}	β_{IXIC}	
HAR-RV-LOCAL+ REGIONAL+GLOBAL	-8.42E-05***	0.164	0.406***	-0.221**	-0.080	0.041	0.129	0.018	0.052	-0.196*	0.256*	⇒

	β_{VIX}	β_{OIL}	β_{GOLD}	β_{BOND}	β_{EPU}	β_{LIBOR}	$Adj. R^2$
⇒	9.87E-06***	-0.001***	-3.19E-06*	0.0002	-3.31E-08	-0.0004	0.58

Notes: β_1, β_2 and β_3 are the daily, weekly, and monthly components of the HAR-RV model. $\boldsymbol{\beta_X}$ denotes the X (exogenous) variable that is named after within the first column. Asterisk *,**, and *** denote rejections of null hypothesis at 10%, 5%, and 1% significance levels, respectively.

4.Table 3

Full sample volatility estimation results of HAR-RV-X models for the FTSE 100 stock market

Models	$Constant$	β_1	β_2	β_3	β_X	$Adj.R^2$
HAR-RV (Baseline)	2.01E-05***	0.772***	-0.017	0.041	-	0.35
REGIONAL INFORMATION						
HAR-RV-GDAXI	9.06E-06***	-0.024	0.621***	-0.018	0.296***	0.37
HAR-RV-FCHI	1.29E-05***	-0.027	0.601***	-0.011	0.270***	0.37
HAR-RV-FTMIB	5.42E-06	0.0002	0.648***	-0.003	0.241***	0.37
HAR-RV-STOX	1.11E-05***	-0.211***	0.611***	0.018	0.353***	0.38
GLOBAL INFORMATION						
HAR-RV-SPX	2.06E-05***	-0.082**	0.530***	0.031	0.371***	0.39
HAR-RV-DJI	2.07E-05***	-0.062	0.623***	0.013	0.252**	0.38
HAR-RV-IXIC	1.84E-05***	-0.028	0.537***	0.038	0.354***	0.39
HAR-RV-VIX	-0.00012***	0.011	0.626***	-0.249*	1.06E-05***	0.38
HAR-RV-WTI	2.02E-05***	0.027	0.769***	-0.002	-0.001	0.35
HAR-RV-GOLD	-3.31E-05**	0.037	0.756***	-0.069	3.44E-06***	0.36
LOCAL INFORMATION						
HAR-RV-BOND	2.01E-05***	0.019	0.809***	-0.038	-0.0005	0.36
HAR-RV-UKEPU	3.73E-05**	0.044	0.772***	0.008	-6.08E-08	0.35
HAR-RV-LIBOR	2.06E-05***	0.044	0.763***	-0.014	0.0005	0.35

	$Constant$	β_1	β_2	β_3	β_{GDAXI}	β_{FCHI}	β_{FTMIB}	β_{STOXX}	$Adj.R^2$
HAR-RV-REGIONAL	6.14e-06	-0.174**	0.593***	0.011	0.136*	-0.074	0.065	0.268	0.37

	$Constant$	β_1	β_2	β_3	β_{SPX}	β_{DJI}	β_{IXIC}	β_{VIX}	β_{OIL}	β_{GOLD}	$Adj.R^2$
HAR-RV-GLOBAL	-7.03E-05***	-0.062*	0.436***	-0.101	0.507**	-0.289**	0.119	7.15E-06***	-0.0008	-6.76E-07	0.41

	$Constant$	β_1	β_2	β_3	β_{BOND}	β_{EPU}	β_{LIBOR}	$Adj.R^2$
HAR-RV-LOCAL	3.84E-05**	0.025	0.800***	-0.008	-0.0006	-6.23E-08	0.0007	0.36

	$Constant$	β_1	β_2	β_3	β_{GDAXI}	β_{FCHI}	β_{FTMIB}	β_{STOXX}	β_{SPX}	β_{DJI}	β_{IXIC}	
HAR-RV-LOCAL+ REGIONAL+GLOBAL	-6.63E-05***	-0.209***	0.473**	-0.089	0.022	-0.196**	0.049	0.240***	0.328*	-0.159	0.133	⇒

		β_{VIX}	β_{OIL}	β_{GOLD}	β_{BOND}	β_{EPU}	β_{LIBOR}	$Adj.R^2$
HAR-RV-LOCAL+ REGIONAL+GLOBAL	⇒	6.96E-06***	-0.0006	-6.17E-07	-0.0005	-2.32E-08	0.0005	0.42

Notes: β_1, β_2 and β_3 are the daily, weekly, and monthly components of the HAR-RV model. $\boldsymbol{\beta_X}$ denotes the X (exogenous) variable that is named after within the first column. Asterisk *,**, and *** denote rejections of null hypothesis at 10%, 5%, and 1% significance levels, respectively.

4.Table 4

Full sample volatility estimation results of HAR-RV-X models for the GDAXI stock market

Models	$Constant$	β_1	β_2	β_3	β_X	$Adj.R^2$
HAR-RV (Baseline)	1.39E-05***	0.465***	0.386***	0.020	-	0.60
REGIONAL INFORMATION						
HAR-RV-FCHI	1.41E-05***	0.412***	0.374***	0.026	0.057	0.61
HAR-RV-FTMIB	7.95E-06	0.325***	0.369***	0.024	0.185*	0.61
HAR-RV-STOX	1.42E-05***	0.297**	0.366***	0.041	0.136	0.62
HAR-RV-FTSE	1.46E-05***	0.399***	0.383***	0.021	0.068	0.61
GLOBAL INFORMATION						
HAR-RV-SPX	1.79E-05***	0.295***	0.323***	0.063	0.203***	0.63
HAR-RV-DJI	1.59E-05***	0.363***	0.359***	0.044	0.112	0.62
HAR-RV-IXIC	1.55E-05***	0.279***	0.303***	0.100	0.257***	0.64
HAR-RV-VIX	-6.84E-05***	0.399***	0.321***	-0.147	6.46E-06***	0.63
HAR-RV-WTI	1.39E-05***	0.460***	0.371***	0.039	-0.001**	0.61
HAR-RV-GOLD	-2.30E-05**	0.455***	0.378***	-0.033	2.56E-06***	0.61
LOCAL INFORMATION						
HAR-RV-BOND	1.38E-05***	0.464***	0.388***	0.018	-5.23E-05	0.61
HAR-RV-VIXDAX	-0.00011***	0.375***	0.312***	-0.225*	8.45E-06***	0.63
HAR-RV-LIBOR	1.50E-05***	0.471***	0.378***	0.018	0.001	0.61

	$Constant$	β_1	β_2	β_3	β_{FCHI}	β_{FTMIB}	β_{STOXX}	β_{FTSE}	$Adj.R^2$
HAR-RV-REGIONAL	8.04e-06*	0.300***	0.381***	-0.036	-0.177	0.165*	0.194	-0.033	0.62

	$Constant$	β_1	β_2	β_3	β_{SPX}	β_{DJI}	β_{IXIC}	β_{VIX}	β_{OIL}	β_{GOLD}	$Adj.R^2$
HAR-RV-GLOBAL	-4.00E-05**	0.260***	0.232***	-0.002	0.467**	-0.386***	0.170*	3.23E-06**	-0.0008*	1.12E-06**	0.66

	$Constant$	β_1	β_2	β_3	β_{BOND}	β_{VIXDAX}	β_{LIBOR}	$Adj.R^2$
HAR-RV-LOCAL	-0.0001***	0.379***	0.306***	-0.231**	-5.63E-05	8.53E-06***	0.001	0.63

	$Constant$	β_1	β_2	β_3	β_{FTSE}	β_{FCHI}	β_{FTMIB}	β_{STOXX}	β_{SPX}	β_{DJI}	β_{IXIC}	
HAR-RV-LOCAL+ REGIONAL+GLOBAL	-8.69E-05***	0.261**	0.233***	-0.092	-0.051	-0.266	0.140	0.124	0.402*	-0.313**	0.173*	⇒

	β_{VIX}	β_{OIL}	β_{GOLD}	β_{BOND}	β_{VIXDAX}	β_{LIBOR}	$Adj.R^2$
⇒	8.88E-07	-0.0007**	5.09E-07	2.04E-05	5.48E-06***	0.0008	0.68

Notes: β_1, β_2 and β_3 are the daily, weekly, and monthly components of the HAR-RV model. $\boldsymbol{\beta_X}$ denotes the X (exogenous) variable that is named after within the first column. Asterisk *,**, and *** denote rejections of null hypothesis at 10%, 5%, and 1% significance levels, respectively.

4.5.2. Out-of-sample evaluation

The whole sample, which is 2600 trading days, is divided into two subgroups: initial sample and out-of-sample windows. The initial sample period has a fixed length, 400 observations. This is the fact that the initial sample accuracy does not necessarily affect the out-of-sample accuracy. Therefore, we arbitrarily choose the 400 observations considering the length at least one year period to let the regression fit normally and obtain a longer out-of-sample period. This is because the main objective of this work is to evaluate the out-of-sample performance of the models. The length of out-of-sample is 2200. Defining the initial and out-of-sample periods, we apply the rolling window method to produce the one-step-ahead volatility forecasts of the stock markets. The rolling window methodology works the way that the estimation sample is then rolled forward by adding one new observation and dropping the most distant observation. In this way, the size of initial sample window used to estimate the models remains at a fixed length. To evaluate the out-of-sample accuracy of the HAR-RV-X models, we apply three different well established loss functions, namely, QLIKE, HMSE, and HMAE. The MCS procedure that identifies the set of the best models is employed for the further analysis of the predictive ability of competing models. For the robustness check, we repeat this forecasting exercise under two more alternative forecasting windows (e.g. 200 and 600 observations).

Essentially, we investigate the impacts of information flow between the EU and US stock markets. The exogenous volatilities are sorted according to our specific classification, which is called as local, regional, and global. More precisely, the exogenous variables; BOND, EPU, VIXDAX, and LIBOR represent the local information. In a similar vein, the EU stock market indices (GDAXI, FCHI, FTMIB, STOXX50E, and FTSE) refer to the regional information. Lastly, the data of US (SPX, DJI, IXIC, VIX, OIL, and GOLD) are employed to understand the role of global information. Our distinctive classification investigates how important the other exogenous volatilities are at improving the predictive accuracy of the stock markets. In this regard, a wide range of exogenous variables are added to the baseline HAR-RV model. Exogenous variables (in the HAR-RV-X model) are used in various forms including, for example, each individual exogenous variable (separately), forecast combination (simple average of individual forecasts for groups), and Kitchen-Sink model (all additional variables included at once in the model).

4.5.3. Overall findings of loss functions and MCS procedure

Tables 5, 6, and 7 show the out-of-sample one-step-ahead rolling windows forecasting and the MCS results of competing models for the SPX, FTSE, and GDAXI stock market indices respectively. This exercise is performed for the aforementioned three major stock market indices in order to strengthen and also enrich our work. The results of the loss functions and the MCS test are merged in the lined columns. The first columns are the loss functions that are calculated as the difference between proxy measure and forecast. The second and third columns are about the MCS procedure that indicates the p-value and the ranking of the MCS. The winners are selected by considering both the loss functions and the MCS procedure.

Afterwards, the evaluations and discussions of competing models for each indices are given in the same section sequentially.

4.Table 5

Out-of-sample 1-step-ahead rolling window forecasting and MCS results for SPX (Window size:400)

S&P 500	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-8.9981	0.0052	18	2.7107	0.0240	14	1.0334	eliminated	–
REGIONAL INFORMATION									
HAR-RV-GDAXI	-8.9977	1.0000	12	2.3973	0.0124	19	0.9838	0.0016	11
HAR-RV-FCHI	-8.9672	1.0000	13	2.5514	1.0000	12	1.0060	0.0040	9
HAR-RV-FTMIB	-8.9797	1.0000	10	2.5762	0.0204	15	1.0072	0.0042	8
HAR-RV-STOXX50E	-8.9627	1.0000	6	2.6847	1.0000	9	1.0257	0.0030	10
HAR-RV-FTSE	-8.9771	1.0000	16	2.8616	0.0170	17	1.0433	eliminated	–
REGIONAL KITCHEN-SINK	-8.8571	1.0000	15	2.7692	0.0068	21	1.0186	0.0000	13
REGIONAL COMBINATION	-9.0060	0.0010	20	2.5525	0.0138	18	1.0067	0.0008	12
GLOBAL INFORMATION									
HAR-RV-DJI	-8.9656	1.0000	17	2.3077	0.0030	22	0.9972	0.0000	14
HAR-RV-IXIC	-8.8754	1.0000	2	2.2060	1.0000	3	0.9579	0.0452	4
HAR-RV-VIX	-7.2554	1.0000	8	2.6734	1.0000	2	1.0338	0.1990	2
HAR-RV-WTI	-8.9461	1.0000	5	2.8192	0.0286	10	1.0344	eliminated	–
HAR-RV-GOLD	-8.9831	1.0000	11	2.6567	1.0000	6	1.0085	0.0106	5
GLOBAL KITCHEN-SINK	-7.7278	1.0000	3	2.8071	1.0000	4	1.0390	0.0674	3
GLOBAL COMBINATION	**-9.0271**	**1.0000**	**1**	**1.7009**	**1.0000**	**1**	**0.8670**	**1.0000**	**1**
LOCAL INFORMATION									
HAR-RV-BOND	-8.6652	1.0000	9	2.7160	0.0104	20	1.0367	eliminated	–
HAR-RV-USEPU	-8.9961	0.0002	22	2.9241	1.0000	11	1.0620	eliminated	–
HAR-RV-LIBOR	-8.9810	0.0032	19	2.9549	1.0000	7	1.0584	eliminated	–
LOCAL KITCHEN-SINK	-8.8856	1.0000	7	3.0304	1.0000	13	1.0777	eliminated	–
LOCAL COMBINATION	-9.0002	0.0002	21	2.7185	1.0000	5	1.0357	0.0042	7
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-8.2101	1.0000	4	3.3937	0.0202	16	1.1565	eliminated	–
OVERALL COMBINATION	-8.9664	1.0000	14	2.1672	1.0000	8	0.9529	0.0080	6

Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

The objective of the MCS test is to investigate the forecasting accuracy of an initial set of competing models by the help of a specific elimination algorithm. The elimination algorithm examines, at a given level of confidence,[19] which group of models survive. The poorly predictive model(s) are removed from the initial set of competing models, which are labelled as 'eliminated' in the below Tables. There are six different statistics for specifying the set of

[19] The significance level is defined as $\alpha = 0.2$.

superior models. We choose the range statistic among them according to suggestion by Hansen et al. (2003). The superior predicting models are selected according to the common sense of loss functions and MCS ranks. The point for interpreting the results is that the minimum values of loss functions are supposed to have higher p-values (unit p-value) and smaller rankings, indicating often superior predicting models. If the results of the loss functions and MCS totally contradict each other, the findings could be suspicious (but small changes do not necessarily matter).

4.Table 6

Out-of-sample 1-step-ahead rolling window forecasting and MCS results for FTSE (Window size:400)

FTSE 100	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-8.5838	0.0000	20	2.3436	0.0842	16	0.9087	eliminated	–
REGIONAL INFORMATION									
HAR-RV-GDAXI	-8.5947	0.4578	9	2.1983	1.0000	6	0.8341	eliminated	–
HAR-RV-FCHI	-8.4747	0.8848	4	1.8077	1.0000	7	0.7933	0.4756	5
HAR-RV-FTMIB	-8.5993	0.5268	6	1.8782	1.0000	10	0.8056	0.3862	6
HAR-RV-STOXX50E	-8.5166	0.1938	12	1.9507	1.0000	12	0.8369	eliminated	–
REGIONAL KITCHEN-SINK	-8.5866	0.4022	10	2.1549	1.0000	8	0.8318	eliminated	–
REGIONAL COMBINATION	-8.6082	0.7980	5	1.8041	0.0000	9	0.8028	0.3520	7
GLOBAL INFORMATION									
HAR-RV-SPX	-8.6003	0.0110	17	2.5611	0.0244	21	0.8773	eliminated	–
HAR-RV-DJI	-8.5810	0.1988	11	2.5441	eliminated		0.8943	eliminated	–
HAR-RV-IXIC	-8.6081	0.1052	13	2.4727	1.0000	15	0.8647	eliminated	–
HAR-RV-VIX	-7.7259	0.4862	8	1.7674	1.0000	4	0.7546	0.7306	4
HAR-RV-WTI	-8.5527	0.0002	19	2.4896	0.0286	20	0.9262	eliminated	–
HAR-RV-GOLD	-8.3386	0.0734	15	1.4750	1.0000	2	0.7434	0.9872	2
GLOBAL KITCHEN-SINK	-8.2100	0.4950	7	**1.1990**	**1.0000**	**1**	**0.7034**	**1.0000**	**1**
GLOBAL COMBINATION	**-8.6326**	**1.0000**	**1**	1.7744	1.0000	3	0.7568	0.9094	3
LOCAL INFORMATION									
HAR-RV-BOND	-8.5381	0.0912	14	2.4231	0.0410	19	0.9378	eliminated	–
HAR-RV-UKEPU	-8.5759	0.0004	18	2.2753	1.0000	14	0.8826	eliminated	–
HAR-RV-LIBOR	-8.5843	0.0000	21	2.3346	0.0774	17	0.9045	eliminated	–
LOCAL KITCHEN-SINK	-8.4818	0.0294	16	2.2989	0.0542	18	0.9177	eliminated	–
LOCAL COMBINATION	-8.5464	0.0000	22	2.2336	1.0000	13	0.8895	eliminated	–
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-6.6819	0.9970	2	1.8931	1.0000	11	0.8220	0.0036	9
OVERALL COMBINATION	-8.6358	0.9312	3	1.7861	1.0000	5	0.7878	0.1076	8

Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

4.Table 7

Out-of-sample 1-step-ahead rolling window forecasting and MCS results for GDAXI (Window size:400)

GDAXI	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-8.4787	0.0012	16	0.8228	0.0098	18	0.5974	0.0014	18
REGIONAL INFORMATION									
HAR-RV-FCHI	-8.4767	1.0000	15	0.8154	1.0000	16	0.5958	1.0000	17
HAR-RV-FTMIB	-8.4137	1.0000	3	0.8092	1.0000	12	0.5927	1.0000	12
HAR-RV-STOXX50E	-8.4664	0.0000	20	0.8161	1.0000	14	0.5927	1.0000	13
HAR-RV-FTSE	-8.4513	0.0000	21	0.8562	0.0080	19	0.6018	0.0002	22
REGIONAL KITCHEN-SINK	-8.4643	1.0000	11	0.8805	0.0004	22	0.6106	0.0004	20
REGIONAL COMBINATION	-8.4724	0.0002	19	0.8140	1.0000	17	0.5933	1.0000	15
GLOBAL INFORMATION									
HAR-RV-SPX	-8.4790	1.0000	12	0.9644	1.0000	9	0.6083	1.0000	11
HAR-RV-DJI	-8.4793	1.0000	13	0.9699	1.0000	10	0.6080	1.0000	14
HAR-RV-IXIC	-8.4794	1.0000	10	0.8669	1.0000	13	0.6046	1.0000	9
HAR-RV-VIX	-8.2381	1.0000	6	0.8190	1.0000	7	0.5993	1.0000	7
HAR-RV-WTI	-8.4327	1.0000	4	0.8437	0.0022	21	0.6056	0.0002	21
HAR-RV-GOLD	-8.3322	1.0000	2	0.7927	1.0000	8	0.5859	1.0000	8
GLOBAL KITCHEN-SINK	-8.1336	1.0000	14	0.9855	1.0000	11	0.6323	1.0000	10
GLOBAL COMBINATION	-8.4779	1.0000	9	0.7960	1.0000	5	0.5799	1.0000	3
LOCAL INFORMATION									
HAR-RV-BOND	-8.4723	0.0004	18	0.8010	1.0000	15	0.5963	0.0012	19
HAR-RV-VIXGDAXI	-7.8568	1.0000	5	**0.6301**	**1.0000**	**1**	**0.5520**	**1.0000**	**1**
HAR-RV-LIBOR	-8.4789	0.0006	17	0.8424	0.0056	20	0.6024	1.0000	16
LOCAL KITCHEN-SINK	-8.4076	1.0000	7	0.6415	1.0000	3	0.5593	1.0000	4
LOCAL COMBINATION	**-8.4837**	**1.0000**	**1**	0.6625	1.0000	2	0.5554	1.0000	2
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-8.3701	eliminated		0.8323	1.0000	4	0.6070	1.0000	6
OVERALL COMBINATION	-8.4672	1.0000	8	0.7546	1.0000	6	0.5744	0.0080	5

Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

Overall, the results presented in Tables 5, 6, and 7 strengthen the conclusion by a majority that the predictive accuracy of stock market volatility increases where we include the other exogenous volatilities. For instance, baseline HAR-RV models are outperformed by the HAR-RV-X models because the p-values of the baseline model for each loss functions are smaller than 0.2 and therefore dropped from the superior set of models. In particular, it is possible for this exercise to say that the integration of various volatilities, namely combination and Kitchen-Sink models, improves the forecast accuracy much better than the models with a single exogenous variable. However, the findings vary according to the sample stock markets and also the loss functions. In terms of our sample stock markets, the SPX and FTSE draw a parallel conclusion, but the GDAXI relatively gives a different result. For example, the winners of the SPX and FTSE stock markets come from the global information class (global combination and global Kitchen-Sink models respectively), whilst the outperformance of the GDAXI belongs to the local information class (local combination and/or HAR-RV-VIXGDAXI models). In more detail, the top three models of the SPX index are the HAR-GLOBAL-COMBINATION, HAR-VIX, and HAR-GLOBAL-KITCHEN-SINK. The worst model is the HAR-OVERALL-KITCHEN-SINK that includes 13 different exogenous variables in the model at the same time. A possible reason of the worst model is that including many exogenous variables in the same model at once is likely to cause the overfitting issue for the overall Kitchen-Sink model. In terms of the FTSE 100 index, the HAR-GLOBAL-KITCHEN-SINK, HAR-GOLD, and HAR-GLOBAL-COMBINATION are the best performing models respectively, whilst the worst one is not clear across the loss functions and the MCS procedure. Lastly, when we look at the GDAXI index, we see the ranking of top three models as the HAR-VIXGDAXI, HAR-LOCAL-COMBINATION, and HAR-LOCAL-KITCHEN-SINK. The worst model changes over the loss functions and the MCS results as it is in the FTSE index. Due to the variation of the results of our sample stock markets, the next paragraph evaluates and also discusses the findings of each stock markets in more detail one by one. In terms of the loss functions, the HMSE and HMAE give similar results whereas the QLIKE presents a mixed picture. The MCS test results that identify the set of the best models are mostly in the same direction with the loss functions. Therefore, it is worth to note here that we account for the predominant results by neglecting some small differences between the loss functions and the rankings of the MCS.

4.5.4. Findings in more details

In terms of the SPX index, the models of the global information class are capable of providing more accurate forecasts. There is no doubt that the best performing model is by far the global combination under all the three loss functions. An explanation as to why the global combination is the winner model lies to the fact that the volatility of the SPX stock market is not substantially affected by a single exogenous volatility, but it is influenced by the different kinds of volatilities (i.e. the members of global information class). Therefore, the global combination method seems to work well in the case of the SPX stock market. The second, third, and fourth best performing models are the VIX, global Kitchen-Sink, and IXIC respectively, but those three can be questionable and changeable between each other relative to the different loss functions and the MCS rankings.[20] For example, the loss functions point out that the IXIC is the second best model, whilst the MCS rankings suggest the VIX index as a second best model. Nevertheless, this is the fact that the winners are the members of global information class and therefore we discuss the predictive accuracy reasons of these models in the aforementioned order. The second best performing model is the model with the VIX index. The VIX index, which is also known as 'fear index', is defined as investors' expectations about future market volatility. The volatility of the next 30 calendar days is represented using the current prices of the SPX index options with a wide array of strike prices rather than stock options. Due to containing the future expectations of investors, the VIX is capable to give a quick sign before the stock market rally starts. Therefore, the impact of the VIX index on the SPX stock market volatility cannot be ignored. The global Kitchen-Sink method is the third best performing model for the SPX index, which underpins our argument that is about "superior forecasting performance of the integration of various exogenous volatilities". The SPX, IXIC, and DJI are the three most-followed US stock market indices in that sense the interrelationship between them is inevitable. On the one hand, the IXIC index is the fourth best performing model. On the other hand, in the global information class, the DJI is the only exception that yields comparable result with the baseline HAR-RV model. A plausible explanation of DJI's poor performance might be that it only includes thirty large-cap US companies and therefore considered an inadequate representation of the overall US stock market. Having said that the models of local and regional information classes for the SPX index present ambiguous results, which gives comparable results with the benchmark model. Moreover, some of them do not even improve the benchmark model such as the models of the local information. Lastly, while the overall combination model exhibits promising out-of-sample performance, the overall Kitchen-Sink model is outperformed by the baseline model. The overall Kitchen-Sink model includes a great number of exogenous variables in the same model at once that could cause the overfitting issue.

In terms of the FTSE, the winner is the global Kitchen-Sink model that includes all the members of the global group at once. The inclusion of the SPX, DJI, IXIC, VIX, WTI, and GOLD works better than the individual models because of involving much more information simultaneously. Afterwards, the global combination method that simply takes the average of all the individual global forecasts also exhibits a decent out-of-sample performance. In

[20] We consider the MCS rankings for this order because MCS rankings have some important features at selecting the superior set of models, which could be more reliable (see MCS section for those important features).

common with the SPX results, our argument that is about "higher forecast accuracy of the integration of different exogenous volatilities" is also consistent with the results of the FTSE index. Following the global Kitchen-Sink model, the second, third, and fourth best performing models for the FTSE index are the gold, the global combination, and the VIX respectively.[21] In individual basis, it is shown that the two members of global information class, namely the gold volatility index and the VIX contain useful information to help in forecasting the UK stock market volatility. Both of them are known as the two of the most important factors affecting the international stock markets. In terms of the relationship between the prices of gold and stock market, gold is often inversely correlated with the stock market; when stock markets decline, gold price rises. For this reason, during risk-on periods, investors will to diversify their portfolios more on gold in order to compensate their possible losses from stocks. In the case of stock market decline, the leverage effect is also an important matter to be taken into account because the negative returns are most likely to cause sharper spikes than the positive ones in volatility. From this point of view, the asymmetric downside losses of markets could be captured on the increasing price of gold if the gold volatility is added to the model as an exogenous variable. This could be one of the most important reasons why the gold volatility as an auxiliary variable adds one of the most valuable information at improving the predictive accuracy of the FTSE 100 stock market volatility. The VIX index also improves the forecast accuracy of FTSE volatility as it is in the SPX. This means that the US information has an important influence on the volatility of UK stock market. The SPX, DJI, and IXIC do not individually show a promising performance for the FTSE, but the global Kitchen-Sink model as a best performing model includes these three US stock markets. The overall Kitchen-Sink and combination methods for the FTSE index relatively have better forecast accuracy even though they are not between one of the top competing models.

The findings from the FTSE are compatible with the SPX, indicating the superiority of the global information against the local and regional information. In other words, the global information has a significant impact on the volatility of both the SPX and FTSE markets because both the indices consist of the largest internationally-focused companies. A financial implication can be that both the indices are highly sensitive to the global news and the integration of various global information can contain useful information to help in forecasting

[21]The FTSE winners might also be changeable between each other relative to the different loss functions and the MCS rankings as it is in the SPX index. However, we make this order according to the most predominant results in the Table.

the future volatility of the global markets. Only the difference between them appears in the rankings of the global information models. For example, while the FTSE suggests the best model as the global Kitchen-Sink method, the SPX shows that the global combination method is the winner. The regional information for the FTSE is the second best group after the group of the global information. The models of the regional information for the FTSE index (e.g. GDAXI, FCHI, and regional combination) have better performance compared to the models of the regional information for the SPX index. Germany and France are the prominent countries of the EU. Germany is the UK's second biggest export market after the US. France is a neighbour and also a major trading partner to the UK. The econometric findings of this work seem to be consistent with the real life economic relationships of the UK. In terms of the local information, we anticipate a more direct impact of UK data on the FTSE 100. However, the members of the local group show a mediocre performance.

The findings of the GDAXI index mainly point out the superiority of the local information class against the global and regional classes. From this aspect, the GDAXI index differs from the SPX and FTSE indices. However, it is hard to decide inside of the local class that either the local combination or the VIXGDAXI is the best performing model because the loss functions and also the MCS results do not seem very sure in between. Following that the third place is taken by the local Kitchen-Sink model. It is important to note that the VIXGDAXI index is the most important measure of GDAXI stock market volatility on which investors' expectations about the volatility of the next 30 calendar days. In terms of our study, the VIXGDAXI index is one of the main input to obtain the forecasts of the local combination and Kitchen-Sink methods. Therefore, we can intuitively say that higher forecast accuracy of the local group stems from the VIXGDAXI, which can be thought as a key factor for the outperformance. It is most likely true to say that if the VIXGDAXI index uninvolved in the local group, the local combination and Kitchen-Sink methods would not have superior performance as can be understood from the BOND and LIBOR's poor performance. Explicitly, the BOND and LIBOR show quite poor performance although the best group for the GDAXI index is the local information class. The reason might be that the GDAXI index consists of the thirty major German companies that does not necessarily represent the whole economy, implying that the index depends primarily on the situation of the 30 major German companies, not to the macroeconomic indicators such as the BOND and LIBOR. After superior performance of the local information group, the second best performing group is the global information. The gold, the VIX, and the global combination indicate a promising out-of-sample accuracy. An economic implication of this result for the GDAXI index is that major German companies tend to be affected by the global news as the structure of their business organisations are multinational. We believe that superior predictive accuracy of the local information class against the global and regional classes is obtained by the help of the VIXGDAXI index. As it is mentioned earlier if we did not employ the VIXGDAXI as a local exogenous variable, the global information class would prevail instead (as it is the case for the SPX and FTSE indices). It can be concluded that the indices that consist of major international companies hinge on more the global news in comparison with the local and regional news.

4.5.5. Cumulative HMSE value of baseline and winning models

The results of the loss functions and MCS provide convincing evidence that the HAR-RV-X model is the winner and always included in the set of the superior models, whereas the baseline HAR-RV model is never included in the set of the best performing models. It is, thus, worth to show the cumulative HMSE value of the benchmark and winning models whether the outperformance is consistent over the time or there are some periods where forecast errors diverge. The cumulative value is obtained by adding the observations of the HMSE series successively over the out-of-sample time period. The HMSE is a consistent criteria and its incremental value can exhibit divergence clearer compared to the other two loss functions. Figure 2 shows the cumulative value of the one-day-ahead forecast errors, based on the HMSE.[22] The models that produce the lowest cumulative HMSE values over time are the most accurate forecasting model for each indices.

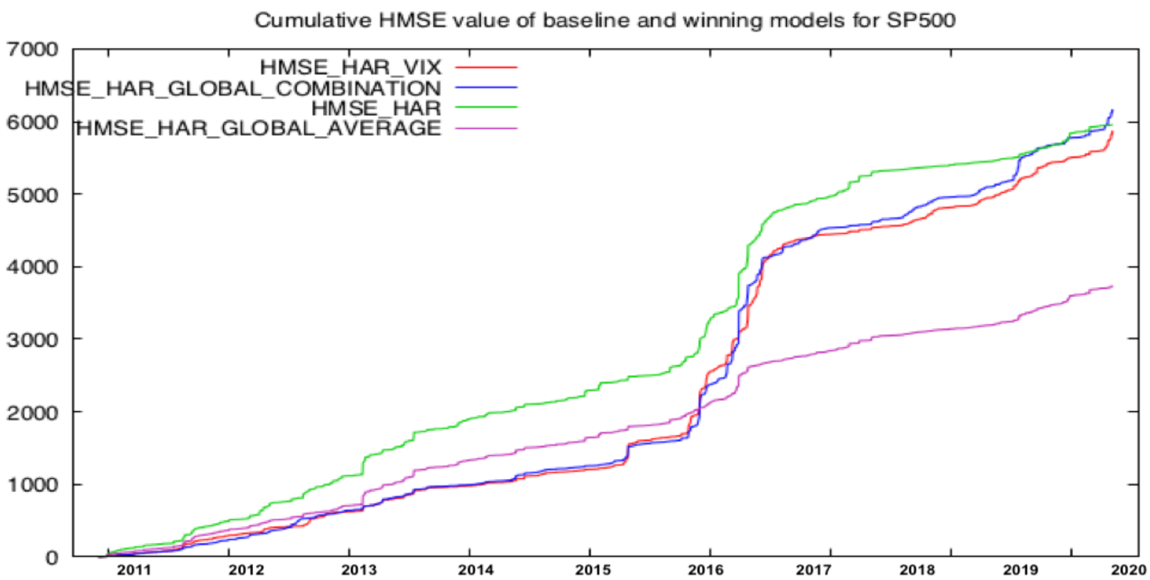
According to Figure 2, the cumulative HMSE values are not very different between the baseline and winning models from 2011 until 2016. However, they begin to show a clear shift in 2016 that lasts until the 2020. Therefore, it is worth to see what happened in the stock markets between 2016 and 2020. The starting point of fluctuation in 2016 where the cumulative HMSE differences between the winning and other models are sizeable for all the three indices corresponds to 2015-16 stock market sell-off period in global markets. This turbulent period includes the effects of the Chinese stock market crash (2015-16), the Greek debt default (mid-2015), oil prices' decline (early 2016), a dramatic increase of global bond markets (early 2016),

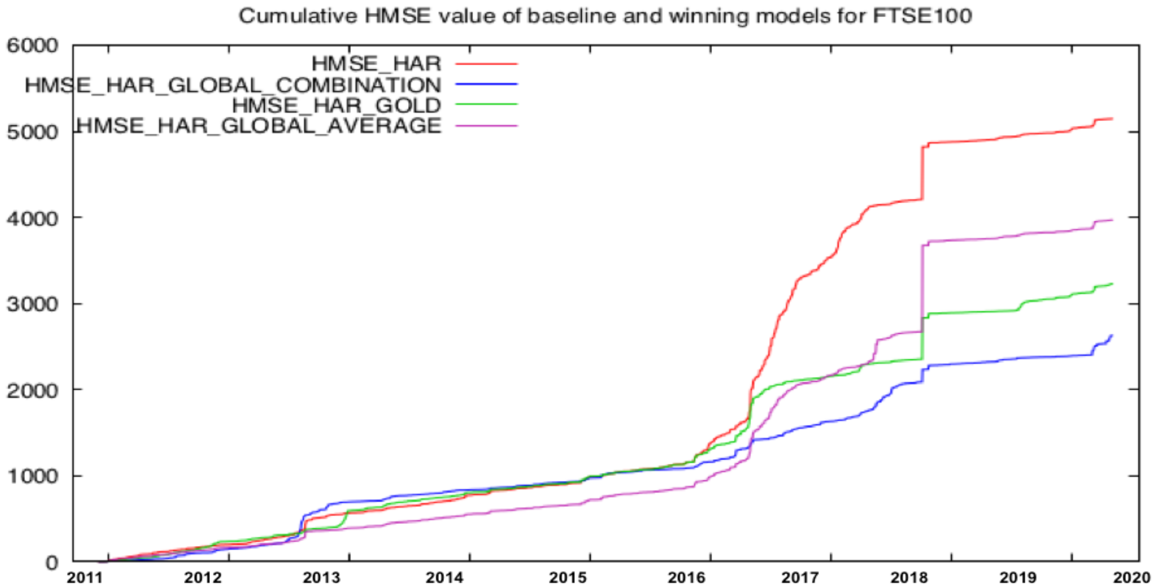
[22] In addition to these graphs, the cumulative HMAE value graphs are inserted in the Appendix (Figure 1), which the divergence is less clear than the HMSE graphs.

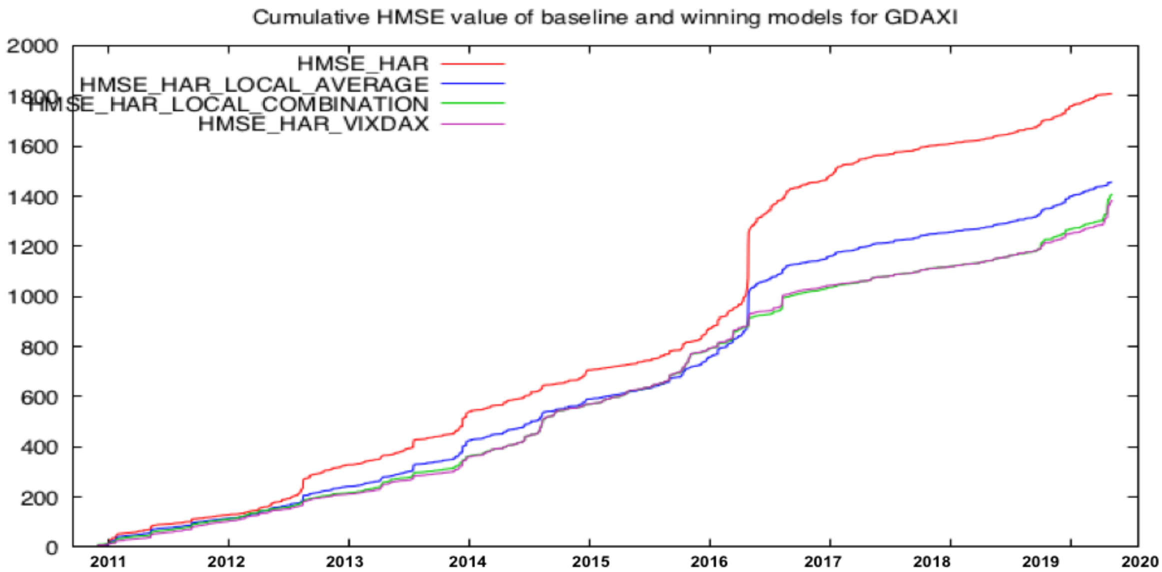
and finally Brexit referendum (June 2016). This study does not reflect the real effects of global pandemic due to the including only first couple of months of the pandemic.

The first graph (Figure 2) for the cumulative forecast error of SPX indicates a turbulent period in early 2016 due to 2015-16 stock market sell-off in global markets. Therefore, the baseline HAR-RV, HAR-RV-VIX, and HAR-RV-GLOBAL-KITCHEN-SINK models' forecast errors start rising dramatically, whereas the best performing model, HAR-RV-GLOBAL-COMBINATION, does not experience any shift throughout the turbulent period. This is to say that the HAR-RV-GLOBAL-COMBINATION model shows quite promising performance in the turbulent times compared to its counterparts. When we look at the second graph that is for the FTSE index, it has a substantial increase in June-2016 because of the Brexit referendum. Afterwards, another sharp increase appears in 2018 which is likely correspond to the impact of a possible no deal Brexit news and the Bank of England's report about recession warning. Unlike the SPX graph, the FTSE graph suggests that the HAR-RV-GLOBAL-KITCHEN-SINK model is more robust to the turbulent periods compared to the HAR-RV-GOLD and HAR-RV-GLOBAL-COMBINATION. On the other hand, the worst performing model is the baseline model. Lastly, the cumulative GDAXI graph also fluctuates with the Brexit vote and the winner is the HAR-RV-VIXGDAXI and/or HAR-RV-LOCAL-KITCHEN-SINK, which is same with our general results. Indeed, the cumulative HMSE differences could experience sizeable divergence between the baseline and winning models, but the best performing models for all the three indices seem to be robust despite the turbulent periods.

4.Figure 2 ***Cumulative HMSE graphs of baseline and winning models for SPX, FTSE, and GDAXI***







4.5.6. Further Robustness Check

Choosing an appropriate forecasting windows is important to the forecasting performance of the models. Rossi and Inoue (2012); Inoue, Jin, and Rossi (2017); Ma, et al. (2017), among others, discuss about the consequences of arbitrary choices of window sizes. However, this is the fact that there is no consensus about how to split the whole sample into in-sample and out-of-sample. Also, the accuracy of in-sample period does not necessarily mean having the same performance in the out-of-sample. In this regard, Wang, et al. (2016) and Ma, et al. (2017) point out that the predictive performance of in-sample changes over time and does not necessarily have an impact on the accuracy of out-of-sample. The out-of-sample performance of a model is more crucial to market participants than in-sample performance, this is because they are more interested in the predictability of future volatility. All in all, in this study, we arbitrarily choose the first 400 observations ($M = 400$) as an initial sample period and the rest of it is the out-of-sample, which is 2200 observations. Along with $M = 400$, we also choose two more alternative forecasting windows such as $M = 200$ and $M = 600$ in order to check the robustness of this study. The results of robustness check are given in the Appendix Table 1-6. Briefly, robustness check findings are consistent with the aforementioned conclusions even if there are some unimportant changes among the rankings that does not make a big difference in overall results. The small changes among the rankings can be summarised as follows. For instance, when we change the forecasting window from 400 to 600 for the SPX, the second and third best models are not clear, but the winner (HAR-RV-GLOBAL-COMBINATION) is still by far the best model. Similarly, if the window size (m=400) is switched to 200 for the FTSE, the HAR-RV-REGIONAL-COMBINATION prevails as the second best model. In terms of the GDAXI, it is still difficult to choose the best model between the HAR-RV-VIXDAX and HAR-RV-LOCAL-COMBINATION despite alternative forecasting windows.

4.6. Conclusion

The aim of this study is to investigate the importance of exogenous volatilities at improving the forecasting accuracy of stock market volatility. In this regard, we examine a wide range of exogenous volatilities whether they improve the predictability of stock market volatility or not. To evaluate the forecasting results from a different perspective, we classify our exogenous volatilities according to different information channels, namely local, regional, and global. Using our specific classification method we attempt to find out which class of models best helps in forecasting stock market volatility. Given the outperformance of the global information over local and regional information, the results are informative to reveal the dynamics of stock markets.

We conduct this volatility forecasting exercise on the SPX, FTSE, and GDAXI stock markets using 'HAR-RV model', but adding a range of different exogenous variables from local, regional, and global information. The time span is from 01 July 2009 to 28 May 2020. In the HAR-RV-X model, exogenous variables are used in various forms including, for example, each individual exogenous variables separately, forecast combinations, and Kitchen-Sink approach. One-day-ahead out-of-sample volatility forecasts are generated using the rolling window forecasting technique and the QLIKE, HMSE, and HMAE criteria measure the forecast losses. For further analysis, we carry out the MCS procedure, forecasts' robustness checks, and the cumulative HMSE difference of the baseline and winning models. In conclusion, we find several noteworthy points. First and foremost, the predictive accuracy of stock market volatility increases where we include the other exogenous volatilities. For instance, baseline HAR-RV models are outperformed by the HAR-RV-X models because the p-values of the baseline model for each loss functions in the MCS procedure are smaller than 0.2 and therefore dropped from the superior set of models. Second, the integration of various volatilities, namely Kitchen-Sink and combination models, gives much better forecasting performance than the models with a single exogenous variable. Third, the global information exhibits superior performance compared to the local and regional information in our sample. The winners of the SPX and FTSE stock markets come from the global information class that are global combination and global Kitchen-Sink models respectively, but the outperformance of the GDAXI belongs to the local information class, which are local combination and HAR-RV-VIXGDAXI models.

Several practical implications from our results can help policy makers, investors, and market participants in the process of both forecasting and asset allocation. Notably, all of them want to know today that what will be the degree of volatility tomorrow? From the forecasting perspective, the gain in forecasting performance is supposed to be economically significant to minimize risk and maximize return. For example, investors and market participants would like to sell their stocks or portfolio before financial markets become too volatile. In a similar vein, policy makers tend to adjust the bid-ask spread less wider to restore market liquidity if the future is expected to be more ambiguous. In terms of portfolio optimization, this study could be beneficial for investors and market participants to align their portfolios by reducing their exposure to various risks (i.e. local, regional, and global risks). This is because, in this work, it is carried out by the inclusion of exogenous variable(s) to the HAR-RV model and the results reveal the degree of interrelation between markets. For future research, the same methodological framework would be applied to other stock indices, including, for example, Asian and/or Middle East stock market indices.

In the end, employing different market data as an exogenous variable the HAR-RV-X model does enable us to analyse the interrelation between markets, which makes our study on the edge of volatility spillover literature. To complement this chapter, the next chapter explores the spillover effects of international financial markets from another perspective.

CHAPTER 5

Complex Network Analysis of Volatility Spillovers between Global Financial Indicators and G20 Stock Markets[23]

Abstract

This paper analyses the dynamic transmission mechanism of volatility spillovers between key global financial indicators and G20 stock markets. To examine volatility spillover relations, we combine a bivariate GARCH-BEKK model with complex network theory. Specifically, we construct a volatility network of international financial markets utilising the spatial connectedness of spillovers (consisting of nodes and edges). The findings show that spillover relations between global variables and G20 markets varies significantly across five identified sub-periods. Notably, networks are much denser in crisis periods compared to non-crisis periods. In comparing two crisis periods, Global Financial Crisis (2008) and Covid-19 Crisis (2020) periods, the network statistics suggest that volatility spillovers in the latter period are more transitive and intense than the former. This suggests that financial volatility spreads more rapidly and directly through key financial indicators to the G20 stock markets. For example, oil and bonds are the largest volatility senders, while the markets of Saudi Arabia, Russia, South Africa, and Brazil are the main volatility receivers. In the former crisis, the source of financial volatility concentrates primarily in the US, Australia, Canada, and Saudi Arabia, which are the largest volatility senders and receivers. China emerges as generally the least sensitive market to external volatility.

---

[23] A shorter version of this chapter is published in a refereed academic journal. Reference: Korkusuz, B., McMillan, D.G. & Kambouroudis, D., (2022). "Complex network analysis of volatility spillovers between global financial indicators and G20 stock markets". *Empirical Economics, pp. 1-21*.

5.1. Introduction

A key concern for investors in seeking to build diversified portfolios are fluctuations in the relations between assets. Further, that fluctuations in a market do not arise solely from internal information but are also affected by external information. The transmission of this external information across markets is known as 'volatility spillover effects' (see, for example, Yu et al., 2015; Rejeb and Arfaoui, 2016; Mensi et al., 2018). This phenomenon is also referred to as fear connectedness by Diebold and Yilmaz (2014).

In the last two decades, the Global Financial Crisis (GFC, 2008) and the Covid-19 Crisis (CVC, 2020) evidence the importance and impact of spillover effects across financial markets. Here, the transmission of volatility across markets is likely to have a profound impact on each economy, varying with the degree of market integration. In the case of such crises, increasing globalisation and financialization of markets allows adverse effects in one market to further intensify existing spillover effects. Consequently, during such crises, investors typically sell-off risky assets on fears of financial contagion that results in a further spread of global risk. Therefore, there is a need to understand and model time-varying spillovers among a range of markets and how this could impact future investment behaviour.

In the spillover literature, there is a lack of empirical research that provides international evidence conducted on a large scale. The existing literature typically concentrates on spillovers between a small number of stock markets, often classified according to their level of development (e.g., emerging or developed) or considers a single stock market with several other assets (see, among others, Zhang et al., 2020; Golosnoy et al., 2015; Piljak and Swinkels, 2017; Yoon et al., 2019; Zhang et al., 2020). As a larger dataset, the G20 stock markets may be thought as a convenient research object to study volatility spillover effects as it account for over 85% of Gross World Product and over 80% of world trade. Thus, this bloc captures large changes to the world economy (Liu et al., 2017; Zhang et al., 2020).

While a range of spillover models exist (e.g., stochastic volatility, Diebold and Yilmaz (2009, 2012) spillover index), an advantage of the GARCH-BEKK model is the ability to capture spillover effects as volatility can be directly computed from the variance-covariance matrix, without imposing any restriction on the conditional correlation structure (Lee et al., 2014). However, when examining the interrelations of multiple series, such as the G20, the GARCH-BEKK is not able handle the high multi-dimensional spillovers of the system as a whole. Therefore, a small number of studies (Liu et al., 2017; An et al., 2020; Zhang et al., 2020) combine the econometric model with complex network theory to construct a network of financial markets. The advantage of this combined approach is to provide a solution to the difficulty encountered by the GARCH-BEKK model when dealing with multi-dimensionality and to provide a visualisation of the complex financial system in a clear way.

To address the identified gaps in the literature, this paper incorporates key global non-stock assets (oil, gold, and bonds) into the G20 market network and extends the analysis to include the Covid-19 period. This allows for a wider perspective including the effect of assets that are considered as both safer in comparison to stocks (e.g., gold and bonds) as well as alternative risky ones (e.g., oil), including across a period of heightened global risk.

The aims of this paper are twofold. First, to compare the network nature of spillovers across the GFC and CVC periods. This will allow a better understanding of the impact of both crises, where no such analysis currently exists. Second, to consider the source of volatility spillovers between G20 stock markets and key global financial assets over different periods.

The main findings and their economic implications can be noted as follows. Volatility spillovers between global financial indicators and G20 stock markets in all sub-periods are significant and exhibit time-variation, with a high level of market interrelation. Importantly, volatility spillover networks are denser during the crisis periods of the GFC and CVC compared to non-crisis period networks. The implication of this finding demonstrates that crisis periods are more transitive in terms of volatility spillovers, which causes volatility to spread more rapidly through major financial indicators to G20 stock markets. These results should be of interest to investors seeking to diversify their portfolios across both asset types, including oil, gold, and bonds, and G20 stock markets. Moreover, the nature of diversification depends on both time and market specific information. Notably, during crisis periods, correlations between all markets increase and diversification opportunities becomes more restricted.

This chapter is organised as follows: ***Section 2*** presents the review of related literature. In ***Section 3***, the methodology is explained in more detail. Afterwards, we give the data description in ***Section 4***. The empirical results and their evaluations are given in ***Sections 5***. Finally, the summary and conclusion are presented in ***Section 6***.

5.2. Literature Review

The transmission of fluctuations across markets is known as 'volatility spillover effects' (see, for example, Yu et al., 2015; Rejeb and Arfaoui, 2016; Mensi et al., 2018). This phenomenon is also referred to as fear connectedness by Diebold and Yilmaz (2014). This is because any fluctuations in the financial markets usually reflect investors' fear. In the last twenty years, the

Global Financial Crisis (GFC, 2008) and the Covid-19 Crisis (CVC, 2020) show the importance and impact of spillover effects across financial markets. In times of crisis, this is the fact that the transmission of volatility across financial markets is likely to have a profound impact on each economy, varying with the degree of market integration. Moreover, increasing globalisation and financialization of markets allows adverse effects in one market to further intensify existing spillover effects. Thus, during such crises, investors typically sell-off risky assets on fears of financial contagion that results in a further spread of global risk. Therefore, there is a need to understand and model time-varying spillovers among a range of markets and how this could impact future investment behaviour.

When we broadly review the literature of volatility spillover effects, most of the relevant studies mainly concentrate their efforts on the emerging markets (e.g. Erten et al., 2012; Maghyereh and Awartani, 2012), developed markets (Golosnoy, Gribisch, and Liesenfeld, 2015), commodity markets (Yoon, Mamum, Uddin, & Kang, 2019), and also sovereign bond markets (Piljak & Swinkels, 2017). To sum up the common point of all, they examine the spillover relationship between several stock markets or just one stock market with other assets. However, capital is mobile across borders and also assets. Therefore, there is a lack of extensive research to provide evidence that whether there is a spillover transmission path between some global assets such as oil, gold, and bond and a wide choice of international stock markets or not. At that point, there is a gap in the existing literature and this work fills the gap in the literature by investigating the spillover relationship between some major assets and G20 stock markets. The reason why we select the G20 markets is that the 85 per cent of Gross World Product and 80 per cent of the World Trading are explained by the G20 countries in 2014 that may represent large changes in the world economy. Therefore, the G20 markets seem to be a convenient research object to explain the markets' spillover effects in the global scale. In a similar vein, the fluctuations of oil, gold, and bond markets already have a significant direct impact on the global economy (especially for stock markets) that motivates us to study the spillover relationship between them in a broader perspective.

In terms of the econometric methodology, the three classes of models in the literature are widely-used to measure the spillover effects. The first model is Stochastic Volatility (SV) model in which the volatility is determined by an unobservable random process. Even if it appears to be a decent method in financial econometrics, it is quite difficult to estimate the parameters of the SV model that makes its usage hard in practice (e.g. Keating & Valcarcel, 2015; Zhang and Zhuang, 2017). The second is the multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) family model. The third model is the Diebold and

Yilmaz (DY) method (2009, 2012) that is based on forecast error variance decompositions in a VAR framework (applied by some recent studies, see, amongst others, Prasad et al., 2018; Caloia et al., 2018).

The MGARCH models are good at capturing the spillover effects as Tse (1999), who is one of the earliest researchers, documents it using a bivariate EGARCH model and find a significant bidirectional volatility transmission between the DJI (Down Jones Industrial Index) and the index futures markets. Afterwards, the GARCH-BEKK model, which was developed by Engle and Kroner (1995), has been applied by many researches (see, among others, Kang et al., 2013; Liu et al., 2017; Zhang et al., 2020; Weiping et al., 2020) to extract the spillover relationships between financial markets. This model is a bi-variate model that can capture volatility spillovers between pairwise markets. In considering multivariate-GARCH models there is a range of alternatives including the CCC (constant conditional correlation; Bollerslev, 1990), DCC (dynamic conditional correlation; Engle, 2002) and the GARCH-VECH (Bollerslev, Engle, and Wooldridge, 1988) models in addition to the GARCH-BEKK. A key advantage of the GARCH-BEKK model is that it does not impose any restriction on the conditional correlation structure between series. In addition, the conditional variances are restricted to ensure they are positive definite, while reducing parameter dimensions. However, it is suggested to use bi-variate form of the GARCH-BEKK model. This is because if a new variable is added into the model, the number of model parameters increases significantly. This issue is worse in terms of the full GARCH-VECH specification as the VECH inherently requires a larger number of parameters than the BEKK to be estimated and this can often lead to non-convergence. The full VECH-GARCH model of Bollerslev et al. (1988) is rarely used in the literature because the number of free parameters increases very fast with the number of variables. In the case of bi-variate VECH-GARCH specification, the model needs to generate 23 model parameters, which is a large number of free parameters for non-linear estimation. The restricted form of VECH or in other words diagonal VECH-GARCH model might be a better alternative to the unrestricted one but the restricted form of VECH does not generate cross-product (spillovers) parameters. This means that the direction of volatility transmission from one market to another cannot be extracted. Therefore, the VECH specification is rarely used and Bollerslev et al. (1988) did not estimate this model in their applications. Similarly, the CCC and DCC GARCH models also do not capture spillover effects from one market to another. Instead, both of the models do extract the magnitude of total spillovers between pairwise markets using econometrics techniques (e.g. see, method of Diebold and Yilmaz, (2009, 2012)). To sum up, except BEKK-GARCH model, the other multivariate GARCH specifications seem not to be suitable for detecting the direction of volatility spillovers.

Compared to the DY method (2009), one of the most important advantages of using the GARCH-BEKK model is that the volatility can be computed directly from the variance covariance matrix that gives us a compelling reason to choose this model. Another advantage of the model is that the GARCH-BEKK does not force any restrictions on the conditional correlation structure between the model variables and also ensures the positive definite matrix. However, the weakness of the model is that it does not give us much more flexibility to choose higher lag structure and more than two variables. Choosing higher lags and/or more than two variables requires to estimate much higher number of model parameters in the model, which could end up with the crash of the model or some calculation issues. Fortunately, a seminal paper of Hansen and Lunde (2005) suggests the choice of first lag phase as optimal for modelling the conditional variance and covariance and therefore we follow them.

When the high dimensional interrelationship in financial markets is considered, the models mentioned above could not solely be enough to reveal the multi-dimensional spillover relationship of financial system as a whole. Therefore, some previous studies (see Liu et al., 2017; An et al., 2020; Zhang et al., 2020) combine the econometric method with the complex network theory in order to construct a network of financial markets. In the financial field, the advantage of this combined method is to be able to provide a solution to the difficulty of the GARCH-BEKK model at dealing with the multi-dimensionality and also provide the visualisation of complex financial system in a simple way. In the most relevant literature, Liu et al. (2017) combine the GARCH-BEKK model with complex network theory to investigate the volatility spillover network in the G20 stock markets. They find that volatility originating in one market rapidly spreads through the network, with the largest volatility receiver and sender being South Korea and Brazil. Weiping et al. (2020) extend this study and divide the G20 stock markets into four different spillover blocks. They highlight the impact of higher tariffs (mid-2018) imposed by the US government on other G20 markets and its economic impact. Both studies note that the interconnectedness of markets peaks during the global financial crisis in 2008, while evidence from the Covid-19 period is not included in the sample period. A further drawback of these studies is that capital is mobile not only across both borders but also asset types. This means that in addition to the G20 markets, other well-established global assets cannot be ignored. For example, this could include evidence of interdependence between stocks and assets, including, oil, gold, and US 10-year Treasury bonds as key indicators of global risk.

Some other relevant studies can be summarised as follows. Kang et al. (2013) investigate the volatility spillover relationship between the spot and futures markets of the KOSPI index employing a bivariate GARCH-BEKK model in the context of high frequency data. They find a strong bidirectional causal relationship between spot and futures markets. Similarly, Kang and Lee (2019) analyse the dynamic spillovers between the global futures markets and find that the highest level of spillover occurs during the global financial crisis in 2008. They also point out that the FTSE 100 index futures is the largest transmitter of the spillover shocks. An et al. (2020) also study the spillover effects among crude oil prices. Their angle is different from the previous works as such; they examine a short term rolling window and its dynamic evolution process of spillover effects. In this way, they identify the most important bridges of network which are found in their study as WTI and Brent. Feng et al. (2018) do not only built the volatility spillover network, but also designed a research framework using the wavelet method for the spillover relationships of Chinese sectors. Zhang, Zhuang, Lu, and Wang (2020) use factor analysis approach to study the spatial linkage of volatility spillovers and its explanation in the G20 stock markets. They introduce the quadratic assignment procedure to identify the major factors which have an impact on the spatial spillovers. They point out that the developed markets are more influential compared to the emerging ones throughout the turbulence periods, whereas emerging markets are more sensitive to the shocks compared to developed markets during any periods. Using the GARCH-BEKK model combined with the network approach Wang, Gao, An, Tang, and Sun (2019) try to capture the spillover relationships of energy stocks and identify influential energy stocks. They find that the top ten influential energy stocks are belong to the industry of power and utilities, in more detail, the most influential sub-industries are petroleum exploitation and petroleum processing. As well as the GARCH-BEKK model-based network approach, some other pioneering works use that combined approach for constructing the Granger-causality network (Billio et al., 2012; Baumohl et al, 2018), the variance decomposition frame-based network (Diebold and Yilmaz, 2014), the tail-risk spillover network (Hardle, Wang, and Yu, 2016), and an extreme risk network (Wang et al., 2017).

If we summarize the existing literature in a few words, it can be said that the volatility spillover has a time-varying feature. This means that the spillover is time and market specific, meanwhile this is the fact that any global information is likely to affect each markets with the increasing globalisation of financial markets (Cardona, Gutierrez, and Agudelo, 2017). In this regard, the purpose of this work is to contribute onto the existing literature by uncovering the spillover relationship between global financial barometers and G20 markets in the network framework, whereas the previous literature covers the spillover relationship between a small number of stock markets or just one stock market with other assets. Furthermore, establishing the spillover networks in different time periods than the previous literature the two most important crisis periods (GFC-2008 and CVC-2020) are compared to each other from the spillover network perspective.

5.3. Empirical Methodology

The first step in the methodology is to extract the volatility spillovers and second, to construct the network and associated statistics.

5.3.1. Spillover Extraction and Construction

Step 1: Synchronize pairwise closing prices and calculate the log return series

The trading days across the G20 stock markets differ between market pairs. Therefore, we need to find the intersection between market closing prices. To obtain the common trading days for any two markets, we synchronize the pairwise closing prices for which both markets have active trading. Having conducted the pairwise synchronization, logarithmic returns are calculated by the following equation (1):

$$R_{i,t} = [\ln(P_{i,t}) - \ln(P_{i,t-1})] * 100 \quad (1)$$

where $P_{i,t}$ is the closing price of stock index for country i at time t and $P_{i,t-1}$ is the closing stock price index for country i at time $t-1$.[24]

Step 2: Employ the bivariate GARCH-BEKK model to extract volatility spillovers

To capture volatility spillovers between pairwise markets, we apply the GARCH-BEKK model of Engle and Kroner (1995). In considering multivariate-GARCH models there is a range of alternatives including the CCC (constant conditional correlation; Bollerslev, 1990), DCC

[24] An alternative approach to the modelling outlined in the section is to use a VAR for the whole network and follow the general method of Diebold and Yilmaz (2014) and, more specifically, the adapted approach of Demirer et al. (2018). However, to use a single VAR would mean matching all series simultaneously (rather than pairwise) and result in a significant loss of data (approx. 45% of original the data would be lost).

(dynamic conditional correlation; Engle, 2002) and the GARCH-VECH (Bollerslev et al., 1988) models in addition to the GARCH-BEKK. A key advantage of the GARCH-BEKK model is that it does not impose any restriction on the conditional correlation structure between series. In addition, the conditional variances are restricted to ensure they are positive definite, while reducing parameter dimensions. In contrast, the CCC and DCC models do not capture spillover effects from one market to another, while the GARCH-VECH requires a large number of parameters (23) to be estimated and this can often lead to non-convergence. Furthermore, the restricted form of the GARCH-VECH, known as the diagonal GARCH-VECH model, may have better convergence properties but it does not generate cross-product (spillovers) parameters. As the GARCH-BEKK does not suffer from these issues and is used within the cogent literature (see, An et al., 2020; Liu et al., 2017; Weiping et al., 2020; Zhang et al., 2020) we conclude that the GARCH-BEKK model provides the preferred estimation approach. We use a single lag for the sake of parsimony.

For the GARCH-BEKK model, the mean equation is given as:

$$R(t) = \begin{bmatrix} R_1(t) \\ R_2(t) \end{bmatrix} = \begin{bmatrix} \mu_1(t) \\ \mu_2(t) \end{bmatrix} + \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix} \begin{bmatrix} R_1(t-1) \\ R_2(t-1) \end{bmatrix} + \begin{bmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{bmatrix} \tag{2}$$

where R_t is the logarithmic return that is a (2×1) vector of market 1 and market 2 at time t, $\mu_1(t)$ and $\mu_2(t)$ represent the long-term drift coefficient, and then $\varepsilon_1(t)$ and $\varepsilon_2(t)$ are the random errors at time t. The variance-covariance equation is then given by:

$$H(t) = C'C + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'H_{t-1}B \tag{3}$$

$$C = \begin{bmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad and \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{4}$$

where $H(t)$ is the conditional covariance matrix (2×2). C is the constant coefficient terms in the form of a lower triangular matrix. The coefficient matrices of the GARCH-BEKK model are given by A and B. A represents the parameters of the conditional residual matrix, and B is the conditional covariance matrix's parameters. The diagonal elements of A and B such as $a_{11}, a_{22}, b_{11}, and\ b_{22}$ measure their own markets' previous shocks (ARCH effect) and volatility (GARCH effect), while the off-diagonal parameters in the matrices $(a_{12}, a_{21}, b_{12}, and\ b_{21})$ quantify the cross-stock market effects of shocks and volatility between stock markets i and j, in other words, the volatility spillovers. The bivariate GARCH-BEKK model is estimated with 17 parameters in total, which does not impose any restriction on the conditional correlation structure between the model variables.

The GARCH-BEKK model is estimated using the maximum likelihood method, with the conditional log likelihood function written as.

$$L(\theta) = -T\, ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}[\, \ln|H_t(\theta)| + \varepsilon_t(\theta)' H_t^{-1} \varepsilon_t(\theta)\,] \quad (5)$$

where T is the number of observations and θ is the vector of parameters to be estimated. The influence of volatility spillover from stock market i to stock market j is calculated as the absolute sum of the off-diagonal values of matrices A and B, which can be defined as follows.

$$e_{1,2} = |a_{12}| + |b_{12}| \quad (6)$$

where a_{12} and b_{12} represent the off-diagonal elements of matrices A and B, respectively. $e_{1,2}$ stands for the magnitude of volatility spillover effect from market 1 to market 2. Similarly, $e_{2,1}$ is the size of volatility spillover from market 2 to market 1. $e_{1,2} + e_{2,1}$ is the total size of bidirectional volatility spillover relation between markets 1 and 2.

STEP 3: Construct the spillover network considering complex network construction rules

Complex network theory takes into account the relations among different parts of a real complex financial system as a network (Hao et al., 2015; An et al., 2014). A complex network is a collection of nodes that are connected by edges. A complex network system is symbolised in equation (7):

$$G=(N, E) \quad (7)$$

where G represents a complex network. N refers to the set of nodes and E is the set of edges between nodes. In the context of this study, our complex networks have some characteristic features such as the small world effect and the superposition phenomenon. The small world effect (Watts and Strogatz, 1998) is a phenomenon in network theory that no node is independent from the network, which means that all nodes are linked to each other either with a direct or indirect tie. Two widely used network statistics, which are the average shortest path length and average clustering coefficient are used to determine the small world effect. Another network characteristic is the superposition phenomenon, which is a principle in physics applying all the linear systems such as height in a water wave, intensity of a light wave or pressure in a sound wave. For example, where two water waves are travelling in opposite directions, the size of combined wave is the sum of both water waves at the intersection point. Similarly, the thickness of an edge between pair nodes is identified by the superposition principle in this work.

To build a complex network, we prepare the nodes (as data frame) and edges (as matrix element) in the form of a square matrix whose main diagonal consists of zeros. Thus, the square matrix of spillover relation for the G20 stock markets can be shown by equation (8):

$$M = \begin{bmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n} \end{bmatrix} \tag{8}$$

where M is the matrix of edges, which creates the complex network. Here, we consider the G20 stock markets and other global assets as the nodes. In a similar vein, the volatility spillovers (e) are considered as unidirectional or bidirectional edges between the nodes. Using the off-diagonal elements of the GARCH-BEKK model, we create a complex network of the global financial system. If the estimated off-diagonal parameters of the GARCH-BEKK are significant, it means that there is a volatility spillover effect from market 1 to market 2 and then the spillover value ($e_{1,2}$) from market 1 to market 2 is entered into the specified cell of matrix.[25] The direction of spillover effect is shown with an arrow mark from market 1 to market 2. If there is also reverse spillover relation (market 2 to market 1), the spillover is bidirectional and arrow marks appear on both edges. In identifying the spillover relations, previous studies (Liu et al., 2017; Weiping et al., 2020; Zhang et al., 2020) use the 10% significance level. They argue, this is because a too strict significance level may miss important spillover relations.

---

[25] If the estimated parameters are insignificant, no spillover relation between pairwise markets and enter 0 into the specified matrix element.

5.3.2. Spillover Network Statistics

Weighted in-degree and weighted out-degree

The weighted degree of the spillover network indicates how strong the effects of the volatility spillovers are. If a node has a larger weighted degree, it has more potential to affect or be affected, depending on the weighted in/out degree, where the weighted-in-degree receives spillovers and the weighted-out-degree causes spillovers. The weighted in and out degrees are expressed by equations (9) and (10) as follows:

$$w_i^{in} = \sum_{j=1}^{m} e_{ij} \tag{9}$$

$$w_i^{out} = \sum_{j=1}^{n} e_{ji} \tag{10}$$

where w_i^{in} and w_i^{out} represent the weighted in-degree and weighted out-degree, respectively. e_{ij} implies the size of volatility spillover from node i to node j and e_{ji} implies the amount of volatility spillover from node j to node i. e is obtained by equation (6). m and n denote the number of edges that node i has with the other nodes of network. A higher weighted degree means stronger volatility spillover relation in a complex network.

Average shortest path length and network diameter

In a complex network, the small world phenomenon (Watts and Strogatz, 1998) means that all nodes are linked either with a direct or indirect tie. For detecting this phenomenon Watt and Strogatz (1998) suggest looking at both the average shortest path length and network diameter statistics. The average shortest path length is defined as the mean of the shortest steps of volatility spillover propagation from node i to node j. It is expressed by equation (11):

$$r = \sum_{i,j} \frac{d_{ij}}{n(n-1)} \quad ; \quad (i \neq j) \tag{11}$$

where r denotes the average shortest path length; the smaller r in a network, the more linkages between nodes, which indicates a denser network (a larger r denotes less linkages between nodes and a looser network). d_{ij} is the shortest distance from node i to node j. The denominator, $n(n-1)$, shows the maximum number of possible edges in the spillover network where n is the total number of nodes. The network diameter is 'the shortest path between the two most distant nodes of the network'. Once the shortest path length of every single node with respect to the other nodes is calculated, the network diameter is the longest one among all the calculated shortest path lengths. The smaller or larger value of diameter is interpreted same as the shortest path length.

Graph density and average clustering coefficient

The graph density indicates how close the number of edges is to the maximum number of possible edges in the network. If the graph density is equal to 1, the network is called a complete graph and includes all the possible edges. The network graph density can be calculated by equation (12) as follows:

$$D = \frac{2|E|}{|n|(|n|-1)} \tag{12}$$

where E is the number of edges between the nodes of network and n is the total number of nodes. This measure amounts to a ratio of actual connections to potential connections. Another network statistic is the average clustering coefficient, which is a similar measure and shows how the nodes are integrated in the network graph. It is calculated by dividing the number of edges connecting a node's neighbours to the total possible number of edges between the node's neighbours.

5.4. Data Description

The G20 is an international forum, consisting of 19 major developed and emerging countries (with the EU as a whole also represented). From a global perspective, the G20 accounts for 85% of Gross World Product and 80% of world trading (2014) and thus, represents an important bloc. However, capital is mobile across asset types, such that other (global) assets should be considered alongside stocks. Therefore, in examining spillovers, we also include oil, bonds and gold market information.

Specifically, we employ the daily closing prices of 19 major stock market indices of the G20 countries; which are S&P MERVAL (Argentina), S&P ASX200 (Australia), BOVESPA (Brazil), GSPTSE (Canada), SHCI (China), CAC40 (France), DAX30 (Germany), SENSEX (India), IDX (Indonesia), Italy 40 (Italy), NIKKEI225 (Japan), KOSPI (Korea), S&P BMV (Mexico), MOEX (Russia), TADAWUL (Saudi Arabia), SOUTH AFRICA TOP40 (South Africa), BIST100 (Turkey), FTSE100 (UK), and SP500 (US).[26] Additionally, crude oil (WTI; West Texas Intermediate), gold returns and the US 10-year Treasury bond yields are included. All the data is extracted from the 'investing.com' website, at the daily frequency and over the sample period between January 8, 2003 and January 4, 2021, which includes both periods of calm and turmoil.

5.Table 1 ***Summary Statistics***

	Mean	Std. Dev.	Skew.	Ex. Kurt.	JB Stat.	ADF	Q(20)	ARCH(1)
Argentina	0.106	2.309	-1.398***	21.537***	82701.4***	-23.549***	41.597***	21.453***
Australia	0.016	1.061	-0.721***	8.408***	13811.1***	-48.044***	49.558***	577.262***
Brasil	0.054	1.777	-0.376***	7.705***	10686.1***	-14.432***	58.796***	540.84***
Canada	0.020	1.107	-1.123***	21.120***	84301***	-12.453***	148.653***	757.594***
China	0.031	1.947	-0.379***	5.954***	4555.17***	-7.823***	53.765***	53.125***
France	0.012	1.394	-0.169***	8.958***	15022.3***	-24.047***	49.609***	222.499***
Germany	0.033	1.401	-0.233***	7.260***	9885.87***	-67.336***	25.465	85.511***
India	0.061	1.454	-0.264***	10.454***	19474.8***	-17.628***	83.539***	211.144***
Indonesia	0.063	1.334	-0.530***	7.308***	9636.64***	-59.179***	63.849***	206.137***
Italy	0.042	0.748	-0.605***	1.735***	177.9***	-33.096***	14.030*	9.372***
Japan	0.028	1.508	-0.202***	10.531***	19474.3***	-23.729***	52.068***	230.224***
Korea	0.035	1.303	-0.455***	7.432***	10161.6***	-12.961***	45.727***	338.596***
Mexico	0.044	1.227	-0.204***	7.157***	9317.0***	-15.878***	63.192***	150.833***
Russia	0.054	1.950	-0.314***	20.147***	73640.3***	-10.338***	104.100***	73.716***
S. Arabia	0.039	1.895	-1.196***	13.869***	25038.7***	-14.263***	71.399***	129.999***
S. Africa	0.043	1.358	-0.162***	5.692***	5744.59***	-15.203***	47.441***	257.982***
Turkey	0.059	1.675	-0.314***	3.997***	2972.02***	-28.556***	38.799***	59.164***
UK	0.010	1.163	-0.357***	9.814***	18034.5***	-13.503***	67.922***	216.446***
US	0.031	1.225	-0.547***	13.831***	35853.9***	-14.511***	207.461***	501.845***
WTI (oil)	0.015	2.644	-0.172***	16.065***	47132.9***	-11.322***	690.461***	1059.56***
10Y bond	-0.032	2.616	0.010***	31.433***	187479***	-11.171***	153.38***	1223.85***
Gold	0.053	1.717	-0.169***	11.603***	18043.4***	-38.333***	117.108***	99.894***

Note: *Asterisk *,**, and *** denote rejections of null hypothesis at 10%, 5%, and 1% significance levels, respectively. The null hypothesis of the third and fourth moments are "Skewness = 0" and "Excess Kurtosis = 3".*

Table 1 presents the descriptive statistics of all series used in this study after synchronising the data as noted above, this leads to approximately 4400 observations for each series. The summary statistics are as expected, with a mean daily return that is close to zero and a larger standard deviation. The skewness and excess kurtosis statistics show that the return series are leptokurtic with higher peak points as well as fatter tails. The Jarque-Bera normality test results indicate that the distribution of each series is not Gaussian. The Augmented Dickey-Fuller (ADF) test results reject the null hypothesis of a unit root at the 1% significance level, with all series stationary. Serial correlation tests for both the mean (Ljung-Box Q-statistic) and variance

[26] In determining the choice of index, we follow those used by Weiping et al. (2020) and Zhang et al. (2020) as well as being based on, arguably, the most recognisable index in each country.

(ARCH-LM test) indicate the presence of such correlation, which supports the use of the models outlined in Equations (2)-(4).

5.5. Empirical Results

This paper seeks to examine volatility spillovers between key global financial assets and G20 stock markets, which compares with the current literature that typically analyses a small selection of stock markets or one stock market with alternative assets. To study this relation, we use a synthesis, first developed by Liu et al. (2017), which combines the bivariate GARCH-BEKK model with the complex network theory.[27] To examine this spillover network of international markets we use daily data over the period 08/01/2003-04/01/2021, which is divided into five sub-periods that cover tranquil and crisis periods. Two important crisis periods within our sample are the Global Financial Crisis (GFC) in 2008 and the Covid-19 Crisis (CVC) in 2020, and these act as the cornerstone of this study.

Specifically, we divide our full sample into sub-periods in accordance with the crisis and non-crisis periods. Period 1 captures the Pre-Crisis period and is from January 8, 2003 to August 9, 2007. Period 2 encompasses the 2008 Global Financial Crisis, covering August 10, 2007 to December 30, 2009. Period 3 (Post-Crisis) takes place between January 4, 2010 and December 16, 2013. Following, Period 4 is the Pre-Pandemic period and covers between December 17, 2013 and December 30, 2019. Period 5 is from January 4, 2020 to January 4, 2021, and is the Covid-19 Crisis period.

In choosing these dates, we follow the work of Weiping et al. (2020) and Zhang et al. (2020) who also use similar sub-sample analysis. Our sub-sample dates differ slightly from these two papers as we use a larger sample, both starting earlier (in 2003 compared to 2006) and ending later (in 2021 compared to 2018). The end of our Period 1 and the dates for Period 2 are the same as these papers, while our Period 3 matches that of Weiting et al. (2020). We extend our Period 4 beyond the sample in each paper, and this allows us to isolate the Covid-19 period. While the choice of sub-sample dates always contains an element of subjectivity, by following the previous literature, we are able to provide some comparability.[28]

---

[27] To estimate the models, we use a range of software. Specifically, R studio (xts package) to match and prepare the pairwise data, Gretl (mgarch package) for GARCH-BEKK model and Gephi for networks construction.
[28] An alternative approach, as considered by Demirer et al. (2018) is to estimate a rolling full-sample VAR. However, as noted above, this would mean a substantial loss of information in aligning the data.

Table 2 presents some widely-used networks statistics. The first statistic, column (i), shows that there are 171, 257, 142, 175, and 297 volatility spillover linkages that are extracted from the 462 edges in the 5 sub-periods, respectively. In column (ii), the total degree indicates the sum of all node sizes in the sub-periods and whose values are highly correlated with the values in column (i). These measures are two of the more important general network statistics and indicate that volatility spillover between global financial indicators and G20 stock markets are present in all sub-periods and that the spillover relations are largely bidirectional. More importantly, volatility spillover networks are more dense during a crisis, as evidenced by the values in Period 2 (GFC in 2008) and Period 5 (CVC in 2020) when compared to the non-crisis Periods 1, 2, and 4. It is also clear that the volatility spillover network is time-varying with the number of linkages changing over the five sub-periods.

5.Table 2 ***Network Statistics***

	Number of edges (i)	Total degree (ii)	Av. Weighted degree (iii)	Av. Shortest Path length (iv)	Network Diameter (v)	Graph density (vi)	Av. Clustering coefficient (vii)
Period 1	171	24.95	1.134	1.463	3	0.541	0.615
Period 2	257	72.28	3.286	1.329	2	0.671	0.699
Period 3	142	36.02	1.638	1.602	3	0.420	0.539
Period 4	175	29.81	1.355	1.424	2	0.576	0.608
Period 5	297	206.9	9.406	1.225	2	0.775	0.808

Note: *Smaller the values of (iv) average shortest path length and (v) network diameter mean tighter networks. The others are meant to be as normal; more precisely, higher values, tighter networks are.*

Of the other statistics, in column (iii), the average weighted degree is the ratio of total degree to the number of nodes. The higher the average weighted degree, the tighter the market interrelations. These values are higher in the crisis periods 2 and 5 (3.286 for GFC and 9.406 for CVC) compared to the non-crisis periods 1, 3, and 4 (1.134, 1.638, and 1.355, respectively). The implication is that crisis periods are likely to deepen market connections and therefore, the potential for financial contagion. Fluctuations occurring in one market can therefore, spread more easily to other markets. For column (iv), the average shortest path length is defined in a network as the mean shortest steps of volatility spillover propagation from node i to node j. A smaller shortest path length means stronger linkages between nodes, which also mean a denser network. This statistic fluctuates between the values 1 and 2, with lower values in the crisis periods (1.329 and 1.225 for periods 2 and 5 respectively) compared to the networks of non-crisis periods (1.463, 1.602, 1.424 for periods 1, 2 and 4 respectively). Hence, volatility spillovers in crisis periods propagate faster compared to those in non-crisis periods.

In column (v) the network diameter is known as the shortest path between the two most distant nodes of the network. The network diameter is at 2 for the second, fourth, and fifth sub-periods, while the same statistic is at 3 for the first and third periods. This means that any spillover in any node may reach the farthest node point in maximum 2 steps in the crisis periods, whilst it is a maximum of 3 steps in normal periods. Again, this is the evidence of stronger spillover transmission during the crisis periods. In column (vi), the graph density is a measure that shows how close the number of edges is to the maximum possible number of edges. If the graph density is equal to 1 in a network, the network is called a complete graph that includes all the possible edges. The graph density statistics are closer to 1 in the turmoil periods 2 and 5 (0.671 and 0.775, respectively) in comparison with the calm periods 1, 3, and 4 (0.541, 0.420, and 0.576, respectively). Column (vii), the average clustering coefficient, is a similar measure to graph density, indicating how the nodes are integrated in a network graph. The graph density and average clustering coefficient move together to a large extent; the two peak points of the average clustering coefficient are approximately 0.7 and 0.8 in the GFC and CVC periods, respectively.

In sum, the network statistics present three key results. First, spillover effects exist through the five different sub-periods. Second, the nature of such spillovers is time-varying across the sample periods. Third, the strength and nature of spillovers increases during a crisis.

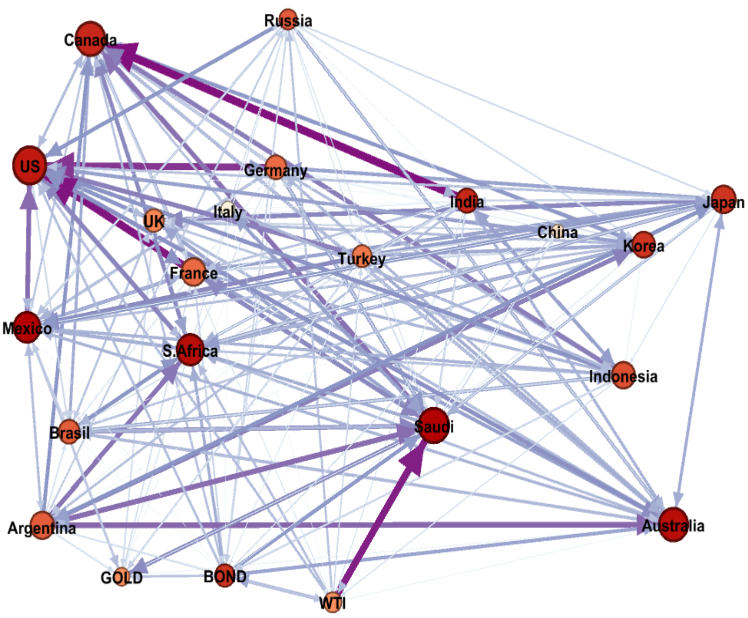
5.5.1. Results by Period

5.5.1.1. Crisis Periods 2 and 5

The crisis periods 2 and 5 cover the GFC and CVC and thus, it is of interest to see the different impact of the two crises. Notably, we can observe a clear difference in the network graphs, presented in Figures 1 and 2, and statistics noted in Table 2 for the two periods. Comparing the two figures, the spillover network graph in Period 5 is tighter and more integrated than the network graph of Period 2. This suggests that volatility spillovers among markets in the CVC period spreads faster than volatility spillovers in the GFC period. This implies that the CVC period exhibits more widespread effects on both key global assets and the G20 markets compared to the GFC period in 2008. This can equally be seen in the number of edges (297 in period 5 against 257 in period 2), the total degree (207 against 72), average weighted degree (9 against 3) and graph density (0.8 against 0.7).

In Period 2 (Figure 1), the rankings of top weighted nodes are US, Saudi Arabia, Canada, and Australia, respectively, while the stock markets of China and Italy are the bottom weighted. More specifically, US, Canada, Saudi Arabia, and Australia have the highest weighted-in-degree that receive spillovers, while Saudi Arabia, Japan, Australia, Argentina, and US are the top weighted-out-degree volatility senders, respectively. Thus, we can see similarity in the markets that are the highest volatility receivers (weighted-in-degree) and senders (weighted-out-degree). For example, the US appears as the most active node with the largest weighted degree.

5.Figure 1 ***Period 2 (10/08/2007-30/12/2009 Global Financial Crisis)***



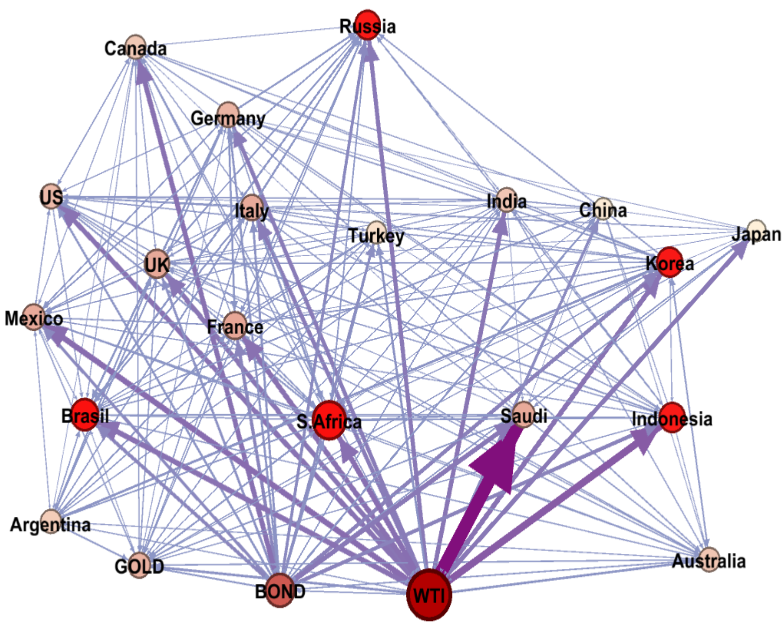
Markets	w. in-degree	w. out-degre	w. degree
US	7.53977	4.53239	12.07216
Saudi	5.83215	4.95354	10.78569
Canada	6.74636	3.57782	10.32418
Australia	5.4197	4.62615	10.04585
S.Africa	5.510416	3.52744	9.037856
Mexico	4.53484	4.27063	8.80547
France	3.74876	3.80245	7.55121
Japan	2.73145	4.74295	7.4744
Argentina	2.72763	4.58876	7.31639
Indonesia	3.08092	3.916216	6.997136
Korea	3.71052	3.02224	6.73276
India	2.42335	3.51356	5.93691
Germany	2.61187	3.18397	5.79584
Brasil	2.75073	3.02663	5.77736
UK	2.76746	2.51186	5.27932
BOND	1.46926	3.66685	5.13611
Turkey	1.97442	2.95419	4.92861
Russia	1.69085	2.4702	4.16105
WTI	1.23364	2.80012	4.03376
GOLD	2.17203	0.93048	3.10251
Italy	0.94476	1.18274	2.1275
China	0.66336	0.48306	1.14642

Note: Darker colours in networks represent larger spillover relationships, while lighter colours indicate weaker spillover relationships; red and bigger nodes show bigger spillover centres, wider and dark purple edges are the strongest linkages. The table on the right hand side is sorted by the average weighted degree values of the markets from largest to smallest.

While the GFC began in the US, we can observe the effect spreading to the rest of the world through interconnections in the global financial system. We can see that Canada and Australia, which rely on commodity exports are notably affected. Equally, a further market affected is the major oil-exporting Saudi Arabia, while South Africa is also a commodity exporter. We can also see the role played by the global risk factors of interest rates, oil, and gold prices play in the transmission. Although, they appear towards the bottom of the weighted list, they play a more significant impact than several G20 markets.

These results do imply the existence of systematic risk on a global scale that can be thought of as restricting the possibility of international portfolio diversification. However, a small number of markets during this period, such as China and Italy, could be used to construct more robust diversified portfolios as the node sizes of these markets are the smallest in the network.

5.Figure 2 ***Period 5 (04/01/2020-04/01/2021 Covid-19 Crisis)***



Markets	w. in-degree	w. out-degree	w. degree
WTI	1.70283	69.488	71.19083
BOND	3.31717	35.49824	38.81541
S.Africa	14.30131	7.48462	21.78593
Korea	13.38907	7.08582	20.47489
Brasil	13.33919	6.7552	20.09439
Russia	14.06039	4.88495	18.94534
Indonesia	12.82614	5.41328	18.23942
France	12.4683	5.55527	18.02357
UK	10.23322	7.39829	17.63151
Saudi	16.26137	1.31707	17.57844
Mexico	12.58052	4.9769	17.55742
Italy	10.93127	6.54846	17.47973
GOLD	9.7947	6.36454	16.15924
Germany	8.56562	7.31533	15.88095
US	11.07091	3.64721	14.71812
Australia	6.93636	5.71164	12.648
India	7.13052	5.49841	12.62893
Canada	8.62799	3.65	12.27799
Argentina	2.45088	9.10638	11.55726
Turkey	6.72688	1.42353	8.15041
China	5.72361	0.74475	6.46836
Japan	4.4892	1.05956	5.54876

Note: Darker colours in networks represent larger spillover relationships, while lighter colours indicate weaker spillover relationships; red and bigger nodes show bigger spillover centres, wider and dark purple edges are the strongest linkages. The table on the right hand side is sorted by the average weighted degree values of the markets from largest to smallest.

In Period 5 (Figure 2), the average weighted degrees are much higher compared to their counterparts in Period 2. A notable difference is that no specific stock market can be regarded as the source of the crisis compared to Period 2 and the US market. In Period 5, the oil and bond markets are shown to be the source of the largest amount of volatility spillovers directed to the G20 markets. In terms of the G20 stock markets themselves, the largest weighted degrees are South Africa, Korea, and Brazil, whereas Japan, China, and Turkey have the smallest weighted degree values, respectively. Looking deeper into the Period 5 results across the weighted degrees, the top volatility senders (out-degree) are different to the receivers (in-degree). As noted, the global financial indicators of oil and bonds are the volatility senders, meaning that they affect but are not affected by other markets. The top receivers are the stock markets of Saudi Arabia, South Africa, Russia, Korea, and Brazil, respectively. This again, notably, highlights the importance of oil within global financial markets, with major oil export markets affected.

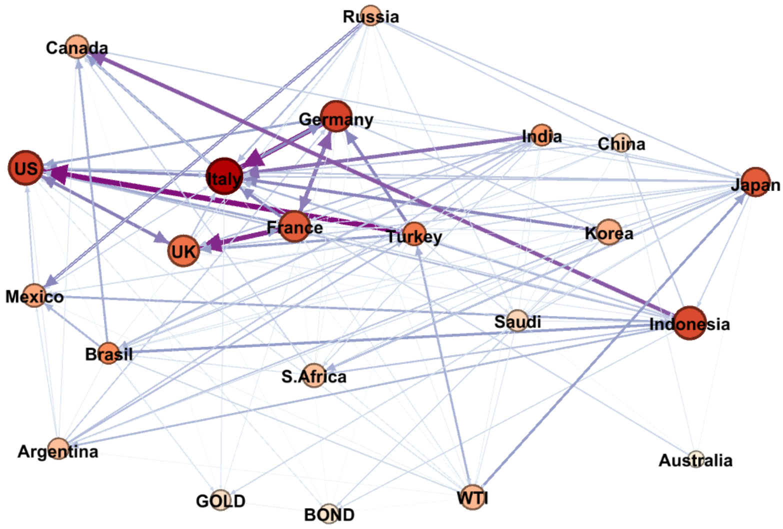
In explaining the large volatility spillovers from oil prices to the G20 markets, we saw reduced global oil demand by 29 million barrels per day during the Covid-19 lockdown period, while the price of West Texas Intermediate (WTI) fell to negative $37 per barrel on April 20, 2020. This oil price crash creates a large spike on own volatility that spread to oil dependent markets as well as globally. Following oil, the node on US 10-year Treasury bond yields is the second largest that has an impact on G20 markets. The bond yields start decreasing when Wuhan lockdown began and reach the lowest level (below 0.6) with the declaration of coronavirus pandemic by the World Health Organisation. In contrast, the spillover effect from gold prices on the G20 stock markets is relatively smaller.

5.5.1.2. Non-Crisis Periods 1, 3, and 4

The network graphs (Figures 3 to 5) of the non-crisis periods are less dense than for the crisis periods. Furthermore, unlike the crisis periods, the influence of the global assets (oil, gold, and bond) on the G20 markets is limited.

In Period 1 (Figure 3), the node on Italy has the largest weighted degree among markets, while its value in Periods 3 and 4 is relatively smaller. This emphasises the time-varying nature of the volatility spillover relations among the G20 markets. The most active nodes in Period 1 following Italy are from the US, Germany, Indonesia, Japan, and France. As a most active node, Italy is the largest volatility receiver, but not the largest volatility sender. The ranking of the largest senders is Indonesia, Japan, and France. The smallest nodes during this first period are Australia, bonds, gold, Saudi Arabia, and China. In Period 1, the impact of oil, bonds, and gold on the G20 stock markets is notably limited.

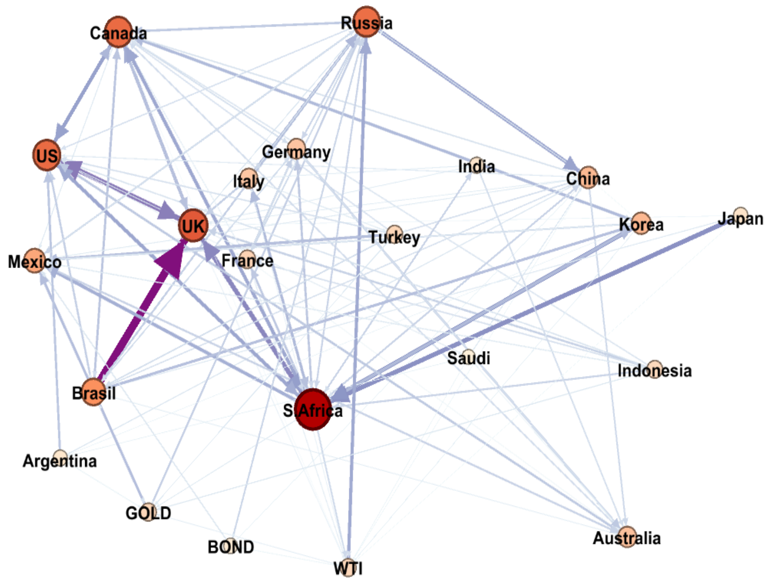
5.Figure 3 ***Period 1 (Pre-Crisis 08/01/2003-09/08/2007)***



Markets	w. in-degree	w. out-degree	w. degree
Italy	3.5682	1.46886	5.03706
US	2.91285	1.02994	3.94279
Germany	1.99071	1.90165	3.89236
Indonesia	1.24253	2.54572	3.78825
Japan	1.11379	2.36277	3.47656
France	1.36577	2.09592	3.46169
UK	1.95541	1.15243	3.10784
Turkey	0.96537	1.99954	2.96491
Brasil	1.27505	1.4191	2.69415
India	0.99191	1.37362	2.36553
Mexico	1.32032	0.7005	2.02082
Canada	1.51089	0.32497	1.83586
Korea	0.64517	1.15586	1.80103
Russia	0.45291	1.2642	1.71711
WTI	0.56119	1.08808	1.64927
Argentina	0.40114	1.06855	1.46969
S.Africa	0.79828	0.61962	1.4179
China	0.53991	0.4297	0.96961
Saudi	0.53014	0.38153	0.91167
GOLD	0.39442	0.27125	0.66567
BOND	0.37218	0.0641	0.43628
Australia	0.04603	0.23626	0.28229

Note: Darker colours in networks represent larger spillover relationships, while lighter colours indicate weaker spillover relationships; red and bigger nodes show bigger spillover centres, wider and dark purple edges are the strongest linkages. The table on the right hand side is sorted by the average weighted degree values of the markets from largest to smallest.

5.Figure 4 ***Period 3 (Post-Crisis 04/01/2010-16/12/2013)***

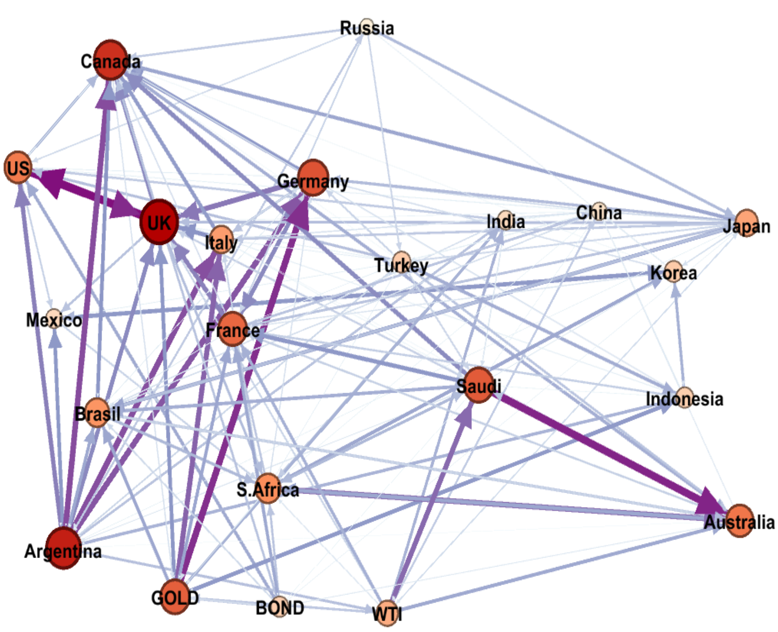


Markets	w.in-degree	w. out-degree	w. degree
S.Africa	4.68359	5.67734	10.36093
UK	4.61137	2.57988	7.19125
US	4.03053	2.53471	6.56524
Canada	3.86337	2.69745	6.56082
Russia	3.18052	3.21998	6.4005
Brasil	1.31492	3.84218	5.1571
Mexico	2.99481	1.0501	4.04491
Korea	1.28196	1.86814	3.1501
China	1.79914	1.33266	3.1318
Australia	1.59701	1.14298	2.73999
Germany	1.5448	1	2.5448
Italy	1.03026	1.39375	2.42401
France	1.16157	0.62334	1.78491
Turkey	0.30306	1.32558	1.62864
GOLD	0.69499	0.92155	1.61654
WTI	0.74969	0.83479	1.58448
Indonesia	0.14001	1.351	1.49101
Japan	0.14654	0.81921	0.96575
India	0.37641	0.51456	0.89097
Argentina	0.21223	0.60741	0.81964
BOND	0.15794	0.4764	0.63434
Saudi	0.15348	0.21519	0.36867

Note: Darker colours in networks represent larger spillover relationships, while lighter colours indicate weaker spillover relationships; red and bigger nodes show bigger spillover centres, wider and dark purple edges are the strongest linkages. The table on the right hand side is sorted by the average weighted degree values of the markets from largest to smallest.

In Period 3 (Figure 4), the ordering of markets changes reflecting the time-variation in spillovers. Here, the most active nodes are South Africa, UK, US, and Canada, while Saudi Arabia, bonds, Argentina, India, and Japan are the least active ones. Notably, South Africa, Brazil, Russia, UK, and US are the largest volatility senders, while South Africa, UK, US, and Canada are the largest volatility receivers.

5.Figure 5 ***Period 4 (Pre-Pandemic 17/12/2013-30/12/2019)***



Markets	w. in-degree	w. out-degree	w. degree
UK	3.08471	1.80924	4.89395
Argentina	0.46217	3.99371	4.45588
Canada	3.7275	0.4948	4.2223
Germany	2.02922	1.72255	3.75177
Saudi	1.35604	2.2856	3.64164
GOLD	0.72721	2.85232	3.57953
France	2.33414	1.13988	3.47402
Australia	2.07343	1.19903	3.27246
US	1.76618	1.43756	3.20374
S.Africa	1.57369	1.34659	2.92028
Brasil	1.73566	1.05871	2.79437
Italy	1.57164	1.0526	2.62424
Japan	0.73658	1.7615	2.49808
WTI	0.39698	1.9414	2.33838
Korea	1.09085	0.64464	1.73549
Turkey	0.64325	1.08105	1.7243
Indonesia	1.15945	0.50399	1.66344
BOND	0.28189	1.37069	1.65258
Mexico	0.93491	0.5232	1.45811
India	0.94509	0.49788	1.44297
China	0.97196	0.2913	1.26326
Russia	0.21262	0.80693	1.01955

Note: Darker colours in networks represent larger spillover relationships, while lighter colours indicate weaker spillover relationships; red and bigger nodes show bigger spillover centres, wider and dark purple edges are the strongest linkages. The table on the right hand side is sorted by the average weighted degree values of the markets from largest to smallest.

In Period 4 (Figure 5), UK, Argentina, Canada, and Germany exhibit the largest set of nodes, whereas Russia and China have the smallest. The biggest volatility senders are Argentina, Saudi Arabia, gold and the UK, with Canada, UK, and France the biggest volatility receivers. In terms of the global financial indicators, gold is a significant volatility sender that affects the G20 markets, but oil and bonds have a moderate effect. Of note, the price of gold saw a gradual increase towards the end of the period, just before the beginning of the CVC. This time period includes the Brexit vote and this may explain the high degree of UK spillovers.

5.5.2. Key Paths in Spillover Networks

The spillover relations are conveyed by the edges among the nodes in a complex network and therefore detecting the thickest linkages between markets may reveal some important information about market interrelations. Table 3 presents the top five edges (strongest linkages) over each of the sub-periods. It is worth noting that the top edges are independent from the size of nodes, meaning that a node could be relatively small, while having the top edge of network. For instance, the node of India is relatively small in Period 2, but it has the largest linkage in the same period, which is from India to Canada.

Examining the results in Table 3, Period 5 has the strongest linkages among all the sub-periods, with spillovers from oil to Saudi Arabia, Indonesia, Mexico, and South Africa. The strongest edge is from oil to Saudi Arabia with a 5.4% weight of the total spillover. This highlights the importance of oil to the Saudi stock market given its major oil-exporting role. As noted above, both oil demand and the oil price fell dramatically at the start of the CVC period. Mexico and Indonesia are also notable for oil exports. In Period 2, the direction of top edges are towards the node of US market. This means that despite the US being the origin of the GFC crisis, the US market is also exposed to significant external volatility, with the top two edges from France and Germany. Again, Saudi Arabia is notably affected as the GFC led to a global recession and a fall in demand for oil. This provides further evidence of a strong relation between oil prices and the Saudi stock market.

5.Table 3 ***Key Paths in Spillover Networks***

PERIOD 1	**Value (%)**	**PERIOD 2**	**Value (%)**	**PERIOD 3**	**Value (%)**
Turkey –> US	0.71 2.8%	India –> Canada	1.02 1.4%	Brasil –> UK	1.55 4.3%
Germany –> Italy	0.66 2.6%	France –> US	1.00 1.3%	UK –> US	0.93 2.5%
France –> UK	0.64 2.5%	OIL –> Saudi	0.96 1.3%	US –> UK	0.90 2.4%
Indonesia–> Canada	0.51 2.0%	Germany –> US	0.87 1.2%	S. Africa –> UK	0.86 2.3%
France –> Germany	0.50 2.0%	Argentina –> Australia	0.69 0.9%	Japan –> S. Africa	0.80 2.2%
PERIOD 4		**PERIOD 5**			
UK –> US	0.70 2.3%	OIL –> Saudi	11.3 5.4%		
Saudi –> Australia	0.65 2.1%	OIL –> Indonesia	5.63 2.7%		
US –> UK	0.65 2.1%	OIL –> Mexico	4.37 2.1%		
GOLD –> Germany	0.62 2.0%	OIL –> S. Africa	4.12 1.9%		
Argentina –> Canada	0.54 1.8%	BOND –> Canada	4.02 1.9%		

Note: *"–>" indicates the spillover direction, the first row (value) is the amount of volatility the second (percent) is the percentage of spillover amount in total spillover.*

In Period 3, the top edge is from Brazil to the UK, while other significant edges take place bidirectionally between the UK and US markets. Japan to South Africa and South Africa to the UK also show notable linkages. The bidirectional spillover relation between the US and UK stock markets returns in Period 4. It is followed by edges from Saudi Arabia to Australia, Gold to Germany, and Argentina to Canada. In Period 4, some economic implications regarding the largest edge between the UK and US markets can be linked to the Brexit referendum (23 June 2016) that has a significant impact on the UK economy and reveals the UK as large volatility source. Afterwards, the 2015-16 stock market sell-off in the US market and the world-wide stock market downturn in 2018 can further explain the strong linkage between the UK and US stock markets. In Period 1, the largest linkages are more mixed, being from Turkey to the US, Germany to Italy, France to the UK, Indonesia to Canada, and France to Germany.

5.5.3. Network Robustness Checks

Testing the robustness of our complex network can ensure its stability and resilience and the reliability of the presented results. To consider robustness in a complex network, we can examine the response of the network to changes in the number of nodes. Therefore, we re-examine each network by, first, adding a further node and, second, by removing the global financial indicators from our sample dataset. Inevitably, changing the number of nodes will alter the results, however, the key question is whether this results in substantial changes to the network statistics and node sizes. The results evidence that no significant changes to the network statistics and node-edge sizes of the included markets is observed across the sub-periods. This underpins the consistency of our results.

Specifically, we undertake two exercises and consider how the network statistics change with a differing number of nodes. First, we increase the number of nodes from 22 to 23 by including the implied volatility index, VIX, based on the US SP500. Here, the average weighted degree, graph density, and clustering coefficient for all sub-periods rise proportionally to indicate that the networks with 23 nodes are slightly tighter compared to those of 22. However, the average shortest path length and network diameter statistics, which are inversely proportional with the other three statistics, decrease slightly. In the second exercise, the number of nodes decreases from 22 to 19 by removing the three global financial indicators. Here, we do not encounter any substantial changes in the network statistics (except for the average weighted degree in Period 5 that falls from 9.406 to 4.407).[29] For the graph density and clustering coefficient, we see a small increase in the first four sub-periods, and a slight decrease in Period 5. Accordingly, the average shortest path length decreases in the first four

[29] This change in Period 5 is linked to the crash in oil prices (20 April 2020) that have a notable impact on the network. Therefore, when we omit the oil, bond, and gold in this exercise, we observe the noted fall.

sub-periods and increasing in the fifth period. The network diameter does not change in general, but the weighted degree results suggest an increase in the first and third periods and a decrease in other periods. To sum, these results do not affect our main conclusions as no significant changes in the node-edge sizes of networks is reported.

5.6. Conclusion

This paper analyses the volatility spillover relations between key global financial barometers (oil, gold, and bond) and G20 stock markets using a network approach. Specifically, a bivariate GARCH-BEKK model that captures spillover relations is combined with a complex network approach. Using this synthesis, we construct the spillover networks of international financial markets between 08/01/2003 and 04/01/2021 and divided this sample into five sub-periods to cover calm and crisis periods including the Global Financial Crisis (GFC) in 2008 and the Covid-19 Crisis (CVC) in 2020.

We detect 171, 257, 142, 175, and 297 volatility spillover linkages across the five sub-periods, respectively. This highlights the time-varying nature of the spillovers between the key global variables and G20 markets. Of note, the volatility spillover networks are much denser during the GFC and CVC crisis periods compared to the networks of non-crisis periods. The crisis periods are more transitive, resulting in volatility that transmits more rapidly and directly through the different assets examined. In Period 5 (CVC), the global financial indicators of oil and bonds appear as the main senders of volatility, indicating that they affect other markets but are not affected by those markets. In contrast, Saudi Arabia, Russia, South Africa, and Brazil are major volatility receivers in this period. An important point here is the role that oil plays in affecting global stock markets and notably oil-exporting markets. In Period 2 (GFC), as the crisis began in the US, it acts as a major source of volatility spillovers along with Australia, Canada, and Saudi Arabia. A further point of interest is that China is among the least sensitive market to external volatility across all the sub-periods. In the non-crisis periods, the influence of the global variables on the G20 is notably lower compared to the case of the crisis periods. Instead, the effect of regional information becomes dominant, such as the Brexit referendum in the UK.

Although the results of this work are time and market specific, the movements of key nodes and edges over time provide important information. For policymakers, investors, and market participants, considering spillover relations in G20 stock markets is important in being able to manage risks and portfolio diversifications. For example, the existence of global risk factors can be thought as a sign to restrict the possibilities of portfolio diversification, especially during crisis periods when the correlations among investment instruments are high. In the non-crisis periods, more diversified portfolios can be constructed depending on time and market specific information.

CHAPTER 6

6. Summary and Concluding Remarks

6.1. Summary

The aim of this thesis is to examine the behaviour of stock market volatility in a selection of international markets, the predictive ability of extended models to provide accurate volatility forecasts, and the nature of the interrelations between markets from the perspective of complex network theory. In particular, this research consists of four substantive chapters in the centre of the conditional volatility, realized volatility, and volatility spillovers.

Focussing on the modelling and forecasting of realized volatility the first chapter empirically contributes to the existing literature by conducting a comprehensive exercise, including 30 different international stock markets and more up-to-date data. We conduct a comparative forecasting exercise for improving out-of-sample volatility forecasts of different stock markets at daily horizon. Two different classes of volatility forecasting models are considered and compared: the conditional volatility models such as GARCH type models and realized volatility models, HAR-RV and ARFIMA-RV. The findings are mostly in line with previous research that HAR-RV is the best-performing model. Moreover, the decomposition of realized volatility into positive and negative realized semivariances (e.g. HAR-PS model) improve the forecast accuracy of HAR-RV model.

In the second chapter, we extend the baseline HAR-RV model with some exogenous (X) variables (symbolized as HAR-RV-X) which are extreme range estimators such as Parkinson, Garman-Klass, Roger-Satchell, and Yang-Zhang. Essentially, this chapter examines whether these extreme range estimators improve the out-of-sample forecast accuracy of future realized volatility in the G7 stock markets. Despite the fact that the findings seem to be inconclusive in the stock markets of group of Seven, Parkinson and Garman-Klass estimators could be employed in the HAR-RV model to generate better forecasts of future realized volatility. To the best of our knowledge, this is the first work which examines the information content of extreme range information at improving the forecasts of realized volatility.

The third substantive chapter uses the HAR-RV-X model from the previous chapter and also further expand it by the Kitchen Sink (KS) strategy. The HAR-RV-KS includes a long list of possible additional variables in the model at once. A wide range of exogenous variables from assets to commodities, implied volatility indices to bond rates are involved in this analysis for an inclusive analysis. Furthermore, those additional variables are classified in respect to different information channels, namely local, regional, and global. In doing so, we aim to investigate which class of models best helps in forecasting the future realized volatility. The results indicate that while the HAR-RV-KS model outperforms the HAR-RV and HAR-RV-X (with only one X variable) specifications, the role of global information at improving the forecasts of future realized volatility is more important than the others.

Unlike the first three chapters, the last chapter handles the transmission of volatility across international markets, namely the interrelation between non-stock assets (oil, gold and bond) and G20 stock markets. In this regard, we consider a bivariate GARCH-BEKK model that captures volatility spillover effects. Moreover, the GARCH-BEKK model is combined with a complex network theory to analyse the volatility spillover relations between key global financial indicators and G20 stock markets. Using this method, we construct the spillover networks of international financial markets under five identified sample sub-periods including crisis and non-crisis periods. The findings contribute to the literature of volatility spillovers from the network theory perspective as follows. The volatility spillover relations between the global indicators such as oil, gold, and bond and G20 markets vary significantly across five identified sub-periods. Notably, networks are much denser in crisis periods than those of non-crisis periods. In comparing two crisis periods, Global Financial Crisis (2008) and COVID-19 Crisis (2020) periods, the network statistics indicate that volatility spillovers in the latter period are more transitive and intense than the former. This suggests that financial volatility spreads more rapidly and directly through key financial indicators to the G20 stock markets, especially in crisis periods.

6.2. Concluding Remarks

In conclusion, this thesis presents the four main contributions. The first substantive work empirically contributes to the volatility forecasting literature by carrying out an extensive exercise with the thirty different international stock markets and more up-to-date data. The second empirical exercise informs that the well-established range based estimators such as Parkinson and Garman-Klass could be employed to generate more accurate forecasts of future realized volatility. Thirdly, the sources of global information are better to improve the future realized volatility forecasts than local and regional information. The final exercise methodologically and empirically contributes to the literature of volatility spillovers. Methodologically, a solution is provided to the difficulty encountered by the bivariate GARCH-BEKK model when dealing with multi-dimensionality issue. In other words, a bivariate GARCH-BEKK specification is combined with complex network theory to build a network of financial markets, consisting of nodes and edges. Empirically, we find that the networks of international financial markets are much denser in crisis periods than that of non-crisis periods and financial volatility spreads more rapidly and directly through key financial indicators to the G20 stock markets, particularly in crisis periods.

The findings of the first three of empirical studies could provide useful information in improving the forecast accuracy of stock market volatility and determining the most and least relevant parameters of stock market volatility. For example, some well-documented extreme range estimators such as Parkinson and Garman-Klass and global information sources (e.g. global financial indicators and US market information) as exogenous variables can help in forecasting the one-day-ahead volatility of stock markets.

Over the last two decades, the Global Financial Crisis (GFC, 2008) and the COVID-19 Crisis (CVC, 2020) have further indicated the importance and impact of accurate and efficient volatility forecasting. Hence, accurate volatility forecasting is quite important in terms of portfolio and risk management, especially for policy makers, investors, and market participants who have certain levels of risk which they intend not to exceed. Some practical implications from our findings can help them in the process of both forecasting and asset allocation. This is the fact that all financial actors desire to know today, “what will be the degree of volatility tomorrow?”. In terms of forecasting angle, the gain in forecasting accuracy is believed to be economically significant to minimize risk and maximize return. For instance, while investors and market participants would like to rearrange their stocks or portfolio positions before financial markets become too volatile, policy makers would desire to narrow bid-ask spread in order to restore market liquidity if the future is expected to be more volatile. From the perspective of asset allocation, investors and market participants could use our findings to align their portfolios by reducing their exposure to various risks (i.e. local, regional, and global risks).

The results of the last empirical research could also contain important information for policymakers, investors, and market participants who would like to invest in overseas. This is because assessing spillover relations in G20 markets is crucial in being able to manage risks and portfolio diversifications. For instance, the existence of global risk factors can be considered as a sign to restrict the possibilities of portfolio diversification, particularly during crisis periods when the correlations among investment instruments are high. In the non-crisis periods, more diversified portfolios can be constructed depending on time and market specific information.

A plausible avenue of future research to extend Chapter 2 would be to investigate the forecasting ability of wider classes of models by adding simple and stochastic volatility models (or maybe machine learning and deep learning models, which are quite popular these days) as well as using wider stock market indices. In Chapter 3, applying our methodological approach to a wider groups of stock market prices and oil and gold prices could validate the capacity of seven range-based estimators in forecasting realized volatility within a HAR-RV-X framework. Also, employing the range-based estimators derived from higher frequencies could yield noteworthy results. In a similar vein, the proposed methodological framework in Chapter 4 would also be used to identify different risk exposures (e.g. local, regional and global) of other different stock indices, including, for instance, Asian and/or Middle East stock market indices. In Chapter 5, it might be worth to conduct this exercise using the Diebold and Yilmaz (2009, 2012) spillover index under a common sample data (which is possible with weekly frequency data), however, we are unable to do it in our case due to the usage of daily data. Also, within the same chapter, a sectoral extension of this network analysis could be another interesting avenue of research.

References:

Akgiray, V. (1989), Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts, Journal of Business, 62, 55-80.

Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance, 57*(3), 1047–1091.

An, H., Zhong, W., Chen, Y., Li, H., Gao, X. (2014). Features and evolution of international crude oil trade relationships: A trading-based network analysis. *Energy,* 74, 254–259.

An, S., Gao, X., An, H., An, F., Sun, Q., Liu, S. (2020). Windowed volatility spillover effects among crude oil prices. *Energy*, 200, 117521.

Andersen, T. G. & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review, 39(4), 885-905.

Andersen, T. G., & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance, 4, 115–158.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of exchange rate volatility. Journal of the American Statistical Association, 96, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modelling and Forecasting Realized Volatility. Econometrica, 71(2), 579-625.

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

Balaban, E., Bayar, A., and Faff, R. W., (2006), Forecasting stock market volatility: further international evidence, *The European Journal of Finance*, Vol. 12, pp. 171-188.

Bali, T.G., Weinbaum, D., (2005). A comparative study of alternative extreme-value volatility estimators. J. Futures Markets 25, 873–892.

Ball, C., & Torous, W. (1984). The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing. Journal of Business, 57, 97–112.

Bandi, F. M., & Russell, J. R. (2005). Separating microstructure noise from volatility. Journal of Financial Economics, 79, 655–692.

Barndorff-Nielsen, O., S. Kinnebrock, and N. Shephard. (2010). Measuring Downside Risk: Realized Semi-variance. In Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle, T. Bollerslev, J. Russell, and M. Watson, eds. Oxford; New York: Oxford University Press, 117–136.

Bates, J. M., Granger, C. W. (1969). The combination of forecasts. Or. 451 – 468.

Black, F. (1976). Studies of stock price volatility changes. In Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section (pp. 177-181).

Blair, B. J., Poon, S.-H., & Taylor, S. J. (2001). Forecasting S&P100 volatility: the incremental information content of implied volatilities and high-frequency index returns. Journal of Econometrics, 105(1), 5-26. Forecasting and empirical methods in finance and macroeconomics.

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

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate Generalized ARCH model. *Review of Economics and Statistics*, 72, pp. 498-505.

Bollerslev, T. (2009). Glossary to ARCH (GARCH), the NBER.

Bollerslev, T., Engle, R.F. and Wooldridge, J.M. (1988). A capital asset pricing model with time varying covariances. *Journal of Political Economy*, 96, 116–131.

Bollerslev, T., Kretschmer, U., Pigorsch, C., & Tauchen, G. (2009). A discrete-time model for daily S & P500 returns and realized variations: Jumps and leverage effects. Journal of Econometrics, 150(2), 151-166.

Bollerslev, T., Li, S., Zhao, B., (2018). Good volatility, bad volatility and the cross-section of stock returns. J. Finance Quantitative Analysis Forthcoming.

Bollerslev, T.; S. Z. Li; and V. Todorov. (2016). Roughing up Beta: Continuous vs. Discontinuous Betas and the Cross-Section of Expected Stock Returns. *Journal of Financial Economics*, 120, 464–490.

Brailsford, T. J. and R. W. Faff (1996), An Evaluation of Volatility Forecasting Techniques, Journal of Banking and Finance, 20, 419-438.

Brandt, M., & Diebold, F. (2006). A no-arbitrage approach to range-based estimation of return covariances and correlations. Journal of Business, 79, 61–74.

Brandt, M., & Jones, C. (2006). Volatility forecasting with range-based EGARCH models. Journal of Business and Economic Statistics, 24, 470–486.

Brooks, C. (2008). Introductory Econometrics for Finance. Cambridge University Press, second edition.

Bu, R., Rodrigo, H., Izzeldin, M., Murphy, A., and Tsionas, M. (2023). The contribution of jump signs and activity to forecasting stock price volatility. Journal of Empirical Finance, Volume 70, 144-164.

Bucci, A. (2017). Forecasting realized volatility: a review, MPRA paper.

Caporin, M. (2021). The role of jumps and asset liquidity in realized volatility modelling and forecasting. Journal of Financial Econometrics, 1-26.

Celik, S., & Ergin, H. (2014). Volatility forecasting using high frequency data: Evidence from stock markets. Economic Modelling, 36, 176–190.

Chortareas, G., Jiang, Y., & Nankervis, J. (2011). Forecasting exchange rate volatility using high-frequency data: Is the euro different? International Journal of Forecasting, 27, 1089-1107.

Chou, R. (2005). Forecasting financial volatilities with extreme values: the conditional autoregressive range CARR) model, Journal of Money, Credit and Banking, 37, 561–82.

Chou, R. Y., & Liu, N. (2011). Estimating time-varying hedge ratios with a range-based multivariate volatility model, Working paper, Academia Sinica.

Christensen, B. & Prabhala, N. (1998). The relation between implied and realized volatility. Journal of Financial Economics, 50(2), 125-150.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. Econometrica, 41(1), 135-55.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography, Int. J. Forecast. 5(4). 559–583.

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

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

Cumby, R., Figlewski, S., & Hasbrouck, J. (1993). Forecasting volatilities and correlations with EGARCH models. Journal of Derivatives, 1, 51-63.

Day, T. E. & Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. Journal of Econometrics, 52(1 - 2), 267-287.

Degiannakis, S., Filis, G. (2017). Forecasting oil price realized volatility using information channels from other asset classes, J. Int. Money Finance. 76, 28–49.

Demirer, M., Diebold, F. X., Liu, L., and Yilmaz, K. (2018). Estimating global bank network connectedness. *Journal of Applied Econometrics*, 33, 1–15.

Diebold, F.X., Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal*, 119, 158–171.

Diebold, F.X., Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1), 57–66.

Diebold, F.X., Yilmaz, K. (2014). On the network topology of variance decompositions: measuring the connectedness of financial firms, *Journal of Econometrics*, 182, 119–134.

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. Journal of Empirical Finance, 1(1), 83-106.

Ding, Yi and Kambouroudis, Dimos S and McMillan, David G., Forecasting Realised Volatility Using Regime-Switching Models. Available at SSRN: https://ssrn.com/abstract=4415386.

Duan, Y., Chen, W., Zeng, Q., Liu, Z. (2018). Leverage effect, economic policy uncertainty and realized volatility with regime switching, Physica A, vol. 493(C), pages, 148-154.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics*, 20, 339-350.

Engle, R. (2002). New frontiers for arch models. Journal of Applied Econometrics, 17(5), 425-446.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), pp. 987-1007.

Engle, R. F. & Bollerslev, T. (1986). Modelling the persistence of conditional variances. Econometric Reviews, 5(1), 1-50.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalised ARCH. *Econometric Theory*, 11, 122-150.

Engle, R. and G. Lee (1993). A permanent and transitory component model of stock return volatility, Working Paper No. 92-44R, University of California San Diego.

Fama, E. F. (1965). The Behaviour of Stock-Market Prices. The Journal of Business, 38(1), 34-105.

Fang, N., Jiang, W., & Luo, R. (2017). Realized Semivariances and the Variation of Signed Jumps in China's Stock Market. Emerging Markets Finance and Trade, 53(3), 563-586.

Feunou, B., Okou, C., (2018). Good volatility, bad volatility, and option pricing. J. Finance Quantitative Analysis, 1–33.

Garman, M.B., Klass, M.J., (1980). On the estimation of security price volatilities from historical data. J. Bus. 53, 67–78.

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. Technical report.

Golosnoy, V., Gribisch, B., & Liesenfeld, R. (2015). Intra-daily volatility spillovers in international stock markets. *Journal of International Money and Finance, 53*, 95–114.

Hansen, P. R., A. Lunde, and J. M. Nason. (2003). Choosing the best volatility models: The Model Confidence Set Approach. Oxford Bulletin of Economics and Statistics, 65 (1): 839–861.

Hansen, P. R., A. Lunde, and J. M. Nason. (2011). The Model Confidence Set. Econometrica, 79 (2): 453–497.

Hansen, P. R., and Lunde, A., (2005), "A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?", *Journal of Applied Econometrics*, Vol. 20, pp. 873-889.

Hansen, P. R., and Lunde, A., (2010), "Forecasting volatility using high frequency data", A systematic review. pp. 1-37.

Hao X, An H, Zhang L, Li H, Wei G. (2015). Sentiment Diffusion of Public Opinions about Hot Events: Based on Complex Network. *PLoS ONE*, 10, e0140027.

Harvey, A. C. & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. Journal of Business and Economic Statistics, 14(4), 429-34.

Heng-Chih C, David W. (2007). Forecasting volatility on the UK stock market: A test of the conditional autoregressive range model, Journal of Money, Credit and Banking, (10):7-13.

Huang, C., Gong, X., Chen, X., Wen, F. (2013), "Measuring and Forecasting Volatility in Chinese Stock Market Using HAR-CJ-M Model", Research Article, Hindawi Publishing Corporation Abstract and Applied Analysis, Article ID 143194, pp. 1-13.

Inoue, A., Jin, L., Rossi, B. (2017). Rolling window selection for out-of-sample forecasting with time-varying parameters. J. Econometrics, 1961, 55-67.

Izzeldin, M., Kabir, M. K., Pappas, V., and Tsionas, M. (2019). Forecasting Realized Volatility Using ARFIMA and HAR Models. In: Quantitative Finance. Vol. 19, No. 10. pp. 1627-1638.

Izzeldin, Marwan, Muradoglu, Yaz, Pappas, Vasileios and Sivaprasad, Sheeja, (2021). The impact of Covid-19 on G7 stock markets volatility: Evidence from a ST-HAR model, International Review of Financial Analysis, 74, issue C, number S1057521921000144.

Kambouroudis, D. S., D. G. McMillan, and K. Tsakou. (2021). Forecasting Realized Volatility: The Role of Implied Volatility, Leverage Effect, Overnight Returns, and Volatility of Realized Volatility. Journal of Futures Markets 41 (10): 1618–1639.

Kambouroudis, D. S., McMillan, D. G. (2016). Does VIX or volume improve GARCH volatility forecasts? Applied Economics, 48(13), 1210–1228.

Kanas, A. (2013). The risk-return relation and VIX: evidence from the S&P500. Empirical Economics, 44, 1291-1314.

Kanniainen, J., B. Lin, and H. Yang. (2014). Estimating and Using GARCH Models with VIX Data for Option Valuation. *Journal of Banking & Finance* 43: 200–211.

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

Kunitomo, N. (1992). Improving the Parkinson method of estimating security price volatilities. Journal of Business, 65, 295–302.

Lamoureux, C. G. & Lastrapes, W. D. (1993). Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. Review of Financial Studies, 6(2), 293-326.

Lee, Y., Huang, Y., Wu, C. (2014). Dynamic Correlations and Volatility Spillovers between Crude Oil and Stock Index Returns: The Implications for Optimal Portfolio Construction. *International Journal of Energy Economics and Policy,* vol. 4(3), 327-336.

Li, K., & Weinbaum, D. (2001). *The empirical performance of alternative extreme value volatility estimators*. New York University, Salomon Centre, Leonard N. Stern School of Business.

Liu, J., Ma, F., Zhang, Y. (2019). Forecasting the Chinese stock volatility across global stock markets, Physica A, 525, 466-477.

Liu, L. Y., Patton, A. J., & Sheppard, K. (2015). Does anything beat 5-minute realized variance? a comparison of realized measures across multiple asset classes. Journal of Econometrics, 187(1), 293-311.

Liu, X. Y., An, H. Z., Li, H. J., Chen, Z. H., Feng, S. D., & Wen, S. B. (2017). Features of spillover networks in international financial markets: Evidence from the G20 countries. *Physica A, 479*, 265–278.

Ma, F., Li, Y., Liu, L., Zhang, Y. (2017) Are low-frequency data really uninformative? A forecasting combination perspective, North Am. J. Econ. Finance (forthcoming).

Ma, F., Liu, D., Huang, D., Chen, W. (2017). Forecasting the oil futures price volatility: A new approach, Econ. Modell. 64, 560-566.

Ma, F., Wahab, M., Liu, J., Liu, L. (2018). Is economic policy uncertainty important to forecast the realized volatility of crude oil futures? Appl. Econ. 50 (18), 2087–2101.

Mandelbrot, B. (1971). When can price be arbitraged efficiently? A limit to validity of the random walk and martingale models. Review of Economics and Statistics, 53, 225–236.

Mandelbrot, B. B. (1963). New methods in statistical economics. Journal of Political Economy, 71.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. The Journal of Business, 36.

Martens, M., & van Dijk, D. (2007). Measuring volatility with the realized range. Journal of Econometrics, 138, 181–207.

Martens, M., & Zein, J. (2004). Predicting financial volatility: high-frequency time-series forecasts vis-`a-vis implied volatility. Journal of Futures Markets, 24, 1005–1028.

McMillan, D. G. & Speight, A. E. H. (2004). Daily volatility forecasts: reassessing the performance of GARCH models. Journal of Forecasting, 23(6), 449-460.

Mei, D., Liu, J., Ma, F., Chen, W. (2017). Forecasting stock market volatility: Do realized skewness and kurtosis help? Physica A, 481, 153-159.

Mensi, W., Boubaker, F.Z., Al-Yahyaee, K.H., Kang, S.H. (2018). Dynamic volatility spillovers and connectedness between global, regional, and GIPSI stock markets. *Finance Res. Lett.*, 25, 230–238.

Merton, Robert C., (1980). On Estimating the Expected Return on the Market: An Exploratory Investigation, *Journal of Financial Economics,* 8, 1-39.

Mincer, J. A. & Zarnowitz, V. (1969). The evaluation of economic forecasts. In Economic Forecasts and Expectations: Analysis of Forecasting Behaviour and Performance, NBER Chapters (pp. 1-46). National Bureau of Economic Research, Inc.

Miralles-Marcelo, J. L., Miralles-Quirós, J. L., & Miralles-Quirós, M. M. (2013). Improving the CARR model using extreme range estimators. *Applied Financial Economics, 23*(21), 1635–1647.

Molnar, P. (2016). High-low range in GARCH models of stock return volatility. *Applied Economics, 48*(51), 4977–4991.

Molnar, P., (2012). properties of range-based volatility estimators. Int. Rev. Finance Anal. 23, 20–29.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V., Von Weizsäcker, J. E. (1997). Volatilities of different time resolutions, Analysing the dynamics of market components, Journal of Empirical Finance, 4 (2-3) (1997), pp. 213-239.

National Research Council. (2005). Network Science. Washington, DC: The National Academic Press.

Nelson, D. B. (1991). Conditional Heteroskedasticity in asset returns: A new approach. Econometrica, 59(2), pp. 347-370.

Pagan, Adrian R. and G. William Schwert. (1990). "Alternative models for Conditional Stock Market Volatility," j. Econometrics 45:1-2, pp. 267-90.

Pandey, A. (2002). Extreme-value volatility estimators and their empirical performance in Indian capital markets. NSE Research Paper, (14).

Parkinson, M., (1980). The extreme value method for estimating the variance of the rate of return. J. Bus. 53, 61–65.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics, 160(1), 246 - 256.

Patton, A. J. & Sheppard, K. (2009). Optimal combinations of realized volatility estimators. International Journal of Forecasting, 25(2), 218-238.

Patton, A. J. & Sheppard, K. (2015). Good Volatility, Bad Volatility: Signed Jumps and The Persistence of Volatility. The Review of Economics and Statistics, 97(3), 683-697.

Peng, H., Chen, R., Mei, D., Diao, X. (2018) Forecasting the realized volatility of the Chinese stock market: Do the G7 stock markets help? Physica A, 501, 78–85.

Petnehází, G., Gáll, J. (2018). Exploring the predictability of range-based volatility estimators using RNNs. arXiv preprint arXiv:1803.07152.

Piljak, V., & Swinkels, L. (2017). Frontier and emerging government bond markets. *Emerging Markets Review, 30*, 232–255.

Poon, S.-H. & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. Journal of Economic Literature, 41(2), pp. 478-539.

Raju, K., Rangaswamy S. (2017). Forecasting volatility in the Indian equity market using return and range-based models, Applied Economics, 49:49, 5027-5039.

Rejeb, A. B., & Arfaoui, M. (2016). Financial market interdependencies: A quantile regression analysis of volatility spillover. *Research in International Business and Finance, 36*, 140–157.

Rogers, L.C.G., Satchell, S.E., (1991). Estimating variance from high, low, and closing Prices. Ann. Appl. Probab. 1, 504–512.

Rossi, B., Inoue, A. (2012). Out-of-sample tests robust to the choice of window size, J. Bus. Economics and Statistics, 30, 432-453.

Sevi, B. (2014). Forecasting the volatility of crude oil futures using intraday data. European Journal of Operational Research, 235, 643–659.

Shiller, R. J. (1989). Co-movements in Stock Prices and Co-movements in Dividends. The Journal of Finance, 44(3), 719–729.

Slim, S., Tabche, I., Koubaa, Y., Osman, M., & Karathanasopoulos, A. (2023). Forecasting realized volatility of Bitcoin: The informative role of price duration. Journal of Forecasting, 1– 21.

Stock, J. H., Watson, M. W. (2004). Combination forecasts of output growth in a seven-country dataset, J. Forecast. 23 (2004), 405–430.

Todorova, N., Husmann, S., (2012). A comparative study of range-based stock return volatility estimators for the German market. J Futures Markets 32, 560–586.

Tse, Y. K., (1991), "Stock return volatility in the Tokyo stock exchange", *Japan and the World Economy*, Vol. 3, pp. 285-298.

Tse, Y. K., and Tung, S. H., (1992), "Forecasting volatility in the Singapore stock market", *Asia Pacific Journal of Management*, Vol. 9, pp. 1-13.

Walsh, D. & Tsou, G. Y.-G. (1998). Forecasting index volatility: sampling interval and non-trading effects. Applied Financial Economics, 8(5), 477-485.

Wang, H. (2019). VIX and volatility forecasting: A new insight. Physica A, 533, 121951.

Wang, J. N., Hsu, Y. T., & Liu, H. C. (2014). How useful are the various volatility estimators for Improving GARCH-based volatility forecasts? Evidence from the Nasdaq-100 Stock Index. *International Journal of Economics and Financial Issues, 4*(3), 651–656.

Wang, J., Lu, X., He, F., & Ma, F. (2020). Which popular predictor is more useful to forecast international stock markets during the coronavirus pandemic: VIX vs EPU?, *International Review of Financial Analysis*, Volume 72, 1057-5219.

Wang, X., Wu, C., & Xu, W. (2015). Volatility forecasting: The role of lunch-break returns, overnight returns, trading volume and leverage effects. *International Journal of Forecasting*, 31(3), 609-619.

Watts, D., Strogatz, S., (1998). Collective dynamics of 'small-world' networks. *Nature,* 393, 440-442.
Weiping, Z., Zhuang X., Dongmei, W. (2020). Spatial connectedness of volatility spillovers in G20 stock markets: Based on block models analysis. *Finance Research Letters*, 34, 101274.

Wu, X. & Hou, X. (2019). Forecasting realized variance using asymmetric HAR model with time-varying coefficients. Finance Research Letters, Vol. 30, pp. 89-95.

Yang, D., Zhang, Q., (2000). Drift independent volatility estimation based on high, low, open and closes prices. J. Bus. 73, 477–491.

Yang, M. J., and M.-Y. Liu. (2012). The Forecasting Power of the Volatility Index in Emerging Markets: Evidence from the Taiwan Stock Market. *International Journal of Economics and Finance* 4 (2): 217–231.

Yarovaya, L., Brzeczczynski, J., Lau, C.K.M., (2016). Volatility spillovers across stock index futures in Asian markets: evidence from range volatility estimators. Finance Res. Lett. 17, 158–166.

Yoon, S. M., Mamum, M. A., Uddin, G. S., & Kang, S. H. (2019). Network connectedness and net spillover between financial and commodity markets. *North American Journal of Economics and Finance, 48*, 801–818.

Yu, L., Li, J., Tang, L. (2015). Dynamic volatility spillover effect analysis between carbon market and crude oil market a DCC-ICSS approach. *Int. J. Global Energy Issues* 38, 242–256.

Yu, M. (2019). Forecasting Bitcoin volatility: The role of leverage effect and uncertainty. Physica A, 533, 120707.

Zhang, W., Zhuang, X., Lu, Y., Wang, J. (2020). Spatial linkage of volatility spillovers and its explanation across G20 stock markets: A network framework. *International Review of Financial Analysis,* 71, 101454.

Zhang, W., Zhuang, X., Wang, J., Lu, Y. (2020). Connectedness and systemic risk spillovers analysis of Chinese sectors based on tail risk network, *North American Journal of Economics and Finance*, 54, 10124.

Appendices:

Chapter-2

Table 1: QLIKE for rolling windows forecast models

	GARCH	EGARCH	PGARCH	TGARCH	HAR-RV	HAR-PS	HAR-RSV	ARFIMA-RV
AEX	-8.7107	-8.7298*	-8.7437	-8.7120	**-8.8382****	-8.8361	-8.8287	-8.6937
AORD	-9.1450	-9.1610	-9.1707*	-9.1465	-9.2734	**-9.2779****	-9.2577	-8.8708
BFX	-8.6286	-8.7225	-8.7361*	-8.7272	**-8.8446****	-8.8441	-8.8411	-8.7465
BSESN	-8.7091	-8.7404*	-8.7180	-8.6927	**-8.8565****	-8.8540	-8.8489	-8.7369
BVLG	-8.9305	-8.9458	-8.9627*	-8.9592	**-9.2641****	-9.2430	-9.1782	-9.2085
BVSP	-8.0419	-8.0106	-8.0432*	-8.0346	**-8.3046****	-8.2847	-8.2970	-8.1859
DJI	-8.9317	-8.9478	**-8.9838****	-8.9635	-8.9827	-8.9319*	-8.9108	-8.9077
FCHI	-8.3810	-8.4446*	-8.4342	-8.4100	**-8.5473****	-8.5471	-8.5425	-8.3937
FTMIB	-8.0436	-8.0573	-8.0743*	-8.0602	**-8.3467****	-8.3413	-8.3226	-8.2210
FTSE	-8.6696	-8.6517	-8.6737*	-8.6058	-8.6841	**-8.6930****	-8.2896	-8.4948
GDAXI	-8.3645	-8.4054	-8.4090*	-8.2687	-8.5083	**-8.5110****	-8.4845	-8.2662
GSPTSE	-9.2440	-9.3235*	-9.3108	-9.2641	-9.3515	**-9.3571****	-9.3409	-9.2069
HSI	-8.4927	-8.5821	-8.6213*	-8.6151	-8.8635	**-8.8636****	-8.8613	-5.4198
IBEX	-8.0355	-8.0375*	-7.8975	-8.0201	**-8.1109****	-8.1082	-8.1007	-8.0316
IXIC	-8.8146	-8.8373	-8.8676*	-8.7976	**-9.0537****	-8.9816	-8.8783	-8.9988
KS11	-9.0523	-9.0275	-9.0843*	-9.0517	-9.2587	-9.2604	**-9.2641****	-9.2024
KSE	-8.8437*	-8.8180	-8.8246	-8.8152	**-9.0120****	-9.0099	-8.9948	-7.8012
MXX	-8.7058	-8.7425	-8.7554*	-8.7025	-8.7678	**-8.7743****	-8.7488	-8.5656
N225	-8.4473	-8.4494	-8.4768*	-8.4650	**-8.7590****	-8.7362	-8.7518	-8.3652
NSEI	-8.5931	-8.7597	-8.7681*	-8.7075	**-8.8752****	-8.8738	-8.8652	-8.7519
OMXC20	-8.4141	-8.4376	-8.4384*	-8.4344	**-8.5103****	-8.5063	-8.4616	-8.3132
OMXHPI	-8.7209	-8.6853	-8.7198	-8.7267*	-8.9772	**-8.9780****	-8.9693	-8.9271
OMXSPI	-8.8722	-8.9059*	-8.8989	-8.8924	**-9.1111****	-9.1056	-9.0925	-8.5845
OSEAX	-8.5447	-8.5548	-8.5657*	-8.5161	**-8.6173****	-8.3806	-8.5349	-8.3525
RUT	-8.6318	-8.6819*	-8.6745	-8.6440	-8.9418	**-8.9420****	-8.9102	-8.8375
SMSI	-8.1292	-8.1448	-8.1474*	-8.1018	**-8.2256****	-7.9472	-8.1851	-8.1601
SPX	-8.9751	-8.9787	-9.0417*	-8.9568	**-9.0802****	-9.0671	-8.9199	-8.9044
SSEC	-8.1701	-8.1341	-8.1699	-8.1868*	**-8.3591****	-8.3450	-8.2361	-8.2106
SSMI	-8.9251	-8.9198	-8.9378*	-8.8715	**-9.0732****	-9.0680	-9.0413	-8.9806
STOXX50E	-8.2579	-8.3149	-8.2995*	-8.0975	-8.3473	**-8.3586****	-8.0329	-8.1420

Notes: *** with bold indicates the best performing model, underlined with * is the best performing GARCH model, regardless of the HAR and ARFIMA models.*

Table 2: MSE for rolling windows forecast models

	GARCH	EGARCH	PGARCH	TGARCH	HAR-RV	HAR-PS	HAR-RSV	ARFIMA-RV
AEX	1.32E-08	9.77E-09*	1.19E-08	1.31E-08	**8.27E-09****	8.45E-09	8.46E-09	9.01E-09
AORD	4.15E-09	3.85E-09*	4.36E-09	4.05E-09	2.46E-09	2.43E-09	**2.35E-09****	2.79E-09
BFX	1.11E-08	8.73E-09*	8.99E-09	1.23E-08	**4.68E-09****	4.96E-09	5.01E-09	5.03E-09
BSESN	4.85E-09	4.64E-09*	5.69E-09	5.06E-09	2.99E-09	**2.98E-09****	3.09E-09	3.30E-09
BVLG	9.80E-09	1.16E-08	9.87E-09	9.76E-09*	**1.12E-09****	1.16E-09	1.37E-09	1.48E-09
BVSP	2.93E-08*	3.63E-08	4.00E-08	3.31E-08	**1.01E-08****	1.04E-08	1.03E-08	1.04E-08
DJI	2.45E-08	2.35E-08	3.00E-08	**2.27E-08****	2.84E-08*	2.96E-08	3.98E-08	3.41E-08
FCHI	2.03E-08	1.45E-08*	2.02E-08	2.00E-08	1.01E-08	**1.00E-08****	**1.00E-08****	1.14E-08
FTMIB	5.83E-08	4.97E-08*	9.06E-08	7.77E-08	**1.17E-08****	1.26E-08	1.25E-08	1.70E-08
FTSE	2.50E-08	2.42E-08	2.42E-08	**2.40E-08****	2.62E-08*	2.78E-08	2.83E-08	3.64E-08
GDAXI	1.84E-08	1.45E-08*	1.68E-08	1.69E-08	1.27E-08	1.32E-08	1.37E-08	**1.26E-08****
GSPTSE	8.57E-09	**7.68E-09****	8.40E-09	8.05E-09	9.43E-09*	1.01E-08	1.39E-08	1.15E-08
HSI	1.71E-08	1.41E-08*	1.54E-08	1.55E-08	**4.51E-09****	4.55E-09	4.64E-09	5.42E-09
IBEX	4.38E-08	4.10E-08	3.80E-08*	4.89E-08	**2.88E-08****	3.09E-08	3.08E-08	3.68E-08
IXIC	1.38E-08	9.88E-09*	1.24E-08	1.39E-08	6.36E-09	6.19E-09	**6.09E-09****	8.29E-09
KS11	1.17E-08	8.11E-09*	9.73E-09	1.22E-08	**3.92E-09****	4.39E-09	4.55E-09	4.17E-09
KSE	6.71E-09*	7.75E-09	8.13E-09	9.42E-09	2.74E-09	**2.69E-09****	2.77E-09	3.75E-09
MXX	2.14E-08	**1.97E-08****	2.08E-08	2.11E-08	2.09E-08*	2.43E-08	3.28E-08	2.10E-08
N225	3.78E-08	2.93E-08*	3.95E-08	4.01E-08	**1.46E-08****	1.63E-08	1.64E-08	1.78E-08
NSEI	6.08E-09	5.87E-09*	7.01E-09	6.63E-09	3.96E-09	**3.95E-09****	4.06E-09	4.50E-09
OMXC20	4.12E-08	4.06E-08*	4.18E-08	4.22E-08	**3.94E-08****	4.33E-08	4.28E-08	4.12E-08
OMXHPI	1.70E-08	1.87E-08	1.98E-08	1.66E-08*	**3.71E-09****	3.73E-09	3.80E-09	3.84E-09
OMXSPI	1.65E-08	1.25E-08*	1.26E-08	1.73E-08	**7.34E-09****	7.83E-09	7.93E-09	8.52E-09
OSEAX	2.05E-08	**1.71E-08****	1.91E-08	1.99E-08	1.73E-08	**1.72E-08****	1.73E-08	2.07E-08
RUT	2.81E-08	1.35E-08*	2.07E-08	2.11E-08	1.05E-08	9.47E-09	9.67E-09	**9.27E-09****
SMSI	6.47E-08	5.96E-08*	6.24E-08	6.55E-08	**4.95E-08****	5.26E-08	5.46E-08	5.16E-08
SPX	1.52E-08	**1.24E-08****	1.46E-08	1.49E-08	1.55E-08*	1.64E-08	1.84E-08	1.71E-08
SSEC	6.27E-08*	6.73E-08	7.16E-08	6.84E-08	4.61E-08	**4.46E-08****	4.58E-08	6.39E-08
SSMI	2.56E-08	2.20E-08*	7.42E-08	9.42E-08	**1.54E-08****	1.55E-08	1.99E-08	2.25E-08
STOXX50E	3.40E-08	3.10E-08*	3.48E-08	3.27E-08	**2.64E-08****	2.71E-08	2.79E-08	2.76E-08

Notes: *** with bold indicates the best performing model, underlined with * is the best performing GARCH model, regardless of the HAR and ARFIMA models.*

Table 3: MAE for rolling windows forecast models

	GARCH	EGARCH	PGARCH	TGARCH	HAR-RV	HAR-PS	HAR-RSV	ARFIMA-RV
AEX	6.05E-05	5.38E-05*	5.65E-05	6.03E-05	3.31E-05	**3.30E-05****	3.38E-05	4.19E-05
AORD	4.01E-05	3.91E-05*	6.06E-05	4.01E-05	2.20E-05	**2.18E-05****	2.26E-05	2.69E-05
BFX	5.88E-05	5.45E-05	5.32E-05*	5.66E-05	2.81E-05	**2.80E-05****	2.82E-05	3.34E-05
BSESN	4.53E-05	4.39E-05*	4.68E-05	4.58E-05	2.62E-05	**2.61E-05****	2.67E-05	3.13E-05
BVLG	7.00E-05	6.92E-05	6.60E-05	6.50E-05*	**1.66E-05****	1.69E-05	1.90E-05	1.89E-05
BVSP	0.000130	0.000139	0.000134	0.000129*	**4.54E-05****	4.59E-05	4.64E-05	5.30E-05
DJI	4.99E-05	4.50E-05*	4.56E-05	4.87E-05	**4.47E-05****	4.53E-05	5.15E-05	5.08E-05
FCHI	8.17E-05	7.33E-05*	8.07E-05	8.18E-05	4.20E-05	**4.19E-05****	4.27E-05	5.42E-05
FTMIB	0.000157	0.000150*	0.000157	0.000154	5.07E-05	**5.05E-05****	5.15E-05	7.03E-05
FTSE	5.04E-05	**4.57E-05****	4.88E-05	4.86E-05	4.80E-05	4.79E-05	4.84E-05	5.74E-05*
GDAXI	7.89E-05	6.97E-05*	7.38E-05	7.87E-05	4.57E-05	**4.54E-05****	4.61E-05	5.93E-05
GSPTSE	3.54E-05	3.05E-05*	3.39E-05	3.33E-05	2.59E-05	**2.58E-05****	2.85E-05	3.23E-05
HSI	8.86E-05	8.35E-05*	8.47E-05	8.57E-05	**2.73E-05****	2.74E-05	2.78E-05	3.44E-05
IBEX	9.47E-05	9.16E-05*	9.23E-05	9.68E-05	**6.31E-05****	6.38E-05	6.34E-05	8.18E-05
IXIC	6.76E-05	6.13E-05*	6.26E-05	7.00E-05	**2.95E-05****	3.02E-05	3.04E-05	3.35E-05
KS11	5.50E-05	4.97E-05	4.93E-05*	5.52E-05	**1.95E-05****	1.99E-05	2.03E-05	2.30E-05
KSE	5.84E-05*	6.20E-05	6.39E-05	6.54E-05	2.71E-05	**2.70E-05****	2.72E-05	3.50E-05
MXX	5.22E-05	4.81E-05*	5.05E-05	4.95E-05	**4.12E-05****	4.24E-05	4.67E-05	4.60E-05
N225	0.000126	0.000118*	0.000128	0.000125	**4.43E-05****	4.60E-05	4.57E-05	5.67E-05
NSEI	4.87E-05	4.70E-05*	5.10E-05	4.95E-05	2.74E-05	**2.73E-05****	2.81E-05	3.40E-05
OMXC20	7.44E-05	7.20E-05*	7.33E-05	7.49E-05	**5.57E-05****	5.67E-05	5.77E-05	6.24E-05
OMXHPI	8.28E-05	9.09E-05	8.72E-05	8.06E-05*	2.67E-05	**2.66E-05****	2.70E-05	3.01E-05
OMXSPI	7.18E-05	6.62E-05*	6.81E-05	7.01E-05	**2.70E-05****	2.71E-05	2.90E-05	3.59E-05
OSEAX	6.75E-05	6.19E-05*	6.52E-05	6.58E-05	4.87E-05	**4.85E-05****	4.91E-05	6.05E-05
RUT	9.45E-05	8.03E-05*	8.65E-05	8.76E-05	3.43E-05	**3.30E-05****	3.36E-05	3.92E-05
SMSI	0.000102	9.41E-05*	0.000104	9.93E-05	6.58E-05	**6.41E-05****	6.52E-05	7.83E-05
SPX	5.14E-05	4.47E-05*	4.81E-05	5.27E-05	**4.06E-05****	4.08E-05	4.40E-05	4.73E-05
SSEC	0.00012*	0.00013	0.00013	0.00013	7.39E-05	**7.19E-05****	7.39E-05	9.38E-05
SSMI	5.69E-05	5.04E-05*	5.91E-05	6.25E-05	2.62E-05	**2.59E-05****	2.88E-05	3.24E-05
STOXX50E	8.40E-05	7.44E-05*	8.31E-05	8.35E-05	5.77E-05	**5.65E-05****	5.68E-05	6.86E-05

Notes: *** with bold indicates the best performing model, underlined with * is the best performing GARCH model, regardless of the HAR and ARFIMA models.*

Table 4: Annualised volatility and geographically distributed indices

	Annualised Conditional volatility	Annualised Realized volatility	**Europe**	**Asia**	**America**	**Australia**
AEX	0.16	0.13	AEX	BSESN	BVSP	AORD
AORD	*0.12*	*0.10*	BFX	HSI	DJI	
BFX	*0.16*	*0.12*	BVLG	KS11	GSPTSE	
BSESN	*0.15*	*0.12*	FCHI	KSE	NASDAQ	
BVLG	*0.16*	*0.10*	FTMIB	N225	MXX	
BVSP	*0.22*	*0.15*	FTSE	NSEI	SPX	
DJI	*0.14*	*0.13*	GDAXI	SSEC		
FCHI	*0.19*	*0.15*	IBEX			
FTMIB	*0.24*	*0.17*	OMXC20			
FTSE	*0.14*	*0.14*	OMXHPI			
GDAXI	*0.19*	*0.15*	OMXSPI			
GSPTSE	*0.12*	*0.10*	OSEAX			
HSI	*0.18*	*0.12*	RUT			
IBEX	*0.21*	*0.19*	SMSI			
IXIC	*0.17*	*0.12*	SSMI			
KS11	*0.14*	*0.10*	STOXX50E			
KSE	*0.15*	*0.11*				
MXX	*0.14*	*0.12*				
N225	*0.20*	*0.13*				
NSEI	*0.15*	*0.12*				
OMXC20	*0.17*	*0.14*				
OMXHPI	*0.18*	*0.13*				
OMXSPI	*0.17*	*0.12*				
OSEAX	*0.17*	*0.14*				
RUT	*0.19*	*0.13*				
SMSI	*0.21*	*0.18*				
SPX	*0.14*	*0.13*				
SSEC	*0.21*	*0.18*				
SSMI	*0.15*	*0.11*				
STOXX50E	*0.19*	*0.17*				

Table 5: GW test results for the rolling windows

ROLLING	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
AEX							
GARCH	0.000 (+)	0.134 (+)	0.047 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.097 (-)	0.192 (-)	0.272 (-)
ARFIMA-RV	-	-	-	-	-	0.211 (+)	0.138 (+)
HAR-RSV	-	-	-	-	-	-	0.000 (+)
AORD							
GARCH	0.000 (+)	0.000 (-)	0.006 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.009 (-)	0.050 (+)	0.025 (+)
ARFIMA-RV	-	-	-	-	-	0.046 (+)	0.014 (+)
HAR-RSV	-	-	-	-	-	-	0.029 (-)
BFX							
GARCH	0.004 (+)	0.023 (+)	0.014 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.001 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.334 (-)	0.424 (-)
ARFIMA-RV	-	-	-	-	-	0.035 (+)	0.016 (+)
HAR-RSV	-	-	-	-	-	-	0.220 (+)
BSESN							
GARCH	0.009 (+)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.017 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.046 (-)	0.059 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.046 (+)

ROLLING	*EGARCH*	*PGARCH*	*TGARCH*	*HAR-RV*	*ARFIMA-RV*	*HAR-RSV*	*HAR-PS*
BVLG							
GARCH	0.070 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.015 (+)	0.050 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.001 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.000 (-)	0.001 (+)
ARFIMA-RV	-	-	-	-	-	0.001 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.000 (+)
BVSP							
GARCH	0.385 (-)	0.000 (-)	0.149 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.032 (-)	0.300 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.002 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.325 (-)	0.118 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.030 (+)
DJI							
GARCH	0.079 (+)	0.004 (+)	0.440 (+)	0.151 (-)	0.514 (-)	0.067 (-)	0.171 (-)
EGARCH	-	0.192 (-)	0.001 (-)	0.092 (-)	0.319 (-)	0.022 (-)	0.094 (-)
PGARCH	-	-	0.051 (+)	0.119 (-)	0.404 (-)	0.032 (-)	0.128 (-)
TGARCH	-	-	-	0.161 (-)	0.483 (-)	0.049 (-)	0.120 (-)
HAR-RV	-	-	-	-	0.155 (-)	0.122 (-)	0.225 (-)
ARFIMA-RV	-	-	-	-	-	0.225 (-)	0.169 (+)
HAR-RSV	-	-	-	-	-	-	0.384 (+)
FCHI							
GARCH	0.000 (+)	0.009 (+)	0.015 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.033 (+)	0.038 (+)
ARFIMA-RV	-	-	-	-	-	0.013 (+)	0.007 (+)
HAR-RSV	-	-	-	-	-	-	0.018 (+)

ROLLING	*EGARCH*	*PGARCH*	*TGARCH*	*HAR-RV*	*ARFIMA-RV*	*HAR-RSV*	*HAR-PS*
FTMIB							
GARCH	0.114 (+)	0.467 (-)	0.575 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.120 (-)	0.132 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.267 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.507 (-)	0.419 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.115 (-)
FTSE							
GARCH	0.009 (+)	0.076 (+)	0.045 (+)	0.202 (-)	0.246 (-)	0.281 (-)	0.151 (-)
EGARCH	-	0.000 (-)	0.006 (-)	0.251 (-)	0.003 (-)	0.183 (-)	0.217 (-)
PGARCH	-	-	0.000 (+)	0.286 (-)	0.022 (-)	0.324 (-)	0.239 (-)
TGARCH	-	-	-	0.225 (-)	0.027 (-)	0.299 (-)	0.228 (-)
HAR-RV	-	-	-	-	0.173 (-)	0.326 (-)	0.453 (-)
ARFIMA-RV	-	-	-	-	-	0.587 (+)	0.590(+)
HAR-RSV	-	-	-	-	-	-	0.239(+)
GDAXI							
GARCH	0.003 (+)	0.002 (+)	0.030 (+)	0.000 (+)	0.003 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.003 (+)	0.096 (+)	0.005 (+)	0.005 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (+)	0.069 (-)	0.326 (-)
ARFIMA-RV	-	-	-	-	-	0.011 (-)	0.000 (-)
HAR-RSV	-	-	-	-	-	-	0.077 (+)
GSPTSE							
GARCH	0.000 (+)	0.005 (+)	0.003 (+)	0.017 (-)	0.396 (-)	0.245 (-)	0.010 (-)
EGARCH	-	0.000 (-)	0.000 (-)	0.522 (-)	0.132 (-)	0.159 (-)	0.581 (-)
PGARCH	-	-	0.000 (+)	0.090 (-)	0.564 (-)	0.257 (-)	0.080 (-)
TGARCH	-	-	-	0.308 (-)	0.445 (-)	0.253 (-)	0.256 (-)
HAR-RV	-	-	-	-	0.017 (-)	0.161 (-)	0.112 (-)
ARFIMA-RV	-	-	-	-	-	0.275 (-)	0.002 (+)
HAR-RSV	-	-	-	-	-	-	0.404 (+)

ROLLING	*EGARCH*	*PGARCH*	*TGARCH*	*HAR-RV*	*ARFIMA-RV*	*HAR-RSV*	*HAR-PS*
HSI							
GARCH	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.224 (-)	0.095 (-)
ARFIMA-RV	-	-	-	-	-	0.013 (+)	0.004 (+)
HAR-RSV	-	-	-	-	-	-	0.186 (+)
IBEX							
GARCH	0.009 (+)	0.061 (+)	0.003 (-)	0.000 (+)	0.008 (+)	0.002 (+)	0.000 (+)
EGARCH	-	0.055 (+)	0.029 (-)	0.000 (+)	0.054 (+)	0.007 (+)	0.007 (+)
PGARCH	-	-	0.052 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.022 (+)	0.001 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.469 (-)	0.474 (-)
ARFIMA-RV	-	-	-	-	-	0.071 (+)	0.066 (+)
HAR-RSV	-	-	-	-	-	-	0.113 (-)
IXIC							
GARCH	0.000 (+)	0.000 (+)	0.001 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.434 (-)	0.235 (+)	0.178 (+)
ARFIMA-RV	-	-	-	-	-	0.284 (+)	0.310 (+)
HAR-RSV	-	-	-	-	-	-	0.187 (-)
KS11							
GARCH	0.000 (+)	0.000 (+)	0.018 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.223 (-)	0.418 (-)	0.436 (-)
ARFIMA-RV	-	-	-	-	-	0.295 (-)	0.274 (-)
HAR-RSV	-	-	-	-	-	-	0.115 (-)

ROLLING	***EGARCH***	***PGARCH***	***TGARCH***	***HAR-RV***	***ARFIMA-RV***	***HAR-RSV***	***HAR-PS***
KSE							
GARCH	0.000 (-)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.000 (-)	0.010 (+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.009 (+)
MXX							
GARCH	0.000 (+)	0.005 (+)	0.153 (+)	0.001 (+)	0.001 (+)	0.317 (-)	0.015 (-)
EGARCH	-	0.001 (-)	0.000 (-)	0.316 (-)	0.088 (-)	0.123 (-)	0.546 (-)
PGARCH	-	-	0.004 (-)	0.017 (-)	0.001 (-)	0.250 (-)	0.144 (-)
TGARCH	-	-	-	0.012 (+)	0.002 (+)	0.286 (-)	0.105 (-)
HAR-RV	-	-	-	-	0.016 (-)	0.119 (-)	0.568 (-)
ARFIMA-RV	-	-	-	-	-	0.184 (-)	0.169 (-)
HAR-RSV	-	-	-	-	-	-	0.382(+)
N225							
GARCH	0.000 (+)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.191 (-)	0.187 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.090 (+)
NSEI							
GARCH	0.003 (+)	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.115 (-)	0.077 (-)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.097(+)

ROLLING	EGARCH	PGARCH	TGARCH	HAR-RV	ARFIMA-RV	HAR-RSV	HAR-PS
OMXC20							
GARCH	0.000 (+)	0.000 (-)	0.001 (-)	0.000 (+)	0.000 (+)	0.191 (-)	0.047 (-)
EGARCH	-	0.001 (-)	0.000 (-)	0.000 (+)	0.015 (-)	0.414 (-)	0.476 (-)
PGARCH	-	-	0.005 (-)	0.000 (+)	0.000 (+)	0.100 (-)	0.050 (-)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.060 (-)	0.031 (-)
HAR-RV	-	-	-	-	0.226 (-)	0.192 (-)	0.579 (-)
ARFIMA-RV	-	-	-	-	-	0.311 (-)	0.167 (-)
HAR-RSV	-	-	-	-	-	-	0.099 (-)
OMXHPI							
GARCH	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.007 (-)	0.006 (-)	0.000 (-)
ARFIMA-RV	-	-	-	-	-	0.026 (+)	0.003 (+)
HAR-RSV	-	-	-	-	-	-	0.006(+)
OMXSPI							
GARCH	0.000 (+)	0.047 (+)	0.006 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.114 (+)	0.002 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.012 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.150 (-)	0.427 (-)
ARFIMA-RV	-	-	-	-	-	0.005 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.004(+)
OSEAX							
GARCH	0.001 (+)	0.130 (+)	0.000 (+)	0.000 (+)	0.031 (-)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.011 (-)	0.003 (+)	0.107 (-)	0.009 (-)	0.001 (+)
PGARCH	-	-	0.019 (-)	0.000 (+)	0.175 (-)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.170 (-)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.000 (-)	0.050 (-)	0.012(+)
ARFIMA-RV	-	-	-	-	-	0.000 (+)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.095(+)

ROLLING	*EGARCH*	*PGARCH*	*TGARCH*	*HAR-RV*	*ARFIMA-RV*	*HAR-RSV*	*HAR-PS*
RUT							
GARCH	0.000 (+)	0.000 (+)	0.003 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.179 (+)	0.134 (+)	0.233(+)
ARFIMA-RV	-	-	-	-	-	0.222 (-)	0.226 (-)
HAR-RSV	-	-	-	-	-	-	0.267 (+)
SMSI							
GARCH	0.033 (+)	0.004 (+)	0.006 (-)	0.004 (+)	0.014 (+)	0.002 (+)	0.001 (+)
EGARCH	-	0.000 (-)	0.042 (-)	0.003 (+)	0.052 (+)	0.007 (+)	0.001 (+)
PGARCH	-	-	0.037 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.004 (+)	0.005 (+)	0.003 (+)	0.001 (+)
HAR-RV	-	-	-	-	0.046 (-)	0.351 (-)	0.221 (-)
ARFIMA-RV	-	-	-	-	-	0.116 (-)	0.030 (-)
HAR-RSV	-	-	-	-	-	-	0.275(+)
SPX							
GARCH	0.021 (+)	0.000 (+)	0.147 (+)	0.048 (-)	0.079 (-)	0.200 (-)	0.149 (-)
EGARCH	-	0.004 (-)	0.034 (-)	0.208 (-)	0.530 (-)	0.134 (-)	0.285 (-)
PGARCH	-	-	0.064 (-)	0.091 (-)	0.276 (-)	0.214 (-)	0.239 (-)
TGARCH	-	-	-	0.098 (-)	0.047 (-)	0.157 (-)	0.085 (-)
HAR-RV	-	-	-	-	0.354 (-)	0.082 (-)	0.143 (-)
ARFIMA-RV	-	-	-	-	-	0.149 (-)	0.219 (+)
HAR-RSV	-	-	-	-	-	-	0.417 (+)
SSEC							
GARCH	0.002 (-)	0.002 (-)	0.001 (-)	0.000 (+)	0.000 (-)	0.000 (+)	0.000 (+)
EGARCH	-	0.032 (-)	0.000 (-)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
PGARCH	-	-	0.002 (+)	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.000 (+)	0.000 (+)	0.000 (+)
HAR-RV	-	-	-	-	0.089 (-)	0.060 (+)	0.017 (+)
ARFIMA-RV	-	-	-	-	-	0.183 (+)	0.072 (+)
HAR-RSV	-	-	-	-	-	-	0.044 (+)

ROLLING	***EGARCH***	***PGARCH***	***TGARCH***	***HAR-RV***	***ARFIMA-RV***	***HAR-RSV***	***HAR-PS***
SSMI							
GARCH	0.047 (+)	0.590 (-)	0.183 (-)	0.009 (+)	0.005 (+)	0.010 (+)	0.019 (+)
EGARCH	-	0.564 (-)	0.209 (-)	0.095 (+)	0.028 (-)	0.045 (+)	0.097 (+)
PGARCH	-	-	0.078 (-)	0.035 (+)	0.427 (+)	0.389 (+)	0.076 (+)
TGARCH	-	-	-	0.072 (+)	0.158 (+)	0.166 (+)	0.085 (+)
HAR-RV	-	-	-	-	0.216 (-)	0.042 (-)	0.074 (-)
ARFIMA-RV	-	-	-	-	-	0.065 (+)	0.256 (+)
HAR-RSV	-	-	-	-	-	-	0.032(+)
STOXX50							
GARCH	0.000 (+)	0.000 (-)	0.054 (+)	0.006 (+)	0.031 (+)	0.006 (+)	0.003 (+)
EGARCH	-	0.000 (-)	0.000 (-)	0.013 (+)	0.086 (+)	0.018 (+)	0.008 (+)
PGARCH	-	-	0.000 (+)	0.000 (+)	0.003 (+)	0.000 (+)	0.000 (+)
TGARCH	-	-	-	0.000 (+)	0.003 (+)	0.000 (+)	0.000(+)
HAR-RV	-	-	-	-	0.000 (-)	0.012 (-)	0.255 (-)
ARFIMA-RV	-	-	-	-	-	0.003 (-)	0.000 (+)
HAR-RSV	-	-	-	-	-	-	0.058(+)

Chapter-3:

Table 1: RECURSIVE WINDOW FORECASTING RESULTS MEASURED BY QLIKE AND HMSE AND MCS TEST

CAC	**QLIKE**	**p-value**	**Rank**	$MZ\text{-}R^2$	**CAC**	**HMSE**	**p-value**	**Rank**
PK (Model-free)	0.2360	eliminated	–	0.721	**PK (Model-free)**	0.3508	1.0000	3
GK (Model-free)	0.1947	eliminated	–	0.793	**GK (Model-free)**	**0.2168**	**1.0000**	**1**
RS (Model-free)	0.6907	eliminated	–	0.788	**RS (Model-free)**	0.2741	1.0000	2
YZ (Model-free)	0.1798	eliminated	–	0.560	**YZ (Model-free)**	3.6710	eliminated	–
HAR-RV	0.1559	eliminated	–	0.499	**HAR-RV**	0.7151	eliminated	–
HAR-RV-ONV	**0.1427**	**1.0000**	**1**	0.529	**HAR-RV-ONV**	0.6886	1.0000	4
HAR-RV-CC	0.1557	eliminated	–	0.512	**HAR-RV-CC**	0.7192	eliminated	–
HAR-RV-PK	0.1565	eliminated	–	0.530	**HAR-RV-PK**	0.6902	0.9962	6
HAR-RV-GK	0.1558	eliminated	–	0.495	**HAR-RV-GK**	0.7140	eliminated	–
HAR-RV-RS	0.1566	eliminated	–	0.489	**HAR-RV-RS**	0.6983	0.7078	7
HAR-RV-YZ	0.1564	eliminated	–	0.505	**HAR-RV-YZ**	0.7239	eliminated	–
HAR-RV-COMB.	0.1494	eliminated	–	0.499	**HAR-RV-COMB.**	0.6895	1.0000	5
DAX	**QLIKE**	**p-value**	**Rank**	$MZ\text{-}R^2$	**DAX**	**HMSE**	**p-value**	**Rank**
PK (Model-free)	0.2141	eliminated	–	0.731	**PK (Model-free)**	0.3707	1.0000	3
GK (Model-free)	0.1620	1.0000	2	0.742	**GK (Model-free)**	**0.2238**	**1.0000**	**1**
RS (Model-free)	0.4212	eliminated	–	0.715	**RS (Model-free)**	0.2757	1.0000	2
YZ (Model-free)	0.1963	eliminated	–	0.430	**YZ (Model-free)**	4.7830	eliminated	–
HAR-RV	0.1623	eliminated	–	0.473	**HAR-RV**	0.7545	1.0000	4
HAR-RV-ONV	**0.1557**	**1.0000**	**1**	0.494	**HAR-RV-ONV**	0.7574	1.0000	5
HAR-RV-CC	0.1661	eliminated	–	0.486	**HAR-RV-CC**	0.7786	0.6164	11
HAR-RV-PK	0.1687	eliminated	–	0.511	**HAR-RV-PK**	0.7711	0.9390	10
HAR-RV-GK	0.1683	eliminated	–	0.501	**HAR-RV-GK**	0.8034	0.0266	7
HAR-RV-RS	0.1656	eliminated	–	0.474	**HAR-RV-RS**	0.8016	0.0004	9
HAR-RV-YZ	0.1660	eliminated	–	0.461	**HAR-RV-YZ**	0.8083	0.0586	12
HAR-RV-COMB.	0.1614	eliminated	–	0.473	**HAR-RV-COMB.**	0.7684	1.0000	6
FTSE	**QLIKE**	**p-value**	**Rank**	$MZ\text{-}R^2$	**FTSE**	**HMSE**	**p-value**	**Rank**
PK (Model-free)	**0.2464**	**1.0000**	**1**	0.708	**PK (Model-free)**	0.4672	1.0000	2
GK (Model-free)	0.3033	eliminated	–	0.760	**GK (Model-free)**	**0.4543**	**1.0000**	**1**
RS (Model-free)	3.9910	eliminated	–	0.585	**RS (Model-free)**	0.6183	1.0000	3
YZ (Model-free)	0.2728	1.0000	2	0.627	**YZ (Model-free)**	1.8960	1.0000	4
HAR-RV	0.2605	0.9248	5	0.239	**HAR-RV**	2.1279	0.0300	12
HAR-RV-ONV	0.2620	0.7684	6	0.239	**HAR-RV-ONV**	2.0387	1.0000	8
HAR-RV-CC	0.2580	1.0000	2	0.263	**HAR-RV-CC**	2.0605	1.0000	9
HAR-RV-PK	0.2585	1.0000	4	0.277	**HAR-RV-PK**	1.9706	1.0000	5
HAR-RV-GK	0.2612	1.0000	3	0.243	**HAR-RV-GK**	2.0053	1.0000	6
HAR-RV-RS	0.2634	0.6496	7	0.231	**HAR-RV-RS**	2.1116	0.0280	11
HAR-RV-YZ	0.2669	0.0086	8	0.243	**HAR-RV-YZ**	2.0778	1.0000	10
HAR-RV-COMB.	0.2581	1.0000	3	0.239	**HAR-RV-COMB.**	2.0020	1.0000	7
FTSEMIB	**QLIKE**	**p-value**	**Rank**	$MZ\text{-}R^2$	**FTSEMIB**	**HMSE**	**p-value**	**Rank**
PK (Model-free)	0.1537	eliminated	–	0.572	**PK (Model-free)**	2.1890	eliminated	–
GK (Model-free)	**0.1243**	**1.0000**	**1**	0.711	**GK (Model-free)**	1.4530	1.0000	2
RS (Model-free)	0.8281	eliminated	–	0.511	**RS (Model-free)**	**1.3870**	**1.0000**	**1**
YZ (Model-free)	0.2326	eliminated	–	0.600	**YZ (Model-free)**	8.7150	eliminated	–
HAR-RV	0.1480	0.3120	5	0.509	**HAR-RV**	1.8228	0.6864	6
HAR-RV-ONV	0.1453	1.0000	3	0.509	**HAR-RV-ONV**	1.7518	1.0000	4
HAR-RV-CC	0.1462	1.0000	4	0.569	**HAR-RV-CC**	1.7309	1.0000	2
HAR-RV-PK	0.1491	0.2910	6	0.539	**HAR-RV-PK**	1.7104	1.0000	3
HAR-RV-GK	0.1490	eliminated	–	0.494	**HAR-RV-GK**	1.7232	0.7582	7
HAR-RV-RS	0.1484	0.2364	7	0.485	**HAR-RV-RS**	1.8314	0.6504	5
HAR-RV-YZ	0.1503	eliminated	–	0.510	**HAR-RV-YZ**	1.7570	0.7122	9
HAR-RV-COMB.	0.1450	1.0000	2	0.509	**HAR-RV-COMB.**	1.7270	0.8146	8

(Continued)

NIKKEI	QLIKE	p-value	Rank	$MZ\text{-}R^2$	NIKKEI	HMSE	p-value	Rank
PK (Model-free)	0.2079	1.0000	2	0.783	**PK (Model-free)**	0.2869	1.0000	3
GK (Model-free)	**0.1806**	**1.0000**	**1**	0.828	**GK (Model-free)**	**0.1914**	**1.0000**	**1**
RS (Model-free)	1.0070	eliminated	–	0.743	**RS (Model-free)**	0.2583	1.0000	2
YZ (Model-free)	0.3027	eliminated	–	0.592	**YZ (Model-free)**	10.440	eliminated	–
HAR-RV	0.2875	eliminated	–	0.260	**HAR-RV**	1.6937	eliminated	–
HAR-RV-ONV	0.2799	1.0000	3	0.256	**HAR-RV-ONV**	1.5365	1.0000	4
HAR-RV-CC	0.2817	0.7368	4	0.282	**HAR-RV-CC**	1.5472	0.8864	5
HAR-RV-PK	0.3242	eliminated	–	0.223	**HAR-RV-PK**	1.8027	eliminated	–
HAR-RV-GK	0.3271	eliminated	–	0.235	**HAR-RV-GK**	1.8173	eliminated	–
HAR-RV-RS	0.3414	eliminated	–	0.240	**HAR-RV-RS**	1.8239	eliminated	–
HAR-RV-YZ	0.2905	eliminated	–	0.257	**HAR-RV-YZ**	1.6741	eliminated	–
HAR-RV-COMB.	0.2862	0.6535	5	0.260	**HAR-RV-COMB.**	1.6240	0.7564	6

SPTSX	QLIKE	p-value	Rank	$MZ\text{-}R^2$	SPTSX	HMSE	p-value	Rank
PK (Model-free)	0.2150	1.0000	2	0.631	**PK (Model-free)**	0.4000	1.0000	3
GK (Model-free)	0.2167	1.0000	3	0.466	**GK (Model-free)**	**0.2766**	**1.0000**	**1**
RS (Model-free)	1.3330	eliminated	–	0.378	**RS (Model-free)**	0.3477	1.0000	2
YZ (Model-free)	**0.1932**	**1.0000**	**1**	0.474	**YZ (Model-free)**	2.1500	eliminated	–
HAR-RV	0.2668	eliminated	–	0.306	**HAR-RV**	1.9910	eliminated	–
HAR-RV-ONV	0.2245	1.0000	4	0.324	**HAR-RV-ONV**	1.9339	eliminated	–
HAR-RV-CC	0.2475	eliminated	–	0.305	**HAR-RV-CC**	1.9170	eliminated	–
HAR-RV-PK	0.2320	0.4454	6	0.327	**HAR-RV-PK**	1.5610	1.0000	4
HAR-RV-GK	0.2473	eliminated	–	0.310	**HAR-RV-GK**	1.6255	eliminated	–
HAR-RV-RS	0.2609	eliminated	–	0.305	**HAR-RV-RS**	1.7507	eliminated	–
HAR-RV-YZ	0.2608	eliminated	–	0.322	**HAR-RV-YZ**	1.9277	eliminated	–
HAR-RV-COMB.	0.2316	0.4856	5	0.306	**HAR-RV-COMB.**	1.6850	eliminated	–

SPX	QLIKE	p-value	Rank	$MZ\text{-}R^2$	SPX	HMSE	p-value	Rank
PK (Model-free)	0.2114	1.0000	2	0.754	**PK (Model-free)**	0.3105	1.0000	3
GK (Model-free)	0.2332	1.0000	3	0.661	**GK (Model-free)**	**0.2131**	**1.0000**	**1**
RS (Model-free)	2.7350	eliminated	–	0.560	**RS (Model-free)**	0.2929	1.0000	2
YZ (Model-free)	**0.2005**	**1.0000**	**1**	0.402	**YZ (Model-free)**	5.3760	eliminated	–
HAR-RV	0.2905	eliminated	–	0.445	**HAR-RV**	2.9500	eliminated	–
HAR-RV-ONV	0.2861	eliminated	–	0.325	**HAR-RV-ONV**	3.0161	eliminated	–
HAR-RV-CC	0.2846	eliminated	–	0.493	**HAR-RV-CC**	3.0495	eliminated	–
HAR-RV-PK	0.2720	1.0000	4	0.483	**HAR-RV-PK**	2.4647	1.0000	4
HAR-RV-GK	0.2831	eliminated	–	0.453	**HAR-RV-GK**	2.7022	eliminated	–
HAR-RV-RS	0.2889	eliminated	–	0.445	**HAR-RV-RS**	2.8742	eliminated	–
HAR-RV-YZ	0.2875	eliminated	–	0.455	**HAR-RV-YZ**	2.9223	eliminated	–
HAR-RV-COMB.	0.2797	1.0000	5	0.445	**HAR-RV-COMB.**	2.7940	eliminated	–

Note: *QLIKE and HMSE values are obtained by one-step-ahead recursive window forecasting method. MZ-R square stands for Mincer-Zarnowitz regression's R squares. . Lower the values of QLIKE and HMSE is better while higher the value of MZ R-square is better for the comparison of forecasts. Window size is 600 observations that is used as in-sample estimation. The out-of-sample consists of 2180 observations. P-value and rank results are received from the MCS procedure. Bold numbers show the best-performing models for each indices.*

Chapter-4:

Table 1: Out-of-sample 1-step-ahead rolling window forecasting and MCS results for SPX (Window size:600)

S&P 500	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-9.1098	0.0108	14	3.3659	eliminated	–	1.1467	eliminated	–
REGIONAL INFORMATION									
HAR-RV-GDAXI	-9.1286	1.0000	10	2.6016	eliminated	–	1.0448	eliminated	–
HAR-RV-FCHI	-9.1165	1.0000	11	2.9471	eliminated	–	1.0907	eliminated	–
HAR-RV-FTMIB	-9.1120	1.0000	8	2.8067	eliminated	–	1.0731	eliminated	–
HAR-RV-STOXX50E	-9.0986	0.0010	18	3.2899	eliminated	–	1.1402	eliminated	–
HAR-RV-FTSE	-8.7559	0.0032	17	3.4309	eliminated	–	1.1562	eliminated	–
REGIONAL KITCHEN-SINK	-9.1055	1.0000	3	2.7375	eliminated	–	1.0572	eliminated	–
REGIONAL COMBINATION	-9.1149	0.0068	15	2.9559	eliminated	–	1.0974	eliminated	–
GLOBAL INFORMATION									
HAR-RV-DJI	-9.0478	1.0000	6	2.5206	eliminated	–	1.0503	eliminated	–
HAR-RV-IXIC	-9.1264	1.0000	7	2.8105	eliminated	–	1.0468	eliminated	–
HAR-RV-VIX	-8.3660	1.0000	4	5.7578	eliminated	–	1.3617	eliminated	–
HAR-RV-WTI	-9.0651	0.0064	16	3.3454	eliminated	–	1.1415	eliminated	–
HAR-RV-GOLD	-9.1383	1.0000	12	2.3928	eliminated	–	0.9930	eliminated	–
GLOBAL KITCHEN-SINK	-6.8097	1.0000	2	3.5249	eliminated	–	1.2069	eliminated	–
GLOBAL COMBINATION	**-9.1476**	**1.0000**	**1**	**1.6023**	**1.0000**	**1**	**0.8503**	**1.0000**	**1**
LOCAL INFORMATION									
HAR-RV-BOND	-8.8423	1.0000	13	3.6082	eliminated	–	1.1590	eliminated	–
HAR-RV-USEPU	-9.1067	0.0000	21	3.3948	eliminated	–	1.1426	eliminated	–
HAR-RV-LIBOR	-9.0623	0.0010	19	3.4924	eliminated	–	1.1554	eliminated	–
LOCAL KITCHEN-SINK	-8.9117	1.0000	5	3.6591	eliminated	–	1.1737	eliminated	–
LOCAL COMBINATION	-9.0698	0.0000	20	3.3372	eliminated	–	1.1347	eliminated	–
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-8.4560	eliminated		4.2103	eliminated	–	1.3211	eliminated	–
OVERALL COMBINATION	-9.0787	1.0000	9	2.2990	eliminated	–	0.9992	eliminated	–

Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

Table 2: Out-of-sample 1-step-ahead rolling window forecasting and MCS results for SPX (Window size:200)

S&P 500	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-8.9673	0.0028	16	2.3016	0.0012	19	0.9531	0.0000	19
REGIONAL INFORMATION									
HAR-RV-GDAXI	-8.7504	1.0000	5	2.1830	1.0000	16	0.9405	1.0000	11
HAR-RV-FCHI	-8.9255	0.0020	17	2.2989	0.0002	22	0.9603	1.0000	17
HAR-RV-FTMIB	-8.9401	1.0000	8	2.3614	1.0000	12	0.9634	1.0000	9
HAR-RV-STOXX50E	-8.9296	0.0000	20	2.2973	0.0006	20	0.9562	0.0000	22
HAR-RV-FTSE	-8.7301	0.0000	19	2.4246	1.0000	13	0.9571	1.0000	16
REGIONAL KITCHEN-SINK	-8.7007	1.0000	3	2.6892	1.0000	11	0.9973	1.0000	13
REGIONAL COMBINATION	-8.9516	0.0000	18	2.2453	1.0000	17	0.9464	1.0000	18
GLOBAL INFORMATION									
HAR-RV-DJI	-8.9076	1.0000	13	2.2566	1.0000	14	0.9610	1.0000	15
HAR-RV-IXIC	-8.9646	1.0000	14	1.9647	1.0000	4	0.9097	1.0000	3
HAR-RV-VIX	-8.7438	1.0000	11	2.0635	1.0000	2	0.8918	1.0000	2
HAR-RV-WTI	-8.9381	0.0000	21	2.3875	0.0006	21	0.9640	0.0004	21
HAR-RV-GOLD	-8.8381	1.0000	4	2.3475	1.0000	15	0.9620	1.0000	12
GLOBAL KITCHEN-SINK	-8.5882	1.0000	7	2.0268	1.0000	3	0.9290	1.0000	4
GLOBAL COMBINATION	**-8.9839**	**1.0000**	**1**	**1.8100**	**1.0000**	**1**	**0.8837**	**1.0000**	**1**
LOCAL INFORMATION									
HAR-RV-BOND	-8.8972	1.0000	6	2.3808	1.0000	18	0.9605	0.0004	20
HAR-RV-USEPU	-8.9439	1.0000	12	2.5542	1.0000	7	0.9614	1.0000	8
HAR-RV-LIBOR	-8.5372	1.0000	2	2.6312	1.0000	6	0.9708	1.0000	10
LOCAL KITCHEN-SINK	-8.8484	1.0000	10	3.0334	1.0000	5	0.9832	1.0000	7
LOCAL COMBINATION	-8.9653	1.0000	15	2.3840	1.0000	8	0.9497	1.0000	6
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-8.6302	eliminated		2.7047	1.0000	10	1.0347	1.0000	14
OVERALL COMBINATION	-8.9784	1.0000	9	2.0417	1.0000	9	0.9141	1.0000	5

Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

Table 3: Out-of-sample 1-step-ahead rolling window forecasting and MCS results for FTSE (Window size:600)

FTSE 100	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-8.6775	0.0000	17	2.8117	eliminated	–	0.9404	eliminated	–
REGIONAL INFORMATION									
HAR-RV-GDAXI	-8.6919	1.0000	10	1.7138	0.1868	12	0.8343	eliminated	–
HAR-RV-FCHI	-8.6928	1.0000	3	1.5850	0.3648	6	0.8001	0.0844	5
HAR-RV-FTMIB	-8.5275	1.0000	4	1.8222	0.2246	9	0.8213	0.0724	6
HAR-RV-STOXX50E	-8.6860	1.0000	8	2.0081	0.2956	8	0.8335	eliminated	–
REGIONAL KITCHEN-SINK	-8.6380	1.0000	2	1.8139	0.2026	11	0.8332	eliminated	–
REGIONAL COMBINATION	-8.5614	1.0000	5	1.7151	0.3220	7	0.8145	eliminated	–
GLOBAL INFORMATION									
HAR-RV-SPX	-8.6962	0.0000	18	2.8851	eliminated	–	0.9116	eliminated	–
HAR-RV-DJI	-8.6819	1.0000	15	2.8425	eliminated	–	0.9317	eliminated	–
HAR-RV-IXIC	-8.6975	1.0000	13	2.8735	0.2206	10	0.8950	eliminated	–
HAR-RV-VIX	-7.8497	eliminated		2.0964	0.7608	3	0.7700	0.0848	4
HAR-RV-WTI	-8.6564	0.0000	20	2.8899	eliminated	–	0.9591	eliminated	–
HAR-RV-GOLD	-8.5442	0.8716	11	1.2141	0.9924	2	0.7164	0.9264	2
GLOBAL KITCHEN-SINK	-8.1505	1.0000	12	**1.0641**	**1.0000**	**1**	**0.6674**	**1.0000**	**1**
GLOBAL COMBINATION	**-8.4869**	**1.0000**	**1**	1.9837	0.7536	4	0.7730	0.2482	3
LOCAL INFORMATION									
HAR-RV-BOND	-8.4640	0.8424	7	2.9217	eliminated	–	0.9794	eliminated	–
HAR-RV-UKEPU	-8.6719	1.0000	14	2.6554	eliminated	–	0.8796	eliminated	–
HAR-RV-LIBOR	-8.6778	1.0000	16	2.8158	eliminated	–	0.9339	eliminated	–
LOCAL KITCHEN-SINK	-8.1715	0.2064	9	2.7366	eliminated	–	0.9281	eliminated	–
LOCAL COMBINATION	-8.6801	0.0000	19	2.6921	eliminated	–	0.9092	eliminated	–
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-7.9749	eliminated		1.5325	0.0970	13	0.7759	0.0000	8
OVERALL COMBINATION	-8.6952	1.0000	6	1.9766	0.5860	5	0.8061	0.0570	7

Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

Table 4: Out-of-sample 1-step-ahead rolling window forecasting and MCS results for FTSE (Window size:200)

FTSE 100	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-8.5465	0.0188	18	1.9557	0.2086	17	0.8077	eliminated	–
REGIONAL INFORMATION									
HAR-RV-GDAXI	-8.5202	1.0000	8	1.3682	1.0000	5	0.7196	0.8966	7
HAR-RV-FCHI	-8.5115	1.0000	3	1.3703	1.0000	3	0.7039	1.0000	2
HAR-RV-FTMIB	-8.5197	1.0000	5	1.6749	1.0000	11	0.7468	0.7458	8
HAR-RV-STOXX50E	-8.5618	1.0000	7	1.4448	1.0000	6	0.7418	0.2474	12
REGIONAL KITCHEN-SINK	-8.4860	1.0000	17	1.5121	1.0000	7	0.7518	0.4574	10
REGIONAL COMBINATION	-8.5549	1.0000	4	1.3393	1.0000	2	0.7091	0.9960	6
GLOBAL INFORMATION									
HAR-RV-SPX	-8.5354	0.0000	14	2.1214	1.0000	16	0.7810	0.0604	13
HAR-RV-DJI	-8.5254	1.0000	11	1.8870	1.0000	15	0.7857	eliminated	–
HAR-RV-IXIC	-8.5695	1.0000	6	2.3493	1.0000	13	0.7765	0.2672	11
HAR-RV-VIX	-7.8476	1.0000	9	1.4252	1.0000	4	0.7116	1.0000	3
HAR-RV-WTI	-8.5117	0.0000	13	2.4302	1.0000	12	0.8261	eliminated	–
HAR-RV-GOLD	-8.4973	1.0000	12	1.6112	1.0000	9	0.7503	0.6112	9
GLOBAL KITCHEN-SINK	-8.0804	1.0000	10	**1.2804**	**1.0000**	**1**	**0.6942**	**1.0000**	**1**
GLOBAL COMBINATION	**-8.6021**	**1.0000**	**1**	1.5702	1.0000	10	0.7132	1.0000	4
LOCAL INFORMATION									
HAR-RV-BOND	-8.5293	0.0102	21	2.0225	0.1662	19	0.8217	eliminated	–
HAR-RV-UKEPU	-8.5287	0.0176	19	1.9402	0.2010	18	0.8046	eliminated	–
HAR-RV-LIBOR	-8.5451	0.0170	20	1.9420	0.1612	20	0.8087	eliminated	–
LOCAL KITCHEN-SINK	-8.4609	1.0000	15	2.0194	eliminated		0.8283	eliminated	–
LOCAL COMBINATION	-8.4687	1.0000	16	1.9090	1.0000	14	0.7984	eliminated	–
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-8.4038	eliminated		1.7934	0.0888	21	0.7904	0.0142	14
OVERALL COMBINATION	-8.6037	1.0000	2	1.4681	1.0000	8	0.7136	1.0000	5

Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

Table 5:Out-of-sample 1-step-ahead rolling window forecasting and MCS results for GDAXI (Window size:600)

GDAXI	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-8.5680	0.0028	17	0.9281	0.0014	20	0.6322	0.0026	19
REGIONAL INFORMATION									
HAR-RV-FCHI	-8.5694	1.0000	16	0.9042	1.0000	13	0.6241	1.0000	14
HAR-RV-FTMIB	-8.5561	1.0000	7	0.9050	1.0000	14	0.6228	1.0000	11
HAR-RV-STOXX50E	-8.5699	1.0000	15	0.9008	1.0000	9	0.6200	1.0000	8
HAR-RV-FTSE	-8.5679	0.0002	20	0.9525	1.0000	19	0.6351	0.0004	22
REGIONAL KITCHEN-SINK	-8.5643	0.0012	19	0.9514	1.0000	18	0.6331	1.0000	17
REGIONAL COMBINATION	-8.5706	0.0020	18	0.9070	1.0000	16	0.6233	1.0000	15
GLOBAL INFORMATION									
HAR-RV-SPX	-8.5673	1.0000	14	1.0500	1.0000	5	0.6381	1.0000	9
HAR-RV-DJI	-8.5645	1.0000	10	1.0515	1.0000	11	0.6419	1.0000	16
HAR-RV-IXIC	-8.5682	1.0000	8	0.9589	1.0000	10	0.6363	1.0000	7
HAR-RV-VIX	-7.6456	1.0000	3	0.8440	1.0000	4	0.6220	1.0000	6
HAR-RV-WTI	-8.5385	1.0000	9	0.9480	0.0012	21	0.6410	0.0008	20
HAR-RV-GOLD	-8.5516	1.0000	11	0.9524	1.0000	15	0.6423	1.0000	12
GLOBAL KITCHEN-SINK	-8.4427	1.0000	13	1.0617	1.0000	12	0.6592	1.0000	10
GLOBAL COMBINATION	-8.5755	1.0000	6	0.8870	1.0000	7	0.6137	1.0000	4
LOCAL INFORMATION									
HAR-RV-BOND	-8.5667	0.0002	21	0.9311	1.0000	17	0.6344	0.0004	21
HAR-RV-VIXGDAXI	-8.0537	1.0000	5	**0.7164**	**1.0000**	**1**	0.5855	1.0000	2
HAR-RV-LIBOR	-8.5636	1.0000	12	0.9366	0.0010	22	0.6338	1.0000	18
LOCAL KITCHEN-SINK	-3.9831	eliminated		0.7355	1.0000	3	0.5941	1.0000	3
LOCAL COMBINATION	**-8.5839**	**1.0000**	**1**	0.7260	1.0000	2	**0.5733**	**1.0000**	**1**
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-8.3704	1.0000	4	0.9433	1.0000	8	0.6464	1.0000	13
OVERALL COMBINATION	-8.5797	1.0000	2	0.8384	1.0000	6	0.6041	1.0000	5

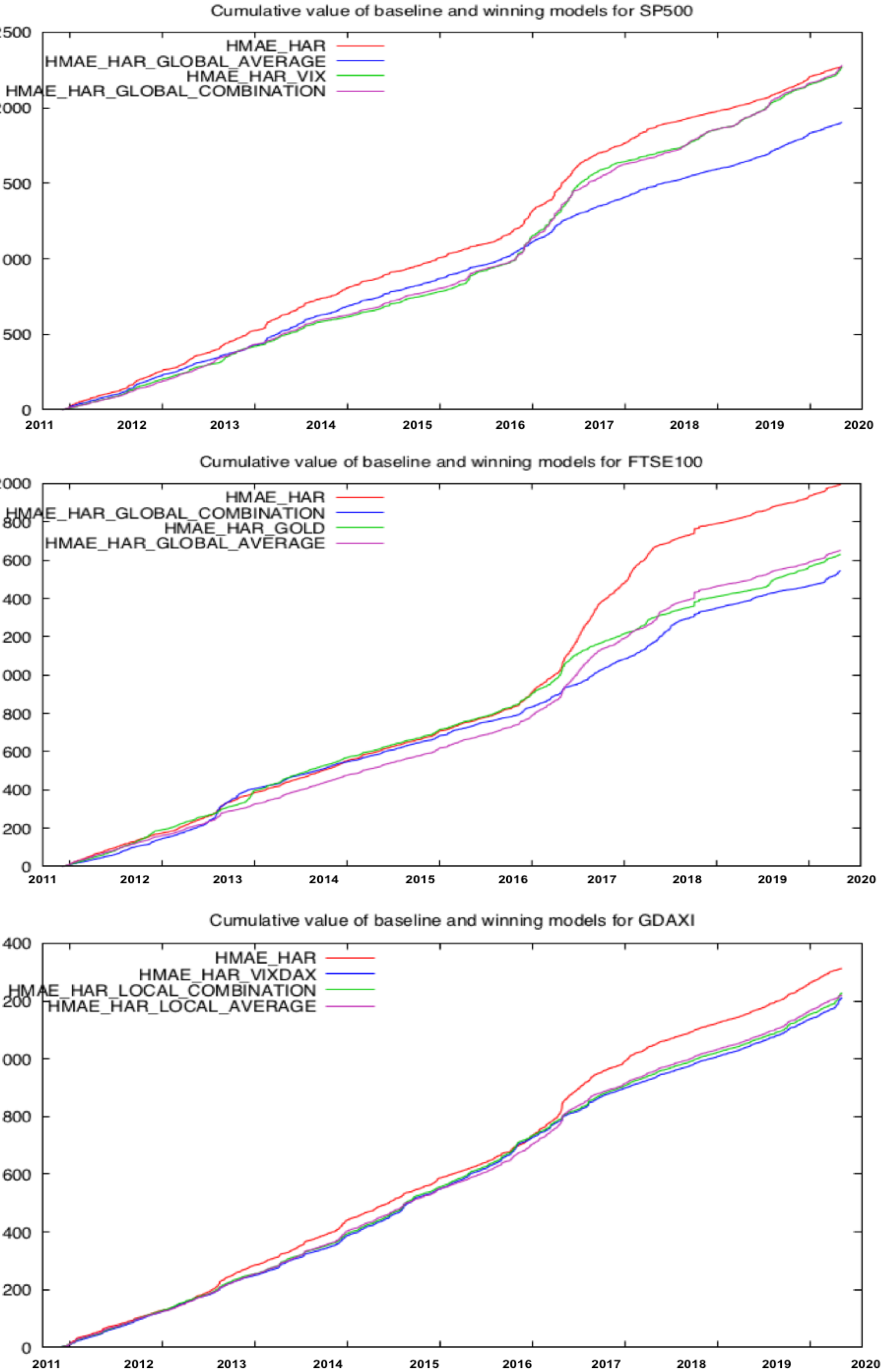
Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

Table 6:Out-of-sample 1-step-ahead rolling window forecasting and MCS results for GDAXI (Window size:200)

GDAXI	QLIKE	p-value	Rank	HMSE	p-value	Rank	HMAE	p-value	Rank
HAR-RV (BASELINE MODEL)	-8.4713	0.0008	17	0.7225	0.0510	15	0.5741	0.0208	15
REGIONAL INFORMATION									
HAR-RV-FCHI	-8.4544	0.0006	19	0.7298	0.0142	18	0.5752	0.0042	17
HAR-RV-FTMIB	-8.4628	0.0006	18	0.7591	1.0000	14	0.5860	1.0000	14
HAR-RV-STOXX50E	-8.4679	1.0000	15	0.7237	1.0000	13	0.5755	0.0196	16
HAR-RV-FTSE	-8.4682	1.0000	16	0.7579	0.0038	21	0.5805	0.0002	21
REGIONAL KITCHEN-SINK	-8.4505	1.0000	13	0.8422	eliminated		0.6068	eliminated	
REGIONAL COMBINATION	-8.4707	0.0004	20	0.7272	0.0194	17	0.5757	0.0038	19
GLOBAL INFORMATION									
HAR-RV-SPX	-8.4686	1.0000	12	0.7643	1.0000	11	0.5767	1.0000	11
HAR-RV-DJI	-8.4699	1.0000	14	0.7866	1.0000	9	0.5785	1.0000	12
HAR-RV-IXIC	-8.4719	1.0000	10	0.7273	1.0000	8	0.5711	1.0000	7
HAR-RV-VIX	-8.4558	1.0000	6	0.6693	1.0000	4	0.5635	1.0000	3
HAR-RV-WTI	-8.4439	1.0000	4	0.7454	0.0126	19	0.5816	0.0014	20
HAR-RV-GOLD	-8.4583	1.0000	8	0.7875	1.0000	10	0.5861	1.0000	10
GLOBAL KITCHEN-SINK	-8.4016	1.0000	7	0.7954	1.0000	12	0.5965	1.0000	9
GLOBAL COMBINATION	-8.4808	1.0000	2	0.6851	1.0000	6	0.5596	1.0000	4
LOCAL INFORMATION									
HAR-RV-BOND	-8.4109	1.0000	11	0.7181	0.0094	20	0.5786	0.0038	18
HAR-RV-VIXGDAXI	-8.4509	1.0000	5	0.6386	1.0000	2	**0.5462**	**1.0000**	**1**
HAR-RV-LIBOR	-8.4710	0.0004	21	0.7430	0.0266	16	0.5753	1.0000	13
LOCAL KITCHEN-SINK	-8.4427	1.0000	9	0.6533	1.0000	3	0.5521	1.0000	6
LOCAL COMBINATION	**-8.4703**	**1.0000**	**1**	**0.6295**	**1.0000**	**1**	0.5488	1.0000	2
OVERALL INFORMATION									
OVERALL KITCHEN-SINK	-8.3969	eliminated		0.7506	1.0000	7	0.5867	1.0000	8
OVERALL COMBINATION	-8.4729	1.0000	3	0.6710	1.0000	5	0.5582	1.0000	5

Note: Bold row in the table is the winner model with the smallest loss functions, unit p-values, and highest MCS ranks.

Figure 1


Cumulative value of baseline and winning models for SP500
HMAE_HAR
HMAE_HAR_GLOBAL_AVERAGE
HMAE_HAR_VIX
HMAE_HAR_GLOBAL_COMBINATION
0
500
1000
1500
2000
2500
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Cumulative value of baseline and winning models for FTSE100
HMAE_HAR
HMAE_HAR_GLOBAL_COMBINATION
HMAE_HAR_GOLD
HMAE_HAR_GLOBAL_AVERAGE
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Cumulative value of baseline and winning models for GDAXI
HMAE_HAR
HMAE_HAR_VIXDAX
HMAE_HAR_LOCAL_COMBINATION
HMAE_HAR_LOCAL_AVERAGE
0
200
400
600
800
1000
1200
1400
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020

Chapter-5:

Table 4: Network statistics for G20 markets without global barometers (19 nodes)

	Number of edges	Average degree	Total degree	Weighted degree	Network diameter	Graph density	Average path length	Clustering coefficient
Period 1	135	7.105	22.257	1.171	2	0.579	1.421	0.636
Period 2	196	10.316	61.006	3.211	2	0.673	1.327	0.718
Period 3	115	6.053	32.317	1.701	3	0.439	1.596	0.592
Period 4	136	7.158	22.726	1.196	2	0.608	1.392	0.651
Period 5	190	10	83.732	4.407	2	0.708	1.292	0.731

Table 5: Top network edges for G20 markets without global barometers (19 nodes)

Only g20								
Period 1			**Period 2**			**Period 3**		
Turkey –>	US	0.71	India –>	Canada	1.025	Brasil –>	UK	1.553
Germany –>	Italy	0.66	France –>	US	1.000	UK –>	US	0.935
France –>	UK	0.64	Germany –>	US	0.876	US –>	UK	0.909
Indonesia –>	Canada	0.51	Argentina –>	Australia	0.695	S. Africa –>	UK	0.861
France –>	Germany	0.50	Mexico –>	US	0.679	Japan –>	S. Africa	0.808
Period 4			**Period 5**					
UK –>	US	0.706	Argentina –>	Korea	1.448			
Saudi –>	Australia	0.650	Italy –>	S. Africa	1.379			
US –>	UK	0.635	Germany –>	Brasil	1.310			
Argentina –>	Canada	0.546	Germany –>	France	1.282			
Argentina –>	Italy	0.539	Brasil –>	Russia	1.213			

Table 6: Network statistics for G20 markets and global barometers with VIX (23 nodes)

	Number of edges	Average degree	Total degree	Weighted degree	Network diameter	Graph density	Average path length	Clustering coefficient
Period 1	199	8.652	37.537	1.632	2	0.545	1.455	0.631
Period 2	282	12.261	82.457	3.585	2	0.676	1.324	0.694
Period 3	153	6.652	38.365	1.668	3	0.419	1.601	0.542
Period 4	192	8.348	33.741	1.467	2	0.589	1.411	0.622
Period 5	332	14.435	223.88	9.734	2	0.787	1.213	0.814

Table 7: Top network edges for G20 markets and global barometers with VIX (23 nodes)

Only g20								
Period 1			**Period 2**			**Period 3**		
Russia –>	Italy	1.53	India –>	Canada	1.025	Brasil –>	UK	1.553
Italy –>	Argentina	1.36	France –>	US	1.000	UK –>	US	0.935
Italy –>	WTI	1.10	WTI –>	Saudi	0.964	US –>	UK	0.909
Italy –>	Mexico	1.02	India –>	VIX	0.922	S. Africa –>	UK	0.861
Brasil –>	Italy	0.92	France –>	VIX	0.914	Japan –>	S. Africa	0.808
Period 4			**Period 5**					
UK –>	US	0.706	WTI –>	Saudi	11.3			
Saudi –>	Australia	0.650	WTI –>	Indonesia	5.63			
VIX –>	UK	0.647	WTI –>	Mexico	4.37			
US –>	UK	0.635	WTI –>	S. Africa	4.12			
Gold –>	Germany	0.623	Bond –>	Canada	4.02			