Essays on the Informational Content of Implied Correlation and Observation-Driven Models

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy, July 2023.

Xiaohang Sun

19900489

Supervised by

Dr. Nikolaos Voukelatos
Dr. Jaideep S Oberoi

Signature _____

Date _____ / _____ / _____

Abstract

This thesis examines the application and comparison of a set of widely adopted parametric and non-parametric approaches in forecasting the realisation of risk measures, such as volatility. In chapter 2, implied correlation is extracted from options and utilised to predict realised correlation after decomposition and re-grouping. I also attempt to forecast market return using the re-organized signal. Starting from chapter 3, I focus on parametric models, more specifically, the observation-driven models represented by GARCH and score-driven models (GAS). The observation-driven models are extensively constructed upon assumptions on innovation terms. Their predictive power concerning realised volatility is evaluated and compared. In addition, the performance of GAS and GARCH models is also compared to implied volatility comprehensively. In chapter 4, I propose to construct GAS models with shifted Gamma (SG-GAS) and shifted negative Gamma (SNG-GAS) innovations for option pricing. Corresponding GARCH models and the Black-Scholes model are built as as benchmarks for comparison. It appears that both GAS models outperform the other models. The superiority of SG-GAS and SNG-GAS is mainly driven by their ability to price out-of-the-money options accurately.

Acknowledgements

I would like to take this opportunity to express my sincere gratitude to the following people. I would not be able to accomplish my research without the support from them.

I would like to thank my supervisors, Dr.Nikolaos Voukelatos and Dr.Jaideep Oberoi. During the past four years, Dr.Voukelatos was always able to provide timely and effective help whenever I was confronted with difficulties in my research. I enjoyed the time spent discussing academic questions with him together. Dr.Oberoi, with his insight, broadened the scope of my research. Apart from this, I received a lot of career guidance from Dr.Oberoi. This makes me feel more confident in job searching.

I would also like to thank the KBS PGR team. I still remember the first day I joined KBS. Since then, I continuously received support from our lovely KBS colleagues on both work and life.

Big thanks to my external examiner, Professor Ekaterini Panopoulou, and my internal examiner, Dr.Iraklis Apergis. Following their constructive suggestions during my viva, the quality of my thesis has been significantly improved.

The past pandemic was depressing. Perhaps nothing is more encouraging than having a newborn during those years. I would like to thank my baby son, who was born in Kent in 2021, for giving me plenty of energy and courage during the past years. Sometimes, when I feel exhausted or depressed, I can instantly get recharged and feel full of power as soon as I watch him. Thank you, little Sun. Thanks to my wife Janet, who sacrificed her career to support my PhD study and to take good care of our baby. Finally, thanks to my parents, who generously offered me support both emotionally and financially.

Contents

List of Tables

List of Figures

Abbreviations

AAA AAA-Rated Corporate Bonds Yields.

ARCH Autoregressive Conditional Heteroskedasticity.

ATM At-the-Money.

BAA BAA-Rated Corporate Bonds Yields.

BM Book-to-Market Ratio.

BS Black-Scholes.

CBR Corporate Bond Returns.

CPR Correlation Risk Premium.

CW Clark-West.

DCS Dynamic Conditional Score.

DJIA Dow Jones Industrial Average.

DM Diebold–Mariano.

dOTM Deep Out-of-the-Money.

EGP Extended Girsanov Principle.

EMD Empirical Mode Decomposition.

EMM Equivalent Martingale Measure.

ETH Efficient Market Hypothesis.

GARCH Generalised Autoregressive Conditional Heteroskedasticity.

GAS Generalised Autoregressive Score.

GED Generalised Error Distribution.

GIM GARCH-in-Mean.

IC Implied Correlation.

ICH High-Frequency Component.

ICL Low-Frequency Component.

IMF Intrinsic Mode Function.

INFL Inflation.

IV Implied Volatility.

LOG Logistic Distribution.

LRNVR Locally Risk-Neutral Valuation Relationship.

LTR Long-Term Rate of Return.

LTY Long-Term Yield.

MAE Mean Absolute Errors.

MAPE Mean Absolute Percentage Error.

MFIV Model-Free Implied Volatility.

MGF Moment Generating Function.

MLE Maximum Likelihood Estimation.

MSFE Mean Square Forecast Error.

NEE Net Equity Expansion.

OTM Out-of-the-Money.

QLIKE Quasi-Likelihood.

RC Realised Correlation.

RV Realised Volatility.

SG Shifted Gamma Distribution.

SIG Shifted Inverse Gaussian Distribution.

SNG Shifted Negative Gamma Distribution.

SNIG Shifted Negative Inverse Gaussian Distribution.

SVAR S&P 500 Index Variance.

SW Shifted Weibull Distribution.

TB T-Bill Rates.

TLS Student's t Location and Scale Distribution.

VPR Variance Risk Premium.

Chapter 1

Introduction

This thesis is composed of three essays on forecasting and option pricing. In the first two essays, emphasis is placed on parametric and non-parametric approaches for predicting realised correlation, market returns and volatility. In the last essay, a parametric approach is applied for option pricing.

Correlation plays a crucial role in finance as an essential measure of dispersion between assets within a basket. In chapter 2, I investigate the predictive power of the average expected correlation between component stock returns that can be extracted from corresponding options plus those written on the index. The predictive power of implied correlation against realised correlation and market returns over a variety of horizons is documented in Driessen et al. (2012), Faria et al. (2022), Hollstein et al. (2018) and Bernales and Valenzuela (2016).

A series of studies have been conducted to inspect the specific channels through which implied correlation works to predict future market returns. For example, Buraschi et al. (2014) propose that correlation risk is priced because it relates to investors' disagreement about the future state of the market. Buss et al. (2017) document that implied correlation is connected with market returns via its link to

non-diversifiable market risk. Buss et al. (2019) suggests that the forecasting power of implied correlation stems from its ability to act as a procyclical state variable that can predict macroeconomic variables and financial risks. It is also reported by Buss et al. (2019) that the forecasting ability of implied correlation is driven by the interplay between its three main determinants, implied market variance, implied idiosyncratic variance and the dispersion of implied market betas. More recently, Bernales and Valenzuela (2016) found that changes in implied correlation tend to cause changes in market volatility and consequently result in changes in the equity premium.

The motivation of this chapter is to discover an alternative channel via which implied correlation forecasts future market returns. More specifically, the innovative part of this chapter is that, in an adaptive sense, the time series of implied correlation is decomposed and re-grouped to form low-frequency and high-frequency components. The relation between the informational content of different components and future market returns at different horizons is investigated. My research is closely connected with Buss et al. (2019) and Branger et al. (2021), who also study the determinants of implied correlation. However, while determinants in previous studies are found to be observable market variables, such as implied market variance, this chapter is the first study to focus on extracted components that are representative of different frequencies. Naturally, these components can be recognised as proxies of expectations from different groups of investors with distinct investment horizons.

The technology I adopt for the decomposition of implied correlation is called Empirical Mode Decomposition (EMD) (Huang et al., 1998). It is not dependent on assumptions on the structure of implied correlation, and it can decompose the original implied correlation time series into components with significantly different properties. It is natural to think that the high-frequency component captures short-term fluctuations while the low-frequency component works better in capturing long-term trends.

This chapter makes several contributions to existing literature. First, I document that the previously reported predictive power of implied correlation over future realisations is predominantly driven by its low-frequency component, which is indicated by consistently higher R-squares in predictive regressions than using either the original time series of implied correlation or its high-frequency component.

Second, I find that the forecasting ability of implied correlation concerning market returns is driven by interactions between its components of different frequencies. overlapping predictive regressions indicate that The predictive power of implied correlation to short-term market returns stems mainly from the informational content of the high-frequency component. On the contrary, as longer forecasting horizons are considered, the low-frequency component dominates. More importantly, it can be observed that the combination of both components is a more efficient predictor of market returns consistently across horizons. This indicates that the decomposition of the informational content of implied correlation is beneficial. These findings are valid after controlling for fundamental variables that predict market returns, as in Goyal and Welch (2008); Fereira and Santa-Clara (2011).

Third, I document that decomposing implied correlation can improve the *out-of-sample* predictability of market returns. More specifically, I find that components of implied correlation outperform implied correlation itself at corresponding forecasting horizons. However, while components are found to outperform in certain scenarios, the benchmark of the historical average is relatively challenging to beat.

In chapter 3, I shift to the parametric methods for volatility modelling and forecasting. I adopt two classes of trending observation-driven models to capture the characteristics of the market returns. The first model is Generalised Autoregressive Conditional Heteroskedasticity (GARCH), which enables the variance of the error term to be time-varying. The second model I apply is the Dynamic Conditional Score (DCS) model, or the score-driven model. As the DCS model is quite compa-

rable to the GARCH model, another name for the DCS model is the Generalised Autoregressive Score (GAS) model. In the rest part of this thesis, the score-driven model is referred to as the GAS model.

This chapter is motivated by the fact that while most of the existing literature in volatility modelling proposes to model financial returns within the framework of GARCH models, GAS models, with appealing features, are less frequently discussed in the application of volatility modelling. Instead of GARCH, it is proposed in this chapter to predict realised volatility within the framework of GAS models. The GAS models are empirically found to have a less extreme response to abnormal shocks in observations and hence are expected to produce fewer extreme volatility forecasts compared to GARCH models. This property is valuable, as the observed large absolute returns might be caused by the fat-tailed nature of the financial returns distribution.

In this study, I make various assumptions on the distribution of the innovation term that brings randomness into the index returns process. More specifically, GARCH models with normal (GARCH), shifted Gamma (SG-GARCH and SNG-GARCH), shifted inverse Gaussian (SIG-GARCH and SNIG-GARCH), shifted Weibull (SNW-GARCH), t-location and scale (TLS-GARCH) and Logistic (LOG-GARCH) innovations are constructed using market returns time series, represented by the Dow Jones Industrial Average (DJIA) index. Following identical assumptions on the distribution of innovation terms, corresponding GAS models are constructed as comparisons to GARCH models. This chapter contributes to the existing literature by implementing comprehensive comparisons between the performance of GARCH and GAS models in volatility forecasting.

Through the comparisons, several findings are documented. First, generally, GARCH models tend to perform better in volatility forecasting than GAS models. Such superiority is reflected by both higher descriptive R^2 from regressing realised volatility

onto model-based predictions and lower prediction errors.

Second, it is documented that while GARCH dominates GAS in most cases, the TLS-GAS model, among all candidate models, tends to have the lowest prediction error at the daily forecasting horizon. Considering that TLS-GAS has a slope closer to 1 and that its R^2 is on the same level as the other top performers, it can be asserted that TLS-GAS models have the overall best performance in forecasting realised volatility at the daily horizon.

Third, I find that at daily and weekly forecasting horizons, GARCH models outperform corresponding GAS models in explaining the realised volatility in most scenarios. The pattern stays unchanged for forecasting accuracy. At the monthly horizon, the superior performance of GARCH against GAS seems to be weakened.

Fourth, using a sub-sample covering the 2007-2008 financial crisis, I find that while the GARCH models still have higher descriptive R^2 from Mincer-Zarnowitz regressions in most cases, GARCH and GAS tend to have comparable performance at forecasting horizons longer than one day. It is also documented that SNW-GAS has the overall best performance in predicting realised volatility at daily horizons during the market turmoil.

Fifth, by comparing the predictive power of model-based volatility predictions to that of option-implied volatility, I find that implied volatility consistently outperforms model-based volatility predictions in describing movements of realised volatility. This is consistent with Becker et al. (2007) and Becker and Clements (2008). However, regarding forecasting accuracy, both Implied Volatility (IV) and Model-Free Implied Volatility (MFIV) have inferior performance than model-based volatility predictions.

The motivation of chapter 4 is similar to that of chapter 3. In terms of pricing options via Monte Carlo simulation approach, the framework of GARCH is a natural

candidate. In comparison, GAS models draw relatively less attentions in this field.

Instead of having closed-form solutions, Duan (1995) and Duan (1999) propose to model the underlying price process with the GARCH model. The GARCH process is risk-neutralised under the framework of the locally risk-neutral valuation relationship (LRNVR) before being used to simulate the dynamics of underlying prices, which is used to compute the expected payoffs. Instead of LNRVR, the Esscher transform is adopted in Christoffersen et al. (2006) for risk-neutralisation due to its capability to handle non-Gaussian innovations.

My study is connected to this strand of literature. The innovative part of this chapter is that, instead of building GARCH models, GAS models with shifted Gamma innovation (SG-GAS and SNG-GAS) are constructed for option valuation via a simulation approach. The Gamma innovation term is chosen because of its capability to account for skewness exhibited in financial returns (Tong et al., 2004). The Esscher transform is adopted for risk-neutralisation.

One attractive feature of the GAS model is that the time-varying parameter that drives the updating equation is not necessarily the variance of the innovation term. Instead, it is usually connected with variance, for example, the scale parameter. Most importantly, I find that if appropriate innovation terms are selected, the GAS model tends to be more accurate when the market is in turmoil. This further supports the empirical findings in the previous chapter.

The primary finding of this study is that, in terms of pricing accuracy measured by the Mean Absolute Percentage Error (MAPE), the GAS models with shifted Gamma innovation (SG-GAS) and shifted negative Gamma innovation (SNG-GAS) significantly dominate other competing models, including the BS model, GARCH model with normal innovation (GARCH) and GAS model with normal innovation (GAS).

Using sub-samples before, during and after the 2007-2008 financial crisis, it is also documented that the superior performance of SG-GAS and SNG-GAS models is even more significant in a normal market environment. However, during the financial crisis,the GAS models with Gamma innovations, while still outperforming the BS model, no longer have the best performance during the unprecedented market turmoil. Instead, models with normal innovation (GARCH and GAS) tend to have better performance during this period of massive market uncertainty. Furthermore, the GAS model significantly dominates its GARCH counterpart with a lower prediction error.

To find out the source of the outperformance of proposed GAS models, the option valuation performance of all models is evaluated across a range of moneyness and a variety of time to maturity. It can be observed that the superior pricing performance of SG-GAS and SNG-GAS is mainly due to their ability to more accurately price Deep Out-of-the-Money (dOTM) options, while these two models have the worst performance on pricing near-the-money options. It is also documented that SG-GAS and SNG-GAS models perform better in pricing options with longer maturities.

I attempt to find out the source of improved performance of SG-GAS and SNG-GAS in pricing Out-of-the-Money (OTM) options by designing a series of Monte Carlo simulation studies. It is documented that, under $\mathbb{Q}$, the density implied by the SNG-GAS model has fatter tails compared to that implied by the GARCH model with normal innovations. Such a pattern can be consistently observed for density simulated 30 days to 360 days from the present. Such behaviour of the SNG-GAS model indicates that it tends to have higher predictions over tail risk, which is related to the price of OTM options. Consequently, SG-GAS and SNG-GAS models are capable of better estimating OTM options prices.

Finally, in chapter 5, I discuss and conclude the findings of all three essays and limitations, based on which potential future research directions are proposed.

Chapter 2

The informational content of implied correlation

2.1 Introduction

Option prices reflect the expectations of market participants about future underlying price movements. According to Bates (1991), option prices contain information that is not fully captured by historical prices. This is consistent with the point proposed by Buss and Vilkov (2012) that option prices respond more rapidly to market shocks than the underlying stocks. It is natural to consider that information extracted from option prices can be utilised to predict future market returns. Such measures include (but are not limited to) implied variance and other higher moments, which can be estimated from equity options or index options. Apart from these, the average pairwise correlation between assets in the market can be extracted using both individual stock options and index options. This chapter focuses on the predictive power of option-implied correlation.

The correlation between assets has been vital in finance for a long time. Changes in

correlation are negatively connected to the consumption and investment opportunity set, as a higher correlation means lower diversification benefits and higher market volatility. Therefore, it is essential to understand how to measure, model and analyse the correlation structure between assets to improve asset pricing, portfolio allocation, and risk management. In this chapter, I inspect the informational content of option-implied correlation at different frequency levels. Emphasis is placed on understanding whether decomposing the information contained in implied correlation can improve its predictive power over market returns.

The recent literature has been paying increasing attention to the predictive power of implied correlation, i.e. the average expected correlation between asset returns that can be inferred from index options and options written on the constituent stocks. More specifically, Driessen et al. (2012), Faria et al. (2022), Hollstein et al. (2018) and Bernales and Valenzuela (2016) show that implied correlation can efficiently predict market returns for horizons of up to one year. Intuitively, since option-implied information is forward-looking by construction, implied correlation reflects market participants' expectations of realised correlations in the future. As a result, implied correlation tends to outperform other predictors in forecasting realised correlation. In addition to its role in predicting its realisations and, thus, market returns, the implied correlation has been found to contain significant information in various other contexts. For example, Driessen et al. (2009) find that correlation risk is priced in the cross-section of index and individual option returns. Buss and Vilkov (2012) use implied correlation to construct option-implied estimates of market betas, which are more accurate predictors of realised betas than historical estimates. Kempf et al. (2015) use implied correlation to obtain more efficient estimators of the covariance matrix, leading to significant improvements in portfolio optimisation. Kelly et al. (2016) extract implied correlation from options on sector indices and individual stocks to measure the implicit government guarantee for the financial sector during the 2007-2009 crisis.

Several alternative frameworks have been proposed to explain the specific channels via which implied correlation is related to future market returns. For instance, Buraschi et al. (2014) argue that correlation risk is priced because it relates to investors' disagreement about the future state of the market. Buss et al. (2017) propose that implied correlation predicts market returns via its link to non-diversifiable market risk, as opposed to diversification risk. Buss et al. (2019) suggests that the forecasting power of implied correlation stems from its ability to act as a pro-cyclical state variable that can predict macroeconomic variables and financial risks. In this context, Buss et al. (2019) show that the forecasting power of implied correlation is driven by the interplay between its three main determinants, namely implied market variance, implied idiosyncratic variance and the dispersion of implied market betas. Furthermore, Bernales and Valenzuela (2016) suggest that implied correlation changes induce market volatility, ultimately leading to changes in the equity premium. Branger et al. (2021) argues that the correlation risk premium contains a premium associated with continuous stock price movements and a separate premium associated with co-jumps, with the former predicting market returns at longer horizons and the latter at shorter horizons.

I contribute to this strand of the literature by exploring an alternative channel through which implied correlation could predict future market returns. In particular, I decompose the time series of implied correlation into a low-frequency and a high-frequency component, and I examine how the specific informational content of each component relates to future market returns at different horizons. My research is more closely associated with Buss et al. (2019) and Branger et al. (2021), who also study the determinants of implied correlation. However, while these previous studies examine a set of determinants that constitute observable market variables, such as implied market variance, this chapter is the first to focus on extracting components that refer to different frequencies. These components are representatives of different groups of investors and are reflections of various expectations of investors from these

groups.

One advantage of the frequency decomposition is that it avoids imposing strong assumptions about the correlation structure and its relationship to other market variables. More importantly, though, the resulting components naturally inherit distinct properties from the original time series, with the high-frequency component capturing short-term fluctuations and the low-frequency component capturing longer-term trends. Therefore, it would be reasonable to expect that a frequency decomposition could explain why implied correlation's predictive power is commonly found to vary considerably across different forecasting horizons. The most robust finding of my study is that the high-frequency component of implied correlation is the most efficient predictor of market returns for shorter horizons. In comparison, the low-frequency component performs better for longer horizons.

The empirical analysis is based on the expected average correlation implied by options written on the S&P500 index and its constituent stocks, with the sample period ranging from 1996 to 2020. Similarly to Driessen et al. (2009), I observe a substantial correlation premium, which is generally larger for longer maturities. To disaggregate the informational content of implied correlation, I apply the Empirical Mode Decomposition (EMD) (Huang et al., 1998) to extract several oscillations with different frequencies from the original time series, which I then use to construct a low-frequency and a high-frequency component.

This study makes several empirical contributions to the literature. First, I find that the previously documented predictive power of implied correlation over future realised correlation is driven predominantly by the former's low-frequency component, which results in consistently higher R-squared in predictive regressions than using either the original time series of implied correlation or its high-frequency component. The superior performance of the low-frequency component is observed across forecasting horizons from one to six months, and it is the strongest for the 3-month

horizon (where the R-squared increases from 22% to 34% when I replace implied correlation with its low-frequency component).

Second, I show that the predictive power of implied correlation over future market returns is driven by the interaction between its two components. To this end, I estimate predictive regressions of market returns against several combinations of implied correlation and its frequency-based components. I estimate these predictive regressions in an overlapping fashion as in Driessen et al. (2012) across several forecasting horizons ranging from one day to one year while controlling for the correlation risk premium, the index's model-free implied variance, and the Bollerslev et al. (2009) variance risk premium. The predictability of market returns at short horizons stems primarily from the informational content of the high-frequency component, which consistently outperforms the original time series of implied correlation and the low-frequency component as a predictor. This superior forecasting performance of the high-frequency component is observed for forecasting horizons of one day up to one month. For instance, at the 1-month horizon, replacing implied correlation with its high-frequency component in the predictive regression more than doubles the R-squared from 4.6% to 10.2%.

As I move towards longer forecasting horizons of up to twelve months, the low-frequency component starts dominating as the primary source of predictive power. In this sense, the empirical findings support the intuitive notion that the high-frequency component contains significant information about short-term fluctuations, which can improve the forecastability of market returns at short horizons. In contrast, the low-frequency component reflects longer-term trends that relate to future returns over longer horizons. Importantly, combining both components consistently results in a more efficient forecast of market returns than that provided by the original series, highlighting the merits of disaggregating the informational content of implied correlation. Although consistent across all forecasting horizons, this improvement in forecasting power is more pronounced for shorter horizons of up to

three months. This finding is particularly important since market returns are typically more difficult to forecast in the short term than in the longer term.

I also examine non-overlapping forecasts where I include several fundamental variables shown to efficiently predict market returns (Goyal and Welch, 2008; Fereira and Santa-Clara, 2011). The improvement in R-squared from decomposing implied correlation is more marginal after accounting for these commonly used predictors of the equity premium. Nevertheless, the high-frequency component remains a robust predictor at shorter horizons, where it outperforms the original time series, while the low-frequency component outperforms at the most extended 12-month horizon. Consistent with my earlier findings, using both components in the place of implied correlation results in higher R-squared across all forecasting horizons, with the improvement being more noticeable for relatively shorter horizons.

Finally, I present evidence that decomposing implied correlation improves the *out-of-sample* predictability of market returns. More specifically, I find that each component outperforms implied correlation at specific forecasting horizons that match the nature of its informational content. When evaluated against the benchmark of the historical average, the high-frequency component results in an out-of-sample R-squared of 3.6% at the 1-month horizon, which is substantially higher than the 0.1% produced by implied correlation and highly significant based on the Clark-West test (Clark and West, 2007). For longer horizons, the low-frequency component has the highest out-of-sample R-squared (reaching 11.2% at the 12-month horizon, compared to 3.4% when using implied correlation), although statistically beating the historical average benchmark becomes more difficult.

The rest of the chapter is organised as follows. Section 2.2 introduces the existing literature on option implied information. Section 2.3 presents the data sources and the construction of the main variables used in the empirical analysis. Section 2.4 describes how the Empirical Mode Decomposition approach is applied to construct

the two components of implied correlation, while Section 2.5 reports the results of predictability regressions for realised correlation. Section 2.6 discusses the empirical results on the predictability of market returns, and Section 2.7 concludes.

2.2 Methodology

2.2.1 Implied Volatility

There is extensive literature on option-implied information, with option-implied volatility, in particular, having attracted significant attention. In my study, implied volatility is one of the building blocks of implied correlation. The latter is often attached to indices and can be estimated using options written on the index and its constituents.

In terms of the methods for extracting option implied volatility, it was originally obtained by backing up from Black-Scholes Model using at-the-money options, provided other inputs are given. The most important innovation comes from the theoretical result of Carr and Madan, Demeterfi et al. (1999b) and Britten-Jones and Neuberger (2000), who obtain implied volatility by exploiting option prices only, without relying on any formula. This approach is thus not restricted by unrealistic assumptions imposed by models like the Black-Scholes Model, and the implied volatility estimated using this method is called model-free implied volatility. CBOE applied this model-free method to develop the VIX or the "fear gauge". Subsequent approaches for computing model-free implied volatility were developed based on this starting point.

One of the most popular methods for calculating implied volatility is proposed by Britten-Jones and Neuberger (2000). This volatility is again independent of any option pricing model. Instead, it comes from the no-arbitrage principle. Britten-

Jones and Neuberger (2000) show that, in a risk-neutral world, underlying assets' volatility can be computed as the integral of call option prices with respect to strike prices. Jiang and Tian (2005) later made this approach more straightforward to implement by providing a discrete approximation to the Britten-Jones & Neuberger formula, in which strikes are continuous. Jiang and Tian (2005) compare the predictive capability of the Black-Sholes model implied volatility to model-free implied volatility. Model-free implied volatility is inferred from the entire set of traded options. Intuitively, it should be able to reflect more information than its counterpart implied by the Black-Scholes model.

Introduced by Martin (2011) and Martin (2016), another way of computing implied variance is utilising simple variance swaps. Buss et al. (2019) adopt this method to compute implied variances, to compute implied correlation extracted from S&P 500 index and investigate the predictive power of implied correlation. To be consistent with existing literature, in my study, the implied correlation data set is obtained from Vilkov's website directly. A detailed discussion of the empirical methodology for computing implied volatility can be found in the data and main variables section.

2.2.2 Implied Correlation

As the main variable of interest in my study, option implied correlation has been found in several studies to contain useful information. The implied correlation was first discussed in the currency market. Consider the US dollar, British Pounds and Japanese Yen. From the triangular arbitrage in the FX market, it can be shown that the correlation between currency pairs can be estimated by variances of these currencies, which can be estimated from option prices. Application of this result is well discussed in Campa et al. (1998) and Lopez and Walter (2000).

Regarding the equity market, Driessen et al. (2009) document that implied correla-

tion within the market index can be obtained by combining index options prices and prices of options written on constituent stocks. They show that implied correlation has strong explanatory power even though it is a biased prediction of realised correlation. CBOE applies a similar method to release the daily implied correlation index based on S&P 500. Driessen et al. (2009) find that the high price of correlation risk can explain the whole index variance premium. Buss and Vilkov (2012) use implied correlation and factor betas to predict future factor betas and find that implied correlation is the most efficient and unbiased predictor of beta. DeMiguel et al. (2013) apply the combination of implied volatility, skewness, correlation and variance risk premium to portfolio selection. They document that implied skewness and variance risk premium help substantially increase the Sharp ratio and certainty equivalent return. Härdle and Silyakova (2011) record that the correlation of a basket of stocks is a vital risk factor. It describes the strength of correlation among the stocks in the basket and thus is a measure of the level of diversification.

An interesting topic on the application of option implied correlation is investigating market return predictability using implied correlation or related factors. Intuitively, a higher correlation indicates higher non-diversifiable risk, and it is natural for investors to demand a higher risk premium. In other words, directly or indirectly, the implied correlation should be positively related to market returns. Such predictive power is discussed in Driessen et al. (2009), Cosemans (2011) and Faria et al. (2022). Buss et al. (2019) trace the economic rationale behind the predictability of implied correlation by checking the predictive power of its determinants to future financial and macroeconomic risks. Implied correlation is empirically found to be significant in predicting future market returns. Yet, it is not the only option-implied variable that can be used to forecast future market returns. Bollerslev et al. (2009), Drechsler and Yaron (2011) and Bollerslev et al. (2014) depict that the variance risk premium is a robust predictor of market returns for horizons up to one quarter. Furthermore, Fan et al. (2022) reports a long-term predictive power of variance risk

premium components.

Another topic that draws extensive attention is correlation risk and its application. With the existence of correlation risk, Buraschi et al. (2010) discuss optimal portfolio selection. Chang et al. (2011) and Buss and Vilkov (2012) attempt to measure systematic risk with option implied correlation. Piatti (2015) documents that correlation risk premium originated from market participants' different views about the likelihood of systematic disasters in the future. Muller, Mueller et al. (2017) focuses on correlations risk embedded in foreign exchange markets.

2.2.3 Decomposition of Information

It is widely agreed that oscillations of asset returns in the financial market are determined by multiple factors, such as macro-economy according to Bansal et al. (2014), the sentiment of investors according to Seo and Kim (2015) and oscillations in relevant markets. These factors make the market more complicated. Some studies attempt to decompose financial time series to better grasp their characteristics. Barndorff-Nielsen et al. (2004) and Barndorff-Nielsen (2005) decompose variance time series into continuous and discrete jump parts. Chen and Ghysels (2011) separate volatility into one part driven by good news and another part stimulated by bad news. Both studies document an improvement in prediction accuracy. Feunou et al. (2017) and Mete Kilic (2018) scrutinise the predictability of variance risk premium from a decomposition perspective. They decompose variance risk premiums into upside semi-variance risk premiums and downside semi-variance risk premiums. Bollerslev et al. (2015) show that the predictive power of variance risk premium can be attributed to jump tail risk.

However, none of these decomposition approaches is able to decompose financial time series into components that can be referred to as proxies of investors with dif-

ferent investment horizons. Intuitively, market participants with varying investment horizons might be represented by series with different frequencies. More specifically, high-frequency components of financial time series should perform better in describing investors with shorter investment horizons. On the contrary, low-frequency components should better capture the expectations or actions of investors with longer investment horizons. Technologies in signal processing have been discussed and applied well in the field of social science studies. Among these technologies, the Empirical Mode Decomposition (EMD) proposed by Huang et al. (1998) is an effective one.

EMD has been extensively applied in the field of signal processing after it was proposed. It also has wide application in analysing non-linear and non-stationary series in the natural sciences, engineering and computing. However, in financial studies, relatively less attention has been paid to EMD. Most existing studies that employ EMD tend to look at commodities markets. In these studies, emphasis is generally placed on event study and making forecasts. Featured studies that correspond to these two aspects include Zhang et al. (2008) which decompose three crude oil price series into several independent intrinsic modes with frequencies from high to low. These modes are reconstructed to form a fluctuating component, a slow-varying component and a trend component via a fine-to-coarse approach. Economic meanings are assigned to these three components as short-term oscillations caused by supply-demand disequilibrium or some other market arbitrage activities, the impact of a significant event shock, and a long-term trend, respectively. As a result, the daily crude oil price can be regarded as a composition of a trend price, a significant event price, plus a fluctuation. Another EMD-based event study by Zhang et al. (2009) adopted EMD to analyse responses of crude oil prices to extreme events, such as the Persian Gulf War in 1991 and the Iraq War in 2003. They follow the fine-to-coarse method to reconstruct intrinsic mode functions and find that the entire impact of an extreme event can be concluded by one or several dominant modes.

Regarding the predictability of commodity prices, Yu et al. (2008) adopt EMD-based neural network methodology to forecast crude oil prices. West Texas Intermediate (WTI) crude oil spot price and Brent crude oil spot price are decomposed into intrinsic modes in the first stage. Then a three-layer feed-forward neural network (FNN) model was constructed to model each mode. The predictions of all modes are used as inputs to an adaptive linear neural network (ALNN). As a result, an ensemble output of the crude oil price series is produced. Yu et al. (2008) find that, generally, the EMD-based methodology outperforms its counterpart, such as single ARIMA or single FNN, with the lowest RMSE and highest D statistic. Other studies that involve the application of EMD include Hua and Jiang (2015), which forecast gold prices using their EMD-obtained lagged components, and Hua et al. (2018), by applying EEMD to decompose the original returns series into the high frequency, low frequency and trend components, propose an option pricing model based on the hybrid generalised autoregressive conditional heteroskedastic (hybrid GARCH)-type model. Their empirical results indicate that this novel model significantly outperforms M-GARCH and Black–Scholes models. It can be concluded from Hua et al. (2018) that decomposing information is consistently helpful in reducing option pricing errors.

According to Yu et al. (2008), EMD has several advantages. First, the decomposition is objective, and the extraction of each sub-part of data is only determined by the original data itself. Second, it helps obtain the trending term and several Intrinsic Mode Function (IMF), resulting in a finer decomposition. Finally, it is more intuitive and can be generally applied. This is a significant advantage compared to Fast Fourier Transformation and Wavelet Analysis, which are subject to the restriction that the data to be decomposed must be stationary. With the application of EMD, I attempt to find out the best proxies of expectations of short-term investors and long-term investors. These proxies are time series formed by intrinsic mode functions.

2.3 Data and Main Variables

2.3.1 Implied volatility

To estimate implied correlations, the first step is to obtain implied volatility for both the S&P 500 index and its constituent stocks. Although the model-free method for estimating implied volatility I discussed in previous sections is probably the most frequently applied, Buss et al. (2019) adopted simple variance swaps, as introduced by Martin (2011) and Martin (2016), to construct the daily implied variance for options.

Given that the prices of call options and put options, denoted by $C_{0,t}$ and $P_{0,t}$, are observable at all strikes denoted by K, the forward price is determined by equalising $C_{0,t}(K) = P_{0,t}(K)$. Under the assumption that the stock is non-dividend paying, $F_{0,T}$ is equal to $S_0 e^{rt}$, where r denotes the risk-free rate. Note that under such conditions, the denominators are increasing as time evolves; also note that if the interest rate is assumed to be zero, the denominators are all equal. (Martin, 2011, 2016) choose V such that there is no money paid or received at the initiation. He shows that, as Δ approximates 0, the 'fair' swap variance can be calculated by

$$V = \frac{2e^{rT}}{F_{0,T}^2}\left\{\int_0^{F_{0,T}} P_{0,T}(K)dK + \int_{F_{0,T}}^{\infty} C_{0,T}(K)dK\right\} \tag{2.1}$$

and the payoff on such a simple variance swap can be replicated by taking both static and dynamic positions in call options, put options and the underlying. It appears that the structure of this expression is quite similar to that proposed by Britten-Jones and Neuberger (2000). [1] However, as stated in Buss et al. (2019), using a different approach to estimate model-free implied volatility does not affect their

[1] Another popular method for estimating implied volatility is proposed by Britten-Jones and Neuberger (2000). Assuming that the no-arbitrage principle is valid and both the dividend rate and risk-free interest rate are equal to zero, in the risk-neutral world, implied volatility can be calculated using the formula

results. Given the same option price data, estimated implied volatility and hence implied correlation should contain identical information and thus have identical predictive power. It is worth mentioning that, while the method in Britten-Jones and Neuberger (2000) is probably the most popular way of computing implied volatility in a model-free sense, and I follow this approach in the next chapters, the method in Martin (2011) and Martin (2016) is adopted in Buss et al. (2019) for implied volatility (and hence implied correlation) estimations.

2.3.2 Implied correlation

Implied correlation was first discussed for the currency market. However, it is much more frequently studied in the equity market. Driessen et al. (2009) and Skintzi and Refenes (2004) introduce implied correlation under the risk-neutral measure, with the later literature referring to it as equicorrelation. Such average implied correlation is model-free and can be calculated by

$$IC_t = \frac{\sigma_{I,t}^2 - \sum_{i=1}^{N} \omega_{i,t}^2 \sigma_{i,t}^2}{\sum_{i=1}^{N} \sum_{i' \neq i} \omega_{i,t} \omega_{i',t} \sigma_{i,t} \sigma_{i',t}} \tag{2.2}$$

where $\sigma_{I,t}^2$ is index's model-free option-implied variance at t, $\sigma_{i,t}^2$ is the model-free implied variance of stock i at time t, and $\omega_{i,t}$ is weight of component stock i at time t. As in Buss et al. (2019) expression (2.2) is used to compute implied correlations. Since each input implied volatility is model-free, the estimated term structure of implied correlation is also regarded as model-free and is thus denoted by ICs. Note that since MFIV estimations are conducted under the $\mathbb{Q}$-measure, the implied cor-

$$E^Q\left[V_{0,T}\right] = 2e^{rT}\left[\int_0^{F_{0.T}} \frac{P(K,T)}{K^2} dK + \int_{F_{0.T}}^{\infty} \frac{C(K,T)}{K^2} dK\right]$$

where E^Q is expected variance over period $[0,T]$ in the risk-neutral world, F is the underlying's forward price at time zero, $P(K,T)$ and $C(K,T)$ are prices of call options and put options written on the same underlying with an identical time to maturity. Finally, K denotes the strike price, while T denotes the option's time-to-maturity.

relations calculated are risk-neutral as well.

The empirical analysis is implemented on the option-implied correlation extracted from S&P500 index and its constituents from January 1996 to December 2020. The implied correlation (IC) series is computed following the methodology presented in Buss et al. (2019) using index options and options written on the constituent stocks, and it refers to the risk-neutral expected average correlation among the index's components. To be consistent with the existing literature, I obtain the S&P500 implied correlation directly from the website of Grigory Vilkov.[2] The data set includes time series of IC across several standardised maturities (30, 90, 180, and 360 days).

2.3.3 Realised correlation

I follow Skintzi and Refenes (2004) to estimate Realised Correlation (RC) as

$$RC_t = \frac{\sum_{1<i<j<N} \omega_{i,t}\omega_{i,t}Corr_{i,j,t}}{\sum_{1<i<j<N} \omega_{i,t}\omega_{i,t}} \tag{2.3}$$

where $Corr_{i,j,t}$ represents the Pearson correlation coefficient between stock i and stock j at time t. The Correlation Risk Premium (CPR) is just the difference between implied and realised correlations.

2.3.4 Control Variables

The time series of the S&P500 index returns and the 1-month T-Bill rate (used as a proxy for the risk-free rate) are obtained from Bloomberg. I also include a set of control variables in the market returns predictive regressions. These control variables have been shown to have significant predictive power over the equity re-

[2]The link of Grigory Vilkov's website is https://www.vilkov.net/codedata.html.

turns (for a more detailed discussion of these factors and their predictive ability, see Goyal and Welch, 2008) and they can be obtained from the website of Amit Goyal. More specifically, the set of control variables includes the Book-to-Market Ratio (BM), T-Bill Rates (TB), AAA-Rated Corporate Bonds Yields (AAA), BAA-Rated Corporate Bonds Yields (BAA), Inflation (INFL), Long-Term Yield (LTY), Net Equity Expansion (NEE), Long-Term Rate of Return (LTR), Corporate Bond Returns (CBR), and the S&P 500 Index Variance (SVAR). [3]

Figure 2.1 plots the time series of the 30-day implied and realised correlation among the components of the S&P500 index during the period January 1996 to December 2020, while Table 2.1 presents a set of summary statistics for Implied Correlation (IC), Realised Correlation (RC) and the correlation risk premium ($CRP = IC - RC$) for maturities of 30, 60 and 90 days. I confirm the existence of a premium for correlation risk. This is indicated by the fact that implied correlation tends to be predominantly higher than corresponding realised correlation, which is consistent with the findings in Driessen et al. (2009). For instance, the mean 30-day implied correlation is equal to 0.37, compared to a mean realised correlation of 0.32. I also observe a term premium in implied correlation, which means that implied correlation is consistently higher for those extracted from options with longer maturities. More specifically, while the mean 30-day implied correlation is 0.37, the 90-day and 180-day ICs are 0.41 and 0.44, respectively. On the contrary, a similar term premium does not appear to be present in realised correlation, which tends to be fairly constant on average across different horizons.

[3]Predicting market return is a wide-ranging topic. Versatile factors that predict market return are being proposed constantly. To be consistent with the fashion of studies in this strand, I stick to the most commonly used factors in existing literature that are closely connected to this chapter. Other factors with predictive power to market return, such as business cycle and historical average return, are not included for consistency concern.

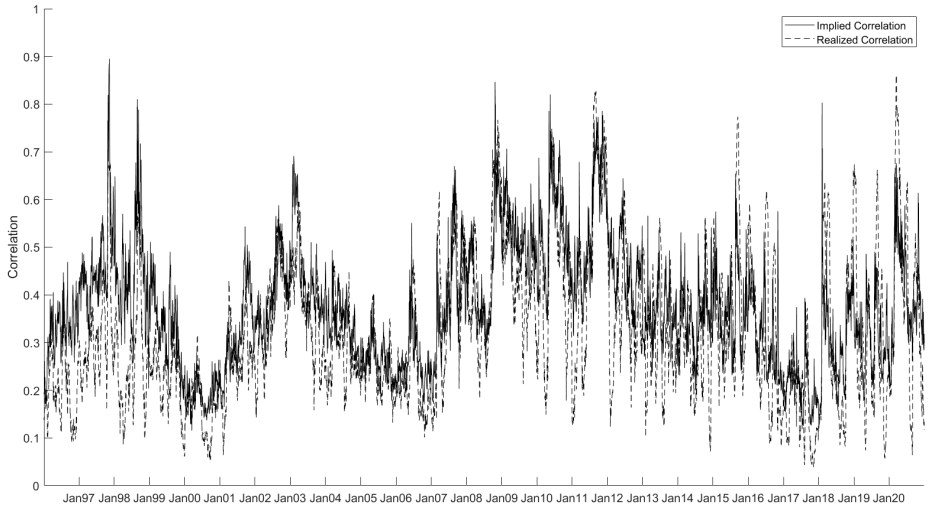

Figure 2.1: Implied correlation versus realised correlation

Notes: This Figure plots the time series of implied correlation and realised correlation among the components of the S&P500 index. Implied correlation is computed via the Driessen et al. (2012) methodology using options with 30 days to maturity, while the realised correlation time series refers to the cross-sectionally weighted average of pairwise realised correlations among the constituent stocks. The sample period runs from January 1996 to December 2020.

Table 2.1: Summary statistics of implied and realised correlation

	30 days			90 days			180 days		
	IC	RC	CRP	IC	RC	CRP	IC	RC	CRP
mean	0.37	0.32	0.05	0.42	0.32	0.09	0.44	0.33	0.11
median	0.36	0.29	0.06	0.41	0.31	0.10	0.44	0.31	0.12
st.dev	0.13	0.15	0.11	0.11	0.13	0.09	0.10	0.12	0.09
min	0.12	0.04	-0.40	0.14	0.06	-0.20	0.17	0.10	-0.21
max	0.89	0.86	0.46	0.86	0.80	0.49	0.75	0.73	0.47
skew	0.60	0.76	-0.38	0.20	0.87	-0.08	-0.12	0.90	-0.39
kurt	0.24	0.18	0.81	0.09	0.83	0.70	0.15	0.75	0.55
5^{th} prctile	0.19	0.11	-0.15	0.23	0.14	-0.06	0.25	0.16	-0.05
95^{th} prctile	0.60	0.62	0.23	0.60	0.60	0.24	0.61	0.61	0.25

Notes: This Table reports a set of summary statistics for the time-series of implied correlation (IC), realised correlation (RC) and the difference between them (CRP) for the components of the S&P500 index. Implied correlation is computed via the Driessen et al. (2012) methodology using options with maturities of 30, 90 and 180 days, while the respective realised correlation time series refers to the cross-sectionally weighted average of pairwise realised correlations among the constituent stocks. The summary statistics reported include the mean, median, standard deviation, minimum, maximum, skewness, and kurtosis, as well as the 5^{th} and 95^{th} percentiles. The sample period runs from January 1996 to December 2020.

2.4 Implied Correlation Decomposition

The decomposition of the implied correlation time series into a low-frequency and high-frequency component is based on the Empirical Mode Decomposition (EMD) technique. First proposed by Huang et al. (1998), the EMD is an adaptive technique that uses the Hilbert-Huang transform to decompose a time series into a small number of oscillatory functions termed Intrinsic Mode Function (IMF). The EMD technique considers the original time series as high-frequency oscillations being superimposed on low-frequency oscillations so that IMFs can be extracted based on the principle of local scale separation. Every IMF needs to satisfy two conditions, namely (1) the number of extrema and zero-crossings are the same or, at most, differ by one, and (2) they are symmetric with respect to a local zero mean.

The EMD is applied as an adaptive process where several IMFs $c_j(t)$ are iteratively extracted from a time-series $x(t)$ according to a pre-specified sifting procedure until the last residual term $r_n(t) = x(t) - \sum_{j=1}^{n} c_j(t)$ becomes a monotonic function or has at most one local extremum. At that final iteration, the EMD technique cannot extract another IMF, and the original time series can be expressed as the sum of extracted IMFs plus the final residual term. This process begins by extracting the IMF with the highest frequency and then proceeds by iteratively extracting IMFs with the next highest frequency from the differences between the original time series and the extracted IMFs until no further IMF is contained in the residual term.

In my empirical analysis, I employ the sifting procedure described in Yu et al. (2008) and Zhang et al. (2009) to extract IMFs from the time series of implied correlation IC. The sifting procedure can be described as in appendix A.[4]

Once the sifting process is complete, the implied correlation time series can be represented as the sum of the n extracted Intrinsic Mode Functions $c_i(t), i = 1, 2, ..., n$

[4]See Yu et al. (2008) and Zhang et al. (2009) for a more detailed discussion on the implementation of the EMD algorithm.

plus the residual term $res(t)$ as

$$IC(t) = \sum_{i=1}^{n} c_i(t) + res(t) \tag{2.4}$$

The EMD technique has certain advantages compared to more traditional decomposition techniques, such as Fourier transform and wavelet analysis. First, EMD is more flexible in terms of allowing for the decomposition of non-stationary time series, compared to the relatively more rigid traditional approaches, which require stationarity. Hence, EMD can decompose a non-stationary time series into components that are locally stationary and oscillating around zero. Second, each IMF is associated with a specific oscillation period and, by extension, with a specific time scale. As a result, the EMD approach can address multi-scaling issues that can pose a particular challenge in forecasting exercises (Di Matteo, 2007). Third, EMD does not require a base filter function to be determined before decomposition, which can be a fairly challenging requirement in wavelet analysis.[5]

When I apply EMD on the time series of 30-day implied correlation, I extract a total of nine IMFs c_j plus the trending term res.[6] Figure 2.2 plots the time series of these IMFs and the trending term extracted from the 30-day implied correlation, with IMFs plotted in descending order of frequency. I then re-group the extracted IMFs to create a high-frequency and a low-frequency component of implied correlation. More specifically, I construct the High-Frequency Component (ICH) as the sum of the first five (highest-frequency) IMFs. In comparison, the Low-Frequency Component (ICL) is computed as the sum of the last four (lowest-frequency) IMFs.[7]

[5]EMD has been extensively used for the analysis of non-linear and non-stationary series in the natural sciences, engineering and computing. However, this technique has yet to receive equivalent attention in finance research. Some notable exceptions include Yu et al. (2008) and Zhang et al. (2009) who use EMD to forecast crude oil prices, Hua and Jiang (2015) who forecast gold prices using their EMD-obtained lagged components, and Hua et al. (2018) who apply EMD in option pricing.

[6]The specific number of IMFs that EMD extracts varies across the different IC time series that refer to options with different maturities (i.e. 30, 90, 180, and 360 days).

[7]I have also explored the informational content of implied correlation components that have been constructed according to the Zhang et al. (2009) fine-to-coarse technique. This approach results in a somewhat different re-grouping of the IMFs, with the two components are given as

The reconstruction of implied correlation into high- and low-frequency components by aggregating different IMFs allows us to identify the informational content of the original time series that corresponds to different time scales, particularly in terms of distinguishing between short-term and long-term effects.

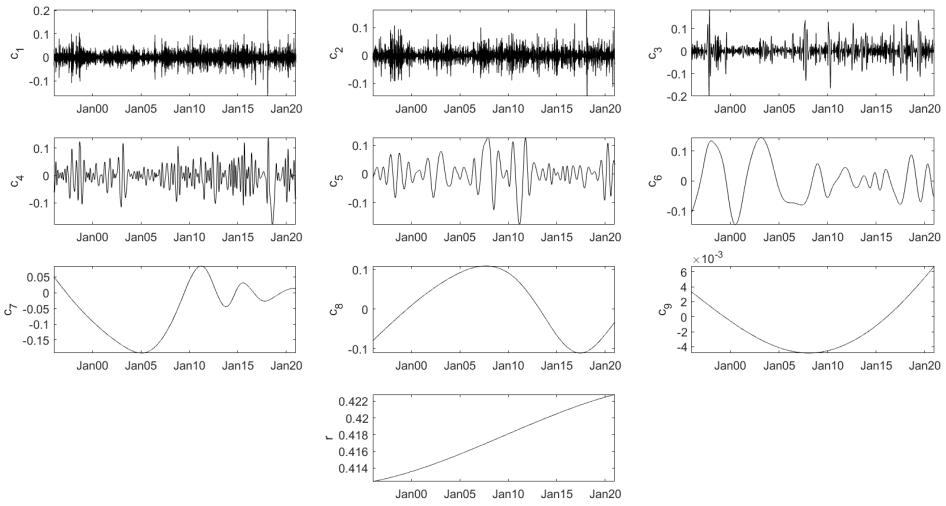

Figure 2.2: Intrinsic Mode Functions of implied correlation

Notes: This Figure plots the time-series of Intrinsic Mode Functions c_j (IMFs) and the residual term res that have been extracted from the time-series of 30-day implied correlation using the Empirical Mode Decomposition approach of Huang et al. (1998). The IMFs are plotted in descending order of frequency, with the residual term plotted on the bottom subplot. The sample period runs from January 1996 to December 2020.

As can be seen from Figure 2.3, the time series of ICH is relatively similar to that of the implied correlation series IC. This high-frequency component exhibits volatility clustering and, more generally, seems to track very closely the short-term fluctuations of the original time series. In contrast, the time series of ICL is much smoother than those of IC and ICH. As such, the low-frequency component appears to track the larger and more "permanent" fluctuations of the implied correlation series. In other words, the high-frequency oscillations in ICH provide useful information about the influence of irregular events that tend to have a short-term or medium-term impact on implied correlation. On the other hand, the low-frequency

$ICH = \sum_{j=1}^{6} c_j$ and $ICL = \sum_{j=7}^{9} c_j$. The subsequent empirical results on the predictability of realised correlation and market returns are qualitatively similar to, albeit not as strong as, those presented in the paper. I omit these additional results for brevity, but they are available upon request.

oscillations that are included in ICL can be thought of as reflecting information about more extreme events that have a long-term impact on the original series, potentially by leading to structural breaks (see also Zhang et al., 2009, 2008).

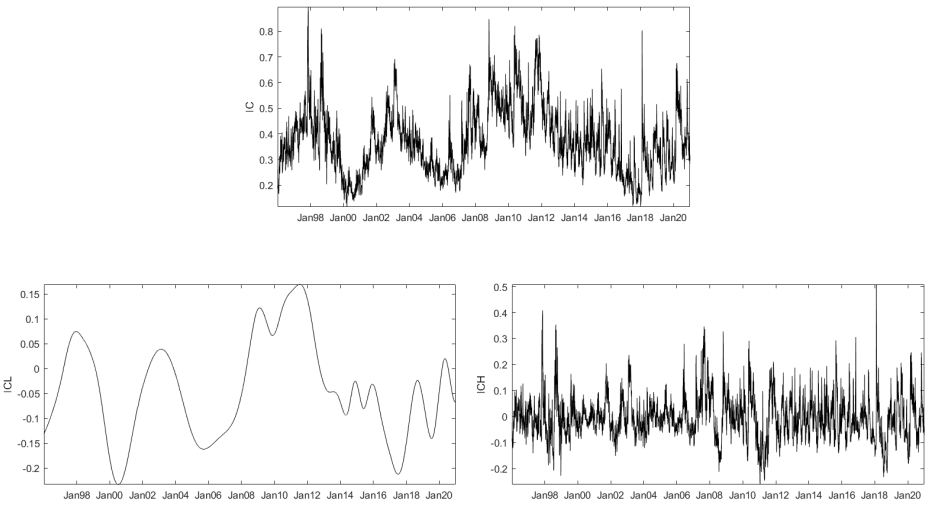

Figure 2.3: Components of implied correlation

Notes: This Figure plots the time series of implied correlation and its components. The implied correlation series has been decomposed into its Intrinsic Mode Functions (IMFs) using the Empirical Mode Decomposition approach of Huang et al. (1998). The upper subplot depicts the time series of 30-day implied correlation IC. The bottom left subplot depicts the time series of the low-frequency component ICL, given as the sum of the trending term plus the last two IMFs. The bottom right subplot depicts the time series of the high-frequency component ICH, given as the sum of the first two IMFs. The sample period runs from January 1996 to December 2020.

As a preliminary examination of how each component contributes to the informational content of implied correlation, I run simple daily regressions of IC separately against ICL and ICH as below.

$$IC_t^h = \alpha + \beta ICH_t^h + \epsilon \tag{2.5}$$

$$IC_t^h = \alpha + \beta ICL_t^h + \epsilon \tag{2.6}$$

Where h stands for the horizon of IC that takes values of 30, 90 and 180. As shown in Table 2.2, both components are highly significant when used to explain the original time series of implied correlation, suggesting that both low-frequency and high-frequency oscillations are likely to be informative. However, it is also clear

that the high-frequency component captures the largest proportion of the variance of implied correlation across all three maturities. More specifically, at the 30-day maturity ICL explains around 38% of the daily variation in implied correlation, while the respective figure for ICH is 58%.

Table 2.2: Regressions of implied correlation against its components

	30 days		90 days		180 days	
constant	0.3762	0.4112	0.4182	0.4137	0.4417	0.4440
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ICL	0.9453		1.0195		1.0135	
	(0.00)		(0.00)		(0.00)	
ICH		0.9645		0.9942		0.9896
		(0.00)		(0.00)		(0.00)
R^2_{Adj}	37.7%	58.0%	17.0%	81.1%	11.4%	79.9%

Notes: This Table reports the results from daily regressions of implied correlation IC against its low-frequency component (ICL) and its high-frequency component (ICH). Implied correlation is computed via the Driessen et al. (2012) methodology using options with maturities of 30, 90 and 180 days. Implied correlation has been decomposed into ICL and ICH using the Empirical Mode Decomposition approach of Huang et al. (1998). The Table reports the estimated coefficients, the associated p-values (in brackets) and the Adjusted R-squared. Statistical significance is based on Newey and West (1987) heteroscedasticity and autocorrelation consistent standard errors. The sample period runs from January 1996 to December 2020.

The difference between the explanatory power of the two components is even more pronounced when I consider longer maturities. For example, regressing IC against its high-frequency component ICH results in an R^2 of 82% at the 90-day maturity, compared to only 17% when using the low-frequency component ICL. This is consistent with Zhang et al. (2009), who document that the information contained in the original time series might be completely subsumed by one or several dominant intrinsic mode functions from EMD.

2.5 Realised Correlation Predictability

The empirical analysis begins with exploring the predictive ability of implied correlation over subsequent realised correlation. The aim is to understand the informational content of implied correlation and its components. More specifically, I perform predictive regressions of realised correlation RC against lagged values of realised correlation, implied correlation IC, and low- and high-frequency components of implied correlation, denoted by ICL and ICH, respectively. For each forecasting

horizon from 30 days to 180 days, the predictive regressions can be described by below equations.

$$RC_{t+h} = \alpha + \beta RC_t + \epsilon_{t+h} \tag{2.7}$$

$$RC_{t+h} = \alpha + \beta IC_t + \epsilon_{t+h} \tag{2.8}$$

$$RC_{t+h} = \alpha + \beta_1 RC_t + \beta_2 IC_t + \epsilon_{t+h} \tag{2.9}$$

$$RC_{t+h} = \alpha + \beta_1 RC_t + \beta_2 ICL_t + \epsilon_{t+h} \tag{2.10}$$

$$RC_{t+h} = \alpha + \beta_1 RC_t + \beta_2 ICH_t + \epsilon_{t+h} \tag{2.11}$$

$$RC_{t+h} = \alpha + \beta_1 RC_t + \beta_2 IC_t + \beta_2 ICL_t \epsilon_{t+h} \tag{2.12}$$

$$RC_{t+h} = \alpha + \beta_1 RC_t + \beta_2 IC_t + \beta_2 ICH_t \epsilon_{t+h} \tag{2.13}$$

$$RC_{t+h} = \alpha + \beta_1 RC_t + \beta_2 ICL_t + \beta_2 ICH_t \epsilon_{t+h} \tag{2.14}$$

Where h denotes the forecasting step size ranging between 30 and 180.

Table 2.3 reports the results of these predictive regressions, with Panels A, B and C referring to horizons of 30, 90 and 180 days, respectively.[8]

When using 30-day predictors, I find that the predictive power over realised correlation can be substantially increased by predicting with implied correlation. As a comparison, when lagged values of realised correlation are the only predictor, it fails to outperform its implied counterpart. In particular, I document that 32% of the variation in RC can be explained by IC, and the R^2 of the bivariate predictive regression where lagged RC is set as the predictor is only 21%. I also find that both lagged RC and IC are positively related to realised correlation, with the associated coefficients being highly significant. This is consistent with the results in Driessen

[8]I use daily data to estimate these predictive regressions for multi-day prediction horizons (namely for realised correlations that are observed over the subsequent 30, 90 and 180 calendar days). Therefore, statistical inference is based on Newey and West (1987) autocorrelation consistent standard errors, with the number of lags determined by the number of overlapping observations (for example, 21 lags in the case of the 30-day horizon).

et al. (2012). However, my empirical results provide stronger support for the informational content of implied correlation than those reported by Driessen et al. (2012), who find that 30-day RC and IC have approximately the same predictive power over subsequent realised correlation (with the R^2 in bivariate predictive regressions being equal to 37% and 36%, respectively). Furthermore, when both RC and IC are used as predictors in a bivariate regression, the resulting R^2 increases marginally to 33% compared to when using only IC, suggesting that the informational content of implied correlation is not only higher than, but it also subsumes to a large extent, that of lagged realised correlation.

Table 2.3: Realised correlation predictability

Panel A: 30 days								
constant	0.1734	0.0659	0.0641	0.0028	0.1714	0.0026	0.0026	0.0026
	(0.00)	(0.00)	(0.00)	(0.61)	(0.00)	(0.64)	(0.64)	(0.64)
RC	0.4608		0.1360	0.1146	0.4659	0.0824	0.0824	0.0824
	(0.00)		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
IC		0.6876	0.5746			0.1776	0.7760	
		(0.00)	(0.00)			(0.00)	(0.00)	
ICL				0.7450		0.5984		0.7760
				(0.00)		(0.00)		(0.00)
ICH					-0.0509		-0.5984	0.4116
					(0.06)		(0.00)	(0.00)
R^2_{Adj}	21.2%	32.1%	33.1%	37.3%	21.2%	37.8%	37.8%	37.8%
Panel B: 90 days								
constant	0.1892	0.1010	0.1098	0.3730	0.1845	0.4705	0.0635	0.3686
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RC	0.4186		0.1993	-0.0592	0.4289	-0.0240	0.1243	-0.0484
	(0.00)		(0.00)	(0.00)	(0.00)	(0.14)	(0.00)	(0.00)
IC		0.5364	0.3608			-0.2448	0.5229	
		(0.00)	(0.00)			(0.00)	(0.00)	
ICL				0.7559		0.9356		0.7454
				(0.00)		(0.00)		(0.00)
ICH					-0.1962		-0.4482	-0.0780
					(0.00)		(0.00)	(0.00)
R^2_{Adj}	17.7%	20.1%	22.0%	34.1%	18.7%	35.1%	26.0%	34.2%
Panel C: 180 days								
constant	0.1967	0.1534	0.1716	0.2943	0.1946	0.3491	0.1553	0.2907
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RC	0.4110		0.3531	0.1539	0.4146	0.1817	0.3275	0.1619
	(0.00)		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
IC		0.3996	0.0994			-0.1380	0.1523	
		(0.00)	(0.00)			(0.00)	(0.00)	
ICL				0.3754		0.4520		0.3678
				(0.00)		(0.00)		(0.00)
ICH					-0.1347		-0.1872	-0.1033
					(0.00)		(0.00)	(0.00)
R^2_{Adj}	16.6%	10.7%	16.9%	21.1%	17.1%	21.5%	17.8%	21.4%

Notes: This Table reports the results of predictive regressions of realised correlation (RC) on lagged realised correlation, implied correlation (IC), the low-frequency component of implied correlation (ICL) and the high-frequency component of implied correlation (ICH). Implied correlation is computed via the Driessen et al. (2012) methodology using options with maturities of 30, 90 and 180 days, with the results presented in Panels A, B and C, respectively. The corresponding realised correlation time series refers to the cross-sectionally weighted average of pairwise realised correlations among the constituent stocks. Implied correlation has been decomposed into ICL and ICH using the Empirical Mode Decomposition approach of Huang et al. (1998). The Table reports the estimated coefficients, the associated p-values (in brackets) and the Adjusted R-squared. Statistical significance is based on Newey and West (1987) heteroscedasticity and autocorrelation consistent standard errors, with the number of lags given by the number of overlapping observations. The sample period runs from January 1996 to December 2020.

I also find that the incremental predictive ability of implied correlation decreases at the 90-day horizon, while it effectively disappears at the even longer 180-day horizon. As can be seen from Panel B of Table 2.3, using the 90-day implied correlation in bivariate predictive regressions results in an R^2 of 20%, compared to an R^2 of 18% when using lagged RC as the single predictor. Using both lagged RC and IC further improves the predictive power to 22%, with the coefficient of implied correlation remaining positive and highly significant. In contrast, the results reported in Panel C indicate that the 180-day IC has a substantially lower predictive power ($R^2 = 11\%$) than that of lagged realised correlation ($R^2 = 17\%$) in bivariate regressions, and it fails to improve the predictive power in bivariate regressions.

Importantly, my results highlight that the low-frequency implied correlation component ICL has a significantly higher predictive power over realised correlation relative to that of the more noisy IC. At the 30-day horizon, using ICL in bivariate regressions increases the R^2 to 37%, compared to 33% when using the original IC series. The ICL component remains a significant predictor resulting in a high R^2 when I also include either IC or ICH in the predictive regression, with the R^2, in this case, reaching 38%. This improvement in RC predictability is even bigger at the 90-day horizon, where the R^2 increases from 22% when using IC to 34% when using ICL in bivariate regressions. Finally, at the longest 180-day horizon, using the low-frequency component ICL in bivariate regressions results in an R^2 of 21%, compared to 17% when using IC.

In contrast, the high-frequency component ICH is not found to have any incremental predictive power over realised correlation at any horizon. For instance, using the 30-day ICH in bivariate predictive regression results in an insignificant slope coefficient (p-value is 0.06) and an R^2 of 21%, which is approximately equal to that obtained when using only the lagged RC as a predictor. More generally, including ICH in the predictive regressions produces negative slope coefficients across all horizons and results in R^2s that are roughly equal to (or, on occasion, lower than) those obtained

from regressions that include only IC and/or ICL.

In summary, my empirical results suggest that the previously documented ability of implied correlation to forecast future realised correlation can be attributed primarily to the informational content of its low-frequency component. The high-frequency component of implied correlation, on the other hand, does not seem to have a significant contribution in terms of forecasting power. In this sense, I show that the previously reported forecastability of realised correlation using option-implied information stems mainly from effectively predicting long-term trends rather than short-term fluctuations.

2.6 Market Return Predictability

2.6.1 Predictability Regressions

Bollerslev et al. (2009) were among the first studies to show that information extracted from index options can substantially improve the predictability of aggregate market returns. More specifically, Bollerslev et al. (2009) find that the Variance Risk Premium (VPR) (computed as the difference between implied and realised index variance) can explain a substantial proportion of subsequent excess index returns, particularly at the quarterly horizon. At the same time, Driessen et al. (2009) show that the variance risk premium that characterises the cross-section of stock returns can be attributed primarily to priced correlation risk. Following this line of thought, Driessen et al. (2012) find that implied correlation can further improve the predictability of market returns in excess of the information already contained in the (Bollerslev et al., 2009) variance risk premium. The role of correlation in forecasting market returns is further supported by the empirical findings of Pollet and Wilson (2010), who report that the average realised correlation has a significant forecasting

power over S&P500 index returns.

Motivated by these empirical findings on the predictability of market returns using implied correlation, I proceed by exploring the relative informational content of the latter's low-frequency and high-frequency components. Similarly to Bollerslev et al. (2009) and Driessen et al. (2012), I run simple linear regressions of S&P500 index returns against a set of lagged predictors. In particular, my set of predictors includes the Driessen et al. (2009) 30-day implied correlation IC, the correlation risk premium CRP, the 30-day model-free implied index variance $MFIV$, the Bollerslev et al. (2009) Variance Risk Premium VRP (computed as $MFIV$ minus the realised index variance given by the sum of high-frequency squared index returns), as well as the low- and high-frequency components of the 30-day implied correlation (ICL and ICH, respectively).

Similarly to Driessen et al. (2012), I begin the analysis by performing these predictive regressions in an overlapping fashion, i.e. I regress the daily observations of excess market returns that have been compounded over several multi-day horizons against the set of lagged predictors that have been observed on every day of my sample period. I control for the resulting autocorrelation issues by computing Newey and West (1987) autocorrelation-consistent standard errors, with the number of lags given by the number of overlapping observations in every case (for example, I use 21 lags when forecasting 1-month market returns). The predictive regressions are implemented following equations below.

$$r^{MKT}_{t-t+h} = \alpha + \beta IC_t + \epsilon_{t+h} \tag{2.15}$$

$$r^{MKT}_{t-t+h} = \alpha + \beta CRP_t + \epsilon_{t+h} \tag{2.16}$$

$$r^{MKT}_{t-t+h} = \alpha + \beta MFIV_t + \epsilon_{t+h} \tag{2.17}$$

$$r^{MKT}_{t-t+h} = \alpha + \beta VRP_t + \epsilon_{t+h} \tag{2.18}$$

$$r^{MKT}_{t-t+h} = \alpha + \beta_1 IC_t + \beta_2 CRP_t + \beta_3 MFIV_t + \beta 4 VRP_t + \epsilon_{t+h} \tag{2.19}$$

$$r^{MKT}_{t-t+h} = \alpha + \beta_1 CRP_t + \beta_2 MFIV_t + \beta 3 VRP_t + \beta_4 ICL_t + \epsilon_{t+h} \tag{2.20}$$

$$r^{MKT}_{t-t+h} = \alpha + \beta_1 CRP_t + \beta_2 MFIV_t + \beta 3 VRP_t + \beta_4 ICH_t + \epsilon_{t+h} \tag{2.21}$$

$$r^{MKT}_{t-t+h} = \alpha + \beta_1 CRP_t + \beta_2 MFIV_t + \beta 3 VRP_t + \beta_4 ICL_t + \beta_5 ICH_t + \epsilon_{t+h} \tag{2.22}$$

Table 2.4 presents the results of these predictive regressions for forecasting horizons of one day, one week and one month, while Table 2.5 presents the respective results for longer forecasting horizons of three, six and twelve months.

At the 1-day horizon, all predictors have significant slope coefficients in bivariate regressions. Consistent with the results in Driessen et al. (2012), implied correlation is found to be positively related to future market returns. However, even though IC is a statistically significant predictor of market returns, with an R^2 of 0.2% it appears to explain less variation than that explained by VRP, which has an R^2 of 0.5% when used as a single predictor, furthermore, including the first four predictors (IC, CRP, $MFIV$, and VRP) jointly in the same specification results in an R^2 that is substantially higher (1.1%) than that obtained from any of the bivariate regressions.

Importantly, I find that decomposing implied correlation improves the predictability of market returns. More specifically, replacing IC with its low-frequency component ICL in the extended regression specification increases the R^2 from 1.1% to 1.4%, while using the high-frequency component ICH in a similar fashion further increases the R^2 to 1.9%. In fact, the highest forecasting power is obtained when using both components in the place of the original IC series, with an associated R^2 of 2.0%. Overall, isolating the effect of short-term and long-term fluctuations in implied correlation almost doubles the proportion of market returns variation that can be explained relative to using the original implied correlation series, although

Table 2.4: Market return predictability

	I	II	III	IV	V	VI	VII	VIII
Panel A: 1-day horizon								
constant	-0.0014	0.0002	-0.0003	0.0002	-0.0010	-0.0015	-0.0003	-0.0008
	(0.01)	(0.39)	(0.18)	(0.32)	(0.06)	(0.00)	(0.23)	(0.02)
IC	0.0045				0.0003			
	(0.00)				(0.87)			
CRP		0.0031			-0.0012	-0.0007	-0.0039	-0.0034
		(0.04)			(0.47)	(0.68)	(0.02)	(0.05)
$MFIV$			0.0153		0.0246	0.0340	0.0173	0.0236
			(0.00)		(0.00)	(0.00)	(0.00)	(0.00)
VRP				0.0309	0.0445	0.0469	0.0416	0.0433
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ICL						-0.0061		-0.0038
						(0.00)		(0.02)
ICH							0.0183	0.0168
							(0.00)	(0.00)
R^2_{Adj}	0.2%	0.1%	0.3%	0.5%	1.1%	1.4%	1.9%	2.0%
Panel B: 1-week horizon								
constant	-0.0048	0.0010	0.0001	0.0010	-0.0043	-0.0039	0.0004	-0.0006
	(0.00)	(0.01)	(0.87)	(0.00)	(0.00)	(0.00)	(0.50)	(0.39)
IC	0.0172				0.0080			
	(0.00)				(0.01)			
CRP		0.0114			-0.0055	-0.0024	-0.0152	-0.0141
		(0.00)			(0.10)	(0.47)	(0.00)	(0.00)
$MFIV$			0.0375		0.0574	0.0981	0.0401	0.0531
			(0.00)		(0.00)	(0.00)	(0.00)	(0.00)
VRP				0.1183	0.1544	0.1652	0.1461	0.1496
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ICL						-0.0178		-0.0080
						(0.00)		(0.02)
ICH							0.0760	0.0729
							(0.00)	(0.00)
R^2_{Adj}	0.8%	0.2%	0.4%	1.9%	3.3%	3.6%	6.2%	6.3%
Panel C: 1-month horizon								
constant	-0.0129	0.0048	0.0047	0.0052	-0.0144	-0.0021	0.0068	0.0079
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.07)	(0.00)	(0.00)
IC	0.0520				0.0505			
	(0.00)				(0.00)			
CRP		0.0326			-0.0081	0.0047	-0.0293	-0.0306
		(0.00)			(0.20)	(0.44)	(0.00)	(0.00)
$MFIV$			0.0468		0.0210	0.1438	0.0239	0.0082
			(0.00)		(0.27)	(0.00)	(0.13)	(0.66)
VRP				0.2563	0.2813	0.3147	0.2723	0.2681
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ICL						-0.0201		0.0097
						(0.00)		(0.12)
ICH							0.2156	0.2194
							(0.00)	(0.00)
R^2_{Adj}	2.0%	0.6%	0.2%	2.5%	4.6%	3.6%	10.2%	10.3%

Notes: This Table reports the results of predictive regressions of S&P500 index returns on a set of explanatory variables, namely implied correlation (IC), the difference between implied and realised correlation (CRP), model-free implied variance ($MFIV$), the variance risk premium (VRP) computed as implied variance minus realised variance, the low-frequency component of implied correlation (ICL), and the high-frequency component of implied correlation (ICH). All implied variables have been computed using options with 30 days to maturity, while realised correlation and realised variance are also computed over 30-day periods. The predictive regressions are estimated using daily data of overlapping market returns compounded over several forecasting horizons (one day, one week and one month). The Table reports the estimated coefficients, the associated p-values (in brackets) and the Adjusted R-squared. Statistical significance is based on Newey and West (1987) heteroscedasticity and autocorrelation consistent standard errors, with the number of lags given by the number of overlapping observations. The sample period runs from January 1996 to December 2020.

the forecasting power remains relatively low at the shortest forecasting horizon of one day.

The results are even stronger when I examine market returns predictability at the 1-week horizon (Panel B of Table 2.4). For instance, regressing market returns against the lagged IC (along with the other three predictors) results in a significantly positive slope and an R^2 of 3.3%, which is considerably higher than that obtained from the 1-day predictive regressions. Replacing implied correlation with its low-frequency component improves the forecasting power to some extent, with the R^2 increasing to 3.6%. More interestingly, though, the corresponding improvement is substantially higher when I replace IC with the high-frequency component ICH, in which case the R^2 of the predictive regression almost doubles to 6.2%.

When I move to 1-month predictive regressions (Panel C of Table 2.4), the results provide even stronger support for the role of the high-frequency component but less so with respect to the low-frequency one. When replacing IC with ICL in the multivariate predictive regressions, the R^2 actually drops from 4.6% to 3.6%. In contrast, replacing IC with ICH results in the R^2 more than doubling to 10.2%, while the forecasting power when using both components in the same specification is even higher (albeit only marginally) at an R^2 of 10.3%.

The superior forecasting performance of the high-frequency component compared to that of the low-frequency component persists until the 3-month horizon, but it seems to disappear after that. As can be seen from Panel A of Table 2.5, the predictive regression's R^2 is much higher when using ICH compared to ICL (12.3% vs 7.6%), although using the original IC series results in a slightly higher R^2 (12.5%) compared to using either of its components. Nevertheless, using both ICL and ICH in the same specification still results in the highest forecasting power among the competing specifications, with an R^2 of 15.4% at the 3-month horizon. In contrast, the high-frequency component loses its significance and forecasting power at the

longer horizons of six and twelve months, while the low-frequency component does relatively better, and it even outperforms IC at the longest 12-month horizon. More specifically, at the 6-month horizon, the R^2 of using ICL is more than double that of using ICH (11.6% vs 5.7%). Furthermore, with an R^2 of 12.8% the original IC series outperforms both components when the latter are used separately, but combining both ICL and ICH in the same specification still produces the highest R^2. Finally, the results from the 12-month predictive regressions suggest that it is the low-frequency component that is driving the predictive power at this longest horizon, as it is found to result in the highest R^2 (11.3%) compared to using either the original IC (10.8%) or the high-frequency component ICH (5.0%).

To further illustrate the two components' relative performance, Figure 2.4 plots the R^2 obtained from regressing market returns against IC, ICL, ICH and $ICL-ICH$ jointly, across horizons ranging from 1 to 252 trading days. Additionally, Figure 2.5 plots these R^2 values in a pairwise sense. In each subplot, IC is compared to one of its components or the latter's combination. Obviously, the high-frequency component of implied correlation results in the highest R^2 in the short- and medium-term, outperforming the original IC series and ICL roughly until the 3-month horizon. However, the forecasting power of the low-frequency component starts to rise substantially when I move to horizons of more than 50 calendar days, with the associated R^2 rising above that of ICH at a horizon of around 80 calendar days and that of IC for horizons longer than 200 calendar days.

The above findings highlight some important features of the well-documented forecasting power of implied correlation over market returns. I confirm that implied correlation has a significant incremental predictive ability in excess of the information already contained in other option-implied predictors. However, while implied correlation remains a very significant and informative predictor of market returns at horizons ranging from one day to one year, the source of this predictability varies. For shorter horizons of up to three months, the high-frequency component of implied

Table 2.5: Market return predictability - Longer horizons

Panel A: 3-month horizon								
constant	-0.0427	0.0133	0.0169	0.0160	-0.0511	0.0140	0.0153	0.0320
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
IC	0.1643				0.1953			
	(0.00)				(0.00)			
CRP		0.1055			-0.0010	0.0362	-0.0069	-0.0284
		(0.00)			(0.92)	(0.00)	(0.50)	(0.01)
$MFIV$			0.0599		-0.1531	0.0721	0.0619	-0.1729
			(0.02)		(0.00)	(0.02)	(0.02)	(0.00)
VRP				0.6007	0.5415	0.6051	0.5827	0.5225
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ICL						0.0920		0.1478
						(0.00)		(0.00)
ICH							0.3373	0.3966
							(0.00)	(0.00)
R^2_{Adj}	7.2%	2.1%	0.1%	4.8%	12.5%	7.6%	12.3%	15.4%
Panel B: 6-month horizon								
constant	-0.0749	0.0300	0.0259	0.0361	-0.0812	0.0401	0.0140	0.0513
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
IC	0.3001				0.3176			
	(0.00)				(0.00)			
CRP		0.1592			0.0283	0.0752	0.0829	0.0353
		(0.00)			(0.06)	(0.00)	(0.00)	(0.02)
$MFIV$			0.3144		-0.1114	0.0372	0.4111	-0.1153
			(0.00)		(0.01)	(0.39)	(0.00)	(0.01)
VRP				0.5460	0.4768	0.5284	0.6093	0.4766
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ICL						0.3045		0.3386
						(0.00)		(0.00)
ICH							0.1151	0.2477
							(0.00)	(0.00)
R^2_{Adj}	10.7%	2.2%	1.3%	1.8%	12.8%	11.6%	5.7%	13.0%
Panel C: 12-month horizon								
constant	-0.0906	0.0647	0.0539	0.0774	-0.0918	0.0772	0.0343	0.0862
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
IC	0.4434				0.4298			
	(0.00)				(0.00)			
CRP		0.2551			0.1246	0.1830	0.2191	0.1512
		(0.00)			(0.00)	(0.00)	(0.00)	(0.00)
$MFIV$			0.5972		-0.0490	0.0789	0.6996	-0.0434
			(0.00)		(0.49)	(0.23)	(0.00)	(0.54)
VRP				0.3765	0.2147	0.2692	0.4073	0.2276
				(0.00)	(0.02)	(0.00)	(0.00)	(0.01)
ICL						0.4685		0.4961
						(0.00)		(0.00)
ICH							0.0091	0.1993
							(0.83)	(0.00)
R^2_{Adj}	9.9%	2.5%	2.1%	0.4%	10.8%	11.3%	5.0%	11.6%

Notes: This Table reports the results of predictive regressions of S&P500 index returns on a set of explanatory variables, namely implied correlation (IC), the difference between implied and realised correlation (CRP), model-free implied variance ($MFIV$), the variance risk premium (VRP) computed as implied variance minus realised variance, the low-frequency component of implied correlation (ICL), and the high-frequency component of implied correlation (ICH). All implied variables have been computed using options with 30 days to maturity, while realised correlation and realised variance are also computed over 30-day periods. The predictive regressions are estimated using daily data of overlapping market returns compounded over several forecasting horizons (three, six and twelve months). The Table reports the estimated coefficients, the associated p-values (in brackets) and the Adjusted R-squared. Statistical significance is based on Newey and West (1987) heteroscedasticity and autocorrelation consistent standard errors, with the number of lags given by the number of overlapping observations. The sample period runs from January 1996 to December 2020.

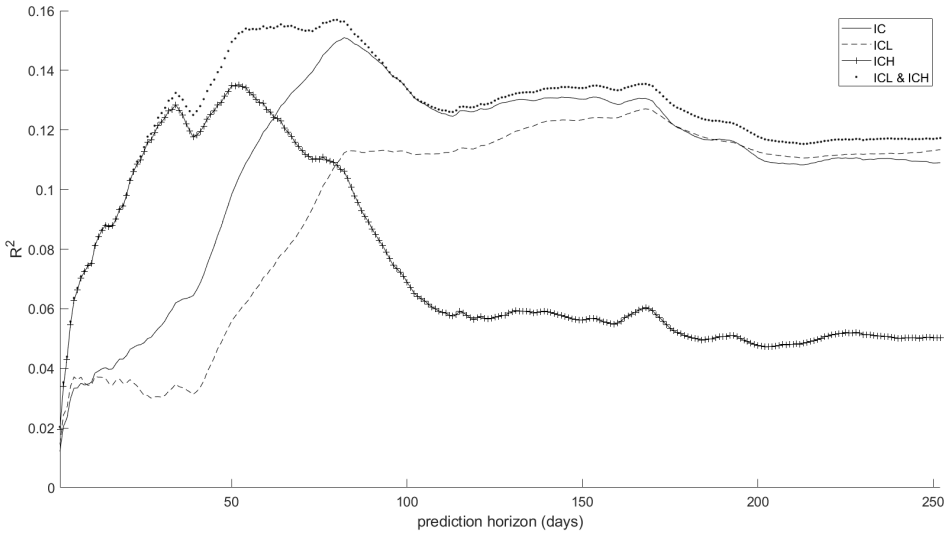

Figure 2.4: R-squared of market predictability regressions

Notes: This Figure plots the R^2 from predictive regressions of market returns against a number of predictors, namely implied correlation (IC), the low-frequency component of implied correlation (ICL), and the high-frequency component of implied correlation (ICH). All predictive regressions also include a set of control predictors, namely the difference between implied and realised correlation (CRP), model-free implied variance ($MFIV$), and the variance risk premium (VRP) computed as implied variance minus realised variance. All implied variables have been computed using options with 30 days to maturity, while realised correlation and realised variance are also computed over 30-day periods. The predictive regressions are estimated using daily data of overlapping market returns compounded over several forecasting horizons, ranging from 1 to 252 days. The sample period runs from January 1996 to December 2020.

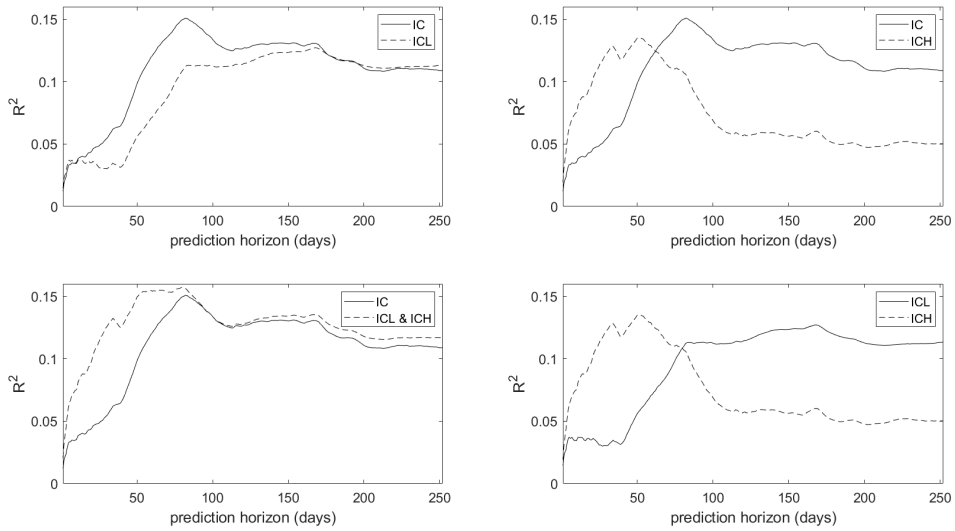

Figure 2.5: R-squared of market predictability regressions

Notes: This Figure plots the R^2 from predictive regressions of market returns against a number of predictors, namely implied correlation (IC), the low-frequency component of implied correlation (ICL), and the high-frequency component of implied correlation (ICH). All predictive regressions also include a set of control predictors, namely the difference between implied and realised correlation (CRP), model-free implied variance ($MFIV$), and the variance risk premium (VRP) computed as implied variance minus realised variance. All implied variables have been computed using options with 30 days to maturity, while realised correlation and realised variance are also computed over 30-day periods. The predictive regressions are estimated using daily data of overlapping market returns compounded over several forecasting horizons, ranging from 1 to 252 days. The sample period runs from January 1996 to December 2020.

correlation outperforms both the original series and the low-frequency component, consistent with the intuition that factors with short-term influences are more relevant when forecasting short-term market returns. In contrast, the low-frequency component dominates when forecasting market returns at horizons of six months or longer, indicating that identifying long-term trends in the predictor is substantially more informative when forecasting long-term market returns.

2.6.2 Controlling for Fundamentals

I proceed by examining the extent to which implied correlation and its components remain significant predictors of market returns after controlling for a set of fundamental variables that have been previously shown to contain important information about subsequent market returns. To this end, I estimate predictive regressions of market returns against IC, ICL, and ICH, while also incorporating in the specification the fundamental variables proposed by Goyal and Welch (2008). More specifically, Goyal and Welch (2008) show that a particular combination of fundamental variables can explain a substantial proportion of subsequent market returns. Therefore, I augment the predictive regressions with the same set of fundamentals, namely the Book-to-Market Ratio (BM), T-Bill Rates (TB), AAA-Rated Corporate Bonds Yields (AAA), BAA-Rated Corporate Bonds Yields (BAA), Inflation (INFL), Long-Term Yield (LTY), Net Equity Expansion (NEE), Long-Term Rate of Return (LTR), Corporate Bond Returns (CBR), and the S&P 500 Index Variance (SVAR).

I perform predictive regressions of market returns that have been compounded over several different horizons (namely one, three, six and twelve months) against implied correlation, its components, and the fundamental variables. Since the fundamentals are available at a monthly frequency, I sample market returns and the set of predictors at the end of each month instead of every day, as in subsection 2.6. The

regression equations, with or without controlling fundamentals, are listed below and

$$r^{MKT}_{t-t+m} = \alpha + \beta IC_t + \epsilon_{t+m} \tag{2.23}$$

$$r^{MKT}_{t-t+m} = \alpha + \beta ICL_t + \epsilon_{t+m} \tag{2.24}$$

$$r^{MKT}_{t-t+m} = \alpha + \beta ICH_t + \epsilon_{t+m} \tag{2.25}$$

$$r^{MKT}_{t-t+m} = \alpha + \beta_1 ICL_t + \beta_2 ICH_t + \epsilon_{t+m} \tag{2.26}$$

$$r^{MKT}_{t-t+m} = \alpha + CTRL_t + \epsilon_{t+m} \tag{2.27}$$

$$r^{MKT}_{t-t+m} = \alpha + CTRL_t + \beta_{11} IC_t + \epsilon_{t+m} \tag{2.28}$$

$$r^{MKT}_{t-t+m} = \alpha + CTRL_t + \beta_{11} ICL_t + \epsilon_{t+m} \tag{2.29}$$

$$r^{MKT}_{t-t+m} = \alpha + CTRL_t + \beta_{11} ICH_t + \epsilon_{t+m} \tag{2.30}$$

$$r^{MKT}_{t-t+m} = \alpha + CTRL_t + \beta_{11} ICL_t + \beta_{12} ICH_t + \epsilon_{t+m} \tag{2.31}$$

where

$$\begin{aligned} CTRL_t = &\beta_1 BM_t + \beta_2 TB_t + \beta_3 AAA_t + \beta_4 BAA_t + + \beta_5 INFL_t + \beta_6 LTY_t + \\ &\beta_7 NEE_t + \beta_8 LTR_t + \beta_9 CBR_t + \beta_{10} SVAR_t + \epsilon_{t+m} \end{aligned} \tag{2.32}$$

The regression results are reported in Table 2.6.

Similarly to my earlier results using daily returns (reported in Tables 2.4 and 2.5), I find that the forecasting power of implied correlation generally increases with the length of the forecasting horizon. For instance, the R^2 of monthly bivariate regressions of market returns against the lagged IC is only 1.6% at the 1-month horizon, compared to 10.7% at the 6-month horizon. Notably, the results from monthly regressions confirm the incremental forecasting power of the high-frequency component of implied correlation for short-term and medium-term forecasting horizons. More specifically, using ICH instead of IC in bivariate predictive regressions in-

Table 2.6: Market return predictability - Controlling for fundamentals

Panel A: 1-month horizon									
constant	-0.0110	0.0047	0.0056	0.0041	-0.0045	-0.0263	-0.0113	-0.0151	-0.0231
	(0.16)	(0.63)	(0.03)	(0.67)	(0.97)	(0.80)	(0.91)	(0.88)	(0.82)
IC	0.0458					0.0856			
	(0.02)					(0.00)			
ICL		0.0051		0.0040			0.0399		0.0464
		(0.84)		(0.87)			(0.24)		(0.16)
ICH			0.1144	0.1144				0.1241	0.1266
			(0.00)	(0.00)				(0.00)	(0.00)
control	no	no	no	no	yes	yes	yes	yes	yes
R^2_{Adj}	1.6%	0.0%	4.1%	3.8%	3.5%	7.3%	3.6%	7.7%	8.1%
Panel B: 3-month horizon									
constant	-0.0428	0.0049	0.0162	0.0024	0.0249	-0.0385	0.0058	-0.0054	-0.0280
	(0.00)	(0.78)	(0.00)	(0.88)	(0.88)	(0.81)	(0.97)	(0.97)	(0.86)
IC	0.1627					0.2488			
	(0.00)					(0.00)			
ICL		0.0396		0.0368			0.1112		0.1290
		(0.37)		(0.37)			(0.04)		(0.01)
ICH			0.3683	0.3678				0.3640	0.3705
			(0.00)	(0.00)				(0.00)	(0.00)
control	no	no	no	no	yes	yes	yes	yes	yes
R^2_{Adj}	7.4%	0.0%	14.6%	14.6%	18.4%	29.7%	19.3%	31.0%	32.4%
Panel C: 6-month horizon									
constant	-0.0691	-0.0151	0.0351	-0.0170	-0.0490	-0.1301	-0.0813	-0.0860	-0.1225
	(0.00)	(0.55)	(0.00)	(0.47)	(0.82)	(0.52)	(0.71)	(0.68)	(0.55)
IC	0.2830					0.3066			
	(0.00)					(0.00)			
ICL		0.1456		0.1385			0.1898		0.2092
		(0.02)		(0.02)			(0.01)		(0.00)
ICH			0.5213	0.5182				0.3943	0.4048
			(0.00)	(0.00)				(0.00)	(0.00)
control	no	no	no	no	yes	yes	yes	yes	yes
R^2_{Adj}	10.7%	1.4%	13.5%	14.8%	37.7%	45.6%	39.2%	44.5%	46.3%
Panel D: 12-month horizon									
constant	-0.0680	-0.0832	0.0777	-0.0845	-0.8568	-0.9246	-0.9385	-0.8581	-0.9420
	(0.03)	(0.03)	(0.00)	(0.02)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
IC	0.38					0.23			
	(0.00)					(0.00)			
ICL		0.4334		0.4292			0.4000		0.4016
		(0.00)		(0.00)			(0.00)		(0.00)
ICH			0.3168	0.3072				0.0135	0.0333
			(0.01)	(0.01)				(0.90)	(0.75)
control	no	no	no	no	yes	yes	yes	yes	yes
R^2_{Adj}	8.6%	6.6%	1.9%	8.4%	42.2%	43.9%	45.3%	42.0%	45.1%

Notes: This Table reports the results of predictive regressions of S&P500 index returns on a set of explanatory variables, namely implied correlation (IC), the low-frequency component of implied correlation (ICL), and the high-frequency component of implied correlation (ICH). All implied variables have been computed using options with 30 days to maturity. In addition to the implied correlation and its components, the set of predictors also includes the fundamental variables from Goyal and Welch (2008). More specifically, these control variables include the dividend-price ratio, default yield spread, Treasury-bill rate, long-term yield, book-to-market, net equity expansion, dividend yield, stock variance, term spread, inflation, and corporate bond returns. The predictive regressions are estimated using end-of-month observations, with non-overlapping market returns compounded over several forecasting horizons (one, three, six and twelve months). The table reports the estimated coefficients, the associated p-values (in brackets) and the Adjusted R-squared. Statistical significance is based on Newey and West (1987) heteroscedasticity and autocorrelation consistent standard errors. The sample period runs from January 1996 to December 2019.

creases the R^2 from 1.6% to 4.1% at the 1-month horizon. A similar improvement in forecastability is obtained at the 3-month horizon (from 7.4% to 14.6%) and the 6-month horizon (from 10.7% to 13.5%).

At the same time, the low-frequency component of implied correlation is again found to contain very little information about future market returns at short horizons. Using ICL in bivariate monthly regressions results in R^2s that are markedly lower than those obtained when using the original IC series. For example, at the 1-month and 3-month horizons, the R^2 is effectively equal to 0% when using ICL, compared to R^2s when using IC of 1.6% and 7.4%, respectively. At the 6-month horizon, ICL seems to have some predictive power over market returns, but with an R^2 of 1.4%, this is substantially lower than that of IC (10.7%) or ICH (13.5%). However, I find that ICL is actually more informative than ICH at the longest 12-month horizon, with R^2s of 6.6% vs 1.9%, respectively. Nevertheless, the original IC series remains the most informative at this longest horizon, with an R^2 of 8.6%.

I confirm that the Goyal and Welch (2008) fundamentals have a significant forecasting power over market returns, especially at longer horizons. In particular, regressing market returns against these fundamental variables results in R^2s that range from 3.5% at the 1-month horizon to 42.2% at the 12-month horizon. These R^2s are substantially higher than those obtained when only using option-implied variables as predictors, with the exception of the 1-month horizon where either implied correlation component seems to be more informative than the set of fundamentals alone.

Furthermore, the significant predictive ability of option-implied variables remains after controlling for fundamentals, consistent with the results in Driessen et al. (2012). In fact, I find that using both implied correlation (or its components) and the Goyal and Welch (2008) fundamentals in predictive regressions can result in a markedly higher R^2 compared to using only one type of predictor, especially

at short-term and medium-term forecasting horizons. More specifically, at the 1-month horizon, the highest R^2 is obtained when the set of predictors includes the fundamentals and both implied correlation components (8.1%), followed by using the fundamentals and only the high-frequency component (7.7%). To put this into context, this R^2 is more than double compared to what would have been obtained when using only the well-established Goyal and Welch (2008) fundamentals (3.5%).

The results are similar at the 3-month horizon, with the highest R^2 obtained when using the fundamentals and both implied correlation components (32.4%), followed by that obtained when using the fundamentals and ICH (31.0%). At these short-term horizons, the low-frequency component ICL seems to contain relatively little information in terms of forecasting future market returns, being consistently outperformed by the high-frequency component as well as by the original IC series. Interestingly, though, adding ICL to the set of predictors still results in a higher R^2 than only using fundamentals, highlighting the general importance of implied information in forecasting market returns.

Consistent with the findings introduced above, I find that the informational content of the high-frequency component diminishes as I move to longer forecasting horizons. At the 6-month horizon, adding ICL to the set of fundamental predictors results in an R^2 of 11.6%, which is noticeably higher than that obtained when adding ICH (5.7%). Moreover, while the original IC series outperforms each of its components when controlling for fundamentals ($R^2 = 12.8\%$), using both components still produces the highest R^2 across all candidate specifications (13.0%). The results are very similar at the 12-month horizon, with ICL outperforming ICH (11.3% vs 5.0%) and using both components resulting in the highest R^2 (11.6%).

Overall, my findings confirm that decomposing implied correlation can significantly improve the predictability of market returns, even after controlling for fundamental variables that have been shown to contain important information. The high-

frequency component of implied correlation can forecast market returns more efficiently than the original series in the short- and medium-term, potentially by providing "cleaner" information about short-term influences. In contrast, the low-frequency component appears to be more efficient in capturing longer-lived, more "permanent" effects, and in any case, replacing the original series of implied correlation with its two components results in significantly higher predictability of market returns at horizons ranging from one month to one year.

2.6.3 Crisis Periods

Implied variables have been shown to be particularly informative during crisis periods (see for instance Kempf et al., 2015). One reason could be that information flows are usually higher during crises, which makes historical information less relevant and favours implied information which is based only on current data. Another reason could be that crisis periods are often characterised by more pronounced information asymmetry, with informed investors being more likely to exploit their private information by trading in the options market relative to the spot market (Chakravarty et al., 2004). Therefore, option prices are expected to be more informative during crisis periods, compared to normal periods when they are expected to have a smaller advantage over historical information.

To better understand the forecasting performance of implied correlation and its components, I turn my attention to the two major crises that are contained in my sample period. The first crisis is the burst of the dot-com bubble, which I define as having begun in March 2000 (when the NASDAQ lost around 9% of its value in the space of just six days) and having ended in April 2003. The second is the global financial crisis, which I define as having started in June 2007 (when issues about Bear Stearns became public) and ended in December 2009.[9] I re-estimate

[9] I use the same beginning and end for each crisis period as in Baule et al. (2015). The results

the monthly predictive regressions of market returns against implied correlation, its components and the set of fundamental variables separately for each of the two crisis periods. Table 2.7 reports the results for forecasting horizons of one and three months.

As can be seen from Panel A of Table 2.7, decomposing implied correlation was even more important during the dot-com bubble burst. At the 1-month horizon, using the original IC series as a single predictor in bivariate regression results in an R^2 of only 0.3%. However, replacing IC with its high-frequency component ICH results in a very substantial increase of the R^2 to 6.7%. Similarly to the results in the full sample, the low-frequency component ICL does not seem to have any significant explanatory power over future market returns at this short-term horizon. When I control for the Goyal and Welch (2008) fundamentals, the high-frequency component remains an essential predictor of market returns, resulting in a markedly higher R^2 compared to the one obtained when using the original IC series (16.4% vs 7.3%).

The results are even stronger when I move to the longer 3-month forecasting horizon. More specifically, using the high-frequency component in bivariate predictive regressions results in an R^2 of 56.5%, compared to only 9.2% when using IC. Incorporating the set of fundamental variables further increases the predictive power, reaching a maximum of 79.5% when both components are used in the same specification (compared to 57.4% when the fundamentals are combined with IC).

The results for the global financial crisis are again more robust than those obtained in the full sample, albeit somewhat weaker than those obtained in the dot-com bubble burst period. As seen from Panel B of Table 2.7, at the 1-month horizon, the original implied correlation series outperforms both components when used as the single predictor. However, after accounting for the set of fundamental variables,

remain qualitatively the same when I use alternative definitions of crisis periods (omitted for brevity but available upon request).

using either component results in a markedly higher R^2 compared to using IC, while using both components results in the highest R^2 among all specifications (60.2%, compared to 48.1% when using IC). Finally, the results are qualitatively similar at the 3-month horizon, where using both ICL and ICH in the place of IC increases the R^2 from 80.9% to 86.7%.

Table 2.7: Market return predictability - Crisis periods

Panel A: Dotcom bubble burst									
1-month horizon									
constant	-0.0353	-0.0100	-0.0137	-0.0146	-5.1345	-5.9916	-3.5491	-5.5834	-4.5133
	(0.15)	(0.39)	(0.11)	(0.21)	(0.04)	(0.03)	(0.16)	(0.06)	(0.10)
IC	0.0703					-0.2556			
	(0.30)					(0.37)			
ICL		0.0209		-0.0105			-1.7856		-2.0757
		(0.81)		(0.90)			(0.73)		(0.05)
ICH			0.2546	0.2581				-0.0959	-0.2610
			(0.07)	(0.07)				(0.07)	(0.34)
control	no	no	no	no	yes	yes	yes	yes	yes
R^2_{Adj}	0.3%	0.0%	6.7%	4.0%	8.0%	7.3%	4.6%	16.4%	16.2%
3-month horizon									
constant	-0.1092	-0.0416	-0.0411	-0.0461	-8.9599	-6.2996	-5.7079	-4.2191	-2.6322
	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.03)	(0.02)	(0.09)	(0.20)
IC	0.2148					0.7303			
	(0.04)					(0.02)			
ICL		-0.0112		-0.0532			-3.3425		-2.6195
		(0.93)		(0.53)			(0.00)		(0.00)
ICH			1.0014	1.0085				0.9360	0.7461
			(0.00)	(0.00)				(0.00)	(0.00)
control	no	no	no	no	yes	yes	yes	yes	yes
R^2_{Adj}	9.2%	0.0%	56.6%	55.7%	46.6%	57.4%	66.6%	67.6%	79.5%
Panel B: Global financial crisis									
1-month horizon									
constant	-0.0954	-0.0086	-0.0153	-0.0215	-1.5240	-3.7816	-9.2317	-5.0927	-7.9651
	(0.12)	(0.50)	(0.23)	(0.16)	(0.67)	(0.18)	(0.01)	(0.07)	(0.02)
IC	0.1761					0.5546			
	(0.14)					(0.00)			
ICL		0.0175		0.1248			-2.7646		-1.5227
		(0.91)		(0.44)			(0.00)		(0.09)
ICH			0.1598	0.2032				0.5279	0.3248
			(0.18)	(0.13)				(0.00)	(0.06)
control	no	no	no	no	yes	yes	yes	yes	yes
R^2_{Adj}	4.1%	0.0%	2.9%	1.5%	8.3%	48.1%	52.9%	55.1%	60.2%
3-month horizon									
constant	-0.3443	-0.0298	-0.0446	-0.0720	11.1759	6.8866	0.1817	4.9102	1.1814
	(0.01)	(0.28)	(0.09)	(0.02)	(0.04)	(0.09)	(0.96)	(0.19)	(0.76)
IC	0.6538					0.7886			
	(0.01)					(0.00)			
ICL		0.1780		0.5748			-3.6424		-2.1493
		(0.58)		(0.08)			(0.00)		(0.07)
ICH			0.4944	0.7117				0.7471	0.4182
			(0.05)	(0.01)				(0.00)	(0.09)
control	no	no	no	no	yes	yes	yes	yes	yes
R^2_{Adj}	19.4%	0.0%	10.4%	17.5%	65.9%	80.9%	84.6%	84.3%	86.7%

Notes: This Table reports the results of predictive regressions of S&P500 index returns on a set of explanatory variables, namely implied correlation (IC), the low-frequency component of implied correlation (ICL), and the high-frequency component of implied correlation (ICH). The predictive regressions are estimated in two subsamples of crisis periods, namely the dot-com bubble burst (March 2000 to April 2003) and the global financial crisis (June 2007 to December 2009). All implied variables have been computed using options with 30 days to maturity. In addition to the implied correlation and its components, the set of predictors also includes the fundamental variables from Goyal and Welch (2008). More specifically, these control variables include the dividend-price ratio, default yield spread, Treasury-bill rate, long-term yield, book-to-market, net equity expansion, dividend yield, stock variance, term spread, inflation, and corporate bond returns. The predictive regressions are estimated using end-of-month observations, with non-overlapping market returns compounded over two forecasting horizons (namely, one and three months). The Table reports the estimated coefficients, the associated p-values (in brackets) and the Adjusted R-squared. Statistical significance is based on Newey and West (1987) heteroscedasticity and autocorrelation consistent standard errors.

Overall, these results highlight an interesting distinction between normal periods and crisis periods. Even though decomposing implied correlation significantly increases the predictability of market returns in the full sample, this improvement seems to be substantially more pronounced during crisis periods.

2.6.4 Out-of-Sample Predictability

Goyal and Welch (2008) show that, although several variables have been found to have significant forecasting power over market returns *in-sample*, this does not always translate to *out-of-sample* predictability too. Therefore, I proceed by exploring the out-of-sample forecasting ability of implied correlation and its low-frequency and high-frequency components. I adopt a recursive monthly forecasting scheme, with an initial estimation period of 36 months (from January 1996 to December 1998), leaving the period from January 1999 to December 2020 for the out-of-sample evaluation.

Starting from January 1999, I produce a forecast of market returns for a given month via the traditional approach of estimating a recursive predictive regression of market returns against a particular predictor (with the initial 36-month estimation period expanding at each iteration). More specifically, I estimate the following bivariate predictive regression

$$r_{t-h\to t} = \beta_0 + \beta_1 x_{t-h} + \epsilon_t \tag{2.33}$$

where $r_{t-h\to t}$ denotes the market returns computed during the holding period $[t-h, t]$, x_{t-h} denotes the predictor observed at time $t-h$, and ϵ_t is a random error term. This regression is estimated separately for each predictor, namely for IC, ICL and ICH.[10] I recursively construct out-of-sample forecasts of market returns

[10] I focus on the predictive performance of single predictors in order to avoid issues of overfitting and model selection (see also Buss et al., 2019), as my emphasis is primarily on the relative predictive ability of implied correlation and its components. I leave the evaluation of multivariate forecasting schemes and forecast combinations for future research.

at $[t, t+h]$ by estimating the regression in (2.33) during the expanding estimation period and then using the in-sample estimated coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ as well as the value of the predictor at time t. This standard approach ensures that out-of-sample forecasts are free from any look-ahead bias since they only use observations that were available when the forecast was produced. The above forecasting exercise is performed separately for forecasting horizons h of 1, 3, 6 and 12 months.

I evaluate the forecasting performance of each predictor by comparing its predictive power against that of a benchmark model. I adopt the natural benchmark of the historical average market returns computed during the estimation period, which Campbell and Thompson (2008) and Goyal and Welch (2008) find to significantly outperform a large set of commonly used predictors out-of-sample. To measure the out-of-sample forecasting performance, I rely on the commonly used out-of-sample R-squared (R^2_{OS}), given by the Mean Square Forecast Error (MSFE) of the candidate model relative to that of the benchmark model as follows.

$$R^2_{OS} = 1 - \frac{MSFE_m}{MSFE_b} \tag{2.34}$$

where $MSFE_m$ and $MSFE_b$ denote the MSFE of model m and that of the benchmark model, respectively. In this setting, a particular model m outperforms the benchmark model if the former's R^2_{OS} is positive. Statistical significance is assessed via the Clark and West (2007) test, with the null hypothesis being that model m produces a higher MSFE than the benchmark, against the alternative hypothesis that model m produces a lower MSFE. Table 2.8 reports the R^2_{OS} and the p-value from the Clark and West (2007) test of each model against the historical average benchmark across forecasting horizons of 1, 3, 6 and 12 months.

The first thing to notice is that implied correlation outperforms the historical average benchmark at the 1-month and 12-month horizons (with out-of-sample R-squared of

0.1% and 3.4%, respectively) but fails to deliver a positive R^2_{OS} at the intermediate horizons of 3 and 6 months. Moreover, the Clark and West (2007) test indicates that IC's superior out-of-sample performance relative to the benchmark is significant only at 10% for the 1-month horizon and insignificant for the longer 12-month horizon. Overall, I find limited evidence of implied correlation being a consistently superior predictor of market returns out-of-sample relative to the historical average benchmark.

Table 2.8: Market return predictability - Out-of-sample R-squared

	historical average			
	1 month	3 month	6 month	12 month
IC	0.1%	-3.6%	-3.2%	3.4%
	(0.07)	(0.89)	(0.74)	(0.17)
ICL	-1.7%	-2.3%	0.3%	11.2%
	(0.66)	(0.45)	(0.30)	(0.10)
ICH	3.6%	-17.5%	-13.1%	-1.0%
	(0.00)	(1.00)	(1.00)	(0.84)

Notes: This Table reports the out-of-sample predictive ability of forecasts of S&P500 index returns using a set of predictors, namely implied correlation IC, its low-frequency component ICL and its high-frequency component ICH. Implied correlation is computed via the Driessen et al. (2012) methodology using options with maturities of 30 days. Each forecasting model is based on recursive monthly regressions of market returns against each lagged predictor separately. Forecasts are produced separately for horizons of one, three, six and twelve months. The Table reports the out-of-sample R-squared (R^2_{OS}), as well as the associated Clark and West (2007) p-values (in brackets). The expanding estimation period starts at 36 months (January 1996 to December 1998), and the out-of-sample evaluation period runs from January 1999 to December 2020.

In contrast, the two components of implied correlation are found to be more efficient predictors of market returns compared to the original series at specific horizons that match the nature of their informational content. More specifically, the high-frequency component produces a positive and highly significant R^2_{OS} (3.6% with a (Clark and West, 2007) p-value of 0.00) at the shortest 1-month horizon. Given that the out-of-sample R-squared of IC at the 1-month horizon was substantially lower (0.1%) and only marginally significant, while that of ICL was negative (-1.7%), it appears that the predictability of short-term market returns can be significantly improved by extracting the information about short-term fluctuations contained in the high-frequency component of implied correlation and discarding the information about longer-term trends contained in the low-frequency component.

As I move towards longer forecasting horizons, the performance of ICH worsens, as evidenced by negative values for the R^2_{OS}. At the same time, the performance of

the low-frequency component begins to improve, producing positive out-of-sample R-squared at the 6-month and 12-month horizons (0.3% and 11.2%, respectively). However, even though the R^2_{OS} of the low-frequency component increases with the forecasting horizon, the (Clark and West, 2007) test generally fails to reject the null that the respective MSFE is lower than that of the benchmark model. For instance, the R^2_{OS} of ICL takes its highest value of 11.2% at the longest 12-month horizon, but the (Clark and West, 2007) p-value of 0.10 suggests that this outperformance of the historical mean benchmark is only marginally statistically significant. [11]

Overall, the out-of-sample evaluation provides additional support for the notion that disaggregating the information contained in implied correlation can lead to a significant improvement in the predictability of market returns.

2.7 Conclusion

A growing literature has been emerging on the informational content of option-implied correlation, especially its ability to predict market returns at various horizons accurately. In this chapter, I explore the source of this informational content. In particular, I show that the predictive ability of implied correlation is driven by the interaction between its low-frequency and high-frequency components.

Theoretically, a frequency decomposition can disaggregate implied correlation into information that refers to different time scales. The high-frequency component is expected to capture short-lived effects on the time series of implied correlation. In contrast, the low-frequency component is likely to pick up more "permanent" fluctuations. Empirically, I provide strong evidence that the high-frequency component

[11] I have also evaluated the out-of-sample forecasting performance of IC and its two components relative to one another by using each predictor in turn as the benchmark model (the full results are unreported for brevity, but they are available upon request). Consistent with my earlier findings, the ICH-based forecast significantly outperforms the IC-based forecast ($R^2_{OS} = 3.5\%$) and the ICL-based forecast ($R^2_{OS} = 5.3\%$) at the shortest 1-month horizon. Moreover, ICL significantly outperforms IC ($R^2_{OS} = 8.1\%$) and ICH ($R^2_{OS} = 12.1\%$) at the longest 12-month horizon.

drives the predictive ability of implied correlation at short horizons and the low-frequency component at longer horizons. More importantly, decomposing implied correlation leads to substantial improvements in the predictability of market returns, especially at shorter horizons where forecasting market returns tends to be more challenging. This improvement in market returns predictability is robust to controlling for a wide set of option-implied and fundamental variables, and it is also documented out-of-sample.

My proposed decomposition of implied correlation has several potential applications, and it can be used to motivate future research. An obvious practical application is in forecasting market variables such as index returns, where the specific forecasting horizon could inform the selection of a particular component in order to reduce noise and improve forecasting accuracy. Another application is in optimal portfolio selection in the presence of correlation risk, with the source of which potentially depending on the specific investment horizon. Another interesting direction for future research is to explore whether each component is associated with a different premium in the cross-section of stock and option returns.

Chapter 3

Volatility Forecasting with GARCH/GAS Models

3.1 Introduction

3.1.1 Observation-driven Models

Volatility modelling and forecasting play a critical role in the finance literature. As a measure for quantifying market risk, volatility is very important in practice. The establishment of financial policies strongly relies on the expectations of volatility in the future. Volatility is also helpful in measuring the sentiment of investors. Furthermore, accurate prediction of volatility leads to improved portfolio construction. Therefore, volatility modelling and forecasting have been an ongoing topic for decades. It is always beneficial to consider and develop more efficient models for volatility.

There are plenty of studies in the literature on volatility forecasting. In terms of discrete volatility models, one of the most influential families of models is the Autore-

gressive Conditional Heteroskedasticity (ARCH) models proposed by Engle (1982) and its generalised version Generalised Autoregressive Conditional Heteroskedasticity (GARCH) submitted by Bollerslev (1986). The GARCH models are advantageous partly in their ability to capture the time-varying property of volatility. However, standard GARCH models have limited capability in characterising the negative skewness and excess kurtosis that are commonly observed in financial returns time series. This motivates the explosive growth of studies in extensions of GARCH models. One crucial extension lies in the distribution of asset returns. According to Bollerslev (1987), distributions with heavier tails, such as the Student's t distribution, can accommodate the excess kurtosis. Another candidate with a similar feature is the Generalised Error Distribution (GED) proposed by Nelson (1991). However, GARCH models specified with these distributions still need to be revised to characterise the negative skewness simultaneously (Nelson, 1991). A solution to this is to introduce asymmetry. This can be realised by either incorporating the "leverage effects" as in Glosten et al. (1993) or simply assuming that the error term follows some skewed distribution, such as the Inverse Gaussian distribution as in Christoffersen et al. (2006).

Volatility forecasting using time-varying parameters models has attracted more and more attention in past years. These models are advantageous because of their superiority in characterising series dynamics. The drivers of time-varying parameters in these models, by Cox et al. (1981), can be either parameters themselves (parameter driven), for example, the state space models, see Durbin and Koopman (2012); Koopman et al. (2000); Saavedra et al. (2020), or empirical observations (observation driven).

In this chapter, I focus on comparing the performance of trending observation-driven models in the application of forecasting volatility of financial time series. The ARCH models and its generalisation GARCH models are effective models for volatility modelling and forecasting. The GARCH models are advantageous partly due to

their ability to capture the time-varying property of volatility. Another advantage of the GARCH model is that it is flexible and extendable. Ever since the GARCH model was proposed, more than one hundred variations that take into consideration different features of observations have been developed. However, the GARCH model is one of many candidates with these appealing characteristics. In recent years, another time-varying parameter model in which the score of the predicted density is adopted as the driver of parameter updating has been becoming increasingly popular. Proposed by Creal et al. (2012) and Harvey (2013), this model is referred to as the Generalised Autoregressive Score (GAS) model, Dynamic Conditional Score (DCS) model, or simply score-driven model. It can be shown that, under specific conditions, the GAS model reduces to the well-established GARCH model. In other words, the GARCH models are nested and are special cases of score-driven models.

Theoretically, there are a few main advantages to applying score-driven models in volatility modelling. First, according to Creal et al. (2012) and Harvey (2013), existing models, such as the stochastic volatility models in Shephard (2005), are not capable of taking into consideration the shape of predicted conditional densities of input data. The GAS models were then developed to bridge this gap. Second, since the updating mechanism of GAS models is driven by conditional scores, the estimation of time-varying parameters can be implemented directly using Maximum Likelihood Estimation (MLE). Third, the GAS models are more flexible in that one can make different assumptions on the distribution of innovation terms. Although the same can be done on GARCH models, the superiority of GAS models is expected to become more significant when it comes to non-Gaussian scenarios.

3.1.2 Model-free volatility forecasting

Option-implied information has been widely discussed. One of the most popular applications of option-implied information is volatility forecasting, where volatility

predictions computed using option prices are often compared to forecasts from historical underlying prices. Christensen and Prabhala (1998), Fleming (1998), and Ederington and Guan (2002) provide empirical evidence that option-implied volatility is a more accurate forecast compared to volatility computed from historical asset prices. Furthermore, Andersen (2000) and Andersen et al. (2003) found that option-implied volatility outperforms high-frequency realised volatility in forecasting future realised volatility. Jiang and Tian (2005), Giot and Laurent (2007) and Blair et al. (2010) verified this empirically for the US stock market.

Regarding the methods for extracting option implied volatility, it was initially obtained by backing up from the Black-Scholes Model using At-the-Money (ATM) options, provided other inputs are given. The most important innovation comes from the theoretical result of Carr and Madan, Demeterfi et al. (1999b) and Britten-Jones and Neuberger (2000), who obtain implied volatility by exploiting option prices only, without relying on any model. Such an approach is thus not restricted by unrealistic assumptions imposed in models like the Black-Scholes Model, and the implied volatility estimated using this method is called the Model-Free Implied Volatility (MFIV). This model-free method was applied by the CBOE to develop the VIX. Later approaches for computing model-free implied volatility were developed based on this starting point. An extensive literature review in implied volatility was conducted by Poon and Granger (2003). They found that, among 34 studies, 26 indicate that option implied volatility outperforms historical averages, and 17 out of 18 studies indicate superior forecasting performance of option implied volatility compared to GARCH model forecasts.

Here I list a few representative studies in volatility forecasting. Liang et al. (2020) construct Realised Volatility (RV) using intra-day high-frequency data from eight stock markets over the world and document that Implied Volatility (IV) is more powerful in predicting future realised volatility. It is also reported that IV significantly improves the forecasting accuracy of RV through horizons from 1 day to one

month. This finding is also confirmed in Kambouroudis et al. (2016). Another study by Martin et al. (2021) extends the framework of the realised GARCH model by introducing VIX, in addition to RV, into the model system. A robust increment in prediction accuracy is documented. Kourtis et al. (2016) adopts a sample containing indices data across ten countries and concludes that implied volatility is preferred for forecasting at the monthly horizon. At the daily horizon, it is the HAR model that dominates. In all circumstances, GARCH models are less powerful in volatility forecasting.

An interesting study by Becker et al. (2007) finds that, as the number of parametric models adopted increases, no incremental information can be detected in implied volatility proxied by VIX. This means that parametric models completely anticipate changes in volatility, and introducing implied volatility does not improve anticipation accuracy. Following this study, Becker and Clements (2008) propose to apply a combination of model-based predictors, confirming the combination model's dominance. The superiority of model-based predictors with respect to implied volatility is once again verified. Focusing on index option implied volatility, in addition to making comparisons between parametric models in volatility modelling, another purpose of this study is to inspect whether implied volatility conveys additional information than model-based volatility predictors.

The most important market of interest for volatility forecasting is the equity market. Many studies have been done in this field, and most agree that option-implied volatility provides valuable information for predicting future volatility in the stock market. Harvey and Whaley (1992) showed that implied volatility helps produce accurate forecasts. However, no arbitrage can be locked on, which is consistent with the market efficiency hypothesis. Canina and Figlewski (1993) document that implied volatility cannot predict future volatility by applying the binomial tree model and taking dividend adjustment and the possibility of early exercise into consideration. Fleming et al. (1995) included an earlier version of VIX from Whaley (1993)

into their study and found that although implied volatility yields biased forecasts, it is still helpful to some extent. Christensen and Prabhala (1998) showed that volatility implied by the Black-Scholes model outperforms historical volatility, Fleming (1998) used a modified binomial tree model to reveal that forecasts made by implied volatility are upward biased. Blair et al. (2010) utilised the VIX method to compute implied volatility and showed that such implied volatility contains all information. Thus, no incremental information can be obtained from intra-day high-frequency data.

Jiang and Tian (2005) made some improvements based on Britten-Jones and Neuberger (2000) and found that model-free implied volatility subsumes the information in the BS implied volatility. Ang et al. (2006) attempted to predict the cross-section of stock returns using implied volatility calculated using the VIX approach and found that shocks in VIX are a priced factor with the negative price of risk.

All these papers set implied volatility as the predictor to predict future volatility. Some other studies, such as Bollerslev et al. (2009), Bekaert et al. (2013), Zhou (2018) used the Variance Risk Premium (VPR), which is extracted as the difference between option implied variance (under $\mathbb{Q}$ measure) and realised variance (under $\mathbb{P}$ measure), to predict the equity risk premium. They found that the VRP is useful in predicting equity risk premiums.

Rather than index options implied volatility Chiras and Manaster (1978), Beckers (1981), and Lamoureux and Lastrapes (1993) focus on the implied volatility of individual stocks, almost all of their studies found that implied volatility better predicts future volatility. Swidler and Wilcox (2002) found that individual implied volatility outperforms historical models in predicting the volatility of bank stocks. More generally, Banerjee et al. (2007) document the predictive power of implied volatility in forecasting returns of characteristic-based portfolios. Diavatopoulos et al. (2008) detected a strong and positive link between idiosyncratic volatility and

future cross-sectional stock returns.

3.1.3 The distribution of financial returns

Ever since it was proposed by Bollerslev (1986), the error term in the GARCH model is assumed to be normally distributed. However, plenty of empirical studies indicate that it is not necessary that financial returns are normally distributed. Instead, asset returns are often skewed and tend to have heavy tails and excess kurtosis. It is beneficial to take such behaviour into account in the process of constructing volatility models. Under the framework of GARCH models, a few studies have focused on making different non-Gaussian assumptions on the distribution of error terms. For example, Bollerslev (1987) document that distributions with heavy tails, such as Student's t distribution, are useful in capturing excess kurtosis observed in financial returns. Another heavy-tailed distribution is adopted in Nelson (1991), known as the Generalised Error Distribution (GED).

However, the fatter-tailed distributions mentioned above are not capable of accounting for skewness. According to Nelson (1991) and Glosten et al. (1993), such asymmetry can be well explained by models that consider the leverage effect. The GJR-GARCH model proposed by Glosten et al. (1993) has a structure such that the volatility responds differently to positive shocks and negative shocks, and hence the leverage effect is incorporated. Another way to introduce skewness into modelling volatility is to assume that the innovation term follows skewed distributions, such as normal inverse Gaussian distribution; see, for example, Forsberg and Bollerslev (2002) and Stentoft (2006), inverse Gaussian distribution Christoffersen et al. (2006), shifted Gamma distribution Tong et al. (2004), generalised error distribution Duan (1999), z-distribution Lanne (2003) and Lanne and Pentti (2007), and α-stable distribution Menn and Rachev (2005). Apart from these univariate distributions, mixture distributions have been becoming increasingly popular in practice. For in-

stance, Badescu et al. (2008) propose an error term that follows a mixture normal distribution. Motivated by the paradigm of making assumptions on the distribution of the error term, in this study, I attempt to construct a series of score-driven models with different error terms.

3.1.4 Contributions

Quite a few studies that investigate volatility modelling and forecasting using financial returns are within the framework of GARCH models. This study is closely connected to this strand of the literature. While variations of GARCH models are being proposed constantly, this chapter proposes to characterise and predict realised volatility within the framework of score-driven models or the GAS models. GAS models are empirically found to have more reasonable responses to abnormal shocks in observations, and they are expected to be more suitable frameworks for modelling financial returns that are distributed in a fat-tailed fashion.

This chapter contributes to the existing literature on volatility forecasting by constructing a series of GARCH and GAS models with differently distributed innovation terms, and by implementing comprehensive comparisons between the performance of GARCH and GAS models in volatility forecasting.

The most prominent finding of this study is that GARCH models tend to have better performance in volatility forecasting than GAS models in most cases. This is indicated by both higher descriptive R^2 from regressing realised volatility onto predictions and lower forecasting errors. It is also documented that the TLS-GAS models have the overall best performance in forecasting realised volatility at the daily horizon. Taking a general comparison between GARCH and GAS, it can be concluded that GARCH outperforms GAS at daily and weekly horizons and that the outperformance becomes weaker at the monthly horizon.

Using a sub-sample that covers the 2007-2008 financial crisis, it can be observed that, while the GARCH models still have higher descriptive R^2 metrics in most cases and across all forecasting horizons, the GAS models tend to have lower forecasting errors in most cases (5 out of 8) at the daily horizon. At weekly and monthly horizons, the previously reported higher accuracy of GARCH models seems to be weakened.

It can be concluded that when the market is experiencing significant uncertainties, GAS models with reasonably selected innovation terms might work better in financial returns modelling, as GAS models are theoretically expected to have lower forecasting errors due to their less fierce reaction to large changes in observations. This is a very useful property, especially in modelling and forecasting volatility of financial returns during market turmoils, as it is likely that the large value is observed due to the heavy-tailed nature of the financial returns data set instead of actual increases in volatility.

Following Becker et al. (2007) and Becker and Clements (2008), the last target of this study is to extensively investigate whether observation-driven models, represented by GARCH and GAS models, provide more information about volatility than option prices. To this point, I document that implied volatility consistently outperforms model-based volatility predictions in terms of explaining realised volatility. However, regarding prediction accuracy, both IV and MFIV are found to have inferior performance than GARCH/GAS-based volatility forecasts.

The rest parts of this chapter are organised as follows: Section 3.2 introduces and derives GARCH models in which the innovation terms are assumed to follow Gaussian distributions and a variety of non-Gaussian distributions. The corresponding GAS models are derived and illustrated. In this section, I also introduce the Implied Volatility (IV) and the Model-Free Implied Volatility (MFIV) and the approach for obtaining them. Section 3.3 introduces all data sets used in this chapter. In Section 3.5, I interpret and analyse predicted volatility from both GARCH and

GAS models. Then a series of statistical tests are conducted to identify superiority and check whether shifting from GARCH to GAS improves forecast accuracy. The comparisons between GARCH and GAS, as well as the comparisons between volatility from observation-driven models and model-free implied volatility, are all implemented over different horizons from one day up to one month. I take a further step into inspecting forecasting performance by testing on sub-sample periods of market turmoil. Section 3.6 concludes.

3.2 Model-based volatility

In this section, I first introduce the framework of the widely used GARCH model. This is followed by an illustration of the mechanism of the GAS model. Based on various assumptions on the distribution of innovation terms, I derive corresponding GARCH and GAS models. More specifically, the error terms are assumed to follow a normal distribution, Shifted Gamma Distribution (SG), Shifted Negative Gamma Distribution (SNG), Shifted Inverse Gaussian Distribution (SIG), Shifted Negative Inverse Gaussian Distribution (SNIG), Shifted Weibull Distribution (SW), Student's t Location and Scale Distribution (TLS), and Logistic Distribution (LOG). I also show that, under some conditions, the GAS model can be converted into the GARCH model.

3.2.1 GARCH Models

As the most representative time-varying parameter model, the GARCH model and its variants are frequently applied in volatility modelling. Prior to GARCH, the ARCH model was first proposed by Engle (1982) for the purpose of predicting conditional variances. The conditional variance in an ARCH(p) model depends on p

order-lagged squared errors. The model can be written down as

$$y_t = \mu_t + \xi_t \tag{3.1}$$

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \xi_{t-i}^2. \tag{3.2}$$

where y_t is the time series of variable of interest, μ_t is a mean value process, ξ is called the innovation term or the error term, h_t represents the conditional variance of ξ (hence y) that changes over time, ω stands for some constant term, and α_i are corresponding coefficients of squared lagged terms of ξ, denoted by ξ_{t-i}.

The ARCH models are capable of capturing some features of volatility, such as volatility clustering (Bera and Higgins, 1993). However, it still has an obvious disadvantage. As widely known, volatility exhibits strong dependency even after long lags. As documented by Abdalla (2012), it can be challenging to estimate a large number of ARCH parameters due to a large number of lagged terms. One significant advantage of the GARCH model is that it avoids the long-lag structure of ARCH models. The GARCH model can be described by the equation below

$$y_t = \mu_t + \xi_t \tag{3.3}$$

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \xi_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}. \tag{3.4}$$

where p and q are number of lags applied with respect to ξ_t and h_t, and β_j are coefficients of lagged variance terms h_{t-j}. The rest of the variables are defined identically as in Equation (3.3) above.

In the application of option pricing, a critical property of the GARCH model is that it is non-Markovian (Duan, 1995). More specifically, the GARCH volatility (hence the GARCH option prices) depends on the information set at present and in the past. This forms the foundation of dynamically making one step ahead volatility

forecast. The GARCH model, when first proposed, was still inadequate to capture some commonly seen features in financial returns time series, such as the asymmetric responses to positive and negative shocks from these returns. Plenty of GARCH extensions have been developed to account for such an effect. Another problematic part of the GARCH model is that the innovation term ξ is not necessarily normally distributed. In fact, returns series tend to be leptokurtic. According to Bollerslev (1987) and Baillie and Bollerslev (1989), the normality assumption fails to accommodate such behaviour of financial asset returns. Tong et al. (2004), within the GARCH framework, propose to assume that the innovation terms ξ follow a shifted Gamma distribution, as it has the ability to capture the skewness of stock returns.

3.2.2 GAS Models

As proposed by Creal et al. (2012), the Generalised Autoregressive Score (GAS) models are often referred to as Dynamic Conditional Score (DCS) models. They offer a framework for accommodating the time-varying property of parametric models. The idea is very straightforward. Denote the conditional density of observation y_t by $p(y_t|f_t)$, where f_t is some time-varying parameter that determines the distribution of y_t. f_t is assumed to update following:

$$f_{t+1} = \omega + \alpha S_t \nabla_t + \beta f_t \tag{3.5}$$

$$\nabla_t = \left[\frac{\partial \ln p\left(y_t \mid f_t\right)}{\partial f_t} \right] \tag{3.6}$$

where ∇_t is the score of $\ln p\left(y_t \mid f_t\right)$. S_t is a function used for scaling the score of the log observation density. The model is innovative in that it uses the scaled score to drive the time-varying parameter f_t. In other words, it connects the dynamics of f_t to all the parameters of the conditional density $p(y_t|f_t)$. It is easy to see that in the GARCH models, the shape of shock density does not matter. It is worth mentioning

that based on the model above, additional lag structures and other dynamics can be easily added to the transition equation for f_{t+1}.

The scaling matrix S_t is usually denoted as:

$$S_t = \mathcal{I}_t^{-d} \tag{3.7}$$

where $\mathcal{I}_t$ is the information matrix of the observation's log density, and $d \in \{0, \frac{1}{2}, 1\}$ is referred to as the scaling coefficient. If $d = 1$, the scaling matrix is equal to the variance of the score of $\ln p(y_t \mid f_t)$ (Creal et al., 2012)[1].

$$S_t = \mathcal{I}_t^{-1} = -E_t \left[\nabla_t \nabla_t^T\right]^{-1} \tag{3.8}$$

In practice, f_t is often re-parameterised in order to guarantee positiveness during the parameter estimation process. In this chapter, I apply the logarithmic link function onto f_t, such that $f_t' = \ln f_t$. The GAS updating process of the re-parameterised time-varying parameters f_t' is then

$$f_{t+1}' = \omega + \alpha S_t' \nabla_t' + \beta f_t' \tag{3.9}$$

Denote the function that maps f_t to f_t' by $g(\cdot)$, its Jacobian is

$$J_t = \frac{\partial g(f_t)}{\partial (f_t)} = \frac{1}{f_t} \tag{3.10}$$

Following Creal et al. (2012), if the scaling coefficient $d = 1$ and the logarithmic link function is applied for re-parameterisation of f_t, it can be shown that the linked

[1] In my study, $\mathcal{I}_t$ is calculated by taking the second-order derivative of $\ln p(y_t \mid f_t)$ with respect to f_t, as they are equivalent given that the log-density is twice differentiable.

scaled score $S_t'\nabla_t'$ in (3.9) is equal to:

$$S_t'\nabla_t' = J_t\mathcal{I}_t^{-1}\nabla_t \tag{3.11}$$

In this chapter, (3.11) is utilised across all GAS models to calculate the linked scaled scores, which are components of the updating equation (3.9).

In terms of option pricing using GAS models, the working paper by Lin and Yu (2017) made a very beneficial attempt. It is also the only study I could find in the literature of this field so far. In their paper, a GAS model with generalised hyperbolic innovations is proposed. Such innovation terms are capable of dealing with features of financial returns like conditional skewness and fat tails. As a result, the so-called "volatility smile" is implied by the model. It is worth mentioning that an additional leverage effect is incorporated into this model as well. With the GAS representation of returns dynamics under $\mathbb{P}$, the model can be estimated by Maximum Likelihood Estimation (MLE). This is followed by the change of measure, usually implemented by methods like the Esscher Transform is applied.

Before I proceed to the construction of GARCH and GAS models, I introduce the process used to describe stock price evolution and the corresponding mean equation μ_t I adopt in my study. Recall that the process of the variable of interest can be expressed by (3.3), where μ_t is some mean process and (3.3) is called the mean equation of the GARCH model. μ_t can have different specifications, such as ARMA structure. In this study, I stick to another common choice of μ_t used in Duan (1995), the GARCH-in-Mean (GIM) model.

To start, suppose that the evolution of asset price can be described by the stochastic differential equation below:

$$S_t = S_{t-1}\exp\left(r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t\right) \tag{3.12}$$

where r is the risk-free rate, λ is a constant coefficient, and usually is a proxy of the unit price of risk.

Taking logs on both sides:

$$\ln\left(\frac{S_t}{S_{t-1}}\right) = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t \tag{3.13}$$

The left-hand side is the logarithmic returns of an asset. Denote it by y_t; the mean equation is finally obtained as:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t \tag{3.14}$$

Equation (3.14) hereafter is applied as the mean equation for both GARCH and GAS models to be derived in my study, while their updating equations are pretty differentiated.

3.2.3 Normal innovation

I depart from the simplest case, where the innovation term ξ_t is normally distributed over time. Note that although $\xi_t \sim N(0, h_t)$ by assumption, according to (3.14), the mean value of stock returns y_t is still time-varying. The conditional density function of stock returns can be written down as follows:

$$f_{\xi_t}(y) = \frac{1}{\sqrt{2\pi h_t}} e^{\left(\frac{-y^2}{2h_t}\right)} \tag{3.15}$$

As a consequence, the stock returns have a mean of $r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t$ and variance of h_t. It is straightforward to calculate the log density of ξ_t as:

$$\ln f_{\xi_t}(y) = -\frac{1}{2}\ln 2\pi h_t - \frac{y^2}{2h_t} \tag{3.16}$$

The log density is needed in parameter estimation via maximum likelihood estimation (MLE). Conveniently, h_t is already part of the updating equation. This means that when h_t is updated each time, the value of log density is updated accordingly. As a result, optimisations for parameter estimation are straightforward. Another application of log density lies in the construction of the GAS model. Recall that in the GAS model, instead of updating variance h_t as in a GARCH system, some factor is specified to be time-varying. The variance in GAS models can be linked to the factors by the property of the assumed distribution of error terms. When ξ_t is normally distributed, the factor that drives the updating of the GAS system is naturally h_t. One beneficial part of the GAS model is its capability to apply link functions to its time-varying factors. It is common sense that volatility (or variance) cannot be negative. However, within the parameter estimation process, the positiveness of h_t may not be guaranteed. Link functions under such circumstances can be used to fix this. In my study, logarithmic link functions are applied for all GAS models to ensure the positiveness of variance in each iteration of the optimisation process. More specifically, I assume that $f_t' = g(f_t) = \ln h_t$. Clearly, $g(\cdot)$ is a mapping from h_t to f_t', and the latter is the factor finally adopted in the updating equation. Calculating the partial difference of $g(\cdot)$ with respect to h_t, the Jacobian of link function $g(\cdot)$ is:

$$\begin{aligned} J\left(h_t\right) &= \frac{\partial g\left(h_t\right)}{\partial h_t} \\ &= \frac{\partial \ln h_t}{\partial h_t} \\ &= \frac{1}{h_t} \end{aligned} \tag{3.17}$$

From log-density function (3.16), the score of the normal distribution can be calculated as:

$$\begin{aligned} \nabla_t &= \frac{\partial \ln f_{\xi_t}(y)}{\partial h_t} \\ &= \frac{y^2}{2h_t^2} - \frac{1}{2h_t} \end{aligned} \tag{3.18}$$

The information matrix can be computed as:

$$\begin{aligned} I_t^{-1} &= -E_t \left[\nabla_t \nabla_t^T \right]^{-1} \\ &= 2h_t^2 \end{aligned} \tag{3.19}$$

With (3.17), (3.18) and (3.19), by equation (3.11), the linked scaled score function can be derived as:

$$S_t' \nabla_t' = J_t I_t^{-1} \nabla_t \tag{3.20}$$

Until now, the components for constructing a GAS system have been obtained. The GAS (1,1) model with Gaussian innovation (GAS) in my study can be written down as follows:

$$y_t = r + \lambda \sqrt{h_t} - \frac{1}{2} h_t + \xi_t, \quad \xi_t \sim N(0, h_t) \tag{3.21}$$

$$f_t' = \ln h_t \tag{3.22}$$

$$f_{t+1}' = \omega + A S_t' \nabla_t' + B f_t' \tag{3.23}$$

It is easy to see that the corresponding GARCH (1,1) model with Gaussian innovation term (GARCH) can be described by:

$$y_t = r + \lambda \sqrt{h_t} - \frac{1}{2} h_t + \xi_t, \quad \xi_t \sim N(0, h_t) \tag{3.24}$$

$$h_{t+1} = \omega + A \xi_t^2 + B h_t \tag{3.25}$$

It is noteworthy that, in terms of the GAS model above, if the link function is not applied, it is straightforward that $f_t = h_t$ and hence $J_t = 1$. The scaled score function is reduced to:

$$\begin{aligned} S_t &= I_t^{-1} \nabla_t \\ &= \left(\frac{y^2}{2h_t^2} - \frac{1}{2h_t} \right) 2h_t^2 \\ &= \xi_t^2 - h_t \end{aligned} \tag{3.26}$$

While the mean equation stays unchanged, equation (3.25) reduces to:

$$\begin{aligned} h_{t+1} &= \omega + A\left(y^2 - h_t\right) + Bh_t \\ &= \omega + A\xi_t^2 + (B-1)\, h_t \end{aligned} \tag{3.27}$$

Obviously, when the innovation term is normally distributed, the GARCH model is nested in the GAS model, and the latter converts to the former if no link function is applied.

3.2.4 Shifted Gamma innovation

The skewed behaviour of stock returns can be captured by the GARCH model with shifted Gamma innovation (Tong et al., 2004). It is natural to consider how the shifted Gamma innovation works in the framework of the GAS model. It is noteworthy that for non-Gaussian distributions like the Gamma distribution, its support lies within $y_t \in (0, \infty)$. This makes it impossible for us to model stock returns using Gamma innovation directly, as returns can be negative. Following Tong et al. (2004) and Zhu and Ling (2015), a shifted Gamma innovation is adopted in modelling instead. The subsequent application of the GAS model with shifted Gamma innovation (SG-GAS) and the corresponding GARCH model (SG-GARCH) follows a similar procedure.

Suppose that $Y_t \sim G(a, b_0)$, where a and b_0 stand for the shape parameter and scale parameter of the Gamma distribution. The probability density function of Y_t is:

$$f_{Y_t}(y) = \frac{1}{\Gamma(a) b_0^a} y^{a-1} e^{-\frac{y}{b_0}} \tag{3.28}$$

From the property of the Gamma distribution, the mean and variance of Y_t are given by:

$$\mu_Y = ab_0 \tag{3.29}$$

and

$$h_Y = ab_0^2 \tag{3.30}$$

Note that in the SG-GAS model, the scale parameter b_0 is set to be a time-varying factor, and thus it is the driver of the updating equation. For simplicity, I have omitted subscripts t. Now let

$$\xi_t = \sqrt{h_t}\left(\frac{Y_t - ab_0}{\sqrt{ab_0^2}}\right) \tag{3.31}$$

Clearly, the shifted innovation process ξ_t has zero mean and variance h_t. Substitute ξ_t back into (3.14):

$$\begin{aligned} y_t &= r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t \\ &= r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t} + \frac{1}{b_0}\sqrt{\frac{h_t}{a}}Y_t \end{aligned} \tag{3.32}$$

By the scaling property of Gamma distribution, $\frac{1}{b_0}\sqrt{\frac{h_t}{a}}Y_t$ still follows a Gamma distribution. The shape parameter a stays unchanged while the scale parameter changes to $\sqrt{\frac{h_t}{a}}$. Thus define $\frac{1}{b}\sqrt{\frac{h_t}{a}}Y_t = \xi_t^* \sim G(a, b_t)$.

From (3.28), the log-density of the shifted innovation term ξ_t^* is given by:

$$\ln f_{\xi_t^*}(y) = (a-1)\ln y - \frac{y}{b_t} - \ln\Gamma(a) - a\ln b_t \tag{3.33}$$

Hence, by setting the scale parameter b as the driver of updating the equation, the score function can be calculated as below:

$$\begin{aligned} \nabla_t &= \frac{\partial \ln f_{\xi_t^*}(y)}{\partial b_t} \\ &= \frac{y}{b_t^2} - \frac{a}{b_t} \end{aligned} \tag{3.34}$$

The logarithmic link function is once more applied so that the driver factor f_t is:

$$f'_t = \ln b_t \tag{3.35}$$

Hence the Jacobian can be computed as:

$$\begin{aligned} J_t(b_t) &= \frac{\partial g(b_t)}{\partial b_t} \\ &= \frac{\partial \ln b_t}{\partial b_t} \\ &= \frac{1}{b_t} \end{aligned} \tag{3.36}$$

The scaling matrix is computed as follows:

$$\begin{aligned} I_t^{-1} &= -E_t[\nabla_t \nabla_t^T]^{-1} \\ &= b_t^2/a \end{aligned} \tag{3.37}$$

With the linked scaled score $S'_t \nabla'_t = J_t I_t^{-1} \nabla_t$, the SG-GAS model is constructed as:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t} + \xi_t^*, \quad \xi_t^* \sim G(a, b_t) \tag{3.38}$$

$$f'_t = \ln b_t \tag{3.39}$$

$$f'_{t+1} = \omega + AS'_t \nabla'_t + Bf'_t \tag{3.40}$$

The corresponding SG-GARCH model can be described by:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t} + \xi_t^*, \quad \xi_t^* \sim G(a, b_t) \tag{3.41}$$

$$\epsilon_t = -\sqrt{ah_t} + \xi_t^* \tag{3.42}$$

$$h_{t+1} = \omega + A\epsilon_t^2 + Bh_t \tag{3.43}$$

Note that while the mean value of ξ_t^* is $ab > 0$, ϵ_t has zero means, and thus it is reasonable to be used in building the GARCH model.

The GAS model with shifted negative Gamma innovation (SNG-GAS) is constructed following an identical procedure. The only difference is that ξ_t^* in SNG-GAS is assumed to follow a negative Gamma distribution. In other words, $-\xi_t^* \sim G(a, b_t)$. The specification of the SNG-GAS model is the same as SG-GAS. The construction of SNG-GARCH and SG-GARCH follows identical procedures.

3.2.5 Shifted Inverse Gaussian innovation

Christoffersen et al. (2006) propose a GARCH model with inverse Gaussian innovation to characterise index returns and document the outperformance of the IG-GARCH model in pricing out-of-the-money put options. Inspired by this study, I construct the GAS model with shifted inverse Gaussian innovation (SIG-GAS) and shifted negative inverse Gaussian innovation (SNIG-GAS). Like the Gamma distribution, the support of inverse Gaussian distribution is $y_t \in (0, \infty)$. This makes it impossible for us to model stock returns using Gamma innovation directly, as returns can be negative. Following an identical method as in the SG-GAS model above, let $Y_t \sim IG(\mu_0, k_0)$, where $IG(\mu_0, k_0)$ stands for inverse Gaussian distribution with a mean of μ_0 and shape parameter k_0. The probability density function of Y_t is:

$$f_{Y_t}(y) = \sqrt{\frac{k_0}{2\pi y^3}} \exp\left(-\frac{k_0(y-\mu_0)^2}{2\mu^2 y}\right) \tag{3.44}$$

By the property of inverse Gaussian distribution, the variance of Y_t is:

$$h_Y = \frac{\mu_0^3}{k_0} \tag{3.45}$$

Different from the SG-GAS model, where the scale parameter is time-varying, in

the SIG-GAS model, the shape parameter k is set to be the driving factor. Let

$$\xi_t = \sqrt{h_t}\left(\frac{Y_t - \mu_0}{\sqrt{\frac{\mu_0^3}{k_0}}}\right) \tag{3.46}$$

Substitute ξ_t back into (3.14):

$$\begin{aligned} y_t &= r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t \\ &= r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{\frac{h_t k_0}{\mu_0}} + \sqrt{\frac{h_t k_0}{\mu_0^3}}Y_t \end{aligned} \tag{3.47}$$

Following the scaling property of inverse Gaussian distribution, $\sqrt{\frac{h_t k_0}{\mu_0^3}}Y_t$ still follows an inverse Gaussian distribution with scaled parameters. Denote the updated innovation term by $\xi_t^* \sim IG(\mu, k_t)$.

The log-density of ξ_t^* is given by:

$$\ln f_{\xi_t^*}(y) = \frac{\ln k_t}{2} - \frac{\ln(2\pi\,(y)^3)}{2} - \frac{k_t\,(y-\mu)^2}{2\mu^2 y} \tag{3.48}$$

Calculating its derivative with respect to k_t, the score function is given below:

$$\begin{aligned} \nabla_t &= \frac{\partial \ln f_{\xi_t^*}(y)}{\partial k_t} \\ &= \frac{1}{2k_t} - \frac{(y-\mu)^2}{2\mu^2 y} \end{aligned} \tag{3.49}$$

The scaling matrix is obtained by taking the second-order derivative

$$\begin{aligned} I_t^{-1} &= -E_t[\nabla_t \nabla_t^T]^{-1} \\ &= 2k_t^2; \end{aligned} \tag{3.50}$$

Similar to the SG-GAS example above, it is easy to see the Jacobian of the SIG-GAS

model is equal to $\frac{1}{k_t}$. The SIG-GAS model is constructed as follows:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{\frac{h_t k_t}{\mu}} + \xi_t^*, \quad \xi_t^* \sim IG(\mu, k_t) \tag{3.51}$$

$$f_t' = \ln k_t \tag{3.52}$$

$$f_{t+1}' = \omega + AS_t'\nabla_t' + Bf_t' \tag{3.53}$$

The corresponding SIG-GARCH model can be constructed as below:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{\frac{h_t k_t}{\mu}} + \xi_t^*, \quad \xi_t^* \sim IG(\mu, k_t) \tag{3.54}$$

$$\epsilon_t = -\sqrt{\frac{h_t k_t}{\mu}} + \xi_t^* \tag{3.55}$$

$$h_{t+1} = \omega + A\epsilon_t^2 + Bh_t \tag{3.56}$$

To develop the GAS model with shifted negative inverse Gaussian innovation (SNIG-GAS) and corresponding SNIG-GARCH, simply assume that $-\xi_t^* \sim IG(\mu, k_t)$.

3.2.6 Shifted Weibull innovation

The Weibull distribution and its variations have been applied in finance for years. For example, Mittnik and Rachev (1993) document that Weibull distribution dominates other *stable* distributions in modelling asset returns. Chen and Gerlach (2013) find that a two-sided Weibull distribution performs more favourably in predicting conditional Value-at-Risk, and such superiority stays unchanged during the financial crisis. Silahli et al. (2019) utilises a two-sided Weibull distribution to estimate the Value-at-Risk of the cryptocurrency portfolio and find that the proposed model significantly outperforms the other competitors. Wang et al. (2018) incorporate two-sided Weibull distribution into the framework of realised GARCH models and

document the outperformance of the proposed model in volatility modelling and tail risk forecasting. Rachev and SenGupta (1993) adopt Laplace-Weibull mixtures in modelling price changes in real estate in France. In the framework of GAS models, Creal et al. (2012) includes Weibull distribution into an extensive Monte Carlo study and document that the GAS models, as member of observation-driven models, outperform the others in terms of prediction accuracy.

Another reason for which I consider Weibull distribution in this study is that one of the main purposes is to find out superior models for financial returns, which tend to be fat-tailed. To this point, the Weibull distribution is potentially a good candidate, as one of its appealing properties is that when the shape parameter k is between 0 and 1, it is a desirable heavy-tailed distribution for financial modelling.

Let $Y_t \sim W(\theta_0, k)$, where θ_0, and k are scale and shape parameters respectively. The probability density function of Y_t is:

$$f_{Y_t}(y) = \frac{k}{\theta_0}\left(\frac{y}{\theta_0}\right)^{k-1} e^{-(y/\theta_0)^k} \tag{3.57}$$

By the property of Weibull distribution, the mean and variance of Y_t are:

$$\mu_Y = \theta_0 \Gamma(1 + 1/k) \tag{3.58}$$

and

$$h_Y = \theta_0^2 \left[\Gamma\left(1 + \frac{2}{k}\right) - \left(\Gamma\left(1 + \frac{1}{k}\right)\right)^2\right] \tag{3.59}$$

Setting the scale parameter as the driving factor and following an identical process as above, the mean equation of stock returns y_t is given by:

$$\begin{aligned} y_t &= r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t \\ &= r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \frac{\sqrt{h_t}\Gamma(1 + 1/k)}{\sqrt{u}} + \frac{\sqrt{h_t}}{\theta_0\sqrt{u}}Y_t \end{aligned} \tag{3.60}$$

where $u = \Gamma\left(1+\frac{2}{k}\right) - \left(\Gamma\left(1+\frac{1}{k}\right)\right)^2$. Let $\xi_t^* = \frac{\sqrt{h_t}}{\theta\sqrt{u}} Y_t \sim W(\theta_t, k)$ denote the shifted Weibull innovation, its log-density is:

$$\ln f_{\xi_t^*}(y) = \ln k + (k-1)\ln y - k\ln\theta_t - \left(\frac{y}{\theta_t}\right)^k \tag{3.61}$$

The score function of ξ_t^* is computed as:

$$\begin{aligned} \nabla_t &= \frac{\partial \ln f_{\xi_t^*}(y)}{\partial \theta_t} \\ &= \frac{k}{\theta_t}\left(\left(\frac{y}{\theta_t}\right)^k - 1\right) \end{aligned} \tag{3.62}$$

The information matrix is computed as:

$$\begin{aligned} I_t^{-1} &= -E_t[\nabla_t \nabla_t^T]^{-1} \\ &= (\frac{\theta_t}{k})^2; \end{aligned} \tag{3.63}$$

Analogous to the SG and SIG cases, the Jacobian of the SW-GAS model is equal to $\frac{1}{\theta_t}$. By (3.11), the linked scaled score is obtained, and the SW-GAS model is constructed as follows:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \frac{\sqrt{h_t}\Gamma(1+1/k)}{\sqrt{u}} + \xi_t^*, \quad \xi_t^* \sim W(\theta_t, k) \tag{3.64}$$

$$u = \Gamma\left(1+\frac{2}{k}\right) - \left(\Gamma\left(1+\frac{1}{k}\right)\right)^2 \tag{3.65}$$

$$f_t' = \ln\theta_t \tag{3.66}$$

$$f_{t+1}' = \omega + AS_t'\nabla_t' + Bf_t' \tag{3.67}$$

The corresponding SW-GARCH model can be constructed as below:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \frac{\sqrt{h_t}\Gamma(1+1/k)}{\sqrt{u}} + \xi_t^*, \quad \xi_t^* \sim W(\theta_t, k) \tag{3.68}$$

$$\epsilon_t = -\frac{\sqrt{h_t}\Gamma(1+1/k)}{\sqrt{u}} + \xi_t^* \tag{3.69}$$

$$h_{t+1} = \omega + A\epsilon_t^2 + Bh_t \tag{3.70}$$

By assuming that $-\xi_t^* \sim W(\theta_t, k)$, SNW-GAS and SNW-GARCH models are constructed. In this study, I only adopt the GAS/GARCH model with SNW innovation, as it appears that models with SW innovation do not fit our data well and characterise volatility reasonably.

3.2.7 T-location and Scale innovation

Bollerslev (1987) adopts Student's t distribution in the construction of the GARCH model to account for excess kurtosis of financial returns. Since then, as one candidate of heavy-tailed distributions, Student's t distribution has been widely used in financial modelling. In my study, I apply t location-scale distribution to construct TLS-GAS and TLS-GARCH models. It is convenient that the support of t location-scale distribution is $y_t \in (-\infty, +\infty)$. Hence assumption can be made directly on ξ_t, and no more shifting operations are needed. The density function of the t location-scale distribution is:

$$f_{Y_t}(y) = \frac{1}{\sqrt{s\nu}B\left(\frac{1}{2}, \frac{\nu}{2}\right)}\left(1 + \frac{(y_t - \mu)^2}{s\nu}\right)^{\frac{-\nu+1}{2}} \tag{3.71}$$

where s and ν are the scale parameter and degree of freedom parameter, respectively, and $B(\cdot)$ is the Beta function. Let $\xi_t \sim T(\mu, s_t, \nu)$, fixing $\mu = 0$, the variance of ξ_t is equal to $s_t\frac{\nu}{\nu-2}$. Let the scale parameter s_t be the time-varying factor, the log-density of ξ_t is:

$$\ln f_{\xi_t}(y) = -\frac{1}{2}\ln \nu s_t - \ln B\left(\frac{1}{2}, \frac{\nu}{2}\right) - \left(\frac{\nu+1}{2}\right)\ln\left(1 + \frac{(y-\mu)^2}{s_t\nu}\right) \tag{3.72}$$

The score function and the information matrix for scaling of ξ_t are:

$$\begin{aligned} \nabla_t &= \frac{\partial \ln f_{\xi_t}(y)}{\partial s_t} \\ &= -\nu \frac{s_t - (y-\mu)^2}{2s_t\left(\nu s_t + (y-\mu)^2\right)} \end{aligned} \tag{3.73}$$

and

$$\begin{aligned} I_t^{-1} &= -E_t\left[\nabla_t \nabla_t^T\right]^{-1} \\ &= \frac{(2s_t^2(\nu+3))}{\nu} \end{aligned} \tag{3.74}$$

Same as above, the Jacobian and the linked scaled scores of TLS-GAS are $\frac{1}{s_t}$ and $S_t'\nabla_t' = JI_t^{-1}\nabla_t$. With these components, the TLS-GAS model and TLS-GARCH models can be written down as follows:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t \sim T(\mu, s_t, \nu) \tag{3.75}$$

$$f_t' = \ln s_t \tag{3.76}$$

$$f_{t+1}' = \omega + AS_t'\nabla_t' + Bf_t' \tag{3.77}$$

and

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t \sim T(\mu, s_t, \nu) \tag{3.78}$$

$$\epsilon_t = \xi_t \tag{3.79}$$

$$h_{t+1} = \omega + A\epsilon_t^2 + Bh_t \tag{3.80}$$

3.2.8 Logistic innovation

Another member of heavy-tailed distributions is the logistic distribution. Assuming logistic innovations, I create LOG-GAS and LOG-GARCH models. The probability

density function of logistic distribution is:

$$f_{Y_t}(y) = \frac{e^{-(y-\mu)/s}}{s\left(1+e^{-(y-\mu)/s}\right)^2} \tag{3.81}$$

where s and μ are scale parameter location parameters, respectively. Let $\xi_t \sim L(\mu, s_t)$, assume that $\mu = 0$, the variance of ξ_t is $\frac{s_t^2\pi^2}{3}$. Setting the scaling coefficient $d = 0$, the log density of ξ_t and its linked score function are given by:

$$\ln f_{\xi_t}(y) = -\frac{y}{s_t} - \ln s_t - 2\ln(1+e^{-\frac{y}{s_t}}) \tag{3.82}$$

and

$$S_t'\nabla_t' = \frac{y}{s_t} - 1 - \frac{2ye^{-\frac{y}{s_t}}}{s_t(1+e^{-\frac{y}{s_t}})} \tag{3.83}$$

The LOG-GAS and corresponding LOG-GARCH models are constructed as follows:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t \sim L(\mu, s_t) \tag{3.84}$$

$$f_t' = \ln s_t \tag{3.85}$$

$$f_{t+1}' = \omega + AS_t'\nabla_t' + Bf_t' \tag{3.86}$$

and

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t \sim L(\mu, s_t) \tag{3.87}$$

$$\epsilon_t = \xi_t \tag{3.88}$$

$$h_{t+1} = \omega + A\epsilon_t^2 + Bh_t \tag{3.89}$$

An overview of these GARCH/GAS models illustrated above can be found in Table 3.1, where the mean equations and link functions that distinguish models are listed.

Table 3.1: GARCH and GAS models with varied innovation terms

	Panel A: GARCH modles	
	Mean equation	Link function
Normal	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t \sim N(0, h_t)$	$\epsilon_t = \xi_t$
SG	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t} + \xi_t^*, \quad \xi_t^* \sim G(a, b)$	$\epsilon_t = -\sqrt{ah_t} + \xi_t^*$
SIG	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{\frac{h_t k}{\mu}} + \xi_t^*, \quad \xi_t^* \sim IG(\mu, k)$	$\epsilon_t = -\sqrt{\frac{h_t k}{\mu}} + \xi_t^*$
SNW	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \frac{\sqrt{h_t}\Gamma(1+1/k)}{\sqrt{u}} + \xi_t^*, \quad \xi_t^* \sim W(\theta, k)$	$\epsilon_t = -\frac{\sqrt{h_t}\Gamma(1+1/k)}{\sqrt{u}} + \xi_t^*$
TLS	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t^* \sim TLS(\mu, s, \nu)$	$\epsilon_t = \xi_t$
LOG	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t \sim LOG(\mu, s)$	$\epsilon_t = \xi_t$
	Panel B: GAS modles	
	Mean equation	Link function
Normal	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t \sim N(0, h_t)$	$f_t = \ln h_t$
SG	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t} + \xi_t^*, \quad \xi_t^* \sim G(a, b)$	$f_t = \ln b_t$
SIG	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{\frac{h_t k}{\mu}} + \xi_t^*, \quad \xi_t^* \sim IG(\mu, k)$	$f_t = \ln k_t$
SNW	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \frac{\sqrt{h_t}\Gamma(1+1/k)}{\sqrt{u}} + \xi_t^*, \quad \xi_t^* \sim W(\theta, k)$	$f_t = \ln \theta_t$
TLS	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t^* \sim TLS(\mu, s, \nu)$	$f_t = \ln s_t$
LOG	$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t, \quad \xi_t \sim LOG(\mu, s)$	$f_t = \ln s_t$

Notes: This table lists GARCH and GAS models constructed in this chapter. Each model consists of a system that includes a mean equation, a link function, and an updating equation, which is the same for GARCH (GAS) models (hence it is not displayed). For SNW-GARCH/GAS, we have $u = \Gamma\left(1 + \frac{2}{k}\right)\left(\Gamma\left(1 + \frac{1}{k}\right)\right)^2$.

3.3 Implied volatility

As one of the main aims of this study, the performance of observation-driven GAS and GARCH models in volatility forecasting is compared to that implied by options prices. I attempt to explore whether parametric models in this study provide additional information compared to options. In this section, I introduce two methodologies applied in extracting expected volatility from options.

The first approach is quite straightforward. The Black-Scholes model can be used inversely to estimate volatility, where ATM option prices are utilised as the inputs. This is also the starting point of research on option implied volatility, as well as other implied information.

On the basis of studies by Breeden and Litzenberger (1978); Derman (1994); Dupire (1994); Neuberger (1994); Carr and Madan; Demeterfi et al. (1999a), Britten-Jones and Neuberger (2000) propose to compute the model-free implied volatility. This volatility is independent of any option pricing model and hence is not constrained by unrealistic assumptions on underlying assets. This is appealing, as popular analytical option pricing models, such as the Black-Scholes model, are founded upon a series of problematic assumptions, such as that of constant volatility.

Demeterfi et al. (1999a), Britten-Jones and Neuberger (2000) claim that, in a risk-neutral world, underlying assets' volatility is integral of call option prices with respect to strike prices. Jiang and Tian (2005) later made this approach easier to apply in a discrete strike setting and compare the predictive capability of the BS model implied volatility to model-free implied volatility. Model-free implied volatility is a direct inspection of options market efficiency. It is inclusive of all types of options. Intuitively, it should be able to reflect more information that the Black-Scholes model implied volatility. Following Demeterfi et al. (1999a), Britten-Jones and Neuberger (2000), assuming that the no-arbitrage principle is valid and both

dividend rate and risk-free interest rate are equal to zero, in a risk-neutral world, estimate MFIV using the formula

$$E^{Q}\left[V_{0,T}\right] = 2\mathrm{e}^{rT}\left[\int_{0}^{F_{0,T}} \frac{P(K,T)}{K^{2}}\,\mathrm{d}K + \int_{F_{0,T}}^{\infty} \frac{C(K,T)}{K^{2}}\,\mathrm{d}K\right] \tag{3.90}$$

where $E^{Q}\left[V_{0,T}\right]$ is the expected variance over period $[0,T]$ in the risk-neutral world, $F_{0,T}$ is underlying's forward price at time zero, $C(K,T)$ and $P(K,T)$ are prices of call options and put options written on same underlying with an identical time to maturity. Finally, K stands for strike prices. All inputs on RHS of Equation (3.90) can be obtained and substituted instantly, except for $F_{0,T}$. Inspired by CBOE's method, for each trading day, define $K_{0,T}$ as the strike price of the 'closest-to-the-money' call option or put option. In other words, $K_{0,T}$ is the strike price of the option whose moneyness is closest to one. To realise this with minimum error, I pick the option with a delta closest to 0.5 in absolute value. Since the market delta is given in our data set, this can be done directly. After selecting $K_{0,T}$ for each day, I compute $F_{0,T}$ as the forward price of $K_{0,T}$ in risk-neutral world.

$$F_{0,T} = \mathrm{e}^{rT}K_{0,T} \tag{3.91}$$

In most cases, $F_{0,T}$ helps separate call option prices and put option prices exactly. Intuitively, the two integral parts in Equation (3.90) will be calculated using put option prices that are lower than $F_{0,T}$ and call option prices that are larger than $F_{0,T}$ on a daily basis. In this way, only out-of-the-money options are used in the computations of implied volatility via Equation (3.90). The reason is that OTM options are typically more liquid in the market and thus tend to be more informative.

In practice, Equation (3.90) will be discretised to be computable. One problem during the implementation is then the limited number of option prices and strike prices. Even without applying filters to original options data, only a few observations of option prices and strike prices are available for each day. In order to obtain

sufficient pairs of option prices and strike prices to approximate integrals in Equation (3.90), it is necessary to fit implied volatility curves (the volatility smile) on each day across a linear space of strikes. The importance of doing so is presented by Jiang and Tian (2007).

The method for fitting implied volatility smile was first proposed by Shimko (1993). Option prices are first translated into implied volatility using the Black-Scholes Model. A continuous function that bridges implied volatility and the strike price is then estimated. Finally, such functions are used to extend ranges of strike prices and corresponding option prices, which are not observable. Malz (1997a) and Malz (1997b) used option delta, instead of strike prices, as the X-axis variable when fitting the implied volatility curve. In terms of curve fitting methods, both Malz (1997b) and Shimko (1993) applied low-order polynomials. As an alternative, Campa et al. (1998) introduced the spline approach. Based on these studies, Bliss and Panigirtzoglou (2002) developed the natural spline method for fitting implied volatility curves. Apart from spline methods, Taylor et al. (2010) propose a constrained weighted optimisation method for calibration of volatility curve parameters.

In this study, I select strike prices as the independent variable again implied volatility. Fitted curves reflect relationships between implied volatility and strike prices. Following Bliss and Panigirtzoglou (2002), I adopt the natural spline method for implied volatility curve fitting. According to Jiang and Tian (2005), although MFIV is advantageous in that it is not constrained by any models and unrealistic assumptions, two types of errors-the truncation error and discretisation error-may still occur during computation. Jiang and Tian (2005) point out that, when doing integrals with respect to strike prices, if $K_{\min} < F_0 - 2\sigma F_0$ and $K_{\max} > F_0 + 2\sigma F_0$, the truncation error is negligible. It is also noteworthy that discontinuity also brings errors. Such discretisation error can be narrowed down by applying smaller and smaller dK in Equation (3.90). When $dK < 0.35\sigma F_0$, this error is negligible. Following Jiang and Tian (2007) and out of discretion concern, I create 500 equally spaced

logarithmic strike prices above $\ln F_{0,T}$ and, correspondingly, 500 equally spaced logarithmic strike prices below $\ln F_{0,T}$. Hence 1001 strike prices and associated option prices are estimated for each time to maturity on each trading day. By doing so, both conditions mentioned above are strictly met. Hence our estimations of MFIV should be accurate. Figure 3.1 depicts the plots of ATM option implied volatility (IV) and corresponding model-free implied volatility (MFIV).

By implementing the procedures above, I obtain implied volatility of the underlying on each day and across a variety of maturities. The next step is to compute 30-day implied volatility via interpolation. For instance, on a specific trading day, there are multiple times to maturity for each option. To compute the implied volatility of that day, I estimate a far-term implied variance, whose time to maturity is longer than 30 days, and a near-term implied variance, whose time to maturity is shorter than 30 days. These two implied variances reflect the market's expectation of future variance of underlying over the coming near term and far term. Our purpose is to estimate the 30-day implied volatility, which can be obtained by interpolation using both near-term and far-term implied variances.

$$MFIV_{30} = \sqrt{\left(T_n \sigma_n^2 \frac{T_f - 30}{T_f - T_n} + T_f \sigma_f^2 \frac{30 - T_n}{T_f - T_n}\right) \frac{365}{30}} \tag{3.92}$$

where T_n and T_f are near-term and far-term time to maturity, while σ_n^2 and σ_f^2 are corresponding implied variances respectively. In this study, $MFIV_{30}$ is the predictor of realised volatility I use in comparison to parametric model-based predictors.

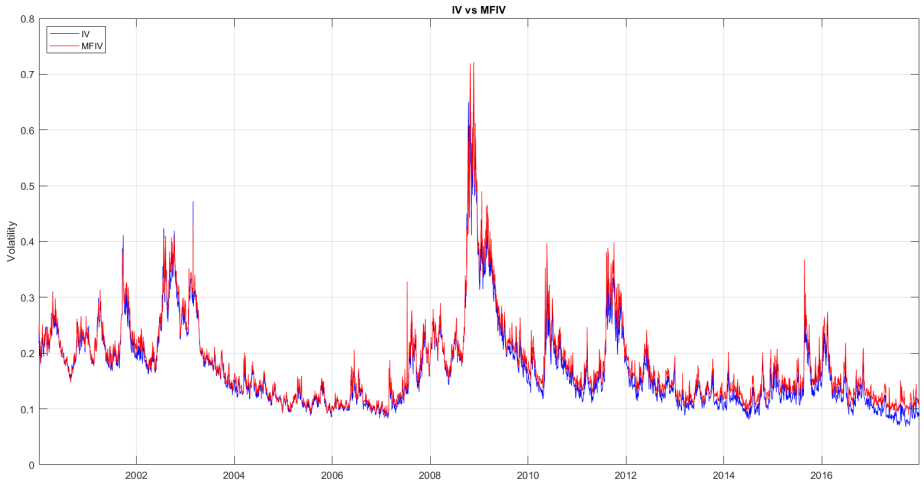

Figure 3.1: IV vs MFIV

Notes: This figure plots a time series of implied volatility, represented by IV (blue line) and MFIV (red line) over a full sample between 01/2001 and 12/2017. While the former is obtained using ATM options, the latter is calculated in a model-free fashion. In both cases, implied volatility is annualised.

3.4 Methodology

3.4.1 Data

In the construction of GAS and GARCH models, I mainly employ daily data of the Dow Jones Industrial Average (DJIA) index between 04/01/1997 and 21/04/2022, obtained from Bloomberg. In both parametric volatility modelling with GAS and GARCH models and implied volatility estimation, the risk-free rate is proxied by 3-month T-bill rates, which are downloaded from the website of the Federal Reserve Banks of St Louis. It is less cost-effective to compute implied volatility, as a large amount of option data is required. Due to limited access to options data, the estimated implied volatility ranges between 04/01/1996 and 31/12/2017. This means that for the purpose of GAS and GARCH modelling and evaluating, a longer sample is adopted, while for calculating implied volatility and making comparisons between parametric models and implied volatility, a shorter sample is applied.

The application of the DJIA index in GAS and GARCH modelling is straightforward. I simply take the logarithmic returns as the training set. Together with risk-free rate time series, a five-year window (1260 observations) is used to dynamically model volatility (or variance), and thus one-step ahead prediction is made step by step after each model fitting.[2] I consider predicting horizons from one day up to one month. For both GAS and GARCH models, one-step-ahead analytical predictions are available. However, when it comes to multi-step ahead predictions, the ways I adopt for GAS and GARCH vary. For GARCH models, it can be shown that multi-step forecasting can still be analytical. The equation below is used to realise it:

$$h_n = \omega + (A + B)\, h_{n-1}, \quad n > 1 \tag{3.93}$$

Where A and B are coefficients of the disturbance term and lagged variance term in the GARCH variance updating equation. It is clear that this analytical forecasting is a point estimation of future variance. On the contrary, for GAS models, as in Ardia et al. (2019), multi-step prediction is usually in the form of density forecast via simulation approaches. For each 5-step ahead and 22-step ahead prediction to be made with GAS models, 5000 simulated paths of driver factors are generated. As a result, we obtain a distribution of the driver factor, which can be converted into a distribution of volatility via the link function. The mean value of the volatility distribution is computed as the volatility prediction at one week and one month forecasting horizons.

The last step before computing IV and MFIV is option data processing. First, I set several filters to drop problematic or less relevant option data. First, I naturally discard option data with negative implied volatility since negative volatility indicates

[2]Note that the fitting of GARCH models is in a rolling window sense. This is different from the recursive fitting process in chapter 2. The reason is that the fitting of the GARCH model only inputs historical returns, and consequently, output predictions correspond to specific return series. It is more like a function between historical returns and predictions. In chapter 2, the fitting is done by regressing market returns onto predictors. Both historical information and future realisations are involved. The prediction is implemented recursively with each fitted linear model.

problematic option prices, which is not capable of providing useful information in estimating MFIV. Second, following Britten-Jones and Neuberger (2000), I only use out-of-the-money (OTM) options data for computing MFIV. In practice, I keep call option data whose delta is smaller than 0.5 and put option data whose delta is larger than -0.5. Third, since an option with zero bid price is often a companion of abnormal volatility, I drop option data with zero bid price. Finally, in order to calculate the 30-day model-free implied volatility, I keep option data with a maturity longer than seven days and shorter than 30 days for interpolation.

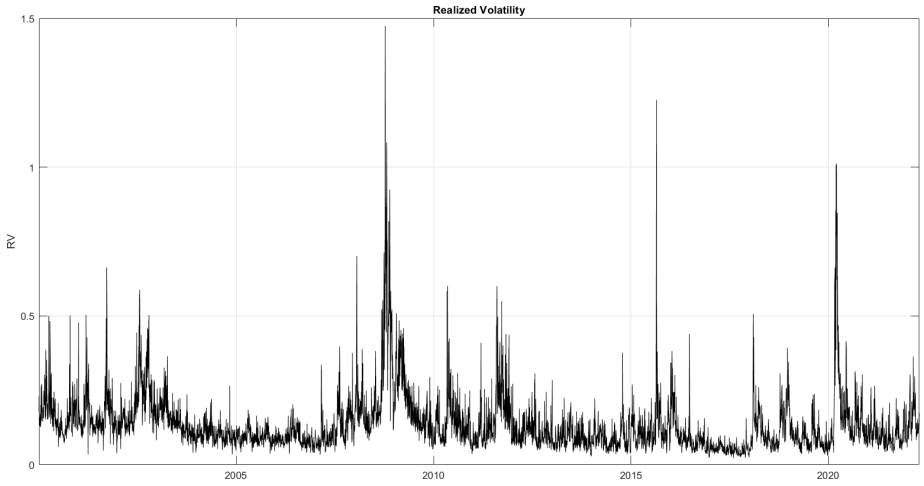

Figure 3.2: Realised Volatility

Notes: This figure plots the time series of realised volatility ranging between 01/2001 and 04/2022. It is computed using 5-min intra-day high-frequency data and is obtained directly from the Ox-Man realised volatility library.

Finally, to evaluate the predictive power of each model, I compare model-predicted volatility to realised volatility (RV) through a series of metrics. The realised volatility is computed using 5-min intra-day index data, which is obtained from the Ox-Man realised volatility library.[3]. The time series of RV and parametric model predicted volatility are plotted in Figures 3.2 and 3.3. Across all eight assumptions on the distribution of innovation terms, the red lines plot the GAS model predicted volatility, while the blue lines plot the GARCH model predictions. Correspondingly,

[3]The realised volatility is computed using high-frequency intra-day data and can be downloaded from the website of Ox-Man Institute of Quantitative Finance. https://oxford-man.ox.ac.uk/resources/the-realised-library/

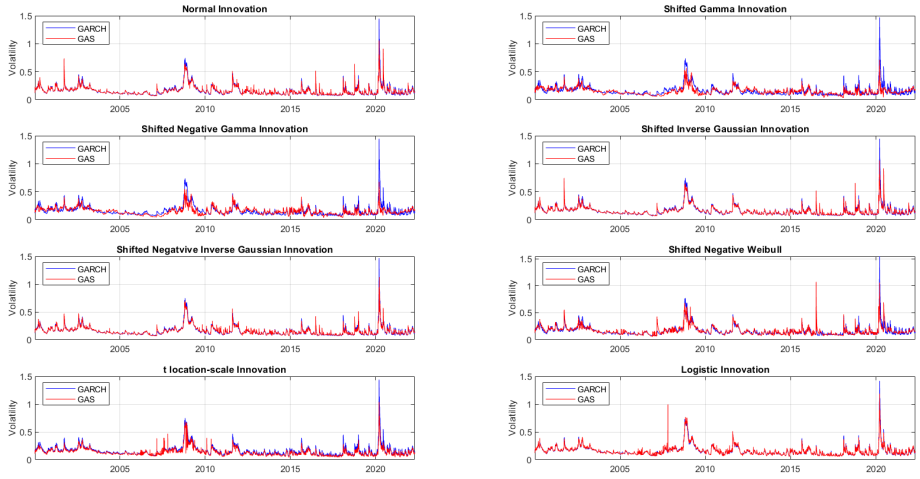

Figure 3.3: GARCH and GAS model-based Volatility Forecast

Notes: This figure depicts comparisons between predictions made by GARCH and GAS models with identically distributed innovations terms. With each GARCH/GAS model, five years' observations are utilised to fit models at each point in time, and one step ahead forecast is made dynamically to form the time series.

summary statistics of both RV and all parametric model predicted volatility can be found in Table 3.2. An analysis of this table will be conducted in the next section.

Table 3.2: Summary statistics of predicted volatility

	Avg.	Max.	Min.	Med.	5th	25th	75th	95th
Panel A: GARCH models								
RV	0.1356	1.4742	0.0220	0.1110	0.0489	0.0778	0.1614	0.2987
Normal	0.1600	1.4485	0.0663	0.1359	0.0817	0.1056	0.1841	0.3102
SG	0.1607	1.4700	0.0687	0.1367	0.0823	0.1061	0.1853	0.3114
SNG	0.1594	1.4477	0.0672	0.1351	0.0816	0.1050	0.1838	0.3105
SIG	0.1600	1.4466	0.0633	0.1360	0.0819	0.1054	0.1843	0.3102
SNIG	0.1593	1.4678	0.0671	0.1352	0.0815	0.1051	0.1834	0.3096
SNW	0.1647	1.5479	0.0680	0.1394	0.0891	0.1100	0.1896	0.3154
TLS	0.1609	1.4419	0.0603	0.1369	0.0797	0.1055	0.1862	0.3117
LOG	0.1596	1.4179	0.0611	0.1360	0.0786	0.1045	0.1851	0.3083
Panel B: GAS models								
	Avg.	Max.	Min.	Med.	5th	25th	75th	95th
RV	0.1356	1.4742	0.0220	0.1110	0.0489	0.0778	0.1614	0.2987
Normal	0.1584	1.0588	0.0650	0.1385	0.0838	0.1085	0.1808	0.2981
SG	0.1498	0.7110	0.0341	0.1387	0.0760	0.1085	0.1774	0.2524
SNG	0.1499	0.6769	0.0353	0.1408	0.0683	0.1059	0.1814	0.2619
SIG	0.1579	1.0596	0.0656	0.1382	0.0839	0.1084	0.1801	0.2966
SNIG	0.1580	1.1244	0.0657	0.1381	0.0834	0.1081	0.1804	0.2984
SNW	0.1651	1.0748	0.0565	0.1438	0.0855	0.1138	0.1926	0.3123
TLS	0.1377	1.0395	0.0423	0.1199	0.0638	0.0895	0.1632	0.2594
LOG	0.1604	1.1869	0.0556	0.1394	0.0763	0.1057	0.1873	0.3065

Notes: This table reports summary statistics of realised volatility (RV) and model-based predictions. More specifically, for each entity, its average, maximum, minimum, and median, together with its 5th, 25th, 75th and 95th percentiles, are displayed. The statistics of GARCH models are reported in panel A, while those of GAS models are reported in panel B.

3.4.2 Model evaluation

An extensive model evaluation is conducted in this study. Each series of predicted volatility from parametric or non-parametric methods is not only compared to its realisations but also to each of its counterparts. The most straightforward comparison is made by running univariate regressions. More specifically, I run the Mincer–Zarnowitz regressions Mincer and Zarnowitz (1969) described by the equation below:

$$RV_{t,t+k} = \alpha + \beta \hat{PV}_{t,t+k} + \epsilon_t \tag{3.94}$$

where $RV_{t,t+k}$ is the realised volatility k days from t, and $\hat{PV}_{t,t+k}$ stands for the corresponding predicted volatility using information available until t.

Generally speaking, volatility predicted by a model is more informative if the intercept α is closer to zero and the slope β, to one. As the predicted volatility is less biased. In the extreme case of zero alpha and a unit beta, the forecast is unbiased in a statistical sense. In terms of R-square, higher R^2 (or, equivalently, a low sum of squared residuals) would indicate the efficiency and accuracy of the forecast volatility series. The results of Mincer-Zarnowitz regressions for forecasting horizons from 1 day up to 1 month are reported in Table 3.3.

The R-squared values from univariate regressions are reasonable measures of how well the realised volatility can be described by each model prediction. However, it might be that, while the trends and changes in volatility can be better captured by some model predictions, and thus result in higher R-squared values, the predicted volatility can be significantly biased in terms of its value. In fact, it can be shown from the nature of GAS and the model that compared to GAS, the GARCH model tends to overestimate volatility. This is because the application of score functions in GAS models makes the dynamics of volatility respond more mildly to large changes in index returns y_t. This is quite an appealing property in financial modelling, as

outliers of index returns might simply originate from its heavy-tailed nature instead of actual increases in realised volatility. It is necessary to inspect how models perform in terms of prediction accuracy. To realise this, I adopt the Mean Absolute Errors (MAE) to evaluate each model prediction. The MAE is computed as:

$$\text{MAE} = \frac{\sum_{t=1}^{n} \left| RV_t - \hat{PV}_t \right|}{n} \tag{3.95}$$

As a robust loss function to noise in volatility modelling, the Quasi-Likelihood (QLIKE) function is widely utilised in practice. In addition to the MAE, the accuracy of model predictions is also evaluated using QLIKE that can be computed as:

$$\text{QLIKE} = \frac{1}{\tau} \sum_{t=1}^{\tau} \left[\ln \left(\hat{PV}_{t:t+k} \right) + \frac{RV_{t:t+k}}{\hat{PV}_{t:t+k}} \right] \tag{3.96}$$

where τ is equal to the number of volatility predictions/realised volatility and k represents the forecasting horizon. I compute the MAE and QLIKE values of each model and report them in Table 3.4 and 3.5, respectively. A detailed discussion of the table is placed in the next section.

The evaluations so far are implemented between model predictions and realisation directly. Given the fact that different models yield different estimates of volatility, I am also interested to know if these differences are significant. In other words, it is beneficial to investigate whether forecast errors from each candidate model are statistically different from other candidates. To conduct pairwise comparisons, I employ the out-of-sample R-squared (R^2_{os}), that can be calculated as:

$$R^2_{os} = 1 - \frac{MAE_m}{MAE_{bmk}} \tag{3.97}$$

where MAE_m is the absolute prediction error from each model. Each time, I fix one of the candidate models as the benchmark and MAE_{bmk} and compute R^2_{os}. In this way, pairwise comparisons between all models are conducted.

Table 3.3: Mincer-Zarnowitz regressions over horizons - full sample

	1 day			5 days			22 days		
	α	β	R^2	α	β	R^2	α	β	R^2
Panel A: GARCH modles									
Normal	0.0058***	0.8113***	0.6069	0.0244***	0.7801***	0.4654	0.0805***	0.5808***	0.1639
SG	0.0059***	0.8068***	0.6060	0.0249***	0.7707***	0.4645	0.0808***	0.5677***	0.1598
SNG	0.0062***	0.8116***	0.6090	0.0242***	0.7826***	0.4718	0.0788***	0.5948***	0.1739
SIG	0.0057***	0.8123***	0.6072	0.0242***	0.7814***	0.4660	0.0802***	0.5830***	0.1651
SNIG	0.0065***	0.8099***	0.6091	0.0246***	0.7798***	0.4723	0.0778***	0.6049***	0.1782
SNW	0.0007	0.8189***	0.6190	0.0188***	0.7924***	0.4830	0.0738***	0.5838***	0.1898
TLS	0.0098***	0.7818***	0.5971	0.0298***	0.7102***	0.4476	0.0836***	0.4639***	0.1525
LOG	0.0091***	0.7924***	0.5999	0.0280***	0.7420***	0.4564	0.0810***	0.5340***	0.1699
Panel B: GAS modles									
	1 day			5 days			22 days		
	α	β	R^2	α	β	R^2	α	β	R^2
Normal	−0.0091***	0.9135***	0.5843	−0.0162***	0.9744***	0.4517	−0.0088**	0.9477***	0.2040
SG	−0.0379***	1.1585***	0.5289	−0.0267***	1.0907***	0.3676	0.0082*	0.8642***	0.1333
SNG	−0.0297***	1.1022***	0.5206	−0.0175***	1.0176***	0.3476	0.0194***	0.7566***	0.1110
SIG	−0.0096***	0.9191***	0.5852	−0.0176***	0.9861***	0.4557	−0.0145***	0.9882***	0.2100
SNIG	−0.0083***	0.9108***	0.5825	−0.0115***	0.9422***	0.4495	0.0255***	0.7168***	0.1625
SNW	−0.0121***	0.8947***	0.5713	−0.0119***	0.9071***	0.4002	0.0062	0.8155***	0.1418
TLS	0.0031*	0.9619***	0.5996	0.0086***	0.7808***	0.4545	0.0296***	0.6507***	0.2111
LOG	0.0019	0.8334***	0.6009	0.0050**	0.8167***	0.4654	0.0293***	0.6682***	0.2133

Notes: This table reports results from Mincer-Zarnowitz regressions over forecasting horizons from one day to one month. A full sample is adopted to dynamically obtain predictions using parametric models and realised volatility is regressed against these predictions. For each regression, intercept α, slope β and goodness of fit R^2 are reported. Statistics of GARCH models are reported in panel A, and those of GAS models are reported in panel B.

Table 3.4: MAE values over horizons - full sample

	Daily	Weekly	Monthly
Panel A: GARCH modles			
Normal	4.6233	4.5307	6.3048
SG	4.6553	4.5590	6.2501
SNG	4.5782	4.4956	6.2182
SIG	4.6173	4.5338	6.2840
SNIG	4.5750	4.4957	6.1774
SNW	4.8495	4.6385	5.9549
TLS	4.6955	4.7907	5.9812
LOG	4.6211	4.6356	6.0952
Panel B: GAS modles			
	Daily	Weekly	Monthly
Normal	4.6380	5.1164	5.9481
SG	4.8680	5.3146	6.0821
SNG	4.9071	5.4573	6.4317
SIG	4.6109	5.0848	5.9237
SNIG	4.6464	5.1250	6.0029
SNW	4.9551	5.5627	6.4441
TLS	3.8626	5.4237	6.4014
LOG	4.6943	5.2415	6.1922

Notes: This table reports the mean absolute errors (MAE) between each observation-driven model-based prediction and realised volatility over the full sample. The MAEs are computed at horizons from one day to one month, and the MAE values are multiplied by 100 for easier comparisons. MAEs of GARCH models are reported in panel A and those of GAS models are reported in panel B.

Table 3.5: Quasi-likelihood values over horizons - full sample

Panel A: GARCH modles			
	Daily	Weekly	Monthly
Normal	-1.0917	-1.0800	-0.4991
SG	-1.0910	-1.0804	-0.5461
SNG	-1.0926	-1.0807	-0.5221
SIG	-1.0918	-1.0799	-0.4852
SNIG	-1.0927	-1.0802	-0.3839
SNW	-1.0854	-1.0779	0.3737
TLS	-1.0913	-1.0782	-0.8738
LOG	-1.0926	-1.0793	-0.7554
Panel B: GAS modles			
	Daily	Weekly	Monthly
Normal	-1.0877	-1.0670	-1.0322
SG	-1.0718	-1.0506	-1.0175
SNG	-1.0715	-1.0489	-1.0125
SIG	-1.0884	-1.0677	-1.0327
SNIG	-1.0875	-1.0672	-1.0318
SNW	-1.0805	-1.0576	-1.0199
TLS	-1.1017	-1.0647	-1.0278
LOG	-1.0897	-1.0679	-1.0311

Notes: This table reports the quasi-likelihood of each model-based prediction against realised volatility over the full sample. The quasi-likelihood is robust to noise contained in volatility and is computed at horizons from one day to one month. More negative quasi-likelihood values indicate superior out-of-sample performance. The statistics of GARCH models are reported in panel A, and those of GAS models are reported in panel B.

Clearly, $R^2_{os} > 0$ indicates that the candidate model performs better than the benchmark, and vice versa. When $R^2_{os} = 1$, the candidate model predicts volatility perfectly. The values of R^2_{os} are reported in Tables 3.6, 3.7 and 3.8. To inspect the statistical significance of the outperformance of each model against the other candidates, the Clark-West (CW) tests (Clark and West, 2007) are implemented and the statistics are reported in bold below associated R^2_{os} entries.

MAE, QLIKE and R^2_{os} are evaluations of one single approach in terms of its forecasting ability. However, I also care about whether the errors between the two forecasts are statistically different. To this end, I additionally adopt the Diebold–Mariano (DM) test (Diebold and Mariano, 1995). Taking into consideration possible autocorrelations, Newey–West standard errors are computed. I report statistics of DM tests and CW tests over forecasting horizons from one day to one month in

Table 3.6: Out-of-sample R^2 at daily horizon - full sample

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	0.00	-0.69	0.97	0.13	1.04	-4.89	-1.56	0.05	-0.32	-5.29	-6.14	0.27	-0.50	-7.18	16.45	-1.54
	NaN	**-2.07**	**4.22**	**1.80**	**2.51**	**-0.13**	**-1.59**	**-0.74**	**1.56**	**3.37**	**3.52**	**1.59**	**1.44**	**2.39**	**4.04**	**1.28**
SG	0.69	0.00	1.66	0.82	1.72	-4.17	-0.86	0.74	0.37	-4.57	-5.41	0.95	0.19	-6.44	17.03	-0.84
	2.95	**NaN**	**5.84**	**2.91**	**6.85**	**0.31**	**-1.25**	**0.01**	**1.59**	**3.31**	**3.45**	**1.62**	**1.51**	**2.36**	**3.99**	**1.40**
SNG	-0.98	-1.68	0.00	-0.85	0.07	-5.93	-2.56	-0.94	-1.31	-6.33	-7.18	-0.71	-1.49	-8.23	15.63	-2.54
	-2.79	**-5.21**	**NaN**	**-1.87**	**0.33**	**-0.60**	**-2.19**	**-1.50**	**1.46**	**3.26**	**3.41**	**1.49**	**1.31**	**2.29**	**3.79**	**1.08**
SIG	-0.13	-0.82	0.85	0.00	0.92	-5.03	-1.69	-0.08	-0.45	-5.43	-6.27	0.14	-0.63	-7.32	16.34	-1.67
	0.30	**-2.33**	**3.19**	**NaN**	**2.00**	**-0.26**	**-1.67**	**-0.87**	**1.55**	**3.35**	**3.50**	**1.58**	**1.43**	**2.38**	**4.04**	**1.27**
SNIG	-1.05	-1.76	-0.07	-0.93	0.00	-6.00	-2.63	-1.01	-1.38	-6.40	-7.26	-0.78	-1.56	-8.31	15.57	-2.61
	-1.05	**-5.45**	**0.88**	**-0.74**	**NaN**	**-0.47**	**-2.40**	**-1.58**	**1.46**	**3.21**	**3.34**	**1.50**	**1.29**	**2.25**	**3.66**	**1.06**
SNW	4.67	4.00	5.59	4.79	5.66	0.00	3.18	4.71	4.36	-0.38	-1.19	4.92	4.19	-2.18	20.35	3.20
	2.36	**1.47**	**2.89**	**2.48**	**2.74**	**NaN**	**0.25**	**1.21**	**2.04**	**3.77**	**3.89**	**2.08**	**1.82**	**2.66**	**5.00**	**1.90**
TLS	1.54	0.86	2.50	1.66	2.57	-3.28	0.00	1.59	1.23	-3.67	-4.51	1.80	1.05	-5.53	17.74	0.03
	2.32	**2.19**	**2.76**	**2.34**	**2.96**	**1.37**	**NaN**	**1.80**	**1.81**	**3.15**	**3.27**	**1.83**	**1.78**	**2.33**	**3.56**	**1.53**
LOG	-0.05	-0.74	0.93	0.08	1.00	-4.94	-1.61	0.00	-0.37	-5.34	-6.19	0.22	-0.55	-7.23	16.41	-1.58
	1.65	**1.21**	**2.00**	**1.72**	**2.12**	**1.00**	**-0.97**	**NaN**	**1.79**	**3.30**	**3.43**	**1.82**	**1.71**	**2.44**	**3.72**	**1.43**
Normal	0.32	-0.37	1.29	0.45	1.36	-4.56	-1.24	0.36	0.00	-4.96	-5.80	0.58	-0.18	-6.84	16.72	-1.21
	2.05	**1.91**	**2.27**	**2.08**	**2.28**	**1.92**	**1.73**	**1.90**	**NaN**	**4.85**	**5.06**	**4.25**	**1.31**	**2.18**	**4.81**	**2.28**
SG	5.03	4.37	5.95	5.15	6.02	0.38	3.54	5.07	4.72	0.00	-0.80	5.28	4.55	-1.79	20.65	3.57
	5.47	**5.39**	**5.51**	**5.45**	**5.48**	**5.23**	**5.30**	**5.44**	**6.76**	**NaN**	**2.17**	**6.80**	**6.29**	**5.74**	**8.29**	**6.15**
SNG	5.78	5.13	6.70	5.90	6.77	1.17	4.31	5.83	5.48	0.80	0.00	6.04	5.31	-0.98	21.28	4.34
	5.59	**5.50**	**5.62**	**5.57**	**5.59**	**5.27**	**5.43**	**5.56**	**6.87**	**3.02**	**NaN**	**6.93**	**6.50**	**5.79**	**8.25**	**6.05**
SIG	-0.27	-0.96	0.71	-0.14	0.78	-5.18	-1.84	-0.22	-0.59	-5.58	-6.42	0.00	-0.77	-7.47	16.23	-1.81
	1.84	**1.70**	**2.05**	**1.86**	**2.07**	**1.71**	**1.57**	**1.72**	**-0.04**	**4.65**	**4.88**	**NaN**	**1.17**	**1.95**	**4.53**	**2.08**
SNIG	0.50	-0.19	1.47	0.63	1.54	-4.37	-1.06	0.55	0.18	-4.77	-5.61	0.76	0.00	-6.64	16.87	-1.03
	2.90	**2.58**	**3.29**	**2.94**	**3.24**	**2.63**	**1.61**	**2.09**	**1.43**	**4.46**	**4.76**	**1.66**	**NaN**	**2.47**	**7.37**	**2.99**
SNW	6.70	6.05	7.61	6.82	7.67	2.13	5.24	6.74	6.40	1.76	0.97	6.95	6.23	0.00	22.05	5.26
	3.25	**3.20**	**3.31**	**3.24**	**3.30**	**2.95**	**3.21**	**3.28**	**2.96**	**2.94**	**2.99**	**2.97**	**3.29**	**NaN**	**4.42**	**3.17**
TLS	-19.69	-20.52	-18.53	-19.54	-18.44	-25.55	-21.56	-19.63	-20.07	-26.03	-27.04	-19.37	-20.29	-28.28	0.00	-21.53
	1.95	**1.93**	**2.01**	**1.98**	**2.00**	**2.72**	**1.33**	**1.38**	**3.53**	**5.61**	**5.86**	**3.66**	**2.10**	**5.24**	**NaN**	**2.07**
LOG	1.51	0.83	2.47	1.64	2.54	-3.31	-0.03	1.56	1.20	-3.70	-4.53	1.78	1.02	-5.56	17.72	0.00
	1.33	**1.18**	**1.48**	**1.36**	**1.47**	**1.21**	**0.59**	**0.89**	**2.93**	**4.93**	**5.00**	**2.99**	**2.39**	**3.65**	**5.50**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions at daily forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using the full sample. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

Table 3.7: Out-of-sample R^2 at weekly horizon - full sample

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	0.00	-0.62	0.78	-0.07	0.77	-2.38	-5.74	-2.31	-12.93	-17.30	-20.45	-12.23	-13.12	-22.78	-19.71	-15.69
	NaN	**-0.60**	**4.48**	**1.89**	**3.59**	**1.85**	**-1.34**	**-1.19**	**1.52**	**2.85**	**2.93**	**1.54**	**1.43**	**2.05**	**1.36**	**1.26**
SG	0.62	0.00	1.39	0.55	1.39	-1.74	-5.08	-1.68	-12.23	-16.57	-19.70	-11.53	-12.42	-22.02	-18.97	-14.97
	1.84	**NaN**	**2.82**	**1.60**	**3.39**	**1.71**	**-1.46**	**-1.09**	**1.46**	**2.68**	**2.75**	**1.48**	**1.39**	**1.92**	**1.23**	**1.20**
SNG	-0.78	-1.41	0.00	-0.85	0.00	-3.18	-6.57	-3.11	-13.81	-18.22	-21.39	-13.11	-14.00	-23.74	-20.64	-16.59
	-2.06	**-2.13**	**NaN**	**-1.32**	**0.95**	**1.39**	**-1.60**	**-1.80**	**1.42**	**2.76**	**2.84**	**1.45**	**1.31**	**1.95**	**1.05**	**1.05**
SIG	0.07	-0.56	0.84	0.00	0.84	-2.31	-5.67	-2.25	-12.85	-17.22	-20.37	-12.16	-13.04	-22.69	-19.63	-15.61
	1.43	**-0.76**	**3.29**	**NaN**	**2.75**	**1.82**	**-1.34**	**-1.22**	**1.53**	**2.83**	**2.92**	**1.55**	**1.44**	**2.04**	**1.37**	**1.23**
SNIG	-0.78	-1.41	0.00	-0.85	0.00	-3.17	-6.56	-3.11	-13.81	-18.22	-21.39	-13.10	-14.00	-23.73	-20.64	-16.59
	-0.23	**-1.84**	**1.43**	**-0.17**	**NaN**	**1.34**	**-1.56**	**-1.72**	**1.41**	**2.69**	**2.75**	**1.44**	**1.28**	**1.90**	**1.07**	**1.04**
SNW	2.32	1.71	3.08	2.26	3.08	0.00	-3.28	0.06	-10.30	-14.58	-17.65	-9.62	-10.49	-19.93	-16.93	-13.00
	1.27	**0.78**	**1.47**	**1.30**	**1.34**	**NaN**	**0.12**	**0.46**	**1.81**	**3.59**	**3.55**	**1.86**	**1.58**	**2.14**	**1.13**	**1.13**
TLS	5.43	4.84	6.16	5.36	6.16	3.18	0.00	3.24	-6.80	-10.94	-13.91	-6.14	-6.98	-16.11	-13.21	-9.41
	2.33	**2.43**	**2.48**	**2.28**	**2.59**	**2.20**	**NaN**	**2.01**	**1.66**	**2.49**	**2.54**	**1.67**	**1.63**	**1.90**	**1.43**	**1.42**
LOG	2.26	1.65	3.02	2.20	3.02	-0.06	-3.35	0.00	-10.37	-14.65	-17.73	-9.69	-10.56	-20.00	-17.00	-13.07
	2.31	**2.26**	**2.55**	**2.25**	**2.76**	**2.17**	**-1.10**	**NaN**	**1.70**	**2.73**	**2.79**	**1.72**	**1.64**	**2.07**	**1.54**	**1.39**
Normal	11.45	10.89	12.13	11.39	12.13	9.34	6.36	9.40	0.00	-3.87	-6.66	0.62	-0.17	-8.72	-6.01	-2.44
	4.94	**4.47**	**5.09**	**4.96**	**4.95**	**5.54**	**3.14**	**3.91**	**NaN**	**3.78**	**3.58**	**2.98**	**1.25**	**0.28**	**2.16**	**2.31**
SG	14.75	14.22	15.41	14.69	15.41	12.72	9.86	12.78	3.73	0.00	-2.68	4.32	3.57	-4.67	-2.05	1.38
	5.57	**5.29**	**5.56**	**5.49**	**5.44**	**5.72**	**4.79**	**5.26**	**6.39**	**NaN**	**-0.04**	**6.51**	**6.54**	**4.60**	**5.11**	**5.53**
SNG	16.98	16.46	17.62	16.92	17.62	15.00	12.21	15.06	6.25	2.61	0.00	6.82	6.09	-1.93	0.62	3.96
	6.00	**5.74**	**5.98**	**5.94**	**5.87**	**5.92**	**5.23**	**5.70**	**6.53**	**4.26**	**NaN**	**6.63**	**6.48**	**5.01**	**5.29**	**5.54**
SIG	10.90	10.34	11.59	10.84	11.59	8.78	5.78	8.84	-0.62	-4.52	-7.32	0.00	-0.79	-9.40	-6.66	-3.08
	4.53	**4.09**	**4.69**	**4.54**	**4.56**	**5.04**	**2.89**	**3.60**	**-0.60**	**3.51**	**3.32**	**NaN**	**0.61**	**-0.23**	**1.85**	**1.93**
SNIG	11.60	11.04	12.28	11.54	12.28	9.49	6.52	9.55	0.17	-3.70	-6.48	0.78	0.00	-8.54	-5.83	-2.27
	3.90	**3.44**	**4.03**	**3.92**	**3.92**	**4.61**	**2.23**	**2.99**	**1.89**	**4.41**	**4.32**	**2.09**	**NaN**	**1.45**	**1.62**	**1.98**
SNW	18.55	18.04	19.18	18.50	19.18	16.62	13.88	16.67	8.02	4.46	1.89	8.59	7.87	0.00	2.50	5.77
	5.84	**5.68**	**5.89**	**5.81**	**5.84**	**5.60**	**5.19**	**5.63**	**4.49**	**3.23**	**2.93**	**4.47**	**4.48**	**NaN**	**4.17**	**4.39**
TLS	16.46	15.94	17.11	16.41	17.11	14.48	11.67	14.53	5.67	2.01	-0.62	6.25	5.51	-2.56	0.00	3.36
	6.38	**5.90**	**6.71**	**6.57**	**6.74**	**6.68**	**3.04**	**4.84**	**4.09**	**4.62**	**4.51**	**4.09**	**4.19**	**3.53**	**NaN**	**4.32**
LOG	13.56	13.02	14.23	13.50	14.23	11.50	8.60	11.56	2.39	-1.40	-4.12	2.99	2.22	-6.13	-3.48	0.00
	5.40	**4.98**	**5.72**	**5.60**	**5.78**	**6.18**	**2.66**	**4.03**	**3.59**	**4.45**	**4.21**	**3.59**	**3.26**	**3.00**	**1.58**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions at the weekly forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using the full sample. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

Table 3.8: Out-of-sample R^2 at monthly horizon - full sample

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	0.00	0.87	1.37	0.33	2.02	5.55	5.13	3.32	5.66	3.53	-2.01	6.04	4.79	-2.21	-1.53	1.79
	NaN	**2.23**	**5.34**	**3.80**	**5.39**	**5.48**	**2.84**	**3.40**	**11.66**	**10.67**	**10.77**	**11.63**	**11.62**	**10.85**	**11.39**	**10.72**
SG	-0.88	0.00	0.51	-0.54	1.16	4.72	4.30	2.48	4.83	2.69	-2.91	5.22	3.95	-3.11	-2.42	0.93
	1.31	**NaN**	**2.59**	**1.63**	**3.70**	**4.55**	**2.02**	**3.05**	**10.44**	**9.64**	**9.86**	**10.40**	**10.42**	**9.90**	**10.85**	**9.97**
SNG	-1.39	-0.51	0.00	-1.06	0.66	4.24	3.81	1.98	4.34	2.19	-3.43	4.74	3.46	-3.63	-2.95	0.42
	-1.09	**-0.68**	**NaN**	**-0.35**	**3.49**	**4.98**	**2.35**	**2.27**	**11.54**	**10.45**	**10.53**	**11.49**	**11.50**	**10.62**	**11.24**	**10.54**
SIG	-0.33	0.54	1.05	0.00	1.70	5.24	4.82	3.00	5.35	3.21	-2.35	5.73	4.47	-2.55	-1.87	1.46
	2.81	**1.58**	**5.33**	**NaN**	**5.25**	**5.33**	**2.79**	**3.33**	**11.49**	**10.40**	**10.51**	**11.44**	**11.43**	**10.64**	**11.28**	**10.57**
SNIG	-2.06	-1.18	-0.66	-1.73	0.00	3.60	3.18	1.33	3.71	1.54	-4.12	4.11	2.82	-4.32	-3.63	-0.24
	-0.60	**-0.63**	**0.91**	**-0.33**	**NaN**	**4.76**	**2.27**	**2.16**	**11.45**	**10.33**	**10.43**	**11.40**	**11.42**	**10.55**	**11.40**	**10.60**
SNW	-5.88	-4.96	-4.42	-5.53	-3.74	0.00	-0.44	-2.36	0.11	-2.14	-8.01	0.52	-0.81	-8.22	-7.50	-3.99
	1.13	**1.34**	**1.23**	**1.13**	**1.26**	**NaN**	**4.56**	**3.21**	**9.97**	**9.44**	**9.28**	**9.97**	**9.84**	**9.15**	**9.03**	**8.56**
TLS	-5.41	-4.50	-3.96	-5.06	-3.28	0.44	0.00	-1.91	0.55	-1.69	-7.53	0.96	-0.36	-7.74	-7.03	-3.53
	1.27	**1.25**	**1.39**	**1.31**	**1.58**	**2.32**	**NaN**	**1.33**	**3.92**	**4.05**	**4.31**	**3.91**	**3.89**	**4.10**	**4.43**	**3.91**
LOG	-3.44	-2.54	-2.02	-3.10	-1.35	2.30	1.87	0.00	2.41	0.21	-5.52	2.81	1.51	-5.72	-5.02	-1.59
	1.03	**1.04**	**1.41**	**1.14**	**1.96**	**3.48**	**0.90**	**NaN**	**8.08**	**7.61**	**7.88**	**8.02**	**8.10**	**7.79**	**9.51**	**8.20**
Normal	-6.00	-5.08	-4.54	-5.65	-3.85	-0.11	-0.56	-2.47	0.00	-2.25	-8.13	0.41	-0.92	-8.34	-7.62	-4.10
	5.63	**5.48**	**5.80**	**5.65**	**5.83**	**4.89**	**6.06**	**6.09**	**NaN**	**1.69**	**0.75**	**1.57**	**1.62**	**-0.77**	**1.50**	**1.13**
SG	-3.66	-2.76	-2.24	-3.32	-1.57	2.09	1.66	-0.22	2.20	0.00	-5.75	2.60	1.30	-5.95	-5.25	-1.81
	7.33	**7.18**	**7.37**	**7.23**	**7.28**	**6.50**	**7.17**	**7.43**	**5.65**	**NaN**	**-1.38**	**5.90**	**5.27**	**4.19**	**4.28**	**4.39**
SNG	1.97	2.82	3.32	2.30	3.95	7.41	7.00	5.23	7.52	5.43	0.00	7.90	6.67	-0.19	0.47	3.72
	8.17	**8.04**	**8.17**	**8.09**	**8.10**	**7.15**	**7.66**	**8.04**	**5.71**	**5.29**	**NaN**	**5.89**	**5.32**	**5.96**	**4.25**	**4.48**
SIG	-6.43	-5.51	-4.97	-6.08	-4.28	-0.53	-0.97	-2.90	-0.41	-2.67	-8.58	0.00	-1.34	-8.79	-8.06	-4.53
	5.42	**5.27**	**5.60**	**5.43**	**5.62**	**4.66**	**5.83**	**5.88**	**0.82**	**1.18**	**0.19**	**NaN**	**0.20**	**-1.72**	**1.12**	**0.78**
SNIG	-5.03	-4.12	-3.59	-4.68	-2.91	0.80	0.36	-1.54	0.91	-1.32	-7.14	1.32	0.00	-7.35	-6.64	-3.15
	3.01	**2.94**	**3.06**	**3.03**	**3.07**	**2.70**	**2.98**	**3.09**	**1.31**	**1.62**	**1.44**	**1.32**	**NaN**	**1.05**	**1.35**	**1.28**
SNW	2.16	3.01	3.51	2.48	4.14	7.59	7.18	5.41	7.70	5.62	0.19	8.08	6.85	0.00	0.66	3.91
	8.12	**7.99**	**8.16**	**8.08**	**8.15**	**7.03**	**7.82**	**8.21**	**5.80**	**5.23**	**3.03**	**5.99**	**5.19**	**NaN**	**3.70**	**4.07**
TLS	1.51	2.36	2.86	1.83	3.50	6.97	6.56	4.78	7.08	4.99	-0.47	7.46	6.22	-0.67	0.00	3.27
	6.95	**6.84**	**7.25**	**7.12**	**7.29**	**6.51**	**5.96**	**7.21**	**4.02**	**3.86**	**3.33**	**3.95**	**3.89**	**2.72**	**NaN**	**4.79**
LOG	-1.82	-0.93	-0.42	-1.48	0.24	3.83	3.41	1.57	3.94	1.78	-3.87	4.34	3.06	-4.07	-3.38	0.00
	5.89	**5.86**	**6.17**	**6.04**	**6.24**	**5.21**	**6.32**	**6.55**	**2.94**	**2.97**	**2.58**	**2.89**	**2.82**	**1.98**	**2.15**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions at the monthly forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using the full sample. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

Tables B.1, B.2 and B.3 in appendix B.

An efficient model for volatility forecasting should have stable performance through different market environments. However, as in the international comparisons of Kourtis et al. (2016), all competing models, including parametric volatility models and implied volatility, suffered significantly weakened performance during the 2008 financial crisis. It is beneficial to inspect how the models work during market turmoils. Therefore, using a sub-sample ranging from 09/2007 to 03/2019 and following an identical procedure as in the full sample case, I also report and evaluate model performance during the financial crisis in Tables 3.9, 3.10, 3.11, B.4, B.5 and B.6 in appendix B.

As introduced above, while most studies indicate that the predictive power of implied volatility is superior compared to those of model-based volatility forecasts, some others suggest the opposite. Becker et al. (2007) and Becker and Clements (2008) generally document that the informational content of implied volatility is nested in the parametric model forecast as more parametric models are applied or combined. The above procedure is implemented to investigate the informational content of the observation-driven models in this study and implied volatility represented by IV computed using ATM options and MFIV. The empirical results are presented in Tables 3.15, 3.16, 3.17, B.7, B.8 and B.9 in appendix B.

3.5 Empirical result analysis

3.5.1 Full sample

A full sample is adopted to train all GARCH and GAS models and make comparisons between them. From Table 3.2, it is easy to see that while GARCH models generally tend to yield higher volatility forecasts compared to those of GAS models, both parametric models overestimate realised volatility in the future, which averages 0.1356. However, it is noteworthy that the mean (0.1377) and median (0.1199) values of the TLS-GAS model are very close to realised volatility. This potentially indicates the superiority of the TLS-GAS model. Apart from that, the maximum values of predicted volatility from GAS models are significantly lower than those from GARCH models, which is consistent with our previous analysis, which shows that GAS responds less fiercely to large changes in index returns compared to GARCH.

From the Mincer-Zarnowitz regression results in Table 3.3, it can be observed that, at the daily horizon, if measured by R^2, the SNW-GARCH model has the best performance overall, with R^2 of 61.90%, indicating shifted negative Weibull could be a good candidate for the innovation term in volatility modelling. Moreover, the slope of TLS-GAS is 0.9619, which is much closer to 1 compared to other competitors, plus that its R^2 is on the same level as SNW-GAS, so it can be deduced that TLS-GAS outperforms the others in volatility modelling and forecasting at the daily horizon. Generally, R^2 of GAS models tend to be slightly lower than those of GARCH models at the daily horizon. This is also the case at the weekly horizon, where SNW-GARCH still has the highest R^2. At the weekly horizon, the slopes of GAS models are significantly closer to 1 compared to those of GARCH models. At the monthly horizon, in most cases, R^2 significantly increases from GARCH to GAS, indicating that GAS, while not necessarily outperforming in volatility forecasting, could be a better model for the purpose of option pricing, as most options tend to

have a maturity longer than one month.

The R^2 measures how well the RVs can be described by model-based forecasts, while prediction accuracy is equally important in volatility forecasting. Hence, I further inspect the accuracy of model prediction by computing mean absolute errors (MAE). The MAEs in Table 3.4 indicate strong outperformance of the TLS-GAS model in terms of prediction accuracy at the daily horizon, with an MAE of 3.86. This is consistent with its lower Quasi-likelihood of -1.10. More importantly, it can be observed that the MAE of TLS-GARCH-based predictions are 4.70, which means that if the innovation term is appropriately selected, the accuracy of model predictions can be improved from GARCH to GAS.

However, for weekly and monthly horizons, it is not the case. From Table 3.4, it is clear that at the weekly horizon, TLS-GAS, with MAE of 5.42, no longer has the best performance and that GARCH models consistently have lower prediction errors than GAS. At the monthly horizon, it is hard to identify domination between GARCH and GAS models, as GARCH is found to have lower prediction errors than GAS in 4 out of 8 scenarios (GARCH with SNG, SNW, TLS and LOG innovations terms). It is noteworthy that according to the quasi-likelihood values reported in Table 3.5, it is likely that GAS model-based volatility forecasts tend to work better in describing realised volatility than GARCH at the monthly horizon. This is indicated by the fact that the GAS model predictions have lower quasi-likelihood statistics below -1, while the statistics of GARCH are all above -1.

Tables 3.6, 3.7 and 3.8 show out-of-sample R^2 (R^2_{os}) of any two model-based RV predictions over different horizons. Corresponding Clark-West test statistics are reported in bold below each entry of out-of-sample R^2. Obviously, at the daily horizon, the domination of the TLS-GAS model is once again confirmed, and this is evidenced by positive R^2_{os} (ranging between 0.16 against SNIG-GARCH and 0.22 against SNW-GAS) against all the other models. Consistently, it can be observed

that GARCH models tend to outperform GAS in most cases. At the weekly horizon, the dominance of GARCH models becomes stronger. When the GARCH models are evaluated against corresponding GAS models with identical innovation terms, all R^2_{os} are found to be positive. It is also noteworthy that SNG-GARCH and SNIG-GARCH models are found to dominate all the other models at the weekly horizon. At the monthly horizon, I find that the dominance of GARCH models becomes weaker. In 4 out of 8 cases, GAS models outperform corresponding GARCH models. It is also noticeable that at the monthly horizon, SIG-GAS consistently outperforms the corresponding SIG-GARCH model. In fact, the SIG-GAS model is found to dominate all parametric models with R^2_{os} ranging from 0.01 (against SNIG-GAS) to 0.08 (against SNW-GAS).

It is worth mentioning that, for multi-step ahead forecasting, the methodologies applied for GARCH and GAS models are different. As introduced in the previous section, while the GARCH models have analytical solutions for making multi-step predictions, a simulation approach is adopted in order to make forecasts of realised volatility within the framework of GAS models. Therefore, it could be more informative to look at the model evaluations at the daily horizon, as the forecasting results are more comparable.

The findings above indicate that TLS-GAS, SNG-GARCH and SNIG-GARCH, SIG-GAS, in terms of prediction accuracy, tend to dominate other competitors at the daily, weekly and monthly horizons, respectively. However, I also care about whether one model indeed makes predictions that are statistically different from the others. Therefore, I apply the DM test to inspect whether forecasting errors from a model are statistically different from those from other candidate models in a pairwise fashion. From Panel A of Table B.1, it is clear that the TLS-GAS model has significant DM statistics (ranging between -7.26 and -4.69) against all the other models. This means that the predictions of realised volatility made by the TLS-GAS model are not only more accurate than those made by other candidate models but also statis-

tically better than other predictions. Additionally, I implement CW tests in order to inspect the relative outperformance between any two models. Consistently, it can be observed that the TLS-GAS model, with CW statistics ranging between 3.56 and 8.29, tends to have significantly lower comparative forecasting errors against other models, indicating its dominance at the daily horizon again.

Similarly, at the weekly horizon, SNG-GARCH and SNIG-GARCH models, while dominating other models in prediction accuracy, are found to make statistically different forecasts compared to other models. The pairwise DM statistics of these two model-based predictions against others are highly significant in most cases (both predictions tend to be statistically different from others in 10 out of 15 scenarios). The CW test results suggest that SNG-GARCH and SNIG-GARCH outperform other models with relatively lower prediction errors in most cases (13 out of 15 cases for both models).

At the monthly horizon, the SIG-GAS model, with relatively lower mean average prediction error and R^2_{os}, is found to produce statistically different predictions compared to others in 13 out of 15 cases (GAS and SNIG-GAS are exceptions). In fact, it can be found from the CW statistics that SIG-GAS tend to have relatively inferior performance, indicated by higher forecasting errors, compared to GAS and SNIG-GAS.

3.5.2 Crisis period

With the above evaluations based on the full sample, GARCH models are found to outperform at shorter forecasting horizons, while they tend to have a similar performance to GAS at the monthly horizon. I am also interested in how GARCH and GAS perform during sub-periods in which the market experiences greater uncertainty. Therefore, utilising a sub-sample during the 2007-2009 financial crisis, I

inspect how GARCH and GAS models perform during market turmoil. From Table 3.9, it can be instantly observed that SNW-GAS has the highest R^2 (68.32%), followed by SNG-GAS (63.07%) and SG-GAS (62.69%), indicating superiority of these models in describing the movements of market returns during crisis time. The outperformance of SNW-GAS continues at the weekly horizon with R^2 of 53.23% and disappears at the monthly horizon. The regression results are consistent with corresponding MAE values and Quasi-loglikelihood values reported in Table 3.10 and 3.11, where SNW-GAS has the lowest forecasting errors at both daily (6.05) and weekly (7.01) horizons. Note that the forecasting errors of the SNW-GARCH model at these horizons are 6.99 and 7.37 at daily and weekly horizons, which are consistently higher than SNW-GAS.

Table 3.9: Mincer-Zarnowitz regressions over horizons - crisis

Panel A: GARCH modles									
	1 day			5 days			22 days		
	α	β	R^2	α	β	R^2	α	β	R^2
Normal	0.0061	0.8988***	0.5932	0.0270**	0.8579***	0.4923	0.1210***	0.5978***	0.1668
SG	0.0066	0.9010***	0.5949	0.0274**	0.8618***	0.4985	0.1222***	0.6012***	0.1715
SNG	0.0058	0.8998***	0.6006	0.0249**	0.8632***	0.5037	0.1139***	0.6218***	0.1808
SIG	0.0066	0.9010***	0.5961	0.0273**	0.8620***	0.4979	0.1167***	0.6265***	0.1835
SNIG	0.0063	0.8995***	0.6027	0.0268**	0.8610***	0.5023	0.1174***	0.6198***	0.1806
SNW	0.0022	0.9016***	0.6164	0.0198*	0.8736***	0.5122	0.1014***	0.6766***	0.1846
TLS	0.0055	0.8781***	0.5957	0.0240**	0.8281***	0.4956	0.1093***	0.5661***	0.1821
LOG	0.0116	0.8554***	0.5762	0.0317***	0.8038***	0.4755	0.1156***	0.5553***	0.1766
Panel B: GAS modles									
	1 day			5 days			22 days		
	α	β	R^2	α	β	R^2	α	β	R^2
Normal	-0.0143	1.0430***	0.5871	-0.0003	1.0188***	0.4857	0.0747***	0.7962***	0.1762
SG	-0.0025	1.2743***	0.6269	0.0110	1.2383***	0.4793	0.0554***	1.1063***	0.1777
SNG	−0.0173*	1.3676***	0.6307	-0.0146	1.4042***	0.4711	0.0136	1.3923***	0.1365
SIG	-0.0154	1.0506***	0.5895	-0.0035	1.0336***	0.4921	0.0729***	0.8099***	0.1744
SNIG	-0.0019	0.9745***	0.5700	0.0101	0.9518***	0.4762	0.0815***	0.7404***	0.1737
SNW	−0.0474***	1.1762***	0.6832	−0.0515***	1.2616***	0.5323	0.0607**	0.9281***	0.0879
TLS	0.0154	0.9534***	0.5639	0.0416***	0.7426***	0.4335	0.1263***	0.4475***	0.1498
LOG	0.0269***	0.7855***	0.5345	0.0412***	0.7390***	0.4476	0.1312***	0.4307***	0.1397

Notes: This table reports results from Mincer-Zarnowitz regressions over forecasting horizons from one day to one month. A sub-sample during the 2007-2008 financial crisis is adopted to obtain predictions using parametric models dynamically, and realised volatility is regressed against these predictions. For each regression, intercept α, slope β and goodness of fit R^2 are reported. Statistics of GARCH models are reported in panel A, and those of GAS models are reported in panel B.

Table 3.10: MAE values over horizons - crisis

Panel A: GARCH modles			
	Daily	Weekly	Monthly
Normal	7.0330	7.4891	9.6913
SG	6.9538	7.3428	9.5196
SNG	6.9463	7.3116	9.2683
SIG	6.9338	7.3459	9.2982
SNIG	6.8974	7.2754	9.3493
SNW	6.9865	7.3686	8.9788
TLS	7.2678	7.7690	9.3518
LOG	7.3273	7.8098	9.4937
Panel B: GAS models			
	Daily	Weekly	Monthly
Normal	6.5771	7.1804	8.9468
SG	7.1151	7.7145	9.4614
SNG	7.2443	8.0477	9.8829
SIG	6.5459	7.1053	8.9268
SNIG	6.7938	7.3437	8.9867
SNW	6.0468	7.0128	9.5403
TLS	6.7364	8.6847	10.8902
LOG	7.7809	8.5491	10.7816

Notes: This table reports the mean absolute errors (MAE) between each observation-driven model-based prediction and realised volatility over a sub-sample that covers the 2007-2008 financial crisis. The MAEs are computed at horizons from one day to one month, and the MAE values are multiplied by 100 for easier comparisons. MAEs of GARCH models are reported in panel A, and those of GAS models are reported in panel B.

Table 3.11: Quasi-likelihood values over horizons - crisis

Panel A: GARCH modles			
	Daily	Weekly	Monthly
Normal	-0.5229	-0.5063	-0.3573
SG	-0.5237	-0.5083	-0.3607
SNG	-0.5240	-0.5101	-0.3973
SIG	-0.5239	-0.5087	-0.3826
SNIG	-0.5247	-0.5098	-0.3767
SNW	-0.5240	-0.5075	0.1458
TLS	-0.5215	-0.5077	-0.4257
LOG	-0.5203	-0.5068	-0.4117
Panel B: GAS modles			
	Daily	Weekly	Monthly
Normal	-0.5248	-0.5071	-0.4240
SG	-0.4921	-0.4686	-0.3635
SNG	-0.4846	-0.4621	-0.3487
SIG	-0.5253	-0.5086	-0.4231
SNIG	-0.5216	-0.5056	-0.4263
SNW	-0.5333	-0.5121	-0.3919
TLS	-0.5163	-0.4857	-0.4181
LOG	-0.5149	-0.4957	-0.4151

Notes: This table reports the quasi-likelihood of each model-based prediction against realised volatility over a subsample that covers the 2007-2008 financial crisis. The quasi-likelihood is robust to noise contained in volatility and is computed at horizons from one day to one month. More negative quasi-likelihood values indicate superior out-of-sample performance. The statistics of GARCH models are reported in panel A and those of GAS models are reported in panel B.

The pairwise R^2_{os}, together with corresponding Clark-West test statistics are reported in Table 3.12 and 3.13 suggest that SNW-GAS dominate all the other models at daily and weekly horizons with positive R^2_{os} against all the other candidates. From the DM test results in Table B.4, it is obvious that SNW-GAS also makes statistically different predictions. However, from Table B.5, it can be found that SNW-GAS-based predictions are not significantly different from others in most cases. The relative outperformance of SNW-GAS is confirmed again by the CW statistics. At daily and weekly horizons, CW statistics of SNW-GAS are found to be highly significant over all the other models. It is also noteworthy that, during the crisis period, the TLS-GAS model has a performance only second to SNW-GAS. This is evidenced by the fact that the R^2_{os} of SNW-GAS against TLS-GAS is the lowest (0.10).

Table 3.12: Out-of-sample R^2 at daily horizon - crisis

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	0.00	1.13	1.23	1.41	1.93	0.66	-3.34	-4.18	6.48	-1.17	-3.00	6.93	3.40	14.02	4.22	-10.63
	NaN	**1.96**	**2.49**	**2.05**	**2.88**	**2.38**	**-0.60**	**-1.11**	**1.58**	**3.58**	**3.59**	**1.64**	**0.53**	**3.47**	**1.97**	**-2.25**
SG	-1.14	0.00	0.11	0.29	0.81	-0.47	-4.51	-5.37	5.42	-2.32	-4.18	5.87	2.30	13.04	3.13	-11.89
	-0.65	**NaN**	**1.22**	**0.93**	**2.20**	**1.87**	**-1.06**	**-1.32**	**1.39**	**3.47**	**3.49**	**1.45**	**0.34**	**3.39**	**1.82**	**-2.43**
SNG	-1.25	-0.11	0.00	0.18	0.70	-0.58	-4.63	-5.49	5.31	-2.43	-4.29	5.76	2.19	12.95	3.02	-12.02
	-1.42	**-0.51**	**NaN**	**-0.12**	**1.32**	**1.80**	**-1.50**	**-1.47**	**1.35**	**3.45**	**3.47**	**1.42**	**0.33**	**3.38**	**1.78**	**-2.62**
SIG	-1.43	-0.29	-0.18	0.00	0.53	-0.76	-4.82	-5.67	5.14	-2.61	-4.48	5.59	2.02	12.79	2.85	-12.22
	-0.73	**-0.04**	**1.06**	**NaN**	**2.52**	**1.88**	**-1.36**	**-1.42**	**1.33**	**3.42**	**3.45**	**1.39**	**0.32**	**3.40**	**1.76**	**-2.52**
SNIG	-1.97	-0.82	-0.71	-0.53	0.00	-1.29	-5.37	-6.23	4.64	-3.16	-5.03	5.10	1.50	12.33	2.33	-12.81
	-1.97	**-1.61**	**-0.27**	**-1.57**	**NaN**	**1.27**	**-1.99**	**-1.69**	**1.24**	**3.42**	**3.45**	**1.31**	**0.21**	**3.35**	**1.70**	**-2.74**
SNW	-0.67	0.47	0.58	0.75	1.27	0.00	-4.03	-4.88	5.86	-1.84	-3.69	6.31	2.76	13.45	3.58	-11.37
	-1.30	**-0.96**	**-0.81**	**-0.92**	**-0.38**	**NaN**	**-2.05**	**-1.58**	**1.27**	**3.54**	**3.55**	**1.35**	**0.19**	**3.33**	**1.79**	**-2.76**
TLS	3.23	4.32	4.42	4.59	5.10	3.87	0.00	-0.82	9.50	2.10	0.32	9.93	6.52	16.80	7.31	-7.06
	1.73	**1.99**	**2.17**	**2.29**	**2.81**	**2.48**	**NaN**	**-0.92**	**1.90**	**3.53**	**3.54**	**1.94**	**1.13**	**3.43**	**2.16**	**-2.07**
LOG	4.02	5.10	5.20	5.37	5.87	4.65	0.81	0.00	10.24	2.90	1.13	10.66	7.28	17.48	8.06	-6.19
	1.49	**1.59**	**1.59**	**1.64**	**1.81**	**1.70**	**1.21**	**NaN**	**1.94**	**3.45**	**3.45**	**1.96**	**1.38**	**3.25**	**2.08**	**-0.18**
Normal	-6.93	-5.73	-5.61	-5.42	-4.87	-6.22	-10.50	-11.41	0.00	-8.18	-10.14	0.47	-3.29	8.06	-2.42	-18.30
	0.54	**0.62**	**0.83**	**0.66**	**0.88**	**1.24**	**0.54**	**0.05**	**NaN**	**3.67**	**3.63**	**2.05**	**-1.13**	**3.16**	**1.85**	**-0.17**
SG	1.15	2.27	2.37	2.55	3.06	1.81	-2.15	-2.98	7.56	0.00	-1.82	8.00	4.52	15.01	5.32	-9.36
	4.46	**4.44**	**4.43**	**4.43**	**4.40**	**4.30**	**4.43**	**4.31**	**4.81**	**NaN**	**-0.81**	**4.82**	**4.48**	**4.60**	**4.49**	**4.49**
SNG	2.92	4.01	4.11	4.29	4.79	3.56	-0.32	-1.15	9.21	1.78	0.00	9.64	6.22	16.53	7.01	-7.41
	4.20	**4.18**	**4.18**	**4.17**	**4.16**	**4.09**	**4.20**	**4.12**	**4.40**	**1.98**	**NaN**	**4.41**	**4.15**	**4.61**	**4.06**	**4.22**
SIG	-7.44	-6.23	-6.12	-5.93	-5.37	-6.73	-11.03	-11.94	-0.48	-8.70	-10.67	0.00	-3.79	7.62	-2.91	-18.87
	0.49	**0.56**	**0.78**	**0.61**	**0.83**	**1.21**	**0.52**	**0.05**	**-0.48**	**3.68**	**3.64**	**NaN**	**-1.28**	**3.14**	**1.74**	**-0.16**
SNIG	-3.52	-2.36	-2.24	-2.06	-1.53	-2.84	-6.98	-7.85	3.19	-4.73	-6.63	3.65	0.00	11.00	0.84	-14.53
	0.70	**0.80**	**0.96**	**0.82**	**1.01**	**1.29**	**0.57**	**0.11**	**2.20**	**3.92**	**3.87**	**2.17**	**NaN**	**3.15**	**2.51**	**-0.16**
SNW	-16.31	-15.00	-14.88	-14.67	-14.07	-15.54	-20.19	-21.18	-8.77	-17.67	-19.80	-8.25	-12.35	0.00	-11.41	-28.68
	0.98	**0.99**	**1.16**	**1.05**	**1.17**	**1.43**	**1.07**	**0.73**	**0.00**	**0.82**	**0.87**	**-0.03**	**0.15**	**NaN**	**0.93**	**0.73**
TLS	-4.40	-3.23	-3.11	-2.93	-2.39	-3.71	-7.89	-8.77	2.37	-5.62	-7.54	2.83	-0.85	10.24	0.00	-15.50
	2.32	**2.39**	**2.41**	**2.39**	**2.48**	**2.64**	**2.26**	**1.83**	**2.38**	**3.24**	**3.28**	**2.39**	**2.00**	**3.66**	**NaN**	**1.83**
LOG	9.61	10.63	10.73	10.89	11.35	10.21	6.59	5.83	15.47	8.56	6.90	15.87	12.69	22.29	13.42	0.00
	1.80	**1.82**	**1.85**	**1.84**	**1.88**	**1.95**	**1.62**	**1.40**	**2.18**	**3.46**	**3.51**	**2.21**	**1.90**	**3.18**	**2.32**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions at daily forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using a sub-sample that covers the 2007-2008 financial crisis. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

Table 3.13: Out-of-sample R^2 at weekly horizon - crisis

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	0.00	1.95	2.37	1.91	2.85	1.61	-3.74	-4.28	4.12	-3.01	-7.46	5.13	1.94	6.36	-15.96	-14.15
	NaN	**2.42**	**2.86**	**2.22**	**2.67**	**2.74**	**-0.17**	**-0.69**	**1.39**	**2.30**	**2.31**	**1.42**	**0.58**	**2.78**	**-0.68**	**-1.56**
SG	-1.99	0.00	0.42	-0.04	0.92	-0.35	-5.80	-6.36	2.21	-5.06	-9.60	3.23	-0.01	4.49	-18.27	-16.43
	-0.55	**NaN**	**1.49**	**0.73**	**1.36**	**1.87**	**-0.74**	**-1.04**	**1.08**	**2.12**	**2.16**	**1.12**	**0.23**	**2.61**	**-1.01**	**-1.92**
SNG	-2.43	-0.43	0.00	-0.47	0.49	-0.78	-6.25	-6.81	1.79	-5.51	-10.07	2.82	-0.44	4.09	-18.78	-16.92
	-1.00	**-0.06**	**NaN**	**-0.07**	**0.85**	**1.59**	**-1.24**	**-1.30**	**1.04**	**2.12**	**2.16**	**1.08**	**0.21**	**2.54**	**-1.37**	**-2.26**
SIG	-1.95	0.04	0.47	0.00	0.96	-0.31	-5.76	-6.32	2.25	-5.02	-9.55	3.28	0.03	4.53	-18.23	-16.38
	-0.10	**1.37**	**1.75**	**NaN**	**1.62**	**1.79**	**-0.72**	**-1.06**	**1.12**	**2.10**	**2.13**	**1.15**	**0.29**	**2.59**	**-0.91**	**-1.94**
SNIG	-2.94	-0.93	-0.50	-0.97	0.00	-1.28	-6.78	-7.35	1.31	-6.03	-10.61	2.34	-0.94	3.61	-19.37	-17.51
	-0.79	**0.00**	**0.99**	**0.00**	**NaN**	**1.62**	**-1.10**	**-1.25**	**1.01**	**2.07**	**2.11**	**1.06**	**0.22**	**2.55**	**-1.21**	**-2.12**
SNW	-1.64	0.35	0.77	0.31	1.27	0.00	-5.43	-5.99	2.55	-4.69	-9.21	3.57	0.34	4.83	-17.86	-16.02
	-0.84	**-0.41**	**-0.10**	**-0.44**	**-0.17**	**NaN**	**-1.30**	**-1.23**	**0.95**	**2.22**	**2.24**	**1.02**	**0.07**	**2.54**	**-1.75**	**-2.26**
TLS	3.60	5.49	5.89	5.45	6.35	5.15	0.00	-0.53	7.58	0.70	-3.59	8.54	5.47	9.73	-11.79	-10.04
	1.69	**1.98**	**2.20**	**2.04**	**2.35**	**2.07**	**NaN**	**-0.62**	**1.76**	**2.52**	**2.53**	**1.76**	**1.20**	**2.69**	**-0.24**	**-2.35**
LOG	4.11	5.98	6.38	5.94	6.84	5.65	0.52	0.00	8.06	1.22	-3.05	9.02	5.97	10.20	-11.20	-9.47
	1.55	**1.68**	**1.69**	**1.69**	**1.79**	**1.68**	**1.22**	**NaN**	**1.80**	**2.44**	**2.45**	**1.80**	**1.38**	**2.57**	**0.72**	**-0.44**
Normal	-4.30	-2.26	-1.83	-2.30	-1.32	-2.62	-8.20	-8.77	0.00	-7.44	-12.08	1.05	-2.27	2.33	-20.95	-19.06
	0.88	**1.05**	**1.25**	**1.06**	**1.15**	**1.51**	**0.84**	**0.39**	**NaN**	**1.62**	**1.52**	**1.98**	**0.01**	**2.51**	**0.66**	**0.21**
SG	2.92	4.82	5.22	4.78	5.69	4.48	-0.71	-1.24	6.92	0.00	-4.32	7.90	4.81	9.10	-12.58	-10.82
	3.82	**3.79**	**3.83**	**3.78**	**3.77**	**3.78**	**3.86**	**3.77**	**4.21**	**NaN**	**-1.68**	**4.25**	**4.04**	**3.94**	**3.92**	**3.89**
SNG	6.94	8.76	9.15	8.72	9.60	8.44	3.46	2.96	10.78	4.14	0.00	11.71	8.75	12.86	-7.92	-6.23
	3.72	**3.71**	**3.73**	**3.69**	**3.69**	**3.68**	**3.78**	**3.74**	**4.01**	**2.56**	**NaN**	**4.04**	**3.84**	**4.04**	**3.82**	**3.81**
SIG	-5.40	-3.34	-2.90	-3.39	-2.39	-3.71	-9.34	-9.92	-1.06	-8.57	-13.26	0.00	-3.36	1.30	-22.23	-20.32
	0.60	**0.75**	**0.97**	**0.76**	**0.89**	**1.25**	**0.66**	**0.25**	**-0.33**	**1.43**	**1.35**	**NaN**	**-0.72**	**2.24**	**0.46**	**0.07**
SNIG	-1.98	0.01	0.44	-0.03	0.93	-0.34	-5.79	-6.35	2.22	-5.05	-9.59	3.25	0.00	4.50	-18.26	-16.41
	0.64	**0.84**	**1.02**	**0.86**	**0.95**	**1.31**	**0.50**	**0.11**	**1.60**	**2.19**	**2.13**	**1.72**	**NaN**	**2.60**	**0.25**	**-0.26**
SNW	-6.79	-4.71	-4.26	-4.75	-3.74	-5.07	-10.78	-11.36	-2.39	-10.01	-14.76	-1.32	-4.72	0.00	-23.84	-21.91
	1.94	**1.99**	**2.10**	**1.99**	**2.07**	**2.10**	**2.04**	**1.84**	**1.39**	**-0.21**	**-0.52**	**1.37**	**1.37**	**NaN**	**1.74**	**1.75**
TLS	13.77	15.45	15.81	15.42	16.23	15.15	10.54	10.07	17.32	11.17	7.34	18.19	15.44	19.25	0.00	1.56
	3.07	**3.23**	**3.28**	**3.25**	**3.33**	**3.33**	**3.01**	**2.63**	**3.02**	**3.31**	**3.26**	**3.00**	**2.79**	**3.41**	**NaN**	**2.15**
LOG	12.40	14.11	14.47	14.07	14.90	13.81	9.13	8.65	16.01	9.76	5.87	16.89	14.10	17.97	-1.59	0.00
	3.26	**3.46**	**3.56**	**3.51**	**3.60**	**3.45**	**3.33**	**2.68**	**2.99**	**3.16**	**3.13**	**2.96**	**2.73**	**3.28**	**2.22**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions at the weekly forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using a sub-sample that covers the 2007-2008 financial crisis. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

Table 3.14: Out-of-sample R^2 at monthly horizon - crisis

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	0.00	1.77	4.37	4.06	3.53	7.35	3.50	2.04	7.68	2.37	-1.98	7.89	7.27	1.56	-12.37	-11.25
	NaN	**1.25**	**3.37**	**3.00**	**2.72**	**4.14**	**2.42**	**2.77**	**4.20**	**1.83**	**1.66**	**3.82**	**4.41**	**1.92**	**1.95**	**0.79**
SG	-1.80	0.00	2.64	2.33	1.79	5.68	1.76	0.27	6.02	0.61	-3.82	6.23	5.60	-0.22	-14.40	-13.26
	1.93	**NaN**	**2.60**	**2.13**	**2.07**	**3.35**	**2.22**	**2.54**	**3.51**	**1.63**	**1.54**	**3.27**	**3.56**	**1.86**	**1.90**	**0.70**
SNG	-4.56	-2.71	0.00	-0.32	-0.87	3.12	-0.90	-2.43	3.47	-2.08	-6.63	3.68	3.04	-2.93	-17.50	-16.33
	0.14	**-1.07**	**NaN**	**0.74**	**0.59**	**3.08**	**1.62**	**1.98**	**3.16**	**1.33**	**1.31**	**2.79**	**3.46**	**1.52**	**1.35**	**0.09**
SIG	-4.23	-2.38	0.32	0.00	-0.55	3.44	-0.58	-2.10	3.78	-1.76	-6.29	3.99	3.35	-2.60	-17.12	-15.95
	1.22	**-0.48**	**2.91**	**NaN**	**1.64**	**3.55**	**2.03**	**2.26**	**3.64**	**1.35**	**1.32**	**3.23**	**3.87**	**1.68**	**1.76**	**0.48**
SNIG	-3.66	-1.82	0.87	0.55	0.00	3.96	-0.03	-1.54	4.31	-1.20	-5.71	4.52	3.88	-2.04	-16.48	-15.32
	1.22	**-0.08**	**2.60**	**1.79**	**NaN**	**3.07**	**2.15**	**2.46**	**3.42**	**1.37**	**1.33**	**3.10**	**3.70**	**1.66**	**1.78**	**0.53**
SNW	-7.94	-6.02	-3.22	-3.56	-4.13	0.00	-4.15	-5.73	0.36	-5.38	-10.07	0.58	-0.09	-6.25	-21.29	-20.08
	-1.33	**-1.99**	**-0.68**	**-1.20**	**-1.00**	**NaN**	**0.86**	**1.04**	**2.60**	**0.86**	**0.84**	**2.23**	**2.61**	**0.99**	**0.89**	**-0.14**
TLS	-3.63	-1.79	0.89	0.57	0.03	3.99	0.00	-1.52	4.33	-1.17	-5.68	4.54	3.90	-2.02	-16.45	-15.29
	0.99	**0.72**	**1.27**	**1.21**	**1.26**	**1.59**	**NaN**	**1.34**	**2.14**	**1.86**	**1.87**	**2.03**	**1.98**	**1.75**	**0.12**	**-1.41**
LOG	-2.08	-0.27	2.37	2.06	1.52	5.42	1.49	0.00	5.76	0.34	-4.10	5.97	5.34	-0.49	-14.71	-13.57
	1.35	**1.10**	**1.58**	**1.50**	**1.55**	**1.92**	**1.80**	**NaN**	**2.41**	**1.90**	**1.89**	**2.27**	**2.35**	**1.89**	**1.22**	**-0.43**
Normal	-8.32	-6.40	-3.59	-3.93	-4.50	-0.36	-4.53	-6.11	0.00	-5.75	-10.46	0.22	-0.45	-6.63	-21.72	-20.51
	-0.87	**-0.99**	**-0.07**	**-0.38**	**-0.30**	**0.48**	**0.97**	**0.90**	**NaN**	**0.05**	**-0.08**	**0.36**	**1.05**	**-0.21**	**1.08**	**0.24**
SG	-2.43	-0.61	2.04	1.73	1.19	5.10	1.16	-0.34	5.44	0.00	-4.45	5.65	5.02	-0.83	-15.10	-13.95
	2.70	**2.47**	**2.84**	**2.66**	**2.62**	**2.94**	**3.27**	**3.23**	**3.76**	**NaN**	**-1.85**	**3.83**	**3.65**	**3.97**	**3.44**	**3.28**
SNG	1.94	3.68	6.22	5.92	5.40	9.15	5.37	3.94	9.47	4.26	0.00	9.67	9.07	3.47	-10.19	-9.09
	2.93	**2.80**	**3.04**	**2.92**	**2.90**	**3.03**	**3.39**	**3.39**	**3.70**	**2.70**	**NaN**	**3.74**	**3.59**	**4.69**	**3.50**	**3.48**
SIG	-8.56	-6.64	-3.83	-4.16	-4.73	-0.58	-4.76	-6.35	-0.22	-5.99	-10.71	0.00	-0.67	-6.87	-21.99	-20.78
	-0.70	**-0.86**	**0.02**	**-0.29**	**-0.18**	**0.51**	**1.03**	**0.96**	**1.42**	**-0.09**	**-0.23**	**NaN**	**0.88**	**-0.39**	**1.14**	**0.35**
SNIG	-7.84	-5.93	-3.13	-3.47	-4.03	0.09	-4.06	-5.64	0.44	-5.28	-9.97	0.67	0.00	-6.16	-21.18	-19.97
	-1.67	**-1.69**	**-0.87**	**-1.06**	**-0.88**	**-0.07**	**0.40**	**0.38**	**0.80**	**0.49**	**0.41**	**0.22**	**NaN**	**0.18**	**0.49**	**-0.38**
SNW	-1.58	0.22	2.85	2.54	2.00	5.89	1.98	0.49	6.22	0.83	-3.59	6.43	5.80	0.00	-14.15	-13.01
	2.25	**2.12**	**2.45**	**2.35**	**2.33**	**2.40**	**2.78**	**2.87**	**2.95**	**1.78**	**-0.84**	**2.95**	**2.80**	**NaN**	**2.67**	**2.61**
TLS	11.01	12.59	14.89	14.62	14.15	17.55	14.13	12.82	17.85	13.12	9.25	18.03	17.48	12.40	0.00	1.00
	2.71	**2.78**	**2.97**	**2.96**	**3.00**	**2.94**	**3.30**	**3.18**	**2.99**	**2.90**	**2.82**	**2.91**	**2.92**	**2.57**	**NaN**	**2.49**
LOG	10.11	11.71	14.04	13.76	13.29	16.72	13.26	11.95	17.02	12.24	8.34	17.20	16.65	11.51	-1.01	0.00
	2.53	**2.55**	**2.74**	**2.71**	**2.79**	**2.68**	**3.08**	**2.95**	**2.75**	**2.66**	**2.60**	**2.67**	**2.67**	**2.37**	**2.17**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions at the monthly forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using a sub-sample that covers the 2007-2008 financial crisis. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

At the monthly horizon, it is observed again that the SIG-GAS model tends to dominate all the other models with the lowest MAE of 8.93. The positive pairwise R^2_{os} values against all the other models once more indicate the dominance of SIG-GAS at the monthly horizon. It is indicated by the DM statistics that predictions from SIG-GAS are significantly different from 11 out of 15 model-based predictions. Finally, from the CW statistics, it can be observed that SIG-GAS-based predictions tend to have lower forecasting errors than 13 out of 15 parametric model-based predictions (the exceptions are again GAS and SNIG-GAS, consistent with the results in the full sample).

It seems that we still have not obtained sufficient evidence to support that GARCH overwhelmingly dominates GAS across all horizons, or the opposite, during the financial crisis period. However, it is clear that during market turmoil, the GAS models outperform their GARCH counterparts in most cases at daily horizons, and they tend to have comparable performance to GARCH models at weekly and monthly horizons. It is also documented that, at the daily horizon, for models with specific innovations like Gaussian, SIG, SNIG, SNW and TLS, the predictive power is greatly improved from GARCH to GAS. This underlines the importance of correctly choosing the innovation term's distribution in the process of volatility modelling and forecasting. If the error term is better specified, GAS may have greater potential to more accurately predict volatility when uncertainty in the market is high.

3.5.3 Comparison to IV

As introduced in Section 3.1, quite a few studies make comparisons between the performance of parametric model-based volatility forecast and implied volatility. To this end, in this chapter, I also examine how the observation-driven models perform in volatility forecasting compared to implied volatility. The implied volatility is calculated using options data between 04/01/1996 and 31/12/2017, therefore, a

different sample period is applied for making comparisons between model-based volatility forecast and implied volatility. I depict comparisons between MFIV and GAS model-based volatility forecast in Figure 3.4. The comparison between MFIV and GAS models is plotted in Figure 3.5.

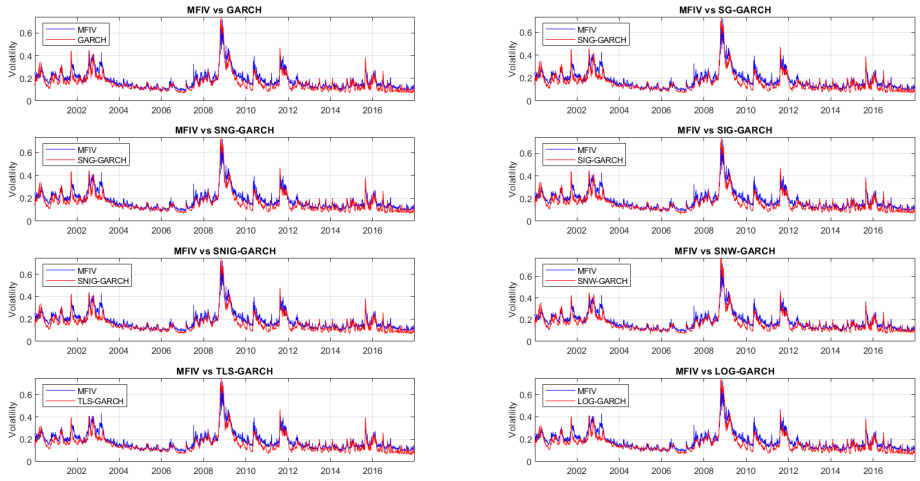

Figure 3.4: MFIV vs GARCH

Notes: This figure plots the comparison between MFIV and predictions made by all GARCH models. At each point in time, MFIV is computed as 30 days ahead prediction in a model-free sense, and GARCH models are fitted on a rolling window basis using 5 years' observations.

From the Mincer-Zarnowitz regressions in Table 3.15, it is straightforward to see the domination of IV ($R^2 = 0.66$) and MFIV ($R^2 = 0.64$) in describing the movement of realised volatility. The superiority of IV and MFIV becomes more significant at longer forecasting horizons. For example, at the daily horizon, the R^2 of MFIV is around 4% higher than that of GARCH, while at the monthly horizon, this number increases to 10%. This indicates that, at longer forecasting horizons, implied volatility is a better candidate for capturing the dynamic trends of realised volatility. However, in terms of prediction accuracy, it can be another story. IV and MFIV are not necessarily the best performers.

The MAE values in Table 3.16 indicate that TLS-GAS, with a prediction error of 3.88, is still the best performer among all competitors at the daily horizon. Generally speaking, it can be observed that parametric models tend to have higher forecasting

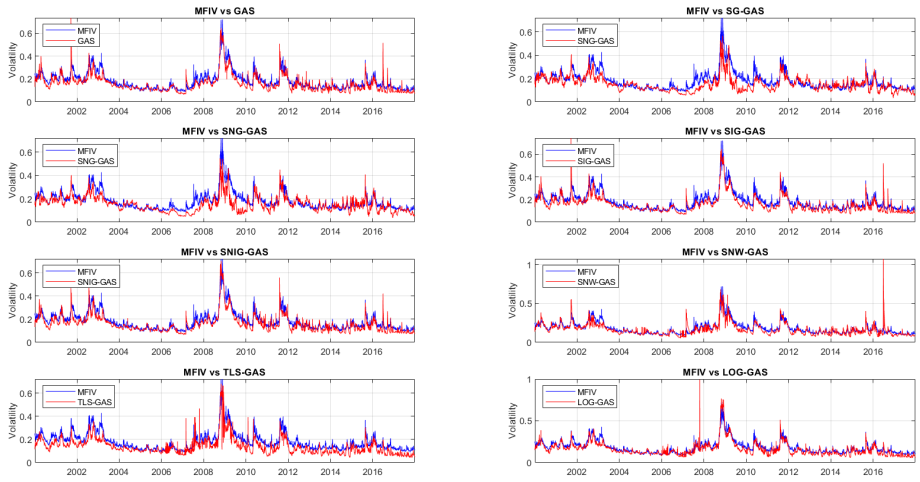

Figure 3.5: MFIV vs GAS

Notes: This figure plots the comparison between MFIV and predictions made by all GAS models. At each point in time, MFIV is computed as 30 days ahead prediction in a model-free sense, and GAS models are fitted on a rolling window basis using 5 years' observations.

Table 3.15: Mincer-Zarnowitz regressions over horizons - including IV

	1 day			5 days			22 days		
	α	β	R^2	α	β	R^2	α	β	R^2
Panel A: Implied volatility									
IV	-0.0365^{***}	1.0177^{***}	0.6580	-0.0195^{***}	0.9172^{***}	0.5341	0.0133^{***}	0.7228^{***}	0.3311
MFIV	-0.0430^{***}	0.9861^{***}	0.6401	-0.0251^{***}	0.8876^{***}	0.5184	0.0091^{***}	0.6989^{***}	0.3212
Panel B: GARCH modles									
	1 day			5 days			22 days		
	α	β	R^2	α	β	R^2	α	β	R^2
Normal	-0.0092^{***}	0.9200^{***}	0.6090	0.0097^{***}	0.8826^{***}	0.4986	0.0673^{***}	0.6800^{***}	0.2214
SG	-0.0096^{***}	0.9200^{***}	0.6096	0.0090^{***}	0.8836^{***}	0.4991	0.0667^{***}	0.6804^{***}	0.2187
SNG	-0.0086^{***}	0.9201^{***}	0.6120	0.0095^{***}	0.8826^{***}	0.5046	0.0644^{***}	0.6975^{***}	0.2353
SIG	-0.0093^{***}	0.9209^{***}	0.6094	0.0096^{***}	0.8830^{***}	0.4981	0.0667^{***}	0.6826^{***}	0.2225
SNIG	-0.0091^{***}	0.9227^{***}	0.6135	0.0093^{***}	0.8826^{***}	0.5048	0.0619^{***}	0.7194^{***}	0.2432
SNW	-0.0115^{***}	0.9077^{***}	0.6152	0.0070^{***}	0.8520^{***}	0.5149	0.0478^{***}	0.7396^{***}	0.2873
TLS	-0.0076^{***}	0.9099^{***}	0.6087	0.0100^{***}	0.8615^{***}	0.4954	0.0662^{***}	0.6299^{***}	0.2280
LOG	-0.0059^{***}	0.9002^{***}	0.6020	0.0124^{***}	0.8534^{***}	0.4905	0.0680^{***}	0.6361^{***}	0.2301
Panel C: GAS models									
	1 day			5 days			22 days		
	α	β	R^2	α	β	R^2	α	β	R^2
Normal	-0.0178^{***}	0.9787^{***}	0.5784	-0.0197^{***}	0.9937^{***}	0.4728	-0.0219^{***}	1.0090^{***}	0.2620
SG	-0.0328^{***}	1.1205^{***}	0.5192	-0.0234^{***}	1.0601^{***}	0.3924	0.0002	0.9066^{***}	0.1700
SNG	-0.0235^{***}	1.0541^{***}	0.5056	-0.0131^{***}	0.9792^{***}	0.3623	0.0123^{***}	0.7933^{***}	0.1410
SIG	-0.0189^{***}	0.9897^{***}	0.5815	-0.0213^{***}	1.0067^{***}	0.4786	-0.0287^{***}	1.0557^{***}	0.2709
SNIG	-0.0157^{***}	0.9655^{***}	0.5788	-0.0155^{***}	0.9637^{***}	0.4770	0.0182^{***}	0.7473^{***}	0.2039
SNW	-0.0131^{***}	0.9054^{***}	0.5573	-0.0128^{***}	0.9088^{***}	0.4231	-0.0087^{*}	0.8936^{***}	0.1877
TLS	-0.0012	0.9936^{***}	0.5822	0.0043^{*}	0.8177^{***}	0.4747	0.0247^{***}	0.6765^{***}	0.2616
LOG	-0.0006	0.8575^{***}	0.5792	0.0045^{*}	0.8198^{***}	0.4782	0.0238^{***}	0.6884^{***}	0.2635

Notes: This table reports results from Mincer-Zarnowitz regressions over forecasting horizons from one day to one month. A sample of option prices between 01/1996 and 12/2017 is used to estimate IV and MFIV. Corresponding historical DJIA returns are adopted to obtain predictions using parametric models dynamically. realised volatility is regressed against these predictions. For each regression, intercept α, slope β and goodness of fit R^2 are reported. Statistics of IV and MFIV are reported in panel A, statistics of GARCH predictions are reported in panel B, and statistics of GAS predictions are reported in panel C.

accuracy compared to implied volatility. At the daily horizon, the former models are found to have forecasting errors between 3.88 and 4.98, which are significantly lower than those from IV and MFIV (with MAEs of 4.95 and 5.86 respectively). It is also the case for IV and MFIV at weekly and monthly forecasting horizons, where both implied volatility are consistently found to have prediction errors higher than those from GARCH models.

The R^2_{os} values in Table 3.18 confirm the domination of TLS-GAS in terms of forecasting accuracy at the daily horizon, with all positive R^2_{os} values. More importantly, the inferior performance of implied volatility against parametric model-based volatility forecasts is once again verified. It is also worth pointing out that IV seems to have higher forecasting accuracy compared to MFIV across all horizons. This finding is supported by positive R^2_{os} of IV against MFIV (15.55%, 13.70% and 11.25% for daily, weekly and monthly horizons). At weekly and monthly horizons, it turns out that the SNG-GARCH and SNW-GARCH models dominate all the other models, including IV and MFIV. It is noteworthy that the strongest outperformance is found to be between TLS-GAS and MFIV, with R^2_{os} of 0.34. It can be concluded so far that implied volatility tends to have lower forecasting accuracy compared to parametric models at horizons from one day up to one month.

For robustness concerns, I additionally implement DM and CW tests. The results are reported in Table B.7, B.8 and B.9. From the DM tests, it can be easily observed that, at the daily horizon, IV is statistically different from model-based forecasts in 6 out of 16 cases, while MFIV forecasts are statistically different from almost all parametric model-based predictions (The only exception is SNW-GAS). At the weekly horizon, IV is significantly different from model-based forecasts in 12 out of 16 cases, and MFIV is found to be different from all model-based predictions. At the monthly horizon, IV is different from 13 out of 16 competing predictions, and for MFIV, this ratio is 9 out of 16.

The CW test is helpful in identifying the model that has relatively lower forecasting errors. The statistics from CW tests generally indicate consistency with previous findings that implied volatility tends to have lower prediction accuracy. When IV and MFIV are set as benchmarks, it is clear that model-based predictions are consistently found to have lower forecasting errors compared to implied volatility.

Table 3.16: MAE values over horizons - including IV

Panel A: Implied volatility			
	Daily	Weekly	Monthly
IV	4.9517	5.4253	6.2678
MFIV	5.8631	6.2866	7.0622
Panel B: GARCH models			
	Daily	Weekly	Monthly
Normal	4.4901	4.3514	5.8452
SG	4.5093	4.3632	5.8058
SNG	4.4471	4.3212	5.7147
SIG	4.4835	4.3555	5.8180
SNIG	4.4428	4.3249	5.6364
SNW	4.7165	4.5554	5.1585
TLS	4.4976	4.4671	5.6493
LOG	4.4955	4.4268	5.7121
Panel C: GAS modles			
	Daily	Weekly	Monthly
Normal	4.5758	5.0338	5.9337
SG	4.8807	5.2717	6.0610
SNG	4.9770	5.4685	6.4406
SIG	4.5432	4.9961	5.9042
SNIG	4.5936	5.0364	5.9991
SNW	4.9238	5.4893	6.4307
TLS	3.8805	5.1932	6.2912
LOG	4.6620	5.1598	6.2025

Notes: This table reports the mean absolute errors (MAE) between each realised volatility and parametric and non-parametric predictions made by GARCH/GAS and IV/MFIV. The sample adopted ranges between 01/1996 and 12/2017. The MAEs are computed at horizons from one day to one month, and the MAE values are multiplied by 100 for easier comparisons. MAEs of IV and MFIV are reported in panel A, MAEs of GARCH models are reported in panel B and those of GAS models are reported in panel C.

Table 3.17: Quasi-likelihood values over horizons - including IV

Panel A: Implied volatility			
	Daily	Weekly	Monthly
IV	-1.0722	-1.0591	-1.0361
MFIV	-1.0488	-1.0372	-1.0168
Panel B: GARCH models			
	Daily	Weekly	Monthly
Normal	-1.0812	-1.0768	-0.7148
SG	-1.0807	-1.0769	-0.7342
SNG	-1.0820	-1.0776	-0.7551
SIG	-1.0813	-1.0767	-0.6986
SNIG	-1.0821	-1.0773	-0.6341
SNW	-1.0752	-1.0761	-0.4201
TLS	-1.0815	-1.0743	-0.8937
LOG	-1.0817	-1.0753	-0.8516
Panel C: GAS models			
	Daily	Weekly	Monthly
Normal	-1.0764	-1.0582	-1.0282
SG	-1.0613	-1.0446	-1.0113
SNG	-1.0585	-1.0403	-1.0056
SIG	-1.0772	-1.0591	-1.0289
SNIG	-1.0759	-1.0585	-1.0278
SNW	-1.0697	-1.0493	-1.0145
TLS	-1.0899	-1.0592	-1.0250
LOG	-1.0783	-1.0598	-1.0267

Notes: This table reports the quasi-likelihood of IV and MFIV against realised volatility over the full sample. As comparisons, the quasi-likelihood of model-based predictions against RV is also reported. The sample adopted ranges between 01/1996 and 12/2017. The quasi-likelihood is robust to noise contained in volatility and is computed at horizons from one day to one month. More negative quasi-likelihood values indicate superior out-of-sample performance. The statistics of IV and MFIV are reported in panel A, the statistics of GARCH models are reported in panel B and those of GAS models are reported in panel C.

Table 3.18: Out-of-sample R^2 at daily horizon - including IV

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	0.00	-0.43	0.96	0.15	1.05	-5.04	-0.17	-0.12	-1.91	-8.70	-10.84	-1.18	-2.31	-9.66	13.58	-3.83	-10.28	-30.58
	NaN	**-0.51**	**4.54**	**1.50**	**4.78**	**-2.63**	**0.69**	**-0.25**	**0.70**	**4.92**	**4.91**	**0.89**	**-0.17**	**2.43**	**6.14**	**-1.49**	**5.02**	**1.94**
SG	0.43	0.00	1.38	0.57	1.47	-4.59	0.26	0.31	-1.47	-8.24	-10.37	-0.75	-1.87	-9.19	13.94	-3.39	-9.81	-30.02
	3.06	**NaN**	**5.39**	**3.32**	**5.80**	**-2.53**	**1.13**	**0.02**	**0.78**	**4.91**	**4.89**	**0.96**	**-0.01**	**2.40**	**6.25**	**-1.16**	**5.01**	**1.88**
SNG	-0.97	-1.40	0.00	-0.82	0.10	-6.06	-1.14	-1.09	-2.90	-9.75	-11.92	-2.16	-3.30	-10.72	12.74	-4.83	-11.35	-31.84
	-3.49	**-4.44**	**NaN**	**-2.95**	**2.07**	**-4.31**	**-0.77**	**-0.96**	**0.37**	**4.74**	**4.74**	**0.55**	**-0.58**	**2.29**	**5.77**	**-2.21**	**4.83**	**1.70**
SIG	-0.15	-0.58	0.81	0.00	0.91	-5.20	-0.32	-0.27	-2.06	-8.86	-11.01	-1.33	-2.46	-9.82	13.45	-3.98	-10.44	-30.77
	0.64	**-1.41**	**4.06**	**NaN**	**5.34**	**-2.98**	**0.51**	**-0.36**	**0.65**	**4.79**	**4.80**	**0.82**	**-0.20**	**2.38**	**5.97**	**-1.57**	**4.97**	**1.90**
SNIG	-1.06	-1.50	-0.10	-0.92	0.00	-6.16	-1.23	-1.19	-2.99	-9.86	-12.02	-2.26	-3.39	-10.83	12.66	-4.93	-11.45	-31.97
	-3.43	**-4.46**	**-0.26**	**-3.72**	**NaN**	**-4.42**	**-1.35**	**-1.22**	**0.35**	**4.72**	**4.71**	**0.54**	**-0.68**	**2.25**	**5.60**	**-2.48**	**4.74**	**1.55**
SNW	4.80	4.39	5.71	4.94	5.80	0.00	4.64	4.69	2.98	-3.48	-5.52	3.67	2.61	-4.40	17.72	1.15	-4.99	-24.31
	5.50	**5.10**	**7.11**	**5.73**	**6.88**	**NaN**	**4.01**	**2.27**	**2.77**	**5.49**	**5.41**	**2.97**	**1.98**	**2.81**	**7.74**	**1.15**	**5.21**	**1.60**
TLS	0.17	-0.26	1.12	0.31	1.22	-4.87	0.00	0.05	-1.74	-8.52	-10.66	-1.01	-2.13	-9.47	13.72	-3.65	-10.10	-30.36
	1.48	**1.14**	**2.37**	**1.68**	**3.00**	**-0.85**	**NaN**	**-0.53**	**1.24**	**4.48**	**4.47**	**1.39**	**0.42**	**2.43**	**5.28**	**-2.18**	**4.58**	**1.79**
LOG	0.12	-0.31	1.08	0.27	1.17	-4.92	-0.05	0.00	-1.79	-8.57	-10.71	-1.06	-2.18	-9.53	13.68	-3.70	-10.15	-30.42
	1.20	**1.09**	**1.49**	**1.26**	**1.72**	**0.23**	**1.03**	**NaN**	**1.30**	**4.24**	**4.23**	**1.39**	**0.72**	**2.41**	**4.39**	**-0.52**	**4.28**	**2.03**
Normal	1.87	1.45	2.81	2.02	2.91	-3.07	1.71	1.76	0.00	-6.66	-8.77	0.71	-0.39	-7.60	15.19	-1.88	-8.21	-28.13
	2.77	**2.71**	**3.04**	**2.80**	**3.03**	**2.11**	**2.65**	**2.32**	**NaN**	**5.50**	**5.63**	**4.98**	**1.13**	**1.45**	**5.40**	**2.07**	**4.57**	**2.54**
SG	8.00	7.61	8.88	8.14	8.97	3.36	7.85	7.89	6.25	0.00	-1.97	6.92	5.88	-0.88	20.49	4.48	-1.45	-20.13
	6.83	**6.80**	**6.83**	**6.79**	**6.81**	**6.12**	**6.56**	**6.55**	**7.68**	**NaN**	**1.50**	**7.77**	**7.21**	**5.59**	**9.19**	**6.50**	**7.12**	**5.47**
SNG	9.78	9.40	10.65	9.91	10.73	5.23	9.63	9.67	8.06	1.93	0.00	8.72	7.70	1.07	22.03	6.33	0.51	-17.81
	6.30	**6.27**	**6.30**	**6.27**	**6.29**	**5.71**	**6.08**	**6.10**	**7.06**	**3.69**	**NaN**	**7.15**	**6.63**	**5.61**	**8.35**	**5.92**	**6.74**	**5.15**
SIG	1.17	0.75	2.12	1.31	2.21	-3.82	1.00	1.05	-0.72	-7.43	-9.55	0.00	-1.11	-8.38	14.59	-2.62	-8.99	-29.05
	2.28	**2.21**	**2.56**	**2.31**	**2.58**	**1.71**	**2.25**	**1.94**	**-0.48**	**5.22**	**5.42**	**NaN**	**0.75**	**1.17**	**4.98**	**1.73**	**4.29**	**2.31**
SNIG	2.25	1.83	3.19	2.40	3.28	-2.68	2.09	2.14	0.39	-6.25	-8.35	1.10	0.00	-7.19	15.52	-1.49	-7.80	-27.64
	3.54	**3.42**	**3.89**	**3.51**	**3.88**	**2.24**	**3.12**	**2.51**	**2.13**	**6.54**	**6.46**	**2.88**	**NaN**	**2.21**	**8.65**	**2.32**	**5.46**	**3.21**
SNW	8.81	8.42	9.68	8.94	9.77	4.21	8.65	8.70	7.07	0.88	-1.08	7.73	6.71	0.00	21.19	5.32	-0.57	-19.08
	2.65	**2.61**	**2.70**	**2.64**	**2.70**	**2.42**	**2.67**	**2.65**	**2.57**	**2.34**	**2.34**	**2.59**	**2.61**	**NaN**	**3.56**	**2.57**	**2.25**	**1.57**
TLS	-15.71	-16.20	-14.60	-15.54	-14.49	-21.54	-15.90	-15.85	-17.92	-25.77	-28.25	-17.08	-18.38	-26.88	0.00	-20.14	-27.60	-51.09
	3.34	**3.39**	**3.44**	**3.38**	**3.48**	**3.47**	**3.15**	**2.71**	**2.89**	**5.56**	**5.63**	**3.00**	**2.53**	**4.54**	**NaN**	**2.48**	**6.31**	**5.91**
LOG	3.69	3.28	4.61	3.83	4.70	-1.17	3.53	3.57	1.85	-4.69	-6.76	2.55	1.47	-5.61	16.76	0.00	-6.21	-25.76
	2.52	**2.46**	**2.65**	**2.53**	**2.68**	**2.06**	**2.37**	**2.17**	**2.62**	**4.42**	**4.40**	**2.69**	**2.26**	**3.24**	**4.47**	**NaN**	**4.06**	**2.50**
IV	9.32	8.93	10.19	9.46	10.28	4.75	9.17	9.21	7.59	1.43	-0.51	8.25	7.23	0.56	21.63	5.85	0.00	-18.41
	6.64	**6.54**	**6.88**	**6.68**	**6.89**	**5.22**	**6.33**	**5.78**	**6.28**	**6.21**	**6.23**	**6.52**	**5.98**	**5.34**	**10.30**	**5.41**	**NaN**	**-10.4**
MFIV	23.42	23.09	24.15	23.53	24.22	19.56	23.29	23.33	21.96	16.76	15.11	22.51	21.65	16.02	33.82	20.49	15.55	0.00
	10.72	**10.66**	**10.88**	**10.71**	**10.89**	**9.55**	**10.28**	**9.87**	**10.77**	**9.84**	**9.61**	**10.94**	**10.54**	**8.77**	**13.29**	**9.76**	**12.86**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions, including IV and MFIV, at daily forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using a sample that ranges between 01/1996 and 12/2017. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

Table 3.19: Out-of-sample R^2 at weekly horizon - including IV

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	0.00	-0.27	0.69	-0.09	0.61	-4.69	-2.66	-1.73	-15.68	-21.15	-25.67	-14.81	-15.74	-26.15	-19.35	-18.58	-24.68	-44.47
	NaN	**1.37**	**3.70**	**1.58**	**3.36**	**1.83**	**-0.25**	**-0.42**	**0.63**	**3.59**	**3.44**	**0.75**	**0.18**	**2.00**	**0.36**	**-0.67**	**4.74**	**3.51**
SG	0.27	0.00	0.96	0.18	0.88	-4.40	-2.38	-1.46	-15.37	-20.82	-25.33	-14.50	-15.43	-25.81	-19.02	-18.26	-24.34	-44.08
	2.01	**NaN**	**3.94**	**1.55**	**3.35**	**1.44**	**-0.34**	**-0.46**	**0.42**	**3.47**	**3.31**	**0.54**	**-0.04**	**1.84**	**0.13**	**-0.91**	**4.58**	**3.27**
SNG	-0.70	-0.97	0.00	-0.79	-0.09	-5.42	-3.38	-2.44	-16.49	-22.00	-26.55	-15.62	-16.55	-27.03	-20.18	-19.41	-25.55	-45.48
	-1.09	**-2.34**	**NaN**	**-1.75**	**1.46**	**0.71**	**-1.44**	**-1.17**	**0.21**	**3.38**	**3.24**	**0.33**	**-0.32**	**1.74**	**-0.39**	**-1.41**	**4.43**	**3.14**
SIG	0.09	-0.18	0.79	0.00	0.70	-4.59	-2.56	-1.64	-15.57	-21.03	-25.55	-14.71	-15.63	-26.03	-19.23	-18.46	-24.56	-44.34
	2.22	**1.44**	**3.73**	**NaN**	**3.21**	**1.55**	**-0.26**	**-0.45**	**0.64**	**3.44**	**3.31**	**0.75**	**0.21**	**1.91**	**0.40**	**-0.75**	**4.61**	**3.45**
SNIG	-0.61	-0.89	0.09	-0.71	0.00	-5.33	-3.29	-2.36	-16.39	-21.89	-26.44	-15.52	-16.45	-26.92	-20.08	-19.30	-25.44	-45.36
	0.04	**-0.81**	**1.47**	**-0.49**	**NaN**	**0.44**	**-1.08**	**-0.96**	**0.11**	**3.30**	**3.15**	**0.25**	**-0.44**	**1.61**	**-0.47**	**-1.54**	**4.42**	**3.02**
SNW	4.48	4.22	5.14	4.39	5.06	0.00	1.94	2.82	-10.50	-15.72	-20.05	-9.67	-10.56	-20.50	-14.00	-13.27	-19.10	-38.00
	3.58	**3.45**	**4.09**	**3.46**	**4.05**	**NaN**	**2.11**	**1.75**	**0.78**	**3.74**	**3.40**	**0.93**	**0.09**	**1.29**	**-0.44**	**-1.23**	**3.68**	**1.05**
TLS	2.59	2.33	3.27	2.50	3.18	-1.98	0.00	0.90	-12.69	-18.01	-22.42	-11.84	-12.74	-22.88	-16.25	-15.51	-21.45	-40.73
	2.84	**2.69**	**3.39**	**2.87**	**3.60**	**2.47**	**NaN**	**0.30**	**1.13**	**3.30**	**3.15**	**1.19**	**0.73**	**1.88**	**0.22**	**-1.13**	**4.24**	**2.88**
LOG	1.70	1.44	2.39	1.61	2.30	-2.90	-0.91	0.00	-13.71	-19.09	-23.53	-12.86	-13.77	-24.00	-17.31	-16.56	-22.56	-42.01
	1.72	**1.64**	**1.88**	**1.68**	**2.06**	**1.84**	**0.89**	**NaN**	**1.24**	**3.16**	**3.05**	**1.27**	**0.95**	**1.93**	**0.87**	**-0.15**	**3.99**	**3.07**
Normal	13.56	13.32	14.16	13.47	14.08	9.50	11.26	12.06	0.00	-4.73	-8.64	0.75	-0.05	-9.05	-3.17	-2.50	-7.78	-24.89
	6.86	**6.78**	**6.99**	**6.81**	**6.84**	**5.88**	**5.11**	**4.84**	**NaN**	**4.08**	**3.56**	**3.05**	**2.31**	**0.38**	**2.70**	**2.11**	**4.23**	**1.77**
SG	17.46	17.23	18.03	17.38	17.96	13.59	15.26	16.03	4.51	0.00	-3.73	5.23	4.46	-4.13	1.49	2.12	-2.91	-19.25
	6.91	**6.82**	**6.87**	**6.82**	**6.80**	**6.35**	**6.32**	**6.48**	**6.21**	**NaN**	**-0.78**	**6.33**	**6.11**	**4.26**	**5.52**	**5.37**	**6.01**	**4.76**
SNG	20.43	20.21	20.98	20.35	20.91	16.70	18.31	19.05	7.95	3.60	0.00	8.64	7.90	-0.38	5.03	5.65	0.79	-14.96
	6.57	**6.49**	**6.53**	**6.51**	**6.48**	**5.97**	**6.07**	**6.25**	**6.03**	**4.93**	**NaN**	**6.13**	**5.84**	**4.69**	**5.26**	**5.14**	**5.85**	**4.59**
SIG	12.90	12.67	13.51	12.82	13.43	8.82	10.59	11.39	-0.75	-5.52	-9.46	0.00	-0.81	-9.87	-3.95	-3.28	-8.59	-25.83
	6.28	**6.18**	**6.42**	**6.23**	**6.25**	**5.27**	**4.58**	**4.34**	**-0.91**	**3.75**	**3.24**	**NaN**	**1.01**	**-0.16**	**2.18**	**1.59**	**3.80**	**1.32**
SNIG	13.60	13.37	14.20	13.52	14.13	9.55	11.30	12.10	0.05	-4.67	-8.58	0.80	0.00	-8.99	-3.11	-2.45	-7.72	-24.82
	6.83	**6.70**	**6.93**	**6.79**	**6.73**	**5.53**	**4.94**	**4.65**	**2.12**	**4.68**	**4.19**	**2.54**	**NaN**	**1.14**	**2.36**	**1.61**	**4.14**	**1.50**
SNW	20.73	20.51	21.28	20.65	21.21	17.01	18.62	19.35	8.30	3.96	0.38	8.99	8.25	0.00	5.39	6.00	1.17	-14.52
	5.17	**5.12**	**5.21**	**5.13**	**5.16**	**4.73**	**4.96**	**5.04**	**3.85**	**2.97**	**2.57**	**3.85**	**3.88**	**NaN**	**3.71**	**3.76**	**3.25**	**2.20**
TLS	16.21	15.98	16.79	16.13	16.72	12.28	13.98	14.76	3.07	-1.51	-5.30	3.80	3.02	-5.70	0.00	0.64	-4.47	-21.05
	6.17	**6.15**	**6.39**	**6.23**	**6.42**	**5.41**	**5.98**	**5.79**	**3.62**	**4.43**	**4.08**	**3.63**	**3.47**	**3.17**	**NaN**	**3.50**	**4.74**	**2.77**
LOG	15.67	15.44	16.25	15.59	16.18	11.71	13.42	14.20	2.44	-2.17	-5.98	3.17	2.39	-6.39	-0.65	0.00	-5.15	-21.84
	6.83	**6.84**	**7.14**	**6.98**	**7.19**	**5.72**	**6.93**	**6.62**	**3.40**	**4.02**	**3.72**	**3.39**	**3.25**	**2.84**	**3.30**	**NaN**	**4.63**	**2.57**
IV	19.79	19.58	20.35	19.72	20.28	16.04	17.66	18.40	7.22	2.83	-0.80	7.91	7.17	-1.18	4.28	4.89	0.00	-15.87
	9.30	**9.14**	**9.40**	**9.17**	**9.38**	**8.27**	**8.37**	**8.06**	**6.01**	**5.08**	**4.40**	**6.08**	**5.98**	**3.59**	**5.74**	**5.04**	**NaN**	**-11.6**
MFIV	30.78	30.59	31.26	30.72	31.20	27.54	28.94	29.58	19.93	16.14	13.01	20.53	19.89	12.68	17.39	17.92	13.70	0.00
	12.45	**12.34**	**12.57**	**12.36**	**12.63**	**12.22**	**11.73**	**11.48**	**9.88**	**8.06**	**7.24**	**9.93**	**10.00**	**7.34**	**10.34**	**9.64**	**13.95**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions, including IV and MFIV, at the weekly forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using a sample that ranges between 01/1996 and 12/2017. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

Table 3.20: Out-of-sample R^2 at monthly horizon - including IV

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	0.00	0.67	2.23	0.47	3.57	11.75	3.35	2.28	-1.51	-3.69	-10.19	-1.01	-2.63	-10.02	-7.63	-6.11	-7.23	-20.82
	NaN	**2.64**	**5.80**	**3.76**	**6.19**	**7.22**	**4.43**	**4.22**	**10.14**	**8.79**	**8.80**	**10.10**	**10.03**	**8.98**	**9.30**	**8.62**	**9.26**	**9.39**
SG	-0.68	0.00	1.57	-0.21	2.92	11.15	2.69	1.61	-2.20	-4.40	-10.93	-1.69	-3.33	-10.76	-8.36	-6.83	-7.96	-21.64
	1.75	**NaN**	**4.21**	**2.22**	**5.06**	**6.85**	**3.74**	**3.41**	**9.77**	**8.47**	**8.48**	**9.73**	**9.66**	**8.68**	**8.98**	**8.31**	**9.05**	**9.19**
SNG	-2.28	-1.59	0.00	-1.81	1.37	9.73	1.14	0.05	-3.83	-6.06	-12.70	-3.32	-4.98	-12.53	-10.09	-8.53	-9.68	-23.58
	-2.13	**-2.14**	**NaN**	**-1.58**	**4.41**	**6.76**	**3.04**	**2.46**	**9.88**	**8.37**	**8.37**	**9.82**	**9.78**	**8.54**	**9.03**	**8.26**	**9.10**	**9.23**
SIG	-0.47	0.21	1.77	0.00	3.12	11.33	2.90	1.82	-1.99	-4.18	-10.70	-1.48	-3.11	-10.53	-8.13	-6.61	-7.73	-21.39
	2.74	**1.90**	**6.18**	**NaN**	**6.11**	**7.05**	**4.46**	**4.06**	**9.97**	**8.48**	**8.51**	**9.91**	**9.86**	**8.75**	**9.20**	**8.47**	**9.14**	**9.30**
SNIG	-3.70	-3.00	-1.39	-3.22	0.00	8.48	-0.23	-1.34	-5.28	-7.53	-14.27	-4.75	-6.44	-14.09	-11.62	-10.04	-11.20	-25.30
	-2.30	**-2.42**	**-0.18**	**-2.02**	**NaN**	**7.07**	**2.37**	**1.57**	**10.28**	**8.42**	**8.35**	**10.21**	**10.20**	**8.62**	**9.30**	**8.43**	**9.26**	**9.36**
SNW	-13.31	-12.55	-10.78	-12.78	-9.26	0.00	-9.51	-10.73	-15.03	-17.50	-24.85	-14.45	-16.30	-24.66	-21.96	-20.24	-21.50	-36.90
	-1.93	**-2.30**	**-1.73**	**-2.02**	**-1.80**	**NaN**	**0.02**	**-0.32**	**8.51**	**6.38**	**6.05**	**8.42**	**8.31**	**5.97**	**6.24**	**5.16**	**7.28**	**7.13**
TLS	-3.47	-2.77	-1.16	-2.99	0.23	8.69	0.00	-1.11	-5.03	-7.29	-14.01	-4.51	-6.19	-13.83	-11.36	-9.79	-10.95	-25.01
	0.70	**0.62**	**1.75**	**0.99**	**2.91**	**5.94**	**NaN**	**1.45**	**8.38**	**6.92**	**6.82**	**8.27**	**8.35**	**6.85**	**8.02**	**7.27**	**8.17**	**8.13**
LOG	-2.33	-1.64	-0.05	-1.85	1.32	9.69	1.10	0.00	-3.88	-6.11	-12.75	-3.36	-5.03	-12.58	-10.14	-8.59	-9.73	-23.64
	1.18	**1.14**	**2.13**	**1.36**	**3.21**	**6.13**	**3.48**	**NaN**	**8.96**	**7.50**	**7.43**	**8.85**	**8.93**	**7.50**	**8.70**	**7.92**	**8.67**	**8.74**
Normal	1.49	2.16	3.69	1.95	5.01	13.06	4.79	3.74	0.00	-2.14	-8.54	0.50	-1.10	-8.38	-6.02	-4.53	-5.63	-19.02
	6.92	**6.62**	**7.17**	**6.90**	**7.26**	**8.18**	**6.32**	**6.83**	**NaN**	**1.62**	**0.71**	**1.57**	**1.60**	**-0.56**	**1.46**	**1.01**	**3.67**	**1.89**
SG	3.56	4.21	5.71	4.01	7.01	14.89	6.79	5.76	2.10	0.00	-6.26	2.59	1.02	-6.10	-3.80	-2.33	-3.41	-16.52
	7.88	**7.66**	**7.92**	**7.75**	**7.85**	**8.04**	**7.05**	**7.49**	**5.41**	**NaN**	**-1.30**	**5.66**	**5.02**	**4.48**	**4.07**	**4.15**	**4.90**	**4.06**
SNG	9.24	9.86	11.27	9.67	12.49	19.91	12.29	11.31	7.87	5.89	0.00	8.33	6.85	0.15	2.32	3.70	2.68	-9.65
	8.25	**8.06**	**8.22**	**8.15**	**8.15**	**7.79**	**7.26**	**7.78**	**5.44**	**5.09**	**NaN**	**5.62**	**5.06**	**6.07**	**4.08**	**4.26**	**4.66**	**3.76**
SIG	1.00	1.67	3.21	1.46	4.53	12.63	4.32	3.25	-0.50	-2.66	-9.09	0.00	-1.61	-8.92	-6.55	-5.05	-6.16	-19.61
	6.74	**6.42**	**6.99**	**6.71**	**7.10**	**8.18**	**6.07**	**6.60**	**0.71**	**1.11**	**0.15**	**NaN**	**0.12**	**-1.50**	**1.03**	**0.64**	**3.31**	**1.54**
SNIG	2.57	3.22	4.74	3.02	6.05	14.01	5.83	4.79	1.09	-1.03	-7.36	1.58	0.00	-7.19	-4.87	-3.39	-4.48	-17.72
	3.09	**3.01**	**3.12**	**3.09**	**3.10**	**2.86**	**2.92**	**3.09**	**1.30**	**1.58**	**1.41**	**1.31**	**NaN**	**1.08**	**1.35**	**1.25**	**1.92**	**1.49**
SNW	9.10	9.72	11.13	9.53	12.35	19.78	12.15	11.18	7.73	5.75	-0.15	8.19	6.71	0.00	2.17	3.55	2.53	-9.82
	8.40	**8.19**	**8.43**	**8.33**	**8.44**	**8.12**	**7.49**	**8.08**	**5.36**	**4.97**	**2.68**	**5.56**	**4.77**	**NaN**	**3.47**	**3.72**	**4.21**	**2.94**
TLS	7.09	7.72	9.16	7.52	10.41	18.00	10.20	9.20	5.68	3.66	-2.38	6.15	4.64	-2.22	0.00	1.41	0.37	-12.26
	6.95	**6.98**	**7.28**	**7.13**	**7.35**	**6.61**	**7.63**	**7.66**	**3.47**	**3.40**	**2.96**	**3.41**	**3.33**	**2.43**	**NaN**	**3.90**	**4.69**	**3.08**
LOG	5.76	6.40	7.86	6.20	9.13	16.83	8.92	7.91	4.33	2.28	-3.84	4.81	3.28	-3.68	-1.43	0.00	-1.05	-13.86
	6.47	**6.50**	**6.82**	**6.65**	**6.90**	**6.02**	**7.42**	**7.37**	**2.86**	**2.86**	**2.50**	**2.80**	**2.76**	**1.99**	**2.62**	**NaN**	**4.10**	**2.49**
IV	6.74	7.37	8.82	7.18	10.07	17.70	9.87	8.87	5.33	3.30	-2.76	5.80	4.29	-2.60	-0.37	1.04	0.00	-12.67
	9.06	**8.95**	**9.23**	**9.07**	**9.35**	**9.19**	**8.70**	**9.03**	**5.51**	**4.39**	**3.42**	**5.39**	**5.41**	**2.77**	**4.65**	**3.60**	**NaN**	**-9.24**
MFIV	17.23	17.79	19.08	17.62	20.19	26.96	20.01	19.12	15.98	14.18	8.80	16.40	15.05	8.94	10.92	12.17	11.25	0.00
	11.12	**11.04**	**11.30**	**11.14**	**11.42**	**11.25**	**10.91**	**11.18**	**7.52**	**6.05**	**4.87**	**7.40**	**7.58**	**4.45**	**8.05**	**6.80**	**11.65**	**NaN**

Notes: This table reports out-of-sample R-squared R^2_{os} from pairwise comparisons between any two model-based predictions, including IV and MFIV, at the monthly forecasting horizon. Associated statistics from Clark-West tests are reported in bold below each entry of R^2_{os}. The results are obtained using a sample that ranges between 01/1996 and 12/2017. Vertically, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows. Positive statistics indicate the relative outperformance of models against the benchmarks.

3.6 Conclusion

In this study, I explore a variety of assumptions on the distribution of innovation term ξ_t that brings randomness into the index returns process. Driven by the variance of the innovation term, at each point in time, GARCH models are constructed to make the returns have different types of distributions so that the returns are characterised by corresponding GARCH processes. The GAS model has been receiving increasing attention in recent years. Also known as conditional score models, they are observation based and can be converted into GARCH models under specific conditions. Following identical assumptions on the distribution of innovation terms, a series of GAS models are constructed as comparisons to GARCH models.

The emphasis in this study is placed on evaluating the predictive power of both types of parametric models in forecasting realised volatility. Additionally, the performance of implied volatility is compared to that of parametric models. The most prominent contribution of this study is that it is proposed in this chapter to model and predict volatility within the framework of GAS models. With a variety of assumptions on innovation terms, a series of observation-driven GARCH and GAS models are fitted, and extensive comparisons between the performance of GARCH and GAS are conducted so that the first systematic evaluation of these two representative observation-driven models is implemented in this study, from which several new findings are documented.

First, I find that GAS models tend to have lower predictions of realised volatility compared to those produced by GARCH models. More importantly, TLS-GAS predictions on average are closer to RVs compared to other model-based predictions. Second, while the SNW-GARCH model is found to be a better candidate for describing the movement of market returns, the TLS-GAS model seems to be the overall best performer in predicting RV at the daily horizon. Third, it can be observed that while descriptive R^2 of GAS models are generally lower than GARCH

models at daily and weekly horizons, GARCH models have lower R^2 than GAS models at the monthly horizon. Furthermore, GARCH models are generally found to have significantly lower forecasting errors measured by MAE compared to GAS, indicating the superiority of GARCH over GAS in volatility modelling. However, it is noteworthy that TLS-GAS has the lowest prediction error of 3.86 at the daily horizon. Such outperformance is found to be statistically different and it is confirmed by out-of-sample R^2 and CW test statistics. Fourth, pairwise out-of-sample R-squared comparisons between models confirm the dominance of TLS-GAS at the daily horizon.

Adopting a sub-sample during the 2007 - 2008 financial crisis, it is documented that the SNW-GAS overwhelmingly dominates other models at the daily horizon, indicating that the Weibull distribution is a more suitable candidate for financial returns modelling during high market uncertainty. The dominance of SNW-GAS lasts at the weekly horizon and disappears at the monthly horizon.

The predictive power of model-based forecasts is also compared to that of implied volatility represented by IV and MFIV. Consistent with existing literature, the dominance of implied volatility in describing the changes in realised volatility is observed. Such superior predictive power is robust across horizons from 1 day up to 1 month and it is found to be more robust as the forecasting horizon extends. Concerning the prediction accuracy measured by MAE, IV and MFIV fail to outperform parametric models, and such inferior performance becomes weaker as the forecasting horizon stretches. The DM test results indicate that, at the daily horizon, IV is not statistically different from other model-based predictions in most cases (10 out of 16), while MFIV is significantly different from almost all other predictions (15 out of 16). This pattern continues at weekly horizons but becomes the opposite at monthly horizons. Finally, the pairwise CW test statistics indicate that model-based forecasts consistently have lower prediction errors than IV and MFIV. Such superiority of parametric models in prediction accuracy lasts for forecasting horizons from one

day up to one month.

Chapter 4

Option Pricing with GARCH/GAS Models

4.1 Introduction

4.1.1 Option pricing models

The seminal work of Black and Scholes (1973) and Merton (1973) was followed by a significant increase in pricing studies that attempted to determine the fair value of contingent claims. It is also the answer to how to assign these contingent claims a fair value. As one of the most common derivatives, options play an essential role in financial markets. The value of an option presently is determined by its possible payoffs at the maturity date. Following the non-arbitrage principle, Black and Scholes (1973) proposed the closed-form BS option pricing model. It is worth mentioning that the BS model is consistent with the Efficient Market Hypothesis (ETH), where there exists a unique risk-neutral probability measure, such that the discounted evolution of the underlying price is a martingale. In that case, the value of an option (or other contingent claims) is simply the expectation of future payoffs

discounted using risk-free rates. More discussions on the non-arbitrage principle can be found in Harrison and Pliska (1981).

Volatility is proportional to the level of option prices. Hence it plays a vital role in option pricing. It is natural to think that high-quality volatility forecasting improves option pricing accuracy. In the practice of volatility forecasting, it is more convenient to consider discrete-time models. In this field, a dominating group is the Autoregressive Conditional Heteroskedasticity (ARCH) models developed by Engle (1982) and its generalised version Generalised Autoregressive Conditional Heteroskedasticity (GARCH) proposed by Bollerslev (1986). The GARCH models are advantageous partly in their ability to capture the time-varying property of volatility. However, standard GARCH models have limited capability in characterising the negative skewness and excess kurtosis that are commonly observed in financial returns time series. This motivates the explosive growth of studies in extensions of GARCH models. One important extension lies in the distribution of asset returns. According to Bollerslev (1987), distributions with heavier tails, such as the Student's t distribution, can accommodate the excess kurtosis. Another candidate with a similar feature is the Generalised Error Distribution (GED) proposed by Nelson (1991). Another well-documented characteristic of volatility is its different responses to positive and negative underlying return shocks. GARCH models specified with these distributions are still insufficient to characterise the negative skewness at the same time. A solution to this is to introduce asymmetry. This can be realised by incorporating the "leverage effects" using GARCH variations, such as threshold GARCH, exponential GARCH and the GJR-GARCH (Glosten et al., 1993).

With reasonably estimated volatility, a variety of options pricing approaches exist for practitioners to choose from. Among these, the Black-Scholes (BS) model is the most popular one in practice. As it is analytical, the pricing can be accomplished in milliseconds using modern computers. However, the problematic part of the BS model is that it is developed based on a series of unrealistic assumptions. Among

these, one of the most problematic ones is probably that the volatility, as one of the inputs to the model, is assumed to be constant over time to maturity and across the spectrum of strike prices. The concept of the "volatility smile" is proposed in Rubinstein (1985) to describe such volatility behaviour. Empirically, it can be observed that volatility forms a surface against time to maturity and strikes. Apart from this, the distribution of underlying asset returns is not necessarily normal. It is reasonable that models that take into consideration the nature of volatility might have improved performance.

In the derivation of the BS model, stock prices are assumed to follow a geometric Brownian motion such that

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{4.1}$$

where S_t denotes the stock price at time t, μ is called the drift term that assigns S_t some directional trend over time, σ is constant volatility and dW_t is some Wiener process.

Quite a few studies propose an improved option pricing model by relaxing the constant volatility assumption. The critical logic lies in this strand of research is to let volatility σ be stochastic. To the best of my knowledge, the first attempt that takes into consideration stochastic volatility in the option pricing model was made by Heston (1993).

$$dS_t = \mu S_t dt + \sqrt{\nu_t} S_t dW_t^S \tag{4.2}$$

$$d\nu_t = \kappa\left(\theta - \nu_t\right) dt + \xi\sqrt{\nu_t} dW_t^\nu \tag{4.3}$$

where ν_t is the instantaneous variance at time t, θ is the average variance in the long run and ξ is usually referred to as the volatility of volatility. It determines the variance of ν_t. κ is the reversion rate of ν_t towards θ. It can be shown that as $t \sim \infty$, $\nu_t \sim \theta$. It is worth mentioning that if the condition $2\kappa\theta > \xi^2$ is satisfied, the

process of ν_t will be strictly positive. This condition is called the Feller condition.

Following the above assumption on the evolution of asset prices, Heston (1993) shows that the price of a European call option can be estimated using a formula analogous to the BS model in terms of model structure.

$$C = S_0 P_1 - e^{-rT} K P_2 \tag{4.4}$$

where S_0 stands for the spot price, K is the strike price, P_1 is the option delta and P_2 is the risk-neutral probability that the call option expires in-the-money, that is $\mathrm{P}\left(S_T > K\right)$. One problem is that such density is not instantly available in closed form. However, it can be calculated via the one-to-one mapping relationship between the probability density and characteristic functions. By adopting inverse Fourier transformation, Heston (1993) shows that

$$P_1 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi \ln(K)} \psi(\phi - i)}{i\phi \psi(-i)}\right] d\phi \tag{4.5}$$

and

$$P_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left[\frac{e^{-i\phi \ln(K)} \psi(\phi)}{i\phi}\right] d\phi \tag{4.6}$$

where $\psi(\phi)$ is the characteristic function of the log price of the underlying asset. Equations (4.4), (4.5) and (4.6) constitute the Heston (1993) option pricing model. This approach, despite having a quasi-closed form, is still semi-analytical. This is because the parameters that define the Heston stochastic process still need to be estimated before using them to compute the characteristic function $\psi(\phi)$. The relative advantage of this model is that once the parameters are obtained from historical data, the option pricing can be done in a fraction of a second. Note that some option approaches adopt Monte Carlo simulation to compute the expected payoff of underlying, which can be very time-consuming.

In the framework of the GARCH model, Heston and Nandi (2000) assume that

the log price of the underlying asset follows the GARCH process and proposes the closed-form GARCH option valuation model

$$\begin{aligned} C =& e^{-r(T-t)} E_t[\text{Max}(S(T) - K, 0)] \\ =& \frac{1}{2} S(t) + \frac{e^{-r(T-t)}}{\pi} \int_0^\infty \text{Re}\left[\frac{K^{-i\phi}\psi(i\phi+1)}{i\phi}\right] d\phi \\ & - \text{K}e^{-r(T-t)} \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re}\left[\frac{K^{-i\phi}\psi(i\phi)}{i\phi}\right] d\phi\right) \end{aligned} \tag{4.7}$$

This discrete time model resembles the Heston (1993) continuous time model, and it can be shown that the latter is the former's limit when time is continuous (Christoffersen et al., 2006).

Instead of having closed-form solutions, some other methods emphasise finding out the process that describes the evolution of the underlying under the risk-neutral measure $\mathbb{Q}$. To the best of my knowledge, Duan (1995) was the first to systematically build a theoretical framework for option pricing under the context of the GARCH model. The most important work in Duan (1995) is generalising the concept of risk-neural pricing by introducing the Locally Risk-Neutral Valuation Relationship (LRNVR). This enables us to select an equivalent martingale under the GARCH model with normal innovation. However, this approach does not apply to GARCH models with non-normal innovations. Adopting a different risk-neutralisation approach, Duan (1999) proposes a model with innovation terms following generalised error distribution (GED). Stentoft (2008) uses the same method to find the risk-neutral representation for the GARCH model with normal inverse Gaussian innovation. The normalised density function of GED is given by

$$f_\varphi(\xi) = \frac{\varphi}{2^{1+\frac{1}{\varphi}}\theta(\varphi)\Gamma\left(\frac{1}{\varphi}\right)} \exp\left(-\frac{1}{2}\left|\frac{\xi}{\theta(\varphi)}\right|^\varphi\right) \tag{4.8}$$

where

$$\theta(\varphi) = \left(\frac{2^{-\frac{2}{\varphi}} \Gamma\left(\frac{1}{\varphi}\right)}{\Gamma\left(\frac{3}{\varphi}\right)} \right)^{\frac{1}{2}} \tag{4.9}$$

and where $\Gamma(\cdot)$ is the Gamma function. ξ has a zero skewness and kurtosis of

$$\kappa(\varphi) = \frac{\Gamma\left(\frac{5}{\varphi}\right) \Gamma\left(\frac{1}{\varphi}\right)}{\Gamma\left(\frac{3}{\varphi}\right)^2} \tag{4.10}$$

It can be shown that when $\varphi = 2$ $\kappa = 0$, and $f(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right)$. This means that the standard normal distribution is nested in GED, as a special case when $\varphi = 2$.

Neither Duan (1995) nor Duan (1999) provide closed form solution as in Heston (1993) and Heston and Nandi (2000). Instead, in these approaches, option pricing is done via simulations after obtaining the risk-neutralised representation of the underlying price dynamics. This is less cost-effective and less accurate occasionally. Another study that follows a similar fashion is Christoffersen et al. (2006). They propose a GARCH model with inverse Gaussian innovation. The difference is that, in their study, the framework of the Esscher transform is adopted for risk-neutralisation, after which option pricing is implemented again via Monte Carlo simulations. Other beneficial attempts on making different assumptions on the innovation term within the GARCH framework include Shifted Gamma distribution in Tong et al. (2004) under GARCH framework and Badescu et al. (2008) in which a mixture normal innovation is adopted.

4.1.2 Risk-neutralisation

Harrison and Pliska (1981) documented a valid mathematical illustration of the relation between the no-arbitrage principle and asset pricing under risk-neutral measures. According to their study, if no arbitrage opportunities can be found, it

means that there exists an Equivalent Martingale Measure (EMM). Consequently, the unique prices for any contingent claims can be estimated by discounting expected payoffs at maturity under such martingale measure. However, the complete market assumption is problematic, and there is normally more than one EMM and hence a range of contingent claim prices. In this case, it is important to select one EMM under which the contingent claim price is economically justifiable. To this point, Gerber and Shiu (1994) propose an elegant approach for EMM selection via the Esscher transform. The latter is a widely applied tool in actuarial science, developed by Escher (1932), in an incomplete market setting. The most appealing part of their approach is that it provides more flexibility for practitioners to choose from a range of parametric models, such as the observation-driven GARCH and GAS models. These models can be used for option pricing in a convenient way. It is worth mentioning that the option prices given by the Esscher transform are consistent with those from the maximisation of the expected power utility of an economy. On the basis of Gerber and Shiu (1994), Bühlmann et al. (1996) propose a generalisation of the Esscher transform in the context of a stochastic process. The concept of conditional Esscher transforms was introduced in order to incorporate semi-martingale under the no-arbitrage condition.

The Esscher transform is not the only way to identify EMM. To obtain these risk-neutral measures, one can also resort to the previously introduced LRNVR proposed by Duan (1995). However, one pre-requisite of applying this approach is that the distribution of innovation terms must be normal. Since the emphasis of our study is to explore and scrutinise alternative assumptions on error terms other than normality and to make comparisons between GARCH models and corresponding GAS counterparts, a more flexible method for computing risk-neutral measures is appealing. The Extended Girsanov Principle (EGP), proposed by Elliott and Madan (1998) fixes the gap. Consistent with Tong et al. (2004), in this study, the Esscher transform is applied to the parameters calibrated under the $\mathbb{P}$ measure to obtain

their counterparts under the $\mathbb{Q}$ measure, under which the option valuations are implemented.

4.1.3 Contributions

My study contributes to the strand of literature represented by Duan (1995) and Duan (1999), where, in order to value options, simulations are applied after risk-neutralising GARCH models. The first innovation of my study is that, instead of GARCH models, I construct GAS models based of different assumptions on the distribution of market returns. The second innovation of my study is that the GAS models' representations under risk-neutral measure are found out by adopting the Esscher transform, as in (Tong et al., 2004). This is because the Esscher transform has higher flexibility in dealing with non-normally distributed innovations. As the proposed GAS models in this study have innovation terms that follow shifted Gamma process, the Esscher transform is a good candidate for risk-neutralisation. The Gamma innovation term is chosen because of its capability in accounting for the skewness exhibited in financial returns (Tong et al., 2004).

The primary finding of this study is that, in terms of pricing accuracy measured by Mean Absolute Percentage Error (MAPE), the GAS models with shifted Gamma innovation (SG-GAS) and shifted negative Gamma innovation (SNG-GAS) significantly outperform other competing models including the BS model, GARCH model with normal innovation (GARCH) and GAS model with normal innovation (GAS).

It is also reported in this study that while all the models seem to suffer from drops in pricing accuracy from the pre-crisis period to the post-crisis period, the superior performance of SG-GAS and SNG-GAS models stays unchanged. In fact, it is even more significant in a normal market atmosphere. During the 2007-2008 financial crisis, the GAS models with Gamma innovations are no longer the best performers.

Instead, the observation-based models with normal innovation (GARCH and GAS) tend to have relatively better performance during this massive market uncertainty.

To investigate the source of the outperformance of GAS models with shifted Gamma innovations, the option valuation performance of all models is tested across a range of moneyness levels, and across a variety of times to maturity. I find that the superior pricing performance of SG-GAS and SNG-GAS is mainly due to their ability to more accurately price Deep Out-of-the-Money (dOTM) options. Given that the number of near-the-money options is 10924 and the number of OTM options is 31773, it is reasonable to think that it is the superiority of SG-GAS and SNG-GAS models in pricing OTM options that drives their overall outperformance in option pricing. It can be observed that SG-GAS and SNG-GAS models perform better in pricing options with longer maturities as well.

Designed Monte Carlo simulation studies, instantly show that under the $\mathbb{Q}$ measure, the density implied by SNG-GAS model tends to have fatter tails compared to that implied by the GARCH model with normal innovations. Such a pattern can be consistently observed for density simulated 30 days to 360 days from now. Such behaviour of the SNG-GAS model indicates that it tends to overestimate tail risk, which is related to the price of OTM options. Note that it is reported that the BS model is flawed due to its inferior accuracy in pricing OTM options because of its tendency to underestimate tail risk. The GARCH model is found to have a similar performance as the BS model.

The rest of this chapter is organised as follows: Section 4.2 introduces the framework of the GARCH model. Section 4.3 illustrates the mechanism of the GAS model, including how the re-parameterisation is implemented via the link function and how the parameters are estimated via Maximum Likelihood Estimation (MLE). Section 4.4 briefly introduces the Esserch transform and how it can be applied to option pricing. In Section 4.5 I construct a GAS model with normal innovation

(GAS) and adopt the Esscher transform to obtain the GAS model for option pricing under $\mathbb{Q}$ measure. Following a similar procedure, the GAS option pricing model with shifted Gamma (SG-GAS) and shifted negative Gamma (SNG-GAS) innovations are constructed. Furthermore, GARCH models with normal and shifted Gamma innovations are also constructed as comparisons to GAS models. In Section 4.6, I design and conduct a simulation study using all 4 parametric models. It can be shown in this section that, from $\mathbb{P}$ measure to $\mathbb{Q}$ measure, the distribution of the terminal index level simulated by the GAS model with normal innovation is shifted leftwards by the Esscher transform, while the distribution of the terminal index level simulated by SNG-GAS is skewed by the transform. I also show that the SNG-GAS model, due to its ability in simulating more skewed index levels, could be a more reasonable choice for option pricing, especially for those deep out-of-the-money ones. I also present and compare in this section the volatility smile characterised by all observation-driven models over different horizons. In Section 4.7, I use DJIA index historical data option data to train GARCH and GAS models and evaluate model performance by computing the MAPE of estimated option prices against corresponding option quotes in the real world. Section 4.8 concludes.

4.2 GARCH Models

The ARCH model was first developed by Engle (1982) to forecast conditional variances. In an ARCH(p) model, the conditional variance is dependent on p-order lagged squared errors. The model can be written down as

$$\begin{aligned} y_t &= \mu_t + \xi_t \\ h_t &= \omega + \sum_{i=1}^{p} \alpha_i \xi_{t-i}^2. \end{aligned} \tag{4.11}$$

where y_t in our study is log-returns of an asset, μ_t is the mean value of y_t that changes over time, and ξ is called the innovation term or error term, h_t is the conditional variance of the error term, ω is some constant value, α_i is defined as the coefficients of lagged terms of ξ.

The ARCH models are able to capture some features of volatility, such as volatility clustering (Bera and Higgins, 1993). However, it still has an obvious disadvantage. As it is widely known, volatility exhibits strong dependency even after long lags. As documented by Abdalla (2012), it can be difficult to estimate a large number of ARCH parameters due to a large number of lagged terms. The GARCH model, as a generalisation of the ARCH model, avoids the long-lag structure of ARCH models. The GARCH model can be described by the equation below:

$$y_t = \mu_t + \xi_t \tag{4.12}$$

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \xi_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}. \tag{4.13}$$

In the application of option pricing, one important property of the GARCH model is that it is non-Markovian (Duan, 1995). More specifically, the GARCH volatility (hence the GARCH option prices) is dependent on the information set at present and in the past. This forms the foundation of dynamically making one-step-ahead volatility forecasts. The GARCH model, when first proposed, was still inadequate to capture some commonly observed features in financial returns time series, such as the asymmetric responses to positive and negative returns shocks. Plenty of GARCH extensions have been developed to account for such an effect. Another problematic part of the GARCH model is that the innovation term ξ is not necessarily normally distributed. In fact, returns series tend to be leptokurtic. According to Bollerslev (1987) and Baillie and Bollerslev (1989), the normality assumption fails to accommodate such behaviour of asset returns. Tong et al. (2004), within the GARCH framework, propose to assume that the innovation terms ξ follow a shifted Gamma

distribution, as it has the ability to account for the skewness of stock returns.

4.3 GAS Models

Time-varying parameters are appealing in describing stochastic processes. The GARCH model is one of the examples of time-varying parameter models. Apart from GARCH, the stochastic volatility model is another candidate; see Shephard (2005) for a review. According to Creal et al. (2014) and Harvey (2013), two challenges are confronted by these models. First, it can be difficult to estimate the parameters. Second, the shape parameter that also characterises a distribution is not appropriately take into consideration by these models. Creal et al. (2014) and Harvey (2013) propose the Generalised Autoregressive Score (GAS) model, where the score of the conditional density function is set as the driver of changes in parameters over time. It is worth mentioning that another name for the GAS model is the Dynamic Conditional Score (DCS) model.

4.3.1 The Framework of GAS

Denote the random variable we attempt to model by y_t, and let f_t denote the time-varying parameter that is relevant to the distribution of y_t. Further denote the information set until time t as I_t, the GAS model supposes that the data generating process y_t is dependent on past information I_{t-1} and f_t plus a set of static hyperparameters θ. The variable y_t is predicted to have a distribution as below:

$$y_t \sim p\left(I_{t-1}, f_t; \theta\right) \tag{4.14}$$

For a GAS(1,1) model, the updating process for time-varying parameter f_t can be represented as

$$f_{t+1} = \omega + \alpha S_t \nabla_t + \beta f_t \tag{4.15}$$

where ω is a vector of constants that usually take values of unconditional expectations of f_t. S_t is a scaling matrix of the log density of observations. It is computed as the score of $\ln p\left(y_t \mid I_t, f_t; \theta\right)$

$$\nabla_t = \left[\frac{\partial \ln p\left(y_t \mid I_t, f_t; \theta\right)}{\partial f_t}\right] \tag{4.16}$$

The coefficient matrices α, β, together with ω are functions of static parameter θ. Clearly, the innovative part of the representation is that the scaled score function is adopted as the driver in the updating process of f_t. It bridges all the parameters that define the conditional distributions of observations to changes in f_t. The scaling matrix S_t can be written as

$$S_t = \mathcal{I}_t^{-d} \tag{4.17}$$

where $\mathcal{I}_t$ is the information matrix of the observation density. d is called the scaling coefficient, and its value is usually chosen between 0, 0.5 or 1. In this chapter, d is set to 1. Under this condition, following Creal et al. (2012), the scaling matrix S_t can be calculated as:

$$S_t = \mathcal{I}_t^{-1} = -E_t\left[\nabla_t \nabla_t^T\right]^{-1} \tag{4.18}$$

As illustrated above, when modelling using GAS models, the shape of observation density matters. If the time series is inclusive of changes in the shape of its distribution at some points in time, such changes can be captured by the time-varying parameters during the updating process of f_t. Figure 4.1 is an illustration of the nature of the GAS model, where it can be observed that the shape of distributions indeed changes over time. This indicates more flexibility is provided by GAS models and hence possibly improved power in characterising volatility movements.

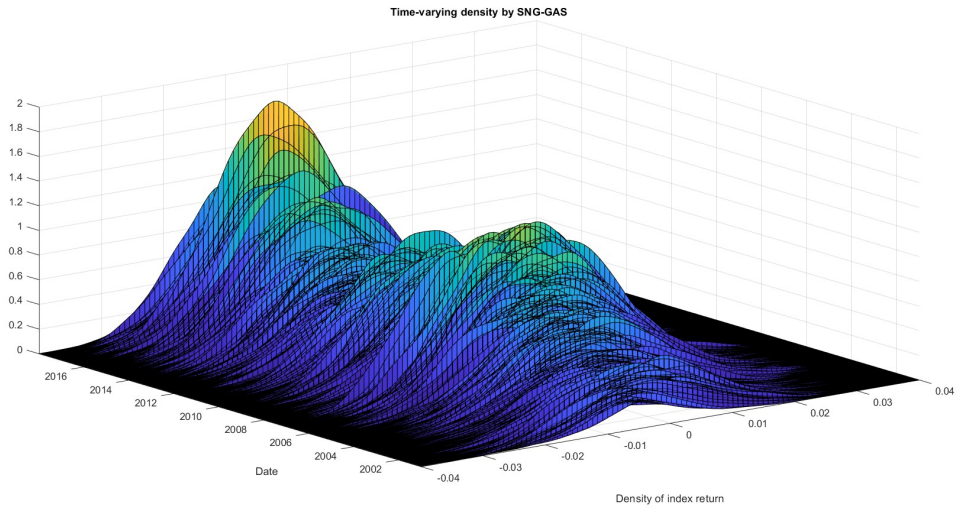

Figure 4.1: Time-varying density implied by SNG-GAS

Notes: This figure plots time-varying kernel densities implied by SNG-GAS. The X-axis represents the dates considered, and the Y-axis depicts the SNG-GAS simulated distribution of logarithmic returns.

4.3.2 Re-parameterisation

It is clear that the time-varying parameters, f_t, do not have boundaries from equation (4.15). In practice, restrictions are occasionally set on f_t, such that the updating process does not produce unreasonable values. For example, if the innovation term is assumed to follow a normal distribution and assumes that the variance h_t is time-varying, it is then necessary to make sure $f_t = h_t$ is positive, as variance cannot be negative. However, it is not necessary that the positiveness of f_t is guaranteed at each stage during the updating recursion within the estimation process. Under such circumstances, one solution we can resort to is re-parameterisation with respect to f_t via some link functions. The procedure is proposed and illustrated in Creal et al. (2012).

Here I provide an example in which the innovation term ξ_t is normally distributed by assumption, that is, $\xi_t \sim N(\mu_t, \sigma_t^2)$. Further assume that $\mu_t = 0$ and $f_t' = \ln \sigma_t^2$. Clearly, since $\sigma_t^2 = \exp f_t'$, the positiveness of σ_t^2 is ensured in the estimation process. The updated GAS updating process using re-parameterised time-varying parameters

f'_t is then

$$f'_{t+1} = \omega + \alpha S'_t \nabla'_t + \beta f'_t \tag{4.19}$$

where

$$\nabla'_t = \left[\frac{\partial \ln p\left(y_t \mid I_t, f_t; \theta\right)}{\partial f'_t} \right] \tag{4.20}$$

It is obvious that the original recursion in equation (4.15) is impacted by re-parameterisation from f_t to f'_t. Note that, after re-parameterisation, the coefficients ω, α and β are still functions of the static parameter θ. Define the mapping from f_t to f'_t as $\mathcal{M}$, and let $\mathcal{M}(f_t) = \ln \sigma_t^2$. The Jacobian of $\mathcal{M}$ can be calculated as:

$$\mathcal{M}'_t = \frac{\partial \mathcal{M}\left(f_t\right)}{\partial (f_t)} = \frac{1}{\sigma_t^2} \tag{4.21}$$

In accordance with Creal et al. (2012), if the logarithmic link function is applied for re-parameterisation of f_t given that $d = 1$, it can be shown that

$$S'_t \nabla'_t = \mathcal{M}'_t S_t \nabla_t \tag{4.22}$$

In this chapter, (4.22) is used to compute linked scaled scores after re-parameterisation.

4.3.3 Maximum Likelihood Estimation

One important step before option valuation is to produce forecasts of volatility using the GAS model. Essentially, the aim is to dynamically obtain estimations of the static hyperparameters θ via Maximum Likelihood Estimation (MLE) using historical data on a rolling window basis.

Mathematically,

$$\hat{\theta}_t = \arg\max_{\theta} \sum_{t=1}^{N} \ln p\left(y_t \mid f_t, I_{t-1}; \theta\right) \tag{4.23}$$

where the initial values $\hat{\theta}_1 = \frac{\omega}{1-\beta}$ is equal to the unconditional expectation of f'_t.

Obtaining the log-likelihood functions is quite straightforward. At each point in time, the updating equation outputs a conditional parameter that determines the density of innovation term ξ_t at each estimation. Connecting the conditional parameter to the log-likelihood function, the next step is to search for a set of θ that maximises the value of the log-likelihood function.

There are two challenging setbacks in the procedure of fitting GAS models via MLE. First, according to Catania and Billé (2017), the maximum likelihood estimation of the GAS model is an ongoing research topic. See for example, Harvey (2013), Blasques et al. (2014b), Blasques et al. (2014a) for some generally reported results. Harvey (2013) introduces the conditions for deriving the asymptotic distribution of the maximum likelihood estimations. Similarly, Blasques et al. (2016) proposes the conditions under which the GAS models are invertible, which is important to ensure asymptotic normality of the maximum likelihood estimations. Second, although the maximum likelihood estimators are asymptotically Gaussian, the numerical maximisation of the log-likelihood function can be challenging since appropriate starting values need to be selected with caution.

4.4 The Esscher Transform

The Esscher transform has been a popular tool in actuarial science. An overview can be found in (Tong et al., 2004). The Esscher Transform is adopted by Gerber and Shiu (1994) to find out an Equivalent Martingale Measure (EMM), which can be efficiently used in option valuation. The generalised version of the Esscher Transform for stochastic processes was proposed by Bühlmann et al. (1996). The advantageous feature of the Esscher Transform that distinguishes it from other methods, such as the LRNVR by Duan (1995) is that the former is more flexible in accounting for different distributions of the innovation term within observation-driven models like

GARCH and GAS, while the latter usually requires a normality assumption on the innovation term.

To show how the Esscher Transform works, let δ_t be a stochastic process and let I_t be the information set at time t. Suppose that δ_t is known given I_{t-1}. The conditional Moment Generating Function (MGF) of the stochastic process of the variable of interest y_t under the probability measure $\mathbb{P}$ can be expressed as:

$$M_{y_t|\Phi_{t-1}}(z) := E_{\mathbb{P}}\left(e^{zy_t} \mid \Phi_{t-1}\right) \tag{4.24}$$

Note that the moment generating function $M_{y_t|\Phi_{t-1}}(z)$ exists if and only if $E_{\mathbb{P}}\left(e^{zy_t} \mid \Phi_{t-1}\right) < \infty$. According to Bühlmann et al. (1996), the Radon-Nikodym derivative can be calculated by the sequence below:

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{I_t} = \prod_{k=1}^{t} \frac{e^{\theta_k y_k}}{M_{y_k|\Phi_{k-1}}(\theta_k)} \tag{4.25}$$

Define the sequence with Λ_t, it is clear that $\Lambda_0 = 1$. Following Tong et al. (2004), by the martingale property of Λ_t, the risk-neutralised moment generating function can be computed as follows:

$$M_{y_t|\Phi_{t-1}}(z;\theta_t) = \frac{M_{y_t|\Phi_{t-1}}(z+\theta_t)}{M_{y_t|\Phi_{t-1}}(\theta_t)} \tag{4.26}$$

where δ_t is called the conditional Esscher parameter. The Esscher Parameters can be dynamically solved from the equation below:

$$r = \ln\left\{M_{y_t|\Phi_{t-1}}(1;\delta_t^q)\right\} \tag{4.27}$$

After the sequence of δ_t^q is obtained, they have substituted back into (4.26) to finally compute the moment generating function under the $\mathbb{Q}$ measure. Under the equivalent martingale measure, option prices are simply discounted payoffs of the

underlying at the risk-free rate.

4.5 GAS Models for Option Pricing

This section introduces our models of interest. Within the GAS framework, the innovation term is assumed to follow a Gaussian distribution, shifted Gamma distribution and shifted negative Gamma distribution to construct corresponding GAS models. I show that if the error term is a Gaussian process, the derived model is consistent with that of Duan (1995), where LRNVR is the adopted method for risk-neutralisation. When the innovation term follows shifted Gamma (SG) or shifted negative Gamma distribution conditionally by assumption, the risk-neutralised error term of Gamma-process in Tong et al. (2004) and the error term of IG-GARCH in Christoffersen et al. (2006) can be retrieved. All models in this study are specified with GAS(1,1) structure as it is the most widely applied. The models are used to predict conditional volatility in the first stage and the predictions, together with the fitted static parameters, are utilised to price options via Monte Carlo simulations.

4.5.1 GAS with Gaussian Innovation

Recall that the process of the variable of interest can be expressed by (4.12) and (4.13), where μ_t is some mean process and (4.12) is called the mean equation of GARCH model. μ_t can have different specifications, such as ARMA structure. In this study, the common choice of μ_t in Duan (1995) is applied.

To start, suppose that the evolution of asset price is described by the stochastic differential equation below:

$$S_t = S_{t-1} \exp\left(r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t\right) \tag{4.28}$$

where r is the risk-free rate, λ is a constant and is normally a proxy of the unit price of risk. Clearly, the function of a stochastic process ξ_t here is to introduce shocks into the evolution of asset prices.

Taking logs on both sides:

$$\ln\left(\frac{S_t}{S_{t-1}}\right) = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t \tag{4.29}$$

Clearly, the left-hand side is the logarithmic returns of an asset. Denote it by y_t, the mean equation of the model is obtained as:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t \tag{4.30}$$

In the GAS environment, the updating process of factor f_t can be described by:

$$f_{t+1} = \omega + \alpha S_t \nabla_t + \beta f_t \tag{4.31}$$

Applying the logarithmic link function so that $f'_t = \ln f_t$, and the updating process changes to:

$$f'_{t+1} = \omega + \alpha S'_t \nabla'_t + \beta f'_t \tag{4.32}$$

To construct the Normal GAS model, assume that ξ_t is a Gaussian process with time-varying variance h_t, that is, $\xi_t \sim N(0, h_t)$. Note that under such circumstances, the asset logarithmic returns are also normally distributed with time-varying mean $\mu_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t$. The probability density function of ξ_t is:

$$f_{\xi_t}(y) = \frac{1}{\sqrt{2\pi h_t}} e^{\left(-\frac{y^2}{2h_t}\right)}, \quad \mu \in \mathbb{R}, \quad h_t \in \mathbb{R}^+ \tag{4.33}$$

Hence the log density can be calculated as follows:

$$\ln f_{\xi_t}(y) = \frac{-1}{2}\ln 2\pi h_t - \frac{y^2}{2h_t} \tag{4.34}$$

and the score function where variance is set as the time-varying parameter is:

$$\nabla_t = \frac{\partial \ln f_{\xi_t}(y)}{\partial h_t} = \frac{-1}{2h_t}\left(1 - \frac{y^2}{h_t}\right) \tag{4.35}$$

Recall that to ensure the positiveness of h_t, the logarithmic link function is applied. As a consequence, the Jacobian of the mapping function $\mathcal{M}$ in the normal GAS case is calculated as:

$$\mathcal{M}'_t = \frac{1}{h_t} \tag{4.36}$$

Given that the value of the scaling coefficient is set to 1, the scaling matrix of the score function is computed as:

$$S_t = \mathcal{I}_t^{-1} = -E_t\left[\nabla_t \nabla_t^T\right]^{-1} = 2h_t^2 \tag{4.37}$$

By (4.22), the parameterised recursion equation (4.32) can be rewritten as:

$$f'_{t+1} = \omega + \alpha \mathcal{M}'_t S_t \nabla_t + \beta f'_t \tag{4.38}$$

Equation (4.30) and (4.38) compose the normal GAS(1,1) model under the $\mathbb{P}$ measure. The parameters of the model can be estimated via MLE.

Since the aim is derivative pricing, it is still necessary to find out the risk-neutralised parameters. Following Tong et al. (2004), the first step is to compute the conditional Esscher parameter δ_t^q. Let $\sigma_t^2 = h_t$, the conditional moment generating function (MGF) of the normal distribution is:

$$M_{y_t|\Phi_{t-1}}(z) = \exp\left(\mu_t z + \sigma_t^2 z^2/2\right) \tag{4.39}$$

From (4.26) and (4.27):

$$\begin{aligned}\exp(r) &= M_{y_t|\Phi_{t-1}}(1;\theta_t^q)\\ &= \frac{M_{y_t|\Phi_{t-1}}(1+\delta_t^q)}{M_{y_t|\Phi_{t-1}}(\delta_t^q)}\\ &= \frac{\exp\left[(1+\delta_t^q)\left(r+\lambda\sqrt{h_t}-\frac{1}{2}h_t\right)+\frac{1}{2}(1+\delta_t^q)^2 h_t\right]}{\exp\left[\delta_t^q\left(r+\lambda\sqrt{h_t}-\frac{1}{2}h_t\right)+\frac{1}{2}\delta_t^{q2}h_t\right]}\end{aligned} \tag{4.40}$$

Solving for δ_t^q, we have $\delta_t^q = -\frac{\lambda}{\sqrt{h_t}}$. Substitute it back into (4.26), the conditional moment generating function under the $\mathbb{Q}$ measure can be written down as follows:

$$\begin{aligned}M_{y_t|\Phi_{t-1}}(z;\theta_t^q) &= \frac{M_{y_t|\Phi_{t-1}}(z+\theta_t^q)}{M_{y_t|\Phi_{t-1}}(\theta_t^q)}\\ &= \frac{\exp\left[(z+\theta_t^q)\left(r+\lambda\sqrt{h_t}-\frac{1}{2}h_t\right)+\frac{1}{2}(z+\theta_t^q)^2 h_t\right]}{\exp\left[\theta_t^q\left(r+\lambda\sqrt{h_t}-\frac{1}{2}h_t\right)+\frac{1}{2}\theta_t^{q2}h_t\right]}\\ &= \exp\left[z\left(r-\frac{1}{2}h_t\right)+\frac{1}{2}z^2 h_t\right]\end{aligned} \tag{4.41}$$

It can be observed from the result above that the conditional distribution of y_t under $\mathbb{Q}$ is still normal. The mean, compared to that under $\mathbb{P}$ is left shifted to $r-\frac{1}{2}h_t$, while the variance h_t remains unchanged. It is necessary to note that the shift is essentially made with respect to the original error term. In other words, the conditional mean of error term ξ_t under $\mathbb{Q}$ is $-\lambda\sqrt{h_t}$. Let $\xi_t^* = \xi_t + \lambda\sqrt{h_t}$ be the updated innovation term under $\mathbb{Q}$, we have $\xi_t^* \sim N(0,h_t)$. The GAS(1,1) model under risk-neutral probability measure can be described as:

$$y_t = r - \frac{1}{2}h_t + \xi_t^*, \quad \xi_t^* \sim N(0,h_t) \tag{4.42}$$

$$f_t^* = \ln h_t \tag{4.43}$$

$$f_{t+1}^* = \omega + \alpha \mathcal{M}'_t S_t \nabla_t + \beta f_t^* \tag{4.44}$$

4.5.2 GAS with Gamma Innovation

Tong et al. (2004) document that the GARCH model with shifted Gamma distributed innovation can be used to better deal with the skewed characteristic of asset returns. Assuming that the innovation process can be described by conditional shifted Gamma distribution, I construct the SG-GAS model. In the same way, I also assume that the innovation term follows conditional shifted negative Gamma distribution and fits the SNG-GAS model accordingly.

Let X_t be a stochastic process under $\mathbb{P}$ measure. Suppose that for any $t \in T$, X_t are independently distributed. Let $G(a, b)$ define a Gamma distribution with shape parameter a and scale parameter b. I assume that $X_t \sim G(a, b_t)$. Note that the scale parameter b_t is set to be changing over time in the GAS mechanism. For simplicity, I omit the subscript t in the model derivation below. The probability density function of X_t is given by:

$$f_{X_t}(y) = \frac{1}{\Gamma(a)b^a} y^{a-1} e^{-\frac{y}{b}} \tag{4.45}$$

Furthermore, the mean of X_t is:

$$\mu = ab \tag{4.46}$$

and the variance can be computed as:

$$h = ab^2 \tag{4.47}$$

It should be noted that the support of the Gamma distribution is $y \in (0, \infty)$. However, one fact of asset returns is that it can be both positive and negative. In this case, I consider applying shifted Gamma distribution to model the innovation term. Essentially, the innovation term is shifted such that only positive values are considered. Following Tong et al. (2004), first define the innovation term of asset

returns by ξ_t, which can be calculated as:

$$\xi_t = \sqrt{h_t}\left(\frac{X_t - ab}{\sqrt{ab^2}}\right) \tag{4.48}$$

It is easy to see that under $\mathbb{P}$, the shifted innovation process ξ_t is a standardised Gamma distribution that has a zero mean and time-varying variance h_t. Under such circumstances, the logarithmic returns process y_t can be re-written as:

$$\begin{aligned} y_t &= r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \xi_t \\ &= r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t} + \frac{1}{b}\sqrt{\frac{h_t}{a}}X_t \end{aligned} \tag{4.49}$$

From the scaling property of Gamma distribution, if

$$X_t \sim G(a, b) \tag{4.50}$$

then

$$cX_t \sim G(a, cb) \tag{4.51}$$

It is easy to see that the last component of (4.49) follows Gamma distribution with shape parameter a and scale parameter $\sqrt{\frac{h_t}{a}}$. The moment-generating function of y_t can be written as:

$$\begin{aligned} M_{y_t}(\theta) &= (1 - b\theta)^{-a} \\ &= \left(\frac{\sqrt{\frac{a}{h_t}}}{\sqrt{\frac{a}{h_t}} - \theta}\right)^a \exp\left[\left(r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t}\right)\theta\right] \end{aligned} \tag{4.52}$$

The moment generating function of y_t after the Esscher Transform is given by:

$$M_{y_t|\Phi_{t-1}}(z;\theta_t) = \left(\frac{\sqrt{\frac{a}{h_t}} - \theta_t}{\sqrt{\frac{a}{h_t}} - \theta_t - z}\right)^a \exp\left[\left(r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t}\right)z\right] \tag{4.53}$$

Following (4.26) and (4.27), the conditional Esscher parameter under measure $\mathbb{Q}$

can be solved for as:

$$\theta_t^q = \sqrt{\frac{a}{h_t}} - \left[1 - \exp\left(\frac{\lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t}}{a}\right)\right]^{-1} \tag{4.54}$$

Recall that b_t is the scale parameter of y_t under $\mathbb{P}$. Following Tong et al. (2004), let b_t^q denote the corresponding scale parameter under $\mathbb{Q}$, then b_t^q is defined as:

$$\begin{aligned} b_t^q &= b_t - \theta_t^q \\ &= \left[1 - \exp\left(\frac{\lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t}}{a}\right)\right]^{-1} \end{aligned} \tag{4.55}$$

It is obvious that, under $\mathbb{Q}$, y_t again is a shifted Gamma process with shape parameter a, scale parameter b_t^q and shift parameter $-r - \lambda\sqrt{h_t} + \frac{1}{2}h_t + \sqrt{ah_t}$. That is, from $\mathbb{P}$ to $\mathbb{Q}$, the Esscher transform solely works on modifying the scale parameter. Denote the error term of y_t under $\mathbb{Q}$ by $\xi_t^* \sim G(a, b_t^q)$. The density function of ξ_t^* is:

$$f_{\xi_t^*}(y) = \frac{1}{\Gamma(a)(b_t^q)^a} y^{a-1} e^{-\frac{y}{b_t^q}} \tag{4.56}$$

The corresponding logarithmic density function is given by:

$$\ln f_{\xi_t^*}(y) = (a-1)\ln y - \frac{y}{b_t^q} - \ln\Gamma(a) - a\ln b_t^q \tag{4.57}$$

Let $f_t^* = \ln b_t^q$, the score function, the Jacobian of the mapping function and the scaling function are given by:

$$\nabla_t^* = \frac{\partial \ln f_{\xi_t^*}(y)}{\partial b_t^q} = \frac{y}{(b_t^q)^2} - \frac{a}{b_t^q} \tag{4.58}$$

$$\mathcal{M}'^*_t = \frac{1}{b_t^q} \tag{4.59}$$

and

$$S_t^* = -E_t\left[\nabla_t^*(\nabla_t^*)^T\right]^{-1} = \frac{(b_t^q)^2}{a} \tag{4.60}$$

Hence, the GAS model with shifted Gamma innovation (SG-GAS) under $\mathbb{Q}$ is given by:

$$y_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t - \sqrt{ah_t} + \xi_t^*, \quad \xi_t^* \sim G(a, b_t^q) \tag{4.61}$$

$$f_t^* = \ln b_t^q \tag{4.62}$$

$$f_{t+1}^* = \omega + \alpha \mathcal{M}_t'^* S_t^* \nabla_t^* + \beta f_t^* \tag{4.63}$$

For the purpose of option pricing, (4.61), (4.62), and (4.63) constitute the process applied to simulate the evolution of underlying asset prices under $\mathbb{Q}$. It is noteworthy that the risk-neutralisation of shifted Gamma innovation term is different from that in the Gaussian case, where the Esscher transform does not change the time-varying parameter h_t, hence functions $\mathcal{M}'$, S_t and ∇_t obtained under $\mathbb{P}$ can be used directly to generate underlying prices. The only operation needed is to shift the mean of y_t. In contrast, the scale parameter of shifted Gamma innovation changes to b_t^q under $\mathbb{Q}$. This means that the functions $\mathcal{M}'^*$, S_t^* and ∇_t^* under $\mathbb{Q}$ are different from their counterparts under $\mathbb{P}$.

In terms of the GAS model with shifted negative Gamma error term (SNG-GAS), the derivation procedure is quite similar. The difference is that, at the very beginning of model construction, the Gamma process $X_t \sim G(a, b)$ changes to $X_t \sim -G(a, b)$. As a result, the scaled error term follows a negative Gamma distribution denoted as $-G(a, \sqrt{\frac{h_t}{a}})$.

4.6 Simulation Study Design

Inspired by Tong et al. (2004) and Zhu and Ling (2015), in this section, I implement a series of Monte Carlo simulation studies to investigate and compare the properties of the models of interest. Four models are involved: GARCH, GAS, SG-GAS and SNG-GAS. First, using each model, I simulate the underlying price process under

both $\mathbb{P}$ and $\mathbb{Q}$. The aim is to show how risk-neutralisation makes a difference to the distributions of simulated underlying prices. Second, under $\mathbb{Q}$, I make comparisons between the GAS model with Gaussian innovation and the GAS model with shifted negative Gamma innovation. Third, I inspect and compare the implied volatility smile given by these models.

4.6.1 Innovation from $\mathbb{P}$ to $\mathbb{Q}$

Using the logarithmic returns of the Dow Jones Industrial Average index between 31/12/2012 and 31/12/2017 and the corresponding risk-free rate proxied by 3-month T-bill rates, I train GARCH, GAS, SG-GAS and SNG-GAS models, and obtain a set of static parameters for each model. These parameters are used to generate 20000 paths of the underlying price evolution under $\mathbb{P}$ and $\mathbb{Q}$ respectively. Conventionally, the risk-free rate is set to zero, and the number of steps (or time to maturity in option pricing language) is set to 30 days. For each model, I keep and plot the distribution of its terminal values from simulations. Figure 4.2 exhibits comparisons between terminal price distributions under $\mathbb{P}$ and $\mathbb{Q}$ for each model.

First, for the GARCH and GAS with normal innovations, our finding is consistent with existing literature, that the risk-neutralisation by the Esscher transform is essentially making a leftwards horizontal shift to the original distribution under $\mathbb{P}$. It can also be observed that, compared to the GAS model, the GARCH model tends to generate more extreme values or outliers. This is reflected by its significantly longer tails. Second, consistent with our previous illustration, GAS models with shifted Gamma innovations tend to be more negatively skewed under $\mathbb{Q}$ compared to those under $\mathbb{P}$. This is because the Esscher transform only affects the scale parameter b and the skewness of the Gamma distribution is equal to $\frac{2}{\sqrt{b}}$, which indicates that the skewness of the Gamma distribution is solely determined by its scale parameter.

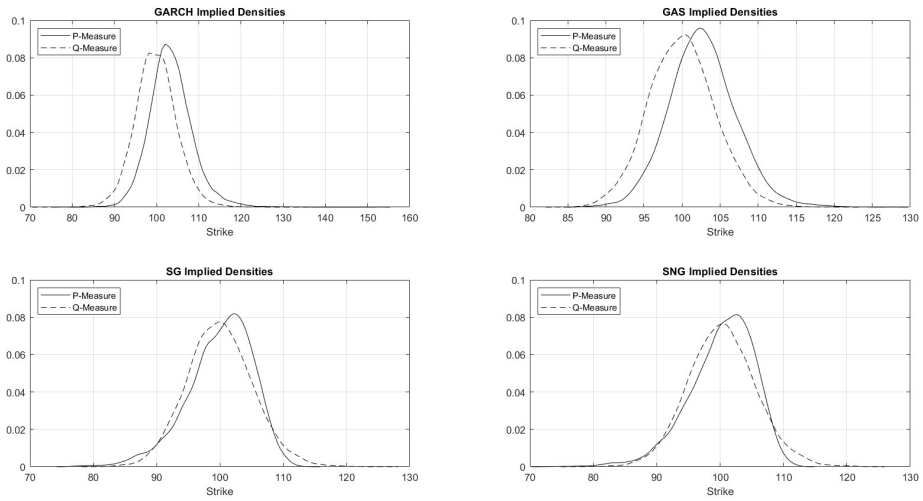

Figure 4.2: Simulated density under $\mathbb{P}$ and $\mathbb{Q}$

Notes: This figure plots densities of underlying prices simulated by GARCH, GAS, SG-GAS and SNG-GAS. By convention, the risk-free rate is set to zero, and the number of simulation steps is set to 30. For each model, 20000 paths of underlying price evolution are generated under P and Q respectively.

4.6.2 Innovation From Gaussian to SNG

Adopting an identical sample as in the above subsection, I investigate behaviours of the GAS model with normal innovation and the GAS model with shifted negative Gamma innovation under risk-neutral measures. I do not include SG-GAS because its behaviour is quite similar to that of SNG-GAS. In terms of the number of simulation steps, in addition to 30 days, I also include 90 days, 180 days and 360 days so that it is possible to inspect how assumptions on the distribution of asset returns differentiate models in simulating asset price processes across different horizons. From Figure 4.3, It can be observed that under the GAS mechanism, simulated ending prices from both GAS and SNG-GAS models tend to be identically widely distributed. Furthermore, as the number of simulation steps increases, the ending prices become more and more widely distributed, which is consistent with intuition. Second, it is obvious that, compared to the GAS model, the distribution of simulated ending values from the SNG-GAS model has significantly fatter tails. In other words, if the SNG-GAS model is adopted to run the simulations, simulated

prices that are more deviated from the starting value are more frequently observed. This is very important, as this means that the GAS models with shifted Gamma innovation tend to overestimate the tail risk, which is underestimated by BS and GARCH models, and thus overestimate prices of out-of-the-money options that are related to tail risk. Hence, the proposed SNG-GAS model tends to have the overall best performance.

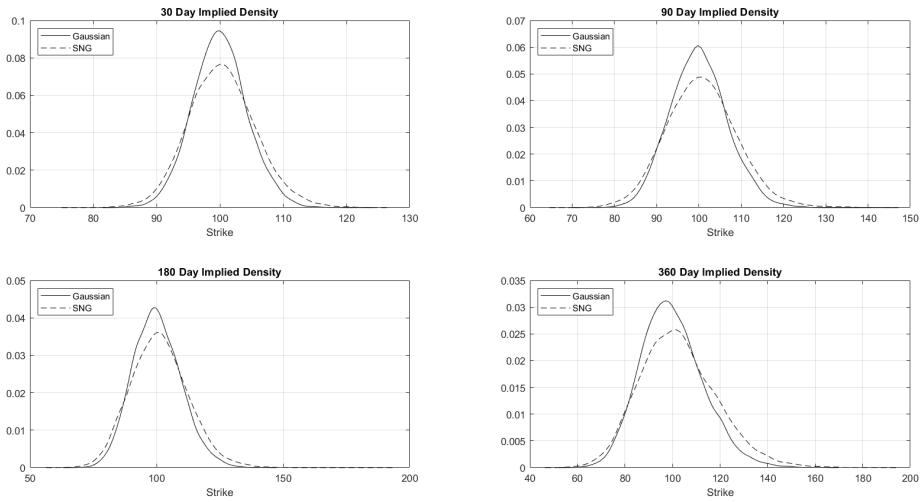

Figure 4.3: Simulated density by models under $\mathbb{Q}$

Notes: This figure depicts densities of underlying prices simulated by GARCH and SNG-GAS under $\mathbb{Q}$ measure. By convention, the risk-free rate is set to zero, and the number of simulation steps is set to 30, 90, 180, and 360. For each model, 20000 paths of underlying price evolution are generated. The solid lines represent the density implied by GARCH and the dashed lines, SNG-GAS.

4.6.3 Volatility Smile

In this subsection, I take a further step into checking how implied volatility changes in accordance with different assumptions on innovation terms and hence different models. Following the procedure in Zhu and Ling (2015), to obtain to parameters for running simulations, I first fit GARCH, GAS, SG-GAS and SNG-GAS models using a historical data set that consists of DJIA log-returns and corresponding 3-month T-bill rates. Within the data set, 5538 observations ranging from 03/01/1996 to 31/12/2017 are included. The estimated parameters are used to estimate European

call option prices via Monte Carlo simulations. Similarly to Duan (1995) and Zhu and Ling (2015), in the simulation procedure, the risk-free rate is set as $r = 0$, and the strike price is set as $K = 1$ so that by adjusting the input of spot price S_0, a series of moneyness S/K ranging from 0.9 to 1.1 is obtained. The time to maturity (TTM) is between 30 days, 90 days and 180 days, and the starting value of variance that is used to initiate the simulation process is $0.8h_0$ or h_0, where h_0 is the variance of log-returns series adopted in the model training process.

Applying each model, I run 20000 simulations for each S/K and for each time to maturity to obtain estimated call option prices. Then I inversely apply the Black-Scholes model to convert these call option prices into implied volatility. Figures 4.4, 4.5 and 4.6 depict implied volatility given by different option pricing models across different times to maturity and starting from different initial variances. I document a few findings from these plots. First, the most significant difference between these models' implied volatility is that, while others tend to show volatility smile to some extent, the SNG-GAS implied volatility exhibits a volatility smirk across all maturities and initial variances. Zhu and Ling (2015) find that the implied volatility curve is U-shaped when the innovation term is assumed to follow the normal distribution, while the curve tends to be skewed when the innovation terms are assumed to follow a distribution, such as shifted Gamma distribution, that takes into consideration the leverage effect of asset returns to volatility. Our finding is consistent with Zhu and Ling (2015) and it indicates that the SNG-GAS model may overestimate out-of-the-money call options.

Second, models assumed to have Gamma innovation tend to show higher implied volatility compared to those with Gaussian innovations. More specifically, the SNG-GAS model implied volatility is the highest among all candidate models, this is followed by SG-GAS implied volatility. Such behaviour is observed across TTM from 30 days to 180 days, and as TTM stretches, SNG-GAS and SG-GAS start to imply more and more significantly higher volatility compared to GARCH and GAS

models.

Third, consistent with Zhu and Ling (2015), I document as well that the level of implied volatility curves is determined by the input of initial variance. The curves shift upwards as the initial variance increases. Fourth, I report that the U-shape of each implied volatility curve gradually disappears as the time to maturity extends. This is again documented in Zhu and Ling (2015).

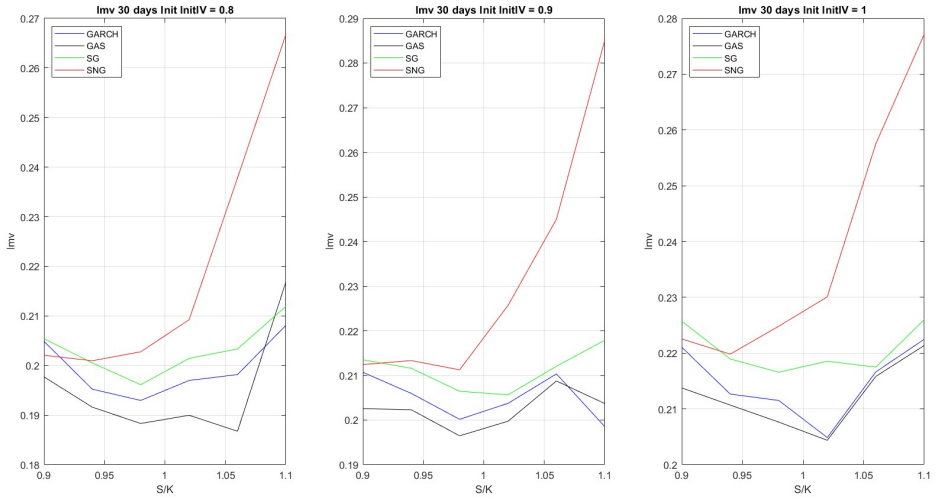

Figure 4.4: Implied volatility smile at 30-day horizon

Notes: This figure plots the curve of 30-day volatility implied by GARCH, GAS, SG-GAS, and SNG-GAS. Consistent with existing literature, the risk-free rate is set to be zero, and initial variance is set to be 80%, 90% and 100% of long-term variance. The strike price is set to be 1, such that in each scenario, the implied volatility is computed across moneyness ranging from 0.9 to 1.1.

4.7 Model Application

4.7.1 Data

In this section, I adopt real-world data to test the performance of our models. I obtain the DJIA logarithmic returns series between 04/01/1996 and 31/12/2017 from Bloomberg, and the corresponding risk-free rate proxied by 3-month T-bill rates are obtained from the Federal Reserve at St. Louis. This data set functions

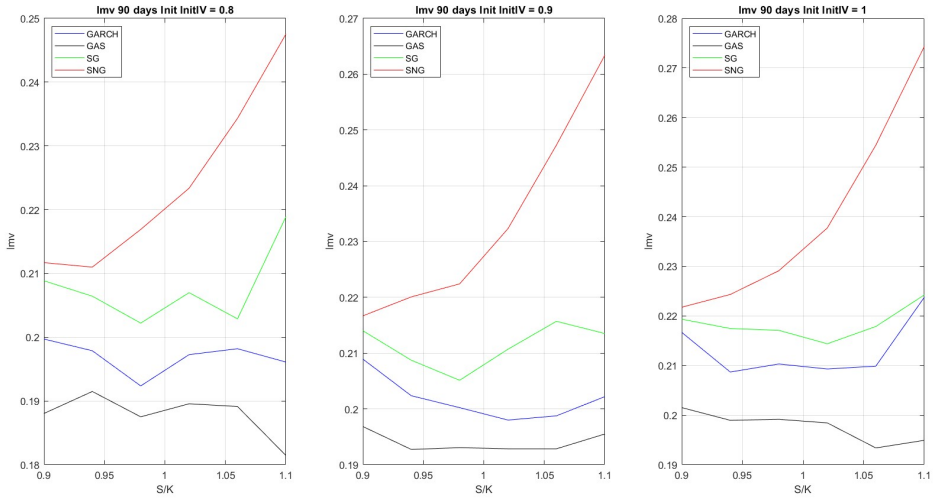

Figure 4.5: Implied volatility smile at 90-day horizon

Notes: This figure plots the curve of 90-day volatility implied by GARCH, GAS, SG-GAS, and SNG-GAS. Consistent with existing literature, the risk-free rate is set to be zero, and initial variance is set to be 80%, 90% and 100% of long-term variance. The strike price is set to be 1, such that in each scenario, the implied volatility is computed across moneyness ranging from 0.9 to 1.1.

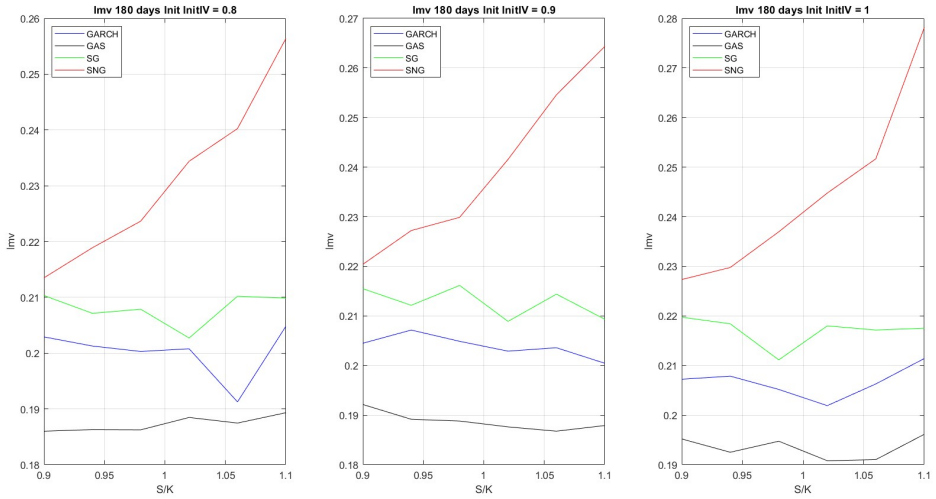

Figure 4.6: Implied volatility smile at 180-day horizon

Notes: This figure plots the curve of 180-day volatility implied by GARCH, GAS, SG-GAS, and SNG-GAS. Consistent with existing literature, the risk-free rate is set to be zero, and initial variance is set to be 80%, 90% and 100% of long-term variance. The strike price is set to be 1, such that in each scenario, the implied volatility is computed across moneyness ranging from 0.9 to 1.1.

as the training set of our models. On a rolling window basis, I use log-returns and risk-free rates over the past 5 years to dynamically make predictions of daily parameters that define the distributions of innovation terms, which are then adopted for pricing options via a simulation approach. To evaluate the performance of our models, I introduce real option price data over the same sample period. The options data is obtained from OptionMetrics. By convention, I only keep option prices on each Wednesday between 04/01/1996 and 31/12/2017 for making comparisons. Furthermore, I remove prices of in-the-money options and options with time to maturity shorter than one week or longer than 2 years are also removed. After applying these filters, we are left with 42697 options. I apply SG-GAS and SNG-GAS models to estimate the prices of these options and compare our estimates to option prices in the real world. In addition to the two GAS models above, I also set the Black-Sholes model as the benchmark and apply GARCH and GAS models as comparisons. For each model, I quantify its accuracy in pricing options by computing its mean absolute percentage error (MAPE) as:

$$\text{MAPE} = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \tag{4.64}$$

where A_t represents the actual option prices and F_t stands for corresponding estimated option prices from the models.

I inspect the pricing accuracy of our models across moneyness ranging from 0.96 to 1.06. By converting put option prices into call option prices, moneyness is a way for us to distinguish puts from calls. To investigate how models perform as expiration differentiates, I also evaluate model performance across time to maturity from 30 days to 720 days. The performance of our models may vary over time. It is more appealing to us to find out how the models perform during market turmoil. Hence, in addition to the full sample, I also divide it into sub-periods before, during and after the 2008 financial crisis.

4.7.2 Empirical Results

The pricing errors over the full sample are concluded in Table 4.1. First of all, it is obvious that all observation-driven models significantly outperform the benchmark BS model with significantly lower pricing errors. To be more specific, the GAS models with shifted Gamma innovation (SG-GAS) and shifted negative Gamma innovation (SNG-GAS) significantly dominate other candidate models. The overall pricing error calculated using the full sample (42697 observations) is 61.60% using the BS model, and this number falls to 55.67% when the GARCH model is utilised. For the GAS model, the pricing error decreases to 53.51%. It is noteworthy that the proposed SG-GAS and SNG-GAS models, with a MAPE of 51.51% and 48.94%, tend to have the overall best performance in option valuation. This is also one of the main findings of this study.

Table 4.1: MAPE of estimated option prices - full sample

	No. obs	BS	GARCH	GAS	SG-GAS	SNG-GAS
M<=0.96	6466	1.1015	1.0723	0.8043	0.7081	0.6834
0.96<M<=0.98	3916	0.7208	0.6660	0.6522	0.7727	0.6249
0.98<M<=1.02	10924	0.2956	0.2604	0.2567	0.3594	0.3124
1.02<M<=1.04	4541	0.3697	0.2978	0.3059	0.3491	0.3144
1.04<M<=1.06	3412	0.5035	0.4070	0.4254	0.3972	0.3862
M>1.06	13438	0.7241	0.6432	0.7030	0.5597	0.5858
TTM<=30	14609	0.7712	0.6031	0.5693	0.6260	0.5428
30<TTM<=60	12852	0.5516	0.5177	0.4993	0.4978	0.4595
60<TTM<=90	4064	0.5386	0.5293	0.5319	0.4490	0.4463
90<TTM<=180	6302	0.5258	0.5551	0.5374	0.4270	0.4615
180<TTM<=360	3494	0.4982	0.5567	0.5365	0.4076	0.4770
TTM>360	1376	0.5109	0.5167	0.5018	0.3701	0.4888
CALLS	14910	0.7888	0.7469	0.6215	0.6607	0.5940
PUTS	27787	0.5233	0.4547	0.4887	0.4370	0.4333
TOTAL	42697	0.6160	0.5567	0.5351	0.5151	0.4894

Notes: This table reports the mean absolute percentage errors between model-based option price estimations (BS, GARCH, GAS, SG-GAS, and SNG-GAS) and real option prices between 04/01/1996 and 31/12/2017. The pricing performance is evaluated separately by segmenting options according to moneyness ranging from below 0.96 to above 1.06 and time to maturity ranging from less than 30 days to more than 360 days. The pricing performance of all models is also evaluated with respect to calls and puts respectively.

It is also documented in this study that the superiority of SG-GAS and SNG-GAS models seems to be enhanced during a normal market environment. Using a sub-sample of options prices before the 2007-2008 financial crisis (18442 observations), the MAPEs are computed and reported in Tables 4.2 and 4.3. It can be found

that the pricing errors of SG-GAS and SNG-GAS models tend to be significantly smaller (43.24% and 39.48% respectively) than those from the BS model, GARCH and GAS models (42.09%, 46.95% and 48.34% respectively). This pattern can still be observed using the post-crisis sub-sample (14766 observations), where SG-GAS and SNG-GAS tend to have pricing errors of 54.86% and 50.16% respectively, while all the other competing models are found to have pricing errors above 60% and the highest pricing error of 65.83% is produced by the BS model. In a timeline sense, comparing the pricing error of the BS model before the financial crisis to its performance after the crisis, it is easy to see that the MAPE increases from 42.09% to 65.83%.

Note that while other parametric models all witnessed weakened pricing performance from pre-crisis time to post-crisis time, the BS model seems to have the most significant drop in terms of pricing accuracy. This indicates that the BS model might be relatively more accurate in earlier years, especially before the 2007-2008 financial crisis. After the financial crisis, when it comes to years that are closer to the present, option valuation generally becomes more difficult. This is indicated by systematic drops in pricing accuracy across all models. Notably, the BS model seems to suffer a bigger loss in pricing accuracy compared to other models.

It is interesting to look at the empirical results obtained using a sub-sample during the 2007-2008 financial crisis (9489 observations). The pricing results are reported in Table 4.4. It is obvious that the proposed SG-GAS and SNG-GAS models fail to beat the other candidate models under such circumstances. The pricing errors of the two proposed models are 62.35% and 65.42% respectively, and the only model they outperform is the BS model, with a high MAPE of 92.94%. The GARCH and GAS models are found to have lower pricing errors of 58.07% and 51.25% respectively. This indicates that, during market turmoil, normal innovation could be a better choice for financial returns modelling compared to Gamma innovation which introduces skewness. It is potentially indicated that the selection of innovation

terms is of great importance in order to better facilitate parametric models for the purpose of option valuation. It can also be reported that the GAS model, with a pricing error of 51.25%, significantly outperforms its GARCH counterpart, with a pricing error of 58.07%. This is possible because of the less fierce response of GAS models to shocks from extreme observations compared to GARCH models.

To find out the source of the outperformance of SG-GAS and SNG-GAS models, option valuation errors from all models are calculated and reported across a variety of moneyness levels, denoted as the quotient between spot prices and corresponding strikes, and across time to maturity from less than 30 days to more than 360 days. It can be observed instantly that the improved pricing performance of SG-GAS and SNG-GAS mainly originates from their ability to more accurately price out-of-the-money options, especially those Deep Out-of-the-Money (dOTM) options. To see this, for dOTM calls with moneyness below 0.96, SG-GAS and SNG-GAS tend to have much lower pricing errors of 70.81% and 68.34% respectively. This is compared to MAPEs of 110.15%, 107.23% and 80.43% from the BS, GARCH and GAS models respectively. For dOTM puts with moneyness above 1.06, the MAPEs of proposed SG-GAS and SNG-GAS models are 55.97% and 58.58%. In comparison, the pricing errors of the BS, GARCH and GAS models are 72.41%, 64.32% and 70.30% respectively. Note that this is consistent with Black (1975) and Duan (1995), in which it is documented that the BS model tends to underestimate deep OTM options.

Clearly, the GAS models with Gamma innovations tend to perform better in pricing OTM options, while it is the opposite for near-the-money options with moneyness ranging from 0.98 to 1.02. In comparison, the SG-GAS and SNG-GAS models tend to have the worst overall performance, with pricing errors of 35.94% and 31.24%. Note that the BS, GARCH and GAS models under such circumstances have obviously lower pricing errors of 29.56%, 26.04% and 25.67% respectively. It is worth noticing that for other OTM options that are further out-of-the-money, the out-

performance of SG-GAS and SNG-GAS is still observed. For example, for OTM puts with moneyness between 1.04 and 1.06, the pricing errors of SG-GAS and SNG-GAS models are 39.72% and 38.62% respectively, while the MAPEs of the BS, GARCH and GAS models are 50.35%, 40.70% and 42.54%. Given that the number of near-the-money options is 10924 and the number of OTM options is 31773, it is reasonable to think that it is the superiority of SG-GAS and SNG-GAS models in pricing OTM options that drives their overall outperformance in option pricing.

Option valuations across different times to maturity are also inspected in this study. It can be observed from Table 4.1 that SG-GAS and SNG-GAS models perform relatively better in pricing options with longer maturities. For instance, for options that expire within 30 days, the pricing errors of SG-GAS and SNG-GAS are 62.60% and 54.28%, this is compared to 77.12%, 60.31% and 56.93% from BS, GARCH and GAS models respectively. The pricing error of SG-GAS and SNG-GAS models with respect to options that expire in more than 360 days are 37.01% and 48.88%. which is significantly lower than 51.09%, 51.67% and 50.18% from BS, GARCH and GAS models. Note that from within 30-day time to maturity to longer than 360-day time to maturity, the pricing errors of these two proposed models drop from 62.60% and 54.28% to 37.01% and 48.88%. These are more remarkable improvements compared to the BS, GARCH and GAS models. Hence, it can be inferred that SG-GAS and SNG-GAS models tend to have a comparative advantage in pricing options with longer lifespans.

In terms of performance in pricing calls and puts, it can be observed from Table 4.1 that, while all models tend to perform better in pricing put options, SG-GAS and SNG-GAS consistently outperform BS, GARCH and GAS models in pricing puts. For the valuation of call options, these two models fail to beat all the other models at the same time.

Table 4.2: MAPE of estimated option prices - pre-crisis

	No. obs	BS	GARCH	GAS	SG-GAS	SNG-GAS
M<=0.96	2834	0.5067	0.7487	0.7401	0.6086	0.4940
0.96<M<=0.98	1677	0.3622	0.4968	0.5398	0.6799	0.4362
0.98<M<=1.02	4395	0.1858	0.2229	0.2236	0.2996	0.2257
1.02<M<=1.04	1893	0.2746	0.2822	0.2671	0.2893	0.2670
1.04<M<=1.06	1522	0.4096	0.3748	0.3633	0.3184	0.3327
M>1.06	6121	0.6142	0.5911	0.6323	0.4511	0.5140
TTM<=30	4867	0.4943	0.4931	0.4938	0.5705	0.4457
30<TTM<=60	5436	0.3949	0.4604	0.4825	0.4526	0.3837
60<TTM<=90	1975	0.4042	0.4791	0.5145	0.3755	0.3764
90<TTM<=180	3350	0.4167	0.4754	0.4883	0.3415	0.3724
180<TTM<=360	2069	0.3714	0.4454	0.4548	0.3178	0.3787
TTM>360	745	0.3318	0.3957	0.3961	0.2616	0.3380
CALLS	5902	0.4087	0.5725	0.5840	0.6006	0.4423
PUTS	12540	0.4266	0.4210	0.4360	0.3533	0.3725
TOTAL	18442	0.4209	0.4695	0.4834	0.4324	0.3948

Notes: This table reports the mean absolute percentage errors between model-based option price estimations (BS, GARCH, GAS, SG-GAS, and SNG-GAS) and real option prices before the 2007-2008 financial crisis. The pricing performance is evaluated separately by segmenting options according to moneyness ranging from below 0.96 to above 1.06 and time to maturity ranging from less than 30 days to more than 360 days. The pricing performance of all models is also evaluated with respect to calls and puts respectively.

Table 4.3: MAPE of estimated option prices - post-crisis

	No. obs	BS	GARCH	GAS	SG-GAS	SNG-GAS
M<=0.96	1397	1.3379	1.9101	1.4895	0.9810	0.9527
0.96<M<=0.98	1469	1.0450	1.0361	0.9803	1.0519	0.8987
0.98<M<=1.02	4845	0.3508	0.3125	0.2957	0.4078	0.3670
1.02<M<=1.04	1859	0.4018	0.3022	0.3110	0.3400	0.2773
1.04<M<=1.06	1238	0.5377	0.4384	0.4616	0.3789	0.3349
M>1.06	3958	0.8094	0.7056	0.7493	0.5327	0.5173
TTM<=30	6279	0.7345	0.6634	0.6399	0.6624	0.5527
30<TTM<=60	4402	0.5540	0.5874	0.5461	0.4738	0.4217
60<TTM<=90	1183	0.6372	0.6483	0.6141	0.4488	0.4231
90<TTM<=180	1691	0.6368	0.6724	0.6372	0.4393	0.4861
180<TTM<=360	853	0.6597	0.7796	0.7029	0.4859	0.6021
TTM>360	358	0.7715	0.7871	0.6805	0.4676	0.6815
CALLS	5468	0.8270	0.9622	0.8245	0.7872	0.7231
PUTS	9298	0.5590	0.4668	0.4905	0.4083	0.3714
TOTAL	14766	0.6583	0.6503	0.6142	0.5486	0.5016

Notes: This table reports the mean absolute percentage errors between model-based option price estimations (BS, GARCH, GAS, SG-GAS, and SNG-GAS) and real option prices after the 2007-2008 financial crisis. The pricing performance is evaluated separately by segmenting options according to moneyness ranging from below 0.96 to above 1.06 and time to maturity ranging from less than 30 days to more than 360 days. The pricing performance of all models is also evaluated with respect to calls and puts respectively.

Table 4.4: MAPE of estimated option prices - during-crisis

	No. obs	BS	GARCH	GAS	SG-GAS	SNG-GAS
M<=0.96	2235	1.7078	0.9587	0.4574	0.6635	0.7554
0.96<M<=0.98	770	0.8833	0.3282	0.2713	0.4422	0.5136
0.98<M<=1.02	1684	0.4236	0.2084	0.2313	0.3765	0.3812
1.02<M<=1.04	789	0.5224	0.3249	0.3869	0.5144	0.5156
1.04<M<=1.06	652	0.6578	0.4227	0.5016	0.6161	0.6086
M>1.06	3359	0.8240	0.6644	0.7772	0.7894	0.7974
TTM<=30	3463	1.2267	0.6485	0.5473	0.6381	0.6612
30<TTM<=60	3014	0.8309	0.5195	0.4612	0.6142	0.6514
60<TTM<=90	906	0.7028	0.4830	0.4625	0.6096	0.6289
90<TTM<=180	1261	0.6668	0.6093	0.5341	0.6379	0.6651
180<TTM<=360	572	0.7162	0.6266	0.5840	0.6156	0.6462
TTM>360	273	0.6579	0.4921	0.5556	0.5383	0.6476
CALLS	3540	1.3632	0.7051	0.3703	0.5653	0.6475
PUTS	5949	0.6713	0.5067	0.5972	0.6582	0.6582
TOTAL	9489	0.9294	0.5807	0.5125	0.6235	0.6542

Notes: This table reports the mean absolute percentage errors between model-based option price estimations (BS, GARCH, GAS, SG-GAS, and SNG-GAS) and real option prices during the 2007-2008 financial crisis. The pricing performance is evaluated separately by segmenting options according to moneyness ranging from below 0.96 to above 1.06 and time to maturity ranging from less than 30 days to more than 360 days. The pricing performance of all models is also evaluated with respect to calls and puts respectively.

4.8 Conclusion

While the majority of existing studies in option pricing with parametric models tend to develop models within the framework of GARCH, this chapter innovatively proposes to build time-varying parameter models for option valuations within the framework of GAS models. One attractive feature of the GAS model is that different from the widely applied GARCH models, in GAS models, the time-varying parameter that drives the updating mechanism is not necessarily the variance of innovation term. Instead, the driving parameter of the GAS model is usually connected with variance via some channel that is closely connected with variance, such as the scale parameter. Another appealing feature of the GAS model is that it tends to respond less fiercely to large values of observations, as the impact from extreme values can be absorbed by the driving score functions. This feature is very useful in financial returns modelling, as the occurrence of large observations might be simply due to the fat-tailed nature of financial returns, and it does not necessarily cause an increase in volatility. Such property of GAS models leads to a more appropriate forecast

of volatility and consequently improved accuracy in option valuations. During periods of market stress, like the 2007 - 2008 financial crisis, the advantage of GAS models becomes more prominent, as during great market uncertainties, fat-tailed distributions of asset returns are more frequently observed.

For the purpose of option valuation, I construct GAS models with shifted Gamma and shifted negative Gamma innovations. The proposed models are hence named SG-GAS and SNG-GAS models. Consistent with Tong et al. (2004), I choose to incorporate Gamma innovation terms in order to take into consideration the skewness exhibited in financial returns. The general procedure of option valuation in this study is analogous to Duan (1995), Duan (1999) and Tong et al. (2004), where historical financial returns are used to train parametric models in the first stage. Then in the second stage, with the estimated parameters under the physical probability measure $\mathbb{P}$, the Esscher transform is applied to find out corresponding model specifications under risk-neutral measure $\mathbb{Q}$, following the methodology in Gerber and Shiu (1994) and Tong et al. (2004). Finally, in the third stage, the parametric models under the $\mathbb{Q}$ measure are used to generate simulated paths of underlying price evolution, and the expected payoff is calculated and discounted to obtain option price estimations.

Before model evaluation and comparison using real-world option data, a series of Monte Carlo simulation studies are implemented in order to reveal the features of the proposed models and to show that the proposed GAS models are potentially better candidates for option pricing. From these simulations, it is easy to observe that, consistent with Duan (1995) and Tong et al. (2004), by applying the Esscher transform to parametric models under $\mathbb{P}$ measure, the density of normally distributed innovation terms are essentially shifted leftwards under $\mathbb{Q}$. In contrast, the Esscher transform, instead of shifting the Gamma innovations, is skewing these innovations so that from $\mathbb{P}$ to $\mathbb{Q}$, they become more negatively skewed. It can also be revealed from the simulation studies that the underlying price distribution simulated by the

SNG-GAS model tends to have fatter tails compared to that simulated using the GARCH model with normal innovations. Such a pattern can be observed for densities simulated 30 days to 360 days from now. This indicates that the SNG-GAS model tends to overestimate tail risk, which is related to the price of OTM options. Note that the BS model is reported to have inferior accuracy in pricing OTM options due to its tendency in underestimating tail risk and that the GARCH model is found to have similar performance as the BS model. They all tend to give OTM option prices that are less than market prices. Therefore, it is natural to expect that the SNG-GAS and SG-GAS models perform better in pricing OTM options, especially those dOTM ones.

Following a similar process as in Zhu and Ling (2015), I inspect the implied volatility smile of each model using a set of static parameters estimated using a full sample. Consistent with Zhu and Ling (2015) and Duan (1995), the risk-free rate is set to zero and the strike price K is assumed to be 1. So that by changing the values of spot prices, the moneyness has a range between 0.9 and 1.1. To be more comprehensive, times to maturity from 30 days to 180 days and starting values of 80% sample variance and 100% variance used to initiate the simulation process are considered. It can be concluded that while other models tend to show volatility smiles to some extent, the SNG-GAS model tends to show volatility smirks across maturities and initial variances. This is again evidence that the SNG-GAS model should outperform in pricing OTM calls, as volatility is positively correlated to option prices. It is also obvious that the SNG-GAS model implied volatility is consistently higher than the volatility implied by other models. Such a pattern becomes more and more significant as time to maturity extends. Last but not least, the U-shaped volatility curves are found to be flattened as time to maturity stretches. All these reported findings are consistent with Zhu and Ling (2015).

Using option data in the real world, the main empirical finding of this study is that the SG-GAS and SNG-GAS models significantly outperform the BS, GARCH

and GAS models in terms of option pricing accuracy. The findings in simulation studies are thus verified. More specifically, the pricing errors of these two proposed models are 51.51% and 48.94%, while the BS, GARCH and GAS models are found to have pricing errors of 61.60%, 55.67% and 53.51%. It is also documented that such superiority of SG-GAS and SNG-GAS disappears during the 2007-2008 financial crisis. However, during the crisis, it is still the GAS model that dominates all the other candidate models. This means that the selection of innovation terms indeed matters in financial returns modelling, especially when the market environment is unusual. Apart from the financial crisis period, the two proposed models are found to consistently dominate other models, and such dominance becomes more prominent in the post-crisis era.

I also test all models across a range of moneyness and times to maturity. It can be concluded that the outperformance of SG-GAS and SNG-GAS models is mainly from their improved accuracy in pricing deep out-of-the-money (dOTM) options. To see this, for near-the-money options with moneyness between 0.98 and 1.02, these two GAS models are found to have the worst performance, with pricing errors of 35.94% and 31.24%. While the BS, GARCH and GAS models have significantly lower pricing errors of 29.56%, 26.04% and 25.67%. For dOTM calls (moneyness below 0.96), SG-GAS and SNG-GAS have much lower pricing errors (70.81% and 68.34%) compared to those from competitors (110.15%, 107.23% and 80.43% for BS, GARCH and GAS). Such outperformance of SG-GAS and SNG-GAS stays unchanged for dOTM puts with moneyness above 1.06. The MAPE of these two models is 55.97% and 58.58%. In comparison, the pricing errors of the BS, GARCH and GAS models are 72.41%, 64.32% and 70.30% in this case. Given that the number of near-the-money options is 10924 and the number of OTM options is 31773, it is reasonable to think that the outperformance of SG-GAS and SNG-GAS models is due to their ability in pricing OTM options more accurately.

Finally, I document that SG-GAS and SNG-GAS models perform better in pricing

options with longer maturities. For instance, for options that expire within 30 days, the pricing errors of these two models are 62.60% and 54.28%, compared to 77.12%, 60.31% and 56.93% from BS, GARCH and GAS models. However, the pricing errors of the proposed models with respect to options that expire in more than 360 days are 37.01% and 48.88%, which is significantly lower than 51.09%, 51.67% and 50.18% from BS, GARCH and GAS models.

Chapter 5

Conclusion

This thesis consists of three essays on forecasting and option pricing. In the first two essays, I attempt to evaluate the performance of parametric and non-parametric approaches in predicting future realisations. More specifically, Chapter 2 is about the predictive power of option-implied correlation over realised correlation and market returns. The time series of implied correlation is obtained model-free, hence it is classified as non-parametric forecasting. In Chapter 3, I evaluate the performance of trending observation-driven models in capturing the dynamics of the Dow Jones Industrial Average (DJIA) returns and predicting realised volatility.

The first model I consider is GARCH, which enables the variance of the error term to be time-varying. As a comparison, I also construct the Dynamic Conditional Score (DCS), or simply score-driven models. The DCS model is also called the Generalised Autoregressive Score (GAS) model due to its similar structure to GARCH. In this thesis, the score-driven model is referred to as the GAS model. A series of GARCH and corresponding GAS models are constructed following different assumptions on innovation terms. These models are fitted before being utilised to make forecasts to volatility or other parameters that are connected to variance, such as the scale parameter, that defines the innovation terms on a rolling window basis. In 3, emphasis

is placed on evaluating the forecasting performance of model-based predictions to realised volatility. Such forecasting performance is also compared to that of implied volatility. Consequently, in chapter 4, option valuation is implemented by running simulations with these time-varying parameters to obtain expected payoffs under risk-neutral measure.

In each of these essays, I have a few important findings that are connected and contributive to existing literature. In Chapter 2, I inspect the informational content of implied correlation in the sense of decomposing it into components with different frequencies. In particular, I report that the predictive power of implied correlation stems from the interaction between its low-frequency and high-frequency components. Intuitively, the high-frequency component is expected to capture short-term changes of implied correlation, while the low-frequency component tends to pick up more lasting fluctuations. The empirical analysis in Chapter 2 indicates that the high-frequency component drives the predictive power of implied correlation at shorter horizons, while the low-frequency component, at longer horizons. More importantly, I find that the predictability of market returns can be substantially improved by decomposing implied correlation into components. This improvement is more significant at shorter horizons, where it is more challenging to predict market returns. The improvement in market return predictability is robust after controlling for a set of option-implied and fundamental variables, and it can be observed out-of-sample as well.

In Chapter 3, using a sample of the Dow Jones Industrial Average index historical prices ranging between 07/1997 and 04/2022, I construct a series of GARCH/GAS models by making assumptions on the distribution of innovation term ξ_t. Extensive comparisons are made between these two groups of models in terms of their ability in predicting realised volatility. As the first study that implements systematic evaluation of these two representative observation-driven models, few findings are documented in it. First of all, it is observed that GARCH models, with higher

R^2 and lower prediction error, tend to have better performance than corresponding GAS models in predicting realised volatility across all forecasting horizons.

Second, I find that, at daily and weekly horizons, the R^2 from regressing RV onto GAS model-based predictions are generally lower than that of GARCH models. However, at the monthly horizon, it is the opposite. In terms of prediction accuracy, GARCH models tend to have significantly lower forecasting errors measured by MAE compared to GAS. This evidence indicates the outperformance of GARCH over GAS in volatility forecasting.

Third, it is documented that TLS-GAS has the lowest prediction error of 3.8626 at the daily horizon, given that its R^2 from Mincer-Zarnowitz regressions is on the same level as the best performer at daily horizons and that the slope of TLS-GAS is closer to 1, it can be concluded that TLS-GAS has the overall best performance in prediction RV at the daily horizon. The additionally reported out-of-sample R^2 and CW test statistics confirm the superiority of TLS-GAS, and the DM statistics indicate that the TLS-GAS prediction is statistically different from others.

Fourth, from the pairwise out-of-sample R-squared R^2_{os}, it can be concluded that GARCH models outperform GAS in most cases at the weekly horizon. In contrast, GAS tends to dominate GARCH at the monthly horizon. In terms of specific models, I find that SNG-GARCH (together with SNIG-GARCH) and SIG-GAS consistently outperform other candidates at weekly and monthly horizons respectively.

Another topic in this chapter is evaluations of parametric models' performance in volatility forecasting during market turmoils. To this end, I adopt a sub-sample during the 2007-2008 financial crisis to fit models and implement evaluations. Several key findings are reported. First, instead of TLS-GAS, SNW-GAS strongly dominates other models at the daily horizon with both higher R^2 and lower forecasting error. The SNW-GAS prediction is statistically different to other forecasts, according to DM test results, and its superiority is verified again by CW statistics. In addition,

SNG-GAS and SG-GAS are found to have a performance only second to SNW-GAS. It appears that shifted Gamma or Weibull distributions are more appropriate candidates for financial returns modelling during financial crises. Second, I find that at the monthly horizon, the outperformance of SNW-GAS vanishes. Instead, the SIG-GAS model is found to dominate in this case.

The comparisons between implied volatility, represented by IV and MFIV, and model-based volatility forecasts, in terms of their performance in predicting RV, are also explored in this chapter. The findings are consistent with existing literature. The most straightforward finding is the domination of implied volatility in terms of describing the movements in future realised volatility. Such superiority is observed across horizons from 1 day up to 1 month and is found to be stronger as the forecasting horizon extends. On the contrary, IV and MFIV are found to have inferior forecasting accuracy, measured by MAE, compared to observation-driven models. and such inferior performance becomes weaker as the forecasting horizon stretches. It can be concluded that model-based predictions consistently have higher forecasting accuracy compared to implied volatility at horizons from one day up to one month. This is consistent with the pairwise CW statistics.

From the DM statistics, it can be inferred that, at the daily horizon, IV is not statistically different from other model-based predictions in most cases (10 out of 16). In contrast, MFIV is significantly different from almost all other predictions (15 out of 16). This pattern is observed at weekly horizons but becomes the opposite at monthly horizons, where IV is significantly different from model-based predictions in most cases.

Chapter 4 is connected to Duan (1995), Duan (1999) and Tong et al. (2004), where option pricing is implemented by running simulations under risk-neutral measure. Under the framework of GAS, SG-GAS and SNG-GAS are constructed due to the Gamma innovation's ability in handling the skewness of in financial returns (Tong

et al., 2004).

The pricing accuracy of the models is measured by the mean absolute percentage error (MAPE) through this chapter. Empirically, SG-GAS and SNG-GAS are found to have significantly lower pricing errors compared to the BS model, normal GARCH (GARCH) and normal GAS (GAS). Utilising sub-samples before, during and after the 2007-2008 financial crisis, I find that the outperformance of SG-GAS and SNG-GAS models is more significant in normal market environments (before and after the financial crisis). However, during the financial crisis, SG-GAS and SNG-GAS, while still outperforming the BS model, are no longer the best performers. Instead, GARCH and GAS tend to outperform other candidates during this massive market turmoil. More specifically, GAS is found to significantly dominate its GARCH counterpart, indicating that if the error term is appropriately selected, GAS is potentially a better candidate model for option pricing during big market uncertainties.

Finally, it can be inferred from simulation studies and empirical results that, the outperformance of SG-GAS and SNG-GAS is mainly from their superior ability in pricing Deep Out-of-the-Money (dOTM) options. This is because they tend to produce 'leptokurtic' underlying price distributions. It is also documented that SG-GAS and SNG-GAS have significant outperformance in pricing options with longer maturities.

From an overview through all chapters, it is obvious that the findings are beneficial to both industry and academia. For instance, the findings in chapter 2 enable investors with different investment horizons to better predict market returns using components of implied correlation. The findings in volatility forecasting and option pricing are more extensive. Chapter 3 can be used as a dictionary that tells investors what model can be a better choice under specific circumstances. For example, during high market uncertainties, if an investor is to predict daily volatility, then SNW-GAS could be a reasonable choice. If the investor intends to price an option when

the market is normal, then he may want to apply the SNG-GAS option pricing model. In terms of academia, the findings are inspiring starting points for more in-depth research. For example, the methodology of chapter 2 can be extended to more option implied information, such as implied skewness and kurtosis. Apart from that, chapter 3 and chapter 4 can be seen as fundamental research on GAS models in volatility forecasting and option pricing. They are extendable, as more models can be proposed by assuming different distributions of innovation terms. Besides, some promising findings, such as the dominance of TLS-GAS in volatility forecasting at the daily horizon, lead to a good starting point for building more advanced models.

The first two chapters of this thesis extensively provide fundamental inspections into parametric and non-parametric methodologies for forecasting realisations. It is also a preparation for more in-depth explorations in this field. For example, in Chapter 2, the decomposition of implied correlation has a number of potential applications that can be used to motivate future research. Among these, an obvious one is forecasting market variables at specific horizons by choosing appropriate components from decomposition so that the prediction accuracy is improved due to a reduction in noise. Another promising application lies in optimal portfolio selection in the presence of correlation risk, as such risk might be depending on specific investment horizons. In a more general sense, the implied correlation does not have to be a univariate time series. Instead, motivated by (Numpacharoen and Numpacharoen, 2013), it is possible to extract correlation matrices that dynamically change over time from option prices. To this end, option implied mean-variance portfolio construction is again one natural and appealing application. Finally, it is also an interesting direction for future research to explore how each component of implied correlation (or other similar risk factors) is priced in the cross-section of stock and option returns.

In terms of volatility forecasting, there are also promising topics for further exploration. For instance, it is crucial to have stable optimisation methods in the model

training process. It is also important to select wisely the initial points for optimisations. Some more advanced methods, such as the grid search might be able to equip the parametric models with improved forecasting power. In fact, optimisation itself is a popular topic that has been investigated all the time. It is possible that the model performance can be improved by applying different types of optimisation methods in the process of model fitting.

I see the ongoing topic of finding distributions that are more suitable for financial returns modelling and volatility forecasting as promising as well. As introduced, this study is fundamental in that the distributions adopted in constructions of GARCH/-GAS models are common ones. There exists great space for more discovery. For example, more types of distributions can be considered in model construction so that the model library can be expanded. Potential candidate distributions are the negative inverse Gaussian (NIG) distribution and the mixture normal distribution. The latter, as the entry-level mixture distributions, can also be extended to more complicated scenarios, such as variance-gamma mixture.

As reported, some of the observation-driven models tend to have superior performance in volatility forecasting at specific horizons, such as the TLS-GAS at the daily horizon and the SIG-GAS at the monthly horizon. For the purpose of further improvement in RV prediction, it is natural to consider more sophisticated GAS model specifications, such as the GJR-TLS-GAS that take into consideration the leverage effect. Alternatively, introducing exogenous variables might also help in enhancing model performance. A possible example could be the RV-TLS-GAS model by introducing realised volatility into the model system. I see this topic as practical, as plenty of studies propose variations of GARCH models, and it is always accessible to get hints from GARCH literature before adopting them in the framework of GAS.

Volatility forecasting using parametric models in this study is straightforward. However, this is not the case for option pricing. This is because risk neutralisation is

needed to find out the equations that describe financial return movements in a risk-neutral world. The Esscher transform is advantageous in its capability in handling non-normal distributions. However, it requires an analytical moment-generating function to accomplish the transformation. This makes it inappropriate to consider some distributions in constructing option pricing models, as they do not have closed-form MGFs. Inspired by existing literature, a possible solution could be the application of machine learning tools. All observation-driven models are naturally able to predict volatility or parameters that define density and are linked to variance closely. The main logic is quite straightforward: all these predictions can be used as inputs to feed neural networks that output option prices.

Bibliography

Abdalla, S. Z. S. (2012). Modelling stock returns volatility: Empirical evidence from saudi stock exchange. *International Research Journal of Finance and Economics*.

Andersen, T. G. (2000). Some reflections on analysis of high-frequency data. *Journal of Business & Economic Statistics*, 18(2):146–153.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71:579–625.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61:259–299.

Ardia, D., Boudt, K., and Catania, L. (2019). Generalized autoregressive score models in r: The gas package. *Journal of Statistical Software*, 88.

Badescu, A., Kulperger, R., and Lazar, E. (2008). Option valuation with normal mixture garch models. *Studies in Nonlinear Dynamics & Econometrics*, 12(2).

Baillie, R. T. and Bollerslev, T. (1989). Common stochastic trends in a system of exchange rates. *The Journal of Finance*, 44:167–181.

Banerjee, P. S., Doran, J. S., and Peterson, D. R. (2007). Implied volatility and future portfolio returns. *Journal of Banking Finance*, 31(10):3183–3199.

Bansal, R., Kiku, D., Shaliastovich, I., and Yaron, A. (2014). Volatility, the macroeconomy, and asset prices. *The Journal of Finance*, 69:2471–2511.

Barndorff-Nielsen, O. E. (2005). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4:1–30.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2004). Regular and modified kernel-based estimators of integrated variance: The case with independent noise. *Economics Papers 2004-W28, Economics Group, Nuffield College, University of Oxford*.

Bates, D. S. (1991). The crash of 87: Was it expected? the evidence from options markets. *The Journal of Finance*, 46:1009–1044.

Baule, R., Korn, O., and Saßning, S. (2015). Which beta is best? On the information content of option-implied betas. *European Financial Management*, 22(3):450—483.

Becker, R. and Clements, A. E. (2008). Are combination forecasts of sp 500 volatility statistically superior? *International Journal of Forecasting*, 24:122–133.

Becker, R., Clements, A. E., and White, S. I. (2007). Does implied volatility provide any information beyond that captured in model-based volatility forecasts? *Journal of Banking Finance*, 31:2535–2549.

Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. *Journal of Banking Finance*, 5:363–381.

Bekaert, G., Hoerova, M., and Lo Duca, M. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788.

Bera, A. K. and Higgins, M. L. (1993). Arch models: Properties, estimation and testing. *Journal of Economic Surveys*, 7:305–366.

Bernales, A. and Valenzuela, M. (2016). Implied correlation and market returns. *Available at SSRN 2739757*.

Black, F. (1975). Fact and fantasy in the use of options. *Financial Analysts Journal*.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654.

Blair, B. J., Poon, S.-H., and Taylor, S. J. (2010). Forecasting sp 100 volatility: The incremental information content of implied volatilities and high-frequency index returns. *Handbook of Quantitative Finance and Risk Management*, pages 1333–1344.

Blasques, F., Koopman, S. J., and Lucas, A. (2014a). Maximum likelihood estimation for correctly specified generalized autoregressive score models: Feedback effects, contraction conditions and asymptotic properties. Tinbergen Institute Discussion Papers.

Blasques, F., Koopman, S. J., and Lucas, A. (2014b). Maximum likelihood estimation for generalized autoregressive score models. Tinbergen Institute Discussion Papers.

Blasques, F., Koopman, S. J., Lucas, A., and Schaumburg, J. (2016). Spillover dynamics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics*, 195:211–223.

Bliss, R. R. and Panigirtzoglou, N. (2002). Testing the stability of implied probability density functions. *Journal of Banking Finance*, 26:381–422.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 69:542.

Bollerslev, T., Marrone, J., Xu, L., and Zhou, H. (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence. *Journal of Financial and Quantitative Analysis*, 49:633–661.

Bollerslev, T., Tauchen, G., and Zhou, H. (2009). Expected stock returns and variance risk premia. *Review of Financial Studies*, 22:4463–4492.

Bollerslev, T., Todorov, V., and Xu, L. (2015). Tail risk premia and return predictability. *Journal of Financial Economics*, 118:113–134.

Branger, N., Flacke, R., Middelhoff, F., and Thimme, J. (2021). Jumps and the correlation risk premium: Evidence from equity options. *SSRN Electronic Journal*.

Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *The Journal of Business*, 51:621–651.

Britten-Jones, M. and Neuberger, A. (2000). Option prices, implied price processes, and stochastic volatility. *The Journal of Finance*, 55:839–866.

Bühlmann, H., Delbaen, F., Embrechts, P., and Shiryaev, A. N. (1996). No-arbitrage, change of measure and conditional esscher transforms. *CWI quarterly*, 9:291–317.

Buraschi, A., Pochia, P., and Trojani, F. (2010). Correlation risk and optimal portfolio choice. *The Journal of Finance*, 65:393–420.

Buraschi, A., Trojani, F., and Vedolin, A. (2014). When uncertainty blows in the orchard: Comovement and equilibrium volatility risk premia. *Journal of Finance*, 69(1):101–137.

Buss, A., Schoenleber, L., and Vilkov, G. (2017). Option-implied correlations, factor models, and market risk. INSEAD Working Paper No. 2017/20/FIN.

Buss, A., Schoenleber, L., and Vilkov, G. (2019). Expected correlation and future market returns. CEPR Discussion Papers 12760, C.E.P.R. Discussion Papers.

Buss, A. and Vilkov, G. (2012). Measuring equity risk with option-implied correlations. *Review of Financial Studies*, 25(10):3113–3140.

Campa, J. M., Chang, P., and Reider, R. L. (1998). Implied exchange rate distributions: evidence from otc option markets. *Journal of International Money and Finance*, 17:117–160.

Campbell, J. and Thompson, S. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average. *Review of Financial Studies*, 21(4):1509—1531.

Canina, L. and Figlewski, S. (1993). The informational content of implied volatility. *The Review of Financial Studies*, 6:659–681.

Carr, P. and Madan, D. Towards a theory of volatility trading. In Jouini, E., Cvitanic, J., and Musiela, M., editors, *Option Pricing, Interest Rates and Risk Management*, pages 458–476. Cambridge University Press.

Catania, L. and Billé, A. G. (2017). Dynamic spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Applied Econometrics*, 32:1178–1196.

Chakravarty, S., Gulen, H., and Mayhew, S. (2004). Informed trading in stock and option markets. *Journal of Finance*, 59(3):1235—1258.

Chang, B.-Y., Christoffersen, P., Jacobs, K., and Vainberg, G. (2011). Option-implied measures of equity risk*. *Review of Finance*, 16:385–428.

Chen, Q. and Gerlach, R. H. (2013). The two-sided weibull distribution and forecasting financial tail risk. *International Journal of Forecasting*, 29(4):527–540.

Chen, X. and Ghysels, E. (2011). News—good or bad—and its impact on volatility predictions over multiple horizons. *The Review of Financial Studies*, 24:46–81.

Chiras, D. P. and Manaster, S. (1978). The information content of option prices and a test of market efficiency. *Journal of Financial Economics*, 6:213–234.

Christensen, B. J. and Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, 50:125–150.

Christoffersen, P., Heston, S., and Jacobs, K. (2006). Option valuation with conditional skewness. *Journal of Econometrics*, 131:253–284.

Clark, T. and West, K. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291—311.

Cosemans, M. (2011). The pricing of long and short run variance and correlation risk in stock returns. *University of Amsterdam Business School, Working Paper*.

Cox, D. R., Gudmundsson, G., Lindgren, G., Bondesson, L., Harsaae, E., Laake, P., Juselius, K., and Lauritzen, S. L. (1981). Statistical analysis of time series: Some recent developments [with discussion and reply]. *Scandinavian Journal of Statistics*, 8:93–115.

Creal, D., Koopman, S. J., and Lucas, A. (2012). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28:777–795.

Creal, D., Schwaab, B., Koopman, S. J., and Lucas, A. (2014). Observation-driven mixed-measurement dynamic factor models with an application to credit risk. *Review of Economics and Statistics*, 96:898–915.

Demeterfi, K., Derman, E., Kamal, M., and Zou, J. (1999a). A guide to volatility and variance swaps. *The Journal of Derivatives*, 6:9–32.

Demeterfi, K., Derman, E., Kamal, M., and Zou, J. (1999b). More than you ever wanted to know about volatility swaps. *Goldman Sachs Quantitative Strategies Research Notes*.

DeMiguel, V., Plyakha, Y., Uppal, R., and Vilkov, G. (2013). Improving portfolio selection using option-implied volatility and skewness. *Journal of Financial and Quantitative Analysis*, 48:1813–1845.

Derman, K. (1994). The volatility smile and its implied tree. *Goldman Sachs Quantitative Strategies Research Notes*.

Di Matteo, T. (2007). Multi-scaling in finance. *Quantitative Finance*, 7(1):21—36.

Diavatopoulos, D., Doran, J. S., and Peterson, D. R. (2008). The information content in implied idiosyncratic volatility and the cross-section of stock returns: Evidence from the option markets. *Journal of Futures Markets*, 28:1013–1039.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business Economic Statistics*, 13:253–263.

Drechsler, I. and Yaron, A. (2011). What's vol got to do with it. *The Review of Financial Studies*, 24:1–45.

Driessen, J., Maenhout, P., and Vilkov, G. (2009). The price of correlation risk: Evidence from equity options. *Journal of Finance*, 64(3):1377–1406.

Driessen, J., Maenhout, P., and Vilkov, G. (2012). Option-implied correlations and the price of correlation risk. *SSRN Electronic Journal*.

Duan, J.-C. (1995). The garch option pricing model. *Mathematical Finance*, 5:13–32.

Duan, J.-C. (1999). Conditionally fat-tailed distributions and the volatility smile in options. *Rotman School of Management, University of Toronto, Working Paper*.

Dupire (1994). Pricing with a smile. *Risk*, 7(1):18–20.

Durbin, J. and Koopman, S. J. (2012). Time series analysis by state space methods. *Oxford University Press*.

Ederington, L. and Guan, W. (2002). Is implied volatility an informationally efficient and effective predictor of future volatility? *USF St. Petersburg campus Faculty Publications*.

Elliott, R. J. and Madan, D. B. (1998). A discrete time equivalent martingale measure. *Mathematical Finance*, 8:127–152.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1007.

Escher, F. (1932). On the probability function in the collective theory of risk. *Scandinavian Actuarial Journal*, 1932(3):175–195.

Fan, Z., Xiao, X., and Zhou, H. (2022). Moment risk premia and stock return predictability. *Journal of Financial and Quantitative Analysis*, 57(1):67–93.

Faria, G., Kosowski, R., and Wang, T. (2022). The correlation risk premium: International evidence. *Journal of Banking Finance*, 136:106399.

Fereira, M. and Santa-Clara, P. (2011). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, 100(3):514–537.

Feunou, B., Jahan-Parvar, M. R., and Okou, C. (2017). Downside variance risk premium*. *Journal of Financial Econometrics*, 16:341–383.

Fleming, J. (1998). The quality of market volatility forecasts implied by sp 100 index option prices. *Journal of Empirical Finance*, 5:317–345.

Fleming, J., Ostdiek, B., and Whaley, R. E. (1995). Predicting stock market volatility: A new measure. *Journal of Futures Markets*, 15:265–302.

Forsberg, L. and Bollerslev, T. (2002). Bridging the gap between the distribution of realized (ecu) volatility and arch modelling (of the euro): the garch-nig model. *Journal of Applied Econometrics*, 17:535–548.

Gerber, H. U. and Shiu, E. S. (1994). Martingale approach to pricing perpetual american options. *ASTIN Bulletin*, 24:195–220.

Giot, P. and Laurent, S. (2007). The information content of implied volatility in light of the jump/continuous decomposition of realized volatility. *Journal of Futures Markets*, 27:337–359.

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

Goyal, A. and Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508.

Harrison, J. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215–260.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press.

Harvey, C. R. and Whaley, R. E. (1992). Market volatility prediction and the efficiency of the s p 100 index option market. *Journal of Financial Economics*, 31:43–73.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343.

Heston, S. L. and Nandi, S. (2000). A closed-form garch option valuation model. *The Review of Financial Studies*, 13:585–625.

Hollstein, F., Prokopczuk, M., Tharann, B., and Simen, C. (2018). Predicting the equity market with option-implied variables. *European Journal of Finance*, 25(10):937–965.

Hua, Q. and Jiang, T. (2015). The prediction for London gold price: Improved empirical mode decomposition. *Applied Economics Letters*, 22(17):1404—1408.

Hua, Q., Jiang, T., and Cheng, Z. (2018). Option pricing based on hybrid GARCH-type models with improved ensemble empirical mode decomposition. *Quantitative Finance*, 18(9):1501—1515.

Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., Yen, N.-C., Tung, C., and Liu, H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings: Mathematical, Physical and Engineering Sciences*, 454:903—995.

Härdle, W. K. and Silyakova, E. (2011). Volatility investing with variance swaps. *Handbook of Computational Finance*, pages 203–219.

Jiang, G. J. and Tian, Y. S. (2005). The model-free implied volatility and its information content. *The Review of Financial Studies*, 18:1305–1342.

Jiang, G. J. and Tian, Y. S. (2007). Extracting model-free volatility from option prices: An examination of the vix index. *The Journal of Derivatives*, 14:35–60.

Kambouroudis, D. S., McMillan, D. G., and Tsakou, K. (2016). Forecasting stock return volatility: A comparison of garch, implied volatility, and realized volatility models. *Journal of Futures Markets*, 36:1127–1163.

Kelly, B., Lustig, H., and Nieuwerburgh, S. (2016). Too-systemic-to-fail: What option markets imply about sector-wide government guarantees. *American Economic Review*, 106(6):1278–1319.

Kempf, A., Korn, O., and Saßning, S. (2015). Portfolio optimization using forward-looking information. *Review of Finance*, 19(1):467—490.

Koopman, S., Harvey, A., Doornik, J., and Shephard, N. (2000). *STAMP 6.0 Structural Time Series Analyser, Modeller and Predictor*. Timberlake Consultants Ltd.

Kourtis, A., Markellos, R. N., and Symeonidis, L. (2016). An international comparison of implied, realized, and garch volatility forecasts. *Journal of Futures Markets*, 36:1164–1193.

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

Lanne, M. (2003). Modeling the u.s. short-term interest rate by mixture autoregressive processes. *Journal of Financial Econometrics*, 1:96–125.

Lanne, M. and Pentti, S. (2007). Modeling conditional skewness in stock returns. *The European Journal of Finance*, 13:691–704.

Liang, C., Wei, Y., and Zhang, Y. (2020). Is implied volatility more informative for forecasting realized volatility: An international perspective. *Journal of Forecasting*.

Lin, Q. and Yu, H. (2017). Option pricing with generalized autoregressive score models. *Working paper, University of Hong Kong*.

Lopez, J. A. and Walter, C. (2000). Is implied correlation worth calculating? Evidence from foreign exchange options and historical data. Working Paper Series 2000-02, Federal Reserve Bank of San Francisco.

Malz, A. M. (1997a). Estimating the probability distribution of the future exchange rate from option prices. *The Journal of Derivatives*, 5:18–36.

Malz, A. M. (1997b). Option-implied probability distributions and currency excess returns. Staff Reports 32, Federal Reserve Bank of New York.

Martin, I. (2011). Simple variance swaps. NBER Working Paper No. w16884.

Martin, I. (2016). What is the expected return on the market? *The Quarterly Journal of Economics*.

Martin, V. L., Tang, C., and Yao, W. (2021). Forecasting the volatility of asset returns: The informational gains from option prices. *International Journal of Forecasting*, 37:862–880.

Menn, C. and Rachev, S. T. (2005). A garch option pricing model with -stable innovations. *European Journal of Operational Research*, 163:201–209.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41:867.

Mete Kilic, I. S. (2018). Good and bad variance premia and expected returns. *Management Science*, 65(6):2522–2544.

Mincer, J. A. and Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, pages 3–46. NBER.

Mittnik, S. and Rachev, S. T. (1993). Modeling asset returns with alternative stable distributions*. *Econometric Reviews*, 12:261–330.

Mueller, P., Stathopoulos, A., and Vedolin, A. (2017). International correlation risk. *Journal of Financial Economics*, 126:270–299.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59:347.

Neuberger, A. (1994). The log contract. *The Journal of Portfolio Management*, 20:74–80.

Newey, W. and West, K. (1987). A simple, positive, semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703—708.

Numpacharoen, K. and Numpacharoen, N. (2013). Estimating realistic implied correlation matrix from option prices. *Journal of Mathematical Finance*, 03:401–406.

Piatti, I. (2015). Heterogeneous beliefs about rare event risk in the lucas orchard. *Paris December 2014 Finance Meeting EUROFIDAI - AFFI Paper, Saïd Business School WP 2014-5*.

Pollet, J. and Wilson, M. (2010). Average correlation and stock market returns. *Journal of Financial Economics*, 96(3):364—380.

Poon, S.-H. and Granger, C. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41:478–539.

Rachev, S. T. and SenGupta, A. (1993). Laplace-weibull mixtures for modeling price changes. *Management Science*, 39:1029–1038.

Rubinstein, M. (1985). Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active cboe option classes from august 23, 1976 through august 31, 1978. *The Journal of Finance*, 40:455–480.

Saavedra, R., Bodin, G., and Souto, M. (2020). Statespacemodels.jl: a julia package for time-series analysis in a state-space framework. *arXiv:1908.01757 [math, stat]*.

Seo, S. W. and Kim, J. S. (2015). The information content of option-implied information for volatility forecasting with investor sentiment. *Journal of Banking Finance*, 50:106–120.

Shephard, N. (2005). *Stochastic Volatility: Selected Readings*. Oxford University Press.

Shimko, D. (1993). Bounds of probability. *Risk*, 6(4):33–37.

Silahli, B., Dingec, K. D., Cifter, A., and Aydin, N. (2019). Portfolio value-at-risk with two-sided weibull distribution: Evidence from cryptocurrency markets. *Finance Research Letters*, page 101425.

Skintzi, V. D. and Refenes, A.-P. N. (2004). Implied correlation index: A new measure of diversification. *Journal of Futures Markets*, 25:171–197.

Stentoft, L. (2006). Modelling the volatility of financial assets using the normal inverse gaussian distribution: With an application to option pricing. Technical report, working paper, HEC Montreal.

Stentoft, L. (2008). American option pricing using garch models and the normal inverse gaussian distribution. *Journal of Financial Econometrics*, 6:540–582.

Swidler, S. and Wilcox, J. A. (2002). Information about bank risk in options prices. *Journal of Banking Finance*, 26:1033–1057.

Taylor, S. J., Yadav, P. K., and Zhang, Y. (2010). The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks. *Journal of Banking Finance*, 34:871–881.

Tong, H., Siu, T. K., and Yang, H. (2004). On pricing derivatives under garch models: a dynamic gerber-shiu approach. *North American Actuarial Journal*, 8:17–31.

Wang, C., Chen, Q., and Gerlach, R. (2018). Bayesian realized-garch models for financial tail risk forecasting incorporating the two-sided weibull distribution. *Quantitative Finance*, 19:1017–1042.

Whaley, R. E. (1993). Derivatives on market volatility. *The Journal of Derivatives*, 1:71–84.

Yu, L., Wang, S., and Lai, K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics*, 30(5):2623—2635.

Zhang, X., Lai, K., and Wang, S.-Y. (2008). A new approach for crude oil price analysis based on empirical mode decomposition. *Energy Economics*, 30:905–918.

Zhang, X., Yu, L., Wang, S., and Lai, K. (2009). Estimating the impact of extreme events on crude oil price: An EMD-based event analysis method. *Energy Economics*, 31(5):768—778.

Zhou, H. (2018). Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty. *Annual Review of Financial Economics*, 10:481–497.

Zhu, K. and Ling, S. (2015). Model-based pricing for financial derivatives. *Journal of Econometrics*, 187:447–457.

Appendices

Appendix A

The Sifting Process

1. Initialisation: set $f_0(t) = IC(t), i = 1$
2. Extract the i^{th} IMF
   - 2.1. Initialisation: set $d_0(t) = f_{i-1}(t), k = 1$
   - 2.2. Identify all the extrema of $d_{k-1}(t)$
   - 2.3. Fit a cubic spline to generate the upper and lower envelopes ($e_{min}(t)$ and $e_{max}(t)$, respectively) of $d_{k-1}(t)$
   - 2.4. Calculate the point-by-point mean $m(t) = \frac{e_{min}(t)+e_{max}(t)}{2}$
   - 2.5. Define the difference $d_k(t) = d_{k-1}(t) - m(t)$
   - 2.6. If $d_k(t)$ satisfies the two conditions of an IMF, denote it as the i^{th} IMF, i.e. $c_i(t) = d_k(t)$, and go back to step 2.2.
3. Define $res(t) = f_i(t) - c_i(t)$
4. Check if the following stopping criteria are satisfied
   (a) $res(t)$ contains fewer than 3 extrema
   (b) the amplitudes of $res(t)$ are smaller than the amplitudes of $f_i(t)$ at each point

If either stopping criterion (or both) is not satisfied, then go back to step 2 to extract the next IMF. If both stopping criteria are satisfied, the sifting process is completed, and the final residual term $res(t)$ can be considered as the average trend of $IC(t)$.

Appendix B

Diebold-Mariano and Clark-West Tests

Table B.1: Diebold–Mariano and Clark-West test at the daily horizon - full sample

Panel A: Diebold–Mariano test at daily horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	4.7819	-4.0323	-0.7700	-2.7654	2.3370	5.1739	3.0444	-0.7606	0.6331	0.7704	-0.9744	-0.9244	1.3721	-5.1529	-0.0386
SG	-4.7819	NaN	-6.8543	-5.1908	-6.9861	1.0424	4.3631	1.5671	-1.0885	0.3207	0.4641	-1.2913	-1.3093	1.1345	-5.2654	-0.4061
SNG	4.0323	6.8543	NaN	2.8016	0.3943	3.3306	6.5588	4.3065	-0.4339	0.8889	1.0177	-0.6480	-0.5186	1.5572	-4.8086	0.2893
SIG	0.7700	5.1908	-2.8016	NaN	-2.2150	2.4940	5.4224	3.3046	-0.7033	0.6836	0.8200	-0.9182	-0.8459	1.4125	-5.0990	0.0221
SNIG	2.7654	6.9861	-0.3943	2.2150	NaN	3.2761	7.0506	4.4840	-0.4521	0.8414	0.9702	-0.6588	-0.5356	1.5195	-4.6882	0.2485
SNW	-2.3370	-1.0424	-3.3306	-2.4940	-3.2761	NaN	1.8034	0.1924	-1.3914	0.0707	0.2217	-1.5918	-1.5356	0.9674	-5.5946	-0.7188
TLS	-5.1739	-4.3631	-6.5588	-5.4224	-7.0506	-1.8034	NaN	-3.6378	-1.9667	-0.6164	-0.4610	-2.1328	-2.3950	0.3714	-5.6878	-1.5099
LOG	-3.0444	-1.5671	-4.3065	-3.3046	-4.4840	-0.1924	3.6378	NaN	-1.3886	0.0009	0.1490	-1.5751	-1.7087	0.8812	-5.4810	-0.8027
Normal	0.7606	1.0885	0.4339	0.7033	0.4521	1.3914	1.9667	1.3886	NaN	1.5938	1.7308	-2.8971	0.0160	2.3658	-5.3166	0.6825
SG	-0.6331	-0.3207	-0.8889	-0.6836	-0.8414	-0.0707	0.6164	-0.0009	-1.5938	NaN	0.7947	-1.8167	-1.7029	0.9905	-7.2643	-0.6568
SNG	-0.7704	-0.4641	-1.0177	-0.8200	-0.9702	-0.2217	0.4610	-0.1490	-1.7308	-0.7947	NaN	-1.9432	-1.7903	0.8647	-6.8238	-0.7965
SIG	0.9744	1.2913	0.6480	0.9182	0.6588	1.5918	2.1328	1.5751	2.8971	1.8167	1.9432	NaN	0.3526	2.5415	-4.9836	0.8844
SNIG	0.9244	1.3093	0.5186	0.8459	0.5356	1.5356	2.3950	1.7087	-0.0160	1.7029	1.7903	-0.3526	NaN	2.0614	-6.5283	0.7655
SNW	-1.3721	-1.1345	-1.5572	-1.4125	-1.5195	-0.9674	-0.3714	-0.8812	-2.3658	-0.9905	-0.8647	-2.5415	-2.0614	NaN	-4.7206	-1.3754
TLS	5.1529	5.2654	4.8086	5.0990	4.6882	5.5946	5.6878	5.4810	5.3166	7.2643	6.8238	4.9836	6.5283	4.7206	NaN	5.1419
LOG	0.0386	0.4061	-0.2893	-0.0221	-0.2485	0.7188	1.5099	0.8027	-0.6825	0.6568	0.7965	-0.8844	-0.7655	1.3754	-5.1419	NaN

Panel B: Clark-West test at daily horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	-2.0717	4.2156	1.8025	2.5078	-0.1349	-1.5938	-0.7421	1.5580	3.3701	3.5158	1.5906	1.4416	2.3907	4.0409	1.2848
SG	2.9530	NaN	5.8378	2.9096	6.8485	0.3106	-1.2511	0.0148	1.5914	3.3102	3.4478	1.6210	1.5074	2.3630	3.9929	1.4004
SNG	-2.7946	-5.2102	NaN	-1.8749	0.3257	-0.6017	-2.1922	-1.4998	1.4555	3.2636	3.4056	1.4881	1.3071	2.2924	3.7924	1.0778
SIG	0.3040	-2.3288	3.1864	NaN	1.9965	-0.2626	-1.6699	-0.8662	1.5515	3.3474	3.4972	1.5837	1.4327	2.3792	4.0414	1.2687
SNIG	-1.0496	-5.4460	0.8785	-0.7365	NaN	-0.4711	-2.4015	-1.5779	1.4634	3.2069	3.3411	1.4966	1.2889	2.2532	3.6645	1.0557
SNW	2.3578	1.4701	2.8937	2.4810	2.7409	NaN	0.2519	1.2119	2.0350	3.7710	3.8937	2.0755	1.8157	2.6619	5.0048	1.9007
TLS	2.3228	2.1948	2.7605	2.3428	2.9569	1.3690	NaN	1.7998	1.8070	3.1459	3.2709	1.8310	1.7769	2.3333	3.5555	1.5262
LOG	1.6466	1.2135	2.0001	1.7173	2.1175	1.0021	-0.9693	NaN	1.7901	3.2981	3.4278	1.8165	1.7050	2.4387	3.7195	1.4327
Normal	2.0502	1.9148	2.2652	2.0800	2.2776	1.9167	1.7315	1.9023	NaN	4.8471	5.0596	4.2472	1.3062	2.1768	4.8051	2.2759
SG	5.4745	5.3873	5.5138	5.4525	5.4834	5.2285	5.2981	5.4430	6.7563	NaN	2.1735	6.8023	6.2854	5.7395	8.2924	6.1486
SNG	5.5868	5.5043	5.6181	5.5699	5.5855	5.2688	5.4280	5.5640	6.8655	3.0217	NaN	6.9303	6.5002	5.7895	8.2487	6.0509
SIG	1.8357	1.7039	2.0500	1.8640	2.0729	1.7111	1.5655	1.7243	-0.0351	4.6493	4.8836	NaN	1.1744	1.9546	4.5269	2.0830
SNIG	2.8977	2.5838	3.2865	2.9376	3.2400	2.6318	1.6054	2.0948	1.4258	4.4609	4.7644	1.6620	NaN	2.4735	7.3712	2.9903
SNW	3.2456	3.1999	3.3073	3.2408	3.3008	2.9489	3.2060	3.2761	2.9624	2.9425	2.9851	2.9747	3.2928	NaN	4.4216	3.1741
TLS	1.9503	1.9252	2.0131	1.9812	2.0039	2.7211	1.3265	1.3760	3.5326	5.6104	5.8585	3.6578	2.1035	5.2403	NaN	2.0736
LOG	1.3313	1.1759	1.4763	1.3599	1.4696	1.2128	0.5912	0.8922	2.9258	4.9291	5.0035	2.9925	2.3943	3.6469	5.4965	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions. The statistics are obtained using predictions at the daily horizon, over the full sample. Panel A reports the Diebold–Mariano (DM) statistics and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.

Table B.2: Diebold–Mariano and Clark-West test at the weekly horizon - full sample

Panel A: Diebold–Mariano test at weekly horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	2.2605	-3.1933	-0.3719	-2.1146	-1.3557	5.2187	4.7006	0.9018	3.2966	4.1030	0.5466	1.3294	4.3471	6.6793	2.9821
SG	-2.2605	NaN	-4.9800	-2.3437	-4.6050	-1.8177	5.5279	4.4052	0.4474	2.8590	3.6480	0.1227	0.7816	3.8768	5.5487	2.2692
SNG	3.1933	4.9800	NaN	2.5943	0.3674	-0.5530	5.7673	5.7285	1.3501	3.6212	4.4157	0.9912	1.8178	4.6758	7.4699	3.5778
SIG	0.3719	2.3437	-2.5943	NaN	-2.0104	-1.3197	5.1067	4.4621	0.9933	3.4195	4.2316	0.6271	1.4243	4.4663	6.8936	3.1690
SNIG	2.1146	4.6050	-0.3674	2.0104	NaN	-0.6272	5.8494	5.7363	1.2633	3.5649	4.3611	0.9159	1.6919	4.6025	7.0988	3.3968
SNW	1.3557	1.8177	0.5530	1.3197	0.6272	NaN	3.9150	3.1526	2.3272	4.5098	5.3795	1.8191	2.9678	5.8098	8.5869	5.3106
TLS	-5.2187	-5.5279	-5.7673	-5.1067	-5.8494	-3.9150	NaN	-4.2511	-1.9205	0.2351	0.8986	-2.1281	-1.8638	1.0561	0.5974	-1.0074
LOG	-4.7006	-4.4052	-5.7285	-4.4621	-5.7363	-3.1526	4.2511	NaN	-0.7650	1.6521	2.3945	-1.0370	-0.5821	2.6134	3.2906	0.6639
Normal	-0.9018	-0.4474	-1.3501	-0.9933	-1.2633	-2.3272	1.9205	0.7650	NaN	4.0256	5.1952	-3.0336	0.6868	6.0182	4.3396	2.3169
SG	-3.2966	-2.8590	-3.6212	-3.4195	-3.5649	-4.5098	-0.2351	-1.6521	-4.0256	NaN	5.0137	-4.4874	-3.8463	1.3857	0.2298	-1.6274
SNG	-4.1030	-3.6480	-4.4157	-4.2316	-4.3611	-5.3795	-0.8986	-2.3945	-5.1952	-5.0137	NaN	-5.6467	-4.9536	0.2497	-0.7595	-2.6115
SIG	-0.5466	-0.1227	-0.9912	-0.6271	-0.9159	-1.8191	2.1281	1.0370	3.0336	4.4874	5.6467	NaN	1.5277	6.4577	4.6976	2.7744
SNIG	-1.3294	-0.7816	-1.8178	-1.4243	-1.6919	-2.9678	1.8638	0.5821	-0.6868	3.8463	4.9536	-1.5277	NaN	4.9324	4.5922	2.1314
SNW	-4.3471	-3.8768	-4.6758	-4.4663	-4.6025	-5.8098	-1.0561	-2.6134	-6.0182	-1.3857	-0.2497	-6.4577	-4.9324	NaN	-0.9835	-2.9186
TLS	-6.6793	-5.5487	-7.4699	-6.8936	-7.0988	-8.5869	-0.5974	-3.2906	-4.3396	-0.2298	0.7595	-4.6976	-4.5922	0.9835	NaN	-3.8599
LOG	-2.9821	-2.2692	-3.5778	-3.1690	-3.3968	-5.3106	1.0074	-0.6639	-2.3169	1.6274	2.6115	-2.7744	-2.1314	2.9186	3.8599	NaN

Panel B: Clark-West test at weekly horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	-0.5958	4.4774	1.8851	3.5899	1.8482	-1.3408	-1.1902	1.5221	2.8523	2.9310	1.5444	1.4349	2.0472	1.3611	1.2568
SG	1.8407	NaN	2.8186	1.6003	3.3893	1.7066	-1.4623	-1.0855	1.4630	2.6810	2.7495	1.4833	1.3860	1.9240	1.2335	1.2039
SNG	-2.0574	-2.1338	NaN	-1.3157	0.9527	1.3929	-1.6031	-1.8010	1.4245	2.7599	2.8360	1.4472	1.3053	1.9534	1.0461	1.0486
SIG	1.4255	-0.7592	3.2946	NaN	2.7519	1.8205	-1.3351	-1.2168	1.5325	2.8343	2.9194	1.5545	1.4369	2.0415	1.3651	1.2336
SNIG	-0.2305	-1.8441	1.4337	-0.1700	NaN	1.3405	-1.5584	-1.7156	1.4125	2.6862	2.7528	1.4361	1.2796	1.8997	1.0657	1.0415
SNW	1.2668	0.7826	1.4735	1.3050	1.3417	NaN	0.1196	0.4633	1.8149	3.5945	3.5507	1.8608	1.5778	2.1383	1.1313	1.1344
TLS	2.3348	2.4284	2.4840	2.2842	2.5923	2.1956	NaN	2.0136	1.6610	2.4914	2.5407	1.6739	1.6289	1.9018	1.4293	1.4159
LOG	2.3101	2.2553	2.5486	2.2460	2.7568	2.1729	-1.0966	NaN	1.7027	2.7285	2.7908	1.7176	1.6404	2.0688	1.5444	1.3947
Normal	4.9385	4.4683	5.0922	4.9568	4.9516	5.5439	3.1426	3.9107	NaN	3.7768	3.5848	2.9836	1.2528	0.2810	2.1622	2.3134
SG	5.5655	5.2897	5.5594	5.4897	5.4389	5.7203	4.7939	5.2584	6.3878	NaN	-0.0417	6.5055	6.5370	4.5969	5.1129	5.5275
SNG	5.9997	5.7428	5.9810	5.9355	5.8721	5.9221	5.2328	5.7031	6.5283	4.2643	NaN	6.6345	6.4825	5.0091	5.2944	5.5440
SIG	4.5310	4.0886	4.6864	4.5386	4.5594	5.0442	2.8879	3.5995	-0.6002	3.5091	3.3226	NaN	0.6083	-0.2337	1.8542	1.9336
SNIG	3.9034	3.4394	4.0314	3.9239	3.9177	4.6068	2.2335	2.9875	1.8877	4.4055	4.3192	2.0901	NaN	1.4542	1.6249	1.9795
SNW	5.8427	5.6799	5.8899	5.8065	5.8355	5.6046	5.1916	5.6339	4.4856	3.2260	2.9319	4.4717	4.4840	NaN	4.1680	4.3916
TLS	6.3797	5.9044	6.7073	6.5744	6.7379	6.6788	3.0409	4.8365	4.0902	4.6171	4.5065	4.0883	4.1915	3.5327	NaN	4.3203
LOG	5.3980	4.9780	5.7151	5.6015	5.7765	6.1753	2.6642	4.0260	3.5933	4.4459	4.2079	3.5872	3.2637	2.9981	1.5824	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions. The statistics are obtained using predictions at the weekly horizon, over the full sample. Panel A reports the Diebold–Mariano (DM) statistics and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.

Table B.3: Diebold–Mariano and Clark-West test at the monthly horizon - full sample

Panel A: Diebold–Mariano test at monthly horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	-0.4260	-4.7286	-0.9519	-5.0266	-8.2044	0.3839	-1.6833	-14.631	-9.5404	-6.4727	-15.480	-4.0010	-7.5123	-6.7796	-8.4673
SG	0.4260	NaN	-3.6425	-0.4526	-4.5828	-7.0628	0.5020	-1.8851	-13.803	-9.0814	-6.1791	-14.578	-3.9291	-7.1597	-6.7114	-8.2974
SNG	4.7286	3.6425	NaN	4.0782	-1.8088	-7.0348	1.0137	0.0654	-13.654	-8.5828	-5.6246	-14.489	-3.5937	-6.6289	-6.1867	-7.9446
SIG	0.9519	0.4526	-4.0782	NaN	-4.8119	-7.8270	0.5280	-1.3290	-14.204	-9.2248	-6.2208	-15.033	-3.9032	-7.2331	-6.6402	-8.3552
SNIG	5.0266	4.5828	1.8088	4.8119	NaN	-6.1958	1.2899	0.8103	-13.201	-8.2734	-5.3211	-14.013	-3.4200	-6.2624	-5.7801	-7.4727
SNW	8.2044	7.0628	7.0348	7.8270	6.1958	NaN	2.7893	4.0549	-10.256	-5.6281	-2.8268	-11.105	-2.1000	-3.6781	-3.0271	-4.6479
TLS	-0.3839	-0.5020	-1.0137	-0.5280	-1.2899	-2.7893	NaN	-1.3718	-7.1528	-5.3300	-3.9977	-7.3836	-3.4989	-4.4346	-4.6448	-5.3579
LOG	1.6833	1.8851	-0.0654	1.3290	-0.8103	-4.0549	1.3718	NaN	-11.106	-7.1744	-4.8693	-11.617	-3.4955	-5.6554	-5.9905	-7.4864
Normal	14.631	13.8028	13.654	14.204	13.201	10.256	7.1528	11.106	NaN	5.2295	7.7493	-1.7376	1.4978	8.3707	5.4378	3.9995
SG	9.5404	9.0814	8.5828	9.2248	8.2734	5.6281	5.3300	7.1744	-5.2295	NaN	8.7705	-6.8026	0.2535	4.5224	1.7765	0.5183
SNG	6.4727	6.1791	5.6246	6.2208	5.3211	2.8268	3.9977	4.8693	-7.7493	-8.7705	NaN	-9.1255	-0.7678	-1.5914	-0.2455	-1.4578
SIG	15.480	14.578	14.4889	15.033	14.013	11.105	7.3836	11.617	1.7376	6.8026	9.1255	NaN	1.6735	10.314	5.9140	4.4942
SNIG	4.0010	3.9291	3.5937	3.9032	3.4200	2.1000	3.4989	3.4955	-1.4978	-0.2535	0.7678	-1.6735	NaN	0.5047	0.6250	0.0001
SNW	7.5123	7.1597	6.6289	7.2331	6.2624	3.6781	4.4346	5.6554	-8.3707	-4.5224	1.5914	-10.314	-0.5047	NaN	0.2779	-1.0181
TLS	6.7796	6.7114	6.1867	6.6402	5.7801	3.0271	4.6448	5.9905	-5.4378	-1.7765	0.2455	-5.9140	-0.6250	-0.2779	NaN	-2.7023
LOG	8.4673	8.2974	7.9446	8.3552	7.4727	4.6479	5.3579	7.4864	-3.9995	-0.5183	1.4578	-4.4942	-0.0001	1.0181	2.7023	NaN

Panel B: Clark-West test at monthly horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	2.2280	5.3422	3.7971	5.3942	5.4774	2.8434	3.4033	11.6632	10.6675	10.7662	11.6252	11.6200	10.8488	11.3852	10.7217
SG	1.3113	NaN	2.5854	1.6310	3.6959	4.5516	2.0217	3.0480	10.4419	9.6398	9.8594	10.3972	10.4214	9.9041	10.8496	9.9717
SNG	-1.0921	-0.6800	NaN	-0.3499	3.4883	4.9791	2.3466	2.2720	11.5412	10.4455	10.5258	11.4913	11.4986	10.6156	11.2432	10.5394
SIG	2.8144	1.5848	5.3343	NaN	5.2501	5.3315	2.7882	3.3323	11.4864	10.3960	10.5120	11.4424	11.4304	10.6377	11.2765	10.5716
SNIG	-0.6013	-0.6327	0.9101	-0.3284	NaN	4.7650	2.2716	2.1591	11.4515	10.3339	10.4259	11.4012	11.4178	10.5512	11.3984	10.6035
SNW	1.1326	1.3374	1.2322	1.1346	1.2607	NaN	4.5612	3.2103	9.9721	9.4388	9.2827	9.9702	9.8400	9.1489	9.0328	8.5564
TLS	1.2750	1.2457	1.3927	1.3142	1.5789	2.3234	NaN	1.3264	3.9226	4.0497	4.3074	3.9118	3.8944	4.0965	4.4322	3.9147
LOG	1.0332	1.0432	1.4067	1.1396	1.9636	3.4827	0.9047	NaN	8.0802	7.6123	7.8790	8.0191	8.0972	7.7894	9.5064	8.1976
Normal	5.6300	5.4783	5.8036	5.6476	5.8267	4.8877	6.0571	6.0885	NaN	1.6873	0.7508	1.5694	1.6189	-0.7728	1.5023	1.1333
SG	7.3284	7.1845	7.3731	7.2279	7.2815	6.5013	7.1667	7.4315	5.6517	NaN	-1.3800	5.9046	5.2722	4.1858	4.2812	4.3921
SNG	8.1714	8.0351	8.1667	8.0910	8.0972	7.1509	7.6587	8.0421	5.7104	5.2935	NaN	5.8943	5.3226	5.9576	4.2477	4.4820
SIG	5.4229	5.2653	5.5967	5.4333	5.6231	4.6558	5.8349	5.8820	0.8159	1.1841	0.1940	NaN	0.2025	-1.7209	1.1167	0.7824
SNIG	3.0108	2.9448	3.0580	3.0286	3.0661	2.6960	2.9776	3.0850	1.3085	1.6155	1.4389	1.3156	NaN	1.0469	1.3480	1.2767
SNW	8.1231	7.9911	8.1600	8.0773	8.1471	7.0275	7.8203	8.2116	5.8021	5.2273	3.0336	5.9912	5.1893	NaN	3.6989	4.0669
TLS	6.9487	6.8362	7.2465	7.1166	7.2933	6.5071	5.9606	7.2114	4.0162	3.8578	3.3259	3.9516	3.8908	2.7207	NaN	4.7888
LOG	5.8936	5.8585	6.1713	6.0412	6.2399	5.2148	6.3227	6.5484	2.9438	2.9713	2.5784	2.8890	2.8244	1.9819	2.1468	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions. The statistics are obtained using predictions at the monthly horizon, over the full sample. Panel A reports the Diebold–Mariano (DM) statistics, and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.

Table B.4: Diebold–Mariano and Clark-West test at the daily horizon - crisis

Panel A: Diebold–Mariano test at daily horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	-1.6780	-2.2081	-1.6523	-3.1896	-2.8578	1.8244	2.8537	-0.8889	1.6193	1.9581	-0.9979	0.1619	-3.0851	0.3014	2.2847
SG	1.6780	NaN	-0.8141	-0.4564	-2.3347	-2.1992	2.3855	3.1669	-0.6743	1.7243	2.0547	-0.7892	0.3801	-2.9700	0.4473	2.3683
SNG	2.2081	0.8141	NaN	0.5264	-0.6407	-1.6747	2.7162	3.2483	-0.4401	1.7947	2.1212	-0.5525	0.5294	-2.8752	0.5736	2.4348
SIG	1.6523	0.4564	-0.5264	NaN	-1.9373	-2.1977	3.0313	3.4175	-0.5796	1.7496	2.0834	-0.6923	0.4097	-3.0061	0.4789	2.3913
SNIG	3.1896	2.3347	0.6407	1.9373	NaN	-1.3607	3.8341	3.8716	-0.2858	1.8490	2.1690	-0.4001	0.6667	-2.8041	0.6597	2.4986
SNW	2.8578	2.1992	1.6747	2.1977	1.3607	NaN	4.2508	3.7053	0.0918	1.9398	2.2508	-0.0131	0.9251	-2.5395	0.8775	2.6309
TLS	-1.8244	-2.3855	-2.7162	-3.0313	-3.8341	-4.2508	NaN	2.4042	-1.2489	1.1863	1.5521	-1.3270	-0.4352	-3.4399	-0.1809	2.0200
LOG	-2.8537	-3.1669	-3.2483	-3.4175	-3.8716	-3.7053	-2.4042	NaN	-2.0575	0.6026	0.9862	-2.1096	-1.2948	-3.8054	-0.8778	1.4851
Normal	0.8889	0.6743	0.4401	0.5796	0.2858	-0.0918	1.2489	2.0575	NaN	2.6628	2.8956	-1.8620	2.0482	-2.6729	1.1061	2.3648
SG	-1.6193	-1.7243	-1.7947	-1.7496	-1.8490	-1.9398	-1.1863	-0.6026	-2.6628	NaN	2.6832	-2.7755	-1.7029	-5.3395	-1.7430	0.6920
SNG	-1.9581	-2.0547	-2.1212	-2.0834	-2.1690	-2.2508	-1.5521	-0.9862	-2.8956	-2.6832	NaN	-2.9984	-2.0093	-5.7273	-2.1262	0.3615
SIG	0.9979	0.7892	0.5525	0.6923	0.4001	0.0131	1.3270	2.1096	1.8620	2.7755	2.9984	NaN	2.1756	-2.6397	1.2140	2.3967
SNIG	-0.1619	-0.3801	-0.5294	-0.4097	-0.6667	-0.9251	0.4352	1.2948	-2.0482	1.7029	2.0093	-2.1756	NaN	-2.6663	0.2211	2.0251
SNW	3.0851	2.9700	2.8752	3.0061	2.8041	2.5395	3.4399	3.8054	2.6729	5.3395	5.7273	2.6397	2.6663	NaN	3.1417	3.4285
TLS	-0.3014	-0.4473	-0.5736	-0.4789	-0.6597	-0.8775	0.1809	0.8778	-1.1061	1.7430	2.1262	-1.2140	-0.2211	-3.1417	NaN	1.7950
LOG	-2.2847	-2.3683	-2.4348	-2.3913	-2.4986	-2.6309	-2.0200	-1.4851	-2.3648	-0.6920	-0.3615	-2.3967	-2.0251	-3.4285	-1.7950	NaN

Panel B: Clark-West test at daily horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	1.9567	2.4909	2.0494	2.8794	2.3754	-0.5973	-1.1086	1.5772	3.5828	3.5950	1.6375	0.5293	3.4727	1.9686	-2.2469
SG	-0.6515	NaN	1.2234	0.9263	2.1972	1.8672	-1.0566	-1.3159	1.3919	3.4678	3.4899	1.4526	0.3394	3.3931	1.8188	-2.4344
SNG	-1.4210	-0.5079	NaN	-0.1179	1.3173	1.7953	-1.5001	-1.4714	1.3545	3.4460	3.4695	1.4221	0.3260	3.3829	1.7775	-2.6158
SIG	-0.7314	-0.0417	1.0605	NaN	2.5204	1.8839	-1.3621	-1.4154	1.3299	3.4246	3.4492	1.3910	0.3179	3.3956	1.7632	-2.5176
SNIG	-1.9690	-1.6051	-0.2693	-1.5743	NaN	1.2652	-1.9915	-1.6915	1.2420	3.4207	3.4468	1.3085	0.2126	3.3501	1.7008	-2.7358
SNW	-1.2972	-0.9593	-0.8086	-0.9248	-0.3782	NaN	-2.0478	-1.5791	1.2710	3.5363	3.5509	1.3518	0.1855	3.3299	1.7895	-2.7633
TLS	1.7295	1.9884	2.1672	2.2876	2.8052	2.4776	NaN	-0.9236	1.9000	3.5323	3.5387	1.9375	1.1279	3.4289	2.1615	-2.0659
LOG	1.4905	1.5893	1.5925	1.6393	1.8141	1.7043	1.2114	NaN	1.9438	3.4453	3.4503	1.9605	1.3779	3.2451	2.0754	-0.1756
Normal	0.5363	0.6167	0.8298	0.6601	0.8805	1.2430	0.5379	0.0549	NaN	3.6738	3.6288	2.0494	-1.1340	3.1642	1.8549	-0.1665
SG	4.4623	4.4409	4.4345	4.4258	4.4043	4.3044	4.4292	4.3076	4.8130	NaN	-0.8081	4.8169	4.4807	4.6024	4.4864	4.4941
SNG	4.2037	4.1832	4.1772	4.1728	4.1639	4.0895	4.1972	4.1244	4.4026	1.9765	NaN	4.4141	4.1534	4.6120	4.0603	4.2212
SIG	0.4933	0.5643	0.7836	0.6105	0.8332	1.2111	0.5213	0.0537	-0.4819	3.6792	3.6380	NaN	-1.2756	3.1352	1.7354	-0.1553
SNIG	0.7020	0.7985	0.9643	0.8211	1.0148	1.2897	0.5708	0.1141	2.1965	3.9217	3.8659	2.1740	NaN	3.1479	2.5080	-0.1576
SNW	0.9767	0.9861	1.1573	1.0495	1.1690	1.4341	1.0665	0.7262	-0.0009	0.8161	0.8709	-0.0311	0.1550	NaN	0.9336	0.7330
TLS	2.3221	2.3850	2.4115	2.3916	2.4757	2.6364	2.2617	1.8296	2.3822	3.2421	3.2841	2.3939	2.0029	3.6554	NaN	1.8252
LOG	1.7952	1.8207	1.8479	1.8357	1.8778	1.9544	1.6201	1.4041	2.1818	3.4566	3.5123	2.2078	1.9006	3.1793	2.3210	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions. The statistics are obtained using predictions at the daily horizon, using a sub-sample that covers the 2007-2008 financial crisis. Panel A reports the Diebold–Mariano (DM) statistics and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.

Table B.5: Diebold–Mariano and Clark-West test at the weekly horizon - crisis

Panel A: Diebold–Mariano test at weekly horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	-2.1240	-2.5351	-1.7472	-2.6408	-2.8119	1.8565	2.6189	-0.4902	2.6362	3.1266	-0.7979	0.0843	-1.0091	4.6596	4.7615
SG	2.1240	NaN	-0.9003	0.3051	-1.0639	-1.8496	2.7314	3.1582	-0.0556	2.7948	3.2640	-0.3774	0.6065	-0.7408	5.0059	5.2190
SNG	2.5351	0.9003	NaN	1.0718	0.1765	-1.2878	3.4580	3.4338	0.1425	2.8003	3.2613	-0.1555	0.7669	-0.5934	5.3236	5.5426
SIG	1.7472	-0.3051	-1.0718	NaN	-1.2320	-1.9271	2.7566	3.1961	-0.0933	2.7730	3.2429	-0.4100	0.5544	-0.7552	4.9961	5.2773
SNIG	2.6408	1.0639	-0.1765	1.2320	NaN	-1.5394	3.4229	3.4891	0.1068	2.7930	3.2581	-0.1956	0.7309	-0.6307	5.1947	5.4501
SNW	2.8119	1.8496	1.2878	1.9271	1.5394	NaN	3.7352	3.4518	0.5093	2.8942	3.3340	0.2259	1.1613	-0.3276	5.4986	5.6276
TLS	-1.8565	-2.7314	-3.4580	-2.7566	-3.4229	-3.7352	NaN	2.0418	-1.0707	1.9808	2.5054	-1.2687	-0.7646	-1.3560	4.3807	4.7647
LOG	-2.6189	-3.1582	-3.4338	-3.1961	-3.4891	-3.4518	-2.0418	NaN	-1.7871	1.4835	2.0390	-1.9440	-1.5679	-1.8471	2.8777	3.1220
Normal	0.4902	0.0556	-0.1425	0.0933	-0.1068	-0.5093	1.0707	1.7871	NaN	3.8589	4.3050	-1.6001	1.2311	-1.0174	3.5515	3.5228
SG	-2.6362	-2.7948	-2.8003	-2.7730	-2.7930	-2.8942	-1.9808	-1.4835	-3.8589	NaN	4.1648	-4.1204	-3.2404	-4.9178	-0.1165	-0.2152
SNG	-3.1266	-3.2640	-3.2613	-3.2429	-3.2581	-3.3340	-2.5054	-2.0390	-4.3050	-4.1648	NaN	-4.5311	-3.7192	-5.5766	-0.7856	-0.8677
SIG	0.7979	0.3774	0.1555	0.4100	0.1956	-0.2259	1.2687	1.9440	1.6001	4.1204	4.5311	NaN	2.0212	-0.7602	3.6669	3.6307
SNIG	-0.0843	-0.6065	-0.7669	-0.5544	-0.7309	-1.1613	0.7646	1.5679	-1.2311	3.2404	3.7192	-2.0212	NaN	-1.1875	3.5135	3.5133
SNW	1.0091	0.7408	0.5934	0.7552	0.6307	0.3276	1.3560	1.8471	1.0174	4.9178	5.5766	0.7602	1.1875	NaN	3.4018	3.2530
TLS	-4.6596	-5.0059	-5.3236	-4.9961	-5.1947	-5.4986	-4.3807	-2.8777	-3.5515	0.1165	0.7856	-3.6669	-3.5135	-3.4018	NaN	-0.2297
LOG	-4.7615	-5.2190	-5.5426	-5.2773	-5.4501	-5.6276	-4.7647	-3.1220	-3.5228	0.2152	0.8677	-3.6307	-3.5133	-3.2530	0.2297	NaN

Panel B: Clark-West test at weekly horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	2.4217	2.8566	2.2168	2.6668	2.7413	-0.1651	-0.6917	1.3912	2.2990	2.3147	1.4230	0.5754	2.7783	-0.6798	-1.5623
SG	-0.5473	NaN	1.4873	0.7294	1.3564	1.8653	-0.7400	-1.0413	1.0783	2.1190	2.1571	1.1209	0.2291	2.6065	-1.0125	-1.9175
SNG	-1.0033	-0.0628	NaN	-0.0652	0.8531	1.5871	-1.2416	-1.3024	1.0386	2.1240	2.1580	1.0809	0.2142	2.5446	-1.3697	-2.2562
SIG	-0.1002	1.3687	1.7476	NaN	1.6241	1.7887	-0.7231	-1.0591	1.1174	2.0953	2.1314	1.1541	0.2917	2.5878	-0.9068	-1.9440
SNIG	-0.7852	-0.0007	0.9862	-0.0049	NaN	1.6202	-1.0970	-1.2455	1.0117	2.0740	2.1140	1.0625	0.2211	2.5535	-1.2073	-2.1214
SNW	-0.8387	-0.4061	-0.0974	-0.4413	-0.1663	NaN	-1.2980	-1.2281	0.9549	2.2188	2.2431	1.0190	0.0714	2.5421	-1.7475	-2.2637
TLS	1.6935	1.9832	2.2040	2.0432	2.3541	2.0650	NaN	-0.6206	1.7624	2.5229	2.5253	1.7646	1.1971	2.6884	-0.2354	-2.3487
LOG	1.5455	1.6836	1.6862	1.6939	1.7913	1.6794	1.2179	NaN	1.8019	2.4379	2.4467	1.7966	1.3820	2.5717	0.7203	-0.4357
Normal	0.8806	1.0487	1.2486	1.0555	1.1491	1.5067	0.8443	0.3942	NaN	1.6195	1.5227	1.9765	0.0072	2.5146	0.6613	0.2145
SG	3.8203	3.7936	3.8253	3.7822	3.7672	3.7798	3.8625	3.7686	4.2148	NaN	-1.6835	4.2527	4.0446	3.9356	3.9196	3.8855
SNG	3.7249	3.7055	3.7255	3.6950	3.6913	3.6767	3.7824	3.7351	4.0056	2.5571	NaN	4.0362	3.8422	4.0422	3.8156	3.8053
SIG	0.5965	0.7543	0.9695	0.7617	0.8852	1.2511	0.6619	0.2491	-0.3327	1.4252	1.3542	NaN	-0.7210	2.2371	0.4643	0.0670
SNIG	0.6444	0.8442	1.0219	0.8587	0.9486	1.3085	0.5043	0.1100	1.6016	2.1947	2.1325	1.7220	NaN	2.6029	0.2490	-0.2571
SNW	1.9364	1.9868	2.0953	1.9873	2.0674	2.0972	2.0406	1.8374	1.3920	-0.2119	-0.5240	1.3677	1.3685	NaN	1.7394	1.7463
TLS	3.0739	3.2341	3.2833	3.2452	3.3302	3.3282	3.0123	2.6276	3.0233	3.3127	3.2592	2.9999	2.7862	3.4142	NaN	2.1464
LOG	3.2583	3.4631	3.5591	3.5072	3.5957	3.4521	3.3343	2.6840	2.9938	3.1577	3.1288	2.9588	2.7273	3.2783	2.2161	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions. The statistics are obtained using predictions at the weekly horizon, using a sub-sample that covers the 2007-2008 financial crisis. Panel A reports the Diebold–Mariano (DM) statistics and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.

Table B.6: Diebold–Mariano and Clark-West test at the monthly horizon - crisis

Panel A: Diebold–Mariano test at monthly horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	0.3941	-2.3328	-1.5733	-1.1064	-4.4230	-1.1882	-0.7183	-4.8971	1.5385	2.6145	-4.4731	-5.8439	0.7215	2.3551	2.7140
SG	-0.3941	NaN	-2.9043	-1.9333	-1.4387	-4.8420	-1.4237	-0.9448	-4.4742	1.3699	2.4471	-4.1822	-5.4555	0.5897	2.3782	2.7328
SNG	2.3328	2.9043	NaN	1.0815	1.0560	-2.8748	-0.1338	0.5843	-2.9531	2.1213	3.0393	-2.6541	-3.9767	1.3418	3.2141	3.5846
SIG	1.5733	1.9333	-1.0815	NaN	0.1760	-3.6613	-0.5737	0.0203	-3.7418	1.9267	2.8965	-3.4064	-4.7230	1.1041	2.8081	3.1824
SNIG	1.1064	1.4387	-1.0560	-0.1760	NaN	-3.0408	-0.6492	-0.0774	-3.4751	1.8050	2.7870	-3.1554	-4.2912	1.0149	2.8078	3.1582
SNW	4.4230	4.8420	2.8748	3.6613	3.0408	NaN	1.1167	1.8068	-1.4986	2.9864	3.7467	-1.2563	-2.1421	2.1246	3.5790	3.8673
TLS	1.1882	1.4237	0.1338	0.5737	0.6492	-1.1167	NaN	0.8227	-1.5824	1.6309	2.4868	-1.4234	-1.9272	1.0763	4.3794	4.9350
LOG	0.7183	0.9448	-0.5843	-0.0203	0.0774	-1.8068	-0.8227	NaN	-2.2769	1.4650	2.3915	-2.0632	-2.7628	0.8659	3.4737	4.0410
Normal	4.8971	4.4742	2.9531	3.7418	3.4751	1.4986	1.5824	2.2769	NaN	4.4986	5.0989	0.6627	-0.2599	3.5034	3.5961	3.8234
SG	-1.5385	-1.3699	-2.1213	-1.9267	-1.8050	-2.9864	-1.6309	-1.4650	-4.4986	NaN	5.0761	-4.7399	-4.0050	-2.1196	0.7659	1.0182
SNG	-2.6145	-2.4471	-3.0393	-2.8965	-2.7870	-3.7467	-2.4868	-2.3915	-5.0989	-5.0761	NaN	-5.2907	-4.6380	-6.0943	-0.1462	0.1026
SIG	4.4731	4.1822	2.6541	3.4064	3.1554	1.2563	1.4234	2.0632	-0.6627	4.7399	5.2907	NaN	-0.6115	3.6584	3.4557	3.6840
SNIG	5.8439	5.4555	3.9767	4.7230	4.2912	2.1421	1.9272	2.7628	0.2599	4.0050	4.6380	0.6115	NaN	3.0974	3.9054	4.1458
SNW	-0.7215	-0.5897	-1.3418	-1.1041	-1.0149	-2.1246	-1.0763	-0.8659	-3.5034	2.1196	6.0943	-3.6584	-3.0974	NaN	1.1585	1.3903
TLS	-2.3551	-2.3782	-3.2141	-2.8081	-2.8078	-3.5790	-4.3794	-3.4737	-3.5961	-0.7659	0.1462	-3.4557	-3.9054	-1.1585	NaN	0.7711
LOG	-2.7140	-2.7328	-3.5846	-3.1824	-3.1582	-3.8673	-4.9350	-4.0410	-3.8234	-1.0182	-0.1026	-3.6840	-4.1458	-1.3903	-0.7711	NaN

Panel B: Clark-West test at monthly horizon

	GARCH								GAS							
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG
Normal	NaN	1.2510	3.3709	2.9979	2.7156	4.1448	2.4184	2.7703	4.1963	1.8308	1.6602	3.8203	4.4080	1.9246	1.9503	0.7939
SG	1.9344	NaN	2.6027	2.1295	2.0678	3.3500	2.2209	2.5400	3.5082	1.6301	1.5424	3.2689	3.5637	1.8616	1.9024	0.7033
SNG	0.1356	-1.0716	NaN	0.7388	0.5889	3.0824	1.6178	1.9811	3.1643	1.3334	1.3074	2.7897	3.4647	1.5185	1.3511	0.0916
SIG	1.2159	-0.4795	2.9061	NaN	1.6351	3.5524	2.0280	2.2580	3.6397	1.3549	1.3211	3.2287	3.8720	1.6822	1.7646	0.4819
SNIG	1.2215	-0.0782	2.5990	1.7886	NaN	3.0707	2.1486	2.4616	3.4202	1.3697	1.3271	3.0977	3.7003	1.6601	1.7763	0.5265
SNW	-1.3329	-1.9856	-0.6750	-1.1971	-1.0007	NaN	0.8635	1.0388	2.6028	0.8569	0.8391	2.2342	2.6145	0.9892	0.8910	-0.1404
TLS	0.9857	0.7223	1.2652	1.2079	1.2555	1.5882	NaN	1.3391	2.1417	1.8550	1.8657	2.0250	1.9807	1.7461	0.1200	-1.4135
LOG	1.3481	1.1026	1.5847	1.5027	1.5464	1.9164	1.7996	NaN	2.4101	1.8958	1.8878	2.2729	2.3480	1.8923	1.2171	-0.4301
Normal	-0.8717	-0.9852	-0.0721	-0.3826	-0.3017	0.4797	0.9662	0.8966	NaN	0.0512	-0.0779	0.3567	1.0481	-0.2140	1.0821	0.2402
SG	2.7033	2.4728	2.8433	2.6624	2.6214	2.9355	3.2739	3.2267	3.7563	NaN	-1.8473	3.8328	3.6529	3.9713	3.4360	3.2752
SNG	2.9319	2.7997	3.0372	2.9238	2.8967	3.0330	3.3869	3.3920	3.7025	2.7031	NaN	3.7424	3.5899	4.6937	3.5032	3.4816
SIG	-0.7042	-0.8628	0.0193	-0.2927	-0.1834	0.5091	1.0290	0.9604	1.4182	-0.0902	-0.2278	NaN	0.8816	-0.3916	1.1398	0.3534
SNIG	-1.6669	-1.6854	-0.8677	-1.0575	-0.8849	-0.0723	0.4012	0.3843	0.7996	0.4850	0.4076	0.2235	NaN	0.1815	0.4867	-0.3835
SNW	2.2501	2.1232	2.4505	2.3529	2.3262	2.3972	2.7774	2.8726	2.9486	1.7779	-0.8442	2.9474	2.7995	NaN	2.6699	2.6135
TLS	2.7138	2.7779	2.9729	2.9580	3.0028	2.9373	3.3035	3.1822	2.9852	2.9014	2.8236	2.9097	2.9206	2.5714	NaN	2.4936
LOG	2.5323	2.5468	2.7369	2.7134	2.7905	2.6776	3.0793	2.9484	2.7476	2.6585	2.6040	2.6736	2.6732	2.3667	2.1684	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions. The statistics are obtained using predictions at the monthly horizon, using a sub-sample that covers the 2007-2008 financial crisis. Panel A reports the Diebold–Mariano (DM) statistics and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.

Table B.7: Diebold–Mariano and Clark-West test at the daily horizon - including IV

Panel A: Diebold-Mariano test at daily horizon

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	NaN	2.0972	-4.812	-0.668	-5.443	7.3373	0.6739	1.9454	2.3720	4.0628	4.4306	1.7507	3.3404	3.0161	-2.423	2.9815	1.4278	12.421
SG	-2.097	NaN	-4.887	-2.755	-6.581	6.6354	0.0400	1.5016	2.2064	3.9363	4.3187	1.5771	3.0677	2.9574	-2.546	2.8610	1.2468	12.221
SNG	4.8123	4.8866	NaN	3.2148	-1.008	9.6757	2.5542	3.0647	2.8728	4.4352	4.7750	2.2662	3.9540	3.2058	-1.959	3.3298	1.9800	13.048
SIG	0.6676	2.7550	-3.215	NaN	-5.365	8.0085	1.0187	2.2001	2.4138	4.1247	4.4939	1.8020	3.2739	3.0503	-2.332	3.0300	1.5194	12.461
SNIG	5.4426	6.5813	1.0084	5.3649	NaN	10.904	3.7573	3.8126	2.9614	4.5100	4.8523	2.3668	4.1110	3.2555	-1.801	3.4410	2.1387	13.255
SNW	-7.337	-6.635	-9.676	-8.009	-10.90	NaN	-4.982	-1.962	0.2644	2.6233	3.0838	-0.263	0.4023	2.3524	-3.842	1.7635	-0.409	10.544
TLS	-0.674	-0.040	-2.554	-1.019	-3.757	4.9818	NaN	2.0504	1.8187	3.7335	4.1416	1.2980	2.3658	2.9278	-2.463	2.8990	1.2230	11.792
LOG	-1.945	-1.502	-3.065	-2.200	-3.813	1.9616	-2.050	NaN	1.1517	3.1609	3.5791	0.6585	1.4744	2.6592	-2.842	2.4165	0.5478	10.562
Normal	-2.372	-2.206	-2.873	-2.414	-2.961	-0.264	-1.819	-1.152	NaN	2.6660	3.1469	-3.635	0.0632	2.5284	-3.731	1.3187	-0.525	7.9780
SG	-4.063	-3.936	-4.435	-4.125	-4.510	-2.623	-3.734	-3.161	-2.666	NaN	2.1192	-3.068	-2.677	0.8982	-6.650	-0.581	-3.672	5.0010
SNG	-4.431	-4.319	-4.775	-4.494	-4.852	-3.084	-4.142	-3.579	-3.147	-2.119	NaN	-3.513	-3.111	0.5380	-6.367	-1.025	-4.082	3.9004
SIG	-1.751	-1.577	-2.266	-1.802	-2.367	0.2632	-1.298	-0.659	3.6349	3.0679	3.5133	NaN	0.8347	2.7558	-3.322	1.6009	-0.107	8.3347
SNIG	-3.340	-3.068	-3.954	-3.274	-4.111	-0.402	-2.366	-1.474	-0.063	2.6966	3.1105	-0.835	NaN	2.2961	-4.679	1.4163	-0.590	9.0227
SNW	-3.016	-2.957	-3.206	-3.050	-3.256	-2.352	-2.928	-2.659	-2.528	-0.898	-0.538	-2.756	-2.296	NaN	-3.701	-1.175	-2.494	1.4451
TLS	2.4226	2.5458	1.9586	2.3317	1.8005	3.8423	2.4626	2.8423	3.7313	6.6497	6.3672	3.3223	4.6787	3.7029	NaN	3.8473	3.3020	11.725
LOG	-2.982	-2.861	-3.330	-3.030	-3.441	-1.764	-2.899	-2.417	-1.319	0.5806	1.0254	-1.601	-1.416	1.1747	-3.847	NaN	-1.698	4.3339
IV	-1.428	-1.247	-1.980	-1.519	-2.139	0.4086	-1.223	-0.548	0.5252	3.6723	4.0818	0.1073	0.5900	2.4944	-3.302	1.6975	NaN	21.492
MFIV	-12.42	-12.22	-13.09	-12.46	-13.25	-10.54	-11.79	-10.56	-7.978	-5.001	-3.900	-8.335	-9.023	-1.451	-11.73	-4.334	-21.49	NaN

Panel B: Clark-West test at daily horizon

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	NaN	-0.510	4.5364	1.5038	4.7795	-2.629	0.6915	-0.249	0.7024	4.9164	4.9055	0.8876	-0.169	2.4299	6.1367	-1.487	5.0177	1.9386
SG	3.0554	NaN	5.3921	3.3177	5.8005	-2.530	1.1272	0.0177	0.7762	4.9063	4.8911	0.9567	-0.013	2.4040	6.2499	-1.160	5.0052	1.8822
SNG	-3.492	-4.444	NaN	-2.946	2.0724	-4.308	-0.773	-0.962	0.3709	4.7438	4.7405	0.5526	-0.579	2.2855	5.7690	-2.215	4.8269	1.6978
SIG	0.6366	-1.407	4.0560	NaN	5.3362	-2.982	0.5084	-0.357	0.6462	4.7866	4.7981	0.8224	-0.197	2.3776	5.9715	-1.571	4.9739	1.8965
SNIG	-3.431	-4.456	-0.264	-3.724	NaN	-4.420	-1.350	-1.224	0.3516	4.7152	4.7104	0.5426	-0.679	2.2533	5.5955	-2.479	4.7437	1.5536
SNW	5.4984	5.0953	7.1119	5.7297	6.8790	NaN	4.0127	2.2670	2.7688	5.4948	5.4144	2.9731	1.9794	2.8072	7.7403	1.1521	5.2141	1.6020
TLS	1.4795	1.1399	2.3744	1.6825	3.0038	-0.853	NaN	-0.525	1.2450	4.4822	4.4683	1.3864	0.4161	2.4262	5.2771	-2.180	4.5808	1.7905
LOG	1.2023	1.0853	1.4937	1.2569	1.7208	0.2310	1.0344	NaN	1.3040	4.2429	4.2317	1.3898	0.7186	2.4084	4.3909	-0.520	4.2771	2.0326
Normal	2.7750	2.7110	3.0432	2.8022	3.0339	2.1071	2.6531	2.3204	NaN	5.4993	5.6275	4.9781	1.1280	1.4462	5.4050	2.0708	4.5694	2.5423
SG	6.8266	6.7972	6.8342	6.7887	6.8071	6.1218	6.5562	6.5547	7.6771	NaN	1.5003	7.7661	7.2148	5.5873	9.1869	6.4953	7.1220	5.4739
SNG	6.2996	6.2679	6.3005	6.2724	6.2893	5.7091	6.0753	6.1030	7.0587	3.6855	NaN	7.1498	6.6285	5.6100	8.3494	5.9166	6.7394	5.1484
SIG	2.2824	2.2053	2.5597	2.3119	2.5771	1.7081	2.2456	1.9350	-0.480	5.2222	5.4190	NaN	0.7527	1.1688	4.9809	1.7311	4.2886	2.3110
SNIG	3.5387	3.4191	3.8880	3.5087	3.8769	2.2380	3.1211	2.5068	2.1291	6.5353	6.4561	2.8762	NaN	2.2085	8.6514	2.3239	5.4551	3.2128
SNW	2.6456	2.6122	2.6970	2.6413	2.6965	2.4210	2.6673	2.6546	2.5729	2.3367	2.3398	2.5940	2.6082	NaN	3.5602	2.5694	2.2545	1.5739
TLS	3.3386	3.3905	3.4409	3.3768	3.4821	3.4672	3.1543	2.7058	2.8910	5.5551	5.6256	3.0034	2.5274	4.5390	NaN	2.4761	6.3107	5.9059
LOG	2.5202	2.4619	2.6463	2.5327	2.6849	2.0553	2.3690	2.1676	2.6228	4.4205	4.3967	2.6885	2.2570	3.2441	4.4681	NaN	4.0646	2.5036
IV	6.6403	6.5428	6.8842	6.6802	6.8891	5.2197	6.3317	5.7776	6.2766	6.2141	6.2250	6.5158	5.9765	5.3377	10.299	5.4112	NaN	-10.41
MFIV	10.719	10.658	10.878	10.712	10.892	9.5536	10.277	9.8657	10.769	9.8365	9.6121	10.941	10.543	8.7709	13.288	9.7576	12.863	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions, including IV and MFIV. The statistics are obtained using predictions at the daily horizon, using a sample that ranges between 01/1996 and 12/2017. Panel A reports the Diebold–Mariano (DM) statistics and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.

Table B.8: Diebold–Mariano and Clark-West test at the weekly horizon - including IV

Panel A: Diebold-Mariano test at weekly horizon

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	NaN	0.272	-3.473	0.315	-2.866	2.381	3.514	3.019	7.189	6.148	7.199	6.361	7.923	6.812	9.083	9.520	6.305	15.336
SG	-0.272	NaN	-3.850	0.108	-3.492	2.410	3.465	2.976	7.222	6.136	7.186	6.391	7.975	6.804	9.124	9.662	6.285	15.379
SNG	3.473	3.850	NaN	3.666	0.143	3.869	5.602	4.143	7.555	6.355	7.365	6.752	8.363	7.019	9.880	10.525	6.870	16.029
SIG	-0.315	-0.108	-3.666	NaN	-3.209	2.267	3.508	3.022	7.073	6.117	7.160	6.250	7.812	6.764	9.081	9.701	6.238	15.264
SNIG	2.866	3.492	-0.143	3.209	NaN	4.122	5.250	4.107	7.614	6.375	7.398	6.807	8.429	7.062	9.797	10.490	6.966	16.206
SNW	-2.381	-2.410	-3.869	-2.267	-4.122	NaN	0.456	0.839	5.331	5.432	6.560	4.511	6.040	6.423	8.454	8.763	6.060	16.569
TLS	-3.514	-3.465	-5.602	-3.508	-5.250	-0.456	NaN	0.920	4.366	4.949	6.060	3.626	4.727	5.949	8.541	9.243	4.980	14.218
LOG	-3.019	-2.976	-4.143	-3.022	-4.107	-0.839	-0.920	NaN	3.653	4.604	5.711	2.963	3.863	5.611	7.408	8.063	4.349	12.942
Normal	-7.189	-7.222	-7.555	-7.073	-7.614	-5.331	-4.366	-3.653	NaN	3.970	5.599	-3.219	-0.189	5.144	3.086	2.784	1.498	10.607
SG	-6.148	-6.136	-6.355	-6.117	-6.375	-5.432	-4.949	-4.604	-3.970	NaN	6.519	-4.605	-3.793	1.484	-0.615	-0.948	-2.305	4.310
SNG	-7.199	-7.186	-7.365	-7.160	-7.398	-6.560	-6.060	-5.711	-5.599	-6.519	NaN	-6.174	-5.313	-0.087	-2.061	-2.364	-3.927	2.243
SIG	-6.361	-6.391	-6.752	-6.250	-6.807	-4.511	-3.626	-2.963	3.219	4.605	6.174	NaN	1.445	5.612	3.623	3.357	2.151	11.219
SNIG	-7.923	-7.975	-8.363	-7.812	-8.429	-6.040	-4.727	-3.863	0.189	3.793	5.313	-1.445	NaN	4.707	3.390	3.106	1.566	10.984
SNW	-6.812	-6.804	-7.019	-6.764	-7.062	-6.423	-5.949	-5.611	-5.144	-1.484	0.087	-5.612	-4.707	NaN	-1.899	-2.225	-3.431	2.200
TLS	-9.083	-9.124	-9.880	-9.081	-9.797	-8.454	-8.541	-7.408	-3.086	0.615	2.061	-3.623	-3.390	1.899	NaN	-0.684	-1.814	6.698
LOG	-9.520	-9.662	-10.53	-9.701	-10.49	-8.763	-9.243	-8.063	-2.784	0.948	2.364	-3.357	-3.106	2.225	0.684	NaN	-1.380	7.559
IV	-6.305	-6.285	-6.870	-6.238	-6.966	-6.060	-4.980	-4.349	-1.498	2.305	3.927	-2.151	-1.566	3.431	1.814	1.380	NaN	27.333
MFIV	-15.34	-15.38	-16.03	-15.27	-16.21	-16.57	-14.22	-12.94	-10.61	-4.310	-2.243	-11.22	-10.99	-2.200	-6.698	-7.559	-27.33	NaN

Panel B: Clark-West test at weekly horizon

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	NaN	1.374	3.700	1.575	3.360	1.835	-0.255	-0.423	0.630	3.593	3.441	0.748	0.180	2.004	0.364	-0.670	4.742	3.509
SG	2.010	NaN	3.942	1.553	3.345	1.441	-0.341	-0.459	0.418	3.468	3.315	0.540	-0.040	1.842	0.128	-0.907	4.576	3.274
SNG	-1.087	-2.336	NaN	-1.751	1.460	0.706	-1.443	-1.168	0.210	3.379	3.237	0.333	-0.319	1.742	-0.389	-1.414	4.433	3.143
SIG	2.217	1.439	3.734	NaN	3.209	1.551	-0.264	-0.453	0.643	3.435	3.314	0.754	0.206	1.906	0.400	-0.751	4.614	3.446
SNIG	0.043	-0.812	1.475	-0.490	NaN	0.436	-1.083	-0.964	0.114	3.303	3.151	0.250	-0.440	1.613	-0.470	-1.545	4.418	3.016
SNW	3.579	3.448	4.090	3.463	4.049	NaN	2.105	1.750	0.776	3.736	3.402	0.934	0.086	1.293	-0.435	-1.233	3.678	1.054
TLS	2.843	2.686	3.394	2.871	3.604	2.473	NaN	0.304	1.130	3.298	3.150	1.190	0.728	1.884	0.218	-1.131	4.237	2.878
LOG	1.719	1.638	1.875	1.682	2.059	1.838	0.892	NaN	1.237	3.161	3.048	1.272	0.950	1.930	0.874	-0.151	3.995	3.067
Normal	6.862	6.777	6.994	6.815	6.837	5.877	5.112	4.844	NaN	4.075	3.558	3.047	2.310	0.383	2.701	2.109	4.231	1.765
SG	6.909	6.824	6.870	6.825	6.802	6.355	6.320	6.477	6.211	NaN	-0.784	6.329	6.109	4.260	5.521	5.370	6.007	4.763
SNG	6.571	6.492	6.526	6.508	6.480	5.971	6.071	6.255	6.027	4.934	NaN	6.128	5.843	4.687	5.264	5.144	5.852	4.594
SIG	6.277	6.177	6.420	6.226	6.253	5.269	4.579	4.343	-0.913	3.749	3.242	NaN	1.006	-0.164	2.180	1.590	3.797	1.324
SNIG	6.831	6.697	6.928	6.786	6.729	5.529	4.944	4.651	2.123	4.680	4.191	2.544	NaN	1.135	2.365	1.608	4.141	1.503
SNW	5.173	5.119	5.205	5.130	5.157	4.726	4.965	5.039	3.851	2.965	2.570	3.849	3.882	NaN	3.711	3.761	3.254	2.203
TLS	6.168	6.151	6.385	6.226	6.415	5.413	5.985	5.790	3.621	4.430	4.084	3.629	3.466	3.168	NaN	3.496	4.738	2.770
LOG	6.832	6.843	7.145	6.982	7.193	5.723	6.927	6.621	3.399	4.022	3.715	3.392	3.247	2.842	3.296	NaN	4.630	2.568
IV	9.296	9.142	9.400	9.173	9.381	8.266	8.365	8.058	6.008	5.082	4.397	6.076	5.979	3.590	5.743	5.039	NaN	-11.63
MFIV	12.454	12.341	12.573	12.362	12.633	12.217	11.732	11.480	9.877	8.059	7.240	9.927	9.999	7.345	10.339	9.638	13.949	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions, including IV and MFIV. The statistics are obtained using predictions at the weekly horizon, using a sample that ranges between 01/1996 and 12/2017. Panel A reports the Diebold–Mariano (DM) statistics and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.

Table B.9: Diebold–Mariano and Clark-West test at the monthly horizon - including IV

Panel A: Diebold-Mariano test at monthly horizon

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	NaN	-0.797	-6.091	-1.112	-7.690	-15.02	-4.034	-2.624	-9.296	-4.284	-1.563	-10.27	-1.370	-2.728	-2.492	-3.268	-4.645	0.341
SG	0.797	NaN	-5.752	-0.325	-7.181	-14.97	-3.839	-2.284	-9.021	-4.082	-1.412	-9.984	-1.309	-2.548	-2.415	-3.199	-4.497	0.484
SNG	6.091	5.752	NaN	5.605	-3.271	-13.79	-1.202	0.940	-7.896	-3.035	-0.478	-8.845	-0.884	-1.537	-1.475	-2.258	-3.553	1.517
SIG	1.112	0.325	-5.605	NaN	-7.184	-14.87	-3.730	-2.107	-8.994	-4.038	-1.362	-9.958	-1.284	-2.500	-2.343	-3.137	-4.460	0.542
SNIG	7.690	7.181	3.271	7.184	NaN	-12.74	0.872	2.852	-7.202	-2.364	0.175	-8.167	-0.592	-0.828	-0.778	-1.522	-2.852	2.251
SNW	15.020	14.971	13.787	14.871	12.736	NaN	10.227	11.496	0.459	3.956	5.844	-0.393	1.570	5.631	4.887	4.393	3.595	9.156
TLS	4.034	3.839	1.202	3.730	-0.872	-10.23	NaN	2.509	-6.663	-2.319	-0.121	-7.366	-0.722	-1.032	-1.236	-2.143	-3.300	2.007
LOG	2.624	2.284	-0.940	2.107	-2.852	-11.50	-2.509	NaN	-7.534	-2.976	-0.670	-8.276	-0.991	-1.646	-1.845	-2.747	-3.937	1.286
Normal	9.296	9.021	7.896	8.994	7.202	-0.459	6.663	7.534	NaN	5.018	7.415	-1.770	1.490	7.492	4.546	3.974	3.571	9.760
SG	4.284	4.082	3.035	4.038	2.364	-3.956	2.319	2.976	-5.018	NaN	8.409	-6.601	0.306	3.658	1.155	0.612	-0.465	4.342
SNG	1.563	1.412	0.478	1.362	-0.175	-5.844	0.121	0.670	-7.415	-8.409	NaN	-8.792	-0.663	-2.524	-0.744	-1.278	-2.657	1.864
SIG	10.271	9.984	8.845	9.958	8.167	0.393	7.366	8.276	1.770	6.601	8.792	NaN	1.669	9.425	5.037	4.481	4.333	10.550
SNIG	1.370	1.309	0.884	1.284	0.592	-1.570	0.722	0.991	-1.490	-0.306	0.663	-1.669	NaN	0.265	0.262	-0.008	-0.488	1.527
SNW	2.728	2.548	1.537	2.500	0.828	-5.631	1.032	1.646	-7.492	-3.658	2.524	-9.425	-0.265	NaN	0.011	-0.553	-1.988	3.120
TLS	2.492	2.415	1.475	2.343	0.778	-4.887	1.236	1.845	-4.546	-1.155	0.744	-5.037	-0.262	-0.011	NaN	-1.197	-2.431	4.079
LOG	3.268	3.199	2.258	3.137	1.522	-4.393	2.143	2.747	-3.974	-0.612	1.278	-4.481	0.008	0.553	1.197	NaN	-1.603	5.171
IV	4.645	4.497	3.553	4.460	2.852	-3.595	3.300	3.937	-3.571	0.465	2.657	-4.333	0.488	1.988	2.431	1.603	NaN	26.682
MFIV	-0.341	-0.484	-1.517	-0.542	-2.251	-9.156	-2.007	-1.286	-9.760	-4.342	-1.864	-10.55	-1.527	-3.120	-4.079	-5.171	-26.68	NaN

Panel B: Clark-West test at monthly horizon

	GARCH								GAS									
	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	Normal	SG	SNG	SIG	SNIG	SNW	TLS	LOG	IV	MFIV
Normal	NaN	2.640	5.798	3.758	6.191	7.218	4.432	4.215	10.140	8.792	8.796	10.099	10.026	8.982	9.300	8.617	9.263	9.390
SG	1.754	NaN	4.209	2.216	5.060	6.854	3.744	3.415	9.766	8.467	8.478	9.730	9.657	8.677	8.982	8.307	9.055	9.189
SNG	-2.132	-2.139	NaN	-1.576	4.414	6.761	3.041	2.461	9.877	8.368	8.366	9.815	9.777	8.536	9.032	8.256	9.103	9.227
SIG	2.739	1.904	6.181	NaN	6.111	7.052	4.456	4.058	9.966	8.484	8.513	9.914	9.862	8.750	9.198	8.471	9.144	9.303
SNIG	-2.302	-2.424	-0.177	-2.019	NaN	7.068	2.365	1.566	10.276	8.425	8.353	10.206	10.199	8.620	9.297	8.430	9.257	9.364
SNW	-1.934	-2.300	-1.727	-2.023	-1.802	NaN	0.018	-0.325	8.515	6.381	6.050	8.424	8.313	5.972	6.239	5.155	7.277	7.129
TLS	0.698	0.616	1.748	0.989	2.909	5.938	NaN	1.449	8.376	6.922	6.820	8.268	8.348	6.851	8.018	7.273	8.168	8.134
LOG	1.176	1.137	2.131	1.363	3.208	6.126	3.482	NaN	8.958	7.505	7.433	8.853	8.929	7.499	8.705	7.923	8.674	8.740
Normal	6.923	6.619	7.170	6.903	7.262	8.183	6.319	6.826	NaN	1.623	0.714	1.574	1.604	-0.556	1.460	1.012	3.670	1.893
SG	7.879	7.662	7.917	7.754	7.851	8.037	7.052	7.492	5.407	NaN	-1.302	5.660	5.017	4.483	4.074	4.150	4.902	4.064
SNG	8.246	8.057	8.219	8.145	8.154	7.786	7.260	7.784	5.439	5.086	NaN	5.619	5.055	6.073	4.082	4.260	4.659	3.760
SIG	6.744	6.421	6.994	6.711	7.097	8.180	6.070	6.604	0.712	1.110	0.149	NaN	0.125	-1.502	1.032	0.640	3.311	1.541
SNIG	3.086	3.012	3.116	3.091	3.103	2.859	2.922	3.085	1.300	1.583	1.413	1.308	NaN	1.080	1.347	1.250	1.918	1.491
SNW	8.400	8.190	8.431	8.326	8.444	8.119	7.489	8.084	5.361	4.973	2.682	5.559	4.767	NaN	3.469	3.724	4.207	2.936
TLS	6.948	6.984	7.283	7.133	7.346	6.612	7.632	7.660	3.469	3.400	2.959	3.411	3.334	2.425	NaN	3.901	4.692	3.078
LOG	6.468	6.498	6.823	6.648	6.903	6.016	7.417	7.371	2.856	2.860	2.501	2.801	2.755	1.985	2.622	NaN	4.096	2.486
IV	9.065	8.949	9.226	9.073	9.354	9.189	8.702	9.027	5.507	4.388	3.417	5.386	5.410	2.771	4.652	3.599	NaN	-9.243
MFIV	11.121	11.038	11.297	11.137	11.421	11.251	10.912	11.182	7.522	6.052	4.873	7.396	7.577	4.446	8.048	6.797	11.653	NaN

Notes: This table reports the pairwise Diebold–Mariano and Clark-West statistics computed between any two predictions, including IV and MFIV. The statistics are obtained using predictions at the monthly horizon, using a sample that ranges between 01/1996 and 12/2017. Panel A reports the Diebold–Mariano (DM) statistics and Panel B reports the Clark-West statistic (CW) statistics. In each panel, GARCH models are set as benchmarks in the first 8 rows, and GAS models are set as benchmarks in the last 8 rows.