Three Essays on Macroeconomic Information, Volatility Persistence and Structural Change

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities

2019

Wei Liu

Alliance Manchester Business School
Accounting and Finance Division

Table of Contents

5 Conclusions and Future Research Suggestions 224

This thesis contains 44,541 words, including title page, tables, and footnotes.

Abstract

The University of Manchester
Wei Liu
Doctor of Philosophy (PhD)
Three Essays on Macroeconomic Information,
Volatility Persistence and Structural Change
September 2019

This thesis explores the relationship between macroeconomic information and stock market volatility. Firstly, we employ the GARCH-MIDAS model to investigate the impact of fundamental macroeconomic variables on long-term stock market volatility across three developed countries: U.S, UK and Japan. Secondly, we further investigate the impact of macroeconomic information on long-term persistence and structural changes in volatility. Thirdly, we introduce a discrete-time realized volatility option pricing model which incorporates macroeconomic variable for option valuation.

This thesis consists of five chapters. In the first chapter, we make a brief introduction of my thesis. In the second chapter, we explore the relationship between macroeconomic information and long-term stock market volatility across three developed countries, U.S, UK and Japan. We employ the two-component GARCH-MIDAS model to carry out international analysis and observe the relationship between macroeconomic variables and stock market volatility changes over time, both in magnitude and significance. This time-varying relationship might largely attributes to the time-varying expectations of market participants towards forthcoming monetary policy changes and macroeconomic uncertainty.

In the third chapter, we extend the Heterogeneous Autoregressive Realized Volatility-type models (HAR and Tree-HAR models) that allows macroeconomic information to explain long-term persistence and structural changes in stock volatility, simultaneously. We find that both macroeconomic information and its uncertainty have prominent impacts on stock volatility. Strikingly, macroeconomic information helps to deliver a more elaborate regime-switching structure for U.S stock volatility, which infers a tight link between macroeconomic information and potential structural changes in stock volatility.

In the fourth chapter, we employ our extended HAR model for option pricing domain. We aim to examine whether macroeconomic information, through its influence on conditional volatility, can affect corresponding option prices? Root mean squared errors for both put and call S&P500 Index options show that adding macroeconomic information, in particular unanticipated information, into option pricing process increases option pricing accuracy and mitigates implied volatility biases, relative to traditional Black-Scholes model and GARCH model. In the last chapter, we make a conclusion and possible future directions.

Declaration

I, Wei Liu, declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

First and foremost, I would like to express deeply thanks to my supervisor, Prof. Ian Garrett for his continuous support and insighful advices. Studying with him, I learn how to conduct rigorous and detail research. He was alway patient and helpful to guide me through my PhD studies. Without his encouragement and inspiration, it would be impossible for me to complete this thesis.

I would like to thank Prof. Michael Brennan for being kindness to discuss with me and gives important suggestions on my second chapter. I deeply thank Prof. Hening Liu for his generous support and valuable advices on my third chapter. I thank Prof. Xiaoquan Jiang for his valuable suggestions on my second chapter during the 25th International Conference on Forecasting Financial Markets at the University of Oxford. I extend my thanks to Dr Yoichi Otsubo for sharing me with the high-frequency data set. A special thank goes to Mr Martin Turner for his friendly help and good sense of humor. It was really nice to have a coffee with him.

The completion of this dissertation brings to an end of several years of studying and living in Manchester, where I treat it as my second hometown. These years would not have been as memorable without my colleagues and friends, Xiaolong Chen, Chen Hua, Xinyu Cui, Xiangshang Cai, Shuwen Yang, Fei Xu, Jingyi Meng, Lincon James, Mengxin Peng, Adnan Gazi.

Finally, I am particularly grateful for the unconditional support as well as continuous encouragement of my family. This thesis is in loving memory to my dear mother, who could not be with me but always stay in my mind.

Wei Liu

Beijing China

Chapter 1

Introduction

As mentioned by Bollerslev et al. (1992), the volatility of asset returns in financial markets is time-varying in a persistent manner across different assets, time spans and countries. Accurately modelling asset volatility is of crucial importance for asset pricing, risk management and portfolio allocation. Recent financial crises and the associated large variations in volatility that accompany them indicate the need for deep exploration of the relationship between macroeconomic information and financial market volatility. Consequently, a large body of literature has emerged modelling the link between macroeconomy and financial volatility; see, for instance, Morana and Beltratti (2004), Engle and Rangel (2008), Christiansen et al. (2012), Engle et al. (2013), Conrad and Loch (2015) and Asgharian et al. (2013). In this thesis, we focus on the link between macroeconomic information and stock market volatility, and the application of the link between macroeconomic information and volatility for European Option pricing. Our empirical results provide a contribution to the debate on whether including macroeconomic information in volatility models delivers a better description of stock market volatility, helps our understanding of structural breaks in volatility and what drives them, improves the predictive performance of volatility models and helps in the pricing of options when volatility is not held constant. The following sections expand on these points and discuss the chapters of the thesis in more details.

1.0.1 The Relationship between Macroeconomic Information and Financial Volatility

The relation between macroeconomic information and financial market volatility was initially proposed by Schwert (1989), who raised the question "why does stock market volatility vary over time?". In Schwert (1989)'s cross sectional analysis, aggregated monthly stock volatility is explained by volatility of real and nominal macroeconomic variables and also financial leverage. Empirical results of Schwert (1989) infers a weak relationship between macroeconomic information and stock volatility, which might possibly due to the fact that long-term trend of stock returns (also volatility) are more related with risk premia and relevant macroeconomic information flow (see Chen et al. (1986) and Engle and Lee (1999)). An alternative explanation would be the mismatched frequency between macroeconomic observations (usually observed monthly or quarterly) and asset returns (usually observed daily or intra-daily.) Engle et al. (2013) successfully resolved this mismatch in frequency by introducing the GARCH-MIDAS model which allows volatility to vary over time (the GARCH part) while also allowing for the inclusion of macroeconomic information (the MIDAS part, MIDAS standing for Mixed Data Sampling.) This GARCH-MIDAS model can be nested into a two-component GARCH process that stems from Ding and Granger (1996) and Engle and Lee (1999), where volatility is decomposed into two parts: the long-term volatility component and the short-term volatility component. The short-term volatility component follows a standard GARCH process and slowly evolves around the long-term volatility component. Owing to the Mixed data sampling (MIDAS) approach

(Ghysels et al. (2007)), macroeconomic data observed at different frequencies can be included in the model to explain long-term volatility movements. The empirical results in Engle et al. (2013) reveal that industrial production growth and inflation improve the long-run prediction of U.S. stock market volatility.

In the second Chapter, We use the GARCH-MIDAS model to examine the relationship between macroeconomic variables and stock return volatility for the U.S, the UK and Japan to see whether the findings are consistent across countries and, given our sample includes the global financial crisis and the sovereign debt crisis, how the GARCH-MIDAS model performs over time, both in terms of the significance of the macroeconomic variables and the forecasting performance of the model. We find that macroeconomic variables are important determinants of return volatility in the U.S, the UK and Japan and that the inclusion of macroeconomic variables improves forecasting performance in terms of predicting volatility. However, our sub-sample analysis reveals that while macroeconomic information matters, the significance of the individual macroeconomic variables, the magnitude of their coefficients and even the sign of the coefficients on some of the macroeconomic variables change over time, suggesting that parameters in the GARCH-MIDAS model are not stable over time and that consideration needs to be given to structural breaks in volatility and their determinants.

1.0.2 Macroeconomic Information and Potential Structural Breaks in Financial Volatility

Empirical results from our second chapter suggest that while evaluating the link between macroeconomic information and long-term financial volatility, it is necessary to bring potential structural breaks into consideration. Andersen and Bollerslev (1997) point out that neglecting proper consideration of structural breaks might artificially lead to a strong persistence of volatility in financial market. Engle et al. (2013) also mention that the GARCH-MIDAS with fixed parameters fails to reflect potential structural changes in volatility. One possible solution is to follow Engle et al. (2013) and partition the whole estimation sample into several sub-periods according to existing structural breaks in volatility, and then carry out GARCH-MIDAS analysis across the different sub-samples. Alternatively. we could employ a regime-switching volatility model that is able to account for structural breaks and different regimes in volatility.

Based on the GARCH framework, Audrino and Bühlmann (2001) introduce a regime-switching GARCH process, the Tree-GARCH model, where the number of regime is endogenously determined. Building on this, Audrino (2006) introduces a regime-switching realized volatility model, the Tree-Heterogeneous Autoregressive (HAR) model, where unobserved volatility is proxied by realized volatility. Regime determination embedded in both the Tree-GARCH and Tree-HAR models comes from the binary-tree algorithm of Breiman et al. (1984) in which the optimal number of regimes are endogenously driven by the data. Therefore, a tree-structured volatility model enables the adjustment of the number of regimes so as to fit structural changes in volatility. Consequently, it is able to generate a multi-regime structure, where macroeconomic variables might possibly be used as explanatory variable, explaining volatility movement within each regime, meanwhile identifying structural changes across regimes.

In the third chapter we use the Tree-HAR model to further analyse the relationship between macroeconomic information and stock market volatility for the US stock market. We extend the Tree-HAR model to allow macroeconomic information to not only explain long-term volatility movements within each regime, but also to determine the regime structure of volatility. We find that the macroeconomic variables individually explain not only long-term volatility movements but also different regimes for stock market volatility. These findings suggest that macroeconomic information is important not only for explaining return volatility but for determining (at least some) changes in volatility regime. The Tree-HAR model also suggests that the regime structure is more sophisticated than the more usual high/low volatility regimes. In particular, we find the Tree-HAR identifies a medium-volatility regime in addition to high and low regimes and that macroeconomic information often plays a role in further sub-dividing at least one of these regimes to give a stable four-regime structure in most cases.

1.0.3 Macroeconomic Information and Option Pricing Application

Due to the importance of macroeconomic information in explaining stock volatility, Christoffersen et al. (2009) utilizes the GARCH-MIDAS model with macroeconomic information for option pricing and finds that introducing macroeconomic information leads to some improvements in the pricing of options: pricing errors in the GARCH-MIDAS option pricing model are less than those observed in traditional GARCH and Black-Scholes option pricing models.

As mentioned in the second chapter, volatility in the GARCH-MIDAS model consists of short-term and long-term volatility components. Andersen and Bollerslev (1997) notes that volatility observed during short time spans largely accounts for high-frequency intraday information. Hence, failure to use intraday returns when estimating short-term volatility persistence, something which the GARCH-MIDAS model does not, might artificially lead to strong persistence in the short-term volatility component. Consequently, extreme persistence of the short-term volatility component in the GARCH-MIDAS option pricing model could have a negative impact on pricing of options.

Corsi (2009) develops the Heterogeneous Autoregressive (HAR) realized volatility (HAR) model that describes daily volatility as a sequence of autoregressive volatility components realized over daily, weekly and monthly horizons. Application of the HAR model in option pricing (Corsi et al. (2013)) reveals that it outperforms the GARCH option pricing model, especially for short-maturity European options. The HAR model has two advantages: First, it is able to mimic the long-memory feature of volatility, which benefits the pricing of long-maturity options and second, its multi-component structure makes it possible for low frequency macroeconomic information to be incorporated into a high-frequency volatility model.

In the fourth chapter, we incorporate macroeconomic information into the HAR model and develop a realized volatility option pricing model. Our results suggest that most macroeconomic variables outperform Duan's GARCH model across maturity and moneyness for both put and call options. Empirical results infer a tight link between option and market reactions towards changes in economy. For instance, macroeconomic variables that usually perform as recession indicators, such as term spread and unemployment rate, tends to be less effective for call option. While inflation factor, that is more active during economy expansion, seems to have limited performance for put option. Surprisingly, unexpected macroeconomic information, which is measured by economic uncertainty, outperforms alternative macroeconomic variables for out-of-money (OTM) options with long-maturity. It infers that unexpected macroeconomic shock matters for option valuation, espeicial for long-maturity options.

1.0.4 Thesis Structure

This thesis is structured around three self-contained essays in Chapter 2, 3 and 4. Each chapter has separate introduction, background information, methodology, data analysis, conclusion and reference. The equations, tables, figures, footnotes are sorted in sequential orders throughout the thesis.

The thesis continues as follows: Chapter 2 investigates the relationship between long-term stock market volatility and macroeconomic variables using the GARCH-MIDAS model throughout U.S, UK and Japan stock markets; Chapter 3 examines the impact of macroeconomic information, especially unexpected information, on the long-term persistency and structural changes in the U.S stock market using HAR-type realized volatility model; Chapter 4 introduces a discrete-time realized option pricing model incorporated with macroeconomic information. Chapter 5 makes conclusion of the major findings of the thesis.

Finally, I use third person (we, our) rather than the first person (I, my) throughout the thesis, indicating that three chapters are in the form of working, or submitted, co-authored with my supervisor Ian Garrett at Alliance Manchester Business School.

References

Andersen, T. G. and T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The journal of Finance 52*(3), 975–1005.

Asgharian, H., A. J. Hou, and F. Javed (2013). The importance of the macroeconomic variables in forecasting stock return variance: A garch-midas approach. *Journal of Forecasting 32*(7), 600–612.

Audrino, F. (2006). Tree-structured multiple regimes in interest rates. *Journal of Business & Economic Statistics 24*(3), 338–353.

Audrino, F. and P. Bühlmann (2001). Tree-structured generalized autoregressive conditional heteroscedastic models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63*(4), 727–744.

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics 52*(1-2), 5–59.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen (1984). *Classification and regression trees.* CRC press.

Chen, N.-F., R. Roll, and S. A. Ross (1986). Economic forces and the stock market. *Journal of business*, 383–403.

Christiansen, C., M. Schmeling, and A. Schrimpf (2012). A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics 27*(6), 956–977.

Christoffersen, P., R. Elkamhi, B. Feunou, and K. Jacobs (2009). Option valuation with conditional heteroskedasticity and nonnormality. *The Review of Financial Studies 23*(5), 2139–2183.

Conrad, C. and K. Loch (2015). Anticipating long-term stock market volatility. *Journal of Applied Econometrics 30*(7), 1090–1114.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, nbp001.

Corsi, F., N. Fusari, and D. La Vecchia (2013). Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics 107*(2), 284–304.

Ding, Z. and C. W. Granger (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of econometrics 73*(1), 185–215.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics 95*(3), 776–797.

Engle, R. F. and G. Lee (1999). A long-run and short-run component model of stock return volatility. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, 475–497.

Engle, R. F. and J. G. Rangel (2008). The spline-garch model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies 21*(3), 1187–1222.

Ghysels, E., A. Sinko, and R. Valkanov (2007). Midas regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Morana, C. and A. Beltratti (2004). Structural change and long-range dependence in volatility of exchange rates: either, neither or both? *Journal of Empirical Finance 11*(5), 629 – 658.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The journal of finance 44*(5), 1115–1153.

Chapter 2

Do the Macroeconomic variables contain explanatory information for stock market volatility? International evidence from the GARCH-MIDAS model

We explore the relationship between long-term stock market volatility and fundamental macroeconomic variables using the GARCH-MIDAS model across US, UK and Japan[1]. The GARCH-MIDAS model introduced by Engle et al. (2013) maintains a two-component volatility structure, where volatility is divided into a short-term component and a long-term component. While the short-term volatility component follows a standard GARCH process, the long-term volatility component is directly determined by macroeconomic informa-

[1]The reason for considering these markets maily arises from the fact that, US, UK and Japan stock markets represents the world's major centre for trading. For US, shares being traded in the New York exchange market accounts for 54.5% of total. For UK, shares being traded in the London exchange market roughly accounts for 5.1%. While for Japan, shares being traded in the Tokyo exchange market accounts for 7.7% of total.

tion. This is done through different lags of macroeconomic observations being brought together and transformed into a weighted average value via the Mixed data sampling (MIDAS) approach of Ghysels et al. (2007). Our findings suggest that macroeconomic variables have an uneven (or time-varying) impact on stock return volatility across countries, with the magnitude of impact and the statistical significance of the macroeconomic variables being different across different sub-periods in every financial market, although the term spread and first principal component perform rather well in most cases. This time-varying relationship might largely attributes to the time-varying expectations of market participants towards forthcoming monetary policy changes and macroeconomic uncertainty. We also observe that long-term volatility originating from the US market transmits to the UK and Japanese markets. Consequently, considering US volatility spillovers helps us to capture the real impact of macroeconomic information on local stock volatility, which is complementary to the global volatility transmission from US to local stock volatility.

Keywords: Mixed Data Sampling, Principal Component, Macroeconomic Variables, Model Confidence Set.

2.1 Introduction

It is well known that the estimation and prediction of volatility is essential for asset pricing and risk management. Following the seminal work of Schwert (1989), one area that has received significant attention is the relationship between stock market volatility and macroeconomic information. Schwert (1989) examines changes in volatility over time and the extent to which such changes are driven by the volatility of macroeconomic variables. He finds that macroeconomic information tends to have limited success in explaining changes in volatility over time. One possible explanation for Schwert (1989) limited findings might be: macroeconomic source is more influential on the long-run trend of variations in financial market, rather than short-run variations (see Chen et al. (1986) and Rebonato and Hatano (2018)). Another possible explanation is the mismatch in the frequency with which macroeconomic variables and stock returns are observed. The problem here is that macroeconomic variables are usually only observed at a monthly or quarterly frequency whereas stock returns are typically observed at a daily or intra-daily frequency. To deal with the mismatched frequency between macroeconomic variables and stock returns, Engle et al. (2013) introduce the GARCH-MIDAS model, where macroeconomic variables observed at low frequency are able to be directly incorporated into the high-frequency volatility model.

The GARCH-MIDAS model can be nested into a two-component GARCH-type volatility model initially introduced by Engle and Lee (1999), where stock volatility is decomposed into two components: a long-term volatility component and a short-term volatility component. The short-term volatility component follows a standard GARCH process and evolves around a time-varying long-term volatility component. Engle et al. (2013) introduce the mixed data sampling (MIDAS) approach of Ghysels et al. (2007) into the two-component

GARCH process and create the GARCH-MIDAS model. Under the GARCH-MIDAS framework, macroeconomic data observed at different frequencies are able to be brought up together into one MIDAS filter explaining long-term stock volatility movements, and consequently in turn affect short-term stock variations in the same empirical model. Using the GARCH-MIDAS model, Engle et al. (2013) find that macroeconomic variables partially explain stock volatility in the US market and improve predictive ability relative to other GARCH-type models (see Asgharian et al. (2013) and Conrad and Loch (2015) for further evidence on the relationship between US stock market volatility and macroeconomic information.)

Though much of the evidence to date illustrates the significant relationship between macroeconomic variables and stock market volatility for the US [2], there is little evidence on the effect of macroeconomic variables on stock volatility in other countries and how stock volatility driven by macroeconomic information transmits from the US to other countries. In this paper, we use the GARCH-MIDAS model (we term this model GARCH-MIDAS-X to denote the presence of macroeconomic variables in the model) to examine the relationship between macroeconomic variables and stock market volatility for developed economies, rather than purely for the US market. We are interested in whether relationships documented for the US are observed in other developed economies, in particular the UK and Japan. To examine this, we employ a range of macroeconomic variables that have been widely used in the previous literature, including growth in industrial production; inflation; unemployment rate; term spread, housing starts and exchange rate, to examine whether and how they impact stock market volatility and whether these relationships are stable over time.

[2] see for example Engle and Rangel (2008), Engle et al. (2013), Asgharian et al. (2013) and Conrad and Loch (2015)

We first examine whether macroeconomic information contributes to stock volatility in the GARCH-MIDAS model for the US, UK and Japan. For the US, we reconfirm a counter-cyclical pattern of stock variations response to macroeconomic information. We also find that the response of volatility towards different macroeconomic variables is vary over time, both in magnitude and significance. Industrial production growth, for example, becomes less informative after the Great Moderation of the 1980s, whereas the unemployment rate and the term spread gradually increase their impacts after the Great Moderation. This counter-cyclical volatility pattern in US is consistent with previous studies of Schwert (1989) and Boyd et al. (2001), who argue that large variations in the macroeconomy are expected to promote stock market volatility, whereas large variations of macroeconomy in UK and Japan during the period 1990-2006 is found to depress stock volatility. For instance, stock volatility in the UK experienced unexpected large spikes, which contrasts with the deflationary economy during the Great Moderation. Similarly for Japan, when its economy entered into its deflationary phase after the burst of the real estate bubble, volatility seems to be more volatile than before. One possible explanation is the fundamental change of monetary policy, which might trigger changes in how stock market volatility responds to macroeconomic information. Another possible explanation might be related to market participant's expectation of further change in monetary policy.

We then proceed to examine whether there is any additional macroeconomic effect accounts for stock variation in the UK (or Japanese) stock market, after controlling the volatility spillovers of US. To do this, we estimate two alternative models GARCH-MIDAS-RV-X models (where RV denotes the presence of realized volatility for the US in the model), one that is restricted with no spillover effects and one that is unrestricted with spillover effects. Estimation results reveal that the likelihood ratio test significantly favors the unrestricted model over the restricted one: US volatility spillovers have an effect on stock market volatility in the UK and Japan. Perhaps surprisingly, most macroeconomic variables still perform equally well in term of providing additional impacts on stock markets.

As most of the macroeconomic impacts on stock volatility changes over different sub-samples, both in magnitude and significance, it is difficult to distinguish the dominant macroeconomic variable from alternative competitors. Therefore, we employ principal component analysis (PCA) to derive a "macroeconomic factor" that can be used in one GARCH-MIDAS-X model, avoiding the computational complications and difficulties of including all of the macro variables into one GARCH-MIDAS model. We find that the first principal component (PC_1) of the macroeconomic variables contributes most to explain stock volatility for the full sample across all three countries. However, when looking across different sub-samples for each country, we observe that PC_1 does not always dominate on stock volatility. In fact, its contribution varies across different sub-samples. For the US market, PC_1 gradually increases in its impact, from 21.38% in 1970-1984 to 31.58% in 2007-2014. In contrast, PC_1 for Japan decreases from 25.25% in 1970-1989 to 9.34% in 2007-2014. In terms of additional impact being observed in the GARCH-MIDAS-X model, PC_1 outperforms alternative macroeconomic variables, excluding inflation growth.

Finally, we utilize the Model Confidence Set (MCS), which is suitable for comparisons among a large group of macroeconomic variables across different sub-periods, to evaluate the predictive ability of the GARCH-MIDAS models with different macroeconomic variables. The prediction results reveal that macroeconomic variables seem to be less useful during the global financial crisis (2008-2009) for both the US and UK. Both term spread and PC_1 have superior predictive ability relative to alternative macroeconomic variables in most cases, especially for UK.

The rest of this paper is organized as follows. Section 2.2 provides literature review. Section 2.3 introduces the GARCH-MIDAS model and discusses various specifications of the MIDAS filter. Section 2.4 describes the stock return data and the macroeconomic variables for the US, UK and Japan. Section 2.5 analyses the empirical results, comparing the GARCH-MIDAS models specified with purely realized volatility with those using realized volatility as well as the macroeconomic variables. Section 2.6 evaluates the additional impacts on stock volatility, if any, of US volatility spillovers. In Section 2.7, we employ principal component analysis (PCA) to generate a "macroeconomic factor" to see whether this improves the performance of the GARCH-MIDAS model. In Section 2.8 provides the Model Confidence Set results for prediction ability comparison. Conclusions are presented in section 2.9.

2.2 Literature Review

Prices in financial markets tend to fluctuate when economic news arrives and they might exhibit unusually large fluctuations during major episodes in the macroeconomy. Campbell and Shiller (1988) relate price fluctuations to the variations in macroeconomic state. They show that the dividend-price ratio reflects expectations about future cash flow growth (dividends) and future movements in returns (discount rates) which in turn are part of a larger set of variables that provide information about the state of the macroeconomy. Campbell and Shiller (1988) note that these other variables capturing the state of the economy can be used to forecast returns and cash flows and as such asset prices reflect expectations of the future state of the economy. Fama (1981, 1990) provide empirical evidence to show there exists interactions between the stock market and aggregate economic activities in the US. A large proportion of stock variations are attributed to the forecast of real economic activity, which includes production growth and inflation. Ferson and Harvey (1993) find similar interrelations for other international equity markets.

One strand of the Macro-Finance literature explores the role of financial information, particularly stock returns, in explaining movements in the macroeconomy, investigating whether financial market volatility is useful for predicting future aggregate economic activity, and whether a huge downturn in financial markets in turn predicts recession in the macroeconomy in the future. Harvey (1988, 1989) find that the revision of slope in bond yield curve helps to predict economic growth in the future. Stock and Watson (2003) examine the role of asset prices in forecasting fundamental macroeconomy and observe that it has limited and unstable predictive power for output growth. Ferrara et al. (2014) find out the asset variations have significant predictive power on GDP growth during the Great recession period (2007-2010) over three industrialized countries, the US, the UK and France, while Bellégo and Ferrara (2012) compress financial market information into synthetic factors from a factor model and use it to predict Euro area business cycles, which suggests an lead-lag relationship between synthetic factors and recent one-year economic recessions.

Another strand focuses on evaluating the contribution of the macroeconomy in explaining movements in financial markets. Fama and French (1989) examine the time-varying nature of risk premiums and claim that risk premiums in the stock market are counter-cyclical with respect to different underlying economic states. Schwert (1989) addresses a similar issue for stock volatility, investigating the exogenous sources that drive stock market movements. He employs the volatility of macroeconomic variables to explain aggregate stock market volatility. His findings reveal that stock market volatility exhibits a counter-cyclical pattern with respect to macroeconomic fundamental variables such as inflation and monetary growth. Boyd et al. (2001) and Mele (2007) provide further evidence to support this conclusion.

However, despite the findings in these papers, macroeconomic variables seem to have limited explanatory power in relation to explaining asset volatility. For instance, although Schwert (1989) uses a large set of macroeconomic variables to explain US stock market volatility, none of the variables provide a dominant contribution to explaining stock volatility in the US market: their explanatory power is often limited. More recently, Calvet et al. (2006) arrives at a similar conclusion. One possible explanation for this is that macro variables are more influential in explaining the long-term component of return instead of the entire asset return. If we decompose return into the expected return and noise and if the underlying expected return is persistent (see Ferson et al. (2003), for example), then from the evidence in Chen et al. (1986) that macroeconomic variables explain expected returns, the link between macro variables and expected returns can be viewed as a kind of macro-finance relationship in the long-run.

Similar points hold for long-term trend of volatility when exploring the contribution of macroeconomic variables. Engle and Rangel (2008) isolate long-term volatility through a Spline-GARCH model, then observe a significant relationship between macroeconomic variables and long-term volatility in the US stock market. The way in which Engle and Rangel (2008) isolate long-term volatility is inspired by Andersen and Bollerslev (1998) and Engle and Lee (1999). As noted by Andersen and Bollerslev (1998), in a conventional GARCH model where we can write the return as $r_t = \sigma_t z_t$, the conditional variance reverts to its mean value of σ^2 but it is difficult to model this potentially slow-moving component of return volatility because z_t is noisy. Engle and Lee (1999) suggest a way to do this by decomposing the unconditional volatility σ^2 into a two-component structure: a permanent (slow-moving) component which is interpreted as trend around which the other transitory component fluctuates. Engle and Lee (1999) provide economic interpretations for these two components: the transitory component accounts for short-lived macroeconomic shocks as well as such things as leverage effects that are observed in stock market volatility; the permanent component is interpreted as investors' response to the risk level for the longer-term. Since the transitory component only accounts for instant shocks which die out quickly, the permanent component is closely related to risk premia and can then be priced.

These two seminal works offer insight and some understanding of Engle and Rangel (2008)'s main results in terms of the finding of a significant relationship between macroeconomic variables and the long-term (permanent) volatility component. That is, long-term volatility component viewed as a proxy of long-term risk, is priced in asset returns. As long-lasting risk can be easily affected by underlying economic states, it is reasonable to consider possible linkages between the long-term volatility component and macroeconomic variables. Engle et al. (2013) further extend this model by adding a MIDAS filter into the two-component structure, the GARCH-MIDAS model. The main advantage of the GARCH-MIDAS model is that macroeconomic data with different frequencies are able to be incorporated into the model within one MIDAS filter, allowing macroeconomic variables to potentially explain the permanent volatility component. Their empirical results demonstrate that macroeconomic variables help to improve stock volatility forecasting accuracy in the long-run. In the following chapter, we will introduce the GARCH-MIDAS model in details.

2.3 The GARCH-MIDAS model

In this section, we introduce the GARCH-MIDAS model with three alternative specifications: the realized volatility (MIDAS-RV) model, the macroeconomic variables (MIDAS-X) model and the mixture of RV together with macroeconomic variables (MIDAS-RV-X). Referring back to Engle et al. (2013), the long-term volatility component possesses a high degree of persistence relative to the short-term volatility component. Monthly RV could be one source that drives long-term volatility movements within the MIDAS structure. This leads to the MIDAS-RV specification. In the MIDAS-X specification, macroeconomic variables are used to explain long-term volatility movements in the stock market. Therefore, we are able to directly evaluate the relationship between macroeconomic information and long-term stock market volatility using the MIDAS-X specification. We can also combine RV and macroeconomic variables together and create a new MIDAS-RV-X specification, in order to exmaine to what extent macroeconomic variables provide complementary information to that contained in RV.

2.3.1 GARCH-MIDAS Model with Realized Volatility

The basic GARCH-MIDAS model specified with Realized Volatility (RV) is:

$$r_{i,t} - E_{i-1,t}(r_{i,t}) = \sqrt{\tau_t g_{i,t}}\epsilon_{i,t} \qquad i = 1, 2, ...N_t \tag{2.1}$$

$$\epsilon_{i,t}/I_{i-1,t} \sim N(0,1)$$

where $r_{i,t}$ is daily log return observed on date i in a period of time t, where i indexes the number of trading days within one period of time t, $i = 1, 2..., N_t$. For example, if we were to specify the time interval t to be monthly, N_t would be the total number of trading days in one month t. On the right-hand side, the entire conditional volatility is decomposed into two sub-components, a long-term component τ_t and a short-term component $g_{i,t}$. According to Engle et al. (2013), the short-term component $g_{i,t}$ follows a typical, but more distinctive GARCH(1,1) process:

$$\begin{gathered} g_{i,t} = (1-\alpha-\beta) + \alpha\frac{(r_{i-1,t}-\mu)^2}{\tau_t} + \beta g_{i-1,t} \\ \alpha > 0, \beta > 0,\ \alpha+\beta < 1 \end{gathered} \tag{2.2}$$

where the short-term volatility component $g_{i,t}$ on day i of month t is explained by its previous day's value of $g_{i-1,t}$ and the squared error term $(r_{i-1,t}-\mu)^2$. One distinguishing feature in Equation 2.2 is that the squared error term has been adjusted by its long-term component τ_t, where τ_t keeps a relatively constant value during that period of time t. Please note that, the long-term component could be one month, quarter or even a year. In this thesis, we will specify the long-term component τ_t in a monthly frequency. To understand why Engle et al. (2013) make such an adjustment, multiply both sides of Equation 2.2 by τ_t on both sides, and rearrange to give:

$$g_{i,t}\ \tau_t = (1-\alpha-\beta)\tau_t + \alpha(r_{i-1,t}-\mu)^2 + \beta g_{i-1,t}\ \tau_t \tag{2.3}$$

where the intercept term in Equation 2.3 is allowed to evolve slowly with the long-term component, rather than keeping a constant level. Intuitively, maintaining a relative constant value during, say, a one month period of time enables the long-term component τ_t to accommodate to the underlying economic circumstances during that month. When the economic state switches, the long-term component τ_t will change accordingly. Then, from Equation 2.3, we can infer that the short-term component $g_{i,t}$ actually swings around the long-term component τ_t during month t with available information up to date

$i-1$ in the information set $I_{i-1,t}$.

Following Ghysels et al. (2007) and Engle et al. (2013), the long-term volatility component τ_t is explained by a weighted average value of past realized volatilities in the MIDAS-RV specification:

$$\begin{aligned} \tau_t &= m + \theta \sum_{k=1}^{K} \phi_k(\omega_1, \omega_2) RV_{t-k} \\ RV_t &= \sqrt{\sum_{i=1}^{N_t} r_{i,t}^2} \end{aligned} \tag{2.4}$$

In Equation 2.4, parameter θ measures the impact of aggregated historical realized volatilities on the long-term volatility component τ_t. K denotes the optimal number of lags of past realized volatilities (RV) being added into the MIDAS-RV filter. The Realized Volatility (RV) is a proxy of unobserved stock volatility, which is measured by taking the square root of the sum of squared returns within a fixed time span t. As an alternative to the fixed RV in Equation 2.4, RV can also be calculated with a rolling window, which means realized volatility is recursively calculated:

$$RV_{i,t}^{rw} = \sqrt{\sum_{j=1}^{N'} r_{i-j,t}^2} \tag{2.5}$$

In Equation 2.5, N' denotes the length of the rolling window. On day i, we roll back N' days from day i in period t, then sum up squared returns during that rolling window. In this paper, we set up N' equal to 22. Therefore, RV^{rw} denotes a monthly rolling RV calculated in a daily frequency. Correspondingly, with RV^{rw}, both the long-term component τ and the short-term component g are able to be observed in a daily frequency.

In the MIDAS-RV specification, ϕ_k is a beta weighting scheme that determines the weight attached to each lag of RV and therefore determines the impact each lag of RV has on the long-term component of volatility τ. The beta weight is defined as:

$$\phi_k(\omega_1, \omega_2) = \frac{(k/K)^{\omega_1 - 1}(1 - k/K)^{\omega_2 - 1}}{\sum_{j=1}^{K}(j/K)^{\omega_1 - 1}(1 - j/K)^{\omega_2 - 1}} \tag{2.6}$$

An advantage of the beta weighting scheme is that it governs the shape of the weighting scheme quite flexibly with only two parameters, ω_1 and ω_2. The beta weights being generated from Equation 2.6 satisfy the restrictions that $\phi_k > 0$ and $\sum_{k=1}^{K} \phi_k = 1$. If ω_1 is set equal to 1 the beta weighting scheme exhibits a decaying pattern. In Figure 2.1, the decaying pattern is shown as the black dashed line and it can be seen that with a decaying pattern the most recent observations tend to exert a greater impact on the long-term component τ_t but the impact weakens over time. The decaying speed is captured by ω_2. If the value of ω_2 increases, the decay is more rapid (see the black solid line in Figure 2.1). If ω_1 and ω_2, are unrestricted then a hump shape weighting scheme will result. Such a scheme is shown by the red line in Figure 2.1, where the maximum weight is allocated to a more distant observation rather than most recent observation.

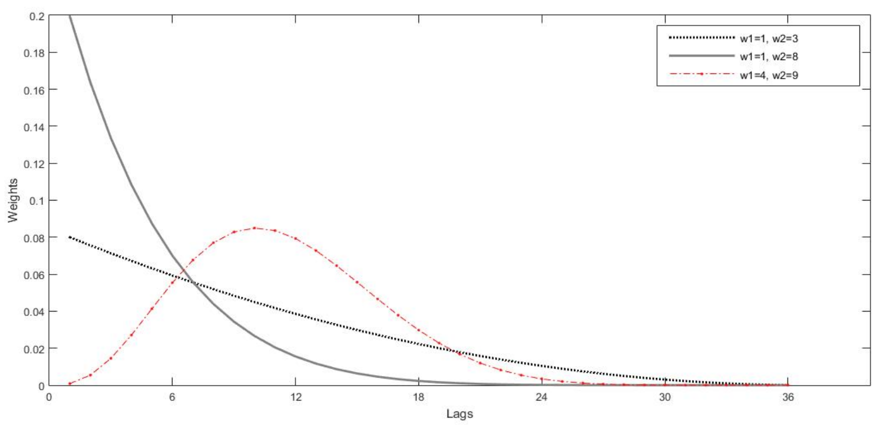


Figure 2.1: Beta Weighting Scheme

Equations 2.1, 2.2, 2.4 and 2.6 jointly constitute the GARCH-MIDAS model with Realized Volatility (GARCH-MIDAS-RV) specification. The parameter space is defined as $\Phi = [\mu, \alpha, \beta, m, \omega_1, \omega_2]$, which is estimated via the maximum likelihood algorithm. The log-likelihood function is:

$$LLF = -\frac{1}{2}\sum_{t=1}^{T}\{log\,(g_t(\Phi)\tau_t(\Phi)) - \frac{(r_t - \mu)^2}{g_t(\Phi)\tau_t(\Phi)}\} \tag{2.7}$$

where $g_t(\Phi)\tau_t(\Phi)$ denotes the conditional variance in the GARCH-MIDAS model.

2.3.2 GARCH-MIDAS Model with Macroeconomic Variables

As an alternative to the realized volatility specification, we can incorporate macroeconomic information (call this X) into the MIDAS filter, leading to the MIDAS-X specification. The MIDAS-X specification allows us to evaluate the relationship between macroeconomic information and stock volatility. The MIDAS-X specification is described below:

$$log\tau_t = m_l + \theta_l \sum_{k=1}^{K_l} \phi_k(\omega_{1,l}, \omega_{2,l}) X_{l,t-k} \tag{2.8}$$

The long-term volatility τ_t in Equation 2.8 is smoothed out by K lags of monthly observations from one macroeconomic variable X. Research associated with the GARCH-MIDAS model has shown that most macroeconomic variables exert significant impacts on stock volatility in the U.S. For example, Engle et al. (2013) observe significant effects of industrial production (IP) and the producer price index (PPI) on US stock volatility. Conrad and Loch (2015) provided encouraging results for the term spread (TS) and housing starts (HS) in the US market as well. Their findings are consistent with Campbell (1999), who documents that the stock volatility could be a result of the impact of macroeconomic factors that vary over different economic states. Not only focusing on the US market, in this paper we examine whether the effects of macroeconomic information on stock volatility observed in the U.S are present in the UK and Japan stock markets.

2.3.3 Measuring the Contribution of Macroeconomic Information

A natural question to ask at this stage is, "How much of stock volatility can be attributed to macroeconomic sources?" To answer this question, we implement two alternative approaches: the first is to examine the incremental contribution, if any, of macroeconomic variables to stock market volatility having controlled for the effects of historical realized volatility; the second is to use the variance ratio introduced by Engle et al. (2013), quantifying the relative contribution of the macroeconomic variables to the entire stock volatility.

A. The GARCH-MIDAS Model with RV and Macroeconomic Variables

To evaluate the additional impacts of macroeconomic variables on stock market volatility, we follow Asgharian et al. (2013) and Conrad and Loch (2015) that combine both realized volatility and macroeconomic variables together within one GARCH-MIDAS model. We term this the GARCH-MIDAS-RV-X model. Hence in the MIDAS-RV-X filter, both macroeconomic information and historical realized volatility jointly account for the stock volatility:

$$\tau_t = m + \theta_{rv} \sum_{k=1}^{K_{rv}} \phi_{rv}(\omega_{1,rv}, \omega_{2,rv}) RV_{t-k} + \theta_x \sum_{k=1}^{K_x} \phi_x(\omega_{1,x}, \omega_{2,x}) X_{t-k} \tag{2.9}$$

where X_t is the macroeconomic variable and captures the additional macroeconomic information that is supplementary to that provided by past realized volatility, supplementary due to the fact that past realized volatilities might already contain information from the macroeconomic variable. Accordingly, θ_x accounts for the additional effect of the macroeconomic variable on long-term volatility τ_t. Under this MIDAS-RV-X specification, past realized volatility (RV) can also be viewed as a benchmark relative to different macroeconomic variables: if θ_x is insignificant, the macroeconomic variable contains no additional information about stock market volatility to that contained in realized volatility.

B. Variance Ratio

Instead of using MIDAS-RV-X filter (2.9) to evaluate the additional impact of macroeconomic variables, we implement an alternative method of The Variance Ratio from Engle et al. (2013). The variance ratio is defined as the fraction of the sample variance of the log long-term volatility component $log(\tau_t^M)$ relative to the entire volatility $log(\tau_t^M g_t^M)$:

$$VR = \frac{\widehat{var} log \tau_t^M}{\widehat{var} log(\tau_t^M g_t^M)} \tag{2.10}$$

where M refers to a typical GARCH-MIDAS model, in which long-term component could be driven by one specified macroeconomic variable x. In the numerator of Equation 2.10, the long-term volatility component τ_t^M is solely explained by one macroeconomic variable x in a MIDAS filter (see Equation 2.8 as an example). In the denominator of Equation 2.10, we combine the long-term component τ_t^M and the short-term component g_t^M together as the total volatility. Therefore, we can see how large a proportion of stock variations is attributed to the variations of macroeconomic variable .

2.3.4 GARCH-MIDAS model with the Principal Component

Previous studies suggest a wide range of macroeconomic variables that have significant impacts on the U.S stock market. For instance, Conrad and Loch (2015) find the term spread and housing starts are significant for the U.S market, Asgharian et al. (2013) find the short-term interest rate performs well in the U.S markets while Engle et al. (2013) find industrial production growth and inflation matter in explaining volatility. One problem these findings raise is that they might result in confusion over which macroeconomic variable contributes most for stock volatility. In fact, Campbell (1999) document that variation in the stock market could be a result of different macroeconomic impacts that vary over different economics states. If this is the case, it is necessary to use a combined macroeconomic indicator to evaluate the dynamic relationship between the macroeconomy and stock market volatility. Asgharian et al. (2013) use Principal Components Analysis (PCA) to address this issue and introduce the principal components into the GARCH-MIDAS model.

PCA offers us an efficient way to bring all the macro variables together into one MIDAS filter, avoiding the significant computational complications and possible convergence problems that arise when they are included in the MIDAS filter as a group. PCA is a dimensionality reduction approach, which constructs principal components to extract combined macroeconomic information from a group of macro variables. The first principal component (PC_1) accounts for the largest proportion of variation from the different macro series. Therefore, we compress a large group of macro variables into one common factor, PC_1, and add it into the MIDAS regression:

$$\tau_t = m_p + \theta_p \sum_{k=1}^{K_p} \phi_k(\omega_{1,p}, \omega_{2,p}) PC_{1,t-k} \tag{2.11}$$

where PC_1 refers to the first principal component. With the MIDAS-PC specification in Equation 2.11, we are able to evaluate combined macroeconomic impacts on stock market volatility efficiently. In line with Section A., we also combine realized volatility and the principal component together into one MIDAS regression:

$$\tau_t = m + \theta_{rv} \sum_{k=1}^{K_{rv}} \phi_k(\omega_{1,rv}, \omega_{2,rv}) RV_{t-k} + \theta_{pc} \sum_{k=1}^{K_{p}c} \phi_k(\omega_{1,pc}, \omega_{2,x}) PC_{t-k} \tag{2.12}$$

We use Equation 2.12 to evaluate the performance of principal components relative to the benchmark of realized volatility. Also, we will make comparisons with different MIDAS-RV-X specifications to identify whether individual macro variables or the principal component have the most prominent effect on stock market volatility.

2.4 Data Description and Summary Statistics

To fully capture the dynamic relationship between macroeconomic information and stock volatility across the different countries in our study, we employ a large set of macroeconomic variables, including industrial production (IP), housing starts (HS), producer price index (PPI), unemployment rate (UEM) and term spread (TS). Both industrial production and housing starts are used as proxies of real activity. The change in the PPI measures average growth in the selling price offered by producers who charge for goods and services. Growth in the PPI Index is the economic indicator for inflation in our study. Following Estrella and Mishkin (1998), we use the term spread, measured as a difference between the ten-year and three-month Treasury yields, and evaluate its impacts on stock volatility, especially during economic recessions. Given the evidence that there are significant spillover effects from the US to Japan (see Ng (2000)), we also include the Japanese Yen Index for Japan. All the macroeconomic data is seasonally adjusted and gathered from Federal Reserve Economic database (FRED) and Datastream. The macroeconomic variables are observed at a monthly frequency and our sample period is January 1970 to December 2014. Excluding the term spread, each macroeconomic variable is converted into a compounded annualized growth rate:

$$\Delta X = ((X_t/X_{t-1})^{12} - 1) \tag{2.13}$$

where ΔX refers to the compounded annualized growth rate of the macroeconomic variable X. Daily log returns on the S&P500 Index (US) and the FTSE All share Index (UK) are collected from WRDS, starting from 02/01/1970 to 31/12/2014. The log return on the Nikkei 225 Index (Japan) is collected from the Global Financial Database.

As mentioned by Liu et al. (2015), RV calculated from 5-minute tick-by-tick data outperforms more sophisticated realized measures in terms of volatility prediction. Therefore, we use 5-minute RV estimated via the two-scale estimator of Zhang et al. (2005) from Oxford-Man Institute Realized library. RV for the S&P 500 Index, FTSE All Share Index and the Nikkei 225 Index are available from 02/01/1970 to 31/12/2014 in a daily frequency.

To allow for potential structural breaks, we partition the whole sample (01/1970 to 12/2014) into three sub-samples for each country. As noted by Schwert (1989) and Stock and Watson (2002), a significant structural break occurred around 1984 in the US, associated with the Great Moderation. Concurrently, the UK started receiving positive feedback from the economy after the implementation of Thatcher's deflationary policy at the beginning of the 1980s. Hence the first split in both the US and UK is around 1984. The first partition in Japan happens around 1989, a known turning point for the Japanese economy (more details will be provided in Section 2.5.3). The second partition of the sample is around 2006 for all three countries. This is associated with the Subprime Mortgage crisis that originated in the US and resulted in a global financial crisis. To summarize, data for the US and UK are divided into three sub-samples: 01/1970-12/1984, 01/1985-12/2006 and 01/2007-12/2014 while the data for Japan is divided into the three sub-samples 01/1970-12/1989, 01/1990-12/2006, 01/2007-12/2014. Table 2.1 reports descriptive statistics for all the data, including mean, standard deviation, skewness and kurtosis.

[**Insert Table 2.1 here**]

Figure 2.2 through 2.4 plot monthly realized volatility and the macroeconomic series for the US, UK and Japan. With the exception of the term spread, most of the time large peaks and troughs in the growth rates of the macroeconomic variables coincide with the local crises in each country as well as the global financial crisis (2007–2009). The term spread appears to behave as a leading economic indicator that signals recessions. As shown in Figures 2.2 through 2.4, substantial declines in the term spread appear to precede recessions (as shown by the shaded areas) across different financial markets. The macroeconomic data in the UK (Figure 2.3) seems to be more volatile before 1980s and then becomes fairly constant up until the global financial crisis. When we take a close look at Japanese market in Figure 2.4, we find stock volatility was relatively stable around 1975-1989, which is associated with the excellent performance of Japanese economy. However, once the asset price bubble collapsed at the end of 1989, Japan's economy abruptly stopped fast growth and stepped into a downturn. Both industrial production growth and inflation became quite constant during 1990-2005, which has been widely recognized as long period of stagnation in the Japanese economy. Bearing this in mind, we undertake an empirical analysis of the relationship between macroeconomic variables and stock market volatility to examine how those changes in financial markets are explained, if at all, by changes in macroeconomic fundamentals.

[**Insert Figures 2.2 through 2.4 here**]

2.5 Empirical Analysis for GARCH-MIDAS model

In this section, we conduct a cross-country study on stock volatility for the US, UK and Japanese stock markets. We initially incorporate monthly RV into the GARCH-MIDAS model, namely as the GARCH-MIDAS-RV model. We use this GARCH-MIDAS-RV model to evaluate the explanatory power of historical stock volatility (approximated by RV) on current long-term stock volatility. We then employ the GARCH-MIDAS-X framework (see Equation 2.8), where macroeconomic information accounts for long-term stock volatility, not only for the full sample but also for the sub-samples. Quite apart from that, we measure the relative performance of the macroeconomic variables via two approaches: one is through examining the impact of the macroeconomic variables on volatility having controlled for historical volatility; the other is through the variance ratio introduced by Engle et al. (2013), quantifying the relative contribution of the macroeconomic variables to stock market volatility. In the MIDAS filter, we allow for two years (24 months; K=24) of macro observations.[3]. Following Engle et al. (2013), we set up a restriction that ω_1 equals to one in the Beta weighting scheme (see Equation 2.6), which provides a decaying weighting curve, where lower weights are assigned to more distant lags

[3]The optimal number of lags K in the MIDAS filter is determined by the maximum Log-likelihood estimation results. The MIDAS filter with 24 monthly lags yields the highest maximum likelihood value for all sub-samples.

of the macro variables, reasonably indicating that the macro variable becomes less informative the further back the observation.

2.5.1 Estimation Results in the GARCH-MIDAS model with Realised Volatility

First, we need to evaluate the explanatory power of historical stock volatility for current stock volatility. To do this, we incorporate only historical RV into the MIDAS filter and examine whether it is significant in explaining the long-term volatility component τ (see Equation 2.4). In a stationary GARCH-MIDAS process, the short-term volatility component g reverts back to the long-term component τ over a long time horizon. The reversion speed is measured by the sum of α and β. As shown in Table 2.2, the sum of the parameters α and β is close to one, which allows us to infer there is a high degree of persistence in stock volatility in all the countries in our sample. The parameter θ measures the impact of historical RV on the current long-term volatility component τ. We find a significant θ with positive sign across all sub-samples, ranging from 0.0183 up to 0.0494, which indicates that lags of RV have a significant influence on current volatility. Looking at the relative performance of RV in the different countries in our sample, in Figure 2.5 we observe that the impact of historical RV on current long-term volatility decays quite fast with the exception of Japan during 1970-1984. In general, most weights are attached to the recent lags of monthly RV (K=1,2,3) for US, UK and Japan. As distance increases, weights sharply decrease and completely vanish after 10 months (K=10). Overall the results indicate that historical realized volatility has a significant short-term impact in the US, UK and Japan. The next question we evaluate is whether there is any significant impact from macroeconomic information on the long-term volatility component τ.

[Insert Table 2.2 here]

[Insert Figure 2.5 here]

2.5.2 Estimation Results in the GARCH-MIDAS model with Macroeconomic Variable for US

For the US market, estimation results for the GARCH-MIDAS-X models are summarized in Tables 2.3 and 2.4. In general, the full-sample (1970-2014) results demonstrate a counter-cyclical pattern between movements in the macroeconomic variables and long-term stock volatility, which is in line with the main findings in Schwarz et al. (1978), Campbell (1999), Engle et al. (2013) and Conrad and Loch (2015). During 1970-2014, inflation and unemployment are positively related with stock market volatility at the 10% and 1% levels of significance respectively. Meanwhile, growth rates of Industrial Production and Housing Starts as well as the term spread are negatively related with stock variations at the 1% significance level. Table 2.5 contains the results for the Variance Ratios. We can see from Table 2.5 that the term spread accounts for 17.49% of stock volatility in the full sample, and it reaches its highest level of 51.09% during the 1985-2006 sub-sample. This is almost twice higher than the other macroeconomic variables during that period. As noted by Harvey (1989, 1991) and Stock and Watson (2003), we find that most peaks in long-term volatility that is driven by term spread (solid magenta line) are followed by peaks being observed in total conditional volatility (dash green line) (see in Figure 2.6). This phenomenon infers that the term spread, as a leading indicator, has better performance in predicting financial crisis than alternative macroeconomic variables.

[Insert Tables 2.3 through 2.5 here]

[Insert Figure 2.6 here]

Looking across all the sub-samples, we observe the linkages between the macroeconomic variables and stock volatility vary across the sub-samples with respect to different underlying conditions in the economy. Both industrial production and housing starts have larger impacts on stock volatility prior to the Great Moderation. After the Great Moderation, industrial production and housing starts become less volatile and their impacts on volatility tends to weaken. Taking industrial production as an example, θ_{IP} observed in Table 2.3 dramatically decreases from -0.9540 in 1970-1984 to -0.0914 in 2007-2014. Similar results can be found in Table 2.5, where the variance ratio of ΔIP is 15.24% during 1970-1984. The contribution of ΔIP to total stock volatility then begins to decrease and continues to decrease. During 2007-2014, it merely accounts for 1.98% of the variation in stock volatility. The decreasing influence of ΔIP actually coincides with the arguments in Schwarz et al. (1978),Blanchard and Simon (2001) and Stock and Watson (2002) that the explanatory power of output growth tends to be weaker after the Great Moderation. This can be seen graphically in Figure 2.7 where prior to the Great Moderation (the top panel) the peaks and troughs of the long-term volatility component τ (the dashed blue line) often coincide with those in ΔIP (the solid magenta line). This does not appear to be the case post the Great Moderation (the bottom panel.) Meanwhile, the contribution of unemployment seems to gradually increase after 1984. The variance ratio of ΔUEM increases from 12.87% to 17.81% across the three sub-samples.

[Insert Figure 2.7 here]

Strikingly, we observe a significant change in the direction of the impact of inflation on stock volatility. In Table 2.3 a positive relationship between inflation and stock volatility can be found during the 1970-1984 and 2007-2014 samples. However, in contrast, the relationship between inflation and stock volatility during the 1985-2006 is negative. The observed change of sign in the relationship between stock volatility and inflation possibly reflects a fundamental change in the economy. Around 1984, the U.S economy ended its boom and bust cycle and stepped into a period with low inflation and positive economic growth, known as the Great Moderation. At the same time, a reduction of instability in monetary policy was also observed. A negative relation during 1985-2006 might also possibly be explained by the inflation illusion hypothesis of Modigliani and Cohn (1979). The hypothesis explains that stock values are undervalued when inflation is high and become overvalued when inflation is low. Consequently, the risk premium is negatively related to inflation.

So far, our results suggest that, when modelled separately, both realized volatility and macroeconomic variables are significant in explaining volatility. We now examine whether the macroeconomic variables contain information above that contained in realized volatility. To do this we use the GARCH-MIDAS-RV-X model, where both realized volatility and macroeconomic variables are incorporated into one MIDAS filter within the GARCH-MIDAS framework (see in Equation 2.9). Results from estimating the GARCH-MIDAS-RV-X model are reported in Table 2.6. θ_{RV} measures the impact of historical RV on current volatility and θ_X measures the impact of the individual macroeconomic variables on volatility. We can see that with the inclusion of the macroeconomic variables, RV still has a positive and statistically significant (at the 1% level) impact on stock volatility, with θ_{rv} ranging from 0.27 (2007-2014) to 0.04 (1985-2006). In terms of the different macroeconomic variables, their significance varies across the different sub samples, with none of the macroeconomic variables having a significant effect on volatility in the 2007-

2014 sub-sample. Within this, the term spread tends to quite consistently have a significant effect on volatility. In the Full sample, θ_{TS} is -0.0585 and significant at the 1% level. Since the weighting function with $\omega_{TS} = 1.0347$ puts 0.0438 on the first lag (the maximum weight) and 0.038 on last lag (at the end of 2nd year), we observe that 1% increase in the term spread during the current month would reduce next month's volatility by $e^{0.0585*0.0438} - 1 \simeq 0.26\%$. This negative impact decays quite slowly, which means that the past 2-year's term spread will still have a significant impact on next month's stock volatility. The results overall suggest that both RV and the macroeconomic variables together significantly explain stock volatility. However, the results also indicate that the significance or otherwise of the macroeconomic variables is dependent on the sub-sample, suggesting that the relationship between macroeconomic variables and volatility changes over time, both in magnitude and significance. This is something we will return to later.

[Insert Table 2.6 here]

2.5.3 Estimation Results in Japan

As noted by Schnabl (2015) and Schnabl and Hoffmann (2008), Japanese stock market experienced boom-and-bust cycle during 1970-2014. Before 1989, stock market and real economy were characterized with boom and start to have bubble around 1985. The economy bubble was largely attributed to the loose monetary and fiscal policy as well as the trading conflict between Japan and US. To resolve trading conflict, US and Japan reached up the Plaza Aggrement in 1985, whereas it adversely leaded to Yen appreciation that far beyond targeted range. Accordingly, central bank tried to reduce appreciation pressure by cutting the short-term interest rate, which lead to a further expansion of economy bubbles. Consistently, we observe a sharp decline of term spread between 1985 and 1987 in Figure 2.4. Later afterward, the Bank of Japan tighten its monetary policy which led to bubble burst in stock and real estate markets. Hence stock market stepped into a prolonged recession.

[Insert Figure 2.4 here]

Tables 2.7-2.8 summarize estimation results for the GARCH-MIDAS-X models for Japan. Generally speaking, industrial production growth, growth in housing starts, the term spread and the Yen index are counter-cyclical, with stock volatility falling as these variables grow (excluding Housing starts in the 1970-1989 sub-sample) while inflation and unemployment are negatively related to stock market volatility over the full sample (1970-2014).

[Insert Tables 2.7 through 2.8 here]

To try and understand the negative relationship between inflation (or unemployment rate) and volatility for the full sample, we look at their sub-sample performances in Tables 2.7 and 2.8 . McQueen and Roley (1993) and Boyd et al. (2005) argue that high inflation (or unemployment rate) can be regarded as a predictor of future changes in monetary policy and corporate profits, which differ in boom and bust periods. During the boom phase (1970-1989), inflation and the unemployment rate have significant (1% level) and negative impact on stock volatility. This negative impact could be a result of market participants' expectation of a tightening of fiscal and monetary policy in the future and the corresponding negative effect this would have on the stock market and hence volatility. During the bust phase (1990-2006), we observe a positive impact of inflation and unemployment on volatility, which implies that market participants might have changed their expectations about future fiscal and monetary policy from tightening to loosening. Comparing these two sub-samples (1970-1989, 1990-2006), we can see that the impact of inflation on volatility decays slowly (it exerts a longer-term impact) during the boom phase ($\omega_{PPI} = 1.0015$), whereas the effect of inflation on volatility dies out quickly during the bust phase ($\omega_{PPI} = 19.5730$). These results suggest that inflation has a more persistent effect on stock volatility during a boom phase than in a bust phase.

Similar to the US, we observe a negative and significant relationship between industrial production and stock volatility in Japan. As can be seen in table 2.7, θ_{IP} takes negative and significant values for all sub-samples, ranging from -0.8090 to -0.5046. θ_{IP} across the sub-samples displays a bell-shape pattern, where it increases from 0.5675 in 1970-1989 to 0.6422 in 1990-2006, then decreases to 0.5046 in 2007-2014. This implies that a one unit increase in the previous month's IP ($\theta_{IP} = -0.5675$) lead to a 2.86% decrease in volatility during the period 1970-1989 and a 2.7% decrease during 1990-2006 ($\theta_{IP} = -0.6422$). Consistent with this, the variance ratio for IP has its maximum value at 11.3% in 1990-2006, relative to 6.65% in 1970-1989 and 6.28% in 2007-2014 (see Table 2.5).

Now we turn to results from estimating the GARCH-MIDAS-RV-X model for Japan. Results from this model are reported in Table 2.9. For the full sample, as for the US, we find that even with the inclusion of the macroeconomic variables, RV still has a positive and statistically significant (at the 1% level) impact on stock volatility. In terms of the significance of the macroeconomic variables for the full sample, we find that with the exception of housing starts, the individual macroeconomic variables still contribute significantly to explaining volatility, that is, they contain information relevant for volatility over that contained in historical realized volatility. However, the sub-sample results show that while RV is always significant, different macroeconomic variables are significant across different sub-samples, with some of the macroeconomic variables experiencing a change in sign, as we found for the GARCH-MIDAS-X models in Tables 2.7 and 2.8.

[Insert Table 2.9 here]

2.5.4 Estimation Results in the GARCH-MIDAS model with Macroeconomic Variable for UK

Figure 2.3 plots movements in the macroeconomic variables for the UK. Both inflation and unemployment growth reach their peaks around 1975-1980, then take their minimums around 1985-1990, which is the time that UK economy enters into a Great Moderation and characterized by low inflation and positive industrial production growth. The Great Moderation is largely attributed to the successful deflationary monetary policy the UK followed. Despite the low inflation, stock market volatility experienced large swings during the Great Moderation. This raises the question, "Why has stability in the macroeconomy not been followed by stability and even a decline in stock volatility?" To address this, we employ the GARCH-MIDAS-X model to, as we did for the US and Japan, examine the impact of individual macroeconomic variables on UK stock market volatility, with particular emphasis on the Great Moderation episode. From the results in Tables 2.11 and 2.12, many of the macroeconomic variables individually have a significant counter-cyclical relationship with stock volatility for the full sample, which is consistent with the full-sample results for the US. Looking across the sub-samples, we observe significant changes in the magnitude of the impact from alternative macroeconomic variables to stock volatility around the turning point of the Great Moderation. Industrial production, for example, has a negative and significant impact ($\theta = -0.4345$) on stock volatility before the Great Moderation in 1970-1984. During the Great moderation, its magnitude dramatically decreases to -0.0454.

[Insert Tables 2.11 through 2.12 here]

We observe that both inflation and unemployment rate experience changes in sign, going from positive to negative during the Great Moderation episode. The unemployment rate has a positive impact on volatility ($\theta = 0.0062$) during 1970-1984. Then it evolves into an insignificant negative one during 1985-2006. This sign change of unemployment is consistent with Boyd et al. (2005)'s main findings. Boyd et al. (2005) state that increasing in unemployment rate is good news for stock market stability during economy expansions (therefore negative macro-volatility relation is expected), and bad news for stock market stability during economy recessions (therefore positive macro-volatility relation is expected).

In terms of inflation, Schwert (1989), Boyd et al. (2001) and Kontonikas and Ioannidis (2005) document that an inflation targeting regime is expected to promote financial market stability (positive relation is expected). However, in our study, inflation has a negative effect during the Great Moderation (1985-2006), which seems to be contrary to their arguments. Borio and Lowe (2002) provide a possible explanation for such a negative relationship during an expansion (1985-2006). Borio and Lowe (2002) argue that low inflation is not a sufficient condition for stock market stability. In fact, lack of inflation pressure might potentially remove the threat of an interest rate increase from the financial market. Consequently, low inflation might result in stock market booms. The foreign exchange rate (USD to British Pound) has a significant effect on stock volatility during 1985-2006. This might be suggestive of a spillover effect from US stock market volatility to the UK market. We turn our attention to investigating spillover effects in the following section.

2.6 Spillover Effect and Macroeconomic Contribution

To have a deeper understanding of the macroeconomic drivers of stock volatility, it is important to take financial integration into consideration. Since the global stock market crash in 1987, a number of studies document significant volatility spillovers across international equity markets. Some empirical studies suggest that both the UK and Japan stock markets are tightly linked with the US stock market. Hamao et al. (1990) investigate the return and volatility spillovers among the US, UK and Japanese markets. They observe significant volatility spillovers originating from the US to the Japanese market and conclude international transmissions reflect fundamental changes in the global macroeconomy. Following on from this study, Lin et al. (1994) suggest that volatility spillovers between US and Japan stock markets are bi-directional. Susmel and Engle (1994), however, observe weak and less persistent volatility spillovers that come from the US stock market into the UK. Hence, we bring global volatility (the US) and local volatility (UK or Japan) into one MIDAS filter evaluating the spillover effects from the US on local stock markets across all sub-samples:

$$\tau_t = m + \theta_L \sum_{k=1}^{K} \phi_L(\omega_L) RV_{L,t-k} + \theta_{US} \sum_{k=1}^{K} \phi_{US}(\omega_{US}) RV_{US,t-k} \tag{2.14}$$

where RV_{US} refers to the historical realized volatility that comes from the US stock market. We calculate weighted average value of RV_{US} as a proxy for the global volatility transmission onto the local long-term volatility component τ_t. RV_L refers to the historical realized volatility from the local stock market (either UK or Japan market). We also calculate the average value of RV_L

to explain the long-term volatility component τ_t. The estimation results are reported in Table 2.15.

[Insert Table 2.15 here]

For the full sample (see in Table 2.15), the parameter θ_{US}^{RV} that measures the spillover effect has a negative impact of -0.0043 on the Japanese market and a positive impact of 0.0041 on the UK market. In terms of sub-samples, US volatility spillovers had a considerably larger impact ($\theta_{US}^{RV} = 0.0079$) during 1985-2006 in UK, which implies that stock volatility in the UK has been increasingly driven by global volatility since the Great Moderation. Interestingly, for Japan we observe a change in the sign of US volatility spillovers after the real estate bubble burst: Before 1989, US volatility spillovers have a negative impact on stock market volatility in Japan; After 1989, this turns into a positive impact. All above results are consistent with Baele (2005): the effect (in terms of magnitude and sign) of US volatility spillovers on UK and Japanese stock market volatilities are regime-dependent (sub-sample dependent).

Given our results concerning spillover effects from the US, we augment the GARCH-MIDAS-RV-X model in Equation 2.9 by bringing local volatility, individual macroeconomic variables and US volatility spillovers (global volatility) together into one MIDAS filter:

$$\tau_t = m + \theta_L \sum_{k=1}^{K} \phi_L(\omega_L) RV_{L,t-k} + \theta_{US} \sum_{k=1}^{K} \phi_{US}(\omega_{US}) RV_{US,t-k} + \theta_X \sum_{k=1}^{K} \phi_X(\omega_X) X_{t-k} \quad (2.15)$$

In such a way, we control for the US spillover effect and re-evaluate the role of macroeconomic variables in terms of volatility explaination. The results are summarized in Table 2.10 for Japan and Table 2.14 for UK, respectively. For the Japanese stock market, the impact from the macroeconomic variables is found to be increased in terms of magnitude (and/or significance level) relative to the GARCH-MIDAS-RV-X estimation results without controlling for volatility spillovers (see Table 2.9). For instance, the negative impact from industrial production growth increases from −0.4819 to −0.2319 after controlling for US volatility spillovers. More importantly, controlling for US volatility spillovers helps to improve the significance of the macroeconomic variables on volatility for the 2007-2014 period. With the exception of the unemployment rate and the term spread, all remaining macroeconomic variables have significant impacts during 2007-2014. We find similar results for the UK (see Table 2.14) once we control for volatility spillovers from the US.

[**Insert Table 2.10 here**]

[**Insert Table 2.14 here**]

2.7 Principal Component Analysis: Combined Macroeconomic contribution on Stock Variations

Our empirical results so far in relation to the macroeconomic variables indicate that the magnitude and significance of the impact for each macroeconomic variable on volatility is uneven across sub-samples in that not all of the variables are significant across all of the sub-samples. Further, our results suggest that alternative macroeconomic variables perform equally well in terms of providing additional contributions to explaining volatility above that provided by historical volatility. This makes it quite difficult for us to select the macroeconomic variables that are the best out of the alternatives we consider. Following Asgharian et al. (2013), we opt for extracting the first principal component from the macroeconomic variables we use and apply it into the GARCH-MIDAS model, evaluating the combined macroeconomic effect on stock market volatility. Following Equation 2.11, we let the long-term volatility component τ_t be driven by the first principal component (hereafter PC_1). Our estimation results are summarized in Tables 2.16.

[Insert Table 2.16 here]

Table 2.16 summarizes estimation results across all sub-samples for each country. For the US stock market, PC_1 has a high correlation with industrial production (92%), the unemployment rate (−87%) and housing starts (50%), respectively. We observe a negative and significant impact of PC_1 on the long-term volatility component, with the exception of the 2007-2014 sub-sample. For the UK stock market, PC_1 is closely related with inflation (79%) and the unemployment rate (75%). As reported in Table 2.11 and Table 2.12, both inflation and unemployment rate experience a change in sign from positive to negative during the UK's Great Moderation because of changes in economic conditions. Hence, we observe that PC_1 in Table 2.16 changes into a negative but insignificant sign during the Great Moderation episode (1985-2006) in the UK.

To try and gain some insight into the additional contribution that PC_1 makes to explaining volatility, we examine the significance of PC_1 in two alternative models: one in which we include PC_1 and RV in the MIDAS filter, which we term the restricted model for the purposes of this discussion (see in Equation 2.12), and one in which we additionally include US volatility spillovers, which we term the unrestricted model. The long-term volatility component τ_t in the restricted GARCH-MIDAS-RV-PC model is:

$$\tau_t = m + \theta_L \sum_{k=1}^{K_{rv}} \phi_L(\omega_L) RV_{L,t-k} + \theta_{pc} \sum_{k=1}^{K_{pc}} \phi_{PC}(\omega_{pc}) PC_{1,t-k}$$

while for the unrestricted GARCH-MIDAS-RV-PC model it is:

$$\tau_t = m + \theta_L \sum_{k=1}^{K} \phi_L(\omega_L) RV_{L,t-k} + \theta_{US} \sum_{k=1}^{K} \phi_{US}(\omega_{US}) RV_{US,t-k} + \theta_{pc} \sum_{k=1}^{K} \phi_{PC}(\omega_{pc}) PC_{1,t-k}$$

where RV_L is local historical realized volatility and RV_{US} is global realized volatility that spillovers from the U.S into local market. PC_1 refers to the first principal component derived from the local market. Results are reported in Table 2.17.

[Insert Table 2.17 here]

Comparing results from these two specifications in Table 2.17, we can see that adding US volatility spillovers into the Japanese stock market helps to deliver a better description of additional impact of PC_1 on stock variations in the sense that the magnitude of the coefficient and its significance increase. There seems to be less improvement on PC_1 when adding US volatility spillovers into the UK stock market, which is consistent with the argument in Susmel and Engle (1994) that there is only weak evidence of volatility spillovers from the US to the UK. In the UK stock market, PC_1 is superior to all the macroeconomic variables bar inflation, in terms of goodness fit across all sub-samples (see in Table 2.14). Taking the full sample results as an example, the log-likelihood ratio for PC_1 is 36525.4, which is larger than that for the other macroeconomic variables bar inflation (36526.7).

2.8 Forecasting Comparison: The Model Confidence Set Approach

In this section, we implement the Model Confidence Set (MCS) approach, as introduced by Hansen et al. (2011), to evaluate the predictive ability of the GARCH-MIDAS models specified with alternative macroeconomic variables. Before introducing the MCS approach, we will describe a recursive forecasting procedure which is used to construct a sequence of loss functions for pairwise forecasting comparisons in the Model confidence set.[4] We use a recursive forecasting procedure to allow for any change/breaks in the parameters of the models we use to forecast. The recursive forecating procedure allows the parameters to be updated to reflect any breaks or time variations in the stock market.

2.8.1 Recursive Forecasting Procedure

We set up a recursive forecasting procedure, which involves a rolling estimation window with fixed length and this estimation window moves forward recursively after one-step ahead predictions over a one-month horizon have been calculated. The recursive prediction procedure is as follows. First, we divide the whole dataset into two sub-periods, as shown in Figure 2.8. There is an initial estimation window from 01/1970 to 12/2009 and a prediction period from 01/2010 to 12/2012. We use available data from the initial estimation window to estimate the GARCH-MIDAS model specified with a macro variable. Second, with the estimated parameters we generate iterated one-day-ahead volatility predictions over a one-month horizon. We then compare the predicted volatilities with the realized volatilities. The daily forecast errors are

[4]The loss function refers to a way of measuring the forecasting error, and it can be interpreted as the distance between predicted and the real values.

averaged over the forecasting month and saved as a loss function. Following Patton (2011), the loss function we use is the mean squared error (MSE) one:

$$L_{i,t} = MSE(\sigma_{i,t}, \hat{\sigma}_{i,t}) = \frac{1}{N_t} \sum_{j_t=1}^{N_t} (\sigma_{i,j_t} - \hat{\sigma}_{i,j_t})^2 \tag{2.16}$$

where $L_{i,t}$ denotes the loss function of model i in month t, $\hat{\sigma}_{i,j_t}$ is the one-day ahead volatility forecast available on date j of month t, σ_{i,j_t} is the realized volatility available on the same date j, j_t indexes trading days during month t and N_t is the total number of trading days in month t. Third, after calculating the loss function, the estimation window automatically moves forward one month so that the sample starts in 02/1970 and ends in 01/2010. The parameters of the GARCH-MIDAS model are then re-estimated. With the updated parameters, we follow the second step above. By repeating this forecasting procedure, a series of loss functions $\{L_{i,t}\}_{t=1}^{T}$ for each month over which we forecast are generated.

We would also like to see how, if at all, macroeconomic information helps to predict the global financial crisis so we repeat the recursive forecasting procedure but this time we use an initial estimation window from 01/1985 to 12/2006, and generate recursive predictions over the period 01/2007 to 12/2009. The time line for this prediction exercise is shown in Figure 2.9.

2.8.2 Model Confidence Set

From the traditional viewpoint, a "best" forecasting model generally should keep its superior position under any circumstance. Hansen et al. (2011), however, argues that model selection in terms of choosing a model based on forecasting performance depends on its informativeness under a certain dataset. When data is informative, it is easy to identify the best forecasting model. When it is not so informative, it might be difficult to distinguish superior models from inferior ones. Hansen et al. (2011) proposed the Model Confidence Set (MCS) to examine the predictive ability of volatility models in a more flexible way that allows more than one model to stay within a certain confidence interval when dealing with uninformative dataset. With regard to our data, a sub-period associated with high volatility in stock market might be more informative about which macro variable contributes most, and vice versa. Therefore, we implement the Model Confidence Set (MCS) to evaluate to what extent does macroeconomic information contribute to stock market volatility in terms of improving the predictive ability of the model.

We now briefly describe the MCS approach to model selection for our case. The main mechanism of MCS in selecting superior models is similar to a hypothesis test of parameter significance under a certain confidence interval. In the first step, we construct an initial set M_0 that contains all the GARCH-MIDAS models specified with the different macroeconomic variables. The null hypothesis ($H_{0,M}$) assumes that all candidate models in the set M_0 have equal predictive ability (EPA). In contrast, the alternative hypothesis ($H_{A,M}$) assumes there is at least one candidate model that is inferior than the other models. More formally,

$$\begin{aligned} &H_{0,M} : d_{i,j} = 0 \qquad\qquad\qquad H_{A,M} : d_{i,j} \neq 0 \\ &d_{i,j} = E(d_{ij,t}) = E(L_{i,t} - L_{j,t}) \qquad \forall i,j \in M \qquad t = 1, 2, ...n \end{aligned} \tag{2.17}$$

The null hypothesis of equal predictive ability can be tested by means of d_{ij}, which infers the relative performance of any two models, i and j. If i and j have equal predictive ability, the null hypothesis is accepted. Otherwise, the null hypothesis is rejected. In Equation 2.17, d_{ij} is calculated as the expectation of loss differential between any two models, i and j.

In the second step, we test the null hypothesis of equal predictive ability in set M_0. To achieve this, we make pairwise comparisons of the relative performance (d_{ij}) between any two models in the set M_0. If the null hypothesis is accepted, we terminate the MCS procedure, setting the confidence set $\hat{M}^*_{1-\alpha}$ equal to M_0, which suggests all candidates models perform equally well under a confidence interval $1 - \alpha$. If the null hypothesis is rejected, the model with the worst performance is eliminated from the initial set M_0. In this case, the initial set M_0 narrows down into set M_1, which satisfies the condition of $M_1 \subset M_0$. In the third step, we repeat the above second step. As long as the null hypothesis is rejected, the model set will continue to be trimmed of the most inferior model.[5]. Once the null hypothesis is accepted, the surviving models with equal predictive ability are saved in a confidence set $\hat{M}^*_{1-\alpha}$. This set $\hat{M}^*_{1-\alpha}$ with the surviving models is denoted as M^*, and it satisfies the following condition:

$$M^* = \{i \in M_0 : E(d_{ij,t} \leq 0) \;\; \forall i, j \in M_0\} \tag{2.18}$$

which indicates that any surviving model i in set M^* must satisfy the condition that it is superior to other eliminated models with $E(d_{ij,t})$ less than zero.

Following Hansen et al. (2011), we employ the range test statistic, T_R, to test the null hypothesis. T_R is constructed based on the t-statistics:

$$\begin{gathered} T_R = max_{i,j \in M} = |t_{ij}| \\ t_{ij} = \frac{\bar{d}_{ij}}{\hat{\sigma}_{\bar{d}_{ij}}} \end{gathered} \tag{2.19}$$

$\bar{d}_{ij}$ in equation 2.19 denotes the loss differential on average between two

[5]As long as the null hypothesis is rejected, the initial Set M_0 will continue to narrow down into a smaller set that satisfies $M_{m0}... \subset M_2 \subset M_1 \subset M_0$.

models i and j. The standard deviation $\hat{\sigma}_{\bar{d}_{ij}}$ can be estimated by using the bootstrap methodology. In Equation 2.19, T_R measures the largest loss difference between two models. In this context, the candidate model that yields the largest value of T_R will be eliminated.

Tables 2.18 through 2.20 report the prediction performance of all the GARCH-MIDAS models specified with the different macroeconomic variables using the MCS equivalent test for the US, UK and Japan, respectively. For each country, we conduct recursive prediction for two periods: 2007-2009 and 2010-2012, which are related to two major events: the global financial crisis and the sovereign debt crisis. Each prediction period covers 36 months that we divide into 6-month sub-periods. For every sub-period, we aggregate monthly loss functions and generate the MCS equivalent test using a significance level of α equal to 25%.

Table 2.18 presents the MCS results for the US stock market. The figures reported are MCS p-values. This p-value can be interpreted as the probability of being kept in the Model Confidence Set. The prediction period of Jul 2010-Dec 2012 seems to be more informative, where most models are left within the confidence set $\hat{M}^*_{75\%}$. The term spread and unemployment rate models in the 2010-2012 period outperform alternatives in terms of obtaining the highest probability of 1.0000 in 4 out of 6 sub-periods, where a p-value of 1.0000 indicates that this candidate model is the last survivor in the set $\hat{M}^*_{75\%}$. In contrast, the prediction period of Jan/2007-Dec/2009 seems to be less informative and only a few candidates are kept within the confidence set. In particular, the model where the macroeconomic variable is the first principal component is superior to alternatives with the highest probability of 1.0000 during the financial crisis episode.

[Insert Table 2.18 here]

Table 2.19 reports MCS results for the UK stock market. Looking across all sub-periods, PC_1 contributes most during the global financial crisis while inflation seems to replace PC_1 and becomes more prominent after 2010. In terms of the Japanese stock market, the MCS p-values in Table 2.20 reveal that term spread and PC_1 perform equally well with p-value no less than 0.8 during 2010-2012.

[**Insert Tables 2.19 and 2.20 here**]

2.9 Conclusion

In this paper, we examine the dynamic relationship between macroeconomic variables and stock market under the framework of the GARCH-MIDAS model. We investigate the sources of macroeconomic information that drive long-run volatility fluctuations in the US, UK and Japanese stock markets. To allow for potential structural breaks and to investigate what impact they may have, if any, we partition our full sample, which runs from 1970-2014 into three different sub-samples and observe the interactions between different macroeconomic variables and stock volatility across the different sub-samples for each country.

For the US, we re-confirm the counter-cyclical relationship between many of the macroeconomic variables and stock volatility found by Engle et al. (2013), Asgharian et al. (2013) and Conrad and Loch (2015). We do find that the counter-cyclical response of stock volatility varies with respect to different underlying economic conditions across all sub-samples. For example, industrial production and housing starts, as indicators of economic activity, become less informative after the Great Moderation. In contrast, the unemployment rate becomes more informative after the Great Moderation.

Though individual macroeconomic variables are found to have significant impacts on stock variations for both UK and Japan markets, macroeconomic influences observed in both UK and Japan are some differences compared to the US. For instance, after the real estate bubble burst at the end of 1989 in Japan, the Japanese economy entered a prolonged deflationary period and therefore we might expect a change in the relationship between stock market volatility and the macroeconomy between the boom phase (1970-1989) and the deflationary period. We find that during the boom phase period, inflation and the unemployment rate have a significantly negative impact on stock volatility. During the deflationary phase, however, we observe a positive relationship between inflation and unemployment and volatility, which implies that market participants might have changed their expectations about future fiscal and monetary policy from tightening to loosening. We find that the change between those two periods can be quite pronounced: the impact of inflation on volatility decays slowly during the boom phase whereas the effect of inflation on volatility dies out quickly during the deflationary phase. For the UK, once the economy entered into the Great Moderation, stock market volatility tends to increase. We also observe significant changes in the magnitude of the impact from alternative macroeconomic variables to stock volatility around the turning point of the Great Moderation. Industrial production, for example, has a negative and significant impact on stock volatility before the Great Moderation. During the Great moderation, its magnitude dramatically decreases. We also observe that both inflation and the unemployment rate experience changes in sign, going from positive to negative during the Great Moderation episode.

Following the literature on volatility spillovers, we examine the robustness of our findings in terms of the significance of macroeconomic variables in explaining stock return volatility for the UK and Japan after controlling for volatility spillovers from the US. We find that US volatility, which we use as a proxy for common volatility, is significant in explaining UK and Japanese stock market volatility. However, the significance of volatility spillovers does not change our conclusions in relation to the role that macroeconomic variables play in explaining stock market volatility. In fact, most macroeconomic variables perform equally well, if not better, once we control for volatility spillovers.

Finally, we examine the forecasting performance of the GARCH-MIDAS-X models using the Model Confidence Set. We find that depending on the forecasting period, the model with the first principal component as the macroeconomic variable performs well across all the countries in our sample, although the models including the term spread and the unemployment rate work well for the US when predicting over the sovereign debt crisis period.

To summarize, using the GARCH-MIDAS model we find evidence that the relationship between macroeconomic variables and stock market volatility that has been documented for the US is still alive and well and is also present in the UK and Japan, although the nature of the relationship in terms of the signs of the coefficients on the macroeconomic variables is not the same across all the countries in our sample. In particular, our sub-sample analysis reveals that the the relationship between the macroeconomy and stock market volatility is not stable over time and may be subject to structural breaks. While any time variation or breaks can be accommodated to some extent when using these models to predict volatility, as the results from our forecasting exercise show, it raises an interesting question as to whether the macroeconomic variables help to explain not only stock market volatility but structural breaks in volatility. We investigate this question in the next chapter.

Table 2.1: Summary Statistics for stock market and macro data

Sample	Variable	**Mean**	**STD**	**Skewness**	**Kurtosis**	Variable	**Mean**	**STD**	**Skewness**	**Kurtosis**
			UK					**U.S**		
1970-2014	R_{FTSE}	0.0004	0.0106	-0.2579	11.1051	$R_{S\&P500}$	0.0003	0.0107	-0.5605	20.5128
	ΔIP	0.9235	7.1513	0.9854	10.4108	ΔIP	2.4074	6.1030	-1.0151	6.2105
	ΔPPI	1.8012	2.3170	1.6995	7.2302	ΔPPI	1.3284	3.8082	0.1029	10.0304
	ΔUEM	3.6641	29.2911	1.8112	7.5175	ΔUEM	0.3889	18.1837	0.5130	4.8061
	TS	1.0764	1.7386	-0.1779	2.9654	TS	1.6463	1.3190	-0.6011	2.9102
	ΔHS	0.1452	18.0300	-0.1500	4.1097	ΔHS	0.9659	13.6814	0.5301	4.2002
1970-1984	R_{FTSE}	0.0006	0.0111	0.2244	8.6460	$R_{S\&P500}$	0.0002	0.0089	0.3179	4.9473
	ΔIP	1.3980	10.6344	0.8519	5.9017	ΔIP	2.7511	8.2581	-0.6246	3.8173
	ΔPPI	3.8068	2.7540	0.8761	4.4771	ΔPPI	2.4021	3.5682	2.2839	13.3845
	ΔUEM	15.2207	34.9393	0.8723	4.1946	ΔUEM	2.1111	22.2477	0.4748	4.2874
	TS	2.0062	1.5804	-0.2107	3.5055	TS	1.1804	1.5104	-0.5781	2.4842
	ΔHS	-1.6382	19.9451	0.2027	2.9629	ΔHS	1.7250	15.7342	0.6028	4.2079
1985-2006	R_{FTSE}	0.0005	0.0093	-0.8705	14.1445	$R_{S\&P500}$	0.0004	0.0104	-1.3139	30.0902
	ΔIP	1.3751	4.1178	0.8601	5.1497	ΔIP	2.8166	3.5980	-0.5217	3.5152
	ΔPPI	0.8029	1.0446	0.3754	3.4118	ΔPPI	0.7602	3.0341	0.1021	5.9474
	ΔUEM	-4.3148	16.6609	1.4610	6.3147	ΔUEM	-1.0985	13.7026	0.2074	3.6709
	TS	0.1704	1.4997	-0.2279	3.1652	TS	1.7373	1.1708	-0.0108	1.8094
	ΔHS	1.1915	11.7441	-0.3354	3.1888	ΔHS	0.5910	10.8558	0.4341	4.2253
2007-2014	R_{FTSE}	0.0002	0.0128	-0.1743	9.7969	$R_{S\&P500}$	0.0003	0.0139	-0.0773	12.5519
	ΔIP	-1.2080	5.1741	-1.5856	5.8594	ΔIP	0.6378	6.7090	-1.7865	5.5515
	ΔPPI	0.7785	1.3116	0.4931	3.9004	ΔPPI	0.9172	5.4125	-1.0694	6.0515
	ΔUEM	3.6641	29.2911	1.8112	7.5175	ΔUEM	1.2500	20.2744	0.3166	3.3333
	TS	1.8247	1.3164	-0.6978	2.2418	TS	2.2694	0.9752	-0.9643	3.6852
	ΔHS	0.6116	26.5282	-0.2535	2.9231	ΔHS	0.5738	16.3602	0.3139	2.5001
						JAPAN				
1970-2014	$R_{Nikkie225}$	0.0002	0.0126	-0.4342	13.4042	ΔIP	0.1633	1.6827	-2.2272	20.6261
	ΔPPI	1.1590	6.9643	4.2493	37.8422	ΔUEM	1.1590	6.9643	4.2493	37.8422
	TS	1.8190	0.8204	0.3683	2.5006	ΔHS	0.1254	6.6542	0.7375	14.6261
	ΔYen	8.9749	42.5824	2.1738	11.7219					
1970-1989	$R_{Nikkie225}$	0.0004	0.0083	-1.1620	16.1157	ΔIP	0.3484	1.3411	-0.1454	3.0379
	ΔPPI	3.2529	8.7972	4.2014	28.5326	ΔUEM	3.9117	13.2949	0.8626	3.7525
	TS	2.3320	0.7517	0.1829	3.1056	$DeltaHS$	0.2934	9.0365	1.0917	11.1978
	ΔYen	10.7724	38.2405	1.7403	7.6240					
1990-2006	$R_{Nikkie225}$	0.0004	0.0130	-0.7702	26.6373	ΔIP	0.0719	1.2650	-0.1631	3.5127
	ΔPPI	-0.7263	2.7418	-0.1918	3.0262	ΔUEM	4.1455	10.1730	0.3969	2.3738
	TS	1.6703	0.6151	0.5616	2.6809	ΔHS	0.1117	4.2497	0.1374	3.8918
	ΔYen	8.1847	46.4783	2.6266	14.8040					
2007-2014	$R_{Nikkie225}$	0.0000	0.0161	-0.5150	10.9457	ΔIP	-0.1049	2.8184	-2.4130	13.3079
	ΔPPI	-0.0698	6.6474	-0.0444	4.8793	ΔUEM	-0.0698	6.6474	-0.0444	4.8793
	TS	0.8880	0.2442	-0.2168	1.9561	ΔHS	-0.1517	6.8724	-1.0755	7.9440
	ΔYen	7.2839	41.9627	1.4340	5.7191					

[a] Table reports statistics for U.S, UK and Japan Stock markets and macroeconomic dataset,

ranging from 01/1970 to 12/2014. For all macroeconomic variables, Δ denotes as the first difference of respective levels. Excluding the Term spread, Δ refers to annualized month-to-month percentage changes as $\Delta X_t = ((X_t/X_{t-1})^{12} - 1)$. The Term Spread is measured as a difference between 10-year Treasury Bond Yield and 3-month T-bill rate. The log-returns for all markets are collected from WRDS database and global financial database. All macroeconomic data is gathered from Federal Reserve Economic dataset (FRED). Realized Volatility (RV) calculated from 5-min tick-by-tick data is gathered from Oxford-Man Institute Realized library for S&P500 Index, FTSE All Share Index and Nikkei 225 Index.

Table 2.2: Parameter Estimates for the GARCH-MIDAS with Realized Volatility

Sample	**MIDAS Filter**	α	β	θ	ω_1	ω_2	**m**	**L-Likelihood**	**AIC**	**BIC**
				Panel A: GARCH-MIDAS with Realized Volatility in U.S						
Full Sample 1970-2014	**Rolling RV**	0.0849 (32.3056)	0.8609 (92.2387)	**0.0281***** (13.8398)	1	**8.4989***** (5.5277)	0.0000 (9.3576)	35608.2042	−71204.4083	−71160.3838
Sub-sample 1970-1984	**Rolling RV**	0.0332 (3.7818)	0.8863 (9.0296)	**0.0355***** (6.0409)	1	46.4365 (1.5355)	0.0000 (2.2153)	10970.9613	−21929.9226	−21892.4803
Sub-sample 1985-2006	**Rolling RV**	0.1101 (26.4980)	0.8002 (49.7525)	**0.0334***** (13.8101)	1	**7.7203***** (5.6323)	0.0000 (6.8383)	16544.5118	−33077.02377	−33037.2943
Sub-sample 2007-2014	**Rolling RV**	0.1280 (7.5937)	0.8235 (30.0502)	**0.0183***** (3.4985)	1	**6.3705*** (2.0817)	0.0001 (3.7695)	4831.0054	−9650.0108	−9616.3635
				Panel B: GARCH-MIDAS with Realized Volatility in UK						
Full Sample 1970-2014	**Rolling RV**	0.1216 (20.6358)	0.7991 (69.3261)	**0.0297***** (17.5896)	1	**10.1055***** (6.9554)	0.0000 (11.0941)	36187.4884	-72362.9768	-72318.8736
Sub-sample 1970-1984	**Rolling RV**	0.1246 (9.1345)	0.7597 (25.3432)	**0.0335***** (11.2577)	1	**13.7498**** (3.2072)	0.0000 (4.8478)	10887.8605	-21763.7210	-21726.0887
Sub-sample 1985-2006	**Rolling RV**	0.1225 (15.7879)	0.8033 (50.9676)	**0.0241***** (8.8746)	1	**9.2504***** (4.3409)	0.0000 (8.0927)	17063.9056	-34115.8113	-34076.0561
Sub-sample 2007-2014	**Rolling RV**	0.1159 (6.1195)	0.7919 (20.5612)	**0.0289***** (5.0499)	1	**12.7572**** (2.3693)	0.0000 (3.1066)	4857.1449	-9702.2898	-9668.6306
				Panel C: GARCH-MIDAS with Realized Volatility in Japan						
Full Sample 1970-2014	**Rolling RV**	0.1605 (45.8341)	0.8038 (144.7237)	**0.0494***** (13.1501)	1	**2.6595***** (8.0819)	0.0000 (5.5630)	33317.6	-66623.2	-66579.3
Sub-sample 1970-1989	**Rolling RV**	0.2850 (42.2200)	0.6209 (45.240)	**0.0337***** (10.1040)	1	**3.5726***** (7.0648)	3.52E-05 (6.6957)	15795.7	-31579.3	-31540.2
Sub-sample 1990-2006	**Rolling RV**	0.0822 (9.0903)	0.8892 (66.6060)	**0.0284***** (3.7630)	1	**1.9649**** (2.0470)	7.97E-05 (2.3567)	10586.7	-21161.4	-21123.4
Sub-sample 2007-2014	**Rolling RV**	0.1198 (8.6405)	0.8224 (29.3864)	0.0107 (1.5889)	1	6.2214 (1.0204)	0.0002 (4.3203)	4141.8	-8271.6	-8238.2

[a] Table reports estimation results for the GARCH-MIDAS model with Realized Volatility, where the long-term volatility component τ_t is smoothed out by a MIDAS filter with 24 lags (K=24) of monthly realized volatility:

$$\tau_{i,t} = m + \theta * \sum_{k=1}^{K} \phi_k(1, \omega_2) RV_{i,t-k}^{rw}$$

We employ the rolling window approach to calculate monthly realized volatility, $RV_{i,t}^{rw}$. $RV_{i,t}^{rw}$ is calculated by recursive rolling back fixed number of days (N'=22) from day i. Consequently, monthly $\tau_{i,t}$ and $RV_{i,t}^{rw}$ is available at each day i in month t. The number in brackets are robust t-statistics. ***, **, * indicate the significants at 1%, 5%, 10% level, respectively.

Table 2.3: Parameter Estimates for GARCH-MIDAS incorpated with Macroeconomic Variables I in U.S

Sample	MIDAS filter	α	β	θ	ω_1	ω_2	m	L-Likelihood	AIC	BIC
				Panel A: GARCH-MIDAS-IP Model						
Full Sample 1970-2014	**IP**	0.0690	0.9310	**−0.5370*****	1	**4.9999*****	−0.2302	34883.2119	−69754.4238	−69710.3993
		(53.6931)	(724.2345)	(−33.3301)		(28.0670)	(−0.6384)			
Sub-sample 1970-1984	**IP**	0.0677	0.9323	**−0.9540*****	1	**4.9998*****	−0.3087	9737.1821	−19462.3643	−19424.9219
		(154.2410)	(2123.9494)	(−57.8164)		(33.0843)	(−0.8069)			
Sub-sample 1985-2006	**IP**	0.0775	0.9134	**−0.1074*****	1	**1.0516*****	−8.6609	16690.4284	−33368.8568	−33329.1275
		(32.6609)	(253.4220)	(−4.0417)		(5.7665)	(−47.7846)			
Sub-sample 2007-2014	**IP**	0.1245	0.8463	**−0.0914*****	1	**4.8899*****	−8.9082	4829.3512	−9646.7025	−9613.0552
		(8.3181)	(48.1014)	(−4.7819)		(2.9936)	(−48.6548)			
				Panel B: GARCH-MIDAS-TS Model						
Full Sample 1970-2014	**TS**	0.0751	0.9170	**−0.1907*****	1	13.1404	−8.5603	35602.3438	−71192.6876	−71148.6631
		(37.4606)	(325.4681)	(−6.2165)		(1.4547)	(−52.5856)			
Sub-sample 1970-1984	**TS**	0.0399	0.9601	**−0.3323***	1	**1.0173*****	−10.2221	10962.1946	−21912.3891	−21874.9468
		(10.4733)	(264.2631)	(−1.7204)		(5.7879)	(−18.6021)			
Sub-sample 1985-2006	**TS**	0.0834	0.8969	**−0.4461*****	1	**1.0010*****	−8.3797	16703.6207	−33395.2415	−33355.5121
		(31.8836)	(178.8744)	(−8.1855)		(23.1092)	(−64.7790)			
Sub-sample 2007-2014	**TS**	0.1474	0.8512	**0.4874***	1	1.0781	−7.7128	4826.6839	−9641.3677	−9607.7205
		(8.9781)	(51.4262)	(1.8609)		(1.3610)	(−8.8168)			
				Panel C: GARCH-MIDAS-PPI Model						
Full Sample 1970-2014	**PPI**	0.05156	0.9484	**0.1622***	1	**1.0476*****	−10.528	35013.5	−70015	−69971
		(48.644)	(742.12)	(1.9007)		(7.6991)	(−25.356)			
Sub-sample 1970-1984	**PPI**	0.0446	0.9434	**0.2184*****	1	**1.0034*****	−10.069	11286.9	−22561.8	−22524.4
		(8.3795)	(130.85)	(3.9859)		(22.012)	(−61.369)			
Sub-sample 1985-2006	**PPI**	0.0684	0.9315	**−0.1421***	1	**4.9999***	−0.2604	16300.7	−32589.4	−32549.7
		(35.658)	(485.3)	(−1.6826)		(1.878)	(−1.1406)			
Sub-sample 2007-2014	**PPI**	0.1388	0.8085	**0.4329****	1	**2.3337*****	−9.749	3809.74	−7607.48	−7573.84
		(6.5502)	(28.752)	(2.845)		(3.7024)	(−51.323)			

[a] Table 2.3 reports estimation results for the GARCH-MIDAS model with alternative macroeconomic variables. Panel A provides parameter estimates for the GARCH-MIDAS-IP model, in which the long-term volatility component τ in a MIDAS filter has been smoothed out by 24 lags (K=24) of Industrial production growth rates in a monthly frequency:
$\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(1, \omega_2) IP_{t-k}$

[b] Panel B and Panel C provide parameter estimates for the GARCH-MIDAS model incorporated with Term Spread and Inflation (PPI). [c] The numbers in parentheses are robust t-statistics, where ***,**,* indicate significant levels at 1%, 5% and 10%, respectively. The number in bold means estimated parameter is significant with at least 10% level. For each specification of the GARCH-MIDAS model, we set up restricted weighting scheme with $\omega_1 = 1$.

Table 2.4: Parameter Estimates for the GARCH-MIDAS incorporated with Macroeconomic Variables II in U.S

Sample	MIDAS filter	α	β	θ	ω_1	ω_2	m	L-Likelihood	AIC	BIC
				Panel A: GARCH-MIDAS-UEM Model						
Full Sample 1970-2014	UEM	0.0728 (37.7527)	0.9127 (292.7543)	**1.1346***** (4.8408)	1	**3.8371***** (3.4969)	−9.2501 (−131.4361)	35785.9330	−71559.8661	−71515.8416
Sub-sample 1970-1984	UEM	0.2154 (14.4671)	0.7846 (52.7054)	**0.5632**** (2.2918)	1	13.8246 (1.3163)	−0.0204 (−0.0412)	10975.5840	−21939.1680	−21901.7257
Sub-sample 1985-2006	UEM	0.1915 (25.3263)	0.8085 (106.9614)	**1.5639***** (4.0651)	1	**9.4033***** (2.8093)	−0.3407 (−0.8799)	16578.7214	−33145.4428	−33105.7135
Sub-sample 2007-2014	UEM	0.1344 (9.1834)	0.8653 (59.1261)	0.8471 (1.2491)	1	**1.0017***** (5.8320)	−5.3063 (−6.2944)	4929.1501	−9846.3003	−9812.6530
				Panel B: GARCH-MIDAS-HS Model						
Full Sample 1970-2014	HS	0.0752 (36.3466)	0.9151 (301.2894)	**−0.1091***** (−5.7428)	1	**1.0026***** (83.5328)	−8.8603 (−73.9977)	35787.9173	−71563.8345	−71519.8100
Sub-sample 1970-1984	HS	0.0465 (8.8704)	0.9459 (149.0078)	0.0055 (1.0588)	1	43.7535 (0.4761)	−9.4708 (−57.7475)	11142.7945	−22273.5890	−22236.1467
Sub-sample 1985-2006	HS	0.0900 (30.2037)	0.8919 (178.2835)	**−0.3111***** (−8.6988)	1	**1.0083***** (132.2393)	−8.9427 (−87.1610)	16704.7243	−33397.4486	−33357.7193
Sub-sample 2007-2014	HS	0.1291 (7.8584)	0.8366 (38.6042)	**−0.1270***** (−4.9136)	1	**1.0010***** (32.3409)	−8.9161 (−53.9686)	4830.8089	−9649.6179	−9615.9706

[a] Table 2.4 reports estimation results for the GARCH-MIDAS model with two alternative specifications, UEM and HS. Panel A provides parameter estimates for the GARCH-MIDAS-UEM model, in which the MIDAS filter smooths out 24 lags (K=24) of monthly growth rates of unemployment rate:

$$\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(1, \omega_2)\Delta UEM_{t-k}$$

[b] Panel B provides parameter estimates for the GARCH-MIDAS-HS model, in which the MIDAS filter smooths out 24 lags of monthly growth rates of Housing Starts:

$$\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(1, \omega_2)\Delta HS_{t-k}$$

[c] The numbers in parentheses are robust t-statistics, where ***,**,* indicate significant levels at 1%, 5% and 10%, respectively. The number in bold means estimated parameter is significant with at least 5% level. For each specification of the GARCH-MIDAS model, we set up restricted weighting scheme with $\omega_1 = 1$.

Table 2.5: Variance Ratios for Macroeconomic Contribution in the stock Markets

Macro Variable	1970–2014	1970-1984	1985-2006	2007-2014
	U.S Stock Market			
Industrial Production Growth (ΔIP)	12.11%	15.24%	9.05%	1.98%
Housing Starts Growth (ΔHS)	16.17%	29.97%	28.31%	21.04%
Unemployment Rate (ΔUEM)	10.01%	12.87%	14.93%	17.81%
Producer Price Index Growth (ΔPPI)	0.68%	13.12%	5.15%	13.06%
Term Spread (TS)	17.69%	26.56%	51.09%	24.07%
1st Principal Component (PC_1)	16.48%	21.38%	23.64%	31.58%
	UK Stock Market			
Industrial Production(ΔIP)	6.40%	7.85%	2.97%	16.65%
Unemployment Rate (ΔUEM)	4.13%	4.68%	13.34%	25.19%
Producer Price Index (ΔPPI)	9.76%	39.21%	17.50%	13.97%
Term Spread (TS)	1.92%	1.63%	11.13%	34.98%
1st Principal Component (PC_1)	32.65%	39.49%	22.0%	21.11%
	1970–2014	**1970-1989**	**1990-2006**	**2007-2014**
	Japan Stock Market			
Industrial Production Growth (ΔIP)	8.24%	6.65%	11.13%	6.28%
Housing Starts Growth (ΔHS)	9.78%	14.24%	35.62%	6.24%
Unemployment Rate (ΔUEM)	21.17%	12.58%	8.55%	12.95%
Producer Price Index Growth (ΔPPI)	27.57%	28.30%	9.09%	30.24%
Term Spread (TS)	39.64%	35.82%	3.52%	1.63%
Japanese Yen Index (Yen)	3.13%	14.10%	8.81%	29.11%
1st Principal Component (PC_1)	41.05%	25.25%	13.26%	9.34%

[a] Table 2.5 summarizes variance ratio for the GARCH-MIDAS model specified with alternative macroeconomic variables and their combined factor, the first principal component.
[b] The variance ratio is calcluated by a sample variance fraction: $\frac{\widehat{var} log\tau_t^M}{\widehat{var} log(\tau_t^M g_t^M)}$, where M refers to a GARCH-MIDAS model specified with one macro variable. $\widehat{var} log\tau_t^M$ denotes the log long-run volatility component that is purely driven by macroeconomic variable x. $\widehat{var} log(\tau_t^M g_t^M)$ denotes that both long-run and short-run components are combined together as an entire volatility.

Table 2.6: Parameter Estimates for GARCH-MIDAS model with Realized Volatility and Macroeconomic Variable in U.S

MIDAS	α	β	m	θ_{RV}	ω_{RV}	θ_X	ω_X	Likelihood	AIC	BIC
					Full-sample 1970-2014					
RV+IP	0.0809	0.7805	2.23e-05	0.0385***	49.074***	-0.0109**	2.8976*	35935	-71853	-71794
	(30.294)	(48.29)	(9.2027)	(26.667)	(7.1594)	(-3.2465)	(2.0313)			
RV+PPI	0.0808	0.7769	1.47e-05	0.0401***	48.779***	0.01181	3.5098	35931	-71846	-71787
	(30.2)	(47.522)	(8.0935)	(29.362)	(7.565)	(1.4151)	(0.68195)			
RV+UEM	0.0812	0.7789	1.94e-05	0.0386***	49.129***	0.0049***	7.0165*	35935	-71854	-71795
	(30.209)	(47.275)	(9.988)	(26.766)	(7.13)	(3.435)	(1.7751)			
RV+HS	0.0808	0.7788	2.17e-05	0.0381***	48.939***	-0.0210***	1.0548***	35937	-71859	-71800
	(29.934)	(47.415)	(10.407)	(26.426)	(7.0358)	(-4.8306)	(7.8091)			
RV+TS	0.0807	0.7782	3.07e-05	0.0374***	48.953***	-0.0585***	1.0347***	35940	-71865	-71806
	(30.331)	(46.608)	(9.7523)	(26.405)	(6.8621)	(-5.9303)	(4.8021)			
					Sub-sample 1970-1984					
RV+IP	0.0319	0.9060	3.16e-05	0.0309***	48.532*	-0.0148**	3.3825*	11242	-22468	-22418
	(4.0717)	(22.001)	(4.3524)	(6.6921)	(1.838)	(-2.9875)	(2.0261)			
RV+PPI	0.0317	0.9001	1.17e-05	0.0339***	47.366*	0.0376*	1.2001	11239	-22463	-22413
	(3.9354)	(21.166)	(2.5842)	(8.2025)	(2.2356)	(2.3206)	(1.1951)			
RV+UEM	0.0320	0.9035	2.01e-05	0.0347***	48.107*	0.0034*	5.0145	11238	-22461	-22411
	(3.9799)	(21.052)	(3.9771)	(8.3566)	(2.0802)	(1.8429)	(0.9257)			
RV+HS	0.0310	0.9062	1.62e-05	0.0369***	47.006*	0.0036	44.37	11239	-22463	-22413
	(3.9895)	(21.265)	(3.3792)	(9.2653)	(2.31)	(1.4451)	(0.62448)			
RV+TS	0.0306	0.8998	3.24e-05	0.0322***	49.637*	-0.0726***	1.2562*	11243	-22469	-22419
	(3.8498)	(18.842)	(4.9127)	(7.952)	(1.8528)	(-3.7795)	(1.9359)			
					Sub-sample 1985-2006					
RV+IP	0.0534	0.7993	2.01e-05	0.0401***	37.307***	-0.0205***	2.2687*	16847	-33679	-33626
	(6.117)	(19.561)	(7.9744)	(25.676)	(9.0867)	(-3.9715)	(2.1499)			
RV+PPI	0.0538	0.7946	1.33e-05	0.0411***	35.202***	-0.0143	7.1095	16844	-33673	-33619
	(6.0464)	(18.553)	(5.9314)	(25.328)	(8.9983)	(-1.2065)	(0.9639)			
RV+UEM	0.0526	0.7967	1.75e-05	0.0389***	42.536***	0.0108***	6.8743**	16849	-33683	-33629
	(5.9639)	(18.745)	(8.1121)	(24.07)	(7.5599)	(4.0308)	(2.609)			
RV+HS	0.0542	0.7959	2.39e-05	0.0372***	46.814***	-0.0706***	1.0023***	16857	-33698	-33645
	(6.1289)	(18.896)	(9.382)	(22.59)	(5.9596)	(-6.7428)	(100.99)			
RV+TS	0.0503	0.7922	3.60e-05	0.0365***	47.384***	-0.0873***	1.0428	16855.3	-33695	-33642
	(5.6709)	(17.325)	(6.9897)	(20.487)	(6.0668)	(-5.5703)	(5.9681)			
					Sub-sample 2007-2014					
RV+IP	0.1283	0.8096	4.83e-05	0.0266***	12.15**	-0.0205	48.441	4979	-9942	-9897
	(7.3047)	(25.615)	(3.4935)	(5.0272)	(2.9278)	(-1.1519)	(0.48833)			
RV+PPI	0.1283	0.8022	3.23e-05	0.0275***	12.006**	0.0732	3.8744	4979	-9943	-9898
	(7.0517)	(23.522)	(3.2843)	(5.6934)	(2.9772)	(1.4019)	(0.97058)			
RV+UEM	0.1305	0.8066	3.66e-05	0.0279***	11.362**	-0.0046	21.48	4978	-9940	-9895
	(7.2667)	(25.33)	(2.759)	(5.0384)	(2.9637)	(-0.5781)	(0.36914)			
RV+HS	0.1277	0.8072	4.03e-05	0.0268***	12.113**	0.0010	19.846	4978	-9939	-9895
	(7.1172)	(24.744)	(3.1766)	(5.2649)	(2.9397)	(0.068962)	(0.060886)			

RV+TS	0.1279	0.8054	5.86e-06	0.0264***	12.634**	0.1412	30.008	4979	-9942	-9897
	(7.114)	(24.195)	(0.27022)	(5.2204)	(2.9211)	(1.602)	(0.18437)			

a Table 2.6 reports paramter estimates for GARCH-MIDAS-RV-X models including lags of RV and macro variable X for the U.S market. In the MIDAS filter, long-term volatility component is regressed by lags of RV and monthly observations from one alternative macro variable:

$$\tau_t = m + \theta_{rv} \sum_{k=1}^{K_{rv}} \phi_k(\omega_{rv}) RV_{t-k} + \theta_x \sum_{k=1}^{K_x} \phi_k(\omega_x) X_{t-k}$$

with K=24. Parameter θ_{rv} and θ_x govern impacts of realized volatility and macro variable, respectively. The numbers in parentheses are robust t-statistics, where ***,**,* indicate significant levels at 1%, 5% and 10%, respectively.

Table 2.7: Parameter Estimates for GARCH-MIDAS incorpated with Macroeconomic Variables I in Japan

Sample	**MIDAS filter**	α	β	θ	ω_1	ω_2	**m**	**L-Likelihood**	**AIC**	**BIC**
Panel A: GARCH-MIDAS-IP Model										
Full Sample	**IP**	0.1391	0.8552	−0.8090***	1	1.0010***	−7.7168	33274	−66536	−66492
1970-2014		(47.2842)	(240.8802)	(−7.0763)		(81.7391)	(−26.2501)			
Sub-sample	**IP**	0.2478	0.7090	−0.5675***	1	1.2112***	−9.0376	11966	-23920	-23883
1970-1989		(47.9922)	(67.0730)	(−3.0051)		(2.6105)	(−61.0008)			
Sub-sample	**IP**	0.0803	0.9022	−0.6422***	1	1.0010***	−8.3289	10708	−21405	−21367
1990-2006		(10.4595)	(96.5828)	(-3.2180)		(32.3668)	(-70.6963)			
Sub-sample	**IP**	0.1139	0.8831	−0.5046*	1	1.0988***	−6.8926	4257	−8501	−8462
2007-2014		(9.4848)	(72.7179)	(−1.7281)		(3.3815)	(−10.0431)			
Panel B: GARCH-MIDAS-TS Model										
Full Sample	**TS**	0.1521	0.8238	−0.8231* * *	1	1.001* * *	−7.0207	33879	−67746	−67702
1970-2014		(48.637)	(176.94)	(−20.134)		(40.897)	(−56.631)			
Sub-sample	**TS**	0.2645	0.6254	−0.9800***	1	1.6486***	−7.1457	11742	-23471	-23434
1970-1989		(35.1218)	(39.4744)	(−14.5275)		(6.5767)	(−40.5935)			
Sub-sample	**TS**	0.0761	0.9024	−0.2178***	1	9.6234	−8.0793	10775	−21538	−21500
1990-2006		(10.6694)	(93.9640)	(−2.8257)		(0.6072)	(−53.4502)			
Sub-sample	**TS**	0.1171	0.8484	0.6167	1	49.9834	−8.8939	4141	−8270	−8237
2007-2014		(8.5402)	(45.4606)	(1.6359)		(0.3002)	(-24.4642)			
Panel C: GARCH-MIDAS-PPI Model										
Full Sample	**PPI**	0.1338	0.8613	−0.1398***	1	1.3764***	−7.7120	33275	−66539	−66495
1970-2014		(40.6430)	(249.2422)	(-8.2286)		(4.9714)	(-24.4094)			
Sub-sample	**PPI**	0.2678	0.7128	−0.0808***	1	1.0015***	−8.4286	10817	−21622	−21585
1970-1989		(46.318)	(81.4120)	(−4.3975)		(18.4330)	(−27.2880)			
Sub-sample	**PPI**	0.1343	0.8641	0.0914**	1	19.5730*	−6.9065	10705	−21398	−21360
1990-2006		(46.6600)	(284.59)	(2.4789)		(1.7214)	(−12.8000)			
Sub-sample	**PPI**	0.1147	0.8253	0.1147***	1	7.8595**	−8.5489	4035	−8058	−8025
2007-2014		(8.3819)	(31.6424)	(3.2959)		(2.2254)	(−78.8703)			

[a] Table 2.7 reports estimation results for the GARCH-MIDAS model with alternative macroeconomic variables. Panel A provides parameter estimates for the GARCH-MIDAS-IP model, in which the long-term volatility component τ in a MIDAS filter has been smoothed out by 24 lags (K=24) of Industrial production growth rates in a monthly frequency:
$\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(\omega_2) IP_{t-k}$

[b] Panel B and Panel C provide parameter estimates for the GARCH-MIDAS model incorporated with Term Spread and Inflation (PPI), respectively: $\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(\omega_2) TS_{t-k}$ $\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(\omega_2) PPI_{t-k}$

[c] The numbers in parentheses are robust t-statistics, where ***,**,* indicate significant levels at 1%, 5% and 10%, respectively. For one GARCH-MIDAS model, we set up restricted weighting scheme with $\omega_1 = 1$ and ω_2, governing shape of beta weights.

Table 2.8: Parameter Estimates for the GARCH-MIDAS incorporated with Macroeconomic Variables II in Japan

Sample	MIDAS filter	α	β	θ	ω_1	ω_2	m	L-Likelihood	AIC	BIC
Panel A: GARCH-MIDAS-UEM Model										
Full Sample 1970-2014	**UEM**	0.1221	0.8768	−0.0449***	1	1.0010***	0.0038	33291	−66571	−66527
		(47.6221)	(328.8445)	(−8.5875)		(265.7575)	(445.46)			
Sub-sample 1970-1989	**UEM**	0.2338	0.7233	−0.0150***	1	5.1837***	0.0008	11890	-23767	-23730
		(45.8524)	(79.5560)	(−3.8077)		(3.0656)	(7.1679)			
Sub-sample 1990-2006	**UEM**	0.0772	0.9053	0.0060**	1	19.981	0.0007	10778	−21544	−21506
		(10.6730)	(102.17)	(2.1361)		(1.6128)	(4.5406)			
Sub-sample 2007-2014	**UEM**	0.1186	0.8358	0.0091*	1	4.6497	0.0008	4256	−8501	−8467
		(9.0703)	(40.4422)	(1.6816)		(1.0690)	(70.4548)			
Panel B: GARCH-MIDAS-HS Model										
Full Sample 1970-2014	**HS**	0.1384	0.8563	−0.1738***	1	1.0010***	0.0006	−32556	−65099	−65056
		(47.2680)	(256.7100)	(−9.4569)		(34.7710)	(7.8260)			
Sub-sample 1970-1989	**HS**	0.2923	0.6963	0.2919***	1	1.7692***	0.0820	11713	−23414	−23377
		(42.2554)	(77.5987)	(4.5244)		(3.1103)	(7.5720)			
Sub-sample 1990-2006	**HS**	0.0743	0.8971	−0.5443***	1	1.0850***	0.0269	10716	−21420	−21382
		(9.5554)	(77.2182)	(−6.1105)		(12.1568)	(1.3307)			
Sub-sample 2007-2014	**HS**	0.1218	0.8202	−1.6938*	1	16.3520*	0.0009	4301	−8591	−8557
		(8.3496)	(27.0280)	(−2.9975)		(1.7956)	(−2.3945)			
Panel B: GARCH-MIDAS-Yen Model										
Full Sample 1970-2014	**Yen**	0.0743	0.9257	−0.2725***	1	4.9689***	1.74e-05	28106	−56200	−56156
		(247.7327)	(3086.3924)	(−110.1697)		(91.8104)	(80.0800)			
Sub-sample 1970-1989	**Yen**	0.2579	0.6588	0.0391***	1	1.0010***	6.28e-06	11722	−23432	−23395
		(32.8015)	(43.6460)	(10.2729)		(50.2913)	(76.4280)			
Sub-sample 1990-2006	**Yen**	0.0713	0.9286	−0.0182**	1	1.1614***	4.7e-05	10718	−21424	−21386
		(19.8464)	(284.8544)	(-2.4721)		(2.7570)	(3.1146)			
Sub-sample 2007-2014	**Yen**	0.1143	0.8409	−0.0671*	1	1.5375*	8.56e-06	4256	−8501	−8467
		(8.8107)	(42.6369)	(−1.7858)		(1.7822)	(−65.0082)			

[a] Table 2.8 reports estimation results for the GARCH-MIDAS model with two alternative specifications, UEM and HS. Panel A provides parameter estimates for the GARCH-MIDAS model, in which the MIDAS filter smooths out 24 lags (K=24) of monthly growth rates of unemployment rate: $\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(\omega_2)\Delta UEM_{t-k}$

[b] Panel B provides parameter estimates for the GARCH-MIDAS model, in which the MIDAS filter smooths out 24 lags of monthly growth rates of Housing Starts: $\tau_t = m+\theta*\sum_{k=1}^{K=24} \phi_k(\omega_2)\Delta HS_{t-k}$

[c] The numbers in parentheses are robust t-statistics, where ***,**,* indicate significant levels at 1%, 5% and 10%, respectively. For one GARCH-MIDAS model, we set up restricted weighting scheme with $\omega_1 = 1$ and ω_2, governing shape of beta weights.

Table 2.9: Parameter Estimates for GARCH-MIDAS model with Realized Volatility and Macroeconomic Variable in Japan

MIDAS	α	β	θ_{RV}	ω_{RV}	θ_X	ω_X	m	Likelihood	AIC	BIC
Full-sample 1970-2014										
RV+IP	0.1381	0.7311	0.0455***	49.438***	−0.2319***	1.7368***	2.44e-05	29504	−58993	−58935
	(18.735)	(46.277)	(27.067)	(7.7357)	(−4.5839)	(4.3188)	(7.2592)			
RV+PPI	0.1383	0.7313	0.0465***	49.288***	−0.0149***	2.192*	1.68e-05	29501	-58987	-58929
	(18.652)	(46.953)	(29.016)	(8.5049)	(-3.5347)	(1.8333)	(8.1497)			
RV+UEM	0.1364	0.7336	0.0469***	49.04***	−0.0035***	4.9675**	1.87e-05	29504	−58992	−58934
	(18.375)	(47.081)	(30.872)	(8.8127)	(−3.2487)	(2.36)	(8.2404)			
RV+HS	0.1387	0.7298	0.0485***	49.076***	−0.0046	25.371	1.24e-05	29499	−58982	−58925
	(18.795)	(46.896)	(30.854)	(8.744)	(−1.0891)	(0.7674)	(7.5874)			
RV+TS	0.1339	0.7387	0.0394***	49.503***	−0.1942***	1.2599***	7.08e-05	29521	−59026	−58968
	(18.442)	(48.702)	(24.006)	(7.1966)	(−9.6352)	(5.624)	(10.551)			
RV+Yen	0.1380	0.7306	0.0476***	49.193***	−0.0013*	7.8907	1.53e-05	29500	−58985	−58927
	(18.5650)	(46.653)	(30.917)	(8.8228)	(−1.9463)	(1.1541)	(8.0509)			
Sub-sample 1970-1989										
RV+IP	0.1820	0.5856	0.0332***	49.499***	−0.0221	34.692	1.58e-05	11963	-23910	-23860
	(10.17)	(16.484)	(15.871)	(4.8618)	(−1.0177)	(0.4781)	(9.4448)			
RV+PPI	0.1837	0.5799	0.0317***	49.495***	−0.0079**	3.9724	1.74e-05	11964	−23911	−23861
	(10.114)	(15.97)	(15.543)	(4.7274)	(−2.3274)	(1.0675)	(9.3427)			
RV+UEM	0.1774	0.5875	0.0318***	49.483***	−0.0021**	6.8459**	1.88e-05	11966	−23917	−23867
	(10.032)	(16.169)	(15.827)	(4.785)	(-2.7542)	(2.0657)	(9.6554)			
RV+HS	0.1834	0.5841	0.0322***	49.507***	0.0222	1.1757	1.61e-05	11963	−23909	−23859
	(10.188)	(16.5)	(15.14)	(4.6107)	(1.325)	(1.3323)	(9.1343)			
RV+TS	0.1840	0.5769	0.0297***	49.503***	−0.0979***	1.408***	4.39e-05	11967	−23917	−23868
	(10.12)	(15.895)	(14.478)	(4.3788)	(−4.3767)	(2.7975)	(6.3377)			
RV+Yen	0.1801	0.5889	0.0325***	49.497***	−0.0007	8.9603	1.66e-05	11963	−23910	−23860
	(10.202)	(16.552)	(15.936)	(4.8758)	(-1.2998)	(0.7796)	(9.4837)			
Sub-sample 1990-2006										
RV+IP	0.0772	0.8284	0.0354***	47.824***	−0.1358	47.329	5.76e-05	10732	−21449	−21398
	(6.7853)	(26.663)	(9.8488)	(3.282)	(−1.5529)	(0.6655)	(4.8748)			
RV+PPI	0.0773	0.825	0.0367***	47.488***	0.0153	20.669	5.28e-05	10730	−21445	−21394
	(6.7831)	(24.991)	(10.847)	(3.4415)	(0.5531)	(0.3134)	(5.1162)			
RV+UEM	0.0768	0.8233	0.0367***	48.51***	0.0096***	17.565**	4.34e-05	10735	−21455	−21404
	(6.6627)	(25.562)	(10.934)	(3.4874)	(3.1845)	(2.0206)	(4.1975)			
RV+HS	0.0751	0.8299	0.0338***	48.006***	−0.2520***	1.0211***	6.11e-05	10732	−21448	−21397
	(6.633)	(26.183)	(9.4851)	(3.2648)	(-2.7794)	(12.2910)	(5.4717)			
RV+TS	0.0754	0.8258	0.0348***	48.067***	-0.1302*	29.254	8.05e-05	10731	−21446	−21396
	(6.4353)	(25.827)	(10.206)	(3.422)	(−1.9401)	(0.2498)	(4.3605)			
RV+Yen	0.0789	0.8269	0.0365***	47.634***	0.0044**	45.401	5.02e-05	10733	−21450	−21399
	(6.7575)	(27.477)	(10.638)	(3.5296)	(2.095)	(0.8177)	(4.829)			
Sub-sample 2007-2014										
RV+IP	0.0952	0.7867	0.0317***	48.343**	0.1004	12.416	6.14e-05	4269	−8522	−8477
	(6.1746)	(19.083)	(7.5756)	(2.1871)	(1.0449)	(0.6107)	(4.2107)			
RV+PPI	0.0939	0.7828	0.0322***	48.576**	0.0619	6.699	6.25e-05	4269	-8523	-8478
	(5.9533)	(17.678)	(7.4142)	(1.9702)	(1.0436)	(0.8477)	(4.0622)			

RV+UEM	0.0965	0.7919	0.0300***	48.573*	0.0047	5.4912	6.81e-05	4268	-8521	-8476
	(6.2748)	(20.247)	(6.6405)	(1.8814)	(0.5477)	(0.4312)	(4.3096)			
RV+HS	0.0954	0.7846	0.0324***	48.395**	0.1096	4.0361	5.96e-05	4268.79	−8522	−8477
	(6.1464)	(19.272)	(7.0259)	(2.1195)	(0.8243)	(0.7442)	(3.4795)			
RV+TS	0.0962	0.7908	0.0315***	48.583*	−0.1429	25.8	7.61e-05	4268	−8521	−8476
	(6.0741)	(19.336)	(6.9175)	(1.9673)	(−0.3805)	(0.0379)	(2.3329)			
RV+Yen	0.0926	0.7809	0.0316***	49.688*	-0.0137**	30.916	6.51e-05	4272	-8528	-8484
	(5.8345)	(16.0510)	(7.4603)	(1.8952)	(-2.0552)	(1.0877)	(4.4275)			

[a] Table 2.9 reports paramter estimates for GARCH-MIDAS-RV-X models including lags of RV and macro variable X for the Japan market. In the MIDAS filter, long-term volatility component is regressed by lags of RV and monthly observations from one alternative macro variable:

$$\tau_t = m + \theta_{rv} \sum_{k=1}^{K_{rv}} \phi_k(\omega_{rv}) RV_{t-k} + \theta_x \sum_{k=1}^{K_x} \phi_k(\omega_x) X_{t-k}$$

with K=24. Parameter θ_{rv} and θ_x govern impacts of realized volatility and macro variable, respectively.

Table 2.10: Parameter Estimates for GARCH-MIDAS-RV-X: Japan (Including RV of US)

Sample	MIDAS Filter	α	β	θ_{RV_J}	ω_{RV_J}	θ_{RV_U}	ω_U	θ_X	ω_X	m	Likelihood	AIC	BIC
1970-2014	$RV_J + RV_U + IP_J$	0.1537	0.8109	0.0585***	2.7249***	-0.0272***	3.5563***	-0.4819***	1.6315***	7.98e-05	29522.2	-59024.5	-58952.4
		(40.197)	(136.07)	(11.269)	(8.0165)	(-5.5421)	(4.2538)	(-3.1965)	(3.2219)	(6.022)			
1970-2014	$RV_J + RV_U + PPI_J$	0.1764	0.6712	0.0497***	44.854***	-0.0016	9.8067	-0.0195***	1.436**	2.06e-05	29514.6	-59009.2	-58937.2
		(35.573)	(54.49)	(26.355)	(7.681)	(-1.0207)	(0.5908)	(-3.9513)	(2.1152)	(6.934)			
1970-2014	$RV_J + RV_U + UEM_J$	0.1525	0.8077	0.0559***	3.2508***	-0.0177***	3.2208***	-0.0053**	6.4247*	5.33e-05	29521.5	-59023.1	-58951
		(38.357)	(131.49)	(13.024)	(8.2825)	(-4.1223)	(3.035)	(-2.2578)	(1.9415)	(6.0423)			
1970-2014	$RV_J + RV_U + TS_J$	0.1677	0.6829	0.0390***	48.792***	0.0021	27.694	-0.2496***	1.0131***	8.57e-05	29539.7	-59059.4	-58987.4
		(33.801)	(54.032)	(19.45)	(5.4531)	(1.2484)	(0.5274)	(-9.8174)	(16.692)	(10.598)			
1970-2014	$RV_J + RV_U + HS_J$	0.1533	0.8098	0.0600***	3.0555***	-0.0225***	3.5267***	0.0194*	24.114*	5.08e-05	29523.1	-59026.3	-58954.2
		(41.044)	(135.4)	(12.722)	(8.3994)	(-4.9624)	(3.5635)	(1.9099)	(1.4817)	(5.8926)			
1970-2014	$RV_J + RV_U + Yen_J$	0.1783	0.6720	0.0523***	38.747***	-0.0031***	12.164	-0.0012	9.4091	1.85e-05	29511.2	-59002.5	-58930.4
		(38.225)	(55.123)	(30.696)	(8.9052)	(-1.9395)	(1.1822)	(-1.7078)	(0.9884)	(6.5073)			
1970-2014	$RV_J + RV_U + PC1_J$	0.1522	0.8075	0.0502***	3.5676***	-0.0167***	3.0932***	0.1818***	1.2283**	6.37e-05	29520.7	-59021.4	-58949.4
		(36.625)	(127.7)	(10.5)	(7.5315)	(-3.7739)	(3.0616)	(3.5718)	(2.8757)	(6.3933)			
Average				0.0526		-0.0176							
1970-1989	$RV_J + RV_U + IP_J$	0.2863	0.4757	0.0392***	47.548***	-0.0046**	17.485	-0.1049	1.4321	2.92e-05	12005.4	-23990.8	-23928.4
		(36.445)	(18.003)	(13.485)	(4.9001)	(-2.7118)	(1.7242)	(-1.5904)	(1.581)	(5.6043)			
1970-1989	$RV_J + RV_U + PPI_J$	0.2859	0.4697	0.0359***	48.103***	-0.0030**	13.414	-0.0124***	3.0997	2.63e-05	12006.7	-23993.3	-23930.9
		(36.898)	(17.779)	(12.677)	(4.6334)	(-1.756)	(1.1776)	(-2.7411)	(1.3788)	(8.3054)			
1970-1989	$RV_J + RV_U + UEM_J$	0.2809	0.4769	0.0372***	47.678***	-0.0036**	13.672	-0.0026**	5.944*	2.78e-05	12008.2	-23996.3	-23933.9
		(36.074)	(17.947)	(14.308)	(4.8326)	(-2.187)	(1.3067)	(-2.3701)	(1.9114)	(8.2926)			
1970-1989	$RV_J + RV_U + TS_J$	0.2807	0.4745	0.0309***	48.568***	-0.0023	9.5239	-0.1584***	1.0891***	6.94e-05	12012.1	-24004.3	-23941.9
		(35.609)	(18.164)	(10.261)	(3.7249)	(-1.3174)	(0.8878)	(-4.9914)	(4.004)	(7.0705)			
1970-1989	$RV_J + RV_U + HS_J$	0.2822	0.4718	0.0346***	47.895***	-0.0032*	18.783	0.0305**	6.2404**	2.55e-05	12008.8	-23997.5	-23935.1
		(36.239)	(17.334)	(11.475)	(4.2795)	(-1.6803)	(1.0652)	(2.2519)	(1.9908)	(8.3175)			

Continued on next page

1970-1989	$RV_J + RV_U + Yen_J$	0.2856	0.4745	0.0386***	47.733***	-0.0037**	16.373	-0.0009	9.9742	2.47e-05	12005.2	-23990.4	-23928
		(36.963)	(17.931)	(14.417)	(5.036)	(-2.1237)	(1.351)	(-1.3106)	(0.8294)	(7.9647)			
1970-1989	$RV_J + RV_U + PC1_J$	0.2841	0.4698	0.0336***	48.294***	-0.0028*	12.142	0.0853***	1.1105**	3.13e-05	12008.7	-23997.4	-23935
		(36.147)	(17.724)	(11.634)	(4.2333)	(-1.7304)	(1.1274)	(3.857)	(2.9807)	(8.5356)			
Average				0.0357		-0.0033							
1990-2006	$RV_J + RV_U + IP_J$	0.0669	0.8307	0.0302***	49.811**	0.0162***	1.6796*	-0.3625***	5.4097*	4.53e-05	10803.7	-21587.4	-21524
		(6.0684)	(23.362)	(7.9692)	(2.7492)	(3.8637)	(1.7275)	(-2.9083)	(1.8782)	(4.1314)			
1990-2006	$RV_J + RV_U + PPI_J$	0.0681	0.8276	0.0317***	49.026***	0.0190***	2.0798	0.15139*	1.0704**	4.51e-05	10801.9	-21583.7	-21520.3
		(6.1351)	(23.46)	(8.8061)	(3.0959)	(4.2887)	(1.7939)	(1.9495)	(2.4287)	(3.8843)			
1990-2006	$RV_J + RV_U + UEM_J$	0.0657	0.8344	0.0317***	49.206***	0.0179***	1.5248**	0.0078***	28.675*	2.69e-05	10808.1	-21596.2	-21532.9
		(5.8056)	(23.463)	(8.6772)	(3.0977)	(4.0269)	(2.0014)	(3.3806)	(1.3024)	(2.799)			
1990-2006	$RV_J + RV_U + TS_J$	0.0678	0.8305	0.0324***	48.926**	0.0195***	1.4568*	0.1397	1.1456	4.20e-05	10801.2	-21582.4	-21519
		(6.3119)	(23.136)	(9.0109)	(3.0402)	(3.7812)	(1.8205)	(1.5405)	(0.5396)	(0.1967)			
1990-2006	$RV_J + RV_U + HS_J$	0.0687	0.8290	0.0325***	49.785**	0.0114**	29.577	-0.4608***	1.0284***	6.77e-05	10808.4	-21596.8	-21533.4
		(5.8697)	(22.066)	(6.8685)	(2.2958)	(2.7281)	(0.7340)	(-4.1718)	(24.41)	(5.0088)			
1990-2006	$RV_J + RV_U + Yen_J$	0.0687	0.8290	0.0325***	48.947***	0.0163***	1.8635	0.0006	10.68	3.36e-05	10800.5	-21581	-21517.6
		(5.9957)	(23.96)	(8.9631)	(3.3144)	(3.7464)	(1.5844)	(0.2483)	(0.1331)	(3.3931)			
1990-2006	$RV_J + RV_U + PC1_J$	0.0686	0.8288	0.0323***	48.966***	0.0161***	1.906	0.0407	7.4243	3.47e-05	10800.7	-21581.3	-21517.9
		(6.3021)	(24.184)	(8.8724)	(3.2468)	(3.5918)	(1.5215)	(0.4243)	(0.1884)	(3.4443)			
Average				0.0319		0.0166							
2007-2014	$RV_J + RV_U + IP_J$	0.1086	0.8174	0.0244***	9.7097**	-0.0032	2.4334	0.3355**	13.133	0.0001	4306.41	-8592.83	-8537.03
		(8.7917)	(27.068)	(3.2921)	(2.1328)	(-0.4306)	(0.3335)	(2.2128)	(1.3656)	(4.1563)			
2007-2014	$RV_J + RV_U + PPI_J$	0.1084	0.7731	0.0085	41.454	0.0207**	23.86	0.1915**	5.2426**	0.0001	4307.93	-8595.86	-8540.06
		(8.5563)	(19.368)	(1.3223)	(0.4968)	(2.3739)	(1.2603)	(2.1191)	(2.0555)	(5.8039)			
2007-2014	$RV_J + RV_U + UEM_J$	0.1139	0.7917	0.0124**	41.784	0.0089	27.182	0.0060	3.5672	0.0001	4303.99	-8587.99	-8532
		(8.7127)	(23.035)	(2.4199)	(0.8295)	(1.2248)	(0.4625)	(0.5370)	(0.4210)	(5.933)			

Continued on next page

2007-2014	$RV_J + RV_U + TS_J$	0.1143	0.7892	0.0133**	47.908	0.0096	9.829	-0.2811	2.8026	0.0001	4303.69	-8587.38	-8531.58
		(8.5779)	(22.231)	(2.5803)	(0.8325)	(1.3989)	(0.9429)	(-0.4547)	(0.1216)	(2.3178)			
2007-2014	$RV_J + RV_U + HS_J$	0.1074	0.8258	0.0151**	4.1023	0.0123	26.967	0.3683**	5.0301*	9.34e-05	4308.39	-8596.78	-8540.99
		(8.7832)	(32.019)	(1.9535)	(1.4597)	(1.4551)	(0.6358)	(2.2947)	(1.8273)	(3.4528)			
2007-2014	$RV_J + RV_U + Yen_J$	0.1089	0.7667	0.0087	10.074	0.0229**	27.756	-0.0466***	5.9556**	0.0001	4310.91	-8601.82	-8546.02
		(8.8672)	(18.487)	(1.2639)	(0.9889)	(2.5099)	(1.2398)	(-3.0668)	(2.3661)	(5.2738)			
2007-2014	$RV_J + RV_U + PC1_J$	0.0998	0.7901	0.0098	8.1227	0.0235**	27.664	-1.1469**	5.045**	0.0002	4312.93	-8605.85	-8550.05
		(8.6694)	(21.233)	(1.5295)	(1.035)	(2.5606)	(1.2532)	(-2.4024)	(2.5136)	(4.9322)			
Average				0.0132		0.0224							

Table 2.10 reports parameter estimates for the GARCH-MIDAS-RV-X models including lags of local RV and macro variable in the Japan market as well as global RV obtained from the U.S market. In the MIDAS filter, the long-term volatility component:

$$\tau_t = m + \theta_{rv}^{UK} \sum_{k=1}^{K} \phi_k(\omega_{rv}^{UK}) RV_{t-k}^{UK} + \theta_{rv}^{US} \sum_{k=1}^{K} \phi_k(\omega_{rv}^{US}) RV_{t-k}^{US} + \theta_x \sum_{k=1}^{K_x} \phi_k(\omega_{1,x}, \omega_{2,x}) X_{t-k}$$

Parameter θ_{rv}^{UK} accounts for the impact of local RV in the UK market. θ_{rv}^{UK} measures the spillover effect that from U.S into Japan market.

Blank Page

Table 2.11: Parameter Estimates for GARCH-MIDAS incorpated with Macroeconomic Variables I in UK

Sample	MIDAS filter	α	β	θ	ω_1	ω_2	m	L-Likelihood	AIC	BIC
				Panel A: GARCH-MIDAS-IP Model						
Full Sample	**IP**	0.0692	0.9308	−0.7217***	1	4.9991***	−0.3773	26490	−52968	−52924
1970-2014		(76.4868)	(1029.4539)	(-65.0043)		(70.0688)	(−1.8223)			
Sub-sample	**IP**	0.0680	0.9320	−0.4345***	1	4.9990***	−0.1314	9528	−19045	−19007
1970-1984		(35.0670)	(480.3149)	(-25.3135)		(26.2200)	(−0.6036)			
Sub-sample	**IP**	0.1011	0.8764	−0.0454**	1	1.0010***	−9.2230	17235	−34458	−34418
1985-2006		(15.6917)	(108.2209)	(−2.2952)		(39.3547)	(−82.4351)			
Sub-sample	**IP**	0.0973	0.8997	−0.1054*	1	3.3551	−7.7859	4964	−9915	−9882
2007-2014		(7.9590)	(69.3764)	(-1.8844)		(1.6210)	(-15.9077)			
				Panel B: GARCH-MIDAS-TS Model						
Full Sample	**TS**	0.0957	0.8885	−0.0676**	1	1.0011***	−9.0370***	36401	−72790	−72746
1970-2014		(22.7639)	(201.2582)	(−2.4733)		(4.0901)	(−100.2728)			
Sub-sample	**TS**	0.0984	0.9005	−0.1448	1	1.2014	−6.6827***	11102	−22192	−22154
1970-1984		(14.4927)	(132.1581)	(−1.2841)		(0.6901)	(−11.6026)			
Sub-sample	**TS**	0.1003	0.8700	−0.1519***	1	1.0419***	−9.3542***	17242	−34472	−34432
1985-2006		(15.9057)	(105.0936)	(−4.6783)		(3.2875)	(−129.2613)			
Sub-sample	**TS**	0.0861	0.9130	−0.1959	1	16.8684	−6.5912***	5010	−10008	−9975
2007-2014		(8.3166)	(87.1692)	(-0.9557)		(0.2820)	(-8.4067)			
				Panel C: GARCH-MIDAS-PPI Model						
Full Sample	**PPI**	0.0965	0.8846	0.1173***	1	19.6527**	−9.3659	36166	−72321	−72277
1970-2014		(23.1796)	(181.3542)	(5.9298)		(2.4946)	(−116.4090)			
Sub-sample	**PPI**	0.0922	0.8623	0.2452***	1	6.1794***	−10.1595	10889	−21767	−21729
1970-1984		(10.0395)	(60.7882)	(9.2200)		(4.7090)	(-76.8303)			
Sub-sample	**PPI**	0.1680	0.8319	−0.3832***	1	1.0252***	−2.6183	16995	−33979	−33939
1985-2006		(18.2050)	(90.1639)	(−3.6621)		(9.8848)	(−6.0152)			
Sub-sample	**PPI**	0.0968	0.8768	0.9468***	1	1.4454***	−9.8526	4854	−9696	−9662
2007-2014		(6.9286)	(49.7906)	(3.7898)		(3.8079)	(−38.6655)			

[a] Table 2.11 reports estimation results for the GARCH-MIDAS model with alternative macroeconomic variables. Panel A provides parameter estimates for the GARCH-MIDAS model, in which the long-term volatility component τ has been smoothed out by 24 lags (K=24) of Industrial production growth rates in a monthly frequency:

$$\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(1, \omega_2) IP_{t-k}$$

[b] Panel B and Panel C provide parameter estimates for the GARCH-MIDAS model incorporated with Term Spread and Inflation (PPI):

$$\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(1, \omega_2) TS_{t-k} \qquad \tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(1, \omega_2) PPI_{t-k}$$

[c] The numbers in parentheses are robust t-statistics, where ***,**,* indicate significant levels at 1%, 5% and 10%, respectively. For one GARCH-MIDAS model, we set up restricted weighting scheme with $\omega_1 = 1$.

Table 2.12: Parameter Estimates for the GARCH-MIDAS incorporated with Macroeconomic Variables II in UK

Sample	MIDAS filter	α	β	θ	ω_1	ω_2	m	L-Likelihood	AIC	BIC
				Panel A: GARCH-MIDAS-UEM Model						
Full Sample 1970-2014	**UEM**	0.0961 (23.1541)	0.8866 (188.0205)	0.0057*** (3.6449)	1	15.7905 (1.5733)	0.0001 (13.0201)	36327	−72642	−72598
Sub-sample 1970-1984	**UEM**	0.0851 (11.6331)	0.8962 (112.3988)	0.0062*** (3.0860)	1	15.3764 (1.2711)	0.0001 (7.1452)	11043	−22075	−22037
Sub-sample 1985-2006	**UEM**	0.1038 (16.5890)	0.8720 (107.3848)	−0.0040 (−1.4254)	1	4.1233 (0.4722)	0.0001 (10.8963)	17234	−34457	−34417
Sub-sample 2007-2014	**UEM**	0.0978 (4.8235)	0.8435 (26.4587)	0.0309*** (4.5731)	1	11.0624** (2.2550)	0.0001 (8.6024)	3863	−7714	−7680
				Panel B: GARCH-MIDAS-HS Model						
Full Sample 1970-2014	**HS**	0.1316 (29.1700)	0.8684 (192.4400)	−0.1728*** (−2.9760)	1	1.0728*** (2.8484)	0.0007 (9.1350)	35603	−71193	−71149
Sub-sample 2007-1984	**HS**	0.0630 (18.2420)	0.9370 (264.9100)	−0.1778* (−1.7787)	1	7.4670 (0.9728)	0.0006 (4.3776)	10882	−21751	−21714
Sub-sample 1985-2006	**HS**	0.1148 (16.6310)	0.8852 (128.2200)	−0.2226** (−1.9796)	1	1.0713* (1.7574)	0.0006 (5.5680)	16453	−32894	−32854
Sub-sample 2007-2014	**HS**	0.0995 (7.0517)	0.8831 (52.8580)	−0.1844* (−1.7284)	1	1.0178*** (4.9165)	0.0006 (2.8142)	4851	−9690	−9656
				Panel C: GARCH-MIDAS-Exchange Rate Model						
Full Sample 1970-2014	**Exchange**	0.0548 (31.932)	0.9452 (609.18)	−1.0506*** (−11.753)	1	1.0027*** (38.527)	0.0004 (7.1930)	35923	−71833	−71789
Sub-sample 1970-1984	**Exchange**	0.0713 (17.3590)	0.9286 (255.6400)	−0.6803*** (−3.6981)	1	4.7977** (2.4328)	0.0002 (1.6458)	10781	−21550	−21513
Sub-sample 1985-2006	**Exchange**	0.0797 (21.7370)	0.9203 (269.5000)	−0.7215*** (−3.3964)	1	1.6612*** (3.7051)	0.0006 (8.1734)	16938	−33864	−33824
Sub-sample 2007-2014	**Exchange**	0.0957 (7.1764)	0.8997 (65.5660)	−0.2021 (−0.2913)	1	2.3593 (0.3100)	0.0006 (−0.4415)	4839	−9666	−9632

[a] Table 2.12 reports estimation results for the GARCH-MIDAS model with alternative specifications, UEM, HS and Exchange rate (USD/GBP). Panel A provides parameter estimates for the GARCH-MIDAS model, in which the MIDAS filter smooths out 24 lags (K=24) of monthly growth of unemployment rate:

$$\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(\omega_2)\Delta UEM_{t-k}$$

[b] Panel B provides parameter estimates for the GARCH-MIDAS model, in which the MIDAS smooths out 24 lags of monthly growth rates of Housing Starts:

$$\tau_t = m + \theta * \sum_{k=1}^{K=24} \phi_k(\omega_2)\Delta HS_{t-k}$$

[c] Panel C provides parameter estimates for the GARCH-MIDAS model specified with Exchange rate. The numbers in parentheses are robust t-statistics, where ***,**,* indicate significant levels at 1%, 5% and 10%, respectively. For one GARCH-MIDAS model, we set up restricted weighting scheme with $\omega_1 = 1$ and ω_2, governing shape of beta weights.

Table 2.13: Parameter Estimates for GARCH-MIDAS model with Realized Volatility and Macroeconomic Variable in UK

MIDAS	α	β	θ_{RV}	ω_{RV}	θ_X	ω_X	m	Likelihood	AIC	BIC
Full-sample1970-2014										
RV+IP	0.1101	0.7762	0.0347***	29.9955***	−0.0031	7.2727	2.33e-05	36447	−72879	−72820
	(16.8971)	(49.6785)	(24.7625)	(9.8371)	(−1.3904)	(0.9297)	(3.4235)			
RV+PPI	0.1096	0.7761	0.0337***	30.2629***	0.0208**	9.0344	2.07e-05	36448	−72880	−72821
	(16.6199)	(49.4703)	(22.4481)	(9.4968)	(2.0162)	(0.6840)	(10.857)			
RV+UEM	0.1096	0.7755	0.0337***	31.5301***	0.0019**	9.5324	2.36e-05	36449	−72882	−72823
	(16.8059)	(49.1876)	(23.0360)	(8.9280)	(2.4167)	(1.1240)	(12.2940)			
RV+TS	0.1099	0.7763	0.0348***	28.9850***	−0.0085	5.6074	2.36e-05	36447	−72878	−72819
	(16.7736)	(49.7636)	(24.4685)	(10.2058)	(−0.9988)	(0.2605)	(12.294)			
RV+FX	0.1096	0.7762	0.0345	30.6045	0.1076	1.2937	0.0006	36448	−72880	−72821
	(16.7180)	(49.2120)	(24.4880)	(9.7065)	(2.1840)	(0.4659)	(8.3058)			
RV+HS	0.1095	0.7917	0.0332***	38.3090***	−0.0046**	1.0643***	2.54E-05	35355	−70693	−70634
	(16.3380)	(55.8100)	(21.973)	(9.2621)	(−2.1174)	(2.7784)	(11.7740)			
Sub-sample 1970-1984										
RV+IP	0.1139	0.7658	0.0352***	17.7814***	−0.0013	3.3673	2.31e-05	11065	−22113	−22063
	(8.2028)	(22.7472)	(11.3774)	(3.4429)	(−0.3019)	(0.2209)	(4.3454)			
RV+PPI	0.1121	0.7478	0.0277***	28.3823***	0.0794***	7.0452*	1.04e-05	11069	−22123	−22073
	(7.7224)	(19.0637)	(7.6433)	(3.1878)	(3.1282)	(1.7947)	(1.7559)			
RV+UEM	0.1140	0.7647	0.0345***	18.3509***	0.0009	7.0793	2.31e-05	11065	−22114	−22064
	(8.0731)	(22.2692)	(10.0895)	(3.4456)	(0.7806)	(0.3652)	(4.2491)			
RV+TS	0.1141	0.7657	0.0351***	17.9119***	0.0137	7.8481	2.31e-05	11065	−22114	−22064
	(8.1127)	(22.6343)	(11.3627)	(3.5414)	(0.6288)	(0.1969)	(4.2491)			
RV+FX	0.1166	0.7967	0.0307***	15.8570***	0.0012	9.8792	1.15e-05	11065	−22113	−22063
	(8.6345)	(30.8040)	(8.4932)	(3.3892)	(1.2346)	(0.1613)	(0.6317)			
RV+HS	0.1204	0.7829	0.0319***	9.2661***	−0.0029	15.7660	2.84e-05	10887	−21757	−21707
	(9.2674)	(30.6390)	(9.7257)	(3.7937)	(−1.2284)	(0.3159)	(4.4401)			
Sub-sample 1985-2006										
RV+IP	0.1077	0.7379	0.0339***	48.9864***	−0.0205***	1.0032***	2.49e-05	17297	−34578	−34525
	(11.7988)	(29.4636)	(18.3645)	(4.9539)	(−3.5066)	(19.7731)	(10.883)			
RV+PPI	0.1094	0.7390	0.0330***	48.9628***	−0.0632***	5.6503	2.79e-05	17296	−34576	−34523
	(12.6192)	(30.3478)	(17.7420)	(4.7716)	(−2.6723)	(1.0149)	(8.6115)			
RV+UEM	0.1099	0.7363	0.0341***	48.8796***	0.0008	7.9153	2.103e-05	17294	−34572	−34519
	(12.4310)	(29.6068)	(18.3882)	(4.9608)	(0.7828)	(0.3193)	(10.248)			
RV+TS	0.1066	0.7359	0.0315***	49.0672***	−0.0389***	7.0147	2.10e-05	17300	−34583	−34530
	(12.1758)	(27.9073)	(17.1450)	(4.7734)	(-4.0671)	(0.9864)	(10.248)			
RV+FX	0.1247	0.7921	0.0254***	12.9140***	0.0245	3.2974	−5.48E-06	17068	−34120	−34067
	(15.1530)	(44.4680)	(9.7692)	(4.2986)	(1.5733)	(0.5582)	(−0.2095)			
RV+HS	0.1248	0.7913	0.0253***	12.9250***	0.0048**	2.9021	3.46e-05	17069	−34121	−34068
	(15.0190)	(44.3500)	(9.5886)	(4.2794)	(2.0853)	(1.5192)	(8.8758)			
Sub-sample 2007-2014										
RV+IP	0.1110	0.7749	0.0339***	27.8893**	−0.0097	2.8563	2.73e-05	4972	−9929	−9884
RV+PPI	0.1100	0.7786	0.0316***	24.218**	0.1562	5.6728	2.04e-05	4974	−9932	−9887
	(5.7735)	(18.426)	(1.4957)	(0.7186)	(6.7298)	(2.6929)	(2.7861)			

RV+UEM	0.1063	0.7756	0.0239***	33.8279*	0.0135***	5.5872	4.73e-05	4976	−9936	−9891
	(5.5862)	(16.9119)	(4.7479)	(1.7301)	(2.9567)	(1.6157)	(4.4399)			
RV+TS	0.1065	0.7774	0.0333***	26.5266***	0.1369	1.8625	4.73e-05	4973	−9931	−9886
	(5.5561)	(17.9704)	(7.3161)	(2.6933)	(1.6368)	(1.2921)	(4.4399)			
RV+FX	0.0949	0.7694	0.0351***	47.3460**	−0.0228	1.4485	0.0004	4979	−9942	−9897
	(4.8314)	(14.4210)	(8.6551)	(1.9975)	(−1.0935)	(1.4509)	(1.1454)			
RV+HS	0.1095	0.7760	0.0310***	29.1340**	−0.0058**	3.6905	3.47E-05	4974	−9931	−9886
	(5.7584)	(17.2550)	(6.3862)	(2.2953)	(−1.6967)	(1.5000)	(3.5247)			

Table 2.13 reports paramter estimates for GARCH-MIDAS-RV-X models including both lags of RV and macro variable X for the UK market. In the MIDAS filter, long-term volatility component is regressed by lags of RV and monthly observations from one alternative macro variable:

$$\tau_t = m + \theta_{rv} \sum_{k=1}^{K_{rv}} \phi_k(\omega_{rv}) RV_{t-k} + \theta_x \sum_{k=1}^{K_x} \phi_k(\omega_x) X_{t-k}$$

with K=24 (months). Parameter θ_{rv} and θ_x govern impacts of realized volatility and macro variable, respectively.

Table 2.14: Parameter Estimates for GARCH-MIDAS-RV-X (Including U.S RV): UK

Sample	MIDAS Filter	α	β	$\theta_{RV_{UK}}$	$\omega_{RV_{UK}}$	$\theta_{RV_{US}}$	ω_{US}	θ_X	ω_X	m	Likelihood	AIC	BIC
1970-2014	$RV_{UK}+RV_{US}+IP_{UK}$	0.1089	0.7576	0.0327 ***	46.259***	0.0038***	4.1734***	-0.0028	1.2427	1.95e-05	36523.1	-73026.3	-72952.8
		(16.274)	(45.742)	(19.064)	(7.152)	(2.6564)	(7.152)	(-0.9421)	(0.6225)	(9.9145)			
1970-2014	$RV_{UK}+RV_{US}+PPI_{UK}$	0.1075	0.7611	0.02988***	48.487***	0.0051***	4.153***	0.0305***	8.4289	1.59e-05	36526.7	-73033.3	-72959.8
		(15.973)	(46.278)	(14.585)	(5.1662)	(3.3431)	(2.5858)	(2.9321)	(1.0863)	(7.7998)			
1970-2014	$RV_{UK}+RV_{US}+UEM_{UK}$	0.1082	0.7593	0.0312***	47.766***	0.0046***	4.9712**	0.0018**	1.0188***	1.88e-05	36524.7	-73029.4	-72955.9
		(16.083)	(46.097)	(17.18)	(6.7397)	(2.9881)	(2.2203)	(2.1902)	(4.9324)	(10.44)			
1970-2014	$RV_{UK}+RV_{US}+TS_{UK}$	0.1087	0.7575	0.0329***	45.902***	0.0036***	3.6955*	-0.0079	5.6588	1.95e-05	36523.3	-73026.5	-72953
		(16.208)	(45.529)	(19.658)	(7.3459)	(2.6512)	(1.9325)	(-1.0933)	(0.26161)	(10.598)			
1970-2014	$RV_{UK}+RV_{US}+HS_{UK}$	0.1086	0.7579	0.0325***	46.736***	0.0039***	3.675**	-0.0007	8.3674	1.88e-05	36523.1	-73026.2	-72952.6
		(16.276)	(45.693)	(19.265)	(6.9473)	(2.9215)	(2.0997)	(-0.7266)	(0.2781)	(10.087)			
1970-2014	$RV_{UK}+RV_{US}+FE_{UK}$	0.1079	0.7592	0.0317***	47.997***	0.0044***	3.6493**	0.0011**	1.321	0.0006	36525	-73031	-72957
		(16.1060)	(45.7430)	(18.1270)	(6.2522)	(3.1472)	(2.3374)	(2.5588)	(0.5491)	(8.1521)			
1970-2014	$RV_{UK}+RV_{US}+PC1_{UK}$	0.1080	0.7606	0.0305***	48.297***	0.0048***	4.218**	0.0455**	9.3158	2.08e-05	36525.4	-73030.9	-72957.4
		(16.161)	(46.528)	(15.123)	(5.1902)	(3.1748)	(2.4548)	(2.3919)	(1.0116)	(9.748)			
Average				0.0316		0.0043							
1970-1984	$RV_{UK}+RV_{US}+IP_{UK}$	0.1192	0.7423	0.0366***	19.636***	0.0035	4.4176	0.0009	5.208	1.50e-05	11142.9	-22265.8	-22203.1
		(8.4011)	(20.506)	(10.078)	(4.4438)	(0.5852)	(0.3767)	(0.2151)	(0.1690)	(1.5512)			
1970-1984	$RV_{UK}+RV_{US}+PPI_{UK}$	0.1177	0.7258	0.029***	29.545***	0.0021	4.1653	0.0764***	6.4537*	6.37e-06	11147.4	-22274.7	-22212
		(7.9711)	(17.749)	(7.1205)	(3.6199)	(0.4878)	(0.2879)	(2.9085)	(1.7482)	(0.8983)			
1970-1984	$RV_{UK}+RV_{US}+UEM_{UK}$	0.1191	0.7406	0.0364***	20.43***	0.0015	5.0969	0.0654	6.3486	1.79e-05	11143	-22266	-22203.3
		(8.3706)	(20.344)	(9.8938)	(3.9963)	(0.2932)	(0.1541)	(0.5121)	(0.2494)	(2.4298)			
1970-1984	$RV_{UK}+RV_{US}+TS_{UK}$	0.1198	0.7412	0.0358***	20.114***	0.0053	4.2854	0.0206	7.7794	1.05e-05	11143.3	-22266.6	-22203.8
		(8.45)	(20.552)	(9.5398)	(4.5394)	(0.8941)	(0.5717)	(0.8630)	(0.2577)	(1.0004)			
1970-1984	$RV_{UK}+RV_{US}+HS_{UK}$	0.1187	0.7422	0.0365***	19.8***	0.0035	3.2844	-0.0014	7.9814	1.56e-05	11143.2	-22266.3	-22203.6
		(8.4132)	(20.319)	(9.9698)	(4.5382)	(0.7052)	(0.4221)	(-0.6955)	(0.2279)	(2.3219)			

Continued on next page

1970-1984	$RV_{UK} + RV_{US} + FE_{UK}$	0.1189	0.7415	0.0360***	20.7870***	0.0069	1.4323	0.0079	7.1730	0.0006	11143.2	−22266.5	−22203.8
		(8.3790)	(20.2070)	(9.7335)	(4.5547)	(0.8631)	(0.8855)	(0.8056)	(0.1078)				
1970-1984	$RV_{UK} + RV_{US} + PC1_{UK}$	0.1184	0.7283	0.0303***	28.389***	0.0002	5.2192	0.1323**	9.8091	2.25e-05	11146.7	-22273.3	-22210.6
		(8.3171)	(18.964)	(7.5678)	(3.4606)	(0.0379)	(0.0198)	(2.8287)	(1.6011)	(3.5074)			
Average				0.0344		0.0033							
1985-2006	$RV_{UK} + RV_{US} + IP_{UK}$	0.0982	0.7604	0.0287***	49.134***	0.0044***	2.5388**	-0.0217***	1.0415***	2.22e-05	17400.9	-34781.9	-34715.6
		(11.149)	(30.955)	(11.892)	(3.5801)	(3.0877)	(2.223)	(-3.3101)	(5.7132)	(9.8125)			
1985-2006	$RV_{UK} + RV_{US} + PPI_{UK}$	0.1003	0.7618	0.0263***	49.151***	0.0053**	8.8633*	-0.0719***	3.681	2.69e-05	17401	-34782	-34715.7
		(12.052)	(32.297)	(7.5923)	(3.05)	(2.1631)	(1.8327)	(-2.8479)	(1.0433)	(7.8966)			
1985-2006	$RV_{UK} + RV_{US} + UEM_{UK}$	0.0996	0.7628	0.0254***	49.068***	0.0069***	9.2058**	0.1654	8.7384	1.91e-05	17398.9	-34777.8	-34711.5
		(11.782)	(31.777)	(7.2426)	(2.9263)	(2.7296)	(2.4005)	(1.2912)	(0.5804)	(9.143)			
1985-2006	$RV_{UK} + RV_{US} + TS_{UK}$	0.0976	0.7584	0.0259***	49.136***	0.0047**	8.924*	-0.0333***	8.6284	2.30e-05	17402.7	-34785.4	-34719.1
		(11.637)	(29.353)	(7.7523)	(3.0783)	(2.0464)	(1.7446)	(-3.6986)	(0.7399)	(9.7249)			
1985-2006	$RV_{UK} + RV_{US} + HS_{UK}$	0.1008	0.7629	0.0255***	49.018***	0.0068***	9.3015**	0.0027	2.0126	1.86e-05	17399.3	-34778.5	-34712.2
		(11.995)	(32.181)	(7.4553)	(2.9536)	(2.7442)	(2.2422)	(1.631)	(1.1918)	(8.5617)			
1985-2006	$RV_{UK} + RV_{US} + FX_{UK}$	0.1017	0.7820	0.0284***	48.3190***	0.0041***	1.0069***	−0.0042	5.4658	1.95e-05	17141	−34263	−34196
		(11.5760)	(33.3780)	(12.3770)	(3.7668)	(3.2334)	(13.8110)	(−0.0381)	(0.1348)	(1.0523)			
1985-2006	$RV_{UK} + RV_{US} + PC1_{UK}$	0.0987	0.7555	0.0271***	49.157***	0.0027	8.1551	-0.1638***	2.9724*	1.46e-05	17403.5	-34787	-34720.8
		(11.679)	(28.911)	(8.3727)	(3.3207)	(1.1616)	(1.0171)	(-4.1704)	(1.7504)	(6.5048)			
Average				0.0268		0.0050							
2007-2014	$RV_{UK} + RV_{US} + IP_{UK}$	0.0995	0.7833	0.0273***	31.999**	0.0069**	1.5081*	0.0525*	1.062***	1.42e-05	5023.53	-10027.1	-9970.97
		(5.4018)	(18.407)	(5.155)	(1.9723)	(2.2399)	(1.9452)	(1.982)	(3.1372)	(2.0093)			
2007-2014	$RV_{UK} + RV_{US} + PPI_{UK}$	0.1063	0.7803	0.0201**	36.813	0.0076	5.7054	0.1779	3.9222	2.09e-05	5024.17	-10028.3	-9972.24
		(5.8351)	(18.638)	(2.3707)	(1.042)	(1.2797)	(1.1875)	(1.4502)	(0.6091)				
2007-2014	$RV_{UK} + RV_{US} + UEM_{UK}$	0.1019	0.7828	0.0201**	41.168	0.0037	8.7967	0.0119*	6.8237	4.58e-05	5024.75	-10029.5	-9973.4
		(5.5861)	(17.277)	(2.0138)	(0.9624)	(0.6334)	(0.4825)	(1.9481)	(1.2112)	(3.3375)			
2007-2014	$RV_{UK} + RV_{US} + TS_{UK}$	0.1005	0.7812	0.0219***	46.113	0.0079*	4.9246*	0.1772**	1.0184***	-1.32e-05	5023.94	-10027.9	-9971.78

Continued on next page

		(5.5362)	(18.507)	(3.1561)	(0.9948)	(1.8323)	(1.8133)	(2.0896)	(4.5218)	(-0.7269)			
2007-2014	$RV_{UK} + RV_{US} + HS_{UK}$	0.1049	0.7807	0.0254***	36.661	0.0049	4.0941	-0.0035	6.3786	3.02e-05	5022.6	-10025.2	-9969.1
		(5.7962)	(18.17)	(3.4609)	(1.2865)	(1.0436)	(1.0943)	(-1.1519)	(0.7317)	(3.2335)			
2007-2014	$RV_{UK} + RV_{US} + FX_{UK}$	0.1064	0.7762	0.0258***	37.0730	0.0050	4.4345	−0.0269	1.3448	0.0004	4974	−9928.9	−9872.8
		(5.6607)	(17.4110)	(3.5167)	(1.3496)	(1.1088)	(0.8829)	(−0.8457)	(1.3746)	(0.8886)			
2007-2014	$RV_{UK} + RV_{UK} + PC1_{UK}$	0.1035	0.7804	0.0189*	40.514	0.0061	8.9772	0.3141*	5.4031	4.55e-05	5024.9	-10029.8	-9973.7
		(5.6965)	(18.112)	(1.9599)	(0.9698)	(1.014)	(0.8529)	(1.8792)	(0.6939)	(3.4461)			
Average				0.0228		0.0060							

Table 2.14 reports parameter estimates for the GARCH-MIDAS-RV-X models including lags of local RV and macro variable in the UK market as well as global RV obtained from the U.S market. In the MIDAS filter, the long-term volatility component is defined as:

$$\tau_t = m + \theta_{rv}^{UK} \sum_{k=1}^{K} \phi_k(\omega_{rv}^{UK}) RV_{t-k}^{UK} + \theta_{rv}^{US} \sum_{k=1}^{K} \phi_k(\omega_{rv}^{US}) RV_{t-k}^{US} + \theta_x \sum_{k=1}^{K} \phi_k(\omega_{1,x}, \omega_{2,x}) X_{t-k}$$

Parameter θ_{rv}^{UK} accounts for the impact of local RV in the UK market. Parameter θ_x accounts for the impact of local macroeconomic variable on UK stock market. θ_{rv}^{UK} measures the spillover effect that from U.S into UK market.

Table 2.15: U.S Volatility Spillover in 1970-2014

Coefficients	θ_L^{RV}	ω_L^{RV}	θ_{US}^{RV}	ω_{US}^{RV}	Likelihood
Panel A: Volatility Spillover from the U.S into UK					
1970-2014	0.0308***	32.1310***	0.0041**	4.2303**	36280.1
	(16.4820)	(7.4831)	(2.5190)	(1.8827)	(0.0000)
1970-1984	0.0352***	26.0770***	0.0048	5.5488	10899.3
	(9.8523)	(3.5510)	(0.9811)	(0.5154)	(0.0001)
1985-2006	0.0225***	48.2680**	0.0079***	17.9320**	17141.1
	(6.1920)	(2.4862)	(3.0603)	(2.3308)	(0.0000)
2007-2014	0.0205***	13.2510*	0.0054	1.1908	4917.17
	(3.4352)	(1.8004)	(1.5532)	(1.4213)	(0.0000)
Panel B: Volatility Spillover from the U.S into Japan					
1970-2014	0.0517***	33.4900***	−0.0043**	13.1500	33373.5
	(28.0650)	(10.3750)	(−2.3792)	(1.4758)	(0.0000)
1970-1989	0.0394***	47.0350***	−0.0033*	17.3220	15850.9
	(15.8150)	(5.4313)	(−1.7905)	(1.1488)	(0.0000)
1990-2006	0.0204***	46.7020**	0.0205***	2.4088*	10597.2
	(5.7642)	(2.2365)	(3.8673)	(1.6733)	(0.0000)
2007-2014	0.0166***	13.1000	0.0040	1.2089	4143.38
	(2.6258)	(1.5828)	(0.7810)	(0.5109)	(0.0000)

[a] Table 2.15 reports volatility spillovers from the U.S into the UK (Panel A) and the Japan (Panel B) stock markets based on the GARCH-MIDAS model, where local volatility RV_L and global volatility RV_{US} are brought into one MIDAS filter shown as below:

$$\tau_t = m + \theta_L \sum_{k=1}^{K} \phi_k(\omega_L) RV_{L,t-k} + \theta_{US} \sum_{k=1}^{K} \phi_k(\omega_{US}) RV_{US,t-k}$$

[b] θ_{US} accounts for the spillover effects from U.S stock volatility onto the local stock markets. ω_{US} governs weights being attached onto each lags of volatility spillover from U.S. θ_L accounts for the local effect from historical volatility onto the local UK or Japanese markets. ω_L governs weights being attached onto lags of local realized volatilities.

Table 2.16: Parameter Estimates for GARCH-MIDAS model with Principal Components

MIDAS	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{m}$	$\boldsymbol{\theta}$	$\boldsymbol{\omega_1}$	$\boldsymbol{\omega_2}$	Likelihood	AIC	BIC
Panel A: Principal Component for U.S									
PC1	0.1374	0.8626	−0.2534	−0.4314***	1	5.7700***	35600.6520	−71189.3	−71145.3
1970-2014	(31.9980)	(200.8751)	(−0.8245)	(−6.5844)		(3.5011)			
PC1	0.0470	0.9387	−9.4986	−0.3160***	1	2.0113**	11225.5	−22439.1	−22401.6
1970-1984	(7.5001)	(112.14)	(-97.935)	(3.1052)		(2.0464)			
PC1	0.0850	0.9007	-9.0182	−0.8912***	1	2.3880***	16700.3	−33388.7	−33348.9
1985-2006	(32.6267)	(211.0995)	(−79.0543)	(−7.0290)		(5.2731)			
PC1	0.1449	0.8531	−6.6857	−0.3667	1	9.6117	4932.5	−9852.9	−9819.3
2007-2014	(8.8919)	(51.0470)	(−8.7526)	(−1.6367)		(1.3605)			
Panel B: Principal Component for JAPAN									
PC1	0.1352	0.8544	0.0003	0.0152***	1	1.0010***	29231.7	−58451.5	−58408.2
1970-2014	(39.925)	(211.2800)	(4.9836)	(4.8714)		(48.8400)			
PC1	0.2621	0.6608	9.21E-05	0.0032***	1	1.2851***	11720.8	−23429.6	−23392.1
1970-1989	(44.6320)	(51.0400)	(12.2530)	(7.9542)		(4.8293)			
PC1	0.0789	0.8990	0.0002	0.0061**	1	5.1260	10585.4	−21158.9	−21120.8
1990-2006	(10.6430)	(91.4200)	(9.1524)	(2.0100)		(1.0053)			
PC1	0.1189	0.8454	0.0003	−0.0043*	1	49.8710	4143	−8274.2	−8240.7
2007-2014	(8.7706)	(42.8070)	(5.7932)	(−1.8408)		(0.8577)			
Panel C: Principal Component for UK									
PC1	0.0975	0.8832	−9.1506	0.2033***	1	30.5579**	36333.7	−72655.5	−72611.3
1970-2014	(22.8954)	(171.5795)	(−126.8185)	(8.8361)		(2.4882)			
PC1	0.0699	0.9301	−10.7505	0.8985***	1	1.7151***	11021.5	−22031	−21993
1970-1984	(17.2097)	(207.8554)	(−12.0652)	(3.4342)		(3.2096)			
PC1	0.0604	0.9396	−10.1515	−0.0415	1	1.0736	17171.5	−34331	−34291.3
1985-2006	(25.5358)	(381.3614)	(−14.4165)	(−0.0699)		(0.1246)			
PC1	0.0968	0.9016	−7.0168	0.8705**	1	5.2061	4964.4	−9916.9	−9883.2
2007-2014	(8.0517)	(73.3330)	(−18.4450)	(1.9866)		(0.8288)			

Table 2.16 reports estimation results for the GARCH-MIDAS model with Principal Component (PC_1), where the long-term volatility component τ_t is smoothed out by a MIDAS filter with 24 lags (K=24) of monthly first principal component:

$$\tau_t = m + \theta * \sum_{k=1}^{K} \phi_k(1, \omega_2) PC_{t-k}$$

Panels A,B and C summarize results of the first principal component for the U.S, UK and Japan markets, respectively.

Table 2.17: Parameter Estimates for GARCH-MIDAS-RV-X model with Principal Components

MIDAS	α	β	m	θ_L	ω_L	θ_X	ω_X	θ_{US}	ω_{US}	Likelihood	AIC	BIC
				Panel A: Principal Component for U.S								
RV_L+PC_1	0.0811	0.7797	2.20E-05	0.0372***	49.1980***	−0.1103**	2.0728***			35939.2	−71862.5	−71803.8
1970-2014	(30.4630)	(47.6110)	(10.8820)	(25.6320)	(6.9159)	(−5.1264)	(3.4295)					
RV_L+PC_1	0.0314	0.9037	2.47E-05	0.0319***	49.6630**	−0.0815***	3.1525*			11241.5	−22466.9	−22417
1970-1984	(4.0017)	(20.7100)	(4.3924)	(7.3257)	(1.8874)	(−3.1182)	(1.7835)					
RV_L+PC_1	0.0513	0.7999	2.04E-05	0.0377***	46.2887***	−0.2021***	2.2840			16855	−33694	−33641.1
1985-2006	(5.9055)	(19.1681)	(9.2150)	(24.1407)	(6.8355)	(−6.6046)	(3.9028)					
RV_L+PC_1	0.1282	0.8069	3.91E-05	0.0271***	11.9072***	0.0270	18.0838			4977.9	−9939.8	−9894.9
2007-2014	(7.2110)	(24.8306)	(2.6167)	(4.9657)	(2.9326)	(0.1703)	(0.1166)					
				Panel B: Principal Component for JAPAN								
RV_L+PC_1	0.1309	0.7569	2.75E-05	0.0418***	41.4290***	0.014***	1.2103***			29311.1	−58606.1	−58548.5
1970-2014	(18.7630)	(62.6480)	(8.9045)	(23.6100)	(8.8420)	(6.9208)	(5.1026)					
RV_L+PC_1	0.2565	0.6330	3.36E-05	0.0291***	5.4097***	0.0063**	4.9371			11731.9	−23447.8	−23397.9
1970-1989	(29.9590)	(40.5960)	(5.5807)	(7.4167)	(5.0180)	(2.0638)	(1.0165)					
RV_L+PC_1	0.0823	0.8399	1.37E-05	0.0268***	46.0030**	0.0026*	7.2945			10592.4	−21168.9	−21118.2
1990-2006	(7.6228)	(30.6730)	(6.0273)	(6.3955)	(2.3297)	(1.8879)	(0.8723)					
RV_L+PC_1	0.1196	0.7868	0.0001	0.0230***	8.8766*	−0.0104*	5.4842*			4148.76	−8281.5	−8236.8
2007-2014	(8.1468)	(20.9440)	(4.4770)	(4.3968)	(1.8573)	(−1.8633)	(1.7753)					

$RV_L+RV_{US}+PC_1$	0.1521	0.8083	6.39E-05	0.0512***	3.4799***	0.0168***	1.3508**	−0.0178**	3.1865***	29520.8	−59021.6	−58949.5
1970-2014	(36.9040)	(129.28)	(6.3502)	(10.3240)	(7.5561)	(3.2649)	(2.5655)	(−3.9156)	(3.1474)			
$RV_L+RV_{US}+PC_1$	0.2839	0.4699	3.14E-05	0.0332***	48.3780***	0.0089***	1.1396	−0.0024	8.8790	12008.5	−23997.1	−23934.7
1970-1989	(36.0410)	(17.7120)	(8.4323)	(11.4620)	(4.1389)	(3.9021)	(2.8563)	(−1.4825)	(1.0071)			
$RV_L+RV_{US}+PC_1$	0.0683	0.8290	3.51E-05	0.0324***	49.7790***	0.0048	6.0861	0.0155***	1.8004	10800.6	−21581.3	−21517.9
1990-2006	(6.2839)	(24.1690)	(3.4199)	(8.9255)	(3.1919)	(0.3888)	(0.2000)	(3.2638)	(1.4664)			
$RV_L+RV_{US}+PC_1$	0.1031	0.7835	0.0002	0.0066	39.5820	−0.1536***	3.7459***	0.0250***	9.4020**	4310.9	−8601.7	−8545.9
2007-2014	(8.5564)	(21.1120)	(5.8156)	(1.1726)	(0.4851)	(−2.8754)	(3.3511)	(3.0758)	(2.0652)			

Panel C: Principal Component for UK

RV_L+PC_1	0.1091	0.7764	2.54E-05	0.0331***	32.7295***	0.0579***	46.3980			36451.1	−72886.1	−72827.3
1970-2014	(16.8192)	(49.5026)	(10.868)	(21.4634)	(8.5638)	(3.2629)	(0.6755)					
RV_L+PC_1	0.1147	0.7420	2.61E-05	0.0271***	33.7527***	0.1431***	39.2856			11071.3	−22126.6	−22076.5
1970-1984	(8.2009)	(19.4823)	(11.0540)	(7.5310)	(2.7563)	(4.0693)	(1.1617)					
RV_L+PC_1	0.1074	0.7364	2.78E-05	0.0299***	49.1048***	−0.1924***	2.0403**			17301.1	−34586.3	−34533.2
1985-2006	(12.5563)	(28.4402)	(5.6087)	(16.1693)	(4.4931)	(−4.6350)	(2.2685)					
RV_L+PC_1	0.1067	0.7747	2.78E-05	0.0263***	30.4557**	0.3637***	4.4901			4976.1	−9936.3	−9891.4
2007-2014	(5.6238)	(17.4815)	(5.6086)	(5.6393)	(2.0727)	(2.6538)	(0.8357)					
$RV_L+RV_{US}+PC_1$	0.1080	0.7606	2.08E-05	0.0305***	48.297***	0.0455**	9.3158	0.0048***	4.218**	36525.4	−73030.9	−72957.4
1970-2014	(16.161)	(46.528)	(9.748)	(15.123)	(5.1902)	(2.3919)	(1.0116)	(3.1748)	(2.4548)			
$RV_L+RV_{US}+PC_1$	0.1184	0.7283	2.25E-05	0.0303***	28.389***	0.1323**	9.8091	0.0002	5.2192	11146.7	−22273.3	−22210.6

1970-1984	(8.3171)	(18.964)	(3.5074)	(7.5678)	(3.4606)	(2.8287)	(1.6011)	(0.0379)	(0.0198)			
RV_L+RV_{US}+PC_1	0.0987	0.7555	1.46E-05	0.0271***	49.157***	−0.1638***	2.9724*	0.0027	8.1551	17403.5	−34787	−34720.8
1985-2006	(11.679)	(28.911)	(6.5048)	(8.3727)	(3.3207)	(−4.1704)	(1.7504)	(1.1616)	(1.0171)			
RV_L+RV_{US}+PC_1	0.1035	0.7804	4.55E-05	0.0189*	40.5140	0.3141*	5.4031	0.0061	8.9772	5024.9	−10029.8	−9973.7
2007-2014	(5.6965)	(18.112)	(3.4461)	(1.9599)	(0.9698)	(1.8792)	(0.6939)	(1.014)	(0.8529)			

Table 2.17 reports estimation results for the GARCH-MIDAS-RV-X model with Principal Component (PC_1) across U.S, UK and Japan markets.

In Panel A: long-term volatility component τ_t is smoothed out by a MIDAS filter specified with 24 lags (K=24) of month local RV_L and first principal component PC_1 in U.S market, which is defined as below:

$$\tau_t = m + \theta_{rv}^{L} \sum_{k=1}^{K} \phi_k(\omega_{rv}^{L}) RV_{t-k}^{L} + \theta_{pc} \sum_{k=1}^{K} \phi_k(\omega_{PC}) PC_{t-k}$$

In Panel B, we consider two different situations to explain the long-term volatility component τ_t:

One way τ_t is smoothed out by a MIDAS filter with local RV_L and first principal component PC_1 in Japan market:

$$\tau_t = m + \theta_{rv}^{L} \sum_{k=1}^{K} \phi_k(\omega_{rv}^{L}) RV_{t-k}^{L} + \theta_{pc} \sum_{k=1}^{K} \phi_k(\omega_{PC}) PC_{t-k}$$

Alternatively, τ_t is smoothed out by a MIDAS filter with local RV_L, first principal component PC_1 in Japan (or UK) market as well as the global RV_{US} from the U.S market.

$$\tau_t = m + \theta_{rv}^{L} \sum_{k=1}^{K} \phi_k(\omega_{rv}^{L}) RV_{t-k}^{L} + \theta_{rv}^{US} \sum_{k=1}^{K} \phi_k(\omega_{rv}^{US}) RV_{t-k}^{US} + \theta_{pc} \sum_{k=1}^{K} \phi_k(\omega_{PC}) PC_{t-k}$$

where θ_{rv}^{US} accounts for the spillover effect from U.S market onto the local market (Japan or UK).

Table 2.18: Model Confidence Set For U.S

Initial Window: 1970-2009	Jan-Jun/2010	Jul-Dec/2010	Jan-Jun/2011	Jul-Dec/2011	Jan-Jun/2012	Jul-Dec/2012
MIDAS-IP	0.1568	0.9669	1.0000	0.7494	0.6984	0.5864
MIDAS-PPI	0.1568	0.9670	0.9441	0.8873	0.7962	0.7904
MIDAS-UEM	1.0000	1.0000	0.9441	0.6433	0.4904	0.3339
MIDAS-TS	0.1568	0.9669	0.9441	1.0000	1.0000	0.9778
MIDAS-HS	0.1354	0.1494	0.1866	0.6433	0.4904	0.3340
MIDAS-PC_1	0.2581	0.9670	0.9441	0.8873	0.9485	1.0000
Initial Window: 1985-2006	Jan-Jun/2007	Jul-Dec/2007	Jan-Jun/2008	Jul-Dec/2008	Jan-Jun/2009	Jul-Dec/2009
MIDAS-IP	0.9357	0.4343	0.3803	0.3425	0.4222	0.4405
MIDAS-PPI	0.5848	0.2528	0.2457	0.3425	0.4507	0.4405
MIDAS-UEM	0.9357	0.2367	0.2457	1.0000	1.0000	1.0000
MIDAS-TS	0.0299	0.2367	0.2457	0.3425	0.4222	0.4405
MIDAS-HS	0.2704	0.1685	0.2457	0.7751	0.4222	0.4405
MIDAS-PC_1	1.0000	1.0000	1.0000	0.7751	0.7030	0.7135

Table 2.18 presents model confidence sets for our selection of GARCH-MIDAS models spicified with alternative macroeconomic variables. Our forecast horizon covers two periods, one is 2007-2009, another one is 2010-2012. For the first forecast period (2007-2009), we construct initial estimation window (1985-2006), which is rolling forward in a monthly frequency. For the second forecast period (2010-2012), the initial estimation window is set up around 1970-2009. Hence each forecasting period covers three years. We divide three-year length into equally length of sub-period with 6-month, and calcluate MCS P-values for each 6-month sub-period. * indicates that the model is in the (1-α) confidence interval of $\hat{M}^*_{1-\alpha}$ using for all comparisons, where α=0.25.

Table 2.19: Model Confidence Set For UK

Initial Window: 1970-2009	Jan-Jun/2010	Jul-Dec/2010	Jan-Jun/2011	Jul-Dec/2011	Jan-Jun/2012	Jul-Dec/2012
MIDAS-IP	0.2284	0.3705	0.3611	0.5931	0.4107	0.1702
MIDAS-PPI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
MIDAS-UEM	0.5352	0.2896	0.2747	0.2527	0.2186	0.1649
MIDAS-TS	0.1842	0.3705	0.3611	0.3194	0.2606	0.1649
MIDAS-PC_1	0.3162	0.3704	0.3611	0.5931	0.4107	0.1702
Initial Window: 1985-2006	Jan-Jun/2007	Jul-Dec/2007	Jan-Jun/2008	Jul-Dec/2008	Jan-Jun/2009	Jul-Dec/2009
MIDAS-IP	0.7637	0.4542	0.3172	0.0781	0.0247	0.0360
MIDAS-PPI	0.0471	0.7403	1.0000	0.1419	0.0400	0.0360
MIDAS-UEM	1.0000	1.0000	0.4603	0.0951	0.0271	0.0296
MIDAS-TS	0.0011	0.2832	0.3600	0.1419	0.0400	0.0360
MIDAS-PC_1	0.0011	0.2282	0.9910	1.0000	1.0000	1.0000

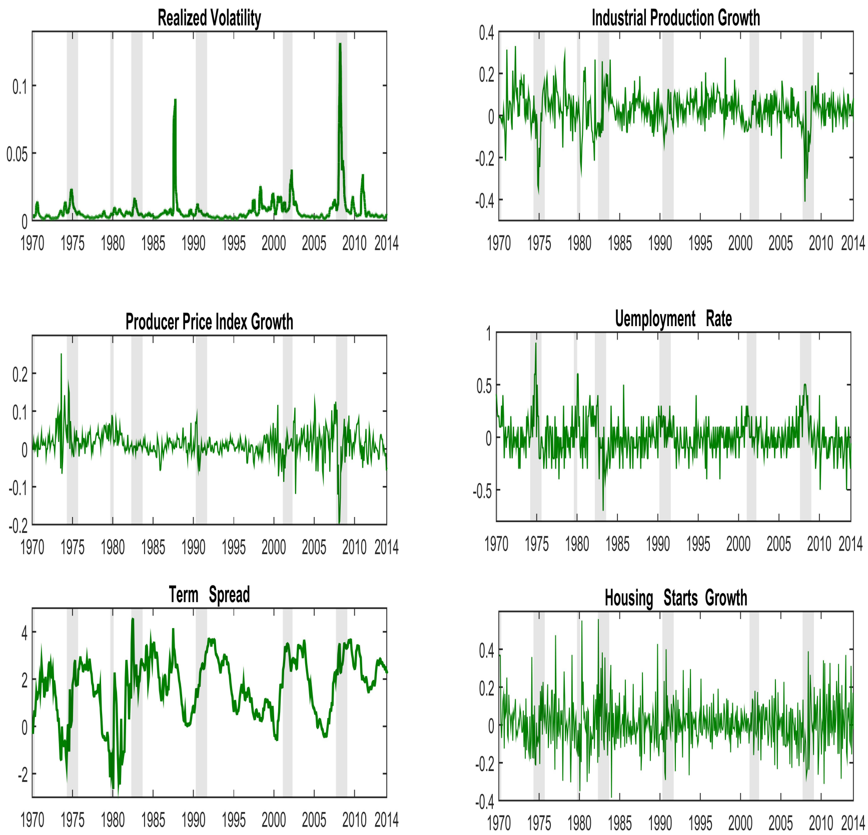
Table 2.19 presents model confidence sets for our selection of GARCH-MIDAS models spicified with alternative macroeconomic variables in the UK market. Our forecast horizon covers two periods, one is 2007-2009, another one is 2010-2012. For the first forecast period (2007-2009), we construct initial estimation window (1985-2006), which is rolling forward in a monthly frequency. For the second forecast period (2010-2012), the initial estimation window is set up around 1970-2009. Hence each forecasting period covers three years. We divide three-year length into equally length of sub-period with 6-month, and calcluate MCS P-values for each 6-month sub-period. * indicates that the model is in the (1-α) confidence interval of $\hat{M}^*_{1-\alpha}$ using for all comparisons, where α=0.25.

Table 2.20: Model Confidence Set For Japan

Initial Window: 1970-2009	Jan-Jun/2010	Jul-Dec/2010	Jan-Jun/2011	Jul-Dec/2011	Jan-Jun/2012	Jul-Dec/2012
MIDAS-IP	0.5941	0.4409	0.4040	0.3244	0.3279	0.3331
MIDAS-PPI	0.3941	0.4409	0.5293	0.4587	0.4719	0.4939
MIDAS-UEM	0.4412	0.4409	0.5293	0.3684	0.3839	0.4108
MIDAS-TS	0.5941	0.4409	1.0000	1.0000	1.0000	1.0000
MIDAS-HS	0.6316	1.0000	0.4053	0.3244	0.3279	0.3331
MIDAS-Yen	1.0000	0.4409	0.4053	0.3244	0.3279	0.3331
MIDAS-PC_1	0.5941	0.4409	0.9167	0.8538	0.8435	0.8315
Initial Window: 1985-2006	Jan-Jun/2007	Jul-Dec/2007	Jan-Jun/2008	Jul-Dec/2008	Jan-Jun/2009	Jul-Dec/2009
MIDAS-IP	1.0000	0.9553	0.8765	0.6507	0.4924	0.4951
MIDAS-PPI	0.8646	1.0000	0.4871	0.7857	0.7649	0.7708
MIDAS-UEM	0.7570	0.4598	0.3524	1.0000	1.0000	1.0000
MIDAS-TS	0.8646	0.8230	0.9442	0.6507	0.4924	0.4951
MIDAS-HS	0.1310	0.1372	0.0510	0.2018	0.2090	0.1985
MIDAS-Yen	0.8646	0.9553	1.0000	0.6507	0.6234	0.6214
MIDAS-PC_1	0.8646	0.9553	0.8765	0.7137	0.6234	0.6214

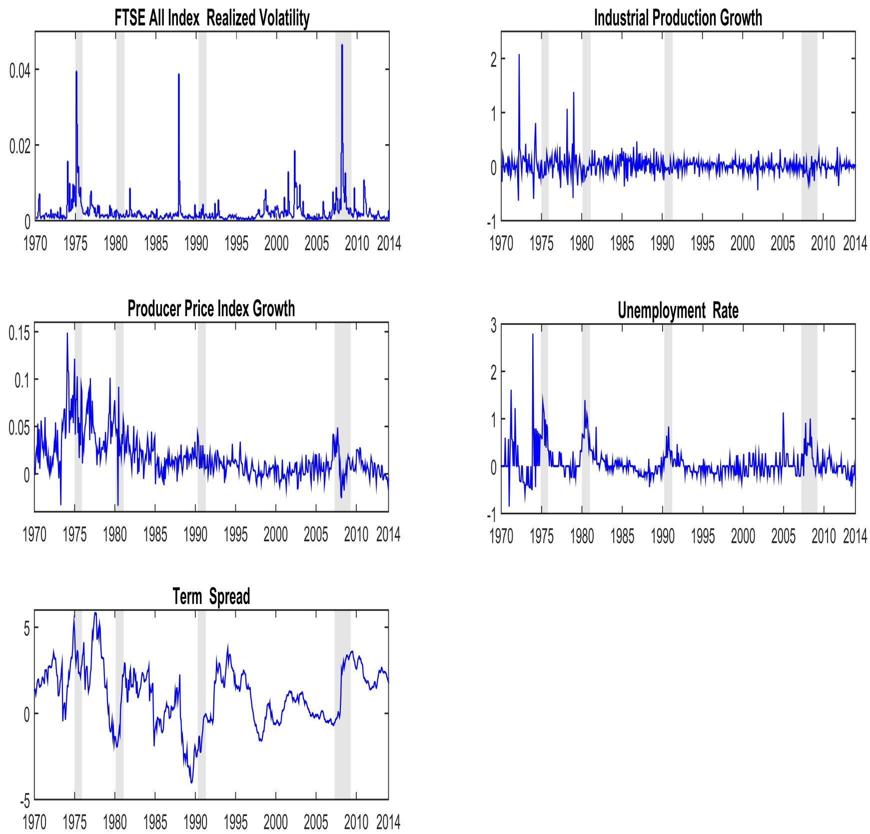
Table 2.20 presents model confidence sets for our selection of GARCH-MIDAS models spicified with alternative macroeconomic variables in the Japan. Our forecast horizon covers two periods, one is 2007-2009, another one is 2010-2012. For the first forecast period (2007-2009), we construct initial estimation window (1985-2006), which is rolling forward in a monthly frequency. For the second forecast period (2010-2012), the initial estimation window is set up around 1970-2009. Hence each forecasting period covers three years. We divide three-year length into equally length of sub-period with 6-month, and calcluate MCS P-values for each 6-month sub-period. * indicates that the model is in the (1-α) confidence interval of $\hat{M}^*_{1-\alpha}$ using for all comparisons, where α=0.25.

Figure 2.2: S&P 500 Realized Volatility and Macroeconomic Variables In The U.S Market (1970-2014)



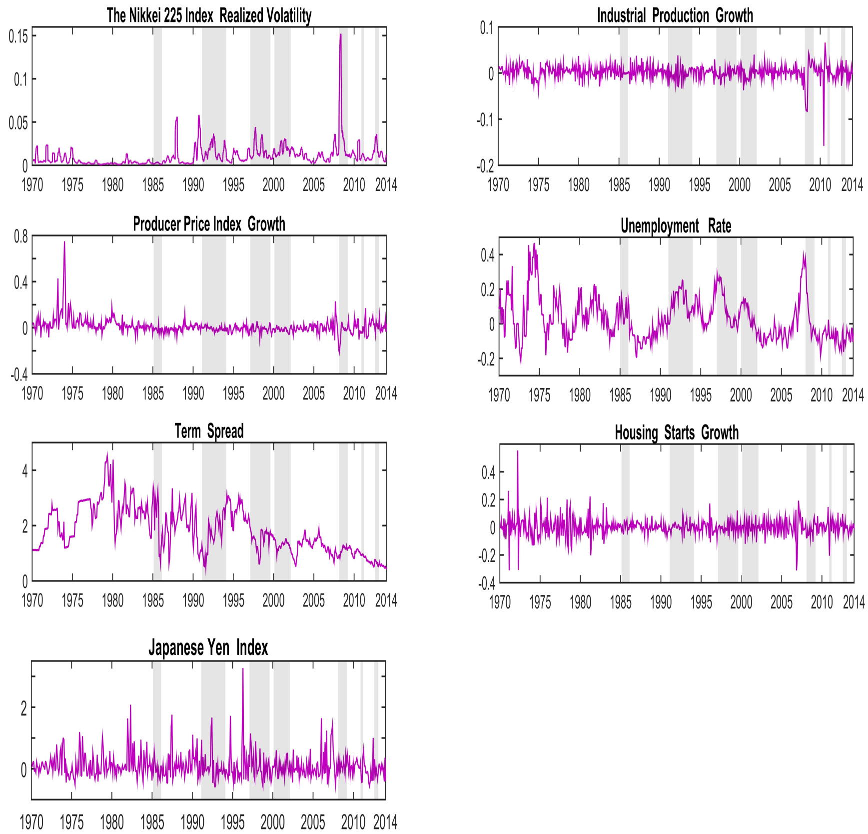
The Figure 2.2 plots Realized Volatility of S&P500 Index and annually growth rates of macroeconomic variables available in a monthly frequency, including Industrial Production (IP), Producer Price Index (PPI), Unemployment Rate (UEM), Term Spread (TS) and Housing Starts (HS). Monthly Macro data ranges from Jan/1970 to Dec/2014. Shade areas represent NBER recessions during 1970–2014.

Figure 2.3: FTSE All Share Index Realized Volatility and Macroeconomic Variables In The UK Market



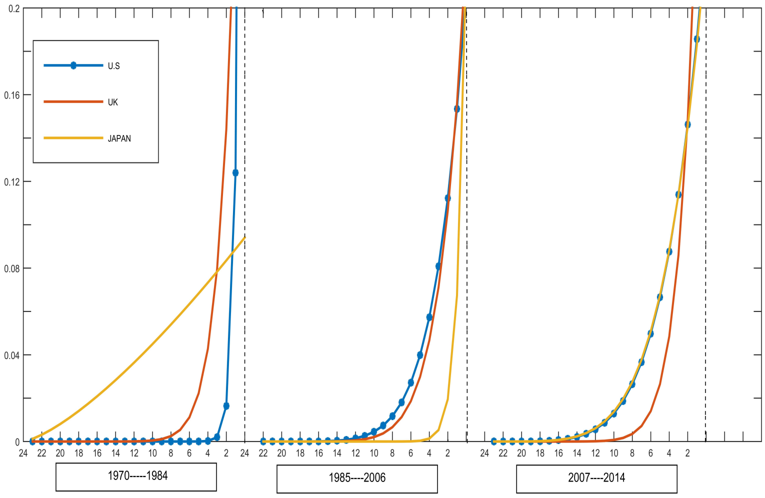
The Figure 2.3 plots Realized Volatility of FTSE All Index and annually growth rates of macroeconomic variables available in a monthly frequency, including Industrial Production (IP), Producer Price Index (PPI), Unemployment Rate (UEM) and Term Spread (TS). Monthly Macro data ranges from Jan/1970 to Dec/2014. Shade areas represent OECD recessions for the UK business cycles from 1970–2014.

Figure 2.4: Nikkei 225 Index Realized Volatility and Macroeconomic Variables In The Japan Market



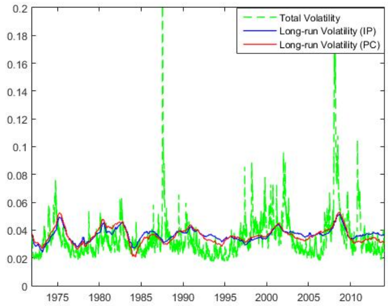
The Figure 2.4 plots Realized Volatility of Nikkei 225 Index and annually growth rates of macroeconomic variables available in a monthly frequency, including Industrial Production (IP), Producer Price Index (PPI), Unemployment Rate (UEM), Term Spread (TS), Housing Starts (HS) and Japanese Yen Index. Monthly Macro data ranges from Jan/1970 to Dec/2014. Shade areas represent Japanese business cycle recessions during 1970–2014.

Figure 2.5: Weights of the GARCH-MIDAS model with Realized Volatility

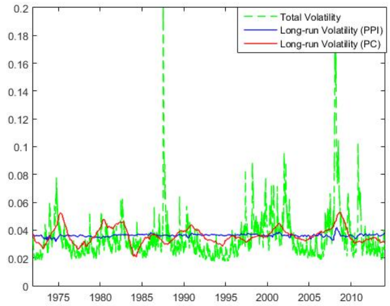


The Figure 2.5 plots beta weighting curves derived from the GARCH-MIDAS model specified with RV throughout different sub-samples of 1970-1984, 1985-2006 and 2007-2014. The blue line depicts the weighting curve of GARCH-MIDAS-RV in the U.S stock market. The red line represents the weighting curve of GARCH-MIDAS-RV in the UK stock market. The yellow line is the weighting curve of GARCH-MIDAS-RV in the Japanese stock market. Weights are assigned to 24 lags (K=24) of observations from monthly realized volatility in the GARCH-MIDAS frame.

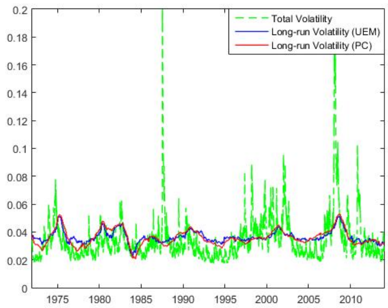
Figure 2.6 Long-run Volatility Component for the GARCH-MIDAS model in U.S



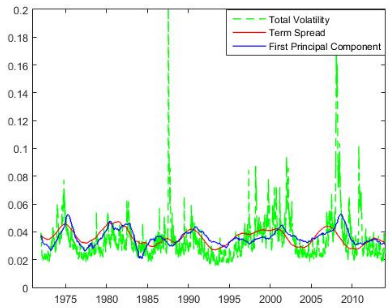
(a) Industrial Production



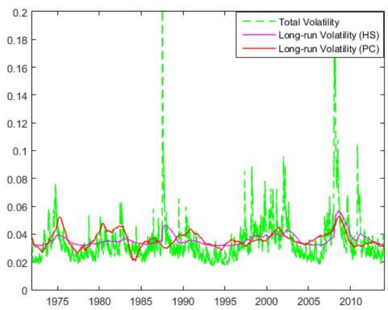
(b) Producer Price Index



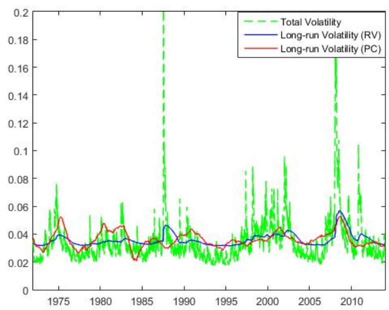
(c) Unemployment Rate



(d) Term Spread



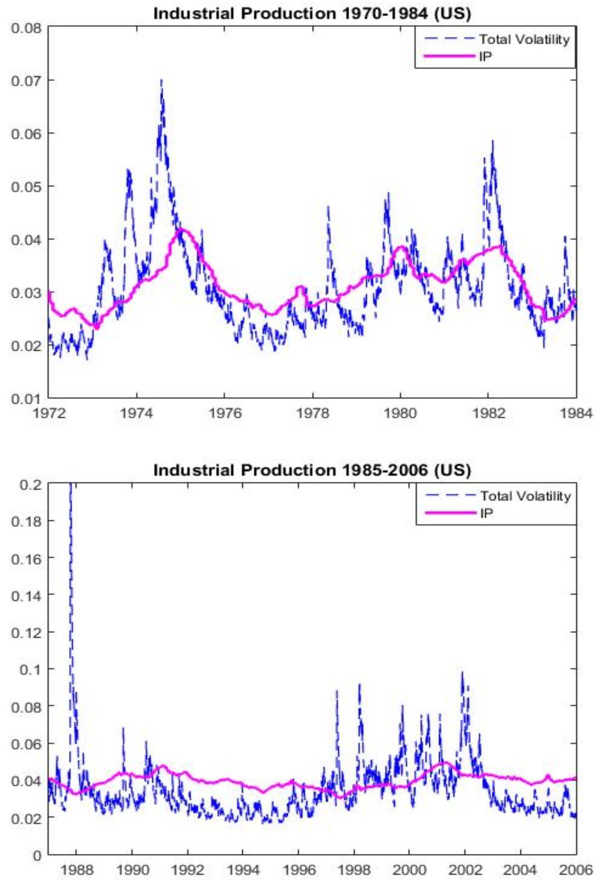
(e) Housing Starts



(f) Realized Volatility

Figure 2.6 depicts the whole conditional volatility $\tau_t g_t$(green dash line) and long-run volatility components τ_t driven by alternative macroeconomic variables (solid magenta line) and principal component (solid blue line) from GARCH-MIDAS model.

Figure 2.7: Sub-sample comparisons of long-term volatility component driven by ΔIP



The Figure 2.7 depicts the whole conditional volatility $\tau_t g_t$ (blue dash line) and long-run volatility component τ_t driven by growth of industrial production ΔIP (solid magenta line) for two sub-periods. Top panel refers to the sub-period 1970-1984 in the U.S market. Bottom panel refers to the sub-period 1985-2006 in the U.S market.

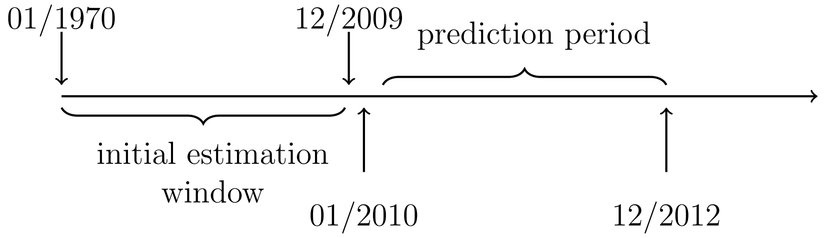


Figure 2.8: Pseudo out-of-sample forecasting for 2010-2012 dataset

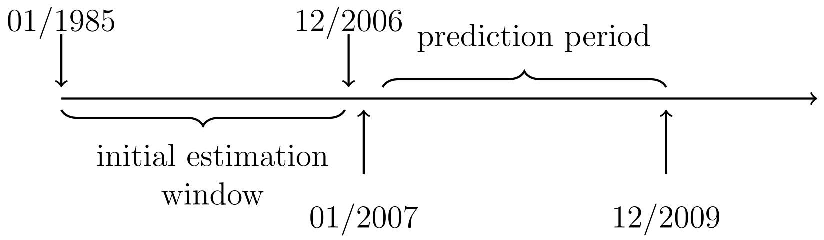


Figure 2.9: Pseudo out-of-sample forecasting for 2007-2009 dataset

References

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

Asgharian, H., A. J. Hou, and F. Javed (2013). The importance of the macroeconomic variables in forecasting stock return variance: A garch-midas approach. *Journal of Forecasting 32*(7), 600–612.

Baele, L. (2005). Volatility spillover effects in european equity markets. *Journal of Financial and Quantitative Analysis 40*(2), 373–401.

Bellégo, C. and L. Ferrara (2012). Macro-financial linkages and business cycles: A factor-augmented probit approach. *Economic Modelling 29*(5), 1793–1797.

Blanchard, O. and J. Simon (2001). The long and large decline in us output volatility. *Brookings papers on economic activity 2001*(1), 135–174.

Borio, C. E. and P. W. Lowe (2002). Asset prices, financial and monetary stability: exploring the nexus.

Boyd, J. H., J. Hu, and R. Jagannathan (2005). The stock market's reaction to unemployment news: Why bad news is usually good for stocks. *The Journal of Finance 60*(2), 649–672.

Boyd, J. H., R. Levine, and B. D. Smith (2001). The impact of inflation on financial sector performance. *Journal of monetary Economics 47*(2), 221–248.

Calvet, L. E., A. J. Fisher, and S. B. Thompson (2006). Volatility comovement: a multifrequency approach. *Journal of econometrics 131*(1-2), 179–215.

Campbell, J. Y. (1999). Asset prices, consumption, and the business cycle. *Handbook of macroeconomics 1*, 1231–1303.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies 1*(3), 195–228.

Chen, N.-F., R. Roll, and S. A. Ross (1986). Economic forces and the stock market. *Journal of business*, 383–403.

Conrad, C. and K. Loch (2015). Anticipating long-term stock market volatility. *Journal of Applied Econometrics 30*(7), 1090–1114.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics 95*(3), 776–797.

Engle, R. F. and G. Lee (1999). A long-run and short-run component model of stock return volatility. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, 475–497.

Engle, R. F. and J. G. Rangel (2008). The spline-garch model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies 21*(3), 1187–1222.

Estrella, A. and F. S. Mishkin (1998). Predicting us recessions: Financial variables as leading indicators. *Review of Economics and Statistics 80*(1), 45–61.

Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *The American economic review 71*(4), 545–565.

Fama, E. F. (1990). Stock returns, expected returns, and real activity. *The journal of finance 45*(4), 1089–1108.

Fama, E. F. and K. R. French (1989). Business conditions and expected returns on stocks and bonds. *Journal of financial economics 25*(1), 23–49.

Ferrara, L., C. Marsilli, and J.-P. Ortega (2014). Forecasting growth during the great recession: is financial volatility the missing ingredient? *Economic Modelling 36*, 44–50.

Ferson, W. E. and C. R. Harvey (1993). The risk and predictability of international equity returns. *Review of financial Studies 6*(3), 527–566.

Ferson, W. E., S. Sarkissian, and T. T. Simin (2003). Spurious regressions in financial economics? *The Journal of Finance 58*(4), 1393–1413.

Ghysels, E., A. Sinko, and R. Valkanov (2007). Midas regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Hamao, Y., R. W. Masulis, and V. Ng (1990). Correlations in price changes and volatility across international stock markets. *The review of financial studies 3*(2), 281–307.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of Financial Economics 22*(2), 305–333.

Harvey, C. R. (1989). Forecasts of economic growth from the bond and stock markets. *Financial Analysts Journal 45*(5), 38–45.

Harvey, C. R. (1991). The term structure and world economic growth. *The Journal of Fixed Income 1*(1), 7–19.

Kontonikas, A. and C. Ioannidis (2005). Should monetary policy respond to asset price misalignments? *Economic Modelling 22*(6), 1105–1121.

Lin, W.-L., R. F. Engle, and T. Ito (1994). Do bulls and bears move across borders? international transmission of stock returns and volatility. *Review of financial studies 7*(3), 507–538.

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

McQueen, G. and V. V. Roley (1993). Stock prices, news, and business conditions. *The Review of Financial Studies 6*(3), 683–707.

Mele, A. (2007). Asymmetric stock market volatility and the cyclical behavior of expected returns. *Journal of financial economics 86*(2), 446–478.

Modigliani, F. and R. A. Cohn (1979). Inflation, rational valuation and the market. *Financial Analysts Journal 35*(2), 24–44.

Ng, A. (2000). Volatility spillover effects from japan and the us to the pacific–basin. *Journal of international money and finance 19*(2), 207–233.

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

Rebonato, R. and T. Hatano (2018). The economic origin of treasury excess returns: A cycles and trend explanation. *Available at SSRN 3183653*.

Schnabl, G. (2015). Monetary policy and structural decline: Lessons from japan for the european crisis. *Asian Economic Papers 14*(1), 124–150.

Schnabl, G. and A. Hoffmann (2008). Monetary policy, vagabonding liquidity and bursting bubbles in new and emerging markets: An overinvestment view. *World Economy 31*(9), 1226–1252.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics 6*(2), 461–464.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The journal of finance 44*(5), 1115–1153.

Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association 97*(460), 1167–1179.

Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature 41*(3), 788–829.

Susmel, R. and R. F. Engle (1994). Hourly volatility spillovers between international equity markets. *Journal of international Money and Finance 13*(1), 3–25.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

Chapter 3

The Role of Macroeconomic Information in High-frequency Realized Volatility: Evaluating Persistence and Structural Change in U.S Stock Market Volatility

This paper investigates the relationship between macroeconomic information and volatility in the US stock market. We examine whether the long-term persistence and structural changes in stock volatility can be explained by relevant macroeconomic information, especially fundamental variables and macroeconomic uncertainty. To examine the effects of macroeconomic information on long-term persistence in stock volatility, we extend the Heterogeneous Autoregressive Realized Volatility (HAR-RV) model of Corsi (2009) to include macroeconomic information by using the Mixed Data Sampling (MIDAS) approach of Ghysels et al. (2007). To investigate the relationship between structural changes in volatility and macroeconomic information, we use and extend the Tree-HAR model of Audrino and Bühlmann (2001). The Tree-HAR model describes volatility as a regime-switching process, where daily realized volatility locally follows a HAR process in each regime, and shifts among regimes are governed by variables that trigger changes in regime after crossing particular threshold values. We allow macroeconomic variables to trigger the changes in regime. Empirical results show that the macroeconomic variables and uncertainty measures we use not only have significant impacts on stock volatility, but also consistently deliver a more elaborate regime structure (typically 3 or 4 regimes) for US stock volatility, relative to the traditional Markov Regime-Switching GARCH (MRS-GARCH) model. In terms of prediction ability, the comparison results generated from the Model Confidence Set (MCS) show that, as the horizon extends, the Tree-HAR models with macro variables outperform the original Tree-HAR model that exclusively depends on volatility, especially during the crisis periods of 2000–2001 and 2007–2008.

Keywords: Macroeconomic Uncertainty, realized volatility, MIDAS, regime switching, long-term volatility, HAR-RV, Tree-HAR, Model Confidence Set.

3.1 Introduction

Volatility is a fundamental measure of risk in financial markets and a large literature has been developed to model and explain volatility and its features. Within this literature, one area of interest is the link between macroeconomic information and volatility, particularly whether macroeconomic information contributes to explaining the long-term component of volatility, persistence in volatility and whether macroeconomic information contributes to explaining structural breaks in volatility. Motivated by Lee and Engle (1993) decomposition of volatility into a short-run (transitory) component and a long-run (permanent) component, several studies have found that macroeconomic information is significant in explaining long-run volatility and its persistence. Examples include the growth rate of industrial production and inflation (Engle et al. (2013)); the term spread and housing starts (Conrad and Loch (2015)); and the short-term interest rate and the first principal component derived from factor analysis of several macroeconomic variables (Asgharian et al. (2013)). With regard to macroeconomic information determining and explaining structural changes in volatility, using Markov-switching GARCH (MS-GARCH) models Hamilton and Lin (1996), Dueker (1997) and So et al. (1998) observe an association between the real economy and structural changes in volatility. In particular, high volatility regimes appear to be triggered by downturns in the real economy. Campbell (1999) also notes that volatility tends to be higher during recessions than expansions.

Though significant progress has been made in examining the relationship between macroeconomic information and financial market volatility, much of the extant literature examines the link between either the macroeconomy and persistence in volatility or the macroeconomy and structural changes in volatility but not both, despite evidence that both are present for at least some markets (see Beine et al. (2001) and Morana and Beltratti (2004) for evidence of both persistence and structural changes in volatility for the foreign exchange market, for example.) The concern here is that failure to consider structural breaks in volatility, for example, might lead to a misinterpretation of the relationship between macroeconomic information and long-term volatility persistence. Andersen and Bollerslev (1997) note that ignoring structural breaks might artificially lead to strong persistence in volatility while Engle et al. (2013) argue that the GARCH-MIDAS volatility model with fixed parameters cannot capture fundamental changes in the real economy over time.

The contribution of this paper is to investigate the relationship between macroeconomic information and US stock market volatility by taking both long-term volatility persistence and structural changes in volatility into consideration, and to examine whether this improves out-of-sample volatility forecasting. We use the Heterogeneous Autoregressive Realized Volatility (HAR-RV) model of Corsi (2009) as the basis of our analysis. Comparing to traditional GARCH-type models, HAR-type models possess potential advantages when dealing with the relationship between macroeconomic information and stock market volatility. Unlike traditional GARCH models, HAR-type models typically have a multi-component volatility structure which means they are suitable to deal with the mismatch in frequencies between macroeconomic and financial data (macroeconomic data is usually observed in a monthly or quarterly frequency while stock returns are typically observed in a daily or intra-daily frequency.) In the HAR-RV model Corsi (2009), for example, daily realized volatility is explained as a function of autoregressive volatility components realized over daily, weekly and monthly horizons. We extend the HAR-RV model to allow for the inclusion of macroeconomic variables as explanatory variables using the Mixed Data Sampling (MIDAS) approach of Ghysels et al. (2007). Using the MIDAS approach allows us to directly observe how macroeconomic information affects volatility movements within the HAR-RV framework. The HAR-RV model can be further extended to incorporate changes in regime by imposing a tree structure on the model. Such a model is known as the Tree-HAR model (Audrino and Bühlmann (2001)). In the Tree-HAR model, volatility is described in a regime-switching model, where volatility locally follows a HAR process in one regime, and shifts across regimes are governed by variables that trigger a change in regime when they cross particular threshold values. We extend the Tree-HAR model to allow macroeconomic information to not only explain long-term volatility movements within each regime, but also to determine the regime structure of volatility.

We utilize several macroeconomic variables in our empirical analysis, including the growth in industrial production growth, inflation, housing starts, unemployment rate and money supply. With the exception of the unemployment rate, all of the macroeconomic variables individually explain not only long-term volatility movements but also different regimes for stock market volatility. For instance, industrial production growth, inflation and housing starts are commonly selected as thresholds that distinguish between three regimes for stock market volatility: low volatility, medium volatility and high volatility while money supply growth is important as a threshold for high volatility regimes. As alternative macroeconomic variables perform equally well in terms of regime identification, we examine whether "broader" measures of macroeconomic information. We use principal components analysis (PCA) and the Aruoba-Diebold-Scotti (ADS) Index as more broad measures of "the macroeconomy" in our HAR-type volatility models to examine the impact of combined macroeconomic information on stock volatility. We also use a measure of macroeconomic uncertainty in our HAR models. Since uncertainty in the macroeconomy is a major concern for investors, and consequently affects risk premia, we investigate the relationship between macroeconomic uncertainty and stock volatility. Following Jurado et al. (2015), we construct a measure of macroeconomic uncertainty by aggregating unexpected information from a broad range of macro variables into an uncertainty index and we use this in our HAR-type models. Our results show that stock market volatility prior to the financial crisis can largely be attributed to macroeconomic uncertainty rather than macroeconomic fundamentals themselves. This suggests that what really matters for market participants, at least outside of times of crisis, is uncertainty about the economy as a whole. Results from our out-of-sample forecasting exercises show that the inclusion of macroeconomic information in the models leads to an improvement in volatility forecasting, particularly over longer-term forecasting horizons.

The rest of the paper is organised as follows. Section 3.2 provides literature reviews. Section 3.3 discusses the models we use in our empirical analysis. Section 3.4 describes the data while sections 3.5 and 3.6 discuss our empirical results by using macroeconomic variables. Section 3.7 discusses empirical results by using macroeconomic uncertainty. Section 3.8 examines forecasting performance while Section 3.9 concludes.

3.2 Literature Review

One typical feature of time-varying volatility in financial markets is volatility clustering and persistence: large changes in the price of an asset are usually followed by other large changes (see Fama (1965), Chou (1988) and Engle and Patton (2007) among many others). Bollerslev (1986) provides an empirical interpretation of volatility persistence as reflecting that news might not be absorbed into prices immediately. Rather, the impact of the arrival of news might last for a while, showing as a hyperbolic decaying pattern in the GARCH process. Ding et al. (1993) argue that long-range persistence might arise from the cross-sectional aggregation of news arrival being observed with different degrees of persistence. Müller et al. (1997) provide a similar theoretical interpretation of volatility persistence from the perspective of heterogeneous trading activities in financial markets. According to the Heterogeneous Market Hypothesis (Müller et al. (1997)), different types of participant have different investment time horizons and consequently different trading activities. Hence these different participants react to the same news differently, which in turn contributes to different frequency volatility components. The interactions among those volatility components can lead to volatility persistence.

Macroeconomic information occupies a large proportion of news. Consequently, there have been many studies that relate volatility persistence to macroeconomic information. Inspired by Lee and Engle (1993)'s decomposition of volatility into short-term and long-term components, Engle et al. (2013) introduce the GARCH-MIDAS model that employs macroeconomic variables to explain long-term volatility movements and its persistence. In the two-component GARCH-MIDAS framework, the long-term component of volatility is interpreted as the trend in volatility around which the short-term component fluctuates. Engle et al. (2013) allow the long-term component to be driven by macroeconomic variables through a MIDAS filter. They find that the macroeconomic variables are significant in explaining volatility and lead to an improvement in predictions of volatility. Following on from Engle et al. (2013) several studies adopt the GARCH-MIDAS framework to examine a the role of a large set of macroeconomic variables in explaining and forecasting stock market volatility. For instance, Conrad and Loch (2015) observe that the term spread and housing starts are able to explain volatility prior to the financial crisis while Asgharian et al. (2013) observe that combining the macroeconomic variables into factors, these factors are able to explain long-term volatility movements. This body of evidence suggests that macroeconomic information is important in explaining stock market volatility, particularly the long-term component.

Another strand of the literature examines long-term persistence in volatility taking structural breaks into consideration. Inclan and Tiao (1994) study structural changes in volatility of asset returns in emerging markets and find those changes match up with local economic crises, political events and the global financial crisis. Hamilton and Lin (1996), Dueker (1997) and So et al. (1998) observe a significant relationship between stock market volatility and macroeconomic conditions under a regime switching model. Their findings highlight that the high volatility regime appears to be triggered by economic downturn. Beltratti and Morana (2006) evaluates the linkage between financial market volatility and macroeconomic volatility taking both persistence and structural breaks into consideration. The empirical evidence in Beltratti and Morana (2006) indicates that the volatility break process for the S&P500 coincides with volatility breaks in the Federal Funds Rate and M1 growth. More importantly, Beltratti and Morana (2006) points out that omitting volatility breaks will lead to misleading results with regard to volatility persistence. Ignoring potential breaks in S&P 500 volatility leads to the artificial emergence of cointegration relationships between output growth, the Federal Funds rate and stock volatility. Granger and Hyung (2004) also highlight the importance of structural breaks when evaluating the long-memory feature of volatility. Both Beltratti and Morana (2006) and Granger and Hyung (2004), therefore, emphasise the importance of taking breaks into consideration when evaluating long-term persistence (or long memory) in volatility: both long-memory and structural breaks characterize time-varying volatility in financial markets (Morana and Beltratti (2004)).

3.3 HAR-type Realized Volatility models

3.3.1 The HAR-RV model

The intuition behind the HAR-RV model comes from the Heterogeneous Market Hypothesis (see, for example, Müller et al. (1997)). The hypothesis suggests that different types of market participants have different time horizons and different trading activities and as a consequence different types of participants react to the same news differently. Short-term traders usually make transactions at a daily or intra-daily frequency. Consequently they are interested in short-term price fluctuations and their trading activities contribute to short-term variations. Long-term traders prefer to hold assets for relatively longer periods of time and consequently are not particularly interested in short-term price movements, focusing more on large price movements. Accordingly, their trading activities contribute to long-term variations in financial market. For the case of foreign exchange market, Müller et al. (1997) and Dacorogna et al. (1997) observe that the long-term variation predicts the short-term variations better than the opposite way around, which infers that return volatilities in financial market actually arise from the interactions of different volatility components realized over different time scales. Following this line, recent empirical works suggest a additive cascade relationship between long-term and short-term asset variations whereby short-term volatility can be influenced by long-term volatility, but not vice versa. (see Arneodo et al. (1998), Muzy et al. (2000), Breymann et al. (2000) and Muzy et al. (2001) as examples.)

Inspired by the Heterogeneous Market Hypothesis of Müller et al. (1997), Corsi (2009) introduces the Heterogeneous Autoregressive Realized Volatility (HAR-RV) model that describes daily volatility as a sequence of autoregressive volatility components realized over daily, weekly and monthly time horizons:

$$\sigma_{t+1}^{d} = \beta_0 + \beta_1 \sqrt{RV_t^d} + \beta_2 \sqrt{RV_t^w} + \beta_3 \sqrt{RV_t^m} + \varepsilon_{t+1} \tag{3.1}$$

where σ_{t+1}^{d} is daily volatility and $\sqrt{RV_t^*}$ is Realized Volatility which is used as a proxy for unobserved volatility component realized over daily (d), weekly (w) and monthly (m) horizons ($^* = d, w, m$). The daily realized variance RV_t^d is calculated as a sum of squared intra-daily returns sampled with tick-by-tick frequency. RV suffers from an estimation bias that will become even worse as sampling frequency increases, and this bias usually comes from microstructure noise in finely intervals. In order to balance the trade-off between microstructure noise and sampling frequency, we employ the two-scale realized volatility method of Zhang et al. (2005) to estimate RV, which is robust to endogeneous microstructure noise under a finely interval. RV_t^w and RV_t^m refer to weekly and monthly realized variances, respectively. These are both calculated using a recursive rolling window with fixed length (one week or one month respectively). RV_t^w and RV_t^m are calculated from non-overlapping daily observations as,

$$RV_t^w = \frac{1}{4} \sum_{i=1}^{4} RV_{t-i}^d \qquad RV_t^m = \frac{1}{17} \sum_{i=5}^{21} RV_{t-i}^d \tag{3.2}$$

3.3.2 The HAR-RV Model Incorporating Macroeconomic Variables (the HAR-MIDAS Model)

Our first contribution in this paper is incorporating macroeconomic variables into the HAR-RV model to create a link between macroeconomic information and stock volatility. Recall from the discussion in Subsection 3.3.1 that a cascade relationship exists among different volatility components such that the short-term (daily) volatility component is influenced by the long-term (weekly and/or monthly) volatility component. We begin by specifying monthly volatility (the long-term component) as a function of lagged realized volatility and macroeconomic variables:

$$\sigma_{t+1}^m = \beta_0^m + \beta_1^m \sqrt{RV_t^m} + \beta_2^m X_t^m + \epsilon_{t+1}^m \tag{3.3}$$

where σ_{t+1}^m is the unobserved monthly volatility component at day $t+1$, which in turn is driven by two streams of information, financial market information represented by monthly realized volatility $\sqrt{RV_t^m}$ and macroeconomic information represented by X_t^m. Given the cascade relationship discussed earlier, we can write weekly volatility as:

$$\sigma_{t+1}^w = \beta_0^w + \beta_1^w \sqrt{RV_t^w} + \beta_2^w E_t[\boldsymbol{\sigma}_{t+1}^m] + \epsilon_{t+1}^w \tag{3.4}$$

where the weekly volatility component σ_{t+1}^w is a function of current realized weekly volatility $\sqrt{RV_t^w}$ and the expectation at time t of the monthly volatility component on day $t+1$, σ_{t+1}^m. In a similar fashion, we can write the following expression for daily volatility:

$$\sigma_{t+1}^d = \beta_0^d + \beta_1^d \sqrt{RV_t^d} + \beta_2^d E_t[\boldsymbol{\sigma}_{t+1}^w] + \epsilon_{t+1}^d \tag{3.5}$$

where the daily volatility component σ_{t+1}^d is a function of current realized daily volatility $\sqrt{RV_t^d}$ and the expectation at time t of the weekly volatility component on day $t+1$.

Substituting 3.3 into 3.4 for σ_{t+1}^m, then substituting 3.4 into 3.5 for σ_{t+1}^w respectively gives,

$$\sigma_{t+1}^d = \beta_0 + \beta_1 \sqrt{RV_t^d} + \beta_2 \sqrt{RV_t^w} + \beta_3 \sqrt{RV_t^m} + \beta_4 X_t^m + \upsilon_{t+1} \tag{3.6}$$

In Equation 3.6 we employ the Mixed Data Sampling (MIDAS) filter of Ghysels et al. (2007) into the HAR-RV model. The MIDAS filter specifies X_t^m as a smoothed weighted function of lagged observations of the macroeconomic variable:

$$X_t^m = \sum_k^K \varphi(\omega_1, \omega_2) x_{t-k}^m \tag{3.7}$$

In the MIDAS filter given by 3.7, X_t^m can be thought of as an weighted average of the available information on the variable x from time $t-k$ to t. The weights allocated to lags of x^m are governed by a beta weighting scheme $\varphi(\omega_1, \omega_2)$ [1]with two parameters, ω_1 and ω_2, which satisfy the conditions $\varphi_k > 0$ and $\sum_{k=1}^K \varphi_k = 1$. The weighting scheme is specified as

$$\varphi(\omega_1, \omega_2) = \frac{(k/K)^{\omega_1 - 1}(1 - k/K)^{\omega_2 - 1}}{\sum_{j=1}^K (j/K)^{\omega - 1}(1 - j/K)^{\omega_2 - 1}} \tag{3.8}$$

This beta weighting scheme is able to deliver flexible weight patterns with only two parameters, ω_1 and ω_2. In particular, if ω_1 is restricted to equal one such that there is only one flexible parameter, ω_2, the weighting scheme can deliver a decaying pattern where higher weights are allocated to more recent lags of the variable.

To complete the model we need to specify an equation for returns. We specify the return-generating equation as

$$r_{t+1} = \mu + \sigma_{t+1}^d \varepsilon_{t+1} \tag{3.9}$$

Equations 3.9 and 3.6 constitute the HAR-MIDAS model that forms the basis

[1]According to Ghysels et al. (2007), the polynomial lag operator ($\varphi(\omega_1, \omega_2)$) in MIDAS filter can be parameterized through alternative ways, such as exponential Beta or Almon lag scheme, Autoregressive distributed lag (ADL) scheme, Distributed lag scheme. In this paper, we follow Engle et al. (2013) that utilize Beta weighting scheme in the MIDAS filter.

of our empirical analysis:

$$\begin{aligned} r_{t+1} &= \mu + \sigma^d_{t+1}\varepsilon_{t+1} \\ \sigma^d_{t+1} &= \beta_0 + \beta_1\sqrt{RV^d_t} + \beta_2\sqrt{RV^w_t} + \beta_3\sqrt{RV^m_t} + \beta_4 X^m_t + \upsilon_{t+1} \end{aligned} \tag{3.10}$$

3.3.3 The Tree-HAR model

Our second contribution is to incorporate macroeconomic information into the Tree-HAR model. The role of the macroeconomic variables in the Tree-HAR model is two-fold, explaining long-term volatility movements and being the threshold variables that identify the regime structure in volatility. The Tree-HAR model introduced by Audrino and Corsi (2010) divides the whole data set G into non-overlapping regimes:

$$G = \cup^K_{j=1} R_j \qquad R_j \cap R_i = \emptyset \tag{3.11}$$

where G should comprises all historical information available at time t and K denotes number of regimes finally being identified in the Tree-HAR model. Reviewing the Tree-HAR model application in Audrino and Corsi (2010), G consists of historical returns and realized volatility for the US market. For each regime R_j, daily volatility σ^d_{t+1} locally follows a HAR-RV process with fixed parameters. Therefore, the Tree-HAR model is a regime switching HAR-RV model, defined as:

$$\begin{aligned} r_{t+1} &= \mu + \sigma^d_{t+1}\varepsilon_{t+1} \\ \sigma^d_{t+1} &= \sum_{j=1}^{K}[\beta_{0,j} + \beta_{1,j}\sqrt{RV^d_t} + \beta_{2,j}\sqrt{RV^w_t} + \beta_{3,j}\sqrt{RV^m_t}] I_{\Phi_t \in R_j} + \upsilon_{t+1} \end{aligned} \tag{3.12}$$

where the dummy variable I_{Φ_t} indicates whether daily realized variance RV_t^d belongs to a particular regime R_j, or not. The shifts among regimes are governed by a set of threshold variables Φ_t that trigger changes in regime once a particular threshold value is crossed. In the case of Audrino and Corsi (2010), stock returns and daily realized volatility are the threshold variables.

3.3.4 The Tree-HAR Model Incorporating Macroeconomic Variables (the Tree-HAR-MIDAS Model)

A multiple-regime structure in the Tree-HAR model can be constructed using the binary tree algorithm of Breiman et al. (1984). The whole data set G can be decomposed into several non-overlapping regimes through a sequence of binary partition steps. In each partition step, the optimal value selected from one threshold variable in set Φ_t serves as an indicator, determining a binary partition in one of the existing regimes. Repeating the binary partition procedure several times, the set G develops into a tree-structure where terminal nodes are perceived as regimes.

Our first extension of the Tree-HAR model is to let macroeconomic information determine daily volatility within each regime:

$$r_{t+1} = \mu + \sigma_{t+1}^d \varepsilon_{t+1}$$
$$\sigma_{t+1}^d = \sum_{j=1}^{K} [\beta_{0,j} + \beta_{1,j}\sqrt{RV_t^d} + \beta_{2,j}\sqrt{RV_t^w} + \beta_{3,j}\sqrt{RV_t^m} + \beta_{4,j}X_t^m] I_{\Phi_t \in R_j} + \upsilon_{t+1} \tag{3.13}$$

where the variables are as defined earlier. In each regime daily volatility locally follows a HAR-MIDAS process. The regime switching is determined by the threshold variables in Φ_t. Our second extension of the Tree-HAR model is to add macroeconomic variables into Φ_t in addition to stock returns and monthly realized volatility so that in our case, $\Phi_t = \{\tilde{r}_t, \quad \widetilde{\sqrt{RV_t^m}}, \quad \widetilde{X}_t\}$. $\tilde{r}_t$ denotes a sequence of stock returns available at time t. Similarly, $\widetilde{\sqrt{RV_t^m}}$ denotes a sequence of monthly realized volatilities and $\widetilde{X}_t$ refers to a series of macroeconomic observations for one macro variable that is available at time t.

The threshold set Φ_t contains α-quantiles of values for all threshold variables where α denotes as i/mesh and i=1,....mesh-1. In terms of mesh points, we simply follow Audrino and Corsi (2010) to choose mesh=8. If we assume the whole data set G is kept in a p-dimensional space (p refers to the number of threshold variables) edged by p coordinates. Therefore, quantiles of threshold values for each threshold variable can be interpreted as mesh points located on each coordinate. In this context, the binary tree algorithm is arranged through a sequence of space-partitions. For each partition step, we search for an optimal binary partition that is triggered by one particular value of one threshold variable by crossing over all possible mesh points of different coordinate. After a sequence of partition steps, the p-dimensional space can be transformed into a multiple-regime strucure. We describe the binary tree algorithm and how we proceed to estimate the Tree-HAR-MIDAS model as below.

A. The Binary Tree Algorithm: Initial Setting

- The data set G in our study consists of historical macroeconomic information, returns and the daily, weekly and monthly realized variances of stock returns: $G = \{\widetilde{X}_t, \quad \tilde{r}_t, \quad \widetilde{RV}_t^d, \quad \widetilde{RV}_t^w, \quad \widetilde{RV}_t^m\}$

- The threshold set contains quantiles of values for returns, monthly realized volatilities and macroeconomic variables observed until time t: $\Phi = \{\tilde{r}_t, \quad \widetilde{\sqrt{RV_t^m}}, \quad \widetilde{X}_t\}$

- The binary partition procedure has m steps. In every step j (j =

$1, 2, \ldots m$), an optimal value d_j and its associated threshold variable ϕ_j are both selected from the threshold set Φ.

- In step j, using the optimal value d_j and its threshold variable ϕ_j, the whole data set G has an optimal regime structure P^j_{opt}, which satisfies $P^j_{opt} = \{R^j_1 \cup R^j_2 ... \cup R^j_k\}$.

- Subsequently in step $j+1$, one regime $R^j_{i^*}$ that comes from the existing structure of P^j_{opt} is selected and split further into two, namely $R_{i^*,left}$ and $R_{i^*,right}$, according to whether $\phi_{j+1} > d_{j+1}$ or $\phi_{j+1} \leq d_{j+1}$, respectively. Thus the whole data set G is re-constructed with an optimal regime structure P^{j+1}_{opt} that satisfies $P^{j+1}_{opt} = \{R^{j+1}_1 \cup ... R^{j+1}_{i-1} \cup R^{j+1}_{i^*,left} \cup R^{j+1}_{i^*,right} \cup R^{j+1}_{i+1} ... \cup R^{j+1}_k\}$.

B. The Binary Tree Algorithm: Partition Steps

The starting point is the Tree-HAR-MIDAS model with no partition of G. This can be nested in the HAR-MIDAS model:

$$\sigma^d_{t+1} = \beta_0 + \beta_1 \sqrt{RV^d_t} + \beta_2 \sqrt{RV^w_t} + \beta_3 \sqrt{RV^m_t} + \beta_4 X^m_t + v_{t+1} \tag{3.14}$$

where the parameter set $\theta = \{\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \omega_1, \omega_2\}$ is estimated via maximum likelihood with the log-likelihood function shown as below:

$$-l(\theta, \Phi) = -\sum_{t=1}^{n} \tfrac{1}{2} \log(2\pi) - \tfrac{1}{2} \sum_{t=1}^{n} \log(\sigma^2_t(\theta)) - \tfrac{1}{2} \sum_{t=1}^{n} \left[\tfrac{r_{t+1} - \mu(\theta)}{\sigma^2_t(\theta)} \right] \tag{3.15}$$

First step: we construct a multi-dimensional grid search that spreads out over the mesh-points of the different threshold variables to find a combination of threshold variable and its corresponding threshold value which delivers a partition structure that minimizes the log-likelihood function in Equation

3.15. With an optimal combination of threshold variable and value $(\phi_1,\ d_1)$, observations of set G available at time t can be allocated into two regimes according to whether $\phi_{1,t} > d_1$ or $\phi_{1,t} \leq d_1$, respectively. Diagrammatically,

$$
\begin{array}{cc}
\multicolumn{2}{c}{R_0} \\
\swarrow & \searrow \\
R_1 = R_{0,left}\{I_\Phi : \phi_1 > d_1\} & R_2 = R_{0,right}\{I_\Phi : \phi_1 \leq d_1\} \\
G = R_1 \cup R_2 & P_{opt}^{(1)} = \{R_1,\ R_2\}
\end{array}
\tag{3.16}
$$

Starting from $j = 0$, G is given by R_0 (see Equation 3.11) and is split into two regimes, $R_{0,left}$ and $R_{0,right}$; This gives the optimal partition structure $P_{opt}^{(1)}$. We rename $R_{0,left}$ and $R_{0,right}$ as R_1 and R_2, respectively. If we stop the binary partition here, we have a Tree-HAR model with two regimes:

$$
\sigma_{t+1}^d = \sum_{R_j \in \{R_1, R_2\}} \left[\beta_{0,j} + \beta_{1,j}\sqrt{RV_t^d} + \beta_{2,j}\sqrt{RV_t^w} + \beta_{3,j}\sqrt{RV_t^m} + \beta_{4,j} X_t^m \right] I_{\Phi_t \in R_j} + \upsilon_{t+1}
\tag{3.17}
$$

Second step: we continue the binary partition into the second step ($j = 2$). As G already has a two-regime structure ($P_{opt}^{(1)}$) from the first step ($j = 1$), we therefore conduct two independent grid searches that go over all available threshold values of the different threshold variables within the two existing regimes, R_1 and R_2, respectively. We then find the new optimal combination of threshold variable (ϕ_2) and its value (d_2) that comes from one of the two existing regimes, either of which could deliver an optimal partition structure. For instance, if the optimal combination of ϕ_2 and d_2 can be found in, say, R_1, then R_1 is split into $R_{1,left}$ and $R_{1,right}$. Diagrammatically,

$$
\begin{array}{ccc}
 & R_0 & \\
 & \swarrow \quad \searrow & \\
R_1 = \{I_\Phi : \phi_1 > d_1\} & & R_2 = \{I_\Phi : \phi_1 \leq d_1\} \\
\swarrow \quad \searrow & & \\
\end{array}
$$

$$
R_3 = R_{1,left}\{I_\Phi : \phi_1 > d_1 \& \phi_2 > d_2\} \quad R_4 = R_{1,right}\{I_\Phi : \phi_1 > d_1 \& \phi_2 \leq d_2\} \tag{3.18}
$$

The two descendants (we rename $R_{1,left}$ and $R_{1,right}$ as R_3 and R_4) that stem from R_1 together with the other regime R_2 constitute a new optimal partition structure $P_{opt}^{(2)}$. If we terminate the binary partition at the second step, we have a Tree-HAR model with three regimes:

$$
\sigma_{t+1}^d = \sum_{R_j \in \{R_2, R_3, R_4\}} \left[\beta_{0,j} + \beta_{1,j}\sqrt{RV_t^d} + \beta_{2,j}\sqrt{RV_t^w} + \beta_{3,j}\sqrt{RV_t^m} + \beta_{4,j}X_t^m \right] I_{\Phi_t \in R_j} + \upsilon_{t+1} \tag{3.19}
$$

***m* steps**: we continue the binary partition until it reaches the mth step. With an optimal partition structure $P_{opt}^{(M-1)}$ obtained in the $(m-1)$th step, we implement independent grid-searches across all existing regimes to find an optimal combination of threshold variable ϕ_m and its threshold value d_m. Assuming this optimal combination can be found in one of the existing regimes R_{j^*}, $j^* \in \{1, 2, 3, ...m\}$, observations in regime R_{j^*} are again sub-divided into two according to whether $\phi_{m,t} > d_1$ or $\phi_{m,t} \leq d_1$, respectively.

The number of partition steps in the Tree-HAR-MIDAS model is important. Too large or too small a tree structure will lead to misspecification. When the tree structure is too fine, the Tree-HAR-MIDAS model is likely to be over-parameterized. When the tree structure is not fine enough, the Tree-

HAR model might fail to capture some regimes. We follow Audrino and Corsi (2010) and Audrino and Bühlmann (2001) and let the sequence of partition steps terminate at 5 ($m = 5$), which seems sufficient empirically to develop a large enough tree structure for financial time series.

C. The Binary Tree Algorithm: Optimal sub-tree structure

Having developed a sufficiently large tree structure it is necessary to prune the tree upward to search for the best subtree structure. This can be done on the basis of Akaike's Information Criterion (AIC). We save all tree nodes into a set Ψ, which includes intermediate partition nodes as well as terminal tree nodes. The intermediate partition nodes are all derived from previous partition steps. We then implement a bottom-up procedure to find all valid combinations from those tree nodes in Ψ that are able to cover entire observations without overlapping. Every valid combination is then defined as a valid sub-tree structure. Among these sub-tree structures, the best sub-tree structure is that which minimizes the AIC.

3.4 Data description

We use daily returns on the S&P500 Index and the daily realized volatility of returns on the S&P500 index over the period January 2nd 1996 to December 31st 2014 as the basis of our analysis. The realized volatility data comes from the Oxford-Man Institute Realized Library, where daily realized volatility is calculated from intra-day returns sampled at the 5-minute frequency. Descriptive statistics for the realized volatilities can be found in Table 3.1, where it can be seen that the realized volatilities for the daily, weekly and monthly horizons are right-skewed. The realized volatilities are all persistent, as shown by the first order autocorrelation coefficient, with the monthly realized volatility being the most persistent. Daily returns and daily realized volatility are plotted in Figure 3.1. The areas shaded grey in Figure 3.1 denote recessions as defined by the NBER. Unsurprisingly, the most volatile period in our sample is the most recent financial crisis which coincides with an economic recession.

[Insert Figure 3.1 and Table 3.1 here]

We are interested in evaluating how different types of macroeconomic information contribute to time-varying volatility. The macroeconomic variables we use are Industrial Production (IP), the Producer Price Index (PPI), the Unemployment Rate (UEM), Housing Starts (HS) and the Term Spread (TS). We use monthly seasonally-adjusted data covering the period January 1996 to December 2014; the data is obtained from the Federal Reserve Economic Database (FRED). The term spread is constructed as the difference between the 10-year treasury bond yield and the 3-month T-bill rate. Bar the Term Spread, we use the annualized growth rate of the macroeconomic variables in our empirical analysis:

$$\Delta X_t = \left[\frac{X_t}{X_{t-1}}\right]^{12} - 1 \tag{3.20}$$

In terms of the variance of the macroeconomic variables, we estimate it through a simplified mean equation:

$$\Delta X_t = \mu + \sigma_t \varepsilon_t \tag{3.21}$$

where ε_t is the innovation that is assumed to follow a standard normal distribution, $\varepsilon \sim N(0,1)$. $\hat{\sigma_t}$ is our estimate of macroeconomic volatility.

In addition to the macroeconomic variables above, we also employ the Aruoba-Diebold-Scotti (ADS) Business Conditions Index and principal components to compress multi-dimensional macroeconomic information into one variable. The ADS Business Conditions Index was introduced by Aruoba et al.

(2009) and combines macroeconomic information, suitably weighted, into an index that proxies underlying economic conditions. The ADS Index is constructed from several macroeconomic series: jobless claims , growth of payroll employment, industrial production growth, real manufacturing and trade sales, real personal income and real GDP. Figure 3.2 plots the annualized growth rates for all the macroeconomic variables. The two shaded areas represent the Dotcom Bubble and the Subprime Crisis. It can be seen from the graphs that there are periods where volatility in the macroeconomic variables quite pronounced, coinciding with one or more of the crises for several of the variables but not restricted to these.

[Insert Figure 3.2 here]

3.5 Empirical Results for HAR-MIDAS model

In this section, we use the HAR-MIDAS model (Equation 3.10) to examine the relationship, if any, between daily realized volatility for returns on the S&P500 Index and macroeconomic information. We consider two sets of macroeconomic information in our empirical analysis: growth rates (bar the term spread) in the macroeconomic variables discussed in the previous section, and the volatility of these macroeconomic variables. We also need to specify any restrictions on parameters in the beta weighting scheme (Equation 3.8). We follow Engle et al. (2013) and restrict $\omega_1 = 1$. The intuition for this restriction is that greater weight is then given to more recent observations on the relevant variable x. We also set the lag length K equal to 6 in Equation 3.8.

[Insert Table 3.2 here]

Table 3.2 reports the estimation results for the HAR-MIDAS model with each of the macroeconomic variables and each of the macroeconomic volatilities. β_1, β_2 and β_3 are the coefficients on the daily, weekly and monthly realized volatility components respectively while β_4 measures the effect of the particular macroeconomic variable on stock volatility. The first point to note is that in all of the models the sum of β_1 through β_3 is close to unity, which implies a high degree of persistence in volatility. We can also see from the results in Panel A of Table 3.2 that the coefficients on the macroeconomic variables are, with the exception of the term spread, statistically significant. This suggests that macroeconomic information is important in explaining realized volatility. Consistent with Engle et al. (2013) and Stock and Watson (2002), there is evidence that stock market volatility is counter-cyclical to the business cycle. The parameters on industrial production growth (ΔIP) and growth in housing starts (ΔHS) negatively affect stock market volatility. In contrast, an increase in the unemployment rate (ΔUEM) and inflation (ΔPPI) increase volatility. With the exception of the term spread, which is statistically insignificant anyway, the estimates of ω_2 in Panel A of Table 3.2 are slightly above, but close to, one. This implies that rather than declining quickly for more distant lags, the weights are quite equally distributed over the six lags in the weighting scheme. The results in Panel B of Table 3.2 indicate that with the exception of volatility in industrial production growth, which is statistically significant at the 1% level, and to a lesser extent the volatility of inflation, volatility in the macroeconomy is not a particularly strong determinant of stock market volatility.

The results in Panel A of Table 3.2 show that several of the macroeconomic variables appear to perform equally well in explaining daily volatility movements under the HAR-MIDAS framework. The individual significance of four out of the five macroeconomic variables (except Term spread) we consider suggests that there is more than one source of macroeconomic information that is important in explaining stock market volatility. However, including all of the macroeconomic variables in the HAR-MIDAS model at the same time leads to overparameterization of the model with the consequence that the parameters are poorly estimated or convergence is not achieved. As an alternative, we estimate the first principal component of the macro variables, and also use the Aruoba-Diebold-Scotti (ADS) Business Conditions Index, to examine the influence of, for want of a better expression, combined macroeconomic information on stock market volatility. The results in Panel C show that the first principal component and the ADS Business Conditions Index are statistically significant at the 1% level, showing that macroeconomic information more broadly significantly contributes to explaining stock market volatility. The results in this section show that macroeconomic information is a significant determinant of stock market volatility. In the next section, using the Tree-HAR-MIDAS model, we turn our attention to an analysis of whether and how the relationship between stock market volatility and macroeconomic information varies as underlying economic conditions vary.

3.6 Regime Structure and Macroeconomic Information

In this section, we employ the Tree-HAR-MIDAS model to investigate to what extent volatility persistence and structural changes in volatility can be attributed to macroeconomic information. As in the previous section 3.5, we restrict ω_1 to equal one in the beta weighting scheme and we set the number of lags K to six. The parameter estimates for all the Tree-HAR-MIDAS models are summarized in Tables 3.3 through 3.7. As a general comment, the results in

Tables 3.3 and 3.4 reveal that the regime structure is more elaborate than that often suggested by Markov Switching volatility models, there being three or four regimes with the Tree-HAR-MIDAS model often identifying low, moderate and high volatility regimes with the macroeconomic variables being important in determining some, but not all, of the regimes. We explore this in more detail in the next subsection.

[Insert Table 3.3 and 3.4 here]

Figures 3.3 through 3.6 show the possible regime structures for daily stock market volatility that the results in Tables 3.3 and 3.4 suggest. When macroeconomic information is not selected as a threshold variable that determines regime structure, which is the case for the growth rate in unemployment (UEM), the term spread (TS) and the ADS Business Conditions Index, daily volatility can be described by a three-regime structure in which the threshold variable is monthly realized volatility RV_m. This situation is shown in Figure 3.3. This three-regime structure is characterized by a high volatility regime (when $RV_m > 0.0101$), a moderate or medium-volatility regime (when $0.0058 < RV_m < 0.0101$) and a low volatility regime (when $RV_m \leq 0.0058$).

[Insert Figures 3.3 through 3.6 here]

When macroeconomic variables are selected as threshold variables, the three-regime structure that we observe when the threshold variable is realized volatility only, now evolves into a more elaborate four-regime structure, where the macroeconomic variables are typically employed to further partition the medium or high volatility regime into two further sub-regimes. Industrial production growth (Figure 3.4) splits the medium volatility regime (the regime where $0.0058 < RV_m < 0.0101$) into two parts, which are characterized by high and low production growth regimes, respectively. Both housing starts (Figure 3.5) and inflation (Figure 3.6) is able to refine the high volatility regime further into two regimes.

The sample period in our study covers several significant financial events including (but not limited to) the Dot-com bubble (2000–2002), the Global Financial crisis (2007–2009) and the subsequent European Sovereign Debt crisis after 2009. These events are known to cause volatility fluctuations in stock markets so it is interesting to interpret these events in the light of our Tree-HAR-MIDAS models. Figures 3.8 through 3.12 plot daily realized volatility with the regime that the particular Tree-HAR-MIDAS model classifies a particular observation into superimposed on realized volatility. With regard to the the Dot-com bubble, for the model with a three regime structure (Figure 3.11), most daily volatility observations tend to stay within two regimes: the high volatility regime ($RV_m > 0.0101$) and the medium volatility regime ($0.0058 < RV_m < 0.0101$). Prior to the Dot-com crash, volatility mostly stays in the medium volatility regime ($0.0058 < RV_m < 0.0101$). Where a macroeconomic variable does play a role in determining regimes, such as growth in inflation (Figure 3.9) for example, it can be seen that at the beginning of the Dot-com bubble, increasing of inflation growth triggers a regime shift from the high volatility regime with low inflation growth (the cluster of volatilities that are edged by $0.0058 < RV_m < 0.0101$ and $\Delta PPI < 0.0095$) to the high volatility regime with high inflation growth (the cluster of volatilities that are edged by $RV_m > 0.0101$ and $\Delta PPI > 0.0095$).

[Insert Figures 3.8 through 3.12 here]

When we look at the global financial crisis period (2007-2008), volatility prior to that period in the three-regime model (Figure 3.11) is mostly in the low volatility regime (shown in yellow $RV_m < 0.0058$). In the run up to the financial crisis the model with the three-regime structure classifies realized volatility as being in the moderate volatility regime (shown in blue $0.0058 < RV_m < 0.0101$) and into the high volatility regime (shown in purple $RV_m > 0.0101$) during the crisis. Where macroeconomic variables do play a role as threshold variables, high volatility during the financial crisis is characterized by higher growth in, for example, high inflation growth with $\Delta PPI > 0.0095$ (Figure 3.9) and high housing starts growth with $\Delta HS > 0.0015$ (Figure 3.10).

In terms of the role macroeconomic information plays in explaining volatility in the Tree-HAR-MIDAS model, the results in tables 3.3 and 3.4 show that the picture is somewhat different for the Tree-HAR-MIDAS models compared to the HAR-MIDAS models. Of particular interest here is not just the significance or otherwise of β_4 but the magnitude of ω_2. Recall from the previous section that in the HAR-MIDAS models (Table 3.2) ω_2 is close to one, suggesting that rather than declining quickly for more distant lags, the weights are quite equally distributed over the lags in the weighting scheme. The results in Tables 3.3 and 3.4 suggest that in the "extreme" regimes (particularly the high volatility regime and to a lesser extent the low volatility regime) ω_2 is quite often much greater than one in at least one of the "extreme" regimes. Figure 3.7 plots the weights attached to the lags (which we term beta lags) observed in each regime, as generated by the Tree-HAR-MIDAS models in Tables 3.3 and 3.4. The speed of decay is determined by ω_2. The larger the value of ω_2, the faster the decaying speed that the lags have. Generally, beta lags generated by the different macro variables in the Tree-HAR-MIDAS models decay faster when stock volatility is located in the "extreme" regimes.[2] This suggests that more distant macroeconomic information has much less impact on volatility when volatility is high and/or low.[3] In contrast, macroeconomic

[2]With housing starts growth, for example, ω_2 is 4.16 for the high volatility regime and 2.5 for the low volatility regime while with respect to medium volatility regime its value turns into 1.45.

[3]In this context, high volatility refers to the regime being classified by threshold variable RV_m that exceeds value of 0.0101 and low volatility refers to the regime being classified by

information becomes prominent and has a longer lasting impact on volatility when volatility is in the medium regime.[4]

[Insert Figures 3.7 here]

In terms of the significance of the macroeconomic variables in explaining daily realized volatility, their role is much less clear cut in the Tree-HAR-MIDAS model than in the HAR-MIDAS models of the previous section. Recall from Table 3.2 that β_4 was often statistically significant, suggesting macroeconomic information has a role to play in explaining daily volatility. Compare those findings to the results in Tables 3.3 and 3.4 where it can be seen that macroeconomic information does not consistently explain daily volatility across all regimes. Growth in industrial production is significant only in the medium volatility regime, for example, while unemployment rate is significant in both medium and low volatility regimes. The results in tables 3.3 and 3.4 suggest that the importance of macroeconomic information in explaining volatility is very much dependent on which variable is being used and which regime volatility is in.

To examine the accuracy of the regime identification in the Tree-HAR-MIDAS model incorporated with macroeconomic variable, we also estimate the Markov Switching GARCH (MS-GARCH) model as the benchmark model for comparison. Following Gray (1996), we employ a two-regime Markov Switching GARCH (hereafter MS-GARCH) model[5]. The model is:

$$\begin{aligned} r_t &= \mu + \sqrt{h_{t,j}}\eta_t \quad \eta_t \sim IID(0,1) \\ h_{t,j} &= \alpha_{0,j} + \alpha_{1,j}\varepsilon_{t-1}^2 + \alpha_{2,j}h_{t-1} \end{aligned} \tag{3.22}$$

threshold variable RV_m that is less than value of 0.0058.

[4] Consistent with empirical results in Table 3.3 and Table 3.4, medium volatility is identified as $0.0058 < RV_m < 0.0101$.

[5] According to the information criteria (AIC, BIC) and log-likelihood ratio, MS-GARCH with two-regime structure seems to be more suitable for S&P 500 Realzied volatility during 1996-2014.

where the parameters $[\alpha_{0,j}, \alpha_{1,j}, \alpha_{2,j}]$ vary across two regimes j, $j = 1, 2$.[6] For the purposes of comparison, we use the first principal component as a combined macroeconomic factor to be incorporated into the Tree-HAR-MIDAS model. Figure 3.13 plots the regime specifications from the Tree-HAR-MIDAS model (top graph) and the MS-GARCH model (bottom graph.) The top graph plots realized volatility with regimes that the particular Tree-HAR-MIDAS model classifies a particular observation into superimposed on realized volatility. The bottom graph plots smoothed probabilities from the MS-GARCH model, where the red line shows the probability of staying in the low-volatility regime while the blue line shows the probability of staying in the high-volatility regime. the plots show that there are episodes of what is quite moderate volatility that the MS-GARCH model characterises as being either in a low or high volatility state while the Tree-HAR-MIDAS model with its more sophisticated regime structure correctly classifies as moderate volatility.

3.7 Macroeconomic Uncertainty

In this section we investigate how macroeconomic uncertainty more generally is transmitted into stock market volatility, if at all. Uncertainty matters because a high degree of uncertainty might depress productivity growth, reduce investment and increase risk premia (Bloom et al. (2018)). Since uncertainty is inherently unobservable, we adopt the measure from Jurado et al. (2015), and examine the response of stock market volatility to macroeconomic uncertainty under the Tree-HAR-MIDAS model. Based on the notion that what matters to economic decision-making is whether the economy has become more or less predictable, Jurado et al. (2015) introduce an aggregated proxy of macroeconomic uncertainty whereby they aggregate the conditional

[6] These regimes can be thought of as high and low volatility regimes.

volatility of pure unexpected components which are realized from h-step-ahead forecast errors from a wide range of macroeconomic variables. Specifically, Jurado et al. (2015) compress a large set of macroeconomic variables into latent common factors (y_j) via a diffusion index model. With this limited number of common factors, they generate h-step-ahead predictions and obtain their associated forecast errors:

$$V^y_{jt+h} = y_{jt+h} - E[y_{jt+h}|I_t] \tag{3.23}$$

The conditional volatility of the prediction errors $E[(V^y_{t+h})^2|I_t]$ then provide a proxy for individual uncertainty, denoted as $\mathcal{U}^y_{jt}(h)$:

$$\mathcal{U}_{jt}(h)^y = \sqrt{E[(V^y_{jt+h})^2|I_t]} \tag{3.24}$$

An economic-wide uncertainty index, which we use as our measure of macroeconomic uncertainty, is then calculated as a weighted average of individual uncertainty measures:

$$\mathcal{U}_{jt}(h) = \sum_{j=1}^{N_y} w_j \mathcal{U}^y_{jt}(h) \tag{3.25}$$

where w_j is the weight attached to each individual uncertainty. N_y indexes the number for each individual uncertainty.

Using Equations 3.24 through 3.25, we construct macroeconomic uncertainty from 132 monthly macroeconomic variables, which cover a broad category of economic activities in the U.S. A detailed summary of the variables can be found in the online appendix of Jurado et al. (2015)[7]. Figure 3.14 plots the estimated macroeconomic uncertainty for three forecasting hori-

[7]In this chapter, we follow Jurado et al. (2015) to construct macroeconomic uncertainty via 132 macro series. The list of 132 macro variables is available through the link: https://assets.aeaweb.org/asset-server/articles-attachments/aer/app/10503/20131193_app.pdf

zons: $h = 1, 3, 12$ months, along with NBER recessions (the shaded areas), and compares it to the corresponding realized monthly volatility (the dashed green line), the first principal component (the solid red line), as well as the Aruoba-Diebold-Scotti Business Conditions Index (the yellow line) for the period 1996-2004. As the forecasting horizon h increases, macroeconomic uncertainty tends to become less volatile. We also observe that uncertainty measures reached their peaks prior to the Dot-com bubble and the global financial crisis, which suggests a lead-lag relationship between macroeconomic uncertainty and volatility.

Table 3.6 reports estimation results for including macroeconomic uncertainty in HAR-MIDAS model. Our measure of macroeconomic uncertainty realized over the 3-month, 6-month and 12-month horizons all have a positive sign, suggesting that the higher is uncertainty the higher will be volatility, and are all significant at the 1% level. As the forecasting horizon used to calculate the measure of macroeconomic uncertainty increases, the magnitude of the impact from uncertainty on stock volatility increases.

In line with the regime identification results in Section 3.6, macroeconomic uncertainty is also selected as a threshold to account for further partition within the moderate volatility regime ($0.0058 < RV_m < 0.0101$), delivering a four-regime structure. This is perhaps not that surprising given we found similar results for some of the individual macroeconomic variables but what is of interest is that regardless of the forecasting horizon used in estimating uncertainty, a four regime structure is always chosen as the best structure: a high volatility regime edged by $RV_m > 0.0101$; a low volatility regime edged by $RV_m < 0.0058$; and a moderate volatility regime ($0.0058 < RV_m < 0.0101$) which is further subdivided into two parts characterized by high and low macroeconomic uncertainty, respectively (see Table 3.7). Macro uncertainty, then, matters most in the moderate volatility regime when uncertainty exceeds its threshold value. For instance, both $\overline{\mathcal{U}}(3)$ and $\overline{\mathcal{U}}(6)$ have positive and significant impacts with magnitude of more than 0.8 on stock volatility within the “moderate volatility and high uncertainty ” regime. More importantly, from Figure 3.12 we can see that this positive impact of uncertainty on volatility in the moderate volatility regime occurs prior to financial crises during which volatility moves into the high volatility regime. This suggests that when macro uncertainty exceeds a certain level, it appears to push up stock volatility, nudging the stock market into the high volatility regime. Once the switch from the moderate volatile regime into the high volatile regime has taken place, the significance of macro uncertainty disappears. In terms of goodness of fit, the Tree-HAR-MIDAS model specified with $\overline{\mathcal{U}}(3)$ yields a higher maximum likelihood ratio relative to the principal component and ADS Index models.

3.8 Out of Sample Forecasting for the Tree-HAR-MIDAS Model

The results in the previous sections have shown the importance of macroeconomic information in explaining stock market volatility and in determining volatility regimes. In this section we investigate whether the inclusion of macroeconomic information improves the accuracy of volatility predictions and whether the role of macroeconomic information in predicting volatility (if any) changes over time.

To evaluate the forecasting performance of the Tree-HAR-MIDAS model with different macroeconomic variables, we use a recursive prediction procedure associated with an increasing window (IW). The initial estimation window starts in January 1996 and ends in December 2006. Once we have estimated the model, we generate our initial one-day ahead ($h = 1$) volatility forecasts over the following month as follows:

$$\sigma_{t+1}^{d}=\sum_{j=1}^{K}[\beta_{0,j}+\beta_{1,j}\sqrt{RV_t^d}+\beta_{2,j}\sqrt{RV_t^w}+\beta_{3,j}\sqrt{RV_t^m}]I_{\Phi_t\in R_j}+\upsilon_{t+1} \quad (3.26)$$

We then extend the estimation window by including the month we have just forecast over. For example, the estimation window now becomes January 1996 through January 2007 rather than January 1996 through December 2006. We then re-estimate the model and use the new parameter estimates to make one-day predictions over the following month. Repeating this procedure, we obtain 96 months of prediction results covering the period from January 2007 until December 2014. For each month, we compute the mean squared error (MSE),

$$L_{i,t}=MSE(\sigma_{i,t}^2,\hat{\sigma}_{i,t}^2)=\frac{1}{n}\sum_{j=1}^{n}(\sigma_{i,j}^2-\hat{\sigma}_{i,j}^2)^2 \quad (3.27)$$

where $L_{i,t}$ denotes the loss function for model i for month t, $\hat{\sigma}_{i,j}^2$ is the one-step ahead forecast at date j in month i and $\sigma_{i,j}^2$ is the realized volatility observed at date j. As noted in Section 3.6, we observe that the impact of the macroeconomic variables on stock volatility vary with respect to which regime volatility is in. Therefore, it is difficult to come to a clear-cut conclusion as to which model dominates in terms of explaining volatility. Accordingly, we use the Model Confidence Set (MCS) analysis of Hansen et al. (2011) to compare relative forecasting performance among the various Tree-HAR-MIDAS models and to identify a set of superior models under a joint confidence interval, rather than picking out one single best model. The main mechanism of MCS is to construct an initial set M_0 containing all competing models. Inferior models are eliminated step-by-step through a sequence of equivalent tests. In the first step, an equivalent test δ_M is applied between any pair of competing models in set M_0. Those pairwise equivalent tests in set M_0 are constructed based on the null hypothesis $H_{0,M}$ that assumes all models perform equally well:

$$\delta_M = \begin{cases} H_{0,M}: & d_{i,j} = 0 \\ H_{A,M}: & d_{i,j} \neq 0 \end{cases} \tag{3.28}$$

$$\bar{d}_{i,j} = E(d_{ij,t}) = E\left(L_{i,t} - L_{j,t}\right) \qquad \forall i, j \in M$$

where $\bar{d}_{i,j}$ refers to the relative performance between any two models, i and j, which is calculated as the difference in the loss functions between i and j on average. If $H_{0,M}$ is accepted for a significance level α, all candidates perform equally well. Otherwise, the worst-performing model (the model which has the highest T_R statistic, defined below) is discarded from the initial set M_0. In this context, T_R measures the greatest loss between two models:

$$T_R = max_{i,j \in M} = \frac{\bar{d}_{i,j}}{\sqrt{var(\bar{d}_{i,j})}} \tag{3.29}$$

Note that the variance of $\bar{d}_{i,j}$ in Equation 3.28 is estimated through the bootstrap procedure. This process test is repeated until the null hypothesis is accepted. At this stage, the initial set M_0 has shrunk and reached the confidence set $\hat{M}^*_{1-\alpha}$ with a significance level of α. The set $\hat{M}^*_{1-\alpha}$ contains the surviving models which have equivalent forecasting performance.

Following Hansen et al. (2011), we use a significant level of α equal to 25%. Our one-day ahead ($h = 1$) prediction results are summarized in Table 3.8, where the standard Tree-HAR model incorporated with RV is included as a benchmark for the purpose of comparison. The standard Tree-HAR model performs particularly well when volatility is high, especially during the global financial crisis of 2007–2008 and the Sovereign Debt crisis at the start of 2010. When comparing the p-value of the MCS with 25% significance level, the standard Tree-HAR model obtains the highest probability of 1.0000 in 10 out of 16 cases[8], where a p-value of 1.0000 indicates that the Tree-HAR model is the last surviving model in set $\hat{M}^*_{75\%}$. By contrast, models that include the macro variables are less informative about future daily volatility during 2007-2008 and 2010. Results in Table 3.8 suggest relatively weak performance of models with macroeconomic variables with the exception of housing starts. This is perhaps because growth in housing starts have long, quite persistent effects in the high volatility regime and the moderate volatility regime (see Figure 3.10 and Table 3.3). This feature pays off in terms of predictive accuracy. During the financial crisis of 2007–2008, ΔHS outperforms alternative candidates when making short-term predictions.

[Insert Table 3.8 here]

We now turn our attention to investigating the predictive performance of the macroeconomic variables over a longer horizon, given that macroeconomic variables tend to be associated with the long-term component of volatility. Following Bollerslev et al. (2016), we extend the forecast horizon to one-week ahead (medium-term) prediction as well as one-month ahead (longer-term) prediction, respectively:

$$RV^w_{t+5/t+1} = \sum_{j=1}^{K}[\beta_{0,j} + \beta_{1,j}\sqrt{RV^d_t} + \beta_{2,j}\sqrt{RV^w_t} + \beta_{3,j}\sqrt{RV^m_t}]I_{\Phi_t \in R_j} + \upsilon_{t+1} \tag{3.30}$$

$$RV^m_{t+22/t+1} = \sum_{j=1}^{K}[\beta_{0,j} + \beta_{1,j}\sqrt{RV^d_t} + \beta_{2,j}\sqrt{RV^w_t} + \beta_{3,j}\sqrt{RV^m_t}]I_{\Phi_t \in R_j} + \upsilon_{t+1} \tag{3.31}$$

[8] We generate MCS comparisons at a biannual frequency. Therefore, MSEs and p-values are shown for 16 cases.

where $RV^w_{t+5/t+1}$ is the one-week-ahead forecast of volatility, which is equivalent to daily realized volatility rolling forward from time $t+1$ to $t+5$. Similarly, $RV^m_{t+22/t+1}$ refers to the one-month-ahead forecast, rolling forward from $t+1$ to $t+22$. Tables 3.9 to 3.10 provide one week ahead and one-month ahead prediction results respectively. Rather surprisingly, the term spread performs relatively well over the longer horizon when volatility is low.[9] Unsurprisingly, as the forecasting horizon increases macroeconomic information becomes much more valuable. Taking the global financial crisis period (2007–2008) as an example, relative to the rather weak performance of the individual macroeconomic variables in short-term prediction (see Table 3.8), we can see that more models that include the macro variables are present in the set $\hat{M}^*_{75\%}$ with higher p-values for long-term prediction. In particular, inflation provides superior long-term predictions in Table 3.10, obtaining p-values of 1.0000 in 5 out of 16 cases.

[Insert Table 3.9 and Table 3.10 here]

Since different macroeconomic variables offer the best forecasting results in different sub-periods, it is hard to identify one dominant variable that outperforms the others in terms of forecasting ability. Christiansen et al. (2012) identify this kind of difficulty as model uncertainty. They point out that it is quite difficult for an economic agent to select the right macro variables that are best suited for volatility prediction. This is where using a composite variable that captures multiple aspects of the macroeconomy, such as the first principal component or the ADS index or macroeconomic uncertainty should help. The results in Table 3.11 show that, especially when forecasting one month ahead, macroeconomic uncertainty outperforms alternative composite

[9] Stock and Watson (2003) and Wheelock et al. (2009) state that the term spread usually performs better for recession prediction rather than output growth.

variables in most of our forecasting periods. In a nutshell, our forecasting results in Section 3.8 reveal that adding macroeconomic information helps to improve volatility prediction for both the long-run and short-run horizons.

[**Insert Table 3.11 here**]

3.9 Conclusion

In this paper, we have proposed the HAR-MIDAS and Tree-HAR-MIDAS models to describe and forecast US stock market volatility. These models allow us to employ macroeconomic information as explanatory variables to both account for movements in stock market volatility (the HAR-MIDAS model) and to determine switches in volatility regimes (the Tree-HAR-MIDAS model.) In the Tree-HAR-MIDAS model, regimes are constructed by using the binary tree partition process, where macroeconomic information together with volatility determine the thresholds for each partition step. With no structural change, adding macroeconomic information into the HAR-MIDAS model provides a better description of stock market volatility relative to the traditional HAR-RV model. When considering potential structural changes in volatility, both expected and unexpected macroeconomic information helps to deliver a more elaborate regime structure, partitioning volatility into low, medium and high volatility regimes with macroeconomic information further partitioning the medium volatility regime regimes. The Tree-HAR-MIDAS model outperforms the Markov-Switching GARCH model and the traditional Tree-HAR model. Finally, we observe a significant improvement in long-term volatility forecasts when using macroeconomic variables in the Tree-HAR-MIDAS model.

Table 3.1: Summary Statistics

	$R_{s\&p500}$	Daily RV	Weekly RV	Monthly RV	Δ IP	Δ PPI	Δ HS	ΔUEM	TS	ADS
Mean	0.0340	0.8921	0.8925	0.8935	2.1342	3.3132	0.5926	0.0000	0.0177	−0.1891
Std	1.2489	0.5875	0.5333	0.4931	7.8749	13.3867	12.6987	16.3299	0.0117	0.6120
Skew	−0.0847	3.1970	3.0011	2.8066	−1.0900	−0.1576	0.2826	0.3082	−0.1749	−2.0429
Kurt	10.8754	21.6864	17.7025	15.2260	7.9621	5.0223	3.2975	4.1275	1.8927	9.6695
Autocorrelations										
1-lag	−0.0742	0.7898	0.9771	0.9978	0.1959	0.3321	−0.3533	0.1908	0.9370	0.9983
5-lag	−0.0503	0.6829	0.8304	0.9705	0.2203	−0.1308	0.0759	0.3470	0.8240	0.9907
10-lag	0.0270	0.6217	0.7611	0.9113	0.0551	0.0017	0.0444	0.1201	0.3424	0.9316

[a] Table 3.1 describes statistics for daily, weekly and monthly Realized Volatility series and annual growth rates for all cited macroeconomic variables. The daily realized volatility is calculated by averaging up 5-mins squared returns over on day horizon, t. The weekly (or monthly) realized volatility observed at date t is calculated by averaging up non-overlapping daily RVs over past one week (or one-month) horizon, which is shown as:

$$RV_t^w = \frac{1}{4}\sum_{i=1}^{4} RV_{t-i}^d \qquad RV_t^m = \frac{1}{17}\sum_{i=5}^{21} RV_{t-i}^d$$

[b] Term spread is the yield spread between 3-month treasury bill and ten-year government bond. ADS refers to the Aruoba-Diebold-Scotti business conditions index, which is designed to track real business condition at a daily frequency.

[c] Series autocorrelations are specified with 1-lag, 5-lag and 10-lag length, respectively. The sample perid is January 1996-December 2014, for a total of 228 macroeconomic observations and 4733 realized volatility observations.

Table 3.2: Parameter Estimates for the HAR-MIDAS model with Macroeconomic Variables and Macroeconomic Variances

Macro-variable specification	β_0	β_1	β_2	β_3	β_4	ω_1	ω_2	Loglikelihood	AIC
Panel A: HAR-MIDAS specified with macroeconomic variable									
Δ *IP*	0.0010***	0.3779***	0.3461***	0.1826***	−0.0397***	1	1.0101***	−19921.5607	−39731
	(4.3499)	(11.0650)	(9.0085)	(5.6694)	(-5.4725)		(2.9973)		
Δ *PPI*	0.0006***	0.3778***	0.3448***	0.1954***	0.0033***	1	1.1198***	−19921.7994	−39727
	(3.0290)	(11.1364)	(8.9920)	(6.1090)	(2.5607)		(4.6141)		
Δ *UEM*	0.0009**	0.3774***	0.3456***	0.1795***	0.0022***	1	1.0111	− 19923.5077	−39729
	(4.0878)	(11.0531)	(8.9862)	(5.4615)	(3.9111)		(1.6459)		
Δ *TS*	0.0007***	0.3794***	0.3477***	0.1945***	0.0022	1	3.9035	−19917.4755	−39717
	(3.2228)	(11.0438)	(8.7332)	(5.8741)	(0.3926)		(1.4027)		
Δ *HS*	0.0008***	0.3783***	0.3468***	0.1859***	−0.0049**	1	1.1005***	− 19920.4367	−39720
	(3.9642)	(11.0712)	(9.0271)	(5.6944)	(-2.0850)		(2.9207)		
Panel B: HAR-MIDAS specified with macroeconomic variance									
Δ *IP* Variance	0.0006***	0.3778***	0.3459***	0.1812***	0.0503***	1	1.0100***	−19921.5887	−39725
	(2.8847)	(11.0478)	(8.9951)	(5.6216)	(5.6308)		(3.6966)		
Δ *PPI* Variance	0.0005**	0.3775	0.3455***	0.1782***	0.0039**	1	1.0167*	− 19922.4781	−39725
	(2.2215)	(11.0939)	(9.0531)	(5.3013)	(1.9668)		(1.7521)		
Δ *UEM* Variance	0.0006***	0.3792***	0.3477***	0.1916***	0.0012	1	1.0361***	−19918.0280	−39718
	(3.0162)	(11.0913)	(9.0402)	(5.9019)	(1.2512)		(3.8114)		
TS Variance	0.0006***	0.3792***	0.3475***	0.1945***	0.5100	1	1.5539	−19917.9546	−39718
	(3.5893)	(11.0678)	(9.0164)	(6.0493)	(0.7914)		(1.7014)		
Δ *HS* Variance	0.0008***	0.3793***	0.3476***	0.1953***	-0.0014	1	2.1164***	-19917.9063	-39717
	(3.2602)	(11.0765)	(9.0302)	(6.0710)	(-1.1451)		(7.0820)		
Panel C: HAR-MIDAS specified with combined economic indicator and its variance									
*PC*1	0.0008***	0.3785***	0.3471***	0.1845***	0.0173***	1	1.0110***	−19919.7800	−39723
	(3.7946)	(11.0774)	(9.0249)	(5.6764)	(3.4836)		(4.1411)		
ADS	0.0009***	0.3772***	0.3452***	0.1752***	-0.0310***	1	1.0100***	−19923.2947	−39728
	(4.0733)	(11.0522)	(8.9716)	(5.3336)	(-4.7282)		(3.5515)		
*PC*1 Variance	0.0006***	0.3790***	0.3477***	0.1879***	0.0134**	1	1.0116*	−19918.3962	−39718
	(3.3780)	(11.0794)	(9.0416)	(5.8191)	(2.8663)		(1.9229)		
ADS Variance	0.0008***	0.3787***	0.3471***	0.1845***	0.0080***	1	1.0108***	− 19919.3889	−37920
	(3.6834)	(11.0719)	(9.0294)	(5.7266)	(3.7519)		(3.8178)		
Benchmark: HAR-RV model									
HAR-RV	0.0005***	0.4302***	0.3643***	0.1498***					
	(4.9583)	(30.3922)	(18.9163)	(8.5977)					

[a] Table 3.2 reports estimation results for the HAR-MIDAS model specified with alternative macroeconomic variables and macroeconomic variances. The HAR-MIDAS model is shown as below:

$$E_t[\sigma_{t+1}^{(d)}|F_t] = \beta_{0j} + \beta_{1j}RV_t^d + \beta_{2j}RV_t^w + \beta_{3j}RV_t^m + \beta_{4j}X_t \qquad X_t = \sum_{k=1}^{K}\phi_k(\omega)x_{i,t-k} \quad X_t = \sum_{k=1}^{K}\phi_k(\omega)\hat{\sigma}_{i,t-k}$$

X_t has two alternative specifications: one is a weighted average value of lagged observations from one macroeconomic variable,x_i. Another one is a weighted average value of lagged volatilities from one macroeconomic variance,$\hat{\sigma}_i$. Panel A reports estimation results for all alternative macroeconomic variables. Panel B reports estimation results for all alternative macroeconomic variances. Panel C provides estimation results for combined macroeconomic indexes, including the first principal component (PC1) and the ADS index. The last row displays estimations of the HAR-RV model, which is defined as:

$$E_t[\sigma_{t+1}^{(d)}|F_t] = \beta_{0j} + \beta_{1j}RV_t^d + \beta_{2j}RV_t^w + \beta_{3j}RV_t^m$$

The numbers in parentheses are robust t-statistics. Statistic significance at the 1%, 5% and 10% level is indicated by ***,**,*, respectively.

Table 3.3: Parameter Estimates for Daily Volatility in the Tree-HAR-MIDAS model with Macroeconomic Variables

Macro-variable	Regime structure	β_0	β_1	β_2	β_3	β_4	ω_1	ω_2	Loglikelihood	AIC
Δ IP	$\boldsymbol{R_1}$			$\boldsymbol{RV_m > 0.0101}$						
	(1178 obs)	0.0013*	0.3804***	0.3224***	0.1922***	0.0019	1	7.1338***	−20583.2584	−40744
		(1.8125)	(7.4337)	(5.8477)	(3.7572)	(0.9371)		(45.7257)		
	$\boldsymbol{R_2}$			$\boldsymbol{0.0058 < RV_m < 0.0101 \;\; \Delta IP > 0.0578}$						
	(636 obs)	0.0094	0.4476***	0.1269	0.0070	−0.0731*	1	3.2156***		
		(2.4834)	(4.1253)	(0.9899)	(0.0524)	(−1.8242)		(3.6062)		
	$\boldsymbol{R_3}$			$\boldsymbol{0.0058 < RV_m < 0.0101 \;\; \Delta IP < 0.0578}$						
	(1719 obs)	0.0013***	0.2832***	0.4142***	0.1461**	−0.0115***	1	1.0100***		
		(2.7927)	(8.9944)	(9.3353)	(2.3665)	(−3.9907)		(4.8284)		
	$\boldsymbol{R_4}$			$\boldsymbol{RV_m \leq 0.0058}$						
	(1178 obs)	0.0021***	0.3191***	0.2335***	0.0690	0.0024	1	4.5709***		
		(4.1582)	(6.6419)	(4.2183)	(0.6423)	(1.1225)		(11.0280)		
Δ PPI	$\boldsymbol{R_1}$			$\boldsymbol{RV_m > 0.0101 \;\; \Delta PPI > 0.0095}$						
	(1178 obs)	0.0051**	0.2186***	0.3670***	0.1257	−0.0315	1	1.0144	−20574.9902	−40727
		(2.5755)	(4.4887)	(4.8391)	(1.4929)	(−1.1547)		(1.3819)		
	$\boldsymbol{R_2}$			$\boldsymbol{RV_m > 0.0101 \;\; \Delta PPI < 0.0095}$						
	(1080 obs)	0.0017**	0.4195***	0.2870***	0.1929***	0.0096	1	2.8384**		
		(1.9648)	(6.4463)	(4.2825)	(3.1960)	(1.3730)		(2.2791)		
	$\boldsymbol{R_3}$			$\boldsymbol{0.0058 < RV \leq 0.0101}$						
	(1275 obs)	0.0015***	0.3494***	0.3230***	0.1299**	0.0100*	1	1.0291*		
		(3.1760)	(7.1919)	(5.7210)	(2.0267)	(1.9240)		(1.7254)		
	$\boldsymbol{R_4}$			$\boldsymbol{RV_m < 0.0058}$						
	(1178 obs)	0.0025***	0.3154***	0.2270***	0.0564	−0.0095**	1	2.9782***		
		(4.8827)	(6.5608)	(4.1034)	(0.5356)	(−2.4925)		(3.0444)		
Δ HS	$\boldsymbol{R_1}$			$\boldsymbol{RV^m > 0.0101 \;\; \Delta HS < -0.0015}$						
	(2051 obs)	0.0049**	0.4058***	0.2969***	0.1460**	0.0502*	1	4.1652**	−20611.2111	−40800
		(2.5342)	(6.0530)	(4.1255)	(2.2517)	(1.8380)		(2.2514)		
	$\boldsymbol{R_1}$			$\boldsymbol{RV_m > 0.0101 \;\; \Delta HS > -0.0015}$						
	(1178 obs)	0.0044**	0.2998***	0.3063***	0.2161***	−0.0644***	1	1.0168***		
		(2.4361)	(5.4676)	(4.1483)	(2.7237)	(−7.1416)		(4.1801)		
	$\boldsymbol{R_3}$			$\boldsymbol{0.0058 < RV_m < 0.0101}$						
	(304 obs)	0.0021***	0.3448***	0.3134***	0.0705	−0.0177***	1	1.4564**		
		(4.2329)	(7.0408)	(5.5647)	(1.0737)	(−4.0360)		(2.2519)		
	$\boldsymbol{R_4}$			$\boldsymbol{RV_m < 0.0058}$						
	(1178 obs)	0.0024***	0.3124***	0.2211***	0.0602	−0.0130***	1	2.5087***		
		(4.7424)	(6.6501)	(3.9587)	(0.5614)	(−2.8195)		(4.2246)		

[a] Table 3.3 reports estimated parameters, regimes and related statistics for the Tree-HAR-MIDAS model specified with alternative macroeconomic variables, including Δ IP, Δ PPI and Δ HS. The estimation period is January 2006-December 2014, for a total of 228 monthly macroeconomic observations and 4733 daily returns and realized volatilities. In Tree-HAR-MIDAS model:

$$\sigma_{t+1}^{(d)} = \sum_{j=1}^{N}[\beta_{0j} + \beta_{1j}RV_t^d + \beta_{2j}RV_t^w + \beta_{3j}RV_t^m + \beta_{4j}X_t]I_{X_t^{pred}\in R_j} + \upsilon_{t+1} \qquad X_t^m = \sum_{k=1}^{K}\phi_k(\omega)x_{i,t-k}$$

X_t is a weighted average value of lagged observations from one typical macroeconomic variable x_t. This macroeconomic variable has been adjusted as annual growth rates. In a Tree-HAR-MIDAS model, the regime structure is constructed by selecting a sequence of optimal threshold values from threshold set $X^{pred} = \{RV_m, r, X_t\}$, where X_T represents alternative Δ IP, Δ PPI and Δ HS. The numbers in parentheses are robust t-statistics. Statistic significance at the 1%, 5% and 10% level is indicated by ***, **, *, respectively.

Table 3.4: Parameter Estimates for Daily Volatility in the Tree-HAR-MIDAS model with Macroeconomic Variables

Macro-variable	Regime structure	β_0	β_1	β_2	β_3	β_4	ω_1	ω_2	Loglikelihood	AIC
Δ UEM	R_1			$RV_m > 0.0101$						
	(1178 obs)	0.0014**	0.3797***	0.3219***	0.1926***	−0.0015	1	9.5014***	−20525.1538	−40627
		(1.9939)	(7.4310)	(5.8452)	(3.7622)	(−1.3852)		(55.4589)		
	R_2			$0.0058 < RV_m < 0.0101$						
	(2355 obs)	0.0023***	0.3467***	0.3155***	0.0498	0.0052***	1	1.0100***		
		(4.5060)	(7.1829)	(5.6428)	(0.6755)	(4.3229)		(2.6686)		
	R_3			$RV_m < 0.0058$						
	(1178 obs)	0.0027***	0.3176***	0.2318***	0.0126	0.0035*	1	1.1258**		
		(4.6841)	(6.7018)	(4.1837)	(0.1132)	(1.7899)		(2.3547)		
TS	R_1			$RV_m > 0.0101$						
	(1178 obs)	0.0014**	0.3808***	0.3222***	0.1928***	−0.0102	1	2.2922***	−20522.7416	−40623
		(1.8652)	(7.4274)	(5.8359)	(3.7571)	(−0.7518)		(7.8215)		
	R_2			$0.0058 < RV_m < 0.0101$						
	(2355 obs)	0.0025***	0.3474***	0.3157***	0.0709	−0.0234***	1	3.2197***		
		(3.7008)	(6.9803)	(5.5450)	(1.1093)	(−2.8338)		(4.2236)		
	R_3			$RV_m < 0.0058$						
	(1178 obs)	0.0024***	0.3181***	0.2326***	0.0458	−0.0084	1	18.4154***		
		(4.5809)	(6.7037)	(4.1786)	(0.4270)	(−1.4280)		(79.2483)		

[a] Table 3.4 reports estimated parameters, regimes and related statistics for the Tree-HAR-MIDAS model specified with alternative macroeconomic variables, including Δ UEM and Δ TS. The estimation period is January 2006-December 2014, for a total of 228 monthly macroeconomic observations and 4733 daily returns and realized volatilities. In Tree-HAR-MIDAS model:

$$\sigma_{t+1}^{(d)} = \sum_{j=1}^{N}[\beta_{0j} + \beta_{1j}RV_t^d + \beta_{2j}RV_t^w + \beta_{3j}RV_t^m + \beta_{4j}X_t]I_{X_t^{pred}\in R_j} + \upsilon_{t+1} \qquad X_t^m = \sum_{k=1}^{K}\phi_k(\omega)x_{i,t-k}$$

X_t is a weighted average value of lagged observations from one typical macroeconomic variable x_t. This macroeconomic variable has been adjusted as annual growth rates. In a Tree-HAR-MIDAS model, the regime structure is constructed by selecting a sequence of optimal threshold values from threshold set $X^{pred} = \{RV_m, r, X_t\}$, where X_T represents alternative Δ UEM, Δ and TS. The numbers in parentheses are robust t-statistics. Statistic significance at the 1%, 5% and 10% level is indicated by ***, **, *, respectively.

Table 3.5: Parameter Estimates for Daily Volatility in the Tree-HAR-MIDAS model with Macroeconomic Variables

Macro-variable	**Regime structure**	β_0	β_1	β_2	β_3	β_4	ω_1	ω_2	**Loglikelihood**	**AIC**
ADS	R_1			$RV_m > 0.0101$						
	(1178 obs)	0.0013*	0.3807***	0.3223***	0.1921***	0.0096	1	19.8077***	−20520.2655	−40617
		(1.8543)	(7.4335)	(5.8405)	(3.7510)	(0.8130)		(186.9733)		
	R_2			$0.0058 < RV_m < 0.0101$						
	(2355 obs)	0.0019***	0.3496***	0.3205***	0.0881	−0.0382***	1	1.0100***		
		(3.9813)	(7.2010)	(5.6982)	(1.3030)	(−3.5726)		(3.6741)		
	R_3			$RV_m < 0.0058$						
	(1178 obs)	0.0022***	0.3179***	0.2316***	0.0593	0.0316*	1	17.5808***		
		(4.3806)	(6.7396)	(4.1493)	(0.5461)	(1.7455)		(115.8082)		
PC_1	R_1			$RV_m > 0.0101$						
	(1178 obs)	0.0012*	0.3809***	0.3218***	0.1965***	0.0092	1	19.9921***	−20619.4034	−40816
		(1.6949)	(7.4350)	(5.8186)	(3.8173)	(0.8124)		(56.1587)		
	R_2			$0.0058 < RV_m < 0.0101$ $PC1 > 0.0070$						
	(570 obs)	0.0012	0.4867***	0.0538	−0.1434	0.3268***	1	19.9945***		
		(1.0344)	(5.8416)	(0.5442)	(−0.9292)	(3.8364)		(8.4553)		
	R_3			$0.0058 < RV_m < 0.0101$ $PC1 < 0.0070$						
	(1785 obs)	0.0014***	0.2588***	0.4346***	0.1377**	0.0262**	1	4.2746***		
		(2.9096)	(6.8632)	(8.6132)	(2.2940)	(2.1569)		(6.3737)		
	R_4			$RV_m < 0.0058$						
	(1178 obs)	0.0024***	0.3148***	0.2241***	0.0326	0.0380**	1	2.0913**		
		(4.7562)	(6.6052)	(4.0410)	(0.3092)	(2.5305)		(2.3218)		
PC_2	R_1			$RV_m > 0.0101$						
	(1178 obs)	0.0012*	0.3809***	0.3218***	0.1965***	0.0092	1	19.9921***	−20619.4034	−40816
		(1.6949)	(7.4350)	(5.8186)	(3.8173)	(0.8124)		(56.1587)		
	R_2			$0.0058 < RV_m < 0.0101$ $PC_2 > 0.0070$						
	(570 obs)	0.0012	0.4867***	0.0538	−0.1434	0.3268***	1	19.9945***		
		(1.0344)	(5.8416)	(0.5422)	(−0.9292)	(3.8364)		(8.4553)		
	R_3			$0.0058 < RV_m < 0.0101$ $PC_2 \leq 0.0070$						
	(1785 obs)	0.0014***	0.2588***	0.4346***	0.1377**	0.0262**	1	4.2746***		
		(2.9096)	(6.8632)	(8.6132)	(2.2940)	(2.1569)		(6.3737)		
	R_4			$RV_m < 0.0058$						
	(1178 obs)	0.00***	0.3174***	0.2303***	0.0262	0.0177**	1	1.5958**		
		(4.7562)	(6.6052)	(4.0410)	(0.3092)	(2.5305)		(2.3218)		

[a] Table 3.5 reports estimated parameters, regimes and related statistics for daily volatility in the Tree-HAR-MIDAS model specified with alternative macroeconomic variables, including PC_1, PC_2 and ADS Index. The estimation period is January 2006-December 2014, for a total of 228 monthly macroeconomic observations and 4733 daily returns and realized volatilities. In Tree-HAR-MIDAS model:

$$\sigma_{t+1}^{(d)} = \sum_{j=1}^{N}[\beta_{0j} + \beta_{1j}RV_t^d + \beta_{2j}RV_t^w + \beta_{3j}RV_t^m + \beta_{4j}X_t]I_{X_t^{pred}\in R_j} + \upsilon_{t+1} \qquad X_t^m = \sum_{k=1}^{K}\phi_k(\omega)x_{i,t-k}$$

X_t is a weighted average value of lagged observations from one typical macroeconomic variable x_t. This macroeconomic variable has been adjusted as annual growth rates. In a Tree-HAR-MIDAS model, the regime structure is constructed by selecting a sequence of optimal threshold values from threshold set $X^{pred} = \{RV_m, r, X_t\}$, where X_T represents alternative PC_1 ,PC_2 and ADS Index. The numbers in parentheses are robust t-statistics. Statistic significance at the 1%, 5% and 10% level is indicated by ***, **, *, respectively.

Table 3.6: Parameter Estimates for the HAR-MIDAS model with Macroeconomic Uncertainty

Macro-Uncertainty	β_0	β_1	β_2	β_3	β_4	ω_1	ω_2	Loglikelihood	AIC
$\overline{\mathcal{U}}$**(3)**	−0.0006*	0.3782***	0.3464***	0.1830***	0.0163***	1	1.0362***	−19920.7186	−39723
	(−1.8489)	(11.0853)	(9.0038)	(5.6201)	(4.0785)		(3.7411)		
$\overline{\mathcal{U}}$**(6)**	−0.0012***	0.3780***	0.3462***	0.1819***	0.0224***	1	1.0146***	−19921.2846	−39727
	(−2.8145)	(11.0714)	(8.9939)	(5.5884)	(4.2348)		(3.7767)		
$\overline{\mathcal{U}}$**(12)**	−0.0027***	0.3776***	0.3457***	0.1797***	0.0370***	1	1.0333***	− 19922.2189	−39729
	(−3.5109)	(11.0663)	(8.9804)	(5.4881)	(4.2210)		(4.4988)		

[a] Table 3.6 reports estimation results for the HAR-MIDAS model specified with alternative macroeconomic uncertainty being realized over 3-month, 6-month and 12-month horizons, respectively. The HAR-MIDAS model is shown as below:

$$E_t[\sigma_{t+1}^{(d)}|F_t] = \beta_{0j} + \beta_{1j}RV_t^d + \beta_{2j}RV_t^w + \beta_{3j}RV_t^m + \beta_{4j}\overline{\mathcal{U}}(h)_t \qquad h = 3, 6, 12$$

The numbers in parentheses are robust t-statistics. Statistic significance at the 1%, 5% and 10% level is indicated by ***,**,*, respectively.

Table 3.7: Parameter Estimates for Daily Volatility in the Tree-HAR-MIDAS model with Macroeconomic Uncertainty

Macro-uncertainty	Regime structure	β_0	β_1	β_2	β_3	β_4	ω_1	ω_2	Loglikelihood	AIC
$\overline{\mathcal{U}}(3)$	R_1	$RV_m > 0.0101$								
	(1178 obs)	0.0029***	0.3788***	0.3205***	0.1920***	−0.0175*	1	11.2891***	−20649.7167	−40775
		(2.5896)	(7.4239)	(5.8213)	(3.7616)	(−1.8054)		(30.7505)		
	R_2	$0.0058 < RV_m < 0.0100$ $\bar{U}(3) > 0.0085$								
	(614 obs)	−0.0797***	0.4699***	0.2578**	0.1271	0.9312***	1	5.6633***		
		(−3.0185)	(6.4826)	(2.3085)	(1.0069)	(3.0780)		(12.8760)		
	R_3	$0.0058 < RV \leq 0.0101$ $\bar{U}(3) \leq 0.0846$								
	(1741 obs)	0.0002	0.2359***	0.3195***	0.1158*	0.0272***	1	1.0825*		
		(0.2171)	(6.6505)	(7.2675)	(1.8077)	(2.8693)		(1.8799)		
	R_4	$RV_m < 0.0846$								
	(1178 obs)	0.0027**	0.3195***	0.2352***	0.0704	−0.0068	1	2.3221		
		(2.7224)	(6.7036)	(4.2353)	(0.6585)	(−0.6614)		(1.5449)		
$\overline{\mathcal{U}}(6)$	R_1	$RV_m > 0.0100$								
	(1178 obs)	0.0034**	0.3790***	0.3206***	0.1923***	−0.0221*	1	11.1610***	−20647.8592	−40771
		(2.4882)	(7.4246)	(5.8202)	(3.7640)	(−1.7632)		(29.6900)		
	R_2	$0.0058 < RV_m < 0.0100$ $\bar{U}(6) > 0.0905$								
	(1235 obs)	−0.0736***	0.4810***	0.2713***	0.1928	0.8039***	1	5.5580***		
		(−2.6833)	(6.9031)	(2.5791)	(1.5490)	(2.7509)		(9.3489)		
	R_3	$0.0058 < RV_m < 0.0100$ $\bar{U}(6) < 0.0905$								
	(1120 obs)	−0.0004	0.2272***	0.3210***	0.1272*	0.0319**	1	1.0972*		
		(−0.3469)	(6.4005)	(7.3349)	(1.9880)	(2.4862)		(1.8335)		
	R_4	$RV_m \leq 0.0058$								
	(1178 obs)	0.0029*	0.3195***	0.2352***	0.0704	−0.0080	1	2.5086		
		(1.7855)	(6.7010)	(4.2418)	(0.6568)	(−0.4616)		(1.6769)		
$\overline{\mathcal{U}}(12)$	R_1	$RV_m > 0.0100$								
	(1178 obs)	0.0046**	0.3791 ***	0.3205***	0.1925***	−0.0330*	1	11.0545***	−20627.8085	−40731
		(2.3037)	(7.4292)	(5.8208)	(3.7668)	(−1.7352)		(61.7309)		
	R_2	$0.0058 < RV_m < 0.0100$ $\bar{U}(12) > 0.0946$								
	(967 obs)	−0.0235**	0.4630***	0.3094***	0.0136	0.2647***	1	5.5690**		
		(−2.5137)	(7.2880)	(3.6967)	(0.1298)	(2.7319)		(2.4833)		
	R_3	$0.0058 < RV \leq 0.0101$ $\bar{U}(12) \leq 0.0946$								
	(1388 obs)	0.0096**	0.2269***	0.3183***	0.0666	−0.0745*	1	18.0179***		
		(2.4375)	(5.5648)	(6.4401)	(0.9179)	(−1.8350)		(16.7050)		
	R_4	$RV_m < 0.0058$								
	(1178 obs)	0.0035	0.3195***	0.2350***	0.0756	−0.0146	1	1.9358***		
		(1.6030)	(6.6977)	(4.2379)	(0.7053)	(−0.6322)		(3.5653)		

[a] Table 3.7 reports estimated parameters, regimes and related statistics for the Tree-HAR-MIDAS model specified with macroeconomic uncertainty being realized over 3-month, 6-month and 12-month horizons, respectively.

[b] The Aggregate Uncertainty is measured by cross averaging individual uncertainty ($\bar{\mathcal{U}}$), $\overline{\mathcal{U}}(h) = \frac{1}{N}\sum_{j=1}^{N}\overline{\mathcal{U}}_j(h)$, where h specifies the uncertainty horizon in a monthly frequency.

[c] The estimation period is January 2006-December 2014, for a total of 228 monthly macroeconomic observations and 4733 daily returns and realized volatilities. In Tree-HAR-MIDAS model:

$$\sigma_{t+1}^{(d)} = \sum_{j=1}^{N}[\beta_{0j} + \beta_{1j}RV_t^d + \beta_{2j}RV_t^w + \beta_{3j}RV_t^m + \beta_{4j}X_t]I_{\overline{\mathcal{U}}_t^{pred} \in R_j} + \upsilon_{t+1} \qquad \overline{\mathcal{U}}_t^m = \sum_{k=1}^{K}\phi_k(\omega)\overline{\mathcal{U}}_{t-k}$$

$\overline{\mathcal{U}}_t$ is a weighted average value of lagged observations from one macroeconomic undertainty factor. In a Tree-HAR-MIDAS model, the regime structure is constructed by selecting a sequence of optimal threshold values from threshold set $X^{pred} = \{RV_m, r, \overline{\mathcal{U}}_t\}$. The numbers in parentheses are robust t-statistics. Statistic significance at the 1%, 5% and 10% level is indicated by ***, **, *, respectively.

Table 3.8: Model Confidence Set for one-day ahead prediction of S&P500 Volatility 2007-2014

one-week ahead forecasting ($h = 1$)	Jan 2007-Jun 2007		Jul 2007-Dec 2007		Jan 2008-Jun 2008		Jul 2008-Dec 2008		Jan 2009-Jun 2009		Jul 2009-Dec 2009		Jan 2010-Jun 2010		Jul 2010-Dec 2010	
	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}
Tree-HAR-MIDAS-IP	4.7875	0.0002	11.3255	0.7776*	17.3170	0.1780	93.4945	0.1265	14.4789	0.8416*	7.0204	0.1681	19.2756	0.2128	6.4681	0.0761
Tree-HAR-MIDAS-PPI	4.5908	0.0229	11.3946	0.7776*	18.2689	0.1780	91.3594	**1.0000***	23.1564	0.0001	6.9986	0.1681	19.0609	0.2128	6.5358	0.0761
Tree-HAR-MIDAS-UEM	5.5498	0.0002	11.5720	0.0201	17.1729	0.8284*	120.4186	0.0173	17.8880	0.0001	9.5827	0.0205	19.2698	0.2128	7.1793	0.0001
Tree-HAR-MIDAS-TS	4.5913	0.0229	11.7455	0.0201	17.1900	0.8284*	98.1227	0.1265	14.0931	**1.0000***	6.9333	0.1681	19.0070	0.2128	7.0029	0.0000
Tree-HAR-MIDAS-HS	4.2387	0.7979*	11.5621	0.0201	17.0193	**1.0000***	94.5172	0.1265	14.6363	0.8416*	6.4158	0.5300*	19.1407	0.2128	6.5059	0.0761
Tree-HAR-RV	4.2137	**1.0000***	11.0307	**1.0000***	22.5604	0.0228	115.3538	0.0173	17.0663	0.0047	6.2667	**1.0000***	17.0058	**1.0000***	5.7879	**1.0000***
	Jan 2011-Jun 2011		Jul 2011-Dec 2011		Jan 2012-Jun 2012		Jul 2012-Dec 2012		Jan 2013-Jun 2013		Jul 2013-Dec 2013		Jan 2014-Jun 2014		Jul 2014-Dec 2014	
	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}
Tree-HAR-MIDAS-IP	8.0616	0.3403*	34.8377	0.2175	6.9749	**1.0000***	7.4471	0.8696*	7.4482	0.0026	4.1753	0.0000	2.8973	0.1165	5.8237	0.0224
Tree-HAR-MIDAS-PPI	8.1997	0.1242	34.7675	0.2445	7.0230	0.9668*	7.4156	0.8696*	7.2629	0.0026	4.0557	0.0000	3.0437	0.0129	5.8101	0.0986
Tree-HAR-MIDAS-UEM	8.8865	0.0873	34.7070	0.2445	7.0030	0.9668*	7.5795	0.6607*	7.2434	0.0026	4.3117	0.0000	4.7328	0.0129	8.5258	0.0005
Tree-HAR-MIDAS-TS	8.1663	0.3403	34.9521	0.2445	7.3172	0.0336	7.6484	0.7677*	7.0265	0.0026	4.0741	0.0000	2.9221	0.0199	5.5353	0.0840
Tree-HAR-MIDAS-HS	7.8793	**1.0000***	35.6229	0.2175	7.0042	0.9668*	7.6097	0.6607*	3.4766	**1.0000***	5.5194	0.0000	3.0391	0.0199	5.3060	0.0986
Tree-HAR-RV	8.4853	0.1242	32.8065	**1.0000***	7.1452	0.9668*	7.2361	**1.0000***	6.8798	0.0026	3.2261	**1.0000***	2.4794	**1.0000***	4.6422	**1.0000***

[a] Table 3.8 summarizes relative forecasting performance for a group of Tree-HAR-MIDAS models associated with different macroeconomic variables, together with the traditional Tree-HAR model (Tree-HAR-RV). Our forecasting horizon covers a period from January 2007 to December 2014. MSEs and MCS p-values are calculated for each sub-period covering 6 months. * indicates that the model is in the (1-α) confidence interval of $\hat{M}^*_{1-\alpha}$ using for all comparisons, where α=0.25.

Table 3.9: Model Confidence Set for one-week ahead prediction of S&P500 Volatility 2007-2014

one-week ahead forecasting ($h=5$)	Jan 2007-Jun 2007		Jul 2007-Dec 2007		Jan 2008-Jun 2008		Jul 2008-Dec 2008		Jan 2009-Jun 2009		Jul 2009-Dec 2009		Jan 2010-Jun 2010		Jul 2010-Dec 2010	
	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}
Tree-HAR-MIDAS-IP	3.6805	0.2008	7.2457	0.9693*	11.9973	0.9485*	72.1839	0.2327	8.9017	0.9591*	3.3942	0.6626*	15.3524	0.1804	3.5939	0.1497
Tree-HAR-MIDAS-PPI	3.5096	**1.0000***	7.2239	0.9693*	13.4695	0.4627*	68.0459	0.5310*	13.6502	0.0038	3.5673	0.1102	14.8117	0.1804	3.5452	0.1497
Tree-HAR-MIDAS-UEM	5.4188	0.2008	7.4293	0.7287*	12.5048	0.4729*	116.9686	0.0710	11.5244	0.0038	5.8905	0.1102	15.1797	0.1804	4.4138	0.0170
Tree-HAR-MIDAS-TS	3.7722	0.2008	7.2347	0.9693*	12.0238	0.9485*	81.9662	0.2131	8.7937	**1.0000***	3.0825	**1.0000***	16.8956	0.1804	3.1472	0.6658*
Tree-HAR-MIDAS-HS	3.6255	0.2008	7.1896	**1.0000***	11.9771	0.9485*	72.9253	0.2327	18.2457	0.0038	3.4241	0.5739*	15.1694	0.1804	3.3386	0.6658*
Tree-HAR-RV	3.5913	0.6188*	7.5362	0.4535*	11.7787	**1.0000***	62.5955	**1.0000***	10.2171	0.0179	3.2127	0.6626*	13.8924	**1.0000***	2.8875	**1.0000***
	Jan 2011-Jun 2011		Jul 2011-Dec 2011		Jan 2012-Jun 2012		Jul 2012-Dec 2012		Jan 2013-Jun 2013		Jul 2013-Dec 2013		Jan 2014-Jun 2014		Jul 2014-Dec 2014	
	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}
Tree-HAR-MIDAS-IP	4.3953	0.4882*	22.1289	0.0121	2.3357	0.0052	3.5554	0.0401	4.4814	0.0080	3.0845	0.0000	2.3276	0.0175	6.7028	0.0305
Tree-HAR-MIDAS-PPI	4.5181	0.4882*	21.8149	0.0152	2.9588	0.0052	4.0149	0.0401	4.3926	0.0087	2.9214	0.0007	2.5348	0.0011	6.8094	0.0305
Tree-HAR-MIDAS-UEM	4.5041	0.4882*	21.4000	0.5605*	2.5732	0.0052	3.8676	0.0165	4.8101	0.0001	4.7562	0.0000	5.6242	0.0011	10.4652	0.0000
Tree-HAR-MIDAS-TS	4.9107	0.0268	23.5257	0.0121	2.5610	0.0052	3.7046	0.0000	4.4201	0.0001	3.4957	0.0000	1.3812	**1.0000***	2.3356	**1.0000***
Tree-HAR-MIDAS-HS	4.3236	0.6218*	22.4029	0.0152	2.1009	0.8351*	3.3553	0.3438*	3.47988	0.7159*	2.8140	0.0142	2.4938	0.0103	7.2276	0.0305
Tree-HAR-RV	3.9996	**1.0000***	21.0538	**1.0000***	2.0703	**1.0000***	2.9812	**1.0000***	3.4080	**1.0000***	1.7516	**1.0000***	1.9658	0.1069	5.1724	0.1848

[a] Table 3.9 summarizes relative forecasting performance for a group of Tree-HAR-MIDAS models associated with different macroeconomic variables, together with the traditional Tree-HAR model (Tree-HAR-RV). Our forecasting horizon covers a period from January 2007 to December 2014. MSEs and MCS p-values are calculated for each sub-period covering 6 months. * indicates that the model is in the (1-α) confidence interval of $\hat{M}^*_{1-\alpha}$ using for all comparisons, where α=0.25.

Table 3.10: Model Confidence Set for one-month ahead prediction of S&P500 Volatility 2007-2014

one-month ahead forecasting ($h = 22$)	Jan 2007-Jun 2007		Jul 2007-Dec 2007		Jan 2008-Jun 2008		Jul 2008-Dec 2008		Jan 2009-Jun 2009		Jul 2009-Dec 2009		Jan 2010-Jun 2010		Jul 2010-Dec 2010	
	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}
Tree-HAR-MIDAS-IP	3.2417	0.0333	9.6553	0.1307	9.3611	**1.0000***	79.5931	0.4059*	15.9929	0.0046	4.2470	0.2429	13.4032	**1.0000***	2.7138	0.2966*
Tree-HAR-MIDAS-PPI	2.4942	**1.0000***	10.4670	0.0275	12.4574	0.0227	71.1596	**1.0000***	8.3767	**1.0000***	2.7127	**1.0000***	13.9208	0.6903*	3.1808	0.1971
Tree-HAR-MIDAS-UEM	9.5760	0.0000	9.8355	0.1449	22.5666	0.0056	267.0487	0.1440	39.4955	0.0013	18.9730	0.0084	19.6521	0.6903*	11.1189	0.0010
Tree-HAR-MIDAS-TS	4.0065	0.0013	8.5996	**1.0000***	9.4405	0.9373*	113.3334	0.2514*	14.5783	0.1790	2.8672	0.7574*	27.8759	0.6903*	3.7180	0.1970
Tree-HAR-MIDAS-HS	3.6920	0.0000	9.8333	0.2943*	10.0309	0.0227	76.1896	0.5807*	97.9316	0.0000	3.8157	0.2429	13.4945	0.7695*	7.8910	0.1971
Tree-HAR-MIDAS-RV	2.9258	0.0909	11.8388	0.0275	12.4560	0.0227	100.0276	0.4059*	8.4179	0.9686*	2.9166	0.5534*	13.8021	0.6903*	2.5585	**1.0000***
	Jan 2011-Jun 2011		Jul 2011-Dec 2011		Jan 2012-Jun 2012		Jul 2012-Dec 2012		Jan 2013-Jun 2013		Jul 2013-Dec 2013		Jan 2014-Jun 2014		Jul 2014-Dec 2014	
	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}	MSE	P_{MCS}
Tree-HAR-MIDAS-IP	3.1392	0.8553*	21.5774	0.2521*	7.5896	0.0000	8.0662	0.0000	8.6957	0.0000	17.5576	0.0003	6.5017	0.0158	18.7198	0.0008
Tree-HAR-MIDAS-PPI	2.8986	**1.0000***	22.9189	0.2521*	2.3474	0.0000	4.2282	0.0000	3.1345	0.0012	2.6059	0.0003	2.4436	0.0434	12.0871	0.7223*
Tree-HAR-MIDAS-UEM	6.4461	0.0256	19.7049	**1.0000***	6.3556	0.0000	6.4422	0.0000	4.5378	0.0000	7.2239	0.0003	9.5715	0.0158	21.7309	0.0265
Tree-HAR-MIDAS-TS	3.5818	0.2063	21.5609	0.5167*	1.5993	0.4354*	3.2472	0.0877	3.5756	0.0012	4.0895	0.0003	3.2267	0.0158	15.2970	0.0265
Tree-HAR-MIDAS-HS	3.3226	0.6989*	21.1130	0.5167*	1.3686	0.4354*	3.4972	0.4277*	2.3904	**1.0000***	2.9147	0.0003	2.5882	0.4796*	11.7066	**1.0000***
Tree-HAR-MIDAS-RV	2.9898	0.8572*	21.2929	0.5167*	1.2875	**1.0000***	2.3040	**1.0000***	2.7300	0.0565	1.9592	**1.0000***	1.8647	**1.0000***	12.2015	0.7223*

[a] Table 3.10 summarizes relative forecasting performance for a group of Tree-HAR-MIDAS models associated with different macroeconomic variables, together with the traditional Tree-HAR model (Tree-HAR-RV). Our forecasting horizon covers a period from January 2007 to December 2014. MSEs and MCS p-values are calculated for each sub-period covering 6 months. * indicates that the model is in the (1-α) confidence interval of $\hat{M}^*_{1-\alpha}$ using for all comparisons, where α=0.25.

Table 3.11: Model Confidence Set: U.S 2007-2014

	P_{MCS}	P_{MCS}	P_{MCS}	P_{MCS}	P_{MCS}	P_{MCS}	P_{MCS}	P_{MCS}
one-week ahead forecasting (h=5)								
Volatility Prediction-Loss=MSE	Jan 2007-Jun 2007	Jul 2007-Dec 2007	Jan 2008-Jun 2008	Jul 2008-Dec 2008	Jan 2009-Jun 2009	Jul 2009-Dec 2009	Jan 2010-Jun 2010	Jul 2010-Dec 2010
Tree-HAR-MIDAS-PC1	0.9211	0.7800	0.8970	0.4931	0.2930	0.8296	0.7561	0.0029
Tree-HAR-MIDAS-ADS	0.9044	0.7341	0.8970	0.4931	**1.0000**	0.8296	0.4184	0.0822
Tree-HAR-MIDAS-$\overline{\mathcal{U}}$	**1.0000**	**1.0000**	**1.0000**	0.3870	0.2933	0.9100	**1.0000**	**1.0000**
Tree-HAR-RV	0.9044	0.1134	0.9791	**1.0000**	0.1206	**1.0000**	0.7561	0.6500
Volatility Prediction-Loss=MSE	Jan 2011-Jun 2011	2011-Dec 2011	Jan 2012-Jun 2012	Jul 2012-Dec 2012	Jan 2013-Jun 2013	Jul 2013-Dec 2013	Jan 2014-Jun 2014	Jul 2014-Dec 2014
Tree-HAR-MIDAS-PC1	**1.0000**	0.6597	0.0009	0.7252	0.3848	0.5247	0.0142	**1.0000**
Tree-HAR-MIDAS-ADS	0.3836	0.7871	0.1894	0.5680	0.0111	0.0000	0.1632	0.0483
Tree-HAR-MIDAS-$\overline{\mathcal{U}}$	0.3836	0.6599	0.3400	0.7500	**1.0000**	0.2340	0.6533	0.5214
Tree-HAR-RV	0.9428	**1.0000**	**1.0000**	**1.0000**	0.8600	**1.0000**	**1.0000**	0.2645
one-month ahead forecasting (h=22)								
Volatility Prediction-Loss=MSE	Jan 2007-Jun 2007	Jul 2007-Dec 2007	Jan 2008-Jun 2008	Jul 2008-Dec 2008	Jan 2009-Jun 2009	Jul 2009-Dec 2009	Jan 2010-Jun 2010	Jul 2010-Dec 2010
Tree-HAR-MIDAS-PC1	0.6952	0.3919	0.1696	0.6850	0.8552	0.1462	0.1596	0.0003
Tree-HAR-MIDAS-ADS	0.2886	0.7326	0.2304	0.5431	0.0123	0.1462	0.7021	0.6543
Tree-HAR-MIDAS-$\overline{\mathcal{U}}$	**1.0000**	**1.0000**	**1.0000**	**1.0000**	0.9322	**1.0000**	**1.0000**	**1.0000**
Tree-HAR-RV	0.8245	0.2560	0.0973	0.6850	**1.0000**	0.3922	0.3176	0.2940
Horizon (month)	Jan 2011-Jun 2011	2011-Dec 2011	Jan 2012-Jun 2012	Jul 2012-Dec 2012	Jan 2013-Jun 2013	Jul 2013-Dec 2013	Jan 2014-Jun 2014	Jul 2014-Dec 2014
Volatility Prediction-Loss=MSE								
Tree-HAR-MIDAS-PC1	0.0655	0.0135	0.0008	0.0119	0.9520	0.1109	**1.0000**	**1.0000**
Tree-HAR-MIDAS-ADS	0.4776	**1.0000**	0.2317	0.7200	0.9662	0.1109	0.4172	0.1131
Tree-HAR-MIDAS-$\overline{\mathcal{U}}$	0.4776	0.3756	**1.0000**	**1.0000**	0.9662	0.0618	0.7320	0.4596
Tree-HAR-RV	**1.0000**	0.9200	0.7112	0.2410	**1.0000**	**1.0000**	0.4280	0.1131

[a] Table 3.11 summarizes relative forecasting performance for a group of Tree-HAR-MIDAS models associated with alternative Principal component PC_1, ADS Index, macroeconomic uncertainty $\overline{\mathcal{U}}$, together with the traditional Tree-HAR model (Tree-HAR-RV). Our recursive one-week (h=5) and one-month (h=22) forecasting horizon entirely covers a period from January 2007 to December 2014. MSEs and MCS p-values are calculated for each sub-period covering 6 months. * indicates that the model is in the (1-α) confidence interval of $\hat{M}^*_{1-\alpha}$ using for all comparisons, where α=0.25.

Figure 3.1: S&P 500 Daily Stock Returns and Daily Realized Volatility

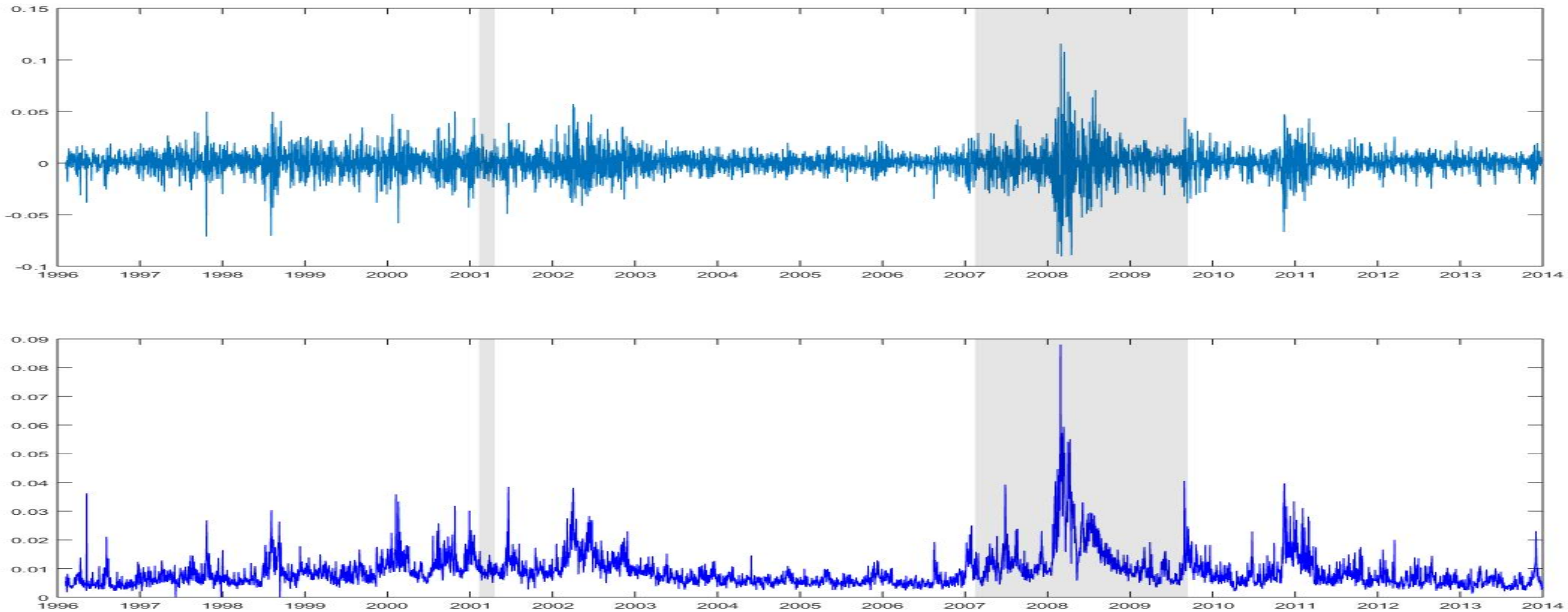
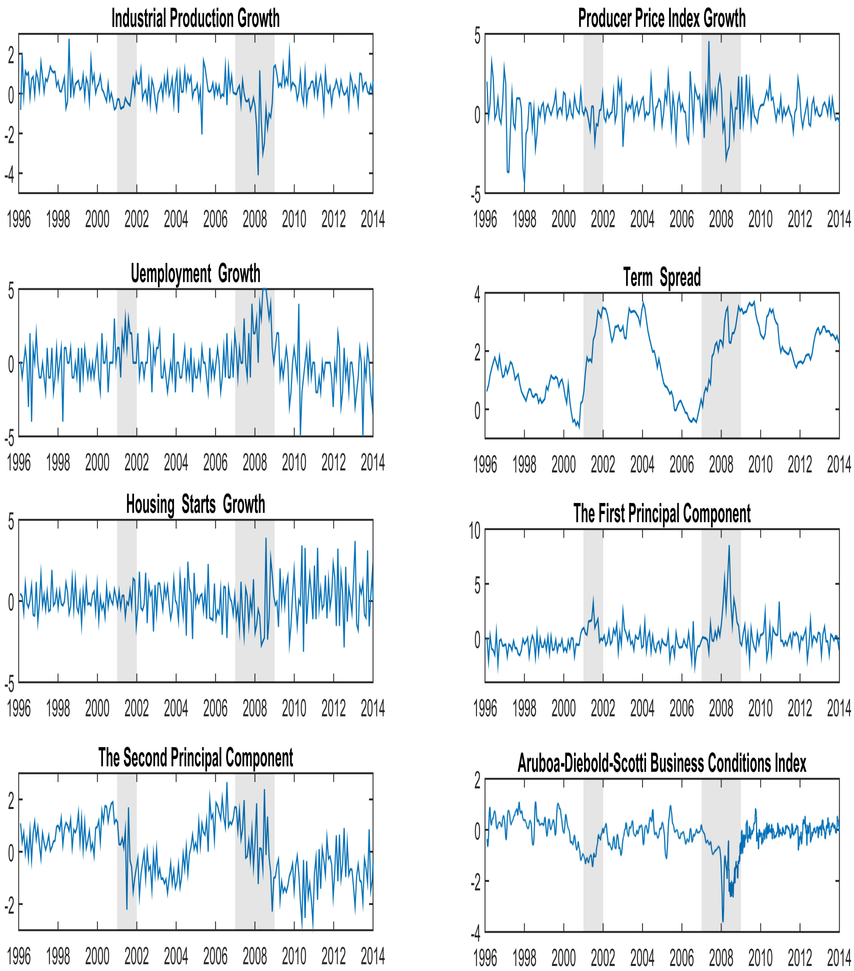


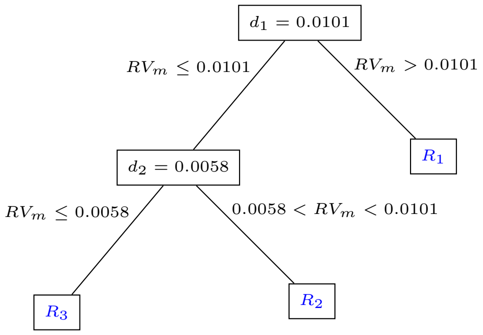
Figure 3.1 shows S&P500 daily stock returns and its daily realized volatility for the period January 1996 to December 2014. Shade areas represent NBER recession periods.

Figure 3.2: Macroeconomic Variables In The U.S Market (1996-2014)



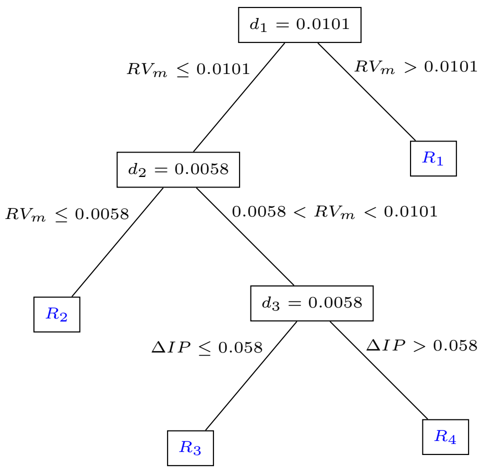
The Figure 3.2 plots monthly growth rates of macroeconomic variables including Industrial Production (IP), Producer Price Index (PPI), Unemployment Rate (UEM), Term Spread (TS) and Housing Starts (HS). Monthly Macro data ranges from Jan/1996 to Dec/2014. The first and second principal components are constructed based on monthly growth of IP, PPI, UEM, TS and HS. Aruboa-Diebold-Scotti Business Conditions (ADS) Index is obtained in a daily frequency. Gray shading areas represent NBER recession periods.

Figure 3.3: Binary Tree Structure in the Tree-HAR-MIDAS model specified with either Δ UEM, TS, or ADS Index



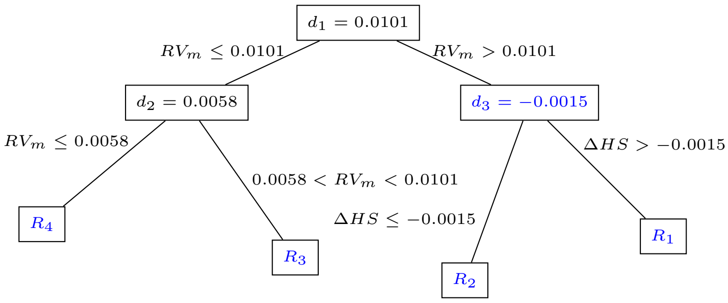
The Tree-HAR-MIDAS specified with the unemployment rate (or term spread, or ADS) maintains a Binary tree structure with three regimes, $G = \{R_1, R_2, R_3\}$. Those three regimes in this model are identified by one threshold variable, RV_m, available from the threshold set $\phi = \{r, RV_m, X\}$. $d_i, i \in \{1, 2, 3\}$ refers to the threshold value being selected from threshold variable RV_m in each partition step.

Figure 3.4: Binary Tree Structure in the Tree-HAR-MIDAS model specified with Δ IP



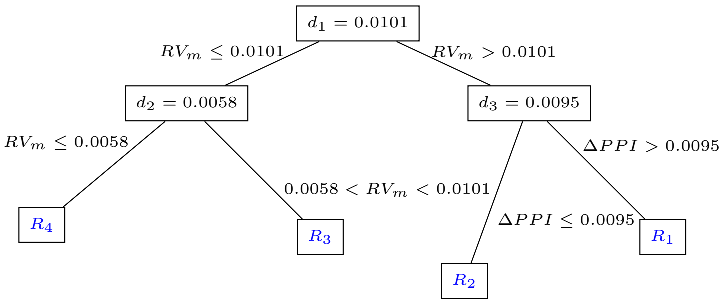
The Tree-HAR-MIDAS specified with Industrial Production growth maintains a Binary tree structure with four regimes, $G = \{R_1, R_2, R_3, R_4\}$. Those four regimes in this model are identified by two threshold variables, RV_m and ΔIP, available from the threshold set $\phi = \{r, RV_m, \Delta IP\}$. $d_i, i \in \{1, 2, 3, 4\}$ refers to the threshold value being selected from RV_m or ΔIP in each partition step.

Figure 3.5: Binary Tree Structure in the Tree-HAR-MIDAS model specified with Δ HS



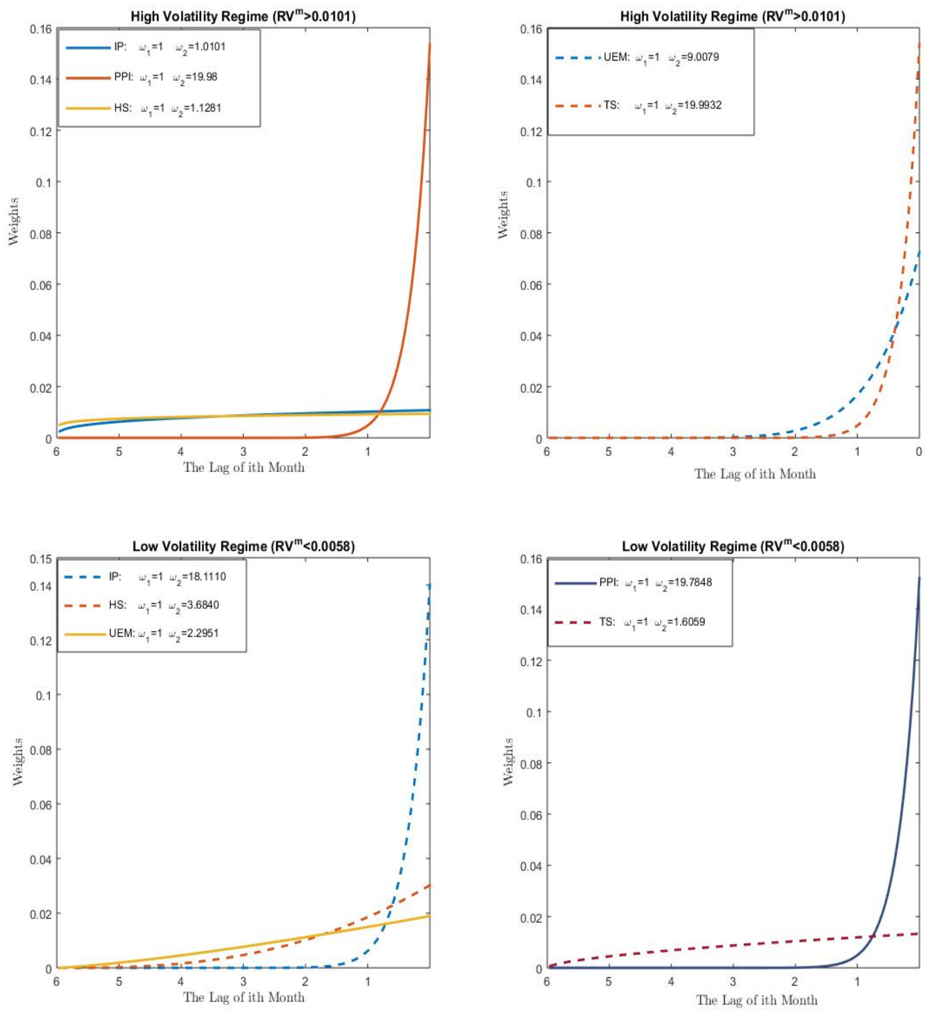
The Tree-HAR-MIDAS specified with the Housing Starts growth rate maintains a Binary tree structure with four regimes, $G = \{R_1, R_2, R_3, R_4\}$. Those four regimes in this model are identified by two threshold variables, RV_m and ΔHS, available from the threshold set $\phi = \{r, RV_m, \Delta HS\}$. $d_i, i \in \{1, 2, 3, 4\}$ refers to the threshold value being selected from RV_m or ΔHS in each partition step. Regime R_1 is edged by $\{RV_m > 0.0101 \& \Delta HS > -0.0015\}$, Regime R_2 is edged by $\{RV_m > 0.0101 \& \Delta HS \leq -0.0015\}$, Regime R_3 is edged by $\{0.0058 < RV_m < 0.0101\}$, Regime R_4 is edged by $\{RV_m \leq 0.0058\}$

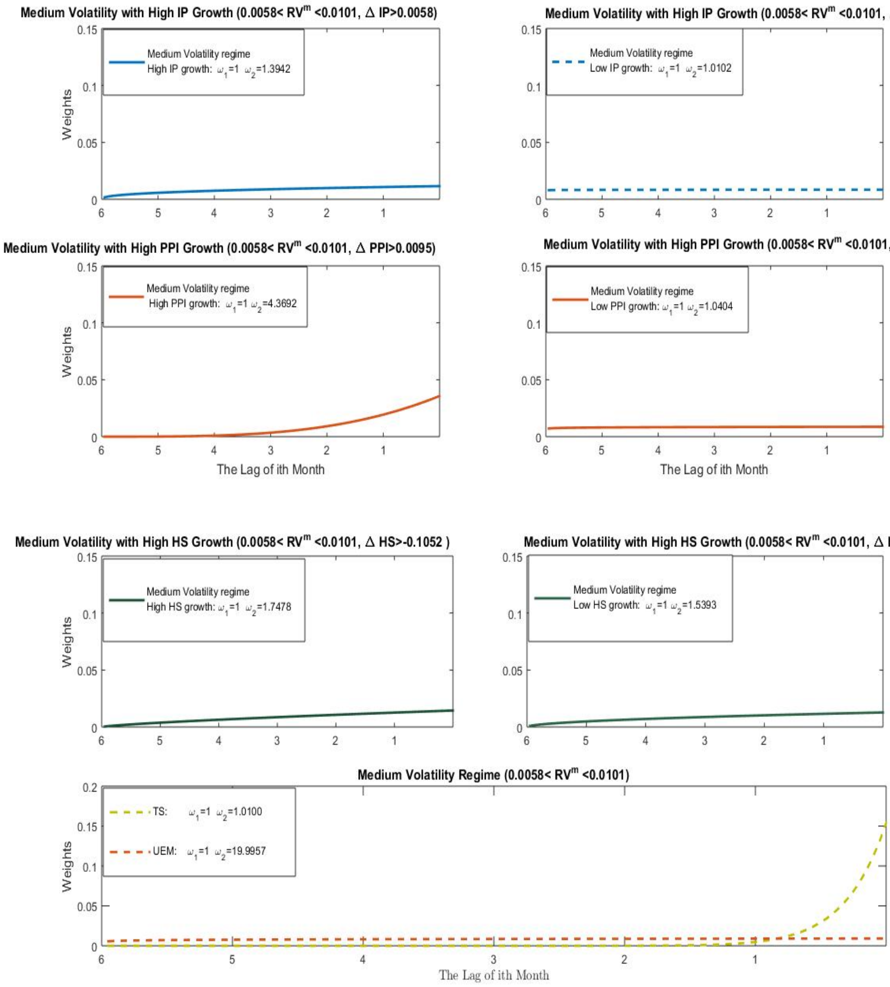
Figure 3.6: Binary Tree Structure in the Tree-HAR-MIDAS model specified with ΔPPI.



The Tree-HAR-MIDAS specified with the Producer Price Index growth rate maintains a Binary tree structure with six regimes, $G = \{R_1, R_2, R_3, R_4, R_5, R_6\}$. Those six regimes in this model are identified by three threshold variables, RV_m, r and ΔPPI available from the threshold set $\phi = \{r, RV_m, PPI\}$. $d_i, i \in \{1, 2, 3...6\}$ refers to the threshold value being selected from either of RV_m, r and ΔPPI in each partition step.Regime R_1 is edged by $\{RV_m > 0.0101 \& \Delta PPI > 0.0095\}$, Regime R_2 is edged by $\{RV_m > 0.0101 \& \Delta PPI \leq 0.0095\}$, Regime R_3 is edged by $\{0.0058 < RV_m < 0.0101\}$, Regime R_4 is edged by $\{RV_m \leq 0.0058\}$

Figure 3.7: Beta Lags in Regimes





In Figure 3.7 depicts restricted beta weighting schemes ($\omega_1 = 1$) from the Tree-HAR-MIDAS models specified with alternative macroeconomic variables within high volatility, low volatility and medium volatility regimes, alternatively. First panel refers to beta weighting schemes for alternative Tree-HAR-MIDAS models within the high-volatility regime, which is edged by $RV^m > 0.0101$. Second panel refers to beta-weighting schemes for alternative Tree-HAR-MIDAS models within the low-volatility regime, which is edged by $RV^m < 0.0058$. Then remaining parts are beta weighting schemes for alternative Tree-HAR-MIDAS models within medium volatility regime, which is edged by $0.0058 < RV^m < 0.0101$.

Figure 3.8: Regime Specification for Daily Volatility in the Tree-HAR-MIDAS model with IP Growth

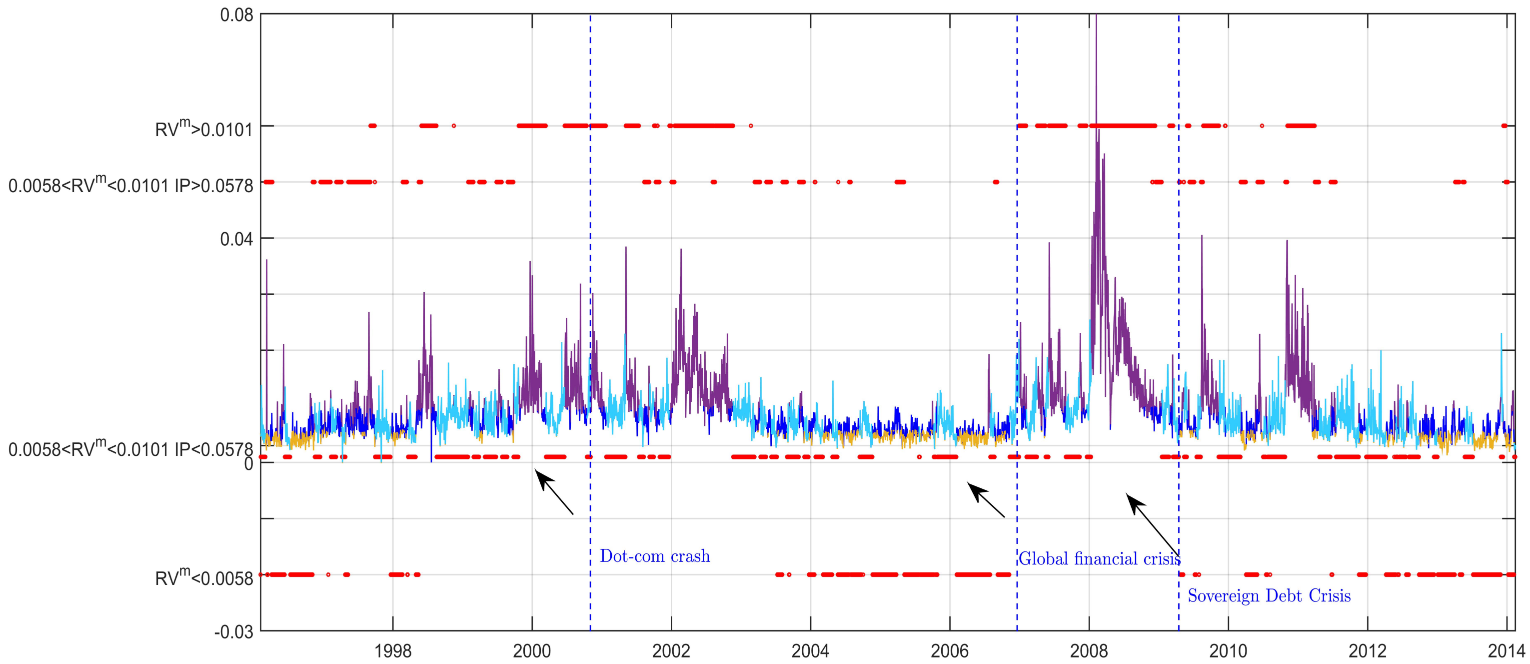


Figure 3.8 depicts daily realized volatility in a Tree-HAR-MIDAS model specified with IP growth. Optimal regimes of this Tree-HAR-MIDAS model are: High volatility regime $RV_m > 0.0101$ in color purple; medium volatility regime with high IP growth $0.0058 < RV_m < 0.0101 \quad \Delta IP > 0.0578$ in light blue; medium volatility regime with low IP growth $0.0058 < RV_m < 0.0101 \quad \Delta IP \leq 0.0578$ in dark blue; low volatility regime $RV_m < 0.0058$ in yellow.

Figure 3.9: Regime Specification for Daily Volatility in the Tree-HAR-MIDAS model with PPI Growth

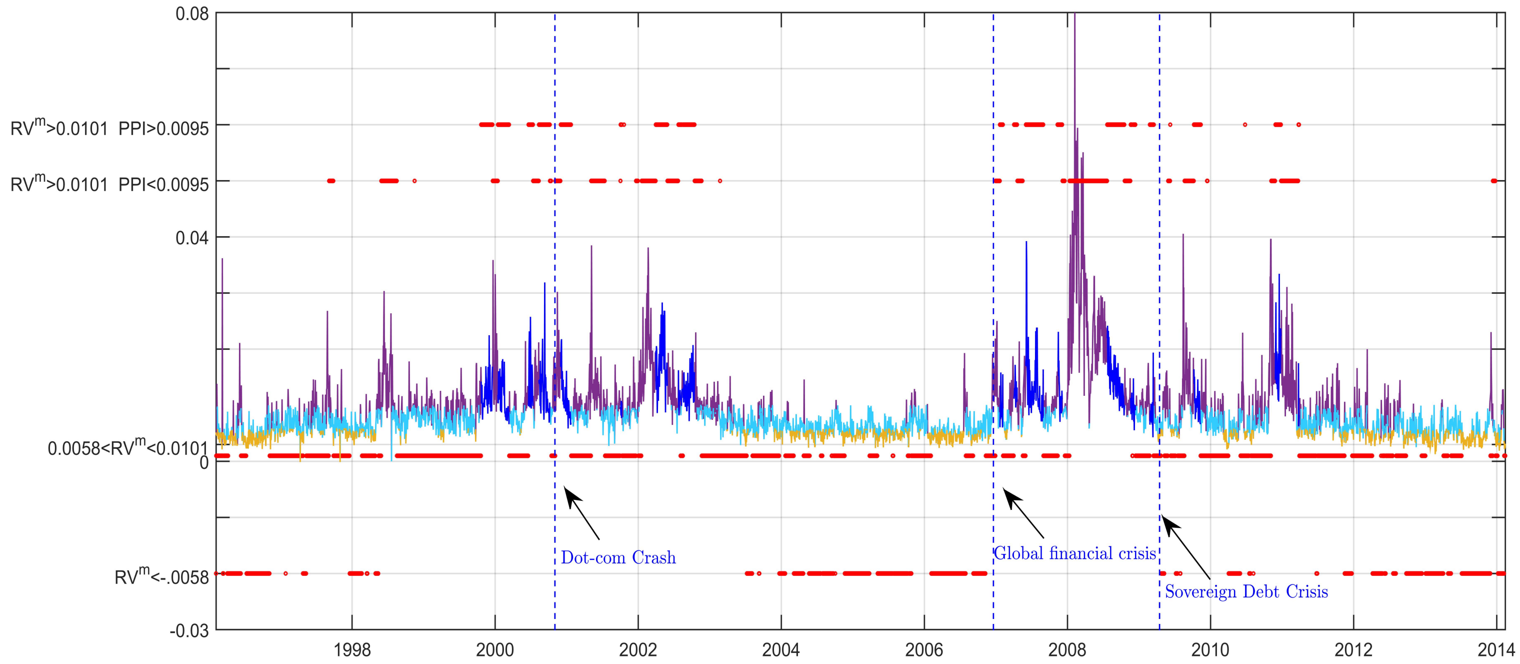


Figure 3.9 depicts daily realized volatility in a Tree-HAR-MIDAS model specified with PPI growth. Optimal regimes of this Tree-HAR-MIDAS model are: High volatility regime with high PPI growth $RV_m > 0.0101 \quad \Delta PPI > 0.0095$ in color purple; high volatility regime with low PPI growth $RV_m > 0.0101 \quad \Delta PPI \leq 0.0095$ in dark blue; medium volatility regime $0.0058 < RV_m < 0.0101$ in light blue; low volatility regime $RV_m < 0.0058$ in yellow.

Figure 3.10: Regime Specification for Daily Volatility in the Tree-HAR-MIDAS model with HS Growth

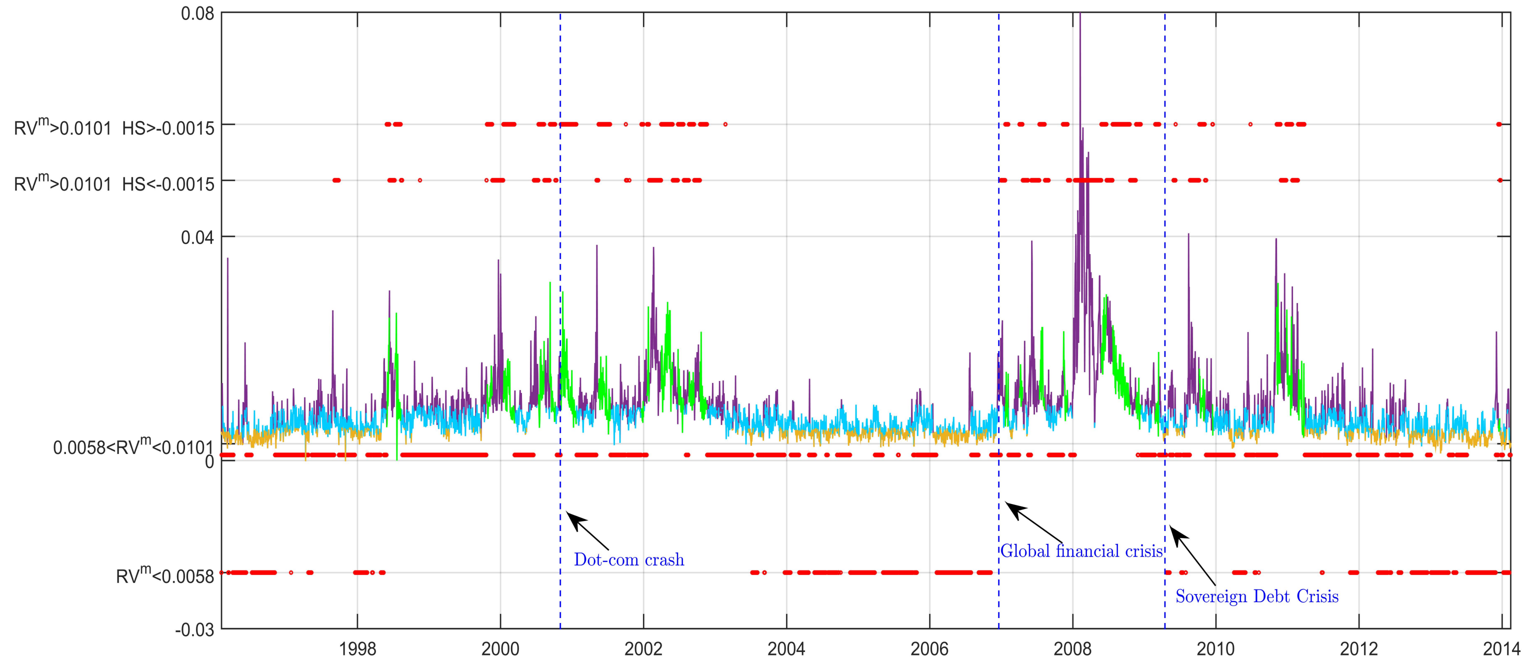


Figure 3.10 depicts daily realized volatility in a Tree-HAR-MIDAS model specified with HS growth. Optimal regimes of this Tree-HAR-MIDAS model are: High volatility regime with high HS growth $RV_m > 0.0101 \quad \Delta HS > 0.0015$ in color purple; high volatility regime with low HS growth $RV_m > 0.0101 \quad \Delta HS \leq 0.0015$ in green; medium volatility regime $0.0058 < RV_m < 0.0101$ in light blue; low volatility regime $RV_m < 0.0058$ in yellow.

Figure 3.11: Regime Specification for Daily Volatility in the Tree-HAR-MIDAS model with either ΔUEM, TS or the ADS Index

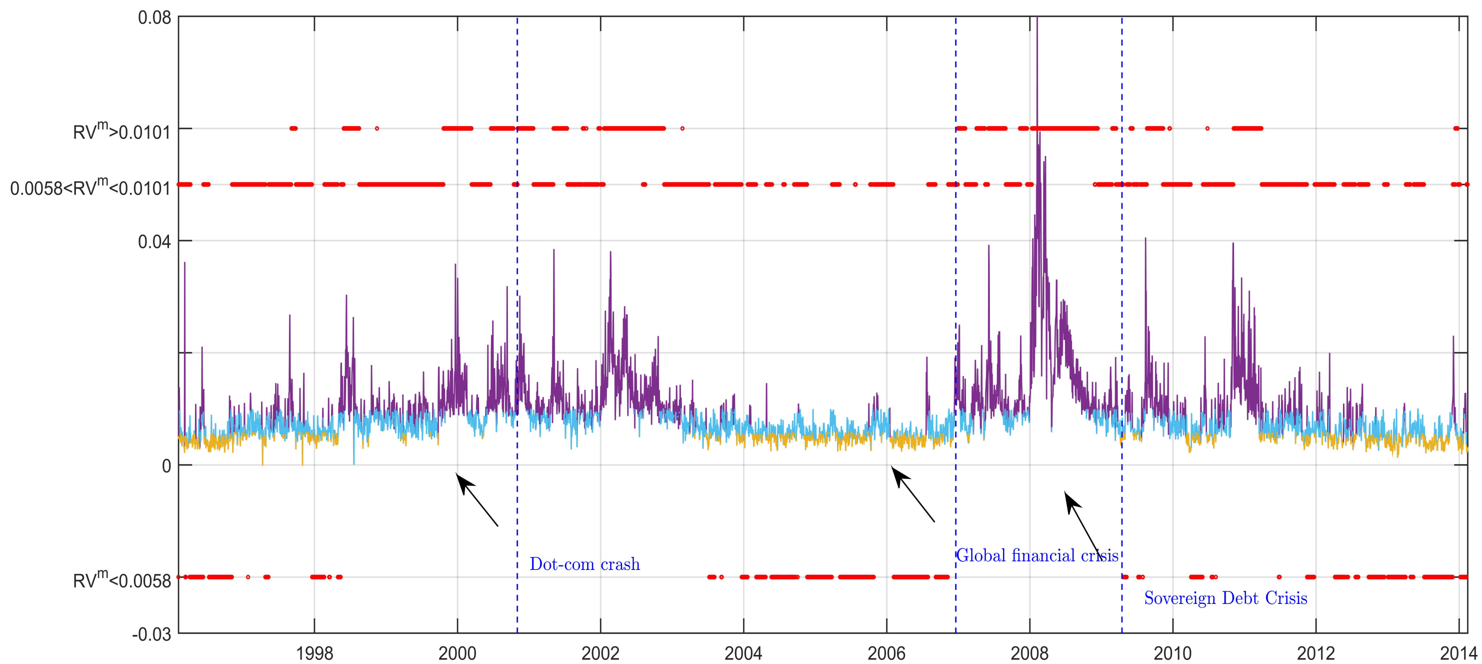


Figure 3.11 depicts daily realized volatility in a Tree-HAR-MIDAS model specified with UEM growth (or TS, ADS Index). Optimal regimes of this Tree-HAR-MIDAS model are: High volatility regime $RV_m > 0.0101$ in color purple; medium volatility regime $0.0058 < RV_m < 0.0101$ in light blue; low volatility regime $RV_m < 0.0058$ in yellow.

Figure 3.12: Regime Specification for Daily Volatility in the Tree-HAR-MIDAS model with Macroeconomic Uncertainty

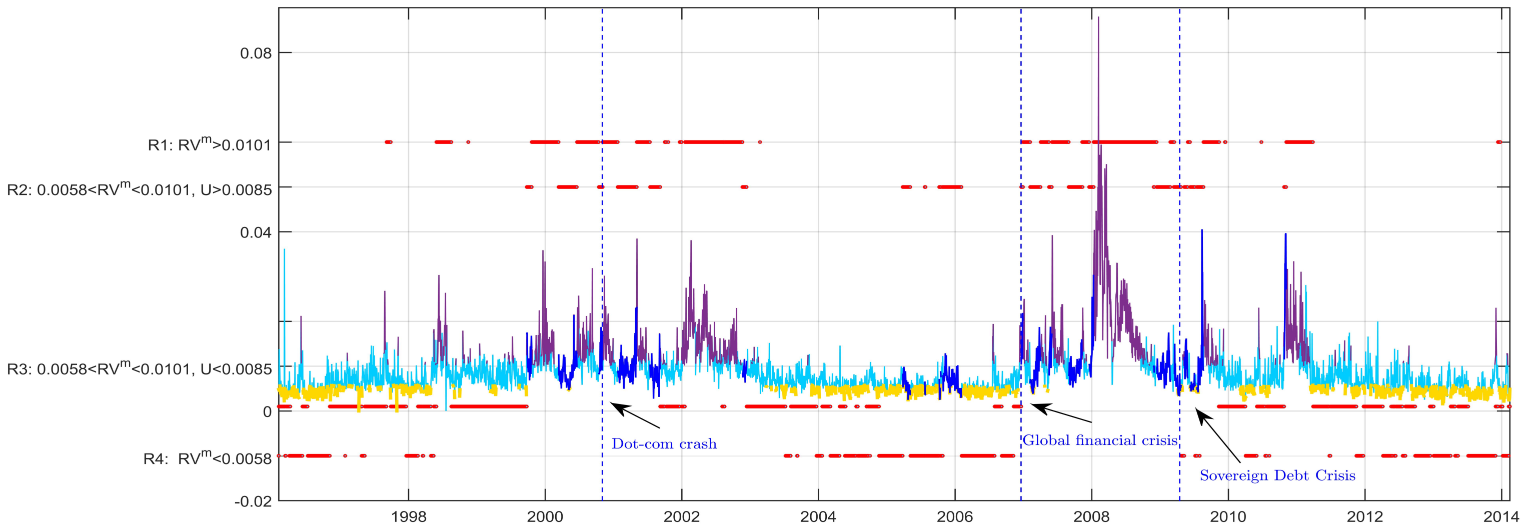
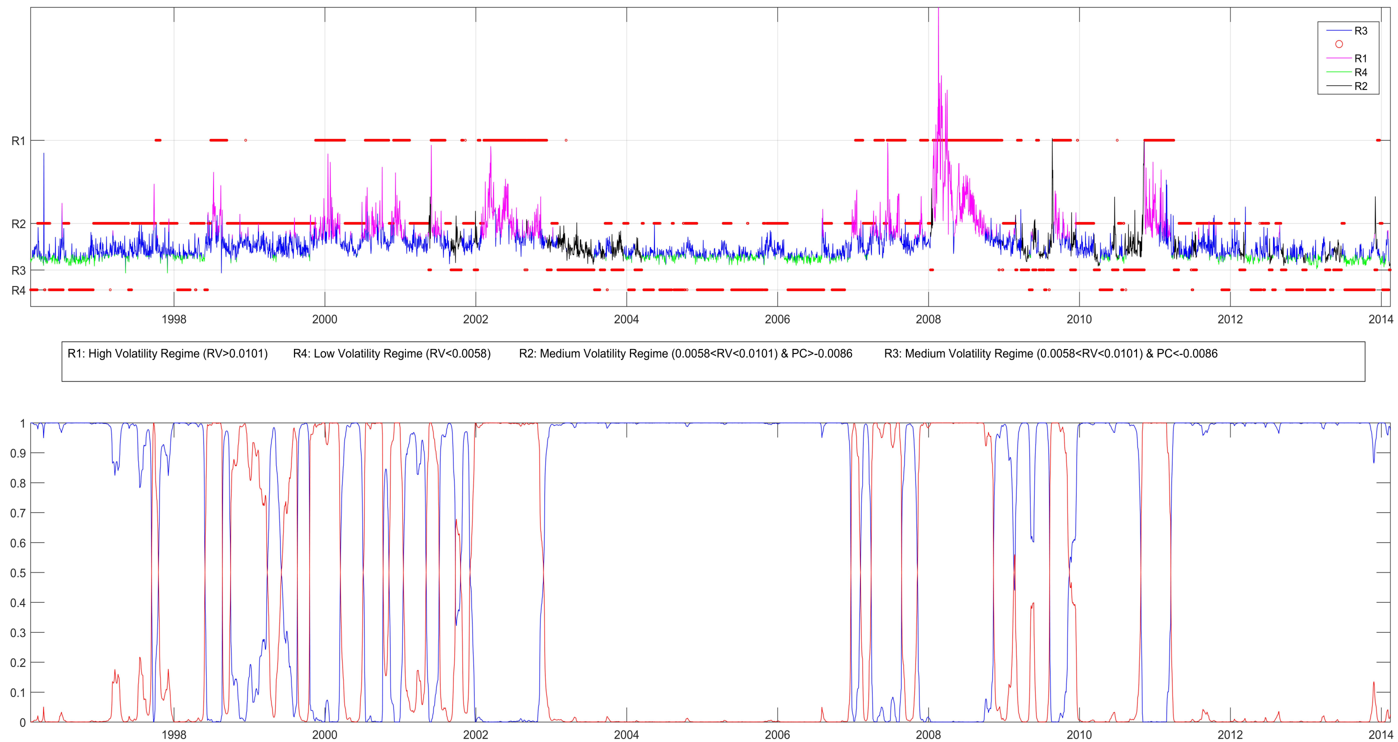


Figure 3.12 depicts daily realized volatility in a Tree-HAR-MIDAS model specified with Macroeconomic Uncertainty. Optimal regimes of this Tree-HAR-MIDAS model are: High volatility regime $RV_m > 0.0101$ in color purple; medium volatility regime with high macro uncertainty $0.0058 < RV_m < 0.0101 \quad U > 0.0085$ in dark blue; medium volatility regime with low macro uncertainty $0.0058 < RV_m < 0.0101 \quad U \leq 0.0085$ in light blue; low volatility regime $RV_m < 0.0058$ in yellow.

Figure 3.13: Regime Structure comparisons between the Tree-HAR-MIDAS model and the MS-GARCH model



Top graph depicts the Regime structure in a Tree-HAR-MIDAS model speicified with first principal component. Red line refers to the realized volatility observations that locate within one identified regime. Bottom plot depicts the Regime structure in a Markov-regime switching GARCH model. It has two regimes, where red line shows the probability of staying in the low-volatility regime while the blue line shows the probability of staying in the high-volatility regime.

Figure 3.14: Macroeconomic Uncertainty In The U.S Market (1996-2014)

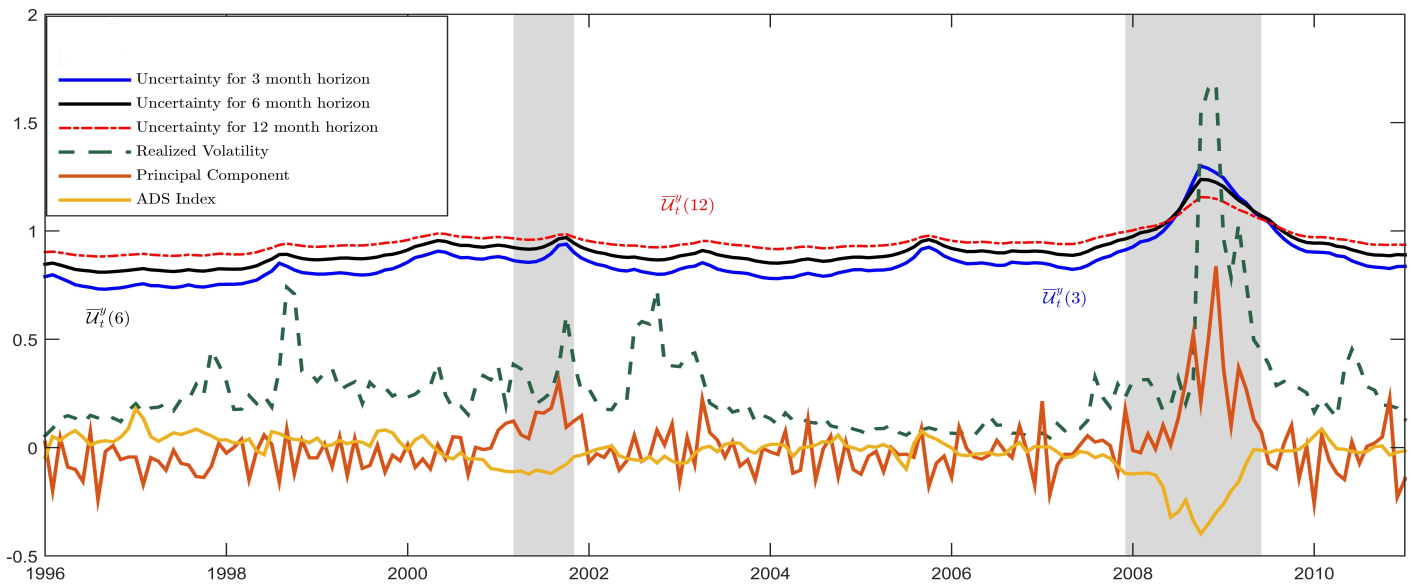


Figure 3.14 plots macroeconomic uncertainties for three forecasting horizons, 3 month, 6 month and 12 month, along with NBER recessions in shading area. The realized volatility is presented in green dot line. The principal component is presented in red line. The ADS Index is presented in yellow line. The data are monthly and span the period 1996 Jan- 2014 Dec.

References

Andersen, T. G. and T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The journal of Finance 52*(3), 975–1005.

Arneodo, A., E. Bacry, S. Manneville, and J. Muzy (1998). Analysis of random cascades using space-scale correlation functions. *Physical review letters 80*(4), 708.

Aruoba, S. B., F. X. Diebold, and C. Scotti (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics 27*(4), 417–427.

Asgharian, H., A. J. Hou, and F. Javed (2013). The importance of the macroeconomic variables in forecasting stock return variance: A garch-midas approach. *Journal of Forecasting 32*(7), 600–612.

Audrino, F. and P. Bühlmann (2001). Tree-structured generalized autoregressive conditional heteroscedastic models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63*(4), 727–744.

Audrino, F. and F. Corsi (2010). Modeling tick-by-tick realized correlations. *Computational Statistics & Data Analysis 54*(11), 2372–2382.

Beine, M. et al. (2001). Structural change and long memory in volatility: new evidence from daily exchange rates. *Developments in Forecast Combination and Portfolio Choice, Wiley Series in Quantitative Analysis, Wiley*, 145–157.

Beltratti, A. and C. Morana (2006). Breaks and persistency: macroeconomic causes of stock market volatility. *Journal of econometrics 131*(1), 151–177.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018). Really uncertain business cycles. *Econometrica 86*(3), 1031–1065.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Bollerslev, T., A. J. Patton, and R. Quaedvlieg (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics 192*(1), 1–18.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen (1984). *Classification and regression trees*. CRC press.

Breymann, W., S. Ghashghaie, and P. Talkner (2000). A stochastic cascade model for fx dynamics. *International Journal of Theoretical and Applied Finance 3*(03), 357–360.

Campbell, J. Y. (1999). Asset prices, consumption, and the business cycle. *Handbook of macroeconomics 1*, 1231–1303.

Chou, R. Y. (1988). Volatility persistence and stock valuations: Some empirical evidence using garch. *Journal of Applied Econometrics 3*(4), 279–294.

Christiansen, C., M. Schmeling, and A. Schrimpf (2012). A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics 27*(6), 956–977.

Conrad, C. and K. Loch (2015). Anticipating long-term stock market volatility. *Journal of Applied Econometrics 30*(7), 1090–1114.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, nbp001.

Dacorogna, M. M., U. A. Müller, O. V. Pictet, and R. B. Olsen (1997). Modelling short-term volatility with garch and harch models. *Available at SSRN 36960*.

Ding, Z., C. W. Granger, and R. F. Engle (1993). A long memory property of stock market returns and a new model. *Journal of empirical finance 1*(1), 83–106.

Dueker, M. J. (1997). Markov switching in garch processes and mean-reverting stock-market volatility. *Journal of Business & Economic Statistics 15*(1), 26–34.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics 95*(3), 776–797.

Engle, R. F. and A. J. Patton (2007). What good is a volatility model? In *Forecasting volatility in the financial markets*, pp. 47–63. Elsevier.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business 38*(1), 34–105.

Ghysels, E., A. Sinko, and R. Valkanov (2007). Midas regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Granger, C. W. and N. Hyung (2004). Occasional structural breaks and long memory with an application to the s&p 500 absolute stock returns. *Journal of empirical finance 11*(3), 399–421.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics 42*(1), 27–62.

Hamilton, J. D. and G. Lin (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics 11*(5), 573–593.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Inclan, C. and G. C. Tiao (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association 89*(427), 913–923.

Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *American Economic Review 105*(3), 1177–1216.

Lee, G. and R. Engle (1993). A permanent and transitory component model of stock return volatility.

Morana, C. and A. Beltratti (2004). Structural change and long-range dependence in volatility of exchange rates: either, neither or both? *Journal of Empirical Finance 11*(5), 629 – 658.

Müller, U. A., M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Pictet, and J. E. von Weizsäcker (1997). Volatilities of different time resolutions—analyzing the dynamics of market components. *Journal of Empirical Finance 4*(2), 213–239.

Muzy, J.-F., J. Delour, and E. Bacry (2000). Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *The European Physical Journal B-Condensed Matter and Complex Systems 17*(3), 537–548.

Muzy, J.-F., D. Sornette, J. Delour, and A. Arneodo (2001). Multifractal returns and hierarchical portfolio theory.

So, M. E. P., K. Lam, and W. K. Li (1998). A stochastic volatility model with markov switching. *Journal of Business & Economic Statistics 16*(2), 244–253.

Stock, J. H. and M. W. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), 147–162.

Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature 41*(3), 788–829.

Wheelock, D. C., M. E. Wohar, et al. (2009). Can the term spread predict output growth and recessions? a survey of the literature. *Federal Reserve Bank of St. Louis Review 91*(5 Part 1), 419–440.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

Chapter 4

Option Valuation using Macroeconomic Information: A Realized Volatility Approach

We develop a discrete-time realized option pricing model where the option price is partly determined by macroeconomic information. In an application using a measure of macroeconomic uncertainty, our new model outperforms alternative benchmarks such as Duan's GARCH and Corsi's HAR Option pricing models, in terms of accuracy in price and implied volatility, especially for Out-of-Money (OTM) S&P 500 Index options. Our results demonstrate that adding macroeconomic information, especially unanticipated information, improves long-term prediction for market participants. Consequently option biases that come from the Black-Scholes are able to be mitigated to some extent.

Keywords: HAR-RV, Mixed Data Sampling Approach, Macroeconomic Uncertainty, Option Pricing.

4.1 Introduction

Traditionally, Black and Scholes (1973) presume that for the same underlying asset, volatility shall be a constant value in the same maturity across different strike prices. However, empirical evidence shows that implied volatilities vary across different strike prices, and takes the shape of a smile or smirk. To address this anomaly, or bias, some studies turn to pricing options under a discrete-time volatility model, in which the conditional variance of asset returns is time-varying (see Duan (1995), Heston and Nandi (2000) for example). The time-varying volatility of asset returns can be interpreted as market response with respect to uncertainty and information flows. Therefore, it is reasonable to expect that volatility model filled with sufficient information could produce better performance in the option pricing domain. In this paper, we endeavor to explore the influence of macroeconomic information that is not perfectly correlated with asset returns on option prices. We propose a discrete-time realized volatility option pricing model, namely the HAR-MIDAS option pricing model, examining to what extent exogenous macroeconomic information, through its influence on conditional volatility of returns, can affect corresponding option prices.

It has been widely confirmed in the literature that macroeconomic information helps improve asset volatility modelling and forecasting. As both Campbell and Shiller (1988) and Schwert (1989) state, in an efficient market, the variation of arriving rates of different information flows is related to heterogeneity in asset return volatility. Macroeconomic information usually being observed in a low frequency, might be closely related with the long-run persistence in asset volatility. In line with this, Engle et al. (2013) develop a two-component GARCH-type model, the GARCH-MIDAS model. In the two-component GARCH-MIDAS model, the short-term volatility component evolves around a time-varying long-term volatility component that is driven by macroeconomic information. Engle et al. (2013) creatively adopt the MIDAS filter of Ghysels et al. (2007) to incorporate macroeconomic data which is used to explain the long-term volatility component. Following this thought, Dorion (2016) applies the GARCH-MIDAS model further for option pricing. Due to the contribution of macroeconomic information, pricing errors in the GARCH-MIDAS option pricing model are less than those observed in traditional GARCH and Black-Scholes pricing models.

Though the GARCH-MIDAS model has allowed for significant progress in linking stock volatility and macroeconomic information because of its two-component structure, studies around the contribution of macroeconomic information to asset volatility and option pricing still can be improved. In the two-component GARCH-MIDAS model, both the short-term and long-term components are measured identically by daily returns. However, the short-term component is typically associated with high-frequency intraday returns. As mentioned by Andersen and Bollerslev (1997), volatility observed during short time spans largely accounts for high-frequency intraday information. Hence, failure to use intraday returns when estimating short-term volatility persistence might artificially lead to strong persistence in the short-term volatility component. Consequently, extreme persistence of the short-term volatility component in the GARCH-MIDAS option pricing model could have a negative impact on the short-maturity option valuation. Using intraday returns for short-term volatility component construction might overcome this. We consider the effect of realized volatility constructed by high-frequency intraday returns on option pricing. Constructed using high-frequency returns, the realized volatility is able to change rapidly according to the market's movements. Hence, it is widely recognized that proper use of realized volatility generates great improvements in volatility modelling and forecasting (see Andersen and

Bollerslev (2003),Andersen et al. (2006) and Barndorff-Nielsen and Shephard (2006)). Excepting the heterogeneous autoregressive realized volatility model (hereafter HAR), to date there has been limited use made of realized volatility for option pricing. Corsi (2009) develops the HAR model that describes daily volatility as a sequence of autoregressive volatility components realized over daily, weekly and monthly horizons. These three realized components are consistent with the heterogeneous asset varations that are driven by the arrival of information at different frequencies. Application of the HAR model in option pricing (Corsi et al. (2013)) reveals that it outperforms the GARCH option pricing model, especially for short-maturity European options. The HAR model has two advantages: First, it is able to mimic the long-memory feature in volatility, which benefits the pricing of long-maturity options. Second, its multi-component structure makes it possible to incorporate low frequency macroeconomic information into a high-frequency volatility model, so as to investigate the macro-volatility relationship properly.

As seen in Chapter 3, we extended Corsi (2009)'s HAR model to the HAR-MIDAS model by including macroeconomic information in the model to explain the monthly realized volatility component. We did this by utilizing the MIDAS technique and introducing it into the original HAR model. Therefore, in a HAR-MIDAS model, the monthly realized volatility component is explained by macroeconomic variables. Empirical results in Chapter 2 show that incorporating macroeconomic information provides a better description of volatility, relative to the original HAR model. As our HAR-MIDAS model in Chapter 2 achieved significant progress in the context of volatility modelling, we use it to further explore the impact of macroeconomic information on option valuation. We are particularly interested in how and to what extent European option price can be affected by macroeconomic information.

Option valuation results suggest that most macroeconomic variables outperform Duan's GARCH model across maturity and moneyness for both put and call options. It infers a close relationship between option and market reactions towards changes in underlying economic conditions. For instance, macroeconomic variables that usually perform as recession indicators, such as term spread and unemployment rate, tends to be less effective for call option. While inflation factor, that is more active during economy expansion, seems to have limited performance for put option. Strikingly, unexpected macroeconomic information, which is measured by economic uncertainty, outperforms alternative macroeconomic variables for out-of-money (OTM) options with long-maturity. It infers that unexpected macroeconomic shock matters for option valuation, espeicially for long-maturity options.

The rest of paper is organized as follows. In section 4.2, we introduce two benchmark option models, Duan's GARCH and Corsi's HAR models. We then formally show our HAR-MIDAS option pricing model under the physical measure, and how to transfer the physical measure into a risk-neutral measure. In section 4.3, we describe the S&P 500 Index Option data and the macroeconomic data. Section 4.4 discusses a new measure of macroeconomic uncertainty due to Jurado et al. (2015). Section 4.5 presents results on option pricing performance for the HAR-MIDAS model specified with alternative macroeconomic variables and macro uncertainty. We then compare them with two benchmark models mentioned in Section 4.2. Finally, Section 4.6 summarizes and concludes.

4.2 Discrete-time Option Pricing model

In this section we first present two discrete-time option pricing models: Duan's GARCH(1,1) model and Corsi's HAR-RV model, both of which serve as benchmark models throughout our study. We then introduce our HAR-MIDAS model as a further extension of the HAR-RV model, in which macroeconomic information is incorporated to explain volatility movements and the pricing kernel.

4.2.1 Duan's GARCH Option Pricing Model

Duan (1995) derives an option pricing model where asset returns follow a GARCH process. Under the physical measure (hereafter P measure), the asset spot price is assumed to be lognormally distributed:

$$ln(S_t/S_{t-1}) = r + \lambda\sqrt{h_t} - \gamma_t + \varepsilon_t \tag{4.1}$$

where excess return $ln(S_t/S_{t-1}) - r$ is determined by the risk premium $\lambda\sqrt{h_t}$. In line with Cox et al. (2005), λ measures a proportion of volatility risk that is associated with excess return. Parameter γ_t serves as a mean correction factor, which satisfies:

$$\exp(\gamma_t) = E_{t-1}\left[\exp(\varepsilon_t)\right] \tag{4.2}$$

Adding γ_t ensures that the conditional expected gross return equals $\exp(r + \lambda\sqrt{h_t})$:

$$E_{t-1}\left[ln(S_t/S_{t-1})\right] = E_{t-1}\left[r + \lambda\sqrt{h_t}\right] \tag{4.3}$$

Therefore γ_t must be equal to $\frac{1}{2}h_t$, and 4.1 under the P measure now becomes:

$$ln(S_t/S_{t-1}) = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \varepsilon_t \tag{4.4}$$

Proof: see Appendix A1

Conditional variance of asset returns h_t is assumed to follow a GARCH(1,1) process:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{4.5}$$

where parameter α_1 determines the kurtosis of distribution for asset return, and the conditional variance satisfies the stationarity constraint $\beta_1 + \alpha_1 \gamma_1^2 \leq 1$.

Following the equivalent martingale measure (EMM), Duan (1995) transfers both log-return (Eq.4.4) and conditional volatility (Eq.4.5) from the physical measure into the risk-neutral measure (the Q meaure). The GARCH option pricing model under the Q measure is:

$$\begin{aligned} \ln(\frac{S_t}{S_{t-1}}) &= r - \frac{1}{2} h_t + \varepsilon_t^* \\ h_t &= \alpha_0 + \alpha_1 (\varepsilon_{t-1}^* - \lambda \sqrt{h_{t-1}})^2 + \beta_1 h_{t-1} \end{aligned} \tag{4.6}$$

Proof: see Appendix A2

where innovation ε_t^* under the Q measure satisfies $\varepsilon_t^* = \varepsilon_t + \lambda$.

4.2.2 Return with decomposed realized volatility

Instead of using latent volatility, Corsi (2009) uses the implied realized volatility. As noted before, realized volatility with high-frequency trading information is able to react quickly with respect to new information arriving and also better capture the long-memory characteristic for option valuation. In this subsection, we will present Corsi's HAR option pricing model and then introduce our extension that incorporates macroeconomic information for option pricing under the HAR framework.

A. The HAR-RV Option Pricing Model

Based on the heterogeneous market hypothesis of Müller et al. (1997), Corsi (2009) models daily volatility with a multi-components structure, in which daily volatility follows an autoregressive process with a sequence of volatility components being realized over daily, weekly and monthly horizons (the Heterogenous Autoregressive Realized Volatility, or HAR-RV, model). Constructed from high frequency intra-daily returns, realized volatility is able to change rapidly with respect to new arrival information in financial markets. Hence, the HAR-RV model outperforms the traditional GARCH volatility model in terms of prediction ability. In addition, the HAR-RV model is able to mimic the long-memory feature of volatility movements. With those advantages, Corsi et al. (2013) utilize the HAR-RV model to develop an option pricing model. Using RV as a proxy for the variance, log-returns in the HAR-RV model evolve as:

$$\ln\left[\frac{S_{t+1}}{S_t}\right] = r + \gamma RV_{t+1} + \sqrt{RV_{t+1}}\varepsilon_{t+1} \tag{4.7}$$

where parameter γ can be decomposed into two parts, $\tilde{\gamma}$ and $-\frac{1}{2}$. $\tilde{\gamma}$ is the price of risk for the conditional variance while $-\frac{1}{2}$ is a mean correction factor as mentioned in Eq.4.3.

Under the physical measure, RV_{t+1} in a HAR-RV process follows a noncentral gamma distribution with shape and scale parameters δ and c:

$$\begin{aligned} &RV_{t+1}|F_t \sim \Gamma(\delta, \beta'(RV_t, L_t), c) \\ &\beta'(RV_t, L_t) = \beta_0 + \beta_1 RV_t^d + \beta_2 RV_t^w + \beta_3 RV_t^m + \beta_4 L_t \end{aligned} \tag{4.8}$$

In Eq.4.8, daily RV_{t+1} conditioned on past information set F_t being available until date t is decomposed into three non-overlapping components: RV_t^d

(short-term variance component measured at the daily frequency), RV_t^w (medium variance component measured at the weekly frequency) and RV_t^m (long-term variance component measured at the monthly frequency) together with a leverage factor L_t that accounts for asymmetric effects. Asymmetric effect refers to a stylized fact that (realized) volatility usually reacts more after negative shock rather than positive shock with same magnitude. Therefore, Corsi et al. (2013) lets $L_t = I_{(y_t<0)} RV_t$, where $I_{(y_t<0)}$ as a dummy variable takes one unit if return y_t takes negative value, otherwise takes zero if return y_t takes non-negative value. Former three components in Eq.4.8 are recursively calculated by aggregating daily RV over a specified time-horizon:

$$RV_t^w = \frac{1}{4}\sum_{i=1}^{4} RV_{t-i}^d \qquad RV_t^m = \frac{1}{17}\sum_{i=5}^{21} RV_{t-i}^d \tag{4.9}$$

As RV_{t+1}^d follows a noncentral gamma distribution, its conditional mean and variance are defined as:

$$\begin{aligned} E_t(RV_{t+1}^d) =& c\delta + c(\beta_1 RV_t^d + \beta_2 RV_t^w + \beta_3 RV_t^m + \beta_4 L_t) \\ V_t(RV_{t+1}^d) =& c^2\delta + 2c^2(\beta_1 RV_t^d + \beta_2 RV_t^w + \beta_3 RV_t^m + \beta_4 L_t) \end{aligned} \tag{4.10}$$

Combining Eq. 4.7 and Eq. 4.8, Corsi et al. (2013) have the Heterogeneous Autoregressive Gamma with Leverage (HARGL) option model. Using a discrete-time exponential affine stochastic discount factor (SDF) as in Gagliardini et al. (2011), parameters in the HARGL process can be mapped into risk-neutral measures as follows:

$$ln(S_{t+1}/S_t)|RV_{t+1} \sim N(r + \gamma RV_{t+1}, \sqrt{RV_{t+1}}),$$

$$RV_{t+1}|F_t \sim \Gamma(\delta^*,\ \beta^*(RV_t, L_t),\ c^*). \tag{4.11}$$

$$\beta^*(RV_t, L_t) = \beta_1 RV_t^d + \beta_2 RV_t^w + \beta_3 RV_t^m + \beta_4 L_t^m.$$

where $\beta^* = \frac{\beta}{(1+c\lambda)^2}$, $\delta^* = \delta$, $c^* = \frac{c}{1+c\lambda}$.

B. HAR-MIDAS option model

We extend the HAR option model (Corsi et al. (2013)) to include macroeconomic information, examining to what extent option pricing biases can be further mitigated, and whether pricing accuracy can be improved. To achieve this, we first assume that the growth in a macroeconomic variable Δx follows a process with constant mean and dynamic variance:

$$x_{t+1} = \mu_x + \sigma_{x,t+1} u_{t+1} \tag{4.12}$$

where $u_{t+1} \sim \mathcal{N}(0, 1)$ are serially uncorrelated innovations with zero mean and unit variance. Second, we include macroeconomic volatility rather than the macro variable itself, into the HAR option model, so as to accommodate the non-centralized gamma distribution the HAR-RV process follows (see Equation 4.11). We employ the mixed data sampling (MIDAS) approach from Ghysels et al. (2007). For our model here, the macroeconomic information is calculated as:

$$
\begin{aligned}
X^m_{t+1} &= \sum_{k=1}^{K} \varphi_k(\omega_1)\hat{\sigma}^m_{x,t+1-k} \\
\varphi_k(\omega_1,\omega_2) &= \frac{(k/K)^{\omega_1-1}(1-k/K)^{\omega_2-1}}{\sum_{j=1}^{K}(j/K)^{\omega_1-1}(1-j/K)^{\omega_1-1}}
\end{aligned}
\tag{4.13}
$$

where $\hat{\sigma}_{x,t+1}$ is the estimated conditional variance of the chosen macro variable. X^m_{t+1} can be thought of as a weighted average of lags of variations in the macroeconomic variable x^m_{t+1} with a monthly frequency. The weights allocated to lags of $\hat{\sigma}^m_{x,t+1}$ are governed by a beta weighting scheme $\varphi_k(\omega_1,\omega_2)$, which satisfies the conditions $\varphi_k > 0$ and $\sum_{k=1}^{K}\varphi_k = 1$. In the beta weighting scheme, parameter ω_1 is restricted to one unit, making sure a decaying pattern for lags of weights. To avoide non-negative value, parameter ω_2 is restricted as $\omega_2 > 1$. [1] Following Christoffersen et al. (2009), we assume shocks to the market come from two main sources: macroeconomic shocks, and "pure-market" shocks that are unrelated to macro information. Therefore, ε_{t+1} is decomposed into:

$$
\varepsilon_{t+1} = \rho u_{t+1} + \sqrt{1-\rho^2} z_{t+1} \tag{4.14}
$$

where the "pure-market" shock, z_{t+1}, has zero mean and unit standard deviation. The parameter ρ refers to the correlation coefficient between market returns and innovations in the macroeconomic variable. We can now generate a discrete-time option pricing model with macroeconomic information under the P-measure, which we call the HAR-MIDAS option model:

[1] Since ω_2 is restricted as $\omega_2 > 1$, the null hypothesis is $H_0 : \omega_2 \leq 1$.

$$
\begin{aligned}
&x_{t+1} = \mu_x + \sigma_{x,t+1}u_{t+1} \\
&ln(S_{t+1}/S_t) = r + \gamma RV_{t+1} + \sqrt{RV_{t+1}}\varepsilon_{t+1} \quad \text{where} \quad \varepsilon_{t+1} = \rho u_{t+1} + \sqrt{1-\rho^2}z_{t+1} \\
&RV^d_{t+1}|F_t \sim \Gamma(\delta,\ \beta'(RV_t, X^m_t),\ c) \\
&\beta'(RV_t, X^m_t) = \beta_1 RV^d_t + \beta_2 RV^w_t + \beta_3 RV^m_t + \beta_4 X^m_t \quad \text{where} \quad X^m_{t+1} = \sum_{k=1}^{K} \varphi_k(1,\omega_2)\hat{\sigma}^m_{x,t+1-k}
\end{aligned}
\tag{4.15}
$$

In Equation 4.15, daily realized variance RV^d_{t+1} is modelled as an autoregressive function $\beta'(RV_t, X^m_t)$ with three realized variance components (RV^d_t, RV^w_t, RV^m_t) together with a weighted average value of variations (X^m_t) from the macroeconomic variable x_t.

4.2.3 Risk Neutralization for the HAR-MIDAS model

In previous literature, the risk neutralization for discrete-time option models such as Duan's GARCH and Heston's GARCH usually follows the equivalent martingale measure (EMM) of Christoffersen et al. (2009), in which only stock return innovations are considered. In our case, however, we need to take both log-return and realized volatility into consideration when generating risk neutralization procedure. Christoffersen et al. (2009) provide the moment generating function (MGF) for a joint process of RV and log-return under the HAR-RV framework:

$$
\begin{aligned}
\varphi^P_{RV}(\eta) =& E\left[exp(-\eta RV_{t+1}|(RV_t))\right] \\
=& exp\left(-\frac{c\eta}{1+c\eta}(\beta'(RV_t) - \delta ln(1+c\eta))\right)
\end{aligned}
\tag{4.16}
$$

where $\eta \in R$ and $\beta'(RV_t)$ refers to the HAR-RV process shown in Equation 4.8.

The MGF of the HAR-RV model in Equation 4.16 allows us to derive a new MGF that transfers the joint process of log-returns, RV and macro volatility from the physical measure into the risk-neutralized measure. First, we set up a joint process $\Phi'_{t+1} = [\beta'(RV_t), \beta'(\hat{\sigma}_{x,t}), y_{t+1}]$ which is a multi-dimensional real-value process of Realized volatility (RV), macro volatility$(\hat{\sigma}_x)$ and log-return (y_{t+1}). Propostion 1 defines a closed-form expression for the MGF of Φ'_{t+1}.

Proposition 1. *In the HAR-MIDAS model, if $RV_{t+1}|F_t \sim \Gamma(\delta,\ \beta'(RV_t, X_t),\ c)$, then the closed-form expression for the MGF of $\Pi'_{t+1} = (\beta'(RV_t), \beta'(X_t), y_{t+1})$ in $\alpha' = (\alpha_1, \alpha_2, \alpha_3)$ under P measure is:*

$$
\begin{aligned}
&E_t^P[\exp\{-\alpha'\Pi_{t+1}\}] \\
&= E_t^P\left[\exp\{-\alpha_1\beta'(RV_t)) - \alpha_2\beta'(X_t) - \alpha_3 y_{t+1}\}\right] \qquad (4.17) \\
&= \phi_{RV}^P(\vartheta_1)\phi_X^P(\vartheta_2)
\end{aligned}
$$

with

$$
\beta'(RV_t, X_t) = \underbrace{\beta_1 RV_t^d + \beta_2 RV_t^w + \beta_3 RV_t^m}_{\beta'(RV_t)} + \underbrace{\beta_4 X_t}_{\beta'(X_t)}
$$

$$
\phi_{RV}^P(\vartheta_1) = \exp\{-b(\vartheta_1) - a(\vartheta_1)\beta'(RV_t)\} \qquad \phi_X^P(\vartheta_2) = \exp\{-b(\vartheta_2) - a(\vartheta_2)\beta'(X_t)\} \tag{4.18}
$$

Proof: see Appendix A3

where $\beta'(RV_t, X_t)$ is divided into two parts: $\beta'(X_t)$ macroeconomic weighted average value X_t, and remaining part of $\beta'(RV_t)$ consists of RV_t^d, RV_t^w and RV_t^m. There are two terms $b(\vartheta_i)$ and $a(\vartheta_i)$ $i = 1, 2$ in Eqs (4.17-4.18) are expressed as below:

$$b(\vartheta) = \delta \ln(1 + c\vartheta) \quad a(\vartheta) = \frac{c\vartheta}{1 + c\vartheta}$$

where $\vartheta_1 = \left(\alpha_1 + \gamma\alpha_3 - \frac{1}{2}\alpha_3^2\right)$ and $\vartheta_2 = \left(\alpha_2 + \gamma\alpha_3 - \frac{1}{2}\alpha_3^2\right)$ in Eq.4.18.

Following Gagliardini et al. (2011) and Corsi et al. (2013), we specify a stochastic discount factor (SDF) that includes log-return, RV and macro volatility:

$$M_{t,t+1} = \frac{M_{t+1}}{E_t^P[M_{t+1}]}$$

$$= \frac{exp\{-m_1 RV_{t+1}^{RV} - m_2 RV_{t+1}^{X} - m_3 y_{t+1}\}}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}$$

$$\phi_{RV}^P(u_1) = \exp\{-b(u_1) - a(u_1)\beta'(RV_t)\} \text{ with } u_1 = m_1 + \gamma m_3 - \frac{1}{2}m_3^2$$

$$\phi_X^P(u_2) = \exp\{-b(u_2) - a(u_2)\beta'(X_t)\} \text{ with } u_2 = m_2 + \gamma m_3 - \frac{1}{2}m_3^2 \tag{4.19}$$

Proof of Equation 4.19: see Appendix A4

where $E_t^P[\,]$ is the conditional expectation at time t under the P measure. In order to satisfy the no-arbitrage restriction:

$$E_t^P[M_{t,t+1}\exp\{y_{t+1}\}] = 1 \tag{4.20}$$

we plug the stochastic discount factor (SDF) Eq.4.19 into the no-arbitrage condition Eq.4.20 to give:

$$m_3 = \gamma + \frac{1}{2} \tag{4.21}$$

Proof of Eq. 4.21: see Appendix A5

We can then convert the HAR-MIDAS option model under the P measure into that under the Q measure:

$$E_t^Q[\exp(-\alpha'\Pi_{t+1})] = E_t^P[M_{t,t+1}\exp(-\alpha'\Pi_{t+1})]$$

$$= E_t^P[M_{t,t+1}\exp\{-\alpha_1\beta'(RV_t) - \alpha_2\beta'(X_t) - \alpha_3 y_{t+1}\}] \tag{4.22}$$

$$= \frac{\phi_{RV}^P(\varpi_1)\phi_X^P(\varpi_2)}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}$$

where

$$u_1 = m_1 + \gamma m_3 - \frac{1}{2}m_3^2$$
$$u_2 = m_2 + \gamma m_3 - \frac{1}{2}m_3^2$$
$$\varpi_1 = (m_1 + \alpha_1) + \gamma(m_3 + \alpha_3) - \frac{1}{2}(m_3 + \alpha_3)^2$$
$$\varpi_2 = (m_2 + \alpha_2) + \gamma(m_3 + \alpha_3) - \frac{1}{2}(m_3 + \alpha_3)^2$$

Plugging $m_3 = \gamma + \frac{1}{2}$ into Eq. 4.22, we have

$$E_t^Q[\exp(-\alpha'\Pi_{t+1})] = E_t^P[M_{t,t+1}exp(-\alpha'\Pi_{t+1})]$$

$$= \frac{\phi_{RV}^P(\varpi_1)\phi_X^P(\varpi_2)}{\phi_{RV}^P(u_1)\phi_X^P(u_2)} \tag{4.23}$$

$$= \frac{\phi_{RV}^P(\varsigma_1+\lambda_1)\phi_X^P(\varsigma_2+\lambda_2)}{\phi_{RV}^P(\lambda_1)\phi_X^P(\lambda_2)}$$

where ϖ_i, $i = 1, 2$ is:

$$\varpi_1 = \underbrace{m_1 + \frac{1}{2}\gamma^2 - \frac{1}{8}}_{\lambda_1} + \underbrace{\alpha_1 - \frac{1}{2}\alpha_3 - \frac{1}{2}\alpha_3^2}_{\varsigma_1}$$

and

$$\varpi_2 = \underbrace{m_2 + \frac{1}{2}\gamma^2 - \frac{1}{8}}_{\lambda_2} + \underbrace{\alpha_2 - \frac{1}{2}\alpha_3 - \frac{1}{2}\alpha_3^2}_{\varsigma_2}$$

Proof of Eq. 4.23: see in Appendix A6

Relying on the SDF which is compatible with the no-arbitrage condition, we can derive the risk-neutral specification of HAR-MIDAS option model:

Proposition 2. *Under the risk neutral probability measure Q, the HAR-MIDAS option pricing model has the following parameters:*

$$\beta'^*_{RV} = \frac{\beta}{1 + c_1^* \lambda_1} \quad with \quad c_1^* = \frac{c}{1 + c\lambda_1} \quad \& \quad \lambda_1 = \alpha_1 - \frac{1}{2}\alpha_3 - \frac{1}{2}\alpha_3^2$$

$$\beta'^*_{X} = \frac{\beta}{1 + c_2^* \lambda_2} \quad with \quad c_2^* = \frac{c}{1 + c\lambda_2} \quad \& \quad \lambda_2 = \alpha_2 - \frac{1}{2}\alpha_3 - \frac{1}{2}\alpha_3^2$$

$$u_t^* = u_t + \rho\frac{1}{4}RV_t \qquad z_t^* = z_t + \sqrt{1-\rho^2}\frac{1}{4}RV_t \tag{4.24}$$

$$\delta^* = \delta \qquad \gamma^* = -\frac{1}{2}$$

$$\varepsilon_t^* = \rho u_t^* + \sqrt{1-\rho^2} z_t^* = \varepsilon_t + RV_t(\frac{1}{2} + \gamma)$$

Proof of Proposition 2: see in Appendix A7

4.2.4 Estimation for the HAR-MIDAS Option Model

Following Gourieroux and Jasiak (2006) and Corsi et al. (2013), the HAR-MIDAS process follows a non-centred gamma distribution which is characterized by three parameters: the degrees of freedom δ, the non-centrality parameter $\beta'(RV_t, X_t)$ and the scale parameter c. This non-centred gamma distribution arises as a Poisson mixture of gamma distributions. The density of the non-centred gamma distribution is as follows:

$$f_{\delta,\beta'(RV_t,X_t),c} = exp\left[-\frac{RV_t}{c}\right]\sum_{k=0}^{\infty}\left[\frac{RV^{\delta+k-1}}{c^{\delta+k}\Gamma(\delta+k)}\frac{\beta'(RV_t,X_t)^k}{k!}\right] \tag{4.25}$$

In line with the PDF of the non-centred gamma in 4.25, the log-likelihood function is:

$$l_t(\theta) = \frac{1}{c}(RV_{t+1} + c\beta'(RV_t, L_t)) + log\left(\sum_{k=0}^{\infty}\frac{RV_{t+1}^{\delta+k-1}}{c^{\delta+k}\Gamma(\delta+k)}\frac{\beta'(RV_t, X_t)^k}{k!}\right) \quad (4.26)$$

where the parameter set $\theta = \{\delta, \beta', c\}$. The log-likelihood function cannot be applied directly since it contains infinite number of terms. Thus, we need to truncate the infinite sum in Eq. 4.25. We follow Corsi et al. (2013) and set the truncation number equal to 90.

Recall from Eq. (4.7) that the log-return is expressed as:

$$\ln\left[\frac{S_{t+1}}{S_t}\right] = r + \gamma RV_{t+1} + \sqrt{RV_{t+1}}\varepsilon_{t+1}$$

To estimate the market price of risk γ, we can rewrite this as

$$\frac{\ln\left[S_{t+1}/S_t\right] - r}{\sqrt{RV_{t+1}}} = \gamma\sqrt{RV_{t+1}} + \varepsilon_{t+1} \quad (4.27)$$

The parameter γ can be estimated through robust ordinary least square (OLS). The log-likelihood function can be written as:

$$l_t(\theta) = -\frac{1}{2}logV(RV_{t+1}/RV_t) - \frac{1}{2}\frac{RV_{t+1} - E((RV_{t+1}/RV_t))^2}{V(RV_{t+1}/RV_t)} \quad (4.28)$$

where

$$E(RV_{t+1}/RV_t) = c\delta + c\beta'(RV_t, X_t)$$

$$V(RV_{t+1}/RV_t) = c^2\delta + 2c^2\beta'(RV_t, X_t)$$

4.3 Data Description

In this section, we will explain the European option data that is written on S&P500 Index in details. Also, we will introduce a set of fundamental macroeconomic variables that are added into the HAR-MIDAS option pricing model.

4.3.1 European Option data

We consider European put and call options that are written on S&P 500 Index. The Option data is collected from the OptionMetrics database for the period from January 1, 1996 to December 31, 2014. Following Christoffersen et al. (2008), we only include options that are traded on Wednesday, which seems to be more liquid than other trading days. In addition, options with trading volume less than 20 are discarded. For maturity, we delete options with time to maturity less than 20 days or more than 160 days. To filter out more volatile options, we only use options with implied volatility less than 60%. For option price, we filter out observations with prices less than 5 dollars. Tables 4.1 and 4.2 present descriptive statistics for S&P500 Put and Call Options. Options are categorized into different groups according to the moneyness (K/S) and time to maturity (τ).

Tables 4.1 and 4.2 present descriptive statistics for S&P500 Put and Call Options, respectively. Options are categorized into different groups according to the moneyness K/S and term to marturity time τ. For the call option, out-of-the-money options (hereafter OTM) are defined as $1.02 < K/S \leq 1.04$, and deep OTM (hereafter DOTM) options are defined as $K/S > 1.04$. In constrast, if $0.94 < K/S \leq 0.96$, the put option is OTM. If $K/S \leq 0.94$, the put option is DOTM. In relation to time to maturity τ, both put and call options are categorized into three groups: short-maturity options with $\tau \leq 20$, medium maturity options with $20 < \tau \leq 60$ and long-maturity options with $60 < \tau \leq 160$. Looking at the implied volatility across different degrees of moneyness (from Deep in-the-money through to Deep out-of-money) for both put and call options, implied volatility exhibits a smirk curve.

[Insert Tables 4.1 and 4.2 here]

As noted in Chapter 3.3.1, the Realized volatility is calculated from intra-daily returns sampled with tick-by-tick 5-mins frequency, in a purpose to balance the trade-off between between microstructure noise and finely sampling frequency (see Liu et al. (2015) for details). Consistent with Chapter 3.3.1, we collect realized variance data from the Oxford-Man Institute Realized Library, where daily realized variance is reported to be estimated via Realized Kernel method (Barndorff-Nielsen et al. (2008)) with 5-mins frequency. All RV data covers a period from Jan 1st 1996 to Dec 31th 2014.

4.3.2 Fundamental Macroeconomic Variables

As our main focus is to investigate the impact of macroeconomic information on volatility, we initially employ a wide range of macroeconomic variables consistent with that being adopted in Chapter 2: Industrial production (IP), Producer Price Index (PPI), Unemployment rate (UEM), Housing Starts (HS) and the Term Spread (TS). We select monthly seasonally-adjusted data that ranges from Jan 1996 to Dec 2014. The data is obtained from the Federal Reserve Economic Database (FRED). All of the macroeconomic variables are converted into annualized growth rates:

$$\Delta X_t = \left[\frac{X_t}{X_{t-1}}\right]^{12} - 1 \tag{4.29}$$

According to Gourieroux and Jasiak (2006) and Corsi et al. (2013), daily volatility in a HAR-type process follows non-central gamma distribution. For consistency, we add the variance of the macroeconomic variables into the HAR-MIDAS option model. To get our estimate of the macroeconomic variance, we let the macroeconomic growth rate follow a linear regression with constant mean and dynamic variance (see in Eq.4.12):

$$x_{t+1} = \mu_x + \sigma_{x,t+1} u_{t+1}$$

We extract σ_x and put it into the HAR-MIDAS option model for pricing measurement.

4.4 An Application Using Macroeconomic Uncertainty Measure

In addition to the macroeconomic fundamental variables discussed in the previous subsection, we consider the influence of macroeconomic uncertainty on stock index options. Macroeconomic uncertainty here is the conditional volatility of a disturbance that is unexpected from the perspective of economic agents. Haddow et al. (2013) and Bloom et al. (2018) mention that high levels of uncertainty might have adverse effects on financial markets, by affecting decision-making in all parts of financial markets as well as that of policymakers. In reality, macroeconomic uncertainty is difficult to quantify due to its latent feature. One attempt to overcome this difficulty is Jurado et al. (2015) and we employ their measure of macroeconomic uncertainty here. Jurado et al. (2015) construct a proxy for macroeconomic uncertainty by aggregating conditional volatility of unexpected components realized via multi-step ahead forecast errors from a wide range of macroeconomic variables. Precisely speaking, macroeconomic uncertainty is constructed by 3 steps. First, a large set of macroeconomic variables are compressed into common factors, denoted by y_j. With a limited number of common factors, the unexpected errors are:

$$V^y_{jt+h} = y_{jt+h} - E[y_{jt+h}|I_t] \tag{4.30}$$

Second, define the conditional volatility of unexpected error as a proxy of individual uncertainty, $\mathcal{U}_{jt}(h)^y$:

$$\mathcal{U}_{jt}(h)^y = \sqrt{E[(V^y_{jt+h})^2|I_t]} \tag{4.31}$$

Third, a macroeconomic uncertainty index is calculated as a weighted average of individual uncertainty measures:

$$\mathcal{U}_{jt}(h) = \sum_{j=1}^{N_y} w_j \mathcal{U}^y_{jt}(h) \tag{4.32}$$

where w_j is the weight attached to each individual uncertainty and N_y is the number of individual uncertainties. Following Equations 4.31 and 4.32, we estimate a macroeconomic uncertainty index from 132 macroeconomic variables, which are consistent with that being used in Jurado et al. (2015). In line with Chapter 2, we employ macroeconomic uncertainty estimated for 3-month horizon, $\mathcal{U}_t(3)$. The macroeconomic uncertainty index is shown in Figure 4.1, which is depicted against the S&P 500 Index returns during 1996-2014. The shading areas highlights economic recessions in U.S market, which is in line with the dates published by the National Bureau of Economic Research (NBER). We can see that large spikes in macroeconomic uncertainty are coincide with financial crisises occurred around 2001-2002, 2007-2008.

[Insert Figure 4.1 here]

4.5 Empirical Comparison with Benchmark models

In this section, we present parameter estimates for the HAR-MIDAS model specified with different macroeconomic variables and macroeconomic uncertainty. To evaluate HAR-MIDAS's performance, we conduct empirical comparisons among the alternative HAR-MIDAS models (i.e. the HAR-MIDAS models using different macroeconomic variables) and two benchmarks, Duan's GARCH and Corsi's HARG models, examining whether macroeconomic information helps improve pricing accuracy.

Following Corsi et al. (2013), we utilize the Monte Carlo simulation method in the HAR-MIDAS volatility model for option pricing. The whole procedure consists of three steps: First, generate maximum likelihood parameter estimates for the HAR-MIDAS model under a physical measure (see in Eq. 4.15). Second, map the parameters in the HAR-MIDAS model from the physical measure P into the risk-neutral measure Q. Third, generate Monte Carlo simulations of both the log-return and conditional volatility, so as to obtain simulated asset prices at maturity under the Q measure. With the underlying spot price S_t, we simulate L paths of asset prices $\{S_t^T\}$ (from t to maturity T) and realized volatilities RV_t^T for each strike price K. With these simulated values, we can simulate the value of call option at maturity: $max\{S_T - K, 0\}$, and the value of put option at maturity: $max\{K - S_T, 0\}$.

Table 4.3 summarizes the parameter estimates for the HAR-MIDAS alternative models as well as two benchmarks, Duan's GARCH model and Corsi's HAR model. For the HAR-MIDAS model, parameter β_4 accounts for the effect of the macroeconomic variable, and ω accounts for the decaying impact the macroeconomic variable has. Excluding inflation (PPI), all the macroeconomic variables have a significant impact on stock volatility, including macroeconomic uncertainty. Comparing the log-likelihood ratio among the different models, the HAR-MIDAS model with macroeconomic information provides a better description of stock volatility relative to the two benchmarks. This general improvement from using the HAR-MIDAS model is largely due to the significant contribution of the macroeconomic variables to explaining volatility, consistent with Engle et al. (2013)'s argument. Given the parameter estimates, for the second step we follow Eq. 4.24 to convert the estimated parameters from the physical measure (P) into the risk-neutral measure (Q). We then generate Monte carlo simulations for both log-return and conditional volatility, and calculate the option prices for both put and call at maturity.

[Insert Table 4.3 here]

We implement the root mean squared error (RMSE) for both price and implied volatility to evaluate the option pricing performances of each HAR-MIDAS model specified with different macroeconomic variables. The $RMSE_p$ is defined as:

$$RMSE_p = \sqrt{\sum_{i=1}^{N} \frac{\tilde{p}_i^{mkt} - \tilde{p}_i^{model}}{N}} \tag{4.33}$$

where N denotes the number of options. For each option i, p_i^{mkt} refers to its market price and p_i^{model} refers to its model simulated price. Following Corsi et al. (2013), we then express the option price relative to underlying asset price S, $\tilde{p}_i = p_i/S$. For implied volatility, the $RMSE_{IV}$ is

$$RMSE_{IV} = \sqrt{\sum_{i=1}^{N} \frac{\widetilde{IV}_i^{mkt} - \widetilde{IV}_i^{model}}{N}} \tag{4.34}$$

where IV^{mkt} refers to the market implied volatility and IV^{model} refers to the model implied volatility. Note that the market implied volatility is the Black-Scholes implied volatility.

Tables 4.4 through 4.12 summarize option performances for both put and call. Initially, we need to look at performance of the two benchmarks in terms of $RMSE_p$ and $RMSE_{IV}$. Panel A of Table 4.4 shows that the HAR option model outperforms Duan's GARCH option model for both short- and medium-maturity call option pricing, which is largely due to the direct use of RV as a proxy for volatility. However, for long-term maturity, the two benchmarks deliver similar results in terms of performance. Similar results can be found in Table 4.9 for put options. The results suggest that the HAR model has limited ability when pricing long-term options, even with the long memory feature of volatility that the HAR model captures. In addition, both Table 4.4 and Table 4.9 demonstrate that the HAR model is superior to Duan's GARCH in terms of the implied volatility root mean squared error, $RMSE_{IV}$ (Panel B), especially when term-to-maturity is less than 160 days. A comparison between Duan's GARCH and the HAR models confirms the importance of utilizing RV for option pricing. As RV contains high-frequency trading information, the HAR model delivers better pricing for both short-term and medium term options. Now we add in macroeconomic information into the HAR model, to see whether it yields better performance for long-term put and call options.

Since we employ several macroeconomic variables, we calculate the relative performance ratio of $RMSE$ for the HAR-MIDAS model against Duan's GARCH and Corsi's HAR models. Tables 4.5 and 4.6 summarize call option pricing relative performance, $RMSE_p$. In table 4.5 we observe that the HAR-MIDAS model incorporating macroeconomic variables has, in general (with the exception of the term spread), superior performance relative to Duan's GARCH model across different degrees of moneyness. That the term spread does not perform particularly well might be explained by the fact that term-spread usually more associated with recession rather than expansion. In particular, there is a body of literature that finds that the term spread, as a leading indicator, predicts recessions well. Therefore, call options, which are more associated with market expectation of economic expansion, are less likely be affected by negative shocks to the term spread (and unemployment.) When we turn to look at the relative performance of $RMSE_p$ for the HAR-MIDAS model against Corsi's HAR model, we find that only macroeconomic uncertainty has better performance when the call option is Out-of-the-Money (OTM) or Deep-Out-of-the-Money (DOTM) (see Table 4.6).

Table 4.10 and 4.11 summarize put option pricing relative performance, $RMSE_p$. Most of the HAR-MIDAS models which include the macro variables provide superior performance to Duan's GARCH model but worse performance than the original HAR model. Inflation seems to have the least impact for put option pricing. Perhaps unsurprisingly, macroeconomic uncertainty outperforms alternative macroeconomic variables and delivers better pricing results for OTM put options.

In terms of implied volatility relative performance, the root mean squared errors ($RMSE_{IV}$) for both call and put are reported in Table 4.8 and 4.12, respectively. Estimation results strongly favor the original HAR model as well as the HAR-MIDAS model specified with macroeconomic uncertainty. Especially when the option has a long term to maturity, the implied volatility errors when using macroeconomic uncertainty are on average 15% less than in the HAR model. The results suggest that adding macroeconomic information, especially macroeconomic uncertainty, into the HAR model pays off in terms of delivering significant improvements for long-term option.

[Insert Tables 4.4 through 4.12 here]

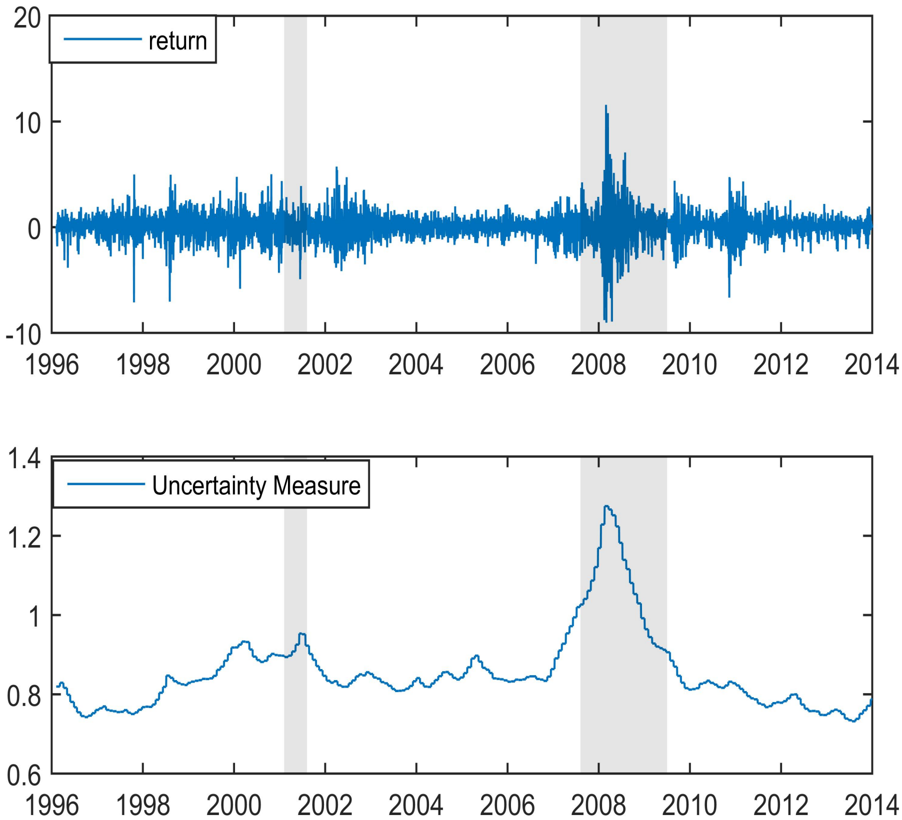
4.6 Conclusion

In this paper, we propose a discrete-time realized option valuation model, the HAR-MIDAS model. In the HAR-MIDAS model, the option price is determined by financial market information and macroeconomic information. We employ the Mixed data sampling (MIDAS) approach, connecting macroeconomic information with stock volatility directly.

The HAR-MIDAS model potentially possesses several advantages for volatility modelling and option pricing. First, latent volatility is directly measured by realized volatility (RV), which is constructed using high-frequency intradaily returns. Embedded within high frequency trading information, HAR-MIDAS model is able to response quickly with respect to market shocks. Second, with its multi-component structure, the HAR-MIDAS model can capture the long memory feature well. Therefore, it might be of benefit in pricing long-maturity options.

Empirically, we consider two sources of macroeconomic information: fundamental macroeconomic variables that reflect past information and a measure of macroeconomic uncertainty that is associated with future expectations. Our empirical results indicate that most macroeconomic variables outperform Duan's GARCH model across maturity and moneyness. In particular, macroeconomic uncertainty has superior performance for OTM options with long-maturity. Precisely speaking, the HARG-MIDAS model incorporated with macroeconomic uncertainty have smaller RMSE results for both Put and Call options when their term to maturities are longer than 60 days, comparing with benchmark models (Duan's GARCH and Corsi's HARG). It suggests that macroeconomic uncertainty plays an important role in mitigating OTM option pricing biases.

Figure 4.1: S&P 500 log return and macroeconomic uncertainty In The U.S Market (1996-2014)



The Figure 4.1 plots daily log return of S&P500 Index, ranging from 1996 to 2014 (shown in upper plot). and estimates of macro uncertainty $U_t(3)$ for 3 month horizon h=3 (shown in lower plot). The macroeconomic uncertainty is avaiable in a monthly frequency from Jan 1996 to Dec 2014. To consistent with daily returns, we extend month macro uncertainty into a daily frequency. The shading areas corresponds to economic recessions in the U.S market.

Table 4.1: S&P500 Call Option Data, 1996–2014

	Maturiy τ		
Moneyness	$\boldsymbol{\tau \leq 20}$	$\boldsymbol{20 < \tau \leq 60}$	$\boldsymbol{60 < \tau \leq 160}$
K/S	**Panel A: Implied Volatility of Call Options**		
K/S≤ 0.94	0.3221	0.2879	0.2762
0.94<K/S≤0.96	0.2101	0.2114	0.2028
0.96<K/S≤0.98	0.1811	0.1885	0.1907
0.98<K/S≤1.02	0.1509	0.1590	0.1742
1.02<K/S≤1.04	0.2090	0.1574	0.1558
1.04<K/S	0.3270	0.2369	0.1799
	Panel B: Implied Volatility Standard Deviation of Call Options		
K/S≤ 0.94	0.1138	0.1046	0.1033
0.94<K/S≤0.96	0.0658	0.0692	0.0554
0.96<K/S≤0.98	0.0609	0.0650	0.0574
0.98<K/S≤1.02	0.0576	0.0633	0.0575
1.02<K/S≤1.04	0.0672	0.0660	0.0572
1.04<K/S	0.0981	0.1026	0.0683
	Panel C: Number of Call Options		
K/S≤ 0.94	1295	5215	3617
0.94<K/S≤0.96	1026	2823	1265
0.96<K/S≤0.98	1887	4333	2005
0.98<K/S≤1.02	5823	15678	8384
1.02<K/S≤1.04	596	5482	4028
1.04<K/S	154	3938	8790
All	10781	37469	28089

[1] In Table 4.1, we provide descriptive statistics for the European Call Option written on the S&P500 Index. Consider about liquidity, we adopt call option prices quoted on each Wednesday available from January 1, 1996 to December 31, 2014. All call options are categorized into different groups according to the moneyness K/S and the term to maturity τ, where moneyness is calculated as the underlying index level divided by strike price. Panels A and B summarize mean values and standard deviations of Implied Volatility for the S&P500 Call Option across different categories. Panel C provides number of observations for the S&P500 Call Option.

Table 4.2: S&P500 Index Put Option Data, 1996–2014

	Maturiy τ		
Moneyness	$\boldsymbol{\tau \leq 20}$	$\boldsymbol{20 < \tau \leq 60}$	$\boldsymbol{60 < \tau \leq 160}$
K/S	**Panel A: Implied Volatility of Put Options**		
1.04<K/S	0.2590	0.2341	0.2525
1.02<K/S≤1.04	0.1576	0.1635	0.1915
0.98<K/S≤1.02	0.1552	0.1651	0.1808
0.96<K/S≤0.98	0.2004	0.1844	0.1808
0.94<K/S≤0.96	0.2742	0.2082	0.1815
K/S≤ 0.94	0.3836	0.2863	0.2255
	Panel B: Implied Volatility Standard Deviation of Put Options		
K/S≤ 0.94	0.1138	0.1046	0.1033
0.94<K/S≤0.96	0.0658	0.0692	0.0554
0.96<K/S≤0.98	0.0609	0.0650	0.0574
0.98<K/S≤1.02	0.0576	0.0633	0.0575
1.02<K/S≤1.04	0.0672	0.0660	0.0572
1.04<K/S	0.0981	0.1026	0.0683
	Panel C: Number of Put Options		
K/S≤ 0.94	1295	5215	3617
0.94<K/S≤0.96	1026	2823	1265
0.96<K/S≤0.98	1887	4333	2005
0.98<K/S≤1.02	5823	15678	8384
1.02<K/S≤1.04	596	5482	4028
1.04<K/S	154	3938	8790
All	10781	37469	28089

In Table 4.2, we provide descriptive statistics for the European Put Option written on the S&P500 Index. Consider about liquidity, we adopt put option prices quoted on each Wednesday available from January 1, 1996 to December 31, 2014. All put options are categorized into different groups according to the moneyness K/S and the term to maturity τ, where moneyness is calculated as the underlying index level divided by strike price. Panels A and B summarize mean values and standard deviations of Implied Volatility for the S&P500 Put Option across different categories. Panel C provides number of observations for the S&P500 Put Option.

Table 4.3: Parameter Estimation on Options, 1996–2014

Parameters	HARG-MIDAS					
	IP	PPI	UEM	HS	TS	$\mathcal{U}$
$\boldsymbol{\beta_1}$	0.0349***	0.0343***	0.0348***	0.0356***	0.0332***	0.0335***
	(12.3581)	(10.2172)	(8.9001)	(12.1425)	(11.3272)	(11.4853)
$\boldsymbol{\beta_2}$	0.0425***	0.0402***	0.0417***	0.0437***	0.0391***	0.0396***
	(11.4969)	(3.1310)	(8.7220)	(11.8993)	(11.5985)	(6.9749)
$\boldsymbol{\beta_3}$	0.0231***	0.0149**	0.0205***	0.0251***	0.0117***	0.0184***
	(6.2958)	(5.1459)	(5.3391)	(6.5908)	(5.0992)	(5.1113)
$\boldsymbol{\beta_4}$	0.8593***	0.6646***	0.7509***	0.7065***	0.6321***	0.7566***
	(38.1678)	(47.5850)	(43.1698)	(34.1472)	(27.4855)	(24.6879)
$\boldsymbol{\omega}$	9.8363***	1.0100	1.0100**	1.7601***	1.5534***	14.7688***
	(36.2673)	(0.3086)	(2.3024)	(10.9075)	(14.6394)	(5.7956)
c	15.5051***	15.5181***	15.5096***	15.5142***	15.5179***	15.5124***
	(56.7290)	(62.7761)	(54.4542)	(60.1484)	(56.3498)	(37.0025)
$\boldsymbol{\delta}$	1.3600***	1.1400***	1.2300***	1.0300***	1.0100***	1.0100***
	(14.9855)	(14.5407)	(15.3981)	(14.7012)	(14.9363)	(19.6156)
Log-likelihood	−12059.0877	−12059.2821	−12059.1273	−12059.2075	−12059.2064	−12059.1106

GARCH		HARG	
α_0	2.05E-06***	β_1	0.0348***
	(5.0587)		(12.1485)
α_1	0.0604***	β_2	0.0423***
	(7.3494)		(11.3767)
β_1	0.8391***	β_3	0.0230***
	(45.7691)		(6.2722)
λ	1.2389***	c	15.5047***
	(9.5852)		(55.3772)
		δ	1.0100***
			(15.1997)
Log-likelihood	−9588.5678		−11086.3421

In Table 4.3, we generate maximum likelihood estimation for the HAR-MIDAS model specified with alternative macroeconomic fundamental variables (or macroeconomic uncertainty) The estimation period ranges from January 1, 1996 to December 31, 2014. In addition, for further comparisions of option pricing performance, we also estimate parameters for two benchmark models, Duan'GARCH model and Corsi's HAR model.

Table 4.4: Call Option Performance for the Benchmark models, 1996–2014

	Maturiy $10 < \tau < 20$		*Maturiy* $20 < \tau < 60$		*Maturiy* $60 < \tau < 160$	
K/S	**Duan-GARCH**	**HARG**	**Duan-GARCH**	**HARG**	**Duan-GARCH**	**HARG**
			Panel A: Option Pricing Performance			
K/S≤0.94	0.0027	0.0028	0.0065	0.0056	0.0110	0.0132
0.94<K/S≤0.96	0.0055	0.0044	0.0099	0.0077	0.0152	0.0147
0.96<K/S≤0.98	0.0075	0.0046	0.0114	0.0083	0.0164	0.0157
0.98<K/S≤1.02	0.0100	0.0047	0.0134	0.0085	0.0171	0.0160
1.02<K/S≤1.04	0.0051	0.0052	0.0119	0.0088	0.0167	0.0163
1.04<K/S	0.0074	0.0054	0.0096	0.0085	0.0125	0.0131
			Panel B: Implied Volatility Performance			
0.94<K/S≤0.96	9.96	10.97	12.69	12.37	9.77	8.29
0.96<K/S≤0.98	14.81	7.35	11.49	9.55	9.50	8.28
0.98<K/S≤1.02	12.76	5.79	11.29	6.21	8.86	7.99
1.02<K/S≤1.04	7.16	6.46	10.86	4.55	8.93	9.00
1.04<K/S	10.31	8.38	8.72	6.21	7.15	8.21

[a] In Table 4.4, we provide root mean squared errors (RMSE) of the Call Option pricing performance for alternative GARCH-type option models (Heston Nandi's GARCH, Duan's GARCH) as well as the HARG model. All $RMSE_p$ are sorted by moneyness and term to maturity.

Table 4.5: Call Option Pricing Relative Performance for the HAR-MIDAS model, 1996–2014

HARG-MIDAS/Duan's GARCH	*Maturiy* $\mathbf{10 < \tau < 20}$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
K/S≤ 0.94	1.1111	1.3333	1.4074	2.7407	1.0741	2.6296
0.94<K/S≤0.96	0.7818	0.7818	0.8909	1.1091	0.8182	0.9091
0.96<K/S≤0.98	0.6533	0.7200	1.4933	1.1467	0.8133	0.6667
0.98<K/S≤1.02	0.5100	0.6500	1.4100	0.6300	0.4900	0.4900
1.02<K/S≤1.04	1.2353	0.8627	0.8627	0.9412	1.0392	0.8824
1.04<K/S	1.2027	1.4054	0.9054	0.8514	1.2027	0.5811
	Maturiy $\mathbf{20 < \tau < 60}$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
K/S≤ 0.94	0.9077	1.2769	1.5231	1.0615	0.9538	1.2615
0.94<K/S≤0.96	0.7879	0.8182	1.2323	1.6263	0.7778	0.9091
0.96<K/S≤0.98	0.7193	1.0789	1.5964	0.8421	0.7193	0.7632
0.98<K/S≤1.02	0.6119	0.9179	1.3582	0.7164	0.6119	0.6493
1.02<K/S≤1.04	0.7227	0.7731	1.3866	0.8403	0.7311	0.6891
1.04<K/S	1.1771	1.4688	1.2708	1.0104	1.1250	0.8438
	Maturiy $\mathbf{60 < \tau < 160}$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
K/S≤ 0.94	1.1455	0.9909	2.2545	1.5182	1.5909	0.9909
0.94<K/S≤0.96	0.8487	0.9145	0.9605	1.0592	0.9605	0.8947
0.96<K/S≤0.98	0.8232	0.9817	1.8476	1.0366	0.9634	1.0122
0.98<K/S≤1.02	1.2923	0.7778	1.8128	1.0234	0.7310	0.7539
1.02<K/S≤1.04	0.7725	0.9281	1.9042	1.0659	0.9281	0.9341
1.04<K/S	0.8560	1.0560	2.1440	1.0880	0.8720	0.9440

[a] In Table 4.5, we provide root mean squared errors (RMSE) of Call Option for the HAR-MIDAS models relative to the benchmark (Duan's GARCH) sorted by moneyness and term to maturity. Maturity is shown in days and monenyness is K/S, where K denotes call option strike price and S denotes underlying market price. Call Option is written on S&P 500 Index from January 1, 1996 to December 31, 2014.

Table 4.6: Call Option Pricing Relative Performance for the HAR-MIDAS model, 1996–2014

HARG-MIDAS/HARG	*Maturiy* $10 < \tau < 20$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
K/S≤ 0.94	1.0714	1.2857	1.3571	2.6429	1.0357	2.5357
0.94<K/S≤0.96	0.9773	0.9773	1.1136	1.3864	1.0227	1.1364
0.96<K/S≤0.98	1.0652	1.1739	2.4348	3.9783	1.3261	1.0870
0.98<K/S≤1.02	1.0851	1.3830	3.0000	1.3404	1.0426	1.0426
1.02<K/S≤1.04	1.2115	0.8462	0.8462	0.9231	1.0192	**0.8654**
1.04<K/S	1.6481	1.9259	1.2407	1.1667	1.6481	**0.7963**
	Maturiy $20 < \tau < 60$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
K/S≤ 0.94	1.0536	1.4821	1.7679	1.2321	1.1071	1.4643
0.94<K/S≤0.96	1.0130	1.0519	1.5844	2.0909	1.0000	1.1688
0.96<K/S≤0.98	1.0241	1.0361	2.4819	1.1687	0.9879	1.0000
0.98<K/S≤1.02	0.9647	1.4471	2.1412	1.1294	0.9647	1.0235
1.02<K/S≤1.04	0.9773	1.0455	1.8750	1.1364	0.9886	**0.9318**
1.04<K/S	1.3294	1.6588	1.4353	1.1412	1.2706	**0.9529**
	Maturiy $60 < \tau < 160$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
K/S≤ 0.94	0.9545	0.8257	1.8787	1.2651	1.3257	0.8257
0.94<K/S≤0.96	0.8776	0.9456	0.9932	1.0952	0.9932	0.9252
0.96<K/S≤0.98	0.8600	1.0255	1.9299	1.0828	1.0064	1.0573
0.98<K/S≤1.02	1.3813	0.8312	1.9375	1.0938	0.7813	0.8063
1.02<K/S≤1.04	0.7914	0.9509	1.9509	1.0920	0.9509	0.9571
1.04<K/S	0.8168	1.0076	2.0458	1.0382	0.8321	0.9008

[a] In Table 4.6,we provide root mean squared errors (RMSE) of Call Option for the HAR-MIDAS models relative to the benchmark (Corsi's HARG) sorted by moneyness and term to maturity. Maturity is shown in days and monenyness is K/S, where K denotes call option strike price and S denotes underlying market price. Call Option is written on S&P 500 Index from January 1, 1996 to December 31, 2014.

Table 4.7: Call Option Implied Volatility Relative Performance for the HARG-MIDAS model, 1996–2014

HARG-MIDAS/GARCH	*Maturiy* $10 < \tau < 20$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
0.94<K/S≤0.96	0.9177	0.8504	1.5502	2.1215	1.1376	1.6044
0.96<K/S≤0.98	0.6962	0.5300	1.3599	1.0635	0.7373	0.5557
0.98<K/S≤1.02	0.4960	0.6110	1.4110	0.6280	0.4640	0.4750
1.02<K/S≤1.04	1.1313	0.7905	0.8408	0.8478	0.9274	0.8450
1.04<K/S	1.2231	1.5936	0.9117	0.8535	1.2182	0.6305
	Maturiy $20 < \tau < 60$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
0.94<K/S≤0.96	0.6060	0.6028	1.1710	1.5658	0.6383	0.9748
0.96<K/S≤0.98	0.6335	0.6423	1.7163	0.7955	0.6301	0.6867
0.98<K/S≤1.02	0.5740	0.8530	1.2923	0.6829	0.5784	0.6395
1.02<K/S≤1.04	0.6464	0.6842	1.3619	0.8057	0.6860	0.6188
1.04<K/S	1.0000	1.6044	1.3383	1.0252	1.1066	0.8073
	Maturiy $60 < \tau < 160$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
0.94<K/S≤0.96	0.7390	0.7462	0.8557	0.9396	0.8557	0.7308
0.96<K/S≤0.98	0.7495	0.8874	1.6832	0.9368	0.8811	0.9074
0.98<K/S≤1.02	1.2607	0.7607	1.7370	0.9865	0.7133	0.7427
1.02<K/S≤1.04	0.7872	0.9619	1.9642	1.1007	0.9574	0.8723
1.04<K/S	0.8923	1.0448	2.4545	1.2014	0.8951	0.9413

[a] In Table 4.7,we provide implied volatility root mean squared errors ($RMSE_{IV}$) of Call Option for the HAR-MIDAS models relative to the benchmark (Duan's GARCH) sorted by moneyness and term to maturity. Maturity is shown in days and monenyness is K/S, where K denotes call option strike price and S denotes underlying market price. Call Option is written on S&P 500 Index from January 1, 1996 to December 31, 2014.

Table 4.8: Call Option Implied Volatility Relative Performance for the HARG-MIDAS model, 1996–2014

HARG-MIDAS/HARG	*Maturiy* $10 < \tau < 20$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
0.94<K/S≤0.96	0.8332	0.7721	1.4075	1.9262	1.0328	1.4567
0.96<K/S≤0.98	1.4027	1.0680	2.7401	2.1429	1.4857	1.1197
0.98<K/S≤1.02	1.0933	1.3472	3.1088	1.3834	1.0225	1.0466
1.02<K/S≤1.04	1.2538	0.8761	0.9319	0.9396	1.0279	0.9365
1.04<K/S	1.5048	1.9606	1.1217	1.0501	1.4988	0.7757
	Maturiy $20 < \tau < 60$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
0.94<K/S≤0.96	0.6217	0.6184	1.2013	1.6063	0.6548	1.0000
0.96<K/S≤0.98	0.7623	0.7728	2.0649	0.9571	0.7581	0.8262
0.98<K/S≤1.02	1.0435	1.5507	2.3494	1.2415	1.0515	1.1626
1.02<K/S≤1.04	1.5429	1.6330	3.2505	1.9231	1.6374	1.4769
1.04<K/S	1.4042	2.2528	1.8792	1.4396	1.5539	1.1337
	Maturiy $60 < \tau < 160$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
0.94<K/S≤0.96	0.8709	0.8794	1.0084	1.1074	1.0084	0.8613
0.96<K/S≤0.98	0.8599	1.0181	1.9312	1.0749	1.0109	1.0411
0.98<K/S≤1.02	1.3980	0.8436	1.9262	1.0939	0.7910	0.8235
1.02<K/S≤1.04	0.7811	0.9544	1.9489	1.0922	0.9500	0.8656
1.04<K/S	0.7771	0.9099	2.1376	1.0463	0.7795	0.8197

[a] In Table 4.8,we provide implied volatility root mean squared errors ($RMSE_{IV}$) of Call Option for the HAR-MIDAS models relative to the benchmark (Corsi's HARG) sorted by moneyness and term to maturity. Maturity is shown in days and moneyness is K/S, where K denotes call option strike price and S denotes underlying market price. Call Option is written on S&P 500 Index from January 1, 1996 to December 31, 2014.

Table 4.9: Put Option Performance for the Benchmark models, 1996–2014

	Maturiy $10 < \tau < 20$		*Maturiy* $20 < \tau < 60$		*Maturiy* $60 < \tau < 160$	
K/S	**Duan-GARCH**	**HARG**	**Duan-GARCH**	**HARG**	**Duan-GARCH**	**HARG**
			Panel A: Put Pricing Performance			
1.04<K/S	0.0039	0.0062	0.0104	0.0119	0.0114	0.0123
1.02<K/S≤1.04	0.0064	0.0065	0.0119	0.0114	0.0159	0.0124
0.98<K/S≤1.02	0.0093	0.0049	0.0131	0.0092	0.0168	0.0133
0.96<K/S≤0.98	0.0067	0.0051	0.0114	0.0083	0.0172	0.0172
0.94<K/S≤0.96	0.0040	0.0063	0.0096	0.0081	0.0185	0.0195
K/S≤ 0.94	0.0052	0.0078	0.0071	0.0072	0.0141	0.0189
			Panel B: Put Implied Volatility Performance			
1.02<K/S≤1.04	13.72	8.69	12.68	10.13	8.70	6.42
0.98<K/S≤1.02	12.09	5.90	10.83	7.33	8.54	6.72
0.96<K/S≤0.98	11.02	6.98	11.30	7.12	9.06	8.60
0.94<K/S≤0.96	6.65	9.81	10.88	8.03	11.10	10.21
K/S≤ 0.94	8.77	13.58	8.58	8.55	8.56	8.29

[a] Table 4.9 summarizes root mean squred errors $RMSE_p$ for the Put Option, which are sorted by moneyness K/S and maturity τ. We use the maximum-likelihood parameter estimates from Table 3.3 to compute the root mean squared error ($RMSE_p$) for alternative discrete volatility models, including Heston-Nandi's GARCH option model (HN-GARCH), Duan's GARCH option model (Duan GARCH) and Corsi's HARG model.

Table 4.10: Put Option Pricing Relative Performance for the HAR-MIDAS model, 1996–2014

HARG-MIDAS/Duan's GARCH	*Maturiy* $10 < \tau < 20$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.04<K/S	1.6154	2.3333	1.3846	2.1026	1.6154	1.3077
1.02<K/S≤1.04	0.9219	1.1250	0.7813	0.9844	1.0625	0.7656
0.98<K/S≤1.02	0.5699	0.5269	0.5806	0.5054	0.6882	0.6129
0.96<K/S≤0.98	0.7761	0.7761	0.6269	0.7164	0.9104	0.5970
0.94<K/S≤0.96	1.3750	1.3000	1.2500	1.0500	1.4750	1.0250
K/S≤ 0.94	0.8846	1.1538	1.3269	0.8269	0.8462	0.8462
	Maturiy $\mathbf{20 < \tau < 60}$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.04<K/S	1.5481	1.2692	1.1731	1.3077	1.4904	1.3750
1.02<K/S≤1.04	0.9748	1.2269	0.8739	0.9076	1.1008	1.0168
0.98<K/S≤1.02	0.6565	1.3740	0.7786	0.6718	0.7863	0.8015
0.96<K/S≤0.98	0.6404	0.6842	1.0789	0.6140	0.8246	0.7368
0.94<K/S≤0.96	0.7604	1.1146	1.2604	0.6771	0.7917	0.7083
K/S≤ 0.94	1.2113	1.0704	1.0000	0.9718	1.3521	0.9155
	Maturiy $\mathbf{60 < \tau < 160}$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.04<K/S	0.9386	1.1754	0.8421	0.9474	1.0438	1.1667
1.02<K/S≤1.04	0.8113	1.1384	0.7547	1.6101	0.8239	0.7925
0.98<K/S≤1.02	1.3155	1.4702	1.2500	0.8929	0.8333	0.8155
0.96<K/S≤0.98	0.7849	1.2151	0.9070	0.9070	0.9593	1.0349
0.94<K/S≤0.96	0.6973	1.2919	0.9622	1.1946	1.1189	0.9946
K/S≤ 0.94	0.8936	2.0426	1.5177	1.6809	2.0496	1.1631

[a] In Table 4.10, we provide root mean squared errors (RMSE) of Put Option for the HAR-MIDAS models relative to the benchmark (Duan's GARCH) sorted by moneyeyness and term to maturity. Maturity is shown in days and monenyness is K/S, where K denotes put option strike price and S denotes underlying market price. Put Option is written on S&P 500 Index from January 1, 1996 to December 31, 2014.

Table 4.11: Put Option Pricing Relative Performance for the HAR-MIDAS model, 1996–2014

HARG-MIDAS/HARG	*Maturiy* **10 < τ < 20**					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.04<K/S	0.5897	0.7692	0.8846	0.5513	0.5641	0.5641
1.02<K/S≤1.04	0.8730	0.8254	0.7937	0.6667	0.9365	0.6508
0.98<K/S≤1.02	1.0196	1.0196	0.8235	0.9412	1.1961	0.7843
0.96<K/S≤0.98	1.0816	1.0000	1.1020	0.9592	1.3061	1.1633
0.94<K/S≤0.96	0.9077	1.1077	0.7692	0.9692	1.0462	0.7538
K/S≤ 0.94	1.0161	1.4677	0.8709	1.3225	1.0161	0.8226
	Maturiy **20 < τ < 60**					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.04<K/S	1.1944	1.0556	0.9861	0.9583	1.3333	0.9027
1.02<K/S≤1.04	0.9012	1.3210	1.4938	0.8025	0.9383	0.8395
0.98<K/S≤1.02	0.8795	0.9398	1.4819	0.8434	1.1325	1.0120
0.96<K/S≤0.98	0.9348	1.9565	1.1087	0.9565	1.1196	1.1413
0.94<K/S≤0.96	1.0175	1.2807	0.9123	0.9474	1.1491	1.0614
K/S≤ 0.94	1.3529	1.1092	1.0252	1.1429	1.3025	1.2017
	Maturiy **60 < τ < 160**					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.04<K/S	0.8699	1.0894	0.7805	0.8780	0.9675	1.0813
1.02<K/S≤1.04	1.0403	1.4597	0.9677	2.0645	1.0565	1.0161
0.98<K/S≤1.02	1.6617	1.8571	1.5789	1.1278	1.0526	1.0301
0.96<K/S≤0.98	0.7849	1.2151	0.9070	0.9070	0.9593	1.0349
0.94<K/S≤0.96	0.6615	1.2256	0.9128	1.1333	1.0615	0.9435
K/S≤ 0.94	0.6667	1.5238	1.1323	1.2539	1.5291	0.8677

[a] In Table 4.11, we provide root mean squared errors (RMSE) of Put Option for the HAR-MIDAS models relative to the benchmark (Corsi's HARG) sorted by moneyeyness and term to maturity. Maturity is shown in days and monenyness is K/S, where K denotes put option strike price and S denotes underlying market price. Put Option is written on S&P 500 Index from January 1, 1996 to December 31, 2014.

Table 4.12: Put Option Implied Volatility Relative Performance for the HARG-MIDAS model, 1996–2014

HARG-MIDAS/HARG	*Maturiy* $10 < \tau < 20$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.02<K/S≤1.04	1.3924	1.9390	0.9217	0.9747	1.0518	0.9321
0.98<K/S≤1.02	1.1068	1.0305	1.1559	0.9729	1.3237	1.1542
0.96<K/S≤0.98	1.0072	1.1848	0.8954	1.0903	1.2149	0.8109
0.94<K/S≤0.96	1.0489	0.9776	0.7523	0.6126	0.9205	0.6045
K/S≤ 0.94	0.6112	0.7644	0.8837	0.5560	0.6038	0.5700
	Maturiy $20 < \tau < 60$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.02<K/S≤1.04	1.1165	1.1629	1.0267	1.0089	1.0760	1.2014
0.98<K/S≤1.02	0.9195	1.9127	1.1214	0.9441	1.0914	1.0764
0.96<K/S≤0.98	0.8455	0.9101	1.4677	0.8371	1.1067	0.9565
0.94<K/S≤0.96	0.8144	1.2341	1.7098	0.7522	0.8518	0.8406
K/S≤ 0.94	1.1123	1.0012	0.9357	0.9263	1.2421	0.8854
	Maturiy $60 < \tau < 160$					
K/S	**IP**	**PPI**	**UEM**	**TS**	**HS**	**U**
1.04<K/S	1.2002	1.0700	0.8046	0.8794	0.9578	1.0446
1.02<K/S≤1.04	1.8069	1.4346	1.0062	2.1916	1.0483	1.0109
0.98<K/S≤1.02	1.0074	1.8259	1.5997	1.1116	1.0432	1.0238
0.96<K/S≤0.98	1.1709	1.1965	0.9314	0.9442	0.9628	1.0826
0.94<K/S≤0.96	1.1861	1.1900	0.9598	1.1117	1.0402	0.9980

[a] In Table 4.12,we provide implied volatility root mean squared errors ($RMSE_{IV}$) of Put Option for the HAR-MIDAS models relative to the benchmark (Corsi's HARG) sorted by moneyness and term to maturity. Maturity is shown in days and moneyness is K/S, where K denotes put option strike price and S denotes underlying market price. Put Option is written on S&P 500 Index from January 1, 1996 to December 31, 2014.

References

Andersen, T. G. and T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The journal of Finance 52*(3), 975–1005.

Andersen, T. G. and T. Bollerslev (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold (2006). Volatility and correlation forecasting. *Handbook of economic forecasting 1*, 777–878.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica 76*(6), 1481–1536.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Variation, jumps and high frequency data in financial econometrics. *Advanced in Economics and Econometrics. Theory and Applications*.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of political economy 81*(3), 637–654.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018). Really uncertain business cycles. *Econometrica 86*(3), 1031–1065.

Campbell, J. Y. and R. J. Shiller (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance 43*(3), 661–676.

Christoffersen, P., R. Elkamhi, B. Feunou, and K. Jacobs (2009). Option valuation with conditional heteroskedasticity and nonnormality. *The Review of Financial Studies 23*(5), 2139–2183.

Christoffersen, P., K. Jacobs, C. Ornthanalai, and Y. Wang (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics 90*(3), 272–297.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, nbp001.

Corsi, F., N. Fusari, and D. La Vecchia (2013). Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics 107*(2), 284–304.

Cox, J. C., J. E. Ingersoll Jr, and S. A. Ross (2005). A theory of the term structure of interest rates. In *Theory of Valuation*, pp. 129–164. World Scientific.

Dorion, C. (2016). Option valuation with macro-finance variables. *Journal of Financial and Quantitative Analysis 51*(4), 1359–1389.

Duan, J.-C. (1995). The garch option pricing model. *Mathematical finance 5*(1), 13–32.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics 95*(3), 776–797.

Gagliardini, P., C. Gourieroux, and E. Renault (2011). Efficient derivative pricing by the extended method of moments. *Econometrica 79*(4), 1181–1232.

Ghysels, E., A. Sinko, and R. Valkanov (2007). Midas regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Gourieroux, C. and J. Jasiak (2006). Autoregressive gamma processes. *Journal of Forecasting 25*(2), 129–152.

Haddow, A., C. Hare, J. Hooley, and T. Shakir (2013). Macroeconomic uncertainty: what is it, how can we measure it and why does it matter? *Bank of England Quarterly Bulletin*, Q2.

Heston, S. L. and S. Nandi (2000). A closed-form garch option valuation model. *The review of financial studies 13*(3), 585–625.

Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *American Economic Review 105*(3), 1177–1216.

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

Müller, U. A., M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Pictet, and J. E. von Weizsäcker (1997). Volatilities of different time resolutions—analyzing the dynamics of market components. *Journal of Empirical Finance 4*(2), 213–239.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The journal of finance 44*(5), 1115–1153.

4.7 Appendix

Appendix A

A1 proof of Equation 4.4

Proof. $\Psi_t(u)$ is defined as the natural logarithm of MGF:

$$E_{t-1}\left[\exp(-u\varepsilon_t)\right] = \exp(\Psi_t(u)) = \exp(\frac{1}{2}h_t u^2)$$

where $\varepsilon_t \sim \mathcal{N}(0, h_t)$. Note that in the normal case we have $\Psi_t(u) = \frac{1}{2}h_t u^2$, and $\Psi_t(-1) = \frac{1}{2}h_t$. As noted by Eq.4.3, expected log-return satisfies:

$$E_{t-1}\left[ln(S_t/S_{t-1})\right] = E_{t-1}\left[r + \lambda\sqrt{h_t}\right]$$

which is equivalent to say:

$$\begin{aligned} E_{t-1}\left[S_t/S_{t-1}\right] &= E_{t-1}\left[\exp(r + \lambda\sqrt{h_t} - \gamma_t + \varepsilon_t)\right] \\ &= E_{t-1}\left[\exp(r + \lambda\sqrt{h_t} - \gamma_t)\right] + E_{t-1}\left[\exp(\varepsilon_t)\right] \\ &= E_{t-1}\left[\exp(r + \lambda\sqrt{h_t} - \gamma_t)\right] + \exp(\Psi_t(-1)) \\ &= E_{t-1}\left[\exp(r + \lambda\sqrt{h_t} - \gamma_t + \frac{1}{2}h_t)\right] \end{aligned}$$

So only when γ_t equals to $\frac{1}{2}h_t$, log-return will converge to its conditional mean of $u_t = r + \lambda\sqrt{h_t}$ □

A2 proof of Equation 4.6

Proof. The risk-neutralization of Duan(1995) is a special case of CEFJ(2010) in which the equivalent martingale measure (EMM) obtains by assuming that the Radon-Nikodym derivative is linear in stock return innovations. That is:

$$\xi_\tau = \frac{dQ}{dP} \mid F_\tau = exp\{-\sum_{t=1}^{\tau}(\eta_t\varepsilon_t + \Psi_t^{\varepsilon}(\eta_t))\} \tag{4.35}$$

Transferring the GARCH specifications from P measure to Q[2] measure, it must satisfies:

[2]Q refers to the risk-neutral measure.

$$E^Q_{t-1}[\frac{S_t}{S_{t-1}}/\frac{B_t}{B_{t-1}}] = E^P_{t-1}[\frac{\xi_t}{\xi_{t-1}} \times \frac{S_t}{S_{t-1}}/\frac{B_t}{B_{t-1}}] = 1 \tag{4.36}$$

where $\frac{\xi_t}{\xi_{t-1}}$ is equivalent to the pricing kernel. $\frac{B_t}{B_{t-1}} = e^r$[3] Rearrange equation, we have:

$$\begin{aligned}
&E^P_{t-1}[e^{-\eta_t v_t - \Psi^\varepsilon_t(\eta_t)} \times \frac{S_t}{S_{t-1}}/\frac{B_t}{B_{t-1}}] = 1\\
&E^P_{t-1}[e^{-\eta_t v_t - \Psi^\varepsilon_t(\eta_t)} e^{r+\lambda\sqrt{h_t}-\frac{1}{2}h_t+\sqrt{h_t}v_t}/e^r] = 1\\
&E^P_{t-1}[e^{-\eta_t v_t - \Psi^\varepsilon_t(\eta_t)} e^{\lambda\sqrt{h_t}-\frac{1}{2}h_t+\sqrt{h_t}v_t}] = 1\\
&E^P_{t-1}[e^{-(\eta_t-\sqrt{h_t})v_t}] e^{-\Psi^\varepsilon_t(\eta_t)} e^{\lambda\sqrt{h_t}-\frac{1}{2}h_t} = 1\\
&\exp\{\Psi_t(\eta_t - \sqrt{h_t}) - \Psi_t(\eta_t) + \lambda\sqrt{h_t} - \frac{1}{2}h_t\} = 1\\
&\Psi_t(\eta_t - \sqrt{h_t}) - \Psi_t(\eta_t) + \lambda\sqrt{h_t} - \frac{1}{2}h_t = 0\\
&\frac{1}{2}(\eta_t - \sqrt{h_t})^2 - \frac{1}{2}\eta_t^2 + \lambda\sqrt{h_t} - \frac{1}{2}h_t = 0\\
&(\lambda - \eta_t)\sqrt{h_t} = 0 \to \lambda = \eta_t
\end{aligned}$$

Replacing η_t by λ in equation 4.4, the Radon-Nikodym derivative turns into:

$$\frac{dQ}{dP} \mid F_t = exp\{(\lambda\varepsilon_t + \frac{1}{2}\lambda^2)\}$$

And the risk-neutral innovation v^*_t:

$$v^*_t = v_t + \lambda$$

Replacing v_t by $v^*_t - \lambda$ in equations 4.4 we now have:

$$\ln(\frac{S_t}{S_{t-1}}) = r + \lambda\sqrt{ht} - \frac{1}{2}h_t + \sqrt{h_t}(v^*_t - \lambda)$$

$$\ln(\frac{S_t}{S_{t-1}}) = r - \frac{1}{2}h_t + \sqrt{h_t}v^*_t \quad \varepsilon^*_t = \sqrt{h_t}v^*_t$$

Replacing v_t by $v^*_t - \lambda$ in equations 4.5 we now have:

$$h_t = \alpha_0 + \alpha_1(v^*_{t-1}\sqrt{h_{t-1}} - \lambda\sqrt{h_{t-1}})^2 + \beta_1 h_{t-1}$$

So now we have the GARCH option model of Duan (1995) under the Q

[3] B_t is the Bond price at time t.

measure:

$$\ln(\frac{S_t}{S_{t-1}}) = r - \frac{1}{2}h_t + \varepsilon_t^*$$
$$h_t = \alpha_0 + \alpha_1(\varepsilon_{t-1}^* - \lambda\sqrt{h_{t-1}})^2 + \beta_1 h_{t-1} \tag{4.37}$$

□

A3 proof of Proposition1

Proof. Let us compute the LT of $\Pi_{t+1}' = \beta'(RV_t), \beta'(X_t), y_{t+1})$ in the HAR-MIDAS option model:

$$\Pi_{t+1} = \begin{pmatrix} \beta'(RV_t) \\ \beta'(X_t) \\ y_{t+1} \end{pmatrix} \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

$$E_t^P\left[\exp\{-\alpha'\Pi_{t+1}\}\right]$$

$$\begin{aligned}
&= E_t^P\left[\exp\{-\alpha_1\beta'(RV_t) - \alpha_2\beta'(X_t) - \alpha_3 y_{t+1}\}\right] \\
&= E_t^P[\exp\{-\alpha_1\beta'(RV_t) - \alpha_2\beta'(X_t) - \alpha_3(\gamma RV_{t+1} + \sqrt{RV_{t+1}}\varepsilon_{t+1})\}] \\
&= E_t^P[\exp\{-\underbrace{(\alpha_1 + \gamma\alpha_3 - \frac{1}{2}\alpha_3^2)}_{\vartheta_1}\beta'(RV_t)\}\exp\{-\underbrace{(\alpha_2 + \gamma\alpha_3 - \frac{1}{2}\alpha_3^2)}_{\vartheta_2}\beta'(X_t)\}] \\
&= \phi_{RV}^P(\vartheta_1)\phi_X^P(\vartheta_2)
\end{aligned}$$

where

$$\phi_{RV}^P(\vartheta_1) = \exp\{-b(\vartheta_1) - a(\vartheta_1)\beta'(RV_t)\}\}$$

.

$$b(\vartheta_1) = \delta\ln(1 + c\vartheta_1) \qquad a(\vartheta_1) = \frac{c\vartheta_1}{1 + c\vartheta_1}$$
$$\vartheta_1 = \alpha_1 + \gamma\alpha_3 - \frac{1}{2}\alpha_3^2$$

$$\phi_X^P(\vartheta_2) = \exp\{-b(\vartheta_2) - a(\vartheta_2)\beta'(X_t)\}\}$$

.

$$b(\vartheta_2) = \delta \ln(1 + c\vartheta_2) \qquad a(\vartheta_2) = \frac{c\vartheta_2}{1 + c\vartheta_2}$$

$$\vartheta_2 = \alpha_2 + \gamma\alpha_3 - \frac{1}{2}\alpha_3^2$$

□

A4 proof of Equation4.19

Proof.

$$M_{t,t+1} = \frac{M_{t+1}}{E_t^P[M_{t+1}]}$$

$$= \frac{\exp\{-m_1\beta'(RV_t) - m_2\beta'(X_t) - m_3 y_{t+1}\}}{E_t^P[\exp\{-m_1\beta'(RV_t) - m_2\beta'(X_t) - m_3 y_{t+1}\}]}$$

$$= \frac{\exp\{-m_1\beta'(RV_t) - m_2\beta'(X_t) - m_3 y_{t+1}\}}{E_t^p[\exp\{-(m_1 + \gamma m_3 - \frac{1}{2}m_3^2)(\beta'(RV_t)\}\exp\{-(m_2 + \gamma m_3 - \frac{1}{2}m_3^2)(\beta'(X_t)\}]}$$

$$= \frac{\exp\{-m_1\beta'(RV_t) - m_2\beta'(X_t) - m_3 y_{t+1}\}}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}$$

Where

$$u_1 = m_1 + \gamma m_3 - \frac{1}{2}m_3^2 \qquad u_2 = m_2 + \gamma m_3 - \frac{1}{2}m_3^2$$

$$\phi_{RV}^P(u_1) = \exp\{-b(u_1) - a(u_1)\beta'(RV_t)\} \qquad \phi_X^P(u_2) = \exp\{-b(u_2) - a(u_2)\beta'(X_t)\}$$

$$b(u_1) = \delta \ln(1 + cu_1) \qquad a(u_1) = \frac{cu_1}{1 + cu_1}$$

$$b(u_2) = \delta \ln(1 + cu_2) \qquad a(u_2) = \frac{cu_2}{1 + cu_2}$$

□

A5 proof of Equation 4.21

Proof. In order to satisfy the no-arbitrage condition in Equation 4.21, we plug stochastic discount factor into Eq.4.21, then we have:

$$E_t^P\left[\frac{\exp\{-m_1\beta'(RV_t)-m_2\beta'(X_t)-m_3y_{t+1}+y_{t+1}\}}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}\right]=1$$

$$E_t^P\left[\frac{\exp\{-(v_1+\gamma(v_3-1)-\frac{1}{2}(v_3-1)^2)\beta'(RV_t)\}\exp\{-(v_2+\gamma(v_3-1)-\frac{1}{2}(v_3-1)^2)\beta'(X_t)\}}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}\right]$$

$$\frac{\phi_{RV}^P(\tilde{u}_1)\phi_X^P(\tilde{u}_2)}{\phi_{RV}^P(u_1)\phi_{RV}^P(u_2)}=1$$

where

$$\phi_{RV}^P(u_1)=\exp\{-b(u_1)-a(u_1)\beta'(RV_t)\}\qquad \phi_X^P(u_2)=\exp\{-b(u_2)-a(u_2)\beta'(X_t)\}$$
$$u_1=m_1+\gamma m_3-\frac{1}{2}m_3^2\qquad u_2=m_2+\gamma m_3-\frac{1}{2}m_3^2$$
$$\phi_{RV}^P(\tilde{u})=\exp\{-b(\tilde{u})-a(\tilde{u})\beta'(RV_t)\}\qquad \phi_X^P(\tilde{u})=\exp\{-b(\tilde{u})-a(\tilde{u})\beta'(X_t)\}$$
$$\tilde{u}_1=m_1+\gamma(m_3-1)-\frac{1}{2}(m_3-1)^2\qquad \tilde{u}_2=m_2+\gamma(m_3-1)-\frac{1}{2}(m_3-1)^2$$

Now we let $\tilde{u}_1=u_1$ & $\tilde{u}_2=u_2$, so as to satisfy right-hand side of Eq.**??**. Hence we have:

$$\begin{aligned}
&m_1+\gamma(m_3-1)-\frac{1}{2}(m_3-1)^2=m_1+\gamma m_3-\frac{1}{2}m_3^2\\
&-\gamma-\frac{1}{2}m_3^2+m_3-\frac{1}{2}=-\frac{1}{2}m_3^2\\
&m_3=\gamma+\frac{1}{2}
\end{aligned}$$

□

A6 proof of Equation 4.23

Proof.

$$E_t^Q\left[exp(-\alpha'\Pi_{t+1})\right] = E_t^P\left[M_{t,t+1}\exp(-\alpha'\Pi_{t+1})\right]$$

$$= E_t^P\left[M_{t,t+1}\exp\{-\alpha_1\beta'(RV_t) - \alpha_2\beta'(X_t) - \alpha_3 y_{t+1}\}\right]$$

$$= \frac{E_t^p\left[\exp\{-(m_1+\alpha_1)\beta'(RV_t) - (m_2+\alpha_2)\beta'(X_t) - (m_3+\alpha_3)y_{t+1}\}\right]}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}$$

$$= \frac{E_t^P\left[\exp\{-(m_1+\alpha_1) + \gamma(m_3+\alpha_3) - \frac{1}{2}(m_3+\alpha_3)^2\beta'(RV_t)\}\right]}{\phi_{RV}^P(u_1)\phi_{RV}^P(u_2)}$$

$$\times \frac{E_t^P\left[\exp\{-(m_2+\alpha_1) + \gamma(m_3+\alpha_3) - \frac{1}{2}(m_3+\alpha_3)^2\beta'(X_t)\}\right]}{\phi_{RV}^P(u_1)\phi_{RV}^P(u_2)}$$

$$= \frac{\phi_{RV}^P(\varpi_1)\phi_X^P(\varpi_2)}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}$$

where

$$\begin{aligned}
\varpi_1 &= (m_1+\alpha_1) + \gamma(m_3+\alpha_3) - \frac{1}{2}(m_3+\alpha_3)^2 \\
&= (m_1+\alpha_1) + \gamma(\frac{1}{2} + \gamma + \alpha_3) - \frac{1}{2}\left[\frac{1}{4} + \gamma + \alpha_3 + 2\alpha_3\gamma + \gamma^2 + \alpha_3^2\right] \\
&= \underbrace{\left(m_1 + \frac{1}{2}\gamma^2 - \frac{1}{8}\right)}_{\lambda_1} + \underbrace{\left(\alpha_1 - \frac{1}{2}\alpha_3 - \frac{1}{2}\alpha_3^2\right)}_{\varsigma_1} \\
&= \lambda_1 + \varsigma_1
\end{aligned}$$

and

$$\begin{aligned}
\varpi_2 &= (m_2+\alpha_2) + \gamma(m_3+\alpha_3) - \frac{1}{2}(m_3+\alpha_3)^2 \\
&= (m_2+\alpha_2) + \gamma(\frac{1}{2} + \gamma + \alpha_3) - \frac{1}{2}\left[\frac{1}{4} + \gamma + \alpha_3 + 2\alpha_3\gamma + \gamma^2 + \alpha_3^2\right] \\
&= \underbrace{\left(m_2 + \frac{1}{2}\gamma^2 - \frac{1}{8}\right)}_{\lambda_2} + \underbrace{\left(\alpha_2 - \frac{1}{2}\alpha_3 - \frac{1}{2}\alpha_3^2\right)}_{\varsigma_2} \\
&= \lambda_2 + \varsigma_2
\end{aligned}$$

Therefore, the LT of HAR-MIDAS under risk-neutral measure Q can be written as:

$$E_t^Q[\exp(-\alpha'\Pi_{t+1})] = E_t^P[M_{t,t+1}\exp(-\alpha'\Pi_{t+1})]$$

$$= \frac{\phi_{RV}^P(\varpi_1)\phi_X^P(\varpi_2)}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}$$

$$= \frac{\phi_{RV}^P(\lambda_1+\varsigma_1)\phi_X^P(\lambda_2+\varsigma_2)}{\phi_{RV}^P(\lambda_1)\varphi_X^P(\lambda_2)}$$

□

A7 proof of Proposition 2

Proof. Since

$$E_t^Q[\exp(-\alpha'\Pi_{t+1})] = E_t^P[M_{t,t+1}\exp(-\alpha'\Pi_{t+1})]$$

$$= \frac{\phi_{RV}^P(\varpi_1)\phi_X^P(\varpi_2)}{\phi_{RV}^P(u_1)\phi_X^P(u_2)}$$

$$= \frac{\phi_{RV}^P(\lambda_1+\varsigma_1)\phi_X^P(\lambda_2+\varsigma_2)}{\phi_{RV}^P(\lambda_1)\varphi_X^P(\lambda_2)}$$

$$= \frac{\exp\{-b(\lambda_1+\varsigma_1)-a(\lambda_1+\varsigma_1)\beta'(RV_t)\}\exp\{-b(\lambda_2+\varsigma_2)-a(\lambda_2+\varsigma_2)\beta'(X_t)\}}{\exp\{-b(\lambda_1)-a(\lambda_1)\beta'(RV_t)\}\exp\{-b(\lambda_2)-a(\lambda_2)\beta'(X_t)\}}$$

Thus

$$b(\lambda_1+\varsigma_1)-b(\lambda_1)=\delta ln\,(1+c(\lambda_1+\varsigma_1))-\delta ln\,(1+c\lambda_1)$$

$$=\delta ln\left[\frac{1+c\lambda_1+c\varsigma_1}{1+c\lambda_1}\right]$$

$$=\delta ln\left[1+\frac{c}{1+c\lambda_1}\varsigma_1\right]$$

where $c_1^*=\frac{c}{1+c\lambda_1}$

$$b(\lambda_2+\varsigma_2)-b(\lambda_2)=\delta ln\,(1+c(\lambda_2+\varsigma_2))-\delta ln\,(1+c\lambda_2)$$

$$=\delta ln\left[\frac{1+c\lambda_2+c\varsigma_2}{1+c\lambda_2}\right]$$

$$=\delta ln\left[1+\frac{c}{1+c\lambda_2}\varsigma_2\right]$$

where $c_2^*=\frac{c}{1+c\lambda_2}$, $\delta^*=\delta$

$$a^*(\varsigma_1)\beta^*_{RV_t}=a(\varsigma_1+\lambda_1)-a(\lambda_1)$$
$$a^*(\varsigma_1)\beta^*_{RV_t}=\left[\frac{c(\varsigma_1+\lambda_1)}{1+c\varsigma_1+c\lambda_1}-\frac{c\lambda_1}{1+c\lambda_1}\right]\beta'_{RV_t}$$

$$=\left[\underbrace{\left(\frac{c}{1+c\lambda_1}\right)}_{c_1^*}\left(\frac{\varsigma_1}{1+c\lambda_1+c\varsigma_1}\right)\right]\beta'_{RV_t}$$

$$=\left[\underbrace{\left(\frac{c}{1+c\lambda_1}\right)}_{c_1^*}\underbrace{(1-c_1^*\lambda_1)\,\beta'_{RV_t}}_{\beta^*_{RV}}\right]$$

where $\beta^*_{RV_t}=\frac{\beta'_{RV_t}}{1+c_1^*\lambda_1}$

$$
\begin{aligned}
a^*(\varsigma_2)\beta^*_{X_t} &= a(\varsigma_2+\lambda_2) - a(\lambda_2) \\
a^*(\varsigma_2)\beta^*_{X_t} &= \left[\frac{c(\varsigma_2+\lambda_2)}{1+c\varsigma_2+c\lambda_2} - \frac{c\lambda_2}{1+c\lambda_2}\right]\beta'_{X_t} \\
&= \left[\underbrace{\left(\frac{c}{1+c\lambda_2}\right)}_{c_2^*}\left(\frac{\varsigma_2}{1+c\lambda_2+c\varsigma_2}\right)\right]\beta'_{X_t} \\
&= \left[\underbrace{\left(\frac{c}{1+c\lambda_2}\right)}_{c_2^*}\underbrace{(1-c_2^*\lambda_2)\,\beta'_{X_t}}_{\beta^*_X}\right]
\end{aligned}
$$

where $\beta^*_{X_t} = \frac{\beta'_{X_t}}{1+c_2^*\lambda_2}$.

□

Chapter 5

Conclusions and Future Research Suggestions

This thesis consists of three papers that focus on evaluating the impacts of macroeconomic information on volatility, its long-term persistence and its structural changes, as well as European option pricing.

In the first paper, we explore the relationship between macroeconomic information and long-term stock market volatility, which has attracted considerable attention after the recent financial crisis and subsequent European debt crisis. We employ the GARCH-MIDAS model to evaluate the impact of macroeconomic information on stock volatility across three developed countries, U.S, UK and Japan. Adopting a wide range of macroeconomic variables for each country, we observe a significant improvement for volatility modelling and forecasting from including macroeconomic information throughout all three stock markets. We also find that the relationship between macro information and stock volatility is time-varying with regard to different underlying economic conditions. It is quite difficult to identify a dominant variable that contributes most in terms of explaining stock market volatility. To address with this we employ principal component analysis to compress the macroeconomic variables into common factors. We then use the the first principal component to examine if there is a "combined", or more general macroeconomic effect on stock volatility. Empirical results in the UK reveal that the first principal component outperforms most individual macroeconomic variables, in terms of goodness of fit across all sub-samples. We also find that there are significant volatility spillovers from the US to the UK and Japan and that including spillovers in the models provides a better understanding of the impact of macro information on local stock volatility.

Our study contributes to the literature on the relationship between the macroeconomy and stock return volatility by providing international evidence on the time-varying relationship between macro information and stock volatility and the importance of global volatility spillovers, proxied by US volatility, in explaining the relationship between volatility and the macroeconomy. Our results in the first chapter also highlight the importance of taking structural breaks into consideration when evaluating the macro influence on long-term volatility persistence. The main shortcoming in the GARCH-MIDAS model is that, without further partition, it can not incorporate potential structural changes in long-term volatility. Hence, further research evaluating the macroeconomic determinants of stock volatility can be carried out under a regime-switching volatility model, where the number of regimes are data driven.

In the second paper, we undertake an analysis of the impact of macroeconomic information on long-term persistence and structural changes in stock volatility. We firstly extend the Heterogeneous autoregressive Realized Volatility (HAR) model by including macroeconomic information using the Mixed data sampling (MIDAS) approach, which we term the HAR-MIDAS model. Our empirical evidence shows that there is a significant improvement in volatility modelling and forecasting via the HAR-MIDAS model specified with macroeconomic variables. We then extend the Tree-HAR model to allow macroeconomic information to determine volatility within local regimes and to perform as a threshold to identify regime-switching in volatility. In addition to the significant explanatory power of macroeconomic information, we also observe that macroeconomic information together with historical volatility jointly trigger structural changes in volatility. Both expected and unexpected macroeconomic information helps to deliver a more elaborate regime structure for stock volatility. In addition to high and low volatility regimes, we also identify a medium-volatility regime that itself can be further split by the macroeconomic threshold variables.

The findings in the second paper contribute to the existing literature on the macroeconomic determinants of structural changes in stock volatility. A limitation of the current literature is the lack of evidence concerning the variation of macroeconomic effects across different regimes in one realized volatility model. According to our estimation results, in the Tree-HAR model most macroeconomic variables re-define the medium-volatility regime into two sub-regions. Strikingly, a substantial amount of magnitude changes of macroeconomic effects can be found across those two sub-regions. Further research can examine whether these variations in the macroeconomic determinants of volatility across different regimes can be related to changes in monetary policy.

In the third paper, we apply our HAR-MIDAS model to the option pricing domain and examine the option valuation aspects of macroeconomic information in a realized volatility model. Our results reveal that macroeconomic variables helps to deliver better option valuation results, relative to the Duan's GARCH-MIDAS model, across maturity and moneyness. Other than that, unexpected information, measured by macroeconomic uncertainty, outperforms alternative macro variables in the sense that it delivers more accurate results in pricing out-of-money (OTM) options with long-maturity.

The findings in our third paper offer several avenues for further research.

Our HAR-MIDAS option model quantifies to what extent changes in economic conditions are reflected in option valuation, which might shed new light on using option data to infer market participants' reactions regarding changes in the economy. In our study, macroeconomic information together with return and realized volatility jointly determine the pricing kernel in the HAR-MIDAS option model. Consequently, the pricing kernel indicates the compensation for two source of risks: one from the stock market (both return and volatility) and one from macroeconomic conditions. Therefore, further research can also focus on the macroeconomic interpretation of the pricing kernel.

In sum, evaluating the relationship between macroeconomic information and asset volatility in this thesis not only enriches the Macro-finance literature but also provides empirical insights that are of benefit for market participants and policy makers. From the market participants' perspective, volatility estimation and prediction are of essential importance for risk management and asset allocation. As various assets behave differently during economic recession (also expansion), investor needs to rebalance the porfolio according to its underlying economic circumstance, so as to maintain a stable returns with lower volatility risk. Kollar (2013) points out that as assets behave differently over different stages of the business cycle, bringing macroeconomic information into consideration in asset allocation and the rebalancing of portfolio positions pays off in terms of more stable performance and lower volatility. Avramov et al. (2011) also state that incorporating macroeconomic information to improve volatility predictability is important in terms of optimal portfolios for hedge funds. In terms of option pricing, our results show that with the incoproation of macroeconomic information, we are able to isolate the long-term volatility component more precisely, somehting which is favourable for option pricing. This is particularly important because, as noted by Amin and Ng (1993), the short-term volatility component does not play an important role for option pricing, especially for long-maturity options. Instead accuracy is largely attributed to the long-term volatility component. Therefore, our study on the relationship between macroeconomic information and long-term volatility movements can surely shed some light on judging the accuracy of option pricing for individual stock options.

In terms of implications for research involving the macroeconomy, consider monetary policy. Its main objective is to achieve price stability and economic growth. Due to the close link between financial market volatility and macroeconomic information, financial market volatility actually reflects participants' expectations about the future direction of the economy and monetary policy. Central banks can make use of this to conduct (or adjust) its monetary policy based on the relationship between and feedback between volatility in the stock market and the macroeconomy. When financial markets are informative, the central bank will promptly adjust its monetary policy. However, when information flows are disrupted, such as during times of financial stress, ineffective policy and the spillover effect from financial markets to the real economy might lead to downturns in the economy (see Bemanke and Gertler (1989)Mishkin (2009)). Our research sheds some light on the linkage between the stock market and the macroeconomy and the role macroeconomic information can play in determining volatility regimes, which in turn provide valuable insight and information about periods of severe market and economic turbulence such as that during and following the most recent financial crisis. The research and results in this thesis suggest that further examination and investigation of the use of macroeconomic variables and macroeconomic information to improve asset allocation strategies, the accuracy of asset pricing models such as those used for pricing options as well as for hedging risk should prove fruitful. Our research also suggests that further research into the real response of different type of trading activities in financial markets to macroeconomic policies such as monetary policy would be beneficial.

References

Amin, K. I. and V. K. Ng (1993). Option valuation with systematic stochastic volatility. *The Journal of Finance 48*(3), 881–910.

Avramov, D., R. Kosowski, N. Y. Naik, and M. Teo (2011). Hedge funds, managerial skill, and macroeconomic variables. *Journal of Financial Economics 99*(3), 672–692.

Bemanke, B. and M. Gertler (1989). Agency costs, net worth, and business fluctuations. *American Economic Review 79*(1), 14–31.

Kollar, M. (2013). A sketch of macro-based asset allocation. *International Journal of Economic Sciences 2*(3), 101–120.

Mishkin, F. S. (2009). Is monetary policy effective during financial crises? *American Economic Review 99*(2), 573–77.