

# Advanced Quantitative Modelling and Analysis of Anomalies on Financial Markets — Feedback Trading and Realized Volatility

Thesis submitted in accordance with the requirements of  
the University of Liverpool for the degree of doctor in Philosophy

by  
Liu Fei

October 26, 2017

# Abstract

This thesis is mainly concerned with two broad topics: i) the impact of country ETF's premiums and discounts over feedback trading; ii) modelling high-frequency realized volatility on liquid assets. Out of the first topic, it investigates whether feedback trading exists in US-listed country ETFs and whether it varies with their observed/forecast premiums and discounts by using a sample of twenty country ETFs for the 2000-2016 window, it shows that feedback trading is present in several of them, particularly those targeting Asia Pacific markets. For the second topic, it analyses the forecastability and tradability of realized volatility on financial markets, specifically, major stock indices and their tradable derivatives are used to help decision makers in taking better hedging or trading positions in the short term. This thesis also extends the study on energy commodities which do not have their own (implied) volatility futures to trade. A heterogenous autoregressive model including jumps is used to model realized volatility, additionally, recurrent neural networks and a hybrid model are also added to the toolbox. It has been noticed, that the linear heterogeneous autoregressive model produces on average the most stable results.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Contents</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vi</b>
<b>1 Introduction and Outline</b>	<b>1</b>
<b>2 The impact of country ETF's premiums and discounts over feedback trading</b>	<b>5</b>
2.1 Introduction . . . . .	5
2.2 Feedback trading and ETFs . . . . .	8
2.3 Data - Methodology . . . . .	14
2.4 Results - Discussion . . . . .	23
2.5 Conclusion . . . . .	33
<b>3 Forecasting and Trading high Frequency Volatility on Large Index</b>	<b>36</b>
3.1 Introduction . . . . .	36
3.2 Literature Research . . . . .	38
3.3 Methodology . . . . .	41
3.3.1 The HAR-RV-J Model . . . . .	41

3.3.2	Recurrent Neural Networks . . . . .	44
3.3.3	Hybrid Model . . . . .	47
3.4	Data . . . . .	48
3.5	Empirical Results . . . . .	48
3.5.1	Statistical Errors . . . . .	52
3.5.2	Trading Efficiency . . . . .	57
3.6	Conclusion . . . . .	63
<b>4</b>	<b>An Intraday Risk Management Support System Based on Volatility</b>	<b>65</b>
4.1	Introduction . . . . .	65
4.2	Literature Review . . . . .	67
4.3	Methodology . . . . .	69
4.3.1	Realized Volatility . . . . .	69
4.3.2	Volatility Forecasting Model . . . . .	71
4.4	Dataset . . . . .	72
4.5	Empirical Results . . . . .	74
4.5.1	Forecast Evaluation . . . . .	74
4.5.2	Economic Significance . . . . .	75
4.5.3	Risk Management Support . . . . .	78
4.6	Discussion and Outlook . . . . .	81
<b>5</b>	<b>Analysing and Modelling High Frequency Realized Volatility of Energy Com-</b>	
	<b>modities</b>	<b>83</b>
5.1	Introduction . . . . .	83

5.2	Literature Research . . . . .	85
5.3	Method and Data . . . . .	87
5.3.1	The HAR Model . . . . .	88
5.3.2	Recurrent Neural Networks . . . . .	90
5.3.3	Hybrid Model . . . . .	91
5.3.4	Data . . . . .	92
5.3.5	Data Preprocessing . . . . .	93
5.4	Results . . . . .	94
5.4.1	Crude Oil . . . . .	95
5.4.2	Natural Gas . . . . .	100
5.4.3	Summary of Results . . . . .	104
5.5	Conclusions . . . . .	106
<b>6</b>	<b>Discussion and Limitations</b>	<b>108</b>
6.1	Discussion of Modelling ETF Premium . . . . .	108
6.1.1	Problem Outline . . . . .	108
6.1.2	Modelling Issues . . . . .	109
6.1.3	Financial Viability . . . . .	110
6.2	Discussion of Modelling Realized Volatility . . . . .	112
6.2.1	Problem Outline . . . . .	112
6.2.2	Data Availability . . . . .	112
6.2.3	Application Domain . . . . .	113
<b>7</b>	<b>Outlook and Recommendations</b>	<b>114</b>

7.1	Conclusions and Outlook . . . . .	114
7.2	Key Recommendations . . . . .	118
<b>A</b>	<b>Abbreviations</b>	<b>120</b>

# Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Dr Athanasios Pantelous for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. Beside my supervisor, I would like to thank the rest of my thesis committee, Prof. Dr. Hans-Jörg von Mettenheim, Dr Vasileios Kallinterakis, for their insightful comments and encouragement.

I would like to convey my thanks to my fellow postgraduate students and the staff in the Institute of Risk and Uncertainty and Department of Mathematical Science, for their support and help during my Ph.D study.

Last but not least, I would like to place my great gratitude to my mum for her everlasting love and support.

# Chapter 1

## Introduction and Outline

My thesis is mainly concerned with, potentially, unearthing anomalies and inefficiencies on financial markets using advanced modelling and analysis methods. With this relatively broad topic, I will mainly deal with two subtopics:

- Investigating the impact of country ETFs' premiums and discounts over feedback trading.
- Modelling high-frequency realized volatility on liquid assets.

For the first topic, a sample of twenty country ETFs for the 2000-2016 window is used to study whether feedback trading exists in these US-listed country ETFs and whether it varies with their observed/forecasted premiums and discounts. It is assessed whether the findings hold before and after the outbreak of the 2008 financial crisis, given earlier evidence on the effect of crises over investors' feedback trading internationally. It further explores whether the presence of feedback trading varies with the size of an ETF's observed/forecast premium/discount. Given that this study entails both observed and forecast premium/discount values, it also tests whether successful premium/discount forecasts are associated with feedback trading patterns in the sample's ETFs. Country ETFs targeting Asia Pacific markets are found to be more prone to feedback trading (compared to those targeting European and Latin American markets).



For the second topic, it analyses the forecastability and tradability of realized volatility on financial markets by applying the Heterogeneous Autoregressive model of Realized Volatility. In Chapter 3, it analyses the forecastability and tradability of volatility on the S&P500 index and the liquid SPY ETF, VIX index and VXX ETN. Even though there is already a huge array of literature on forecasting high frequency volatility, most publications only evaluate the forecast in terms of statistical errors. In practice, this kind of analysis is only a minor indication of the actual economic significance of the forecast that has been developed. For this reason, in my approach, it also includes a test of the forecast through trading an appropriate volatility derivative. As a method parametric and artificial intelligence models are used. This study also combines these models in order to achieve a hybrid forecast. It reports that the results of all three model types are of similar quality. However, it has been observed that artificial intelligence models are able to achieve these results with a shorter input time frame and the errors are uniformly lower comparing with the parametric one. Similarly, the chosen models do not appear to differ much while the analysis of trading efficiency is performed. Finally, it has been noticed that Sharpe ratios tend to improve for the longer forecast horizon.

In Chapter 4, it specifically looks at major stock indices and their tradable derivatives to help decision makers in taking better hedging or trading positions in the short term. As Chapter 3 shows promising results on the S&P500 index and the corresponding Volatility Index VIX it is a sensible step to check, whether these results can be reproduced or improved on other assets. It has been made a point to specifically account for trading efficiency by looking at tradable assets and not just assuming the theoretical index data. Computation effort to build realistic models is significant and different model variants have been checked, including the presence of jumps in the

intraday realized volatility. For the assets in this study the results are encouraging with Sharpe ratios that are better than a buy and hold investment.

As numerous studies have been applied to the task of forecasting daily implied volatility on large stock indices, in Chapter 5, it extends the research on volatility in two ways: firstly, it analyses energy commodities which do not have their own (implied) volatility futures to trade. This means, that a good forecast of implied volatility cannot directly be exploited. Rather, a good forecast would have to be linked to a corresponding trade in the underlying. Secondly, instead of focusing on daily implied volatility this chapter targets high frequency intraday realized volatility. Arguably, this brings us closer to modelling the true volatility process. A heterogeneous autoregressive model including jumps is used to model realized volatility. While this model is refreshingly simple and straightforward to estimate, it features all the stylized facts of volatility and does a good job at this task. Furthermore, recurrent neural networks and a hybrid model are added to the toolbox of useful models. It has been noticed, that the linear heterogeneous autoregressive model produces on average the most stable results. However, the non-linear models are able to produce attractive out-of-sample results if care is taken to build the model properly and especially on short look-back intervals.

Due to the cumulative nature of my thesis (based on academic papers), the research questions dealt with in my thesis are twofold:

- Are actual/forecast tracking errors (premium/discounts) of country ETFs related to their feedback trading?
- Is high-frequency realized volatility predictive and economically exploitable for future realized volatility?

Ultimately, a focus on predictivity can be seen. Nevertheless, I place my thesis in the realm of market anomalies / inefficiencies, because any exploitable predictivity could be regarded as a market anomaly

The remainder of this thesis is structured as follows. The next section detail the first aspect of my two main research areas; that is investigating the impact of country ETF's premiums and discounts over feedback trading. The following three sections present the second aspect of my main research areas, which are modelling realized volatility. I highlight the importance of both problems and put them in the context of today's research. The sixth section discusses limitations of my work and outlines potential future research. Finally, the last section wraps up the thesis and summarizes the work.

## Chapter 2

# The impact of country ETF's premiums and discounts over feedback trading

### 2.1 Introduction

Research (Deville (2008)) on exchange traded funds (ETFs, hereafter) has denoted the presence of significant deviations of US country ETFs' prices from their underlying net asset value (NAV, hereafter), leading these ETFs to document substantial premiums and discounts. This has been fundamentally attributed to the non-synchronicity in trading between these ETFs (traded in the US) and their underlying benchmark portfolios' stocks. As the US market and the markets of these ETFs' underlying benchmarks are not simultaneously open for trading, the deviations of these ETFs' prices from their NAVs cannot be arbitrated away real-time, thus raising the possibility of their exploitation via *ad hoc* designated trading strategies, whose profitability has been confirmed in several studies (Cherry (2004); Jares and Lavin (2004); Engle and Sarkar (2006); Ackert and Tian (2008)). Considering that such strategies are essentially feedback in style, a question arising is whether US country ETFs' premiums/discounts give rise to distinct feedback trading patterns in these ETFs' trading process.

This issue is addressed by drawing on a sample of twenty US-listed country ETFs targeting

a series of markets in the Americas, Asia Pacific and Europe by investigating whether they accommodate feedback trading and whether the latter's presence varies with these ETFs' *observed* price-deviations from their NAVs. What is more, it controls for the possibility that feedback traders condition their trades on predicted (rather than observed) premium/discount values by examining whether feedback trading in US country ETFs varies with their *forecast* price-deviations from their NAVs. It further assesses whether these findings hold before and after the outbreak of the 2008 financial crisis, given earlier evidence on the effect of crises over investors' feedback trading internationally. Moreover, it explores whether the presence of feedback trading varies with the level of an ETF's observed/forecast price-deviations from their NAV; given that this study entails both observed and forecast premium/discount values, it also tests whether successful premium/discount forecasts are associated with feedback trading patterns in the sample's ETFs.

Overall, these results reveal that feedback traders are active in several US-listed country ETFs, with their presence being sensitive to the time period examined (pre versus post 2008 crisis) and the sign and level of the (observed and forecast) price-deviations of each ETF from its NAV (i.e., premiums and discounts); it also has been found very little evidence in support of feedback trading patterns emerging upon the realization of successful premium/discount forecasts. As a general observation, feedback traders are active the most in those ETFs targeting Asia Pacific markets, with little (no) evidence of their presence documented in ETFs targeting markets in Europe (the Americas). These findings are attributed to the noise trading often encountered in Asia Pacific markets, leading ETFs investing there to exhibit feedback trading either due to them mirroring (given their tracking nature) these markets' price trends, or due to these ETFs' investors choosing to feedback trade as a rational response to these markets' noise trading levels. It is also possible

that the results are due to the lack of overlapping trading sessions between the US and Asia Pacific markets leading these ETFs' NAVs to be known and possibly used as reference points - before the start of trading in the US.

This research contributes to the extant ETF literature by showcasing that the widely documented premiums/discounts of US-listed country ETFs are related to feedback trading, particularly for those ETFs targeting markets in Asia Pacific region. Although the literature has mainly focused on country ETFs' actual premiums/discounts, it has been demonstrated that forecast premiums/discounts are associated with feedback trading as much as actual ones, thus demonstrating that forecasting country ETFs' price-deviations from their NAVs (irrespective of forecasts' success) can be used to attain insight into behavioural patterns in their trading. Furthermore, the evidence presented here is of key interest to country ETFs' investors (in particularly those targeting Asia Pacific markets, in view of these results), as it allows them insight into the trading dynamics associated with these ETF's premiums/discounts that can be used to inform their trades. If an investor trading such an ETF discovers, for example, that it accommodates feedback trading for specific observed/forecast premiums/discounts, they can develop a strategy that takes into account the anticipate feedback trading contingent upon the realization of a certain (observed or forecast) premium/discount by that ETF.

The rest of this chapter is organized as follows: the next section presents a review of the literature on feedback trading and ETFs, while section 2.3 introduces the data utilized with descriptive statistics and delineates the methodology employed. Section 2.4 presents and discusses the results and section 2.5 concludes by summarizing the study's main findings and outlining their implications.

## 2.2 Feedback trading and ETFs

Feedback trading is an umbrella term encompassing any trading strategy based on the identification of patterns in historical market data. Feedback traders believe that prices exhibit inertia (Farmer (2002)), are characterized by trends of a repetitive (and, hence, predictable) nature that can be profitably exploited via *ad hoc* trading rules. The prevalence of feedback trading in the market can amplify existing price trends, leading prices to depart from fundamentals (De Long et al. (1990)) and enhance serial correlation (Cutler et al. (1990)) and excess volatility (Farmer (2002); Farmer and Joshi (2002)) in the return-generation process. Feedback traders are distinguished into positive and negative feedback traders and their conduct can be motivated by a notably wide array of factors, both rational, as well as behavioural.

From a rational perspective, *rational speculation* (De Long et al. (1990)) can lead informed investors to exploit their superior informational foresight by entering positions in stocks prior to the release of news, in order to launch price trends in the market and profitably exploit them. Investors also engage in feedback trading believing they can extract useful information from historical prices when the *information risk* of their investments is high. This is the case, for example, when investing in small capitalization stocks, about which little information is normally available (Lakonishok et al. (1992); Wermers (1999); Sias (2004); Voronkova and Bohl (2005)) and foreign stocks (given the perceived information superiority of overseas markets' indigenous traders, (see Brennan and Cao (1997); Lin and Swanson (2008)). *Style investing* (see Bennett et al. (2003)) is a key driver of feedback trading, since several investment styles popular among institutional investors, including momentum and contrarian strategies, (see Galariotis (2014) for an excellent review on both), are based on historical prices. *Technical analysis*, (see e.g. Fong and Yong (2005)) is another key

expression of feedback trading, while the latter can also be driven by traditional trading practices, including *portfolio insurance* (Kodres (1994)), *stop-loss orders* (Osler (2005)), and *margin trading* (Watanabe (2002); Hirose et al. (2009)). *Professional reasons* are conducive to feedback trading as well, with fund managers often buying stocks with positive recent performance in order to generate a positive impression as regards their skills (Lakonishok et al. (1992)).

From a behavioural perspective, investors resort to feedback trading primarily due to *observational learning*: prices provide a statistical summary of market activity (Holmes and Kallinterakis (2014)) that indirectly allows them insight into the trades of other market participants, without the need to actively monitor the latter (Hirshleifer and Teoh (2003); Hirshleifer et al. (2011)). The *representativeness heuristic* can motivate trend-chasing (Barberis et al. (1998)), since it can prompt investors to buy (sell) a stock after only a few days of positive (negative) performance. This can be further reinforced by the *availability bias* (Barberis and Thaler (2003)), according to which more (less) recent events are more (less) easily retrievable by human memory and enjoy a higher (lower) weight in decision-making. *Anchoring* (Barberis and Thaler (2003)) is also relevant here, since using reference points in trading is very common among feedback-style strategies.

Empirical evidence on feedback trading overall confirms its presence internationally across several markets, asset classes and investor types. As far as studies using micro data are concerned, positive feedback trading has been found to be popular among US fund managers, with its magnitude being greater in more recent studies (Sias (2004); Froot and Teo (2008); Choi and Sias (2009)) compared to older ones (Lakonishok et al. (1992); Grinblatt et al. (1995); Wermers (1999)). US retail investors engage less in positive feedback trading compared to their institutional counterparts (Nofsinger and Sias (1999)), while the sign of feedback trading of retail investors in Germany varies



with the order-type they employ (Dorn et al. (2008)). Walter and Weber (2006) report significant positive feedback trading among German mutual funds; conversely, (Kremer and Nautz (2013)) show that German funds are contrarian traders, similar to UK funds (Wylie (2005)). Choe et al. (1999) report significant positive feedback trading for overseas investors in South Korea prior to the Asian crisis, with this feedback trading largely dissipating following the crisis' outbreak. On the other hand, Kim and Wei (2002a,b) find that foreign institutional investors exhibit more positive feedback trading in the South Korea market in the aftermath (as opposed to before) the Asian crisis' outbreak, while Bowe and Domuta (2004) report very limited evidence of feedback trading for foreign and domestic investors in Indonesia before, during and after the Asian crisis. Hung et al. (2010) find that mutual funds tend to negative feedback trade in Taiwan, while Feng and Seasholes (2004) detect no evidence of feedback trading among retail investors in China. Finally, the global study by Choi and Shiba (2015) presents evidence indicating the prevalence of positive feedback trading among institutional investors internationally. Turning now to studies using aggregate data, significant positive feedback trading has been documented for equity (Sentana and Wadhwani (1992); Koutmos (1997); Koutmos and Saidi (2001); Watanabe (2002); Koutmos et al. (2006); Bohl and Siklos (2008); Schuppli and Bohl (2010); Chau and Deesomsak (2015)), currency (Aguirre and Saidi (1999); Laopodis (2005)) and energy (Chau et al. (2015)) markets, while futures markets offer limited evidence of feedback trading presence (Antoniou et al. (2005); Chau et al. (2008)).

The availability of feedback trading opportunities can be seen as evidence going against Eugene Fama's 1960s Efficient Market Hypothesis (EMH). EMH, roughly, states that all information that is currently available is reflected in current prices. There are, however, different variants of EMH,

namely the weak, semi-strong and strong form efficiency. Specifically, weak form efficiency assumes that only all past publicly available information only is already subsumed in the current price. Insofar, we might not necessarily assume that feedback trading opportunities violate weak-form efficiency as the numbers we base our computations on are not necessarily readily publicly available (although this remains, of course, open to debate). Additionally, as also outlined elsewhere, the availability of a theoretical feedback trading opportunity does not necessarily signify that excess returns can be earned on a risk-adjusted basis. Several market restrictions (including spreads, liquidity, short selling restrictions, and more) can make the opportunity unprofitable.

When discussing any type of active trading strategy, it is useful to also assess the tax implications. Insofar, feedback trading is not different from any other active trading strategy. If we consider the amount of activity, feedback trading will be at the higher frequency end of the range as it is expected to potentially trade every day. Profits of such a strategy will, typically, be small. The trader will hope to recoup a share of the premium or discount through the trading activity. Considering the small potential profits, feedback trading will be sensitive to any kind of stamp duty (Tobin tax like) transaction tax. A prime example for this might be found in the Irish market, where a 1 percent tax is levied on stock transactions. A stamp duty following the Irish model would pretty much wipe out any potential gain from a feedback trading strategy. An explanation of these mechanics can be found in Bond et al. (2005)

To a lesser extent, capital gains taxes will also have to be considered. However, the effect will be more mitigated as gains and losses can, typically, compensate each other when considering the tax levied. Additionally, most institutional traders will be able to design their trading in a way that avoids paying capital gains taxes.

Exchange traded funds have gained in popularity during the 1990s. They were designed as a low-cost and more flexible alternative to mutual funds. Indeed, ETFs, typically feature low total expense ratios, because they are, by design, a passive investment. (Some very new ETFs are hybrid, featuring elements of active management. These ETFs are omitted from the discussion here.) The line of thought that pervades most ETFs, is that "markets cannot be beaten consistently". With this as a mantra, pure index tracking strategies make sense. Indeed, if we believe, that we cannot, for example, consistently beat the S&P500 index on a risk-adjusted basis, then an investor could, just as well, invest in a low-cost S&P500 index ETF. The performance of ETFs has been very satisfying in the sense that most liquid ETFs manage to track their index with a low error. Insofar, their performance is satisfying to investors, because they get exactly what they ask for, for minimal costs. Of course, if the belief is that an active manager can perform better than the market, then it might be worth paying a premium for (hopefully) getting a better risk-adjusted return than the market can deliver. We will not discuss the relative merits of these arguments in the context of this short introduction on ETFs. All in all it can be said, that ETFs are nowadays very popular as an index tracking tool among investors. Their low fees and (typically) high liquidity make them attractive to retail and institutional investors alike who use ETFs to implement their market views. In the US, the world's largest ETF market, there exist 1,716 ETFs of multiple types with a combined value of just over USD 2.524 trillion, while the number of ETFs globally amounts to 3,259 with a combined market value of USD 3.306 trillion. ETFs also possess other attractive properties, including instant exposure, transparency, dividend-treatment, risk management, and tax-efficiency, which have been delineated in a series of studies (Gastineau (2001); Kostovetsky (2003); Deville (2008)) and which help explain the wide popularity ETFs have been enjoying among both retail

and institutional investors (Charteris et al. (2014)).

Evidence on the behaviour of ETF-traders has indicated that they subscribed to feedback-style strategies. Drawing on high frequency data from US ETFs during the internet bubble, Madura and Richie (2004) demonstrate the presence of intraday overreaction patterns in their trading dynamics that correct themselves within the same day, thus presenting profitable opportunities to traders with intraday horizons. Chau et al. (2011) find that the US' three largest ETFs ("Spiders"; "Cubes"; "Diamonds") are characterized by significant positive feedback trading, whose presence grows more pronounced during bullish sentiment periods. Chen et al. (2012) show that US institutional investors negative feedback trade when investing in ETFs, while Charteris et al. (2014) report limited evidence of feedback trading among ETFs traded in emerging markets.

A factor capable of encouraging feedback trading in ETFs is their tracking error, namely the deviations between their market price and their net asset value, which can be either positive or negative. The presence of large premiums/discounts in an ETF implies inefficiency in its pricing and, as such, would be expected to be arbitrated away, particularly given the in-kind creation/redemption mechanism discussed previously. However, for arbitrage to be feasible in this case it is necessary for the ETF and its portfolio's underlying assets to be traded simultaneously. This is not the case, though, for the specific category of country ETFs, which are listed predominantly in the US and invest in equities in overseas markets, whose trading times are in the vast majority of cases, thus rendering arbitrage a technical impossibility. In view of country ETFs' substantial premiums/discounts (Harper et al. (2006); Deville (2008); Blitz and Huij (2012)), it is reasonable to expect that investors will attempt to exploit this pricing inefficiency by employing strategies based on these ETFs' historical premium/discount pattern; indeed, evidence to date (Cherry (2004); Jares

and Lavin (2004); Engle and Sarkar (2006); Ackert and Tian (2008)) has confirmed the profitability of such strategies. Being based on historical ETF prices and their deviations from NAVs, these strategies are essentially feedback in style and, although this suggests that they can potentially give rise to distinct feedback trading patterns in country ETFs, the latter has not been empirically assessed to date. This study contributes to the literature by examining this issue drawing on a sample of twenty US country ETFs and produces results indicating that these ETFs' (observed and forecast) premiums/discounts are associated with feedback trading, particularly for those ETFs targeting Asia Pacific markets. The next section provides a detailed presentation of the ETFs included in the sample with detailed descriptive statistics, while also introducing the methodology utilized for the empirical investigation.

## 2.3 Data - Methodology

The data includes daily observations of the closing prices and net asset values of twenty iShares MSCI ETFs, which are presented in Table 2.1 (panel A). The data covers the period between June 20th, 2000 and April 27th, 2016 and has been obtained from Thomson-Reuters DataStream (closing prices) and Black Rock iShares (NAVs), with the observations from both databases matched. The choice of June 20th, 2000 as the starting date of the sample coincides with the launch-date of the iShares MSCI Taiwan ETF (the ETF with the latest launch date out of all twenty ETFs) and the reason for this is that it aimed at including in the sample all US-listed country ETFs launched before 2001 in order to have a sufficiently long pre crisis window when testing for the effect of the 2008 crisis over the results. The names of the ETF will for some include the denomination *capped*. This is just an indication of how ETF portfolios are set up. For some countries the market capitalization of individual companies is so large that just this single company would significantly

skew the returns of the ETF. In this case, the weight of a single company may be capped. This is indicated in the name of the ETF.

Table 2.1 (panel B) provides a series of descriptive statistics (mean; standard deviation; skewness; kurtosis; Jarque-Bera normality test, Ljung-Box test statistic for returns and squared returns for the ten lags) pertaining to the log-differenced returns of the sample ETFs. Sixteen (four) ETFs exhibit negative (positive) skewness, while all twenty ETFs presents us with rather large Jarque-Bera test-statistics and leptokurtosis in their returns' distribution. To gauge whether these departures from normality are the product of temporal dependencies in the series' structures, the Ljung-Box portmanteau test are applied on the first and second moment of all ETFs' returns. All Ljung-Box test-statistics on ETFs' returns are significant (at least at the 5 percent level), indicating the presence of significant autocorrelations in the ETFs' return distributions; this, however, is not in itself evidence in support of feedback trading, since dependencies in the first moment of returns can also be due to market inefficiencies, such as thin trading. In view of the documented (see e.g. Farmer (2002)) ability of feedback traders to accentuate volatility in capital markets, it tests for higher moment temporal dependencies by calculating the Ljung-Box test-statistics for squared returns. As the results indicate, all of these test-statistics are significant (at the 1 percent level) and always higher in value than the Ljung-Box test-statistics calculated previously for returns, thus confirming the presence of time-varying volatility in the ETFs. The presence of significant first- and second- order temporal dependencies in financial time series is well-established in the literature (Bollerslev et al. (1994)) and in the next section it will investigate whether they are related to feedback trading.

Panel C presents some statistics on each ETF's percentage price deviations from its net asset

value contingent on their sign (premiums, if the sign is positive; discounts, if it is negative). The average percentage deviation of ETFs's prices from their NAVs is positive, denoting that US-listed country ETFs traded on average at a premium during the full sample period, with almost all ETFs<sup>1</sup> having traded on average at a premium over the entire sample period. On average, the sample ETFs traded 55.5% (44.5%) of the time at a premium (discount), with emerging country ETFs tending to trade more often at a discount compare to developed country ETFs. The results, broadly, show that on average we observe a premium for most ETFs. The notable exception are two emerging markets ETFs which on average exhibit a discount. (While still other emerging markets ETFs also exhibit a premium.) This is only a rough observation and a more detailed analysis on the causes shall be left for future work. However, there is a plausible reason, why this might be the case, at least for some of the emerging market ETFs (Two emerging markets show this behaviour in our sample.) Indeed, traders in some emerging market ETFs may, rightly, assume that the shares hold within the ETF are illiquid and difficult to trade. Insofar there may be some insecurity as to whereas the displayed prices of the individual shares could be fully realized should the ETF issuer have to sell them. This might lead to a discount to NAV in this ETF. On the other hand the premium in the developed market ETFs (and some emerging market ETFs) is slightly less pronounced. However, we might explain a small premium by the added convenience to be able to buy an index portfolio, instead of individual shares. Indeed, if an investor wanted to replicate a large index portfolio substantial transaction costs might be incurred. This argument, of course, also holds for the emerging market ETFs. However it seems that the pricing insecurity for the emerging markets would be a stronger effect for two ETFs in our sample and lead to the average discount.

To empirically address the research questions of this study, it relies on the model developed

---

<sup>1</sup>The sole exception here is the iShares MSCI Austria Capped ETF, which traded on average at a discount.

by Sentana and Wadhwani (1992), which assumes the interaction of two groups of traders in the market. The first group consists of rational speculators who maximize their expected utility based on a mean-variance framework, as reflected in their demand function below:



Table 2.1: Sample Statistics

Panel A: List of sample ETFs							
iShares MSCI Australia ETF				iShares MSCI Malaysia ETF			
iShares MSCI Austria Capped ETF				iShares MSCI Mexico Capped ETF			
iShares MSCI Belgium Capped ETF				iShares MSCI Netherlands ETF			
iShares MSCI Brazil Capped ETF				iShares MSCI Singapore ETF			
iShares MSCI Canada ETF				iShares MSCI South Korea Capped ETF			
iShares MSCI France ETF				iShares MSCI Spain Capped ETF			
iShares MSCI Germany ETF				iShares MSCI Sweded ETF			
iShares MSCI Italy Capped ETF				iShares MSCI Switzerland Capped ETF			
iShares MSCI HongKong ETF				iShares MSCI Taiwan ETF			
iShares MSCI Japan ETF				iShares MSCI United Kingdom ETF			
Panel B: Descriptive Statistics							
ETF	Mean(%)	Standard Deviation	Skewness	Kurtosis	Jarque-Bera	LB(10)	Lb <sup>2</sup> (10)
iShares MSCI Australia ETF	0.0197	0.0182	-0.1430	12.1400	12350***	56.22***	3417.3***
iShares MSCI Austria Capped ETF	0.0141	0.0182	-0.6369	11.6014	11174***	7.57**	2514.5***
iShares MSCI Belgium Capped ETF	0.0129	0.0164	-0.5508	9.9403	7298***	5.58**	2407.8***
iShares Brazil Capped ETF	-0.0879	0.0201	-0.2495	9.4194	6857***	31.13***	2837.0***
iShares MSCI Canada ETF	0.0242	0.0148	-0.4438	9.4130	6194***	13.60**	3524.7***
iShares MSCI France ETF	0.0082	0.0174	-0.2503	9.1038	5543***	37.29***	2077.4***
iShares MSCI Germany ETF	0.0171	0.0177	-0.0643	10.9100	9244***	26.33***	1699.8***
iShares MSCI Hong Kong ETF	0.0243	0.0167	0.1328	12.2000	12518***	113.56***	3501.4***
iShares MSCI Italy Capped ETF	-0.0104	0.0189	-0.3989	9.0500	5501***	26.01***	1435.3***
iShares MACI Japan ETF	0.0110	0.0144	0.1420	11.8631	11622***	39.87***	2223.4***
iShares MSCI Malaysia ETF	-0.0861	0.0203	-0.2649	22.2801	78469***	28.33***	506.0***
iShares MSCI Mexico Capped ETF	0.0326	0.0198	-0.0259	11.7451	16134***	27.6***	1443.0***
iShares MSCI Netherlands ETF	0.0103	0.0168	-0.4216	9.8133	6965***	24.12***	2614.5***
iShares MSCI Singapore ETF	0.0220	0.0168	0.0757	10.6823	8726***	95.56***	2866.9***
iShares MSCI South Korea Capped ETF	0.0248	0.0226	0.0954	2.2247	14234***	61.08***	2684.0***
iShares MSCI Spain Capped ETF	0.0080	0.0189	-0.3589	9.9694	7254***	27.20***	1302.5***
iShares MSCI Sweden ETF	0.0262	0.0207	-0.2652	9.9811	5328***	42.87***	2784.1***
iShares MSCI Switzerland Capped ETF	0.0243	0.0140	-0.3752	9.0499	5492***	42.27***	2278.8***
iShares MSCI Taiwan ETF	0.0805	0.0266	-0.0931	7.5724	3476***	53.05***	1497.0***
iShares MSCI United Kingdom ETF	0.0102	0.0173	-0.1121	9.3547	7654***	30.12***	1789.0***
Panel C: Statistics on percentage price deviation from NAV							
	Average price deviation (%)	Average premium (%)	Average discount(%)	% of days when ETF trades at premium	% of days when ETF trades at discount		
iShares MSCI Australia ETF	0.1284	0.7660	-0.8398	0.6036	0.3964		
iShares MSCI Austria Capped ETF	-0.0029	0.0057	-0.0066	0.5318	0.4682		
iShares MSCI Belgium Capped ETF	0.0734	0.5676	-0.5664	0.5647	0.4353		
iShares MSCI Brazil Capped ETF	0.0357	0.7200	-0.6100	0.5503	0.4497		
iShares MSCI Canada ETF	0.0669	0.3694	-0.2924	0.5445	0.4555		
iShares MSCI France ETF	0.0537	0.5073	-0.5170	0.5577	0.4423		
iShares MSCI Germany ETF	0.0475	0.4930	-0.5199	0.5607	0.4393		
iShares MSCI Hong Kong ETF	0.0277	0.7404	-0.8570	0.5541	0.4459		
iShares MSCI Italy Capped ETF	0.0534	0.5435	-0.5608	0.5566	0.4434		
iShares MSCI Japan ETF	0.0799	0.8182	-0.8490	0.5576	0.4424		
iShares MSCI Malaysia ETF	0.4426	2.1700	-1.5300	0.4782	0.5218		
iShares MSCI Mexico Capped ETF	0.0542	0.5700	-0.6500	0.4899	0.5101		
iShares MSCI Netherlands ETF	0.0504	0.5091	-0.5494	0.5670	0.4330		
iShares MSCI Singapore ETF	0.0286	0.7259	-0.8846	0.5673	0.4327		
iShares MSCI South Korea Capped ETF	0.0391	1.0100	-1.0500	0.5322	0.4678		
iShares MSCI Spain capped ETF	0.0486	0.5465	-0.5737	0.5559	0.4441		
iShares MSCI Sweden ETF	0.1058	0.6634	-0.6576	0.5784	0.4216		
iShares MSCI Switzerland Capped ETF	0.1585	0.5495	-0.4815	0.6221	0.3779		
iShares MSCI Taiwan ETF	0.1622	1.1300	-0.9900	0.5418	0.4582		
iShares MSCI United Kingdom ETF	0.0544	0.5028	-0.5531	0.5592	0.4408		

<sup>1</sup>The table above contains a series of information on the sample ETFs used in our study. The list of the twenty ETFs employed here is outlined in panel A. Panel B presents a series of descriptive statistics on the log-difference returns of our twenty ETFs; these statistics include the mean, standard deviation, skewness, kurtosis, Jarque-Bera normality test-statistics and Ljung-Box test-statistics at ten lags for the return- and squared return-series of the twenty ETFs. \* \*\* \*\*\* represent significance at the 10%, 5% and 1% levels, respectively. Panel C contains summary statistics on the observed percentage price deviations of each ETF from its NAV; these statistics include the average price deviation (%), the average premium (%), the average discount (%) and the percentage of days for which an ETF has traded at a premium/discounts.

$$Q_t = \frac{E_{t-1}(r_t - \alpha)}{\theta \sigma_t^2}, \quad (2.3.1)$$

In Equation 2.3.1 above,  $E_{t-1}(r_t)$  is the expectation in period  $t - 1$  of the ETF's return,  $r_t$ , in period  $t$ ,  $\alpha$  is the risk-free return,  $\theta$  is the time-invariant coefficient of risk-aversion and  $\sigma_t^2$  is the conditional variance (proxying for risk) at period  $t$ .

The second group comprises of feedback traders, who trade on the premises of historical prices, which means they buy (sell) after the price increase (decrease), their demand function is reflected as follows:

$$Y_t = \gamma r_{t-1}, \quad (2.3.2)$$

As Equation 2.3.2 suggests, feedback traders base their trades on the previous period's return, with the direction of their trades varying, depending on whether they positive (i.e., if  $\gamma > 0$ , in which case, they buy if  $r_{t-1} > 0$  and sell if  $r_{t-1} < 0$ ) or negative (i.e., if  $\gamma < 0$ , in which case, they buy if  $r_{t-1} < 0$  and sell if  $r_{t-1} > 0$ ) feedback trade. For the market to be in equilibrium, all shares must be held, in which case:

$$Q_t + Y_t = 1, \quad (2.3.3)$$

It follows from equation 2.3.1 and 2.3.2 that we obtain:

$$E_{t-1}(r_t) = \alpha + \theta \sigma_t^2 - \theta \gamma \sigma_t^2 r_{t-1}, \quad (2.3.4)$$

To estimate Equation 2.3.4 we convert the expected return,  $E_{t-1}(r_t)$  into a realized one ( $r_t$ ), by assuming the latter's rational expectation  $r_t = E_{t-1}(r_t) + \varepsilon_t$ , where  $\varepsilon_t$  is a stochastic error term:

$$r_t = \alpha + \theta \sigma_t^2 - \theta \gamma \sigma_t^2 r_{t-1} + \varepsilon_t, \quad (2.3.5)$$

As Equation 2.3.5 shows, the first-order return-autocorrelation interacts both with risk ( $\sigma_t^2$ ), and

feedback trading (the first-order autocorrelation sign will be positive if  $\gamma < 0$  and negative if  $\gamma > 0$ ). However, autocorrelation can be the results of both inefficiencies in the market (such as, for example, thin trading) as well as feedback traders and Equation 2.3.5 does not allow us to disentangle between the two possibilities. To that end, Sentana and Wadhwani (1992) suggested the following ad hoc empirical specification of Equation 2.3.5:

$$r_t = \alpha + \theta\sigma_t^2 + (\psi_0 + \psi_1\sigma_t^2)r_{t-1} + \varepsilon_t, \quad (2.3.6)$$

Equation 2.3.6 distinguishes between the part of autocorrelation due to market inefficiencies (denote by  $\psi_0$ ) and that due to feedback trading (denote by  $\psi_1$ ). With  $\psi_1 = -\theta\gamma$ , significantly positive (negative) values for  $\psi_1$  will denote the presence of negative (positive) feedback trading.

To asses the interaction of feedback trading with the observed premiums/discounts of the sample ETFs, this study employs Chau et al. (2011)'s empirical extension of the Sentana and Wadhwani (1992) model:

$$\begin{aligned} r_t = & \alpha_0 D_{t-1} + \alpha_1 (1 - D_{t-1}) + \theta_0 D_{t-1} \sigma_t^2 + \theta_1 (1 - D_{t-1}) \sigma_t^2 + D_{t-1} (\psi_{0,0} + \psi_{1,0} \sigma_t^2) r_{t-1} \\ & + (1 - D_{t-1}) (\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t, \end{aligned} \quad (2.3.7)$$

The term " $D_{t-1}$ " in Equation 2.3.7 is a dummy variable assuming the value of unity if the ETF has posted a discount in period  $t - 1$ , zero otherwise<sup>2</sup>. Equation 2.3.7 allows all terms of Equation 2.3.6 to shift with the observed lagged premiums/discounts of the ETF and permits us to gauge how feedback trading manifests itself when the ETF's price exhibits a positive (the case of a premium)

---

<sup>2</sup>Given the daily frequency of our data, both the closing prices and NAVs employed are day-end observations; as a result, it is not possible for the feedback trader of Equation 2.3.7 to trade on the contemporaneous (period  $t$ ) premium/discount, since he cannot observe it until the session is over (this would have been the case only if we were working on the premises of real-time data), hence we rely on lagged premiums/discounts. This is further supported by the fact that feedback traders in the Sentana and Wadhwani (1992) framework base their trades on the previous day's returns, not the contemporaneous ones (which, given the daily frequency of our data could not be traded on anyway, since that are day-end ones)

or negative (the case of a discount) deviation from its NAV in period  $t - 1$ <sup>3</sup>.

The conditional variance  $\sigma_t^2$  in all of above equations follows a GJR-GARCH (1,1) Glosten et al. (1993) process as:

$$\sigma_t^2 = \omega + \beta\varepsilon_{t-1}^2 + \lambda\sigma_{t-1}^2 + \delta I_{t-1}\varepsilon_{t-1}^2, \quad (2.3.8)$$

In Equation 2.3.8, the parameter  $\delta$  reveals whether volatility responds asymmetrically to positive versus negative shocks.  $I_{t-1}$  is a dummy variable, assuming the value of unity if the lagged shock is negative, zero otherwise; significantly positive estimates for the  $\delta$  denote that volatility is higher following negative (compared to positive) shocks.

Given country ETFs' documented wide premium and discounts, it is possible that feedback traders condition their feedback trading on forecast premiums/discounts when trading these ETFs. To explore this possibility, it assesses the interaction between forecast premiums/discounts and feedback trading, by first assuming that the dynamics of ETFs' percentage price deviations from their NAV follow a standard Ornstein-Uhlenbeck (OU) process<sup>4</sup>, as follows:

$$dX_t = -\rho(X_t - \mu)dt + \xi dW_t, \quad (2.3.9)$$

$X_t$  represents the percentage price deviation of the ETF from its NAV,  $\rho$  is the speed of mean reversion,  $W_t$  is a standard Brownian motion (on some probability space), and  $\mu$  is the long term<sup>5</sup> equilibrium level of the ETF's percentage price deviation from its NAV. The solution of equation 2.3.9 is provided by

$$X_{i+1} = X_i e^{-\rho t} + \mu(1 - e^{-\rho t}) + \xi \sqrt{\frac{1 - e^{-\rho t}}{2\rho}} N_{0,1}, \quad (2.3.10)$$

---

<sup>3</sup>Equation 2.3.7 combines the possibility of NAV-deviations interacting with feedback trading both additively and multiplicatively. As Chau et al. (2011) showed, the additive version of this interaction assumes the following feedback trading function:  $Y_t = \gamma r_{t-1} + \kappa D_t$  in which case the combined function of rational and feedback traders becomes:  $r_t = \alpha_0 D_t + \alpha_1 (1 - D_t) + \theta_0 D_t \sigma_t^2 + \theta_1 (1 - D_t) \sigma_t^2 + (\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t$  The multiplicative version of this interaction assumes the following feedback trading function:  $Y_t = [\gamma D_t + \kappa (1 - D_t)] r_{t-1}$ , in which case the combined function of rational and feedback traders assumes the following form:  $r_t = \alpha + \theta \sigma_t^2 + D_t (\psi_{0,0} + \psi_{1,0} \sigma_t^2) r_{t-1} + (1 - D_t) (\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t$

<sup>4</sup>Applications of the OU-process in finance include Bormetti et al (2010); Griffin (2010)

<sup>5</sup>The long term equilibrium is equivalent here to a window of 252 days (i.e. a year's observations)

where  $t$  denotes the fixed time steps and  $N_{0,1}$  is the standard normal distributions. Then the parameters are estimated using maximum likelihood method, with the conditional probability density function derived as follows:

$$f(X_{i+1}|X_i; \mu, \rho, \xi) = \frac{1}{\sqrt{2\pi\hat{\xi}^2}} \exp \left( -\frac{(x_i - x_{i-1}e^{-\rho t} - \mu(1 - e^{-\rho t}))^2}{2\hat{\xi}^2} \right) \quad (2.3.11)$$

with  $\hat{\xi}^2 = \xi^2 \frac{1-e^{-2\rho t}}{2\rho}$ . The log-likelihood function of a set of observations  $(X_0, X_1, \dots, X_n)$  can be derived as:

$$\begin{aligned} \mathcal{L}(\mu, \rho, \hat{\xi}) &= \sum_{i=1}^n \ln f(X_{i+1}|X_i; \mu, \rho, \hat{\xi}) = -\frac{n}{2} \ln(2\pi) - n \ln(\hat{\xi}^2) - \\ &\quad \frac{1}{2\hat{\xi}^2} \sum_{i=1}^n (X_i - X_{i-1}e^{-\rho t} - \mu(1 - e^{-\rho t}))^2, \end{aligned} \quad (2.3.12)$$

Algebraically, the following equations are derived from the above:

$$\mu = \frac{\sum_{i=1}^n (X_i - X_{i-1}e^{-\rho t})}{n(1 - e^{-\rho t})} \quad (2.3.13)$$

$$\rho = -\frac{1}{t} \ln \frac{\sum_{i=1}^n (X_i - \mu)(X_{i-1} - \mu)}{\sum_{i=1}^n (X_{i-1} - \mu)^2}, \quad (2.3.14)$$

$$\hat{\xi}^2 = \frac{1}{n} \sum_{i=1}^n [(X_i - \mu - e^{-\rho t})(X_{i-1} - \mu)]^2. \quad (2.3.15)$$

To gauge whether the forecast premiums/discounts generated from the OU-process affect feedback trading in the sample ETFs, we employ equation 2.3.16, a close variant of Equation 2.3.7:

$$\begin{aligned} r_t &= \alpha_0 D_t + \alpha_1 (1 - D_t) + \theta_0 D_t \sigma_t^2 + \theta_1 (1 - D_t) \sigma_t^2 + D_t (\psi_{0,0} + \psi_{1,0} \sigma_t^2) r_{t-1} \\ &\quad + (1 - D_t) (\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t, \end{aligned} \quad (2.3.16)$$

In the above equation, " $D_t$ " is equal to one, if the OU-process forecasts a discount for the day  $t$ , zero otherwise. Finally, to test whether these findings hold in view of the outbreak of the 2008 global financial crisis, the sample period is split into a pre (June 20th, 2000 - August 31st, 2008) and a post

(September 1st, 2008<sup>6</sup> - April 27th, 2016) crisis-outbreak period and repeat all of the estimations for both sub periods. This partition in our sample window is motivated mainly by the fact that the outbreak of financial crises has been found (Choe et al. (1999); Kim and Wei (2002a); Charteris et al. (2014)) to produce changes in the market’s directional trend, from upward to downward; with feedback traders extrapolating from historical price trends, any such groundbreaking shift in the market’s direction is bound to affect their trading pattern. Additionally, the fact that crisis-periods encompass extreme price movements is expected to lead to wilder swings of country ETFs’ prices and, quite possibly, amplify their deviations from their NAVs, given that arbitrage is harder to practice during extreme market periods.

## 2.4 Results - Discussion

The discussion is began with the presentation of the results from equation 2.3.6, i.e., the original Sentana and Wadhwani (1992) model. The estimates outlined in Table 2.2 indicate that several US-listed country ETFs exhibit inefficiencies in their returns, as demonstrated by the significantly<sup>7</sup> negative (positive) values of the first-order autocorrelation coefficient  $\psi_0$  for eight (two) ETFs.  $\psi_1$  assumes significantly negative values for six ETFs (iShares MSCI Australia ETF; iShares MSCI Hong Kong ETF; iShares MSCI Japan ETF; iShares MSCI Malaysia ETF; iShares MSCI Singapore ETF; iShares MSCI Taiwan ETF), denoting the presence of positive feedback trading in their dynamics. Although the above indicate that the majority (fourteen) of the sample’s US-listed country ETFs accommodates no feedback trading, it is worth noting that the above six ETFs

---

<sup>6</sup>The choice of September 2008 as the cut-off point in our sample window is motivated by the groundbreaking events that took place in the US during that month (including the bankruptcy of Lehman Brothers and the US government’s decision to place mortgage providers Freddie Mac and Fannie May into conservatorship) and which changed the landscape of the US financial system in the post 2008 years

<sup>7</sup>In the interest of brevity, any reference to statistical significance in this section shall pertain to estimated coefficients, whose p-value are less than 0.1

are all targeting markets in the Asia Pacific region, thus suggesting that feedback trading is more pronounced for country ETFs investing in markets whose trading times do not overlap with those of the US at all.<sup>8</sup> A key issue regarding several Asia Pacific markets is that retail investors command a substantial fraction of their turnover (Chou et al. (2011)), thus amplifying noise trading (Barber et al. (2007, 2009); Kuo et al. (2015)). As a result, the feedback trading documented for country ETFs targeting markets in that region may be due either to these ETFs mirroring these markets' performance or to these ETFs' investors opting for feedback trading as a rational strategy given the noise levels of these markets.<sup>9</sup> The large time difference between the Asia Pacific region and the US further facilitates feedback trading in those ETFs, since US investors trading them will be aware of their underlying benchmarks' NAV for the day well before trading in the US had started, possibly choosing to use these NAVs as reference points. As per the structure of their conditional variance, the significant (at the 1 percent level)  $\lambda$  values indicate that contemporaneous volatility is significantly related to lagged volatility, thus denoting its persistence. The volatility of most (seventeen) ETFs responds significantly to news (as the significant  $\beta$  values indicate), with this response being asymmetric in all cases, since the coefficient  $\delta$  is always significantly positive. Overall, the structure of the conditional variance of this study's ETFs reflects similar properties to that reported in prior studies on ETFs' feedback trading (Chau et al. (2011); Charteris et al. (2014)).

---

<sup>8</sup>The solo ETF in our sample focusing on an Asia Pacific market and not exhibiting any feedback trading is the iShares MSCI South Korea Capped ETF.

<sup>9</sup>The selection of feedback trading as a strategy by investors of country ETFs targeting Asia Pacific markets in this case can be motivated either by rational speculative reasons (to exploit the noise trading patterns in Asia Pacific markets' equity returns via those ETFs) or informational reasons (noise trading renders the public pool of information poorer and feedback trading has been shown - Brennan and Cao (1997) to be an option when trading in markets with informational uncertainty)

Table 2.2: Maximum likelihood estimation from the original Sentana and Wadhwani (1992) model

	AU	AT	BE	BR	CA	FR	DE	HK	IT	JP
$\alpha$	0.0085 (0.07926)	0.0483 (0.1606)	0.0203 (0.4629)	-0.0058 (0.9042)	0.0351 (0.1811)	0.0053 (0.8531)	0.0309 (0.2976)	0.0350 (0.2471)	0.0172 (0.5990)	-0.0157 (0.6380)
$\theta$	0.0047 (0.7292)	-0.0050 (0.7149)	0.0044 (0.7963)	0.0024 (0.8088)	-0.0015 (0.9304)	0.0062 (0.6266)	0.0005 (0.9632)	-0.0019 (0.8988)	-0.0043 (0.7321)	0.0179 (0.3784)
$\psi_0$	-0.0079 (0.7131)	0.0188 (0.3821)	-0.0563 (0.0040)	0.0447 (0.0398)	0.0168 (0.4332)	-0.0728 (0.0011)	-0.0442 (0.0434)	-0.0201 (0.3402)	-0.1014 (0.0000)	-0.0369 (0.1084)
$\psi_1$	-0.0067 (0.0411)	-0.0040 (0.1850)	0.0010 (0.7977)	-0.0028 (0.1569)	-0.0033 (0.5052)	-0.0007 (0.8610)	-0.0027 (0.5269)	-0.0109 (0.0054)	0.0008 (0.8387)	-0.0120 (0.0551)
$\omega$	0.0355 (0.0000)	0.0363 (0.0000)	0.0370 (0.0000)	0.0681 (0.0000)	0.0159 (0.0000)	0.0293 (0.0000)	0.0340 (0.0000)	0.0310 (0.0000)	0.0273 (0.0000)	0.0578 (0.0000)
$\beta$	0.0198 (0.0144)	0.0082 (0.1047)	0.0211 (0.0031)	0.0122 (0.0382)	0.0232 (0.0045)	0.0083 (0.1523)	0.0134 (0.0040)	0.0278 (0.0000)	0.0248 (0.0000)	0.0436 (0.0001)
$\lambda$	0.9252 (0.0000)	0.9385 (0.0000)	0.9075 (0.0000)	0.9290 (0.0000)	0.9290 (0.0000)	0.9274 (0.0000)	0.9251 (0.0000)	0.9100 (0.0000)	0.9255 (0.0000)	0.8815 (0.0000)
$\delta$	0.0753 (0.0000)	0.0703 (0.0000)	0.1040 (0.0000)	0.0881 (0.0000)	0.0738 (0.0000)	0.1006 (0.0000)	0.0907 (0.0000)	0.0730 (0.0000)	0.0792 (0.0000)	0.0819 (0.0000)
	ML	MX	NL	SG	SK	SP	SW	CH	TW	UK
$\alpha$	0.0151 (0.6672)	0.0144 (0.6490)	0.0194 (0.4895)	-0.0070 (0.7957)	-0.0014 (0.9685)	0.0348 (0.2981)	-0.0116 (0.7297)	0.0315 (0.2357)	0.0291 (0.3886)	0.0212 (0.6543)
$\theta$	0.0096 (0.6539)	0.0030 (0.8086)	-0.0005 (0.9693)	0.0082 (0.5457)	0.0082 (0.4100)	-0.0046 (0.7108)	0.0117 (0.2157)	0.0008 (0.9640)	-0.0046 (0.7011)	0.0043 (0.6548)
$\psi_0$	-0.0020 (0.9407)	0.0430 (0.0341)	-0.0398 (0.0560)	-0.0464 (0.0278)	0.0227 (0.3115)	-0.0550 (0.0138)	-0.0256 (0.2327)	-0.0741 (0.0009)	-0.0039 (0.8674)	-0.0573 (0.1004)
$\psi_1$	-0.0321 (0.0001)	-0.0048 (0.1374)	-0.0048 (0.2616)	-0.0083 (0.0485)	-0.0028 (0.2469)	-0.0020 (0.6028)	-0.0047 (0.1008)	-0.0069 (0.2945)	-0.0101 (0.0079)	-0.0016 (0.6592)
$\omega$	0.0148 (0.0000)	0.0418 (0.0000)	0.0276 (0.0000)	0.0158 (0.0000)	0.0252 (0.0000)	0.0320 (0.0000)	0.0403 (0.0000)	0.0250 (0.0000)	0.0197 (0.0000)	0.0352 (0.0000)
$\beta$	0.0289 (0.0000)	0.0014 (0.7802)	0.0130 (0.0527)	0.0235 (0.0001)	0.0283 (0.0001)	0.0114 (0.0394)	0.0115 (0.1240)	0.0275 (0.0000)	0.0246 (0.0002)	0.0184 (0.0005)
$\lambda$	0.9509 (0.0000)	0.9209 (0.0000)	0.9229 (0.0000)	0.9344 (0.0000)	0.9376 (0.0000)	0.9283 (0.0000)	0.9288 (0.0000)	0.9163 (0.0000)	0.9392 (0.0000)	0.9842 (0.0000)
$\delta$	0.0249 (0.0000)	0.1245 (0.0000)	0.1004 (0.0000)	0.0704 (0.0000)	0.0555 (0.0000)	0.0979 (0.0000)	0.0896 (0.0000)	0.0780 (0.0000)	0.0603 (0.0000)	0.0912 (0.0000)

<sup>1</sup>The table presents the estimates from the set of the following equations for the full sample period (20/6/2000 - 27/4/2016):

$$r_t = \alpha + \theta \sigma_t^2 + (\psi_0 + \psi_1 \sigma_t^2) r_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = \omega + \beta \varepsilon_{t-1}^2 + \lambda \sigma_{t-1}^2 + \delta I_{t-1} \varepsilon_{t-1}^2$$

ETFs appear in the table with the following abbreviations: AU (iShares MSCI Australia ETF), AT (iShares MSCI Austria Capped ETF), BE (iShares MSCI Belgium Capped ETF), BR (iShares MSCI Brazil Capped ETF), CA (iShares MSCI Canada ETF), FR (iShares MSCI France ETF), DE (iShares MSCI Germany ETF), HK (iShares MSCI Hong Kong ETF), IT (iShares MSCI Italy Capped ETF), JP (iShares MSCI Japan ETF), ML (iShares MSCI Malaysia ETF), MX (iShares MSCI Mexico Capped ETF), NL (iShares MSCI Netherlands ETF), SG (iShares MSCI Singapore ETF), SK (iShares MSCI South Korea Capped ETF), SP (iShares MSCI Spain Capped ETF), SW (iShares MSCI Sweden ETF), CH (iShares MSCI Switzerland ETF), TW (iShares MSCI Taiwan ETF), UK (iShares MSCI United Kingdom ETF). Parentheses include p-values.

We now turn to assessing whether feedback trading varies in its presence in US-listed country ETFs with the sign of the observed lagged percentage price deviation of an ETF from its NAV. Table 3 presents the results from the estimation of equation 2.3.7 for the sample's ETFs. The estimations, overall, reveal a rather limited presence of feedback trading contingent upon the realization of a lagged premium or discount. More specifically, two ETFs exhibit positive feedback trading when a premium is observed on the previous day, while the iShares MSCI Belgium Capped ETF (iShares



MSCI Malaysia ETF) exhibits negative (positive) feedback trading in the presence of a lagged discount. Again here, it is interesting to note that three of those four ETFs are targeting Asia Pacific markets, with some of these ETFs exhibiting some of the largest average premiums/discounts for the sample period. Several ETFs exhibit inefficiencies in their return-generating process irrespective of the presence of lagged premiums or discounts, as their significant  $\psi_{0,0}$  and  $\psi_{0,1}$  values indicate. When a discount has materialized on the previous day,  $\psi_{0,0}$  is significantly positive (negative) for one (two) ETFs; conversely, the realization of a lagged premium is found to be associated with some cases of significant first-order autocorrelation, as it yields significantly positive (negative)  $\psi_{0,1}$  values for five (two) ETFs. Regarding the volatility's structure, it appears highly persistent and asymmetric for all twenty ETFs, in line with results reported previously.

Table 2.4 presents the estimates from equation 2.3.16, controlling for the presence of a predicted (as opposed to observed) premium or discount. Significantly positive feedback trading exists for predicted premiums for three ETFs (iShares MSCI Hong Kong ETF, iShares MSCI Singapore ETF; iShares MSCI Taiwan ETF) and for predicted discounts for the iShares MSCI Malaysia ETF and the iShares MSCI Taiwan ETF; significant negative feedback trading is reported for the iShares MSCI France ETF and the iShares MSCI Spain Capped ETF for predicted discounts. Once more, the results show that feedback trading tends to be more prevalent among country ETFs investing in Asia Pacific markets, with evidence on its presence among country ETFs targeting European markets being limited. The majority of ETFs exhibit significant first-order autocorrelation for predicted discounts, with several of them doing so for predicted premiums as well, thus confirming the presence of widespread inefficiencies in their returns' structure. Again here, volatility appears highly persistent and asymmetric across all twenty ETFs.

Table 2.3: Maximum likelihood estimation from the Sentana and Wadhwani (1992) model controlling for observed premiums/discounts

	AU	AT	BE	BR	CA	FR	DE	HK	IT	JP
$\alpha_0$	0.0628 (0.2661)	0.1554 (0.0021)	0.0723 (0.0924)	0.1500 (0.0638)	0.0972 (0.0204)	0.0540 (0.2307)	0.0649 (0.1982)	0.1205 (0.0278)	0.1556 (0.0046)	0.0109 (0.8459)
$\alpha_1$	-0.0186 (0.6649)	-0.0311 (0.5743)	0.0108 (0.7898)	-0.1404 (0.0262)	0.0207 (0.5381)	-0.0442 (0.2820)	0.0073 (0.8646)	-0.0098 (0.8131)	-0.0780 (0.0817)	-0.0076 (0.8876)
$\theta_0$	-0.0042 (0.8591)	0.0094 (0.6061)	0.0560 (0.0051)	0.0040 (0.7982)	0.0017 (0.9472)	0.0138 (0.5154)	-0.0072 (0.4024)	0.0144 (0.6008)	0.0188 (0.3541)	0.0419 (0.1998)
$\theta_1$	0.0082 (0.6878)	-0.0256 (0.3302)	-0.0505 (0.0398)	-0.0043 (0.7519)	-0.0060 (0.7861)	0.0044 (0.8254)	-0.0072 (0.7166)	-0.0191 (0.3993)	-0.0232 (0.2186)	-0.0214 (0.5605)
$\psi_{0,0}$	0.0116 (0.7536)	0.0279 (0.3960)	-0.0249 (0.4559)	0.0312 (0.3633)	0.0025 (0.9403)	-0.0814 (0.0210)	-0.0170 (0.6342)	0.0280 (0.4355)	-0.0531 (0.1672)	-0.0122 (0.7363)
$\psi_{1,0}$	-0.0086 (0.1250)	-0.0008 (0.8188)	0.0082 (0.0929)	-0.0024 (0.4460)	0.0017 (0.8416)	-0.0359 (0.6537)	-0.0001 (0.9829)	-0.0098 (0.1804)	0.0027 (0.6727)	-0.0097 (0.2978)
$\psi_{0,1}$	-0.0006 (0.9844)	0.0559 (0.0911)	-0.0355 (0.2433)	0.0871 (0.0028)	0.0592 (0.0416)	-0.0359 (0.2539)	-0.0053 (0.1731)	0.0054 (0.8586)	-0.0602 (0.0878)	-0.0255 (0.5160)
$\psi_{1,1}$	-0.0073 (0.1402)	-0.0057 (0.4820)	0.0020 (0.8072)	-0.0030 (0.3239)	-0.0085 (0.1995)	-0.0056 (0.4230)	-0.0053 (0.4598)	-0.0109 (0.0589)	-0.0018 (0.7903)	-0.0068 (0.6544)
$\omega$	0.0365 (0.0000)	0.0427 (0.0000)	0.0360 (0.0000)	0.0620 (0.0000)	0.0154 (0.0000)	0.0296 (0.0000)	0.0347 (0.0000)	-0.0109 (0.0000)	0.0269 (0.0000)	0.0596 (0.0000)
$\beta$	0.0201 (0.1351)	0.0179 (0.0019)	0.0195 (0.1351)	0.0145 (0.0058)	0.0253 (0.0019)	0.0082 (0.1758)	0.0133 (0.0438)	0.0297 (0.0000)	0.0234 (0.0001)	0.0426 (0.0000)
$\lambda$	0.9250 (0.0000)	0.9280 (0.0000)	0.9108 (0.0000)	0.9332 (0.0000)	0.9292 (0.0000)	0.9270 (0.0000)	0.9254 (0.0000)	0.9174 (0.0000)	0.9270 (0.0000)	0.8792 (0.0000)
$\delta$	0.0751 (0.0000)	0.0684 (0.0000)	0.0751 (0.0000)	0.0802 (0.0000)	0.0698 (0.0000)	0.1008 (0.0000)	0.0913 (0.0000)	0.0734 (0.0000)	0.0787 (0.0000)	0.0860 (0.0000)
	ML	MX	NL	SG	SK	SP	SW	CH	TW	UK
$\alpha_0$	0.1071 (0.0348)	0.0646 (0.1691)	0.0903 (0.0416)	-0.0220 (0.6532)	-0.1115 (0.0300)	0.1439 (0.0173)	0.0036 (0.9513)	0.0411 (0.3735)	0.0263 (0.6135)	0.0658 (0.7654)
$\alpha_1$	-0.0566 (0.2335)	-0.0265 (0.6250)	-0.0122 (0.7642)	0.0085 (0.8182)	0.0391 (0.4640)	-0.0135 (0.7631)	0.0393 (0.0372)	0.0628 (0.0960)	-0.0130 (0.7860)	0.0098 (0.8539)
$\theta_0$	0.0104 (0.7196)	0.0339 (0.0684)	0.0192 (0.3642)	0.0280 (0.2261)	0.0420 (0.0020)	0.0063 (0.7691)	-0.0123 (0.4534)	0.0749 (0.0123)	0.0259 (0.1613)	-0.0089 (0.7643)
$\theta_1$	-0.0275 (0.3588)	0.0440 (0.0662)	-0.0219 (0.3411)	-0.0178 (0.4074)	-0.0152 (0.3155)	-0.0276 (0.1450)	-0.0043 (0.7519)	-0.0621 (0.0345)	-0.0338 (0.0415)	-0.0067 (0.7654)
$\psi_{0,0}$	0.0189 (0.6022)	0.0561 (0.0503)	-0.0068 (0.8423)	-0.0500 (0.1568)	-0.0543 (0.0896)	-0.0018 (0.5074)	-0.0198 (0.5788)	-0.0161 (0.6934)	-0.0139 (0.7013)	-0.0124 (0.7634)
$\psi_{1,0}$	-0.0213 (0.0298)	-0.0013 (0.7648)	-0.0024 (0.6839)	-0.0076 (0.2613)	0.0029 (0.3442)	-0.0018 (0.7475)	-0.0022 (0.6128)	-0.0027 (0.8116)	-0.0056 (0.3219)	-0.0076 (0.4562)
$\psi_{0,1}$	0.0023 (0.9521)	0.0884 (0.0110)	-0.0244 (0.4215)	-0.0326 (0.3101)	0.0026 (0.9409)	-0.0288 (0.3851)	-0.0079 (0.7981)	-0.0793 (0.0170)	0.0581 (0.0878)	-0.0073 (0.1348)
$\psi_{1,1}$	-0.0023 (0.8627)	-0.0026 (0.7457)	-0.0053 (0.4823)	-0.0057 (0.4394)	-0.0034 (0.4772)	-0.0008 (0.9065)	-0.0041 (0.3599)	-0.0030 (0.8170)	-0.0106 (0.0685)	-0.0034 (0.4265)
$\omega$	0.0202 (0.0000)	0.0422 (0.0000)	0.0268 (0.0000)	0.0161 (0.0000)	0.0224 (0.0000)	0.0317 (0.0000)	0.0411 (0.0000)	0.0249 (0.0000)	0.0198 (0.0000)	0.0316 (0.0000)
$\beta$	0.0347 (0.0000)	0.0022 (0.6819)	0.0134 (0.0499)	0.0243 (0.0001)	0.0179 (0.0003)	0.0124 (0.0000)	0.0138 (0.0069)	0.0285 (0.0059)	0.0252 (0.0002)	0.0045 (0.0143)
$\lambda$	0.9397 (0.0000)	0.9208 (0.0000)	0.9237 (0.0000)	0.9338 (0.0000)	0.9458 (0.0000)	0.9281 (0.0000)	0.9272 (0.0000)	0.9158 (0.0000)	0.9391 (0.0000)	0.9321 (0.0000)
$\delta$	0.0314 (0.0000)	0.1221 (0.0000)	0.0986 (0.0000)	0.0696 (0.0000)	0.0603 (0.0000)	0.0968 (0.0000)	0.0879 (0.0000)	0.0773 (0.0000)	0.0604 (0.0000)	0.0527 (0.0000)

<sup>1</sup>The table presents the estimates from the set of the following equations for the full sample period (20/6/2000 - 27/4/2016):

$$\begin{aligned}
 r_t &= \alpha_0 D_{t-1} + \alpha_1 (1 - D_{t-1}) + \theta_0 D_{t-1} \sigma_t^2 + \theta_1 (1 - D_{t-1}) \sigma_t^2 + D_{t-1} (\psi_{0,0} + \psi_{1,0} \sigma_t^2) r_{t-1} + (1 - D_{t-1}) (\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t, \\
 \sigma_t^2 &= \omega + \beta \varepsilon_{t-1}^2 + \lambda \sigma_{t-1}^2 + \delta I_{t-1} \varepsilon_{t-1}^2
 \end{aligned}$$

$D_t$  is a dummy variable assuming the value of unity if the ETF has posted a discount in period  $t - 1$ , zero otherwise. ETFs appear in the table with the following abbreviations: AU (iShares MSCI Australia ETF), AT (iShares MSCI Austria Capped ETF), BE (iShares MSCI Belgium Capped ETF), BR (iShares MSCI Brazil Capped ETF), CA (iShares MSCI Canada ETF), FR (iShares MSCI France ETF), DE (iShares MSCI Germany ETF), HK (iShares MSCI Hong Kong ETF), IT (iShares MSCI Italy Capped ETF), JP (iShares MSCI Japan ETF), ML (iShares MSCI Malaysia ETF), MX (iShares MSCI Mexico Capped ETF), NL (iShares MSCI Netherlands ETF), SG (iShares MSCI Singapore ETF), SK (iShares MSCI South Korea Capped ETF), SP (iShares MSCI Spain Capped ETF), SW (iShares MSCI Sweden ETF), CH (iShares MSCI Switzerland ETF), TW (iShares MSCI Taiwan ETF), UK (iShares MSCI United Kingdom ETF). Parentheses include p-values.

Table 2.4: Maximum likelihood estimation from the Sentana and Wadhwani (1992) model controlling for forecast premiums/discounts

	AU	AT	BE	BR	CA	FR	DE	HK	IT	JP
$\alpha_0$	-0.3744 (0.0000)	-0.3312 (0.0000)	-0.2536 (0.0000)	-0.0613 (0.3823)	-0.1928 (0.0000)	-0.2856 (0.0000)	-0.3288 (0.0000)	-0.3196 (0.0000)	-0.3390 (0.0000)	-0.4218 (0.0000)
$\alpha_1$	0.1146 (0.0042)	0.2537 (0.0000)	0.1205 (0.0010)	0.0297 (0.6605)	0.2189 (0.0000)	0.1580 (0.0000)	0.2449 (0.0000)	0.2459 (0.0000)	0.2631 (0.0000)	0.1625 (0.0019)
$\theta_0$	-0.2115 (0.0000)	-0.0761 (0.0049)	-0.1077 (0.0000)	0.0015 (0.9164)	-0.0758 (0.0030)	-0.0874 (0.0003)	-0.0537 (0.0049)	-0.2215 (0.0000)	-0.0883 (0.0003)	-0.1276 (0.0001)
$\theta_1$	0.2429 (0.0000)	0.1167 (0.0000)	0.1647 (0.0000)	0.0050 (0.7415)	0.0591 (0.0150)	0.0249 (0.0000)	0.0747 (0.0006)	0.2135 (0.0000)	0.0790 (0.0000)	0.2501 (0.0000)
$\psi_{0,0}$	-0.0403 (0.1669)	-0.0769 (0.0025)	-0.0844 (0.0022)	0.0373 (0.2715)	0.0302 (0.2913)	-0.1549 (0.0000)	-0.0892 (0.0018)	-0.1093 (0.0002)	-0.1687 (0.0000)	-0.1179 (0.0001)
$\psi_{1,0}$	0.0010 (0.8875)	0.0052 (0.1961)	0.0064 (0.3121)	-0.0014 (0.6599)	0.0069 (0.3778)	0.0130 (0.0622)	0.0024 (0.7065)	-0.0021 (0.8283)	0.0083 (0.2249)	-0.0030 (0.7663)
$\psi_{0,1}$	-0.0607 (0.0290)	0.0076 (0.8091)	-0.0477 (0.1381)	0.0540 (0.0618)	-0.0321 (0.2606)	-0.0371 (0.2410)	-0.0472 (0.1462)	-0.0793 (0.0025)	-0.1049 (0.0007)	-0.0650 (0.0247)
$\psi_{1,1}$	-0.0108 (0.1166)	-0.0044 (0.5447)	-0.0035 (0.7039)	-0.0037 (0.1716)	0.0002 (0.9733)	-0.0014 (0.8567)	-0.0010 (0.8980)	-0.0162 (0.0320)	0.0051 (0.3988)	-0.0178 (0.1328)
$\omega$	0.0224 (0.0000)	0.0298 (0.0000)	0.0275 (0.0000)	0.0602 (0.0000)	0.0122 (0.0000)	0.0265 (0.0000)	0.2736 (0.0000)	0.0170 (0.0000)	0.0201 (0.0000)	0.0399 (0.0000)
$\beta$	0.0245 (0.0000)	0.0167 (0.0019)	0.0259 (0.0000)	0.0150 (0.0096)	0.0157 (0.0059)	0.0176 (0.0000)	0.0154 (0.0019)	0.0247 (0.0000)	0.0280 (0.0000)	0.0535 (0.0001)
$\lambda$	0.9382 (0.0000)	0.9350 (0.0000)	0.9186 (0.0000)	0.9314 (0.0000)	0.9370 (0.0000)	0.9228 (0.0000)	0.9262 (0.0000)	0.9382 (0.0000)	0.9279 (0.0000)	0.8942 (0.0000)
$\delta$	0.0483 (0.0000)	0.0668 (0.0000)	0.0804 (0.0000)	0.0827 (0.0000)	0.0760 (0.0000)	0.0924 (0.0000)	0.0896 (0.0000)	0.0536 (0.0000)	0.0727 (0.0000)	0.0476 (0.0000)
	ML	MX	NL	SG	SK	SP	SW	CH	TW	UK
$\alpha_0$	0.0008 (0.9800)	0.0372 (0.6775)	-0.3122 (0.0000)	-0.3625 (0.0000)	-0.0313 (0.5246)	-0.3783 (0.0000)	0.3342 (0.0000)	-0.2761 (0.0000)	0.0332 (0.4964)	-0.3167 (0.0000)
$\alpha_1$	-0.0161 (0.6169)	-0.0036 (0.9736)	0.1702 (0.0007)	0.1929 (0.0000)	0.0041 (0.9390)	0.2549 (0.0000)	0.1457 (0.0007)	0.1247 (0.0002)	0.0077 (0.8715)	0.1874 (0.0000)
$\theta_0$	0.0072 (0.5924)	-0.0009 (0.9654)	-0.0997 (0.0000)	-0.1521 (0.0000)	0.0067 (0.6023)	-0.0735 (0.0000)	-0.0883 (0.0000)	-0.1684 (0.0030)	-0.0041 (0.8097)	-0.0637 (0.0009)
$\theta_1$	-0.0025 (0.7950)	0.0133 (0.4547)	0.1391 (0.0000)	0.1752 (0.0000)	0.0145 (0.3996)	0.0958 (0.0000)	0.1220 (0.0000)	0.1790 (0.0000)	-0.0015 (0.9287)	0.0537 (0.0001)
$\psi_{0,0}$	-0.0178 (0.5970)	0.0372 (0.2100)	-0.0917 (0.0007)	-0.1256 (0.0000)	-0.0228 (0.3456)	-0.1895 (0.0000)	0.0977 (0.0001)	-0.0965 (0.0029)	-0.0007 (0.9839)	-0.1145 (0.0428)
$\psi_{1,0}$	-0.0241 (0.0162)	-0.0036 (0.4762)	0.0017 (0.7859)	-0.0040 (0.6163)	-0.0025 (0.3808)	0.0161 (0.0098)	0.0027 (0.6007)	-0.0045 (0.7131)	-0.0097 (0.0677)	0.0101 (0.2398)
$\psi_{0,1}$	-0.0113 (0.6849)	0.0558 (0.0697)	-0.0280 (0.3738)	-0.0872 (0.0023)	-0.0175 (0.6010)	-0.0383 (0.2143)	-0.0167 (0.5842)	-0.0458 (0.1777)	0.0098 (0.7755)	-0.0541 (0.3217)
$\psi_{1,1}$	-0.0016 (0.5990)	-0.0058 (0.2695)	-0.0034 (0.6900)	-0.0156 (0.0482)	-0.0039 (0.4439)	-0.0007 (0.9027)	-0.0066 (0.2282)	-0.0091 (0.5459)	-0.0110 (0.0508)	-0.0023 (0.3478)
$\omega$	0.0063 (0.0000)	0.0534 (0.0000)	0.0225 (0.0000)	0.0092 (0.0000)	0.0260 (0.0000)	0.0216 (0.0000)	0.0307 (0.0000)	0.0198 (0.0000)	0.0209 (0.0000)	0.0539 (0.0000)
$\beta$	0.0234 (0.0000)	0.0024 (0.1538)	0.0231 (0.0002)	0.0310 (0.0001)	0.0216 (0.0013)	0.0116 (0.0000)	0.0060 (0.0034)	0.0345 (0.0059)	0.0227 (0.0002)	0.0236 (0.0002)
$\lambda$	0.9611 (0.0000)	0.9088 (0.0000)	0.9212 (0.0000)	0.9400 (0.0000)	0.9401 (0.0000)	0.9356 (0.0000)	0.9384 (0.0000)	0.9218 (0.0000)	0.9382 (0.0000)	0.9349 (0.0000)
$\delta$	0.0351 (0.0000)	0.1456 (0.0000)	0.0878 (0.0000)	0.0513 (0.0000)	0.0612 (0.0000)	0.0894 (0.0000)	0.0859 (0.0000)	0.0568 (0.0000)	0.0653 (0.0000)	0.0328 (0.0000)

<sup>1</sup>The table presents the estimates from the set of the following equations for the full sample period (20/6/2000 - 27/4/2016):

$$r_t = \alpha_0 D_t + \alpha_1 (1 - D_t) + \theta_0 D_t \sigma_t^2 + \theta_1 (1 - D_t) \sigma_t^2 + D_t (\psi_{0,0} + \psi_{1,0} \sigma_t^2) r_{t-1} + (1 - D_t) (\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = \omega + \beta \varepsilon_{t-1}^2 + \lambda \sigma_{t-1}^2 + \delta I_{t-1} \varepsilon_{t-1}^2$$

$D_t$  is a dummy variable assuming the value of unity if a discount was forecast for the ETF for day  $t$ , zero otherwise. ETFs appear in the table with the following abbreviations: AU (iShares MSCI Australia ETF), AT (iShares MSCI Austria Capped ETF), BE (iShares MSCI Belgium Capped ETF), BR (iShares MSCI Brazil Capped ETF), CA (iShares MSCI Canada ETF), FR (iShares MSCI France ETF), DE (iShares MSCI Germany ETF), HK (iShares MSCI Hong Kong ETF), IT (iShares MSCI Italy Capped ETF), JP (iShares MSCI Japan ETF), ML (iShares MSCI Malaysia ETF), MX (iShares MSCI Mexico Capped ETF), NL (iShares MSCI Netherlands ETF), SG (iShares MSCI Singapore ETF), SK (iShares MSCI South Korea Capped ETF), SP (iShares MSCI Spain Capped ETF), SW (iShares MSCI Sweden ETF), CH (iShares MSCI Switzerland ETF), TW (iShares MSCI Taiwan ETF), UK (iShares MSCI United Kingdom ETF). Parentheses include p-values.

To test whether the results reported in Tables 2.2–2.4 hold when controlling for the outbreak of the 2008 financial crisis, all of the above tests prior to (20/6/2000 -31/8/2008) and after (1/9/2008 - 27/4/2016) the crisis' outbreak are repeated and it reports the results in Tables 2.5 – 2.7. Table 2.5 and 2.6 present the estimates from the original Sentana and Wadhwani (1992) model (equation

2.3.6) pre and post crisis' outbreak, respectively. As the results indicate, feedback trading appears scant within each of the two sub periods, significant positive feedback trading exists in only three ETFs (iShares MSCI Malaysia ETF; iShares MSCI Singapore ETF; iShares MSCI Sweden ETF) before and two (iShares MSCI Australia ETF; iShares MSCI Hong Kong ETF) after the crisis' outbreak. Again here, the presence of feedback trading in country ETFs targeting Asia Pacific markets is confirmed, with four of the above mentioned five ETFs investing in markets from that region. Overall, the varying presence of feedback trading prior to and after the events of September 2008 is in line with the results reported on emerging markets' ETFs by Charteris et al. (2014) and confirms prior evidence (Antoniou et al. (2005); Laopodis (2005); Schuppli and Bohl (2010); Chau and Deesomsak (2015)) on the sensitivity of feedback trading to periods characterized by different market conditions. The presence of autocorrelation in the ETFs' structure is also confirmed, yet, much like with feedback trading, surfaces less frequently within each sub period (for six ETFs pre and five post crisis); as for the structure of the ETFs' volatility, it remains highly persistent and asymmetric<sup>10</sup> during both sub periods. Table 2.7 presents the results on the feedback coefficients ( $\psi_{1,0}$ ;  $\psi_{1,1}$ ) from the tests conditioning feedback trading on the lagged/predicted discounts/premiums before and after the crisis' outbreak. Results suggest limited evidence in favour of feedback trading, the latter being mostly detected among country ETFs targeting Asia Pacific markets.

---

<sup>10</sup>No asymmetry is detected for the volatility of the iShares MSCI Malaysia ETF pre crisis, as  $\delta$  is found to be significantly negative there.

Table 2.5: Maximum likelihood estimation from the original Sentana and Wadhwani (1992) pre crisis' outbreak

	AU	AT	BE	BR	CA	FR	DE	HK	IT	JP
$\alpha$	0.0387 (0.5844)	0.1940 (0.0326)	0.0207 (0.6449)	0.0563 (0.6092)	0.1365 (0.0431)	0.0257 (0.0563)	0.1070 (0.0354)	0.0919 (0.1424)	0.0390 (0.4716)	-0.0023 (0.9642)
$\theta$	0.0189 (0.6437)	-0.0571 (0.3030)	0.0242 (0.3677)	0.0102 (0.6368)	-0.0392 (0.4502)	0.0212 (0.4708)	-0.0171 (0.5482)	-0.0190 (0.5353)	0.0091 (0.8223)	0.0146 (0.7401)
$\psi_0$	0.0122 (0.8265)	-0.0311 (0.8167)	-0.1004 (0.0040)	0.0893 (0.0774)	0.0459 (0.4316)	-0.0829 (0.0479)	-0.0951 (0.0174)	-0.0637 (0.2303)	-0.1132 (0.0138)	0.0262 (0.6709)
$\psi_1$	-0.0318 (0.1094)	-0.0258 (0.3251)	0.0058 (0.5903)	-0.0108 (0.1162)	-0.0422 (0.1884)	-0.0001 (0.9927)	0.0034 (0.7703)	-0.0244 (0.1602)	-0.0049 (0.8287)	-0.0413 (0.1740)
$\omega$	0.0470 (0.0000)	0.0760 (0.0000)	0.0545 (0.0000)	0.2273 (0.0000)	0.0504 (0.0000)	0.0237 (0.0000)	0.0291 (0.0000)	0.0190 (0.0041)	0.0319 (0.0000)	0.0209 (0.0050)
$\beta$	0.0085 (0.5653)	0.0136 (0.0093)	0.0019 (0.1248)	-0.0009 (0.9182)	0.0257 (0.0708)	-0.0009 (0.9250)	0.0205 (0.0040)	0.0302 (0.0003)	0.0131 (0.1884)	0.0462 (0.0001)
$\lambda$	0.9297 (0.0000)	0.9328 (0.0000)	0.8906 (0.0000)	0.9854 (0.0000)	0.8983 (0.0000)	0.9439 (0.0000)	0.9287 (0.0000)	0.9468 (0.0000)	0.9232 (0.0000)	0.9345 (0.0000)
$\delta$	0.0609 (0.0000)	0.0700 (0.0000)	0.1101 (0.0000)	0.1173 (0.0000)	0.0746 (0.0000)	0.0775 (0.0000)	0.0660 (0.0000)	0.0285 (0.0000)	0.0771 (0.0000)	0.0161 (0.0000)
	ML	MX	NL	SG	SK	SP	SW	CH	TW	UK
$\alpha$	0.0104 (0.8609)	0.0673 (0.3192)	0.0319 (0.4637)	0.0009 (0.9889)	0.1004 (0.3105)	0.0822 (0.1350)	0.0237 (0.6837)	0.0727 (0.1322)	0.0598 (0.3418)	0.0946 (0.2389)
$\theta$	0.0126 (0.6931)	-0.0001 (0.9980)	0.0045 (0.8713)	0.0275 (0.3851)	-0.0063 (0.8062)	-0.0008 (0.9822)	0.0151 (0.5173)	-0.0127 (-0.0127)	-0.0090 (0.5337)	-0.0058 (0.8547)
$\psi_0$	-0.0343 (0.3994)	0.0542 (0.2624)	0.0136 (0.2471)	0.0206 (0.7167)	-0.0267 (0.6626)	-0.0829 (0.5148)	-0.0372 (0.3673)	-0.1329 (0.0008)	-0.0410 (0.2670)	-0.0568 (0.5789)
$\psi_1$	-0.0268 (0.0371)	-0.0164 (0.2669)	-0.0068 (0.6059)	-0.0524 (0.0040)	-0.0070 (0.5705)	-0.0289 (0.6028)	-0.0251 (0.0093)	0.0018 (0.9090)	-0.0072 (0.1126)	-0.0158 (0.5479)
$\omega$	0.0080 (0.0000)	0.1101 (0.0000)	0.0229 (0.0000)	0.0335 (0.0000)	0.0583 (0.0000)	0.0273 (0.0000)	0.0355 (0.0000)	0.0244 (0.0000)	0.0398 (0.0000)	0.0457 (0.0000)
$\beta$	0.0320 (0.0000)	-0.0209 (0.0649)	0.0012 (0.8905)	0.0182 (0.0979)	0.0260 (0.0084)	0.0048 (0.2720)	-0.0005 (0.9643)	0.0241 (0.0595)	0.0365 (0.0002)	0.0180 (0.1690)
$\lambda$	0.9691 (0.0000)	0.8907 (0.0000)	0.9319 (0.0000)	0.9379 (0.0000)	0.9369 (0.0000)	0.9383 (0.0000)	0.9452 (0.0000)	0.9271 (0.0000)	0.9324 (0.0000)	0.9457 (0.0000)
$\delta$	-0.0076 (0.0623)	0.1749 (0.0000)	0.0876 (0.0000)	0.0558 (0.0000)	0.0458 (0.0000)	0.0759 (0.0000)	0.0780 (0.0000)	0.0594 (0.0011)	0.0494 (0.0000)	0.0568 (0.0000)

<sup>1</sup>The table presents the estimates from the set of the following equations pre crisis' outbreak (20/6/2000 - 31/8/2008):

$$r_t = \alpha + \theta\sigma_t^2 + (\psi_0 + \psi_1\sigma_t^2)r_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = \omega + \beta\varepsilon_{t-1}^2 + \lambda\sigma_{t-1}^2 + \delta I_{t-1}\varepsilon_{t-1}^2$$

Table 2.6: Maximum likelihood estimation from the original Sentana and Wadhwani (1992) model post crisis' outbreak

	AU	AT	BE	BR	CA	FR	DE	HK	IT	JP
$\alpha$	-0.0302 (0.4500)	-0.0373 (0.4113)	0.0127 (0.7252)	-0.0641 (0.2034)	0.0058 (0.8396)	-0.0357 (0.4040)	-0.0235 (0.5611)	0.0082 (0.8096)	-0.0973 (0.1214)	-0.0357 (0.3541)
$\theta$	0.0007 (0.9610)	0.0006 (0.9642)	-0.0042 (0.8044)	-0.0006 (0.9571)	-0.0041 (0.8077)	0.0076 (0.6010)	0.0069 (0.6161)	0.0047 (0.7780)	0.0141 (0.3693)	0.0232 (0.3232)
$\psi_0$	0.0082 (0.7692)	0.0285 (0.2999)	-0.0301 (0.2539)	0.0388 (0.1659)	0.0252 (0.3587)	-0.0519 (0.0891)	-0.0129 (0.6487)	0.0398 (0.1391)	-0.0651 (0.0690)	-0.506 (0.0897)
$\psi_1$	-0.0059 (0.0822)	-0.0027 (0.3755)	-0.0004 (0.9163)	-0.0016 (0.4544)	-0.0016 (0.7483)	-0.0017 (0.6979)	-0.0041 (0.3890)	-0.0110 (0.0092)	-0.0008 (0.8646)	-0.0101 (0.1332)
$\omega$	0.0273 (0.0000)	0.0344 (0.0000)	0.0247 (0.0000)	0.0343 (0.0003)	0.0079 (0.0022)	0.0453 (0.0000)	0.0416 (0.0000)	0.0378 (0.0000)	0.1083 (0.0000)	0.0842 (0.0000)
$\beta$	0.0206 (0.0338)	0.0275 (0.0010)	0.0244 (0.0065)	0.0114 (0.1449)	0.0028 (0.7569)	0.0237 (0.0074)	0.0104 (0.2155)	0.0263 (0.0003)	0.0399 (0.0000)	0.0324 (0.0003)
$\lambda$	0.9235 (0.0000)	0.9350 (0.0000)	0.9139 (0.0000)	0.9452 (0.0000)	0.9419 (0.0000)	0.9019 (0.0000)	0.9179 (0.0000)	0.9004 (0.0000)	0.8886 (0.0000)	0.8449 (0.0000)
$\delta$	0.0939 (0.0000)	0.0736 (0.0000)	0.1057 (0.0000)	0.0946 (0.0003)	0.0990 (0.0000)	0.1217 (0.0000)	0.1128 (0.0000)	0.1048 (0.0309)	0.0918 (0.0000)	0.1420 (0.0000)
	ML	MX	NL	SG	SK	SP	SW	CH	TW	UK
$\alpha$	-0.0003 (0.9918)	-0.0374 (0.2835)	0.0032 (0.9348)	-0.0252 (0.3704)	-0.0456 (0.2179)	-0.0955 (0.1018)	-0.0522 (0.2311)	.0098 (0.7603)	0.0432 (0.4636)	-0.0679 (0.1674)
$\theta$	0.0026 (0.9086)	0.0080 (0.5588)	-0.0003 (0.9855)	0.0010 (0.9484)	0.0085 (0.4426)	0.0130 (0.3886)	0.0121 (0.3196)	0.0046 (0.8299)	0.0326 (0.4045)	0.0057 (0.3789)
$\psi_0$	-0.0277 (0.4243)	0.0521 (0.0667)	-0.0264 (0.3291)	-0.0112 (0.6692)	-0.0019 (0.9472)	-0.0281 (0.3869)	-0.0311 (0.2803)	-0.0364 (0.2052)	0.0735 (0.0871)	-0.0179 (0.5439)
$\psi_1$	-0.0010 (0.9245)	-0.0018 (0.6501)	-0.0047 (0.3050)	-0.0069 (0.1304)	-0.0027 (0.2917)	-0.0019 (0.6400)	-0.0032 (0.2955)	-0.0088 (0.2177)	-0.0336 (0.1054)	-0.0018 (0.4279)
$\omega$	0.0224 (0.0000)	0.0271 (0.0000)	0.0324 (0.0000)	0.0095 (0.0000)	0.0219 (0.0000)	0.0842 (0.0000)	0.0459 (0.0000)	0.0245 (0.0000)	0.0392 (0.0001)	0.0569 (0.0000)
$\beta$	0.0144 (0.0963)	-0.0125 (0.0649)	0.0296 (0.0048)	0.0095 (0.1783)	0.0089 (0.3041)	0.0306 (0.0002)	-0.0357 (0.0034)	0.0323 (0.0007)	-0.0133 (0.0550)	0.0158 (0.0579)
$\lambda$	0.9285 (0.0000)	0.9313 (0.0000)	0.9048 (0.0000)	0.9407 (0.0000)	0.9399 (0.0000)	0.8961 (0.0000)	0.9041 (0.0000)	0.9075 (0.0000)	0.9349 (0.0000)	0.9459 (0.0000)
$\delta$	0.0927 (0.0000)	0.1404 (0.0000)	0.1089 (0.0000)	0.0925 (0.0000)	0.0860 (0.0000)	0.1098 (0.0000)	0.0966 (0.0000)	0.0915 (0.0000)	0.1047 (0.0000)	0.0935 (0.0000)

<sup>1</sup>The table presents the estimates from the set of the following equations post crisis' outbreak (1/9/2008 - 27/4/2016):

$$r_t = \alpha + \theta\sigma_t^2 + (\psi_0 + \psi_1\sigma_t^2)r_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = \omega + \beta\varepsilon_{t-1}^2 + \lambda\sigma_{t-1}^2 + \delta I_{t-1}\varepsilon_{t-1}^2$$

Table 2.7: **Estimates from the feedback coefficients of the extended Sentana and Wadhwani (1992) model accounting for actual (lagged) and forecast premiums/discounts**

	Pre crisis' outbreak				Post crisis' outbreak			
	$\psi_{1,0}$		$\psi_{1,1}$		$\psi_{1,0}$		$\psi_{1,1}$	
	Actual discount	Forecast discount	Actual premium	Forecast premium	Actual discount	Forecast discount	Actual premium	Forecast premium
iShares MSCI Australia ETF	-0.0173 (0.6247)	0.0732 (0.0913)	-0.0375 (0.3039)	-0.0095 (0.7747)	-0.0099 (0.0763)	0.0007 (0.9150)	-0.0079 (0.1161)	-0.0104 (0.1734)
iShares MSCI Austria Capped ETF	0.0322 (0.3661)	-0.0164 (0.7002)	-0.0884 (0.1802)	0.0113 (0.7600)	-0.0019 (0.5985)	0.0032 (0.4397)	-0.0052 (0.5391)	-0.0062 (0.4221)
iShares MSCI Belgium Capped ETF	0.0320 (0.0786)	0.0084 (0.5320)	0.0038 (0.8477)	0.0027 (0.9004)	0.0038 (0.4730)	0.0066 (0.4152)	-0.0005 (0.9565)	-0.0126 (0.2760)
iShares MSCI Brazil Capped ETF	-0.0099 (0.2805)	-0.0014 (0.8430)	0.0008 (0.9534)	-0.0186 (0.0761)	-0.0019 (0.5627)	-0.0003 (0.9348)	-0.0024 (0.4306)	-0.0031 (0.2904)
iShares MSCI Canada ETF	-0.0441 (0.4292)	0.0406 (0.4399)	-0.0490 (0.3146)	-0.0257 (0.5091)	0.0025 (0.7683)	0.0028 (0.7556)	-0.0066 (0.3458)	0.0017 (0.8175)
iShares MSCI France ETF	-0.0040 (0.8468)	0.0365 (0.0843)	-0.0079 (0.7337)	-0.0204 (0.4350)	0.0022 (0.7557)	0.0087 (0.2853)	-0.0086 (0.2699)	-0.0151 (0.1333)
iShares MSCI Germany ETF	-0.0110 (0.5305)	0.0142 (0.3944)	0.1137 (0.5509)	0.0078 (0.6203)	0.0031 (0.6586)	0.0011 (0.8753)	-0.0097 (0.1857)	-0.0051 (0.5943)
iShares MSCI Hong Kong ETF	-0.0412 (0.1501)	0.0011 (0.9715)	-0.0199 (0.4672)	0.0089 (0.7547)	-0.0070 (0.3783)	-0.0045 (0.6533)	-0.0117 (0.0611)	-0.0177 (0.0285)
iShares MSCI Italy Capped ETF	-0.0657 (0.1055)	0.0500 (0.1603)	0.0512 (0.1625)	0.0041 (0.9109)	0.0064 (0.3689)	0.0093 (0.2544)	-0.0052 (0.5188)	0.0010 (0.8738)
iShares MSCI Japan ETF	0.0064 (0.3689)	0.0093 (0.2544)	-0.0052 (0.5188)	0.0010 (0.8738)	-0.0109 (0.2634)	-0.0050 (0.5839)	-0.0044 (0.7966)	-0.0103 (0.2755)
iShares MSCI Malaysia ETF	0.0140 (0.4918)	-0.0007 (0.9649)	-0.0007 (0.9868)	-0.0338 (0.3025)	-0.0313 (0.0476)	-0.0284 (0.0807)	-0.0012 (0.9245)	-0.0010 (0.9401)
iShares MSCI Mexico Capped ETF	-0.0155 (0.4566)	-0.0168 (0.4325)	-0.0166 (0.5048)	-0.0246 (0.3663)	-0.0007 (0.8910)	0.0016 (0.7806)	-0.0049 (0.5819)	-0.0045 (0.4677)
iShares MSCI Netherlands ETF	-0.0150 (0.4641)	-0.0085 (0.6645)	-0.0026 (0.8848)	0.0048 (0.8104)	-0.0022 (0.7482)	0.0028 (0.7789)	-0.0082 (0.3536)	-0.0275 (0.4279)
iShares MSCI Singapore ETF	-0.0922 (0.0062)	-0.0507 (0.1356)	-0.0306 (0.2732)	-0.0695 (0.0419)	-0.0055 (0.4777)	-0.0047 (0.5735)	-0.0050 (0.5283)	-0.0143 (0.1070)
iShares MSCI South Korea Capped ETF	-0.0001 (0.9945)	-0.0046 (0.8027)	-0.0040 (0.7552)	-0.0205 (0.2719)	-0.0194 (0.3608)	-0.0022 (0.4798)	-0.0536 (0.0532)	-0.0029 (0.5559)
iShares MSCI Spain Capped ETF	-0.0414 (0.2734)	0.0081 (0.8141)	-0.0016 (0.9593)	-0.0338 (0.2021)	-0.0036 (0.6066)	0.0102 (0.1725)	-0.0091 (0.2374)	-0.0120 (0.1815)
iShares MSCI Sweden ETF	-0.0242 (0.1241)	-0.0300 (0.1335)	-0.0231 (0.0918)	-0.0227 (0.1060)	-0.0009 (0.8369)	0.0046 (0.3723)	-0.0021 (0.6628)	-0.0065 (0.2768)
iShares MSCI Switzerland Capped ETF	0.0085 (0.7133)	-0.0216 (0.3921)	0.0283 (0.3105)	-0.0062 (0.8295)	0.0006 (0.9626)	0.0023 (0.8669)	-0.0076 (0.5681)	-0.0230 (0.1639)
iShares MSCI Taiwan ETF	-0.0106 (0.4263)	-0.0171 (0.2220)	-0.0084 (0.4331)	-0.0038 (0.7465)	-0.0039 (0.5629)	-0.0064 (0.2826)	-0.0144 (0.0569)	-0.0119 (0.0642)
iShares MSCI United Kingdom ETF	-0.0210 (0.5987)	0.0246 (0.3589)	0.1269 (0.5985)	0.0097 (0.6298)	0.0028 (0.6035)	0.0046 (0.8309)	-0.0047 (0.1907)	-0.0068 (0.5097)

<sup>1</sup>The table presents the feedback coefficient estimates when feedback trading is conditioned upon actual lagged/forecast ( $\psi_{1,0}$ ) and premiums ( $\psi_{1,1}$ ) based on estimates from the following set of equations before(20/6/2000 - 31/8/2008) and after the crisis' outbreak (1/9/2008 - 27/4/2016):  
Actual(lagged) premiums/discounts:

$$\begin{aligned} r_t &= \alpha_0 D_{t-1} + \alpha_1 (1 - D_{t-1}) + \theta_0 D_{t-1} \sigma_t^2 + \theta_1 (1 - D_{t-1}) \sigma_t^2 + D_{t-1} (\psi_{0,0} + \psi_{1,0} \sigma_t^2) r_{t-1} + (1 - D_{t-1}) (\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t, \\ \sigma_t^2 &= \omega + \beta \varepsilon_{t-1}^2 + \lambda \sigma_{t-1}^2 + \delta I_{t-1} \varepsilon_{t-1}^2 \end{aligned}$$

Forecast premiums/discounts

$$\begin{aligned} r_t &= \alpha_0 D_t + \alpha_1 (1 - D_t) + \theta_0 D_t \sigma_t^2 + \theta_1 (1 - D_t) \sigma_t^2 + D_t (\psi_{0,0} + \psi_{1,0} \sigma_t^2) r_{t-1} + (1 - D_t) (\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t, \\ \sigma_t^2 &= \omega + \beta \varepsilon_{t-1}^2 + \lambda \sigma_{t-1}^2 + \delta I_{t-1} \varepsilon_{t-1}^2 \end{aligned}$$

Parentheses include p-values.

As an additional robustness test, it assessed the effect of various lagged (forecast) premium<sup>11</sup> and discount<sup>12</sup> levels over feedback trading by setting the variable  $D_{t-1}(D_t)$  equal to one for each of these levels in equation 2.3.7(2.3.16) and re-estimating it for the full sample period and the two sub periods (pre-/post-crisis). The estimates reveal the presence of positive feedback trading across

<sup>11</sup>The premium-levels tested for are: +25%; +0.5%; +0.75%

<sup>12</sup>The discount-levels tested for are: -0.25%; -0.5%; -0.75%

several premium/discount levels for country ETFs targeting Asia Pacific markets (particularly for the full sample period and post crisis' outbreak), while several ETFs targeting European markets also furnished us with evidence of (positive and negative) feedback trading.

Overall, this study has shown that feedback traders are active in several US-listed country ETFs, with their presence being sensitive to the time period examined and the sign and level of the (observed and forecast) percentage deviations of each ETF's price from its NAV. The fact that country ETFs targeting Asia Pacific markets are some susceptible to feedback trading, both in its conditional and unconditional versions, raises interesting issues for those investing in these ETFs. Considering the relatively limited evidence of feedback trading for country ETFs targeting markets with complete or partial overlap of trading sessions with the US, it is likely that the time difference involved contributes to this. US investors of ETFs targeting Asia Pacific markets are faced with a non-synchronicity of these ETFs' prices with their NAVs as these ETFs never trade real-time with their underlying benchmarks: they begin their trading in the US with their NAV of the day already known. Although it is possible that this foments feedback tendencies among their clientele, the validity of the latter can only be confirmed using real-time micro data.

An issue of interest to country ETFs' investors, however, that we can examine in the context of this study, is whether there exists a relationship between successful predictions of these ETFs' premiums/discounts and their feedback trading. If, for example, successfully predicted discounts in an ETF are accompanied by significant positive feedback trading, this would suggest that the predictive model (in this case, the Ornstein-Uhlenbeck process) is capable of, indirectly, offering insight into that ETF's trading dynamics as well. This is a rather interesting issue and, to that end, the dummy  $D_t$  in equation 2.3.16 is set to equal to 1 for those days when the predicted sign

of the ETF's percentage price deviation from its NAV equals the actual one<sup>13</sup>, zero otherwise, and estimate the equation for all twenty ETFs for the full sample period, prior to and after the crisis' outbreak. Results from the feedback coefficients of interest  $(\psi_{1,1}; \psi_{1,1})$  are presented in Table 2.8 and indicate that successful premium/discount predictions are accompanied by significant feedback trading on very few occasions:  $\psi_{1,0}$  is significantly positive for the iShares MSCI Belgium Capped ETF for the full sample period and significantly negative for the iShares MSCI Austria Capped ETF pre crisis' outbreak, the iShares MSCI Singapore ETF pre crisis' outbreak and the iShares Taiwan ETF for the full sample period. The significance of  $\psi_{1,1}$  (reflective of unsuccessful predictions) is linked with more (eight) cases of feedback trading significance, while the majority of coefficients reported in Table 2.8 are insignificant, thus showcasing the absence of a widespread relationship between successful predictions of country ETFs' premium/discounts and their feedback trading.

## 2.5 Conclusion

This study investigate whether feedback traders are active in US-listed country ETFs and whether their presence is affected by the significant premiums and discounts that have been documented for these ETFs in the literature. Drawing on a sample of twenty ETFs from that category for the 2000-2016 period it reports significant feedback trading for several of them, with its presence varying with these ETFs' observed/forecast premiums/discounts and the level of the latter, as well as before and after the 2008 crisis' outbreak. Country ETFs targeting Asia Pacific markets are found to be more prone to feedback trading (compared to those targeting European and Latin American markets) and it has been discussed how this might be related to the noise trading often

---

<sup>13</sup>This is the case when the Ornstein-Uhlenbeck process predicts a discount (premium) for the day  $t$  and the ETF posts a discount (premium) on day  $t$



encountered in these markets, as well as the non-synchronicity in trading times between them and their underlying benchmarks.

The previously cited study Charteris et al. (2014) is a precursor to the present work and a strong motivation for the present study. This study differs from Charteris et al. (2014) in the data employed. A much larger sample of country ETFs is taken, both emerging markets and developed markets. Also, an OU process is used to model the premium/discount process. A test on the impact of the magnitude of the lagged returns was initially carried out for the emerging markets sample only. But, as this is not the main focus of the study and the amount of data is considerable, this was left for future research. We add to Charteris et al. (2014) among other by comparing emerging and developed markets.

From a research perspective these findings bear important implications, as they offer novel insights into country ETFs' trading activity, by demonstrating how these ETFs' extensively documented wide premiums and discounts can be related to feedback trading. The evidence presented here is also of key relevance to investors, particularly those focusing on country ETFs, as it could be used to inform their trading strategies, by prompting them to utilize the relationship between feedback trading and country ETFs' premiums/discounts when trading those ETFs

Table 2.8: **Estimates from the feedback coefficients of the extended Sentana and Wadhwani (1992) model accounting for successful premiums/discounts forecast**

	Full Period		Pre crisis' outbreak		Post crisis' outbreak	
	$\psi_{1,0}$	$\psi_{1,1}$	$\psi_{1,0}$	$\psi_{1,1}$	$\psi_{1,0}$	$\psi_{1,1}$
iShares MSCI Australia ETF	0.0037 (0.5266)	-0.0076 (0.1101)	0.0320 (0.3432)	-0.0202 (0.5102)	0.0008 (0.8939)	-0.0071 (0.1294)
iShares MSCI Austria Capped ETF	-0.0002 (0.9575)	-0.0034 (0.5496)	-0.0793 (0.0315)	0.1293 (0.0139)	0.0001 (0.9857)	-0.0059 (0.3252)
iShares MSCI Belgium Capped ETF	0.0099 (0.0929)	-0.0060 (0.4033)	0.0152 (0.3352)	0.0103 (0.5828)	0.0071 (0.3579)	-0.0092 (0.2495)
iShares MSCI Brazil Capped ETF	-0.0019 (0.6275)	-0.0040 (0.1338)	-0.0073 (0.4052)	-0.0141 (0.1825)	-0.0009 (0.8259)	-0.0023 (0.4115)
iShares MSCI Canada	-0.0027 (0.7254)	0.0005 (0.9345)	-0.0133 (0.7926)	-0.0247 (0.5337)	0.0021 (0.7838)	-0.0001 (0.9843)
iShares MSCI France ETF	0.0091 (0.2978)	-0.0032 (0.5627)	0.0005 (0.9794)	0.0040 (0.8672)	0.0093 (0.3539)	-0.0047 (0.4131)
iShares MSCI Germany ETF	0.0043 (0.5812)	-0.0035 (0.5542)	-0.0011 (0.9461)	0.0094 (0.6171)	0.0057 (0.5129)	-0.0064 (0.2966)
iShares MSCI Hong Kong ETF	-0.0084 (0.1725)	-0.0104 (0.1123)	-0.0153 (0.5503)	-0.0015 (0.9583)	-0.0092 (0.1898)	-0.0114 (0.0922)
iShares MSCI Italy Capped ETF	-0.0013 (0.8517)	0.0062 (0.2337)	-0.0104 (0.7887)	0.0128 (0.6646)	-0.0049 (0.5184)	0.0052 (0.3735)
iShares MSCI Japan ETF	-0.0134 (0.2498)	-0.0041 (0.6769)	-0.0470 (0.2866)	0.0073 (0.8641)	0.0164 (0.1774)	-0.0004 (0.9962)
iShares MSCI Malaysia ETF	-0.0192 (0.1536)	-0.0374 (0.0379)	0.0017 (0.9399)	-0.0463 (0.1134)	-0.0023 (0.8653)	-0.0409 (0.0938)
iShares MSCI Mexico Capped ETF	-0.0018 (0.7313)	-0.0062 (0.2380)	-0.0169 (0.3844)	-0.0120 (0.5533)	0.0034 (0.6044)	-0.0052 (0.3568)
iShares MSCI Netherlands ETF	-0.0002 (0.9743)	-0.0047 (0.4531)	-0.0040 (0.8588)	-0.0040 (0.8253)	-0.0021 (0.7773)	-0.0050 (0.4610)
iShares MSCI Singapore ETF	-0.0091 (0.1848)	-0.0048 (0.4626)	-0.0642 (0.0133)	-0.0650 (0.0707)	-0.0066 (0.3871)	-0.0051 (0.4514)
iShares MSCI South Korea Capped ETF	-0.0017 (0.5956)	-0.0036 (0.3959)	-0.0072 (0.6479)	-0.0039 (0.8397)	-0.0005 (0.8704)	-0.0046 (0.3011)
iShares MSCI Spain Capped ETF	0.0013 (0.8349)	0.0014 (0.7882)	-0.0344 (0.2290)	-0.0139 (0.6832)	0.0005 (0.9445)	0.0014 (0.7928)
iShares MSCI Sweden ETF	-0.0022 (0.8787)	-0.0074 (0.0625)	-0.0108 (0.5012)	-0.0218 (0.0936)	0.0016 (0.7797)	-0.0069 (0.1009)
iShares MSCI Switzerland Capped ETF	0.0074 (0.4897)	-0.0512 (0.1467)	-0.0041 (0.8294)	0.0316 (0.2489)	0.0094 (0.4728)	-0.0233 (0.0207)
iShares MSCI Taiwan ETF	-0.0133 (0.0334)	-0.0075 (0.1491)	-0.0094 (0.1884)	-0.0057 (0.3655)	-0.0199 (0.4563)	-0.0489 (0.1218)
iShares MSCI United Kingdom ETF	0.0057 (0.5983)	-0.0078 (0.5087)	-0.0036 (0.9891)	0.0089 (0.6398)	0.0059 (0.5109)	-0.0069 (0.2590)

<sup>†</sup>The table presents the feedback coefficient estimates from the set of following equations for the full sample period (20/6/2000 - 27/4/2016), before (20/6/2000 - 31/8/2008) and after the crisis' outbreak (1/9/2008 - 27/4/2016):

$$\begin{aligned}
 r_t &= \alpha_0 D_t + \alpha_1 (1 - D_t) + \theta_0 D_t \sigma_t^2 + \theta_1 (1 - D_t) \sigma_t^2 + D_t 9(\psi_{0,0} + \psi_{1,0} \sigma_t^2) r_{t-1} + (1 - D_t)(\psi_{0,1} + \psi_{1,1} \sigma_t^2) r_{t-1} + \varepsilon_t, \\
 \sigma_t^2 &= \omega + \beta \varepsilon_{t-1}^2 + \lambda \sigma_{t-1}^2 + \delta I_{t-1} \varepsilon_{t-1}^2
 \end{aligned}$$

$D_t$  is a dummy variable assuming the value of unity in the predicted sign of the ETF's percentage price deviation from its NAV for day  $t$ , equals the actual one, zero otherwise. Parentheses include p-values.

## Chapter 3

# Forecasting and Trading high Frequency Volatility on Large Index

### 3.1 Introduction

Volatility plays central roles in asset pricing and allocation, and in risk management, e.g. value-at-risk and expected shortfall. Modelling and forecasting volatility is important for econometricians, statisticians and practitioners, and for that reason it has gained much interest in the financial and economic literature, however the application of traditional *Generalized AutoRegressive Conditional Heteroskedasticity* (GARCH) and *Stochastic Volatility* (SV) models are not appropriately suited for applications where *high frequency* data has been used. In response to the increasing availability of those financial data, Andersen and Bollerslev (1998) proposed that the daily volatility, which is normally treated as a latent variable in various parametric models, now can be approximated using intraday data, and their new measure was called *Realized Volatility* (RV).

Undoubtedly, high frequency data contains more information of the daily transaction, and are useful not only in measuring volatility, but also in direct model estimation and forecast evaluation, therefore, by increasing the sampling frequency, the RV is considered as a very good proxy of the true volatility under the assumption of no market microstructure noise. However, a higher frequency leads inevitably to a larger microstructure noise, thus Hansen and Lunde (2006) suggested a 5-

min sampling frequency which is now commonly used to compute RV in order to trade off the bias-variance problem.

Many recent studies focusing on high frequency data evaluate the performance of various models of RV. The parametric model termed *Heterogeneous Autoregressive model of Realized Volatility* (HAR-RV) is the commonly used for prediction of RV, which has a simple structure and also considers the long memory property. Additionally, an alternative model which allowing *jumps* or *discontinuities* in the estimation of RV is proposed by Andersen et al. (2007a), which referred as HAR-RV-J in this thesis. Their empirical studies show that incorporating the jumps to the HAR model increase the accuracy of forecasting performance.

The HAR families have been developed to capture certain features of volatility, however, the errors in prediction by using the parametric models are often argued by researchers. This is because the linear models are often based on certain distribution assumptions and the microstructure noise can arise by bid-ask bounce, asynchronous trading, and price discreteness (Barunik and Krehlik (2016)). *Artificial Neural Network* (ANN) models offer a potential improvement to earlier approaches because ANNs have the ability to tolerate data with errors and also find nonlinear associations between the parameters of the model.

This chapter compares the HAR-RV-J with a Recurrent Neural Network (RNN) and the *hybrid* HAR-RV-J-RNN model to forecast volatility, thereby analysing the forecastability. The application of machine learning is increasing in the volatility literature, however the studies also on hybrid models which incorporate parametric model and neural network by using high frequency data are limited, therefore this study is also contributed on this gap. What is more, most of the published papers evaluate the volatility forecasting performance of certain models by using traditional statis-

tical accuracy criteria, e.g. *mean square error*, *mean absolute error* and *mean absolute percentage error*. However, the practitioners select appropriate models base on financial rather than statistical criteria. Therefore, the intuitively appealing idea of this chapter is to investigate the power of forecasting models from both a statistical and economic point of view. In order to do so, a realistic volatility trading strategy is applied by using the first volatility futures ETNs to be issued were Barclays iPath S&P 500 VIX short-term (Ticker: VXX), launched at the beginning of 2009, which tracks the performance of S&P 500 VIX short-term futures index. To the best of our knowledge, this is the first attempt to apply this trading strategy based on the new approach of volatility forecasting for high frequency data.

This chapter is organised as follows. The Section 3.2 gives a brief literature on volatility modelling, forecasting and trading. The Section 3.3 discusses the volatility forecasting models including HAR-RV-J, RNN, and hybrid models. The Section 3.4 presents dataset employed in the empirical study. The Section 3.5 compares the estimation and forecasting results of the models and also introduces the volatility trading strategies. Moreover, it provides detailed trading results and the discussion of their applications. Section 3.6 concludes the whole discussion.

## 3.2 Literature Research

The development of volatility research has had at least three notable stages. The first stage is GARCH model which was proposed by Bollerslev (1986), see also Bollerslev et al. (1994) and Engle and Patton (2001). The second stage is the so called SV model which contributed to the contemporaneous development in Bayesian statistical analysis using the Markov Chain Monte Carlo procedure, see Taylor (1986) and Harvey et al. (1994), also the recent work of Lux and Moreles-Arias (2013). The third stage was followed by the work of Andersen and Bollerslev (1998) and Barndorff-

Nielsen and Shephard (2001), who proposed the use of the sum of the squared intradaily returns at different sampling frequencies as a proxy measure for the corresponding daily volatility. This measure provides a consistent estimator of the latent volatility under an ideal market condition. Barndorff-Nielsen and Shephard (2002a), Andersen et al. (2003) among others, have established some theoretical foundations for RV construction via high frequency data.

Since GARCH-type models were constructed to describe daily data, in high-frequency data environment, they are not suitable to solve this problem. Hansen et al. (2012) considered the so called realized GARCH (RGARCH) model by introducing a measure function to link the latent variances to realized volatility. HAR model became widely used to forecast realized volatility because this model easily capture the long memory property in contrast to RGARCH model. Andersen et al. (2007a) built on the theoretical results of realized variation measures constructed from high frequency returns by involving the so called bipower variation measures. Their study pointed out that volatility jump component is essential and significant jumps were associated with specific macroeconomic news announcements; see also the recent work of Borovkova and Mahakena (2015). In this context, Corsi (2009) found that HAR-RV model is able to reproduce the same volatility persistence observed in the empirical data as well as many from the other main stylized facts of financial data, in spite of its simplicity and the fact that it does not formally belong to the class of long-memory models.

Following Andersen's and his co-authors' works, Celik and Ergin (2014) found that the heterogeneous autoregressive model allowing for discontinuities was the best among high frequency based on the volatility forecasting models. They use Turkey index futures data and proved the superiority of high frequency data based volatility forecasting model over traditional GARCH model. More-

over, Papavassiliou (2016) confirmed the significance of discontinuous jumps in forecasting volatility by studying individual stocks and demonstrated the importance of using high frequency data in model-free, non-parametric financial econometric procedures. For detailed analysis of dynamics of jumps can also be found in Fan and Wang (2007); Lee and Hannig (2010); Lee (2012); Prokopczuk et al. (2015); Borovkova and Mahakena (2015); Boudt and Zhang (2015); Sevi (2014); Bajgrowicz et al. (2016), among others.

Linear models, which are based on restrictive distribution assumptions, have been developed to capture certain properties of volatility, however, changes in market conditions and many microstructure noise lead to complex patterns which cannot be captured. Other tools used in the study of return volatility are ANNs. The application of ANNs to modelling economic conditions has been expanding rapidly the last decades, see for instance Dunis and Huang (2002); Bildirici and Ersin (2009); Hajizadeh et al. (2012); Kotkatvuori-Ornberg (2016). Recent studies on stock markets price forecast using ANNs can also be found by Jammazi and Aloui (2012); Panella et al. (2012); Papadimitriou et al. (2014), among others. Kristjanpoller and Minutolo (2015) applied a hybrid ANN-GARCH model to forecast the gold price volatility and concluded that the overall forecasting performance was improved as compared to a GARCH method alone. However, their study focused on daily returns for forecasting the daily volatility, and used daily squared returns which are calculated from closing prices and therefore cannot capture price fluctuations during day. In high-frequency data context, Barunik and Krehlik (2016) proposed an ANN approach that incorporates realized measures with generalised regression to capture the complex patterns hidden in linear models, and evaluated multiple-step-ahead volatility forecasts of energy markets using several popular high frequency measures and forecasting models, concluding that this newly proposed

methodology yields both statistical and economic gains.

However, it seems that in the literature most papers evaluate forecasting performance by using traditional statistical accuracy criteria, seldom has applied the forecasting results to volatility products trading. Since the financial crisis exchange-traded products have been developed rapidly and become more popular among investors. Carr and Lee (2009) provided an extensive literature on volatility derivatives. Zhang et al. (2010) explored the relationship between the VIX index and VIX futures and showed that the VIX and VIX futures are high correlated by establishing a mean-reverting variance model. The study of Fassas and Siriopoulos (2012) showed that VIX futures prices can be used as an efficient and unbiased estimator for the spot VIX. More recently, Alexander et al. (2015) overviewed the recent developments in the volatility exchange-traded products that are related to implied volatility.

To summarise, the ANN models continue to provide more accurate forecasting performance, nonetheless, there are still room for improving upon the existing models. In the following sections, the specific methodology is presented and empirical data are used to test the model.

### 3.3 Methodology

#### 3.3.1 The HAR-RV-J Model

We consider an  $n$ -dimensional price process defined on a complete probability space,  $(\Omega, \mathcal{F}, P)$ , evolving in continuous time over the interval  $[0, T]$ , where  $T$  denotes a positive integer. Following closely the setup of Andersen et al. (2003, 2007a)'s work, let  $p_t$  denote a logarithmic asset price at time  $t$ , and incorporating also the theoretical framework of Back (1991), the continuous-time semimartingale jump diffusion process used in asset pricing is as follows:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)dq(t), \quad 0 \leq t \leq T, \quad (3.3.1)$$



where  $\mu(t)$  is a continuous and locally bounded variation process,  $\sigma(t)$  is a positive and cadlag stochastic volatility process,  $W(t)$  is a standard Brownian motion,  $q(t)$  a counting process with  $dq(t) = 1$  corresponding to a jump at time  $t$  and  $dq(t) = 0$  otherwise with jump intensity  $\lambda(t)$ , and  $\kappa(t)$  refers the size of the corresponding discrete jumps in the logarithmic price process. The quadratic variation for the cumulative return process,  $r(t) = p(t) - p(0)$ , is given by:

$$[r, r]_t = \int_0^t \sigma^2(s)ds + \sum_{0 < s \leq t} \kappa^2(s). \quad (3.3.2)$$

In the absence of jumps, the quadratic variation  $[r, r]_t$  is equal to integrated volatility  $\int_0^t \sigma^2(s)ds$ , see Andersen and Bollerslev (1998); Andersen et al. (2001, 2003, 2006); Barndorff-Nielsen and Shephard (2001, 2002a,b).

Let denote the sampled  $\delta$ -period returns  $r_{t,\delta} = p(t) - p(t - \delta)$ , then define the daily RV by summing the corresponding  $1/\delta$  high frequency intradaily squared returns:

$$RV_{t+1}(\delta) = \sum_{j=1}^{1/\delta} r_{t+j*\delta,\delta}^2. \quad (3.3.3)$$

By the theory of quadratic variation, see Back (1991); Andersen et al. (2003), the realized variation converges uniformly in probability to the increment of the quadratic variation process as the sampling frequency of the underlying returns go to infinity, that is

$$RV_{t+1}(\delta) \sim \int_t^{t+1} \sigma^2(s)ds + \sum_{t < s \leq t+1} \kappa^2(s). \quad (3.3.4)$$

Thus, in the absence of jumps the realized variation is consistent for the integrated volatility. However, in order to separate the continuous variation and jump components, Barndorff-Nielsen and Shephard (2004) proposed the *Bipower Variation* (BV), which is defined as follows:

$$BV_{t+1}(\delta) = \frac{2}{\pi} \sum_{j=2}^{1/\delta} |r_{t+j*\delta,\delta}| |r_{t+(j-1)*\delta,\delta}|. \quad (3.3.5)$$

As  $\delta \sim 0$ , it is possible to see that:

$$BV_{t+1}(\delta) \sim \int_t^{t+1} \sigma_s^2 ds. \quad (3.3.6)$$

Combining the results in Eqs. (3.3.4) and (3.3.5), the contribution to the quadratic variation process due to jumps in the underlying process can be estimated by:

$$RV_{t+1}(\delta) - BV_{t+1}(\delta) \sim \sum_{t < s \leq t+1} \kappa^2(s). \quad (3.3.7)$$

To prevent the right hand-side of Eq. (3.3.7) from becoming negative, Andersen et al. (2007a) imposed non-negativity truncation on the jump measurements:

$$J_{t+1}(\delta) = \max[RV_{t+1}(\delta) - BV_{t+1}(\delta), 0]. \quad (3.3.8)$$

HAR-RV model is introduced by Corsi (2009), and it can be expressed as:

$$RV_{t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \varepsilon_{t+1}, \quad (3.3.9)$$

$t = 1, 2, \dots, T$ .  $RV_t$ ,  $RV_{t-5}$  and  $RV_{t-22}$  mark daily, weekly (5 business days) and monthly (22 business days) RV, respectively. Weekly and monthly RV is calculated as:  $RV_{t,t+h} = h^{-1}[RV_{t+1} + RV_{t+2} + \dots + RV_{t+h}]$ ,  $h = 1, 2, \dots$ . Andersen et al. (2007a) proposed the new HAR-RV-J model, in which included the jump components. Daily HAR-RV-J model is expressed as:

$$RV_{t,t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \beta_J J_t + \varepsilon_{t,t+1}. \quad (3.3.10)$$

Logarithmic and standard deviation form of HAR-RV-J model is given by:

$$(RV_{t,t+1})^{1/2} = \beta_0 + \beta_D (RV_t)^{1/2} + \beta_W (RV_{t-5,t})^{1/2} + \beta_M (RV_{t-22,t})^{1/2} + \beta_J (J_t)^{1/2} + \varepsilon_{t,t+1}, \quad (3.3.11)$$

and

$$\log(RV_{t,t+1}) = \beta_0 + \beta_D \log(RV_t) + \beta_W \log(RV_{t-5,t}) + \beta_M \log(RV_{t-22,t}) + \beta_J \log(J_t + 1) + \varepsilon_{t,t+1}. \quad (3.3.12)$$

### 3.3.2 Recurrent Neural Networks

ANNs are a powerful non-parametric tool used for signal filtering, recognition of patterns and interpolation, also, can tolerate data with errors and find nonlinear associations between the parameters of the model, see Haykin (2007); Kristjanpoller et al. (2014); Kristjanpoller and Minutolo (2015). In particular, as we discussed it in Section 3.2, ANNs have been applied with increasing success to economic and financial forecasting. Most of econometric models are developed by capturing specific features of time-series, e.g. long memory, or making an assumption of functional relationship among variables, the major advantages of ANNs is that they contain nonlinearities and incorporate all variables.

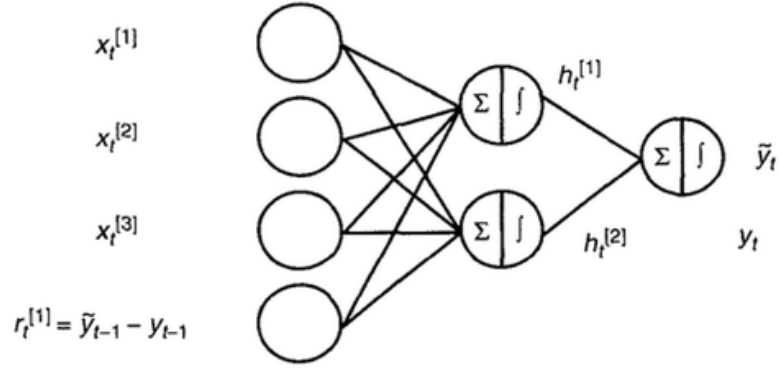
Briefly speaking, see also Haykin (2007) among other classical books, each neural network connects a group of *input* variables  $\{x_1, x_2, \dots\}$  with one or more *output* variables  $\{y_1, y_2, \dots\}$  and zero, one or more *hidden* layers. Neurons are connected between the layers for connections that are activated by reaching a threshold. Each layers can have a different number of neurons. A series of weight vectors  $\{w_{i,j}, w_{2,j}, \dots, w_{n,j}\}$  is associated with the input vectors, each node may additionally have also a bias input  $\theta_j$ , thus the actual outputs of the neurons in the hidden layer is:

$$y_i = \text{sigmoid}[\sum_{i=1}^n x_i * w_{i,j} - \theta_j],$$

and *sigmoid* is the sigmoid activation function  $f(x) = \text{sigmoid}(x) = \frac{1}{1+e^{-x}}$ .

In this thesis it uses *Recurrent Neural Network* (RNN) models which were introduced by Elman (1990), see also an application of RNNs in currency trading by Dunis and Huang (2002). Their only difference from multilayer neural network is that they include a loop back from one layer, either the output or the intermediate layer or the input layer. Figure 3.1 shows a single output

Figure 3.1: Single output Recurrent Neural Network (RNN) model with one hidden layer



RNN model with one hidden layer and two hidden nodes.

When using ANN, generally, we have to carefully consider the topic of data pre-processing. Indeed, for most ANN the output is limited due to the squashing function being either the hyperbolic tangent or the logistic function. These two sigmoid squashing functions as mentioned above are the most commonly used. While the hyperbolic tangent leads to an output in the interval  $] - 1, +1[$ , the logistic function has an output in the interval  $]0, +1[$ . In the empirical application in this thesis, it chooses the latter one, the logistic function that is better adapted to the output domain of the numbers it analyses.

However, the generated data is typically relatively small compared to unity. The typical order of magnitude is between  $10^{-5}$  and  $10^{-4}$ , see Table 3.1. In this case, a simple linear transformation is advised that makes better use of the available output range of the logistic function. To achieve this, we simply multiply all data with  $10^3$  for the linear HAR-RV-J model inputs. As it also trains ANN on the logarithmic deviation of the HAR-RV-J model a slightly different approach is needed. Indeed, the logarithmic transformation will generate data in the range  $[-14, -5]$ , see Table 3.1. This also has to be transformed. An appropriate scaling factor is given by  $5 \times 10^{-2}$  with a shift of  $+1$ . This simple linear transformation does not change the basic interpretation of the data but

makes it easier to be learned by an ANN.

The inputs to the RNN model are the same as to the linear model to allow a fair comparison. That is, the three RV and one jump inputs are included. Also, the question is to the meta-parameters of the neural network that have to be addressed, specifically, the size of the hidden layer and the exact architecture. Determining the number of hidden neurons is often left to experiment, and no single dominant method has emerged. However, taking twice the geometric mean of the input and output layer size is an often used heuristic. Therefore in all of RNN models the number of hidden neurons  $h$  is determined in the following way:

$$h = 2 \times \sqrt{i \times o}, \quad (3.3.13)$$

where  $i$  and  $o$  refer to the size of the input and output layer, respectively. Note, that in the case of the input layer we do not take the bias neuron into account.

Upto this point the RNN looks very similar to a standard three-layer perception. However, storing the output in a separate state layer and feeding this back into the hidden layer makes the network state aware. Therefore, when evaluating the network, we have to be careful to store the present state and to carry out any evaluation in sequential order of time. As the rolling-window approach includes just one forecast per model, there is not much potential for confusion in our specific application. Each model is trained on the rolling-window and then immediately used for the corresponding forecast horizon. However, if the network has to be reused for several forecasts, then this issue has to be considered. For example, if network training times were much longer, it might make sense not to train a new model every day. In this latter case it is compulsory to store the network state for later reuse.

The training of ANN is a topic that has been discussed intensely. The architecture of a neural

network makes it computationally efficient and straightforward to compute the partial derivatives of the error with respect to the network’s weights. Therefore, algorithms that make use of partial derivatives can be used. This includes, for example, simple gradient descent and its variations. However, computing the Hesse matrix of second order partial derivatives is much less straightforward and computationally intensive. Therefore, pure Newton methods are generally avoided. However, quasi-Newton methods can be used instead. In the case of the training of an ANN our goal is twofold. On the one hand, of course, we want to decrease the error. On the other hand, however, we also strive at achieving a robust model. Therefore, the weight set that leads to the absolute smallest error is not necessarily the one to be preferred, if it is just a lone minimum in a steep valley. Rather, we would prefer minima where the neighbouring values also lead to decent results. In the present case, all of the networks are trained using resilient backpropagation, where it only considers the sign of the first partial derivative. This, generally, leads to a robust result.

### 3.3.3 Hybrid Model

The hybrid model is also designed as an RNN. However, as an additional input the *forecast* of the linear model are fed to the RNN. The four basic inputs are kept. Thus, the total number of inputs rises to five in the case of the hybrid model.

All other model parameters are kept the same. Specifically, the number of hidden neurons is determined as above. Also, the model architecture stays identical.

The motivation of using a hybrid model stems from the desire to use each model in a way that exploits its specific abilities. By feeding the linear forecast to the RNN it potentially removes any linear component from the forecasting task. This should leave more room for better matching the non-linear residual of the linear forecast error.

### 3.4 Data

The base dataset consists of tick data from Thomson Reuters Tick History (TRTH) for the S&P500 index that starts on the January, 2nd, 1996 upto June, 2nd, 2016. For the reason it has been mentioned in Section 3.1, see Hansen and Lunde (2006), initially we aggregate it to 5 minutes data. Thus, intraday RV is computed based on these 5 minutes blocks. To filter out any half-holidays, it requires that a trading day has to have a complete history of data from 10am to 3.55pm in order to be included in our computation. The regular trading hours for the index are 9.30am to 4pm. However, as is commonly done, it excludes the first half hour of trading and the very last five minute interval of the regular trading session to avoid any bias that may be caused by the price determination process at the beginning of the trading day or by any rebalancing trades towards the end of the session.

Of course, it is necessary to strike a balance between unwanted noise and making best use of the intraday data. The specific cut-off times are debatable. However, the preliminary experiments showed that the above procedure produced sensible results.

The tick data series from TRTH for the SPY ETF, the VIX index and the VXX ETN start on March, 20th, 1996, January, 2nd, 1996, and January, 30th, 2009, respectively. Among these, only the VXX series starts much later, simply, because the corresponding ETN was introduced only in 2009.

### 3.5 Empirical Results

Table 3.1 summarizes the distributional properties of RV and jump series. It is evident that the realized volatility and jump are highly significant serial correlation. This can be confirmed by the Ljung and Box (1978) statistics for up to tenth-order serial correlation. Variables have kurtosis

greater than 3 indicating leptokurtic distribution, the distribution of logarithmic transformation of RV are closer to normal than RV and standard deviation form of RV. This finding is consistent with the study of Andersen et al. (2007a). Figure 3.2 provides a visual illustration of RV and jumps for S&P500. Also consistent with earlier evidence of Andersen et al. (2007a), many of the largest realized volatility are directly associated jumps in the underlying price process. The largest jump occurred around 2008 when global financial crisis broke out. In the following it presents statistical analysis of various models, it is expected the modelling to yield insight into the actual tradeability of volatility with readily available products.

All computations are rather straightforward and can be carried out in acceptable time on a laptop. Our machine is a Lenovo Thinkpad W530 with an Intel Quad Core i7-3720QM CPU, running at 3.6GHz, with 6MB Level 3 cache, and 1600MHz FSB. Due to the generally fast computation no care was taken to parallelize the computation. Therefore, in the end, only one core was used for all computations. As it computes daily updated rolling-window forecasts a single run through the dataset produces 4448 single models to cover a timespan of approximately 18 years. Two years have to be removed to account for the maximum lookback. For each of the 4448 models three variants have to be computed for the linear model, RNN, and hybrid model. Both again are computed using the basic variant and the *log* variant. Each model finally is computed for five different lookbacks and three different forecast horizons. In total this leads to  $4448 \times 3 \times 2 \times 5 \times 3 = 400320$  different models that are computed. Or, put differently, for each of our three model types 133440 models are computed. This may sound like a lot, however all our models are comparatively small by today's standards and can therefore be computed quickly. The linear model has a closed form solution and just requires linear algebra. The RNN has to be trained numerically. The computational core of an



Table 3.1: **Summary statistics for S&P500 index at 5-minute frequency**

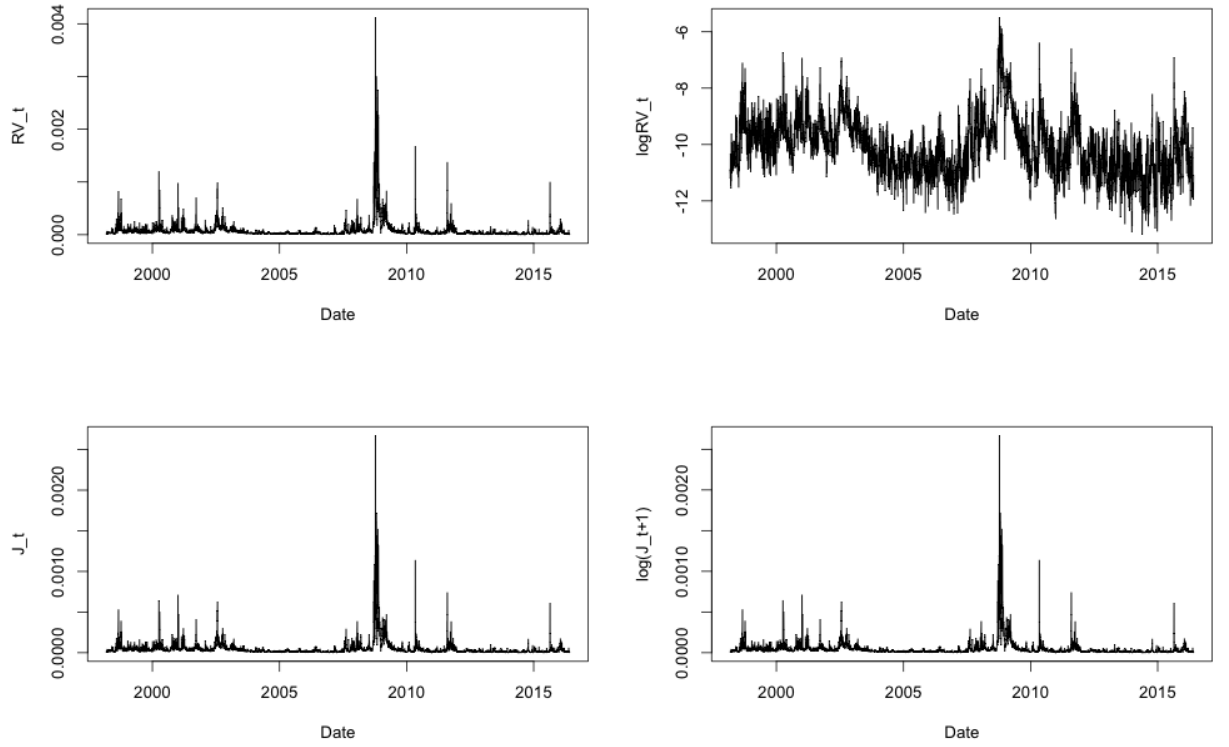
	$RV_t$	$RV_t^{1/2}$	$\log(RV_t)$	$J_t$	$J_t^{1/2}$	$\log(J_t + 1)$
Mean	0.7863	0.0074	-10.1053	0.4898	0.0059	0.4897
St.dev	0.0002	0.0049	1.0367	0.0001	0.0038	0.0001
Skewness	10.0472	3.3765	0.4676	10.2250	3.3436	10.2159
Kurtosis	153.3177	22.9232	3.5630	162.9341	22.7987	162.6297
Min	0.0186	0.0014	-13.1961	0.0113	0.0011	0.0113
Max	0.0041	0.0642	-5.4919	0.0027	0.0517	0.0027
$LB_{10}$	15188	22354	22391	14782	22086	14790

**Note:** The table summarizes the distributional properties of the daily realized volatility.  $RV_t$ ,  $RV_t^{1/2}$  and  $\log(RV_t)$  denote the daily realized volatility, realized standard deviation and realized logarithmic form respectively.  $J_t$ ,  $J_t^{1/2}$  and  $\log(J_t + 1)$  denote the daily jump measures invariance, standard deviation, and logarithmic forms respectively. LB is the statistics of Ljung and Box (1978) Q test for up to tenth-order correlation. The mean and minimum values of  $RV$ ,  $J$ , and  $\log(J + 1)$  are multiplied by  $10^4$ .

ANN is, however, also linear algebra with just a very small amount of computation time dedicated to computing the non-linear squashing function. In all cases, the initial data-preprocessing time can be neglected, as it only has to be carried out once per variant.

In total the computation times are 492 seconds for the linear models, 22217 seconds for the RNN and 56582 seconds for the hybrid model. Adding up the numbers leads to a total computation time of a bit more than 22 hours, for the *entire* model ranges. The computation could even have been sped-up by parallelizing it, as this problem is ideally coarsely parallelizable without interdependencies which would have slowed down the computation. However, for simplicity, this was not done presently. Looking at the numbers, this model seems very suitable for real-time applications. Indeed, for computing a single decision (model) it only needs to carry out the computation once. Dividing the above numbers by the number of models per variant it has been achieved computation times of 3.7 ms per model in the linear case, 166.5 ms per model for the RNN, and, finally, 424.0 ms per model for the hybrid case.

Figure 3.2: Daily S&P500 Realized Volatility and Jumps



**Note:** The top panel shows daily realized volatility and its logarithmic transformation,  $RV_t$  and  $\log(RV_t)$ , respectively. The lower panel graphs the jump components,  $J_t$  and  $\log(J_t + 1)$ , respectively.

### 3.5.1 Statistical Errors

As a first indication with respect to the quality and robustness of the different HAR model implementations, it presents and discusses the usual statistical errors like *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), and *Mean Absolute Percent Error* (MAPE) for the different types of models computed, Hyndman and Athanasopoulos (2014).<sup>1</sup>

Each group of columns in the following Tables 3.2 and 3.3 represents a forecast with the given number of days (lookahead) described by the letter  $l$ . The first group of columns therefore stands for a 1-day forecast, the second column group for a 2-day forecast and the third column group for a 1-week (i.e., 5 business days) forecast. Indeed, while the original HAR model limits the analysis to a 1-day forecast, there is no specific reason to assume, that this model type would be less suitable for forecasts several time steps into the future. For this reason it is of interest to analyse model robustness for different forecast horizons which would allow to identify model types which might be suitable for use in different setups.

Each group of rows in the Tables 3.2 and 3.3, furthermore, represents a different amount of historical data used in the rolling window. Therefore, the first group of rows just makes use of the past 22 business days (around 1 month) of data, while the last group of rows includes 504 business days (around 2 years) of in-sample data for computing the out-of-sample forecast.

Within the groups of rows, it presents the different model types one separate rows. The first row in the group shows the basic linear (or logarithmic) variant, while the second row presents the RNN version (using the same inputs). Finally the third row in each group shows the results of the hybrid model. To recall, the hybrid model is also an RNN that uses the previously mentioned four

---

<sup>1</sup>In Hyndman and Koehler (2006) (and references therein), there is an interesting discussion and comparison among different measures of accuracy of univariate time series forecasts.

inputs, and, additionally, the linear model forecast as a fifth input. The two tables, finally, present grouped results for the basic linear version of the HAR-RV-J approach and for the logarithmic version.

As the different forecast horizons solve different forecasting problems, it makes sense to analyse and discuss the results depending on the forecast horizon. For the HAR base version and a 1-day ahead forecast, it has been noticed that the RNN with just 22 days of in-sample data consistently produces the best results, when looking at the fit (out-of-sample RMSE). This is remarkable, because this best result is achieved with the least amount of data. The linear model manages to come close to the RNN results, but only when using 2 years (504 business) days of in-sample data. It appears, therefore, that the RNN is able to better extract information from a relatively small amount of data, than the linear model. This was rather expected as the RNN has the additional advantage of an implicit storage of state data, while the linear model only explicitly stores state as given by the different (three) lags of RV. Now, this may seem like a trade-off between a more complex model using less data and a simpler model using more data. However, for practical applications, we may be limited in the amount of intraday data available. Let alone that not everyone has an easy access to large historical database of intraday data. In practice, most paid-for data services might only offer a 3 months historical intraday database at an affordable price. Or conversely, the user may be tempted to collect tick data. In all these cases, the RNN provides a practical and robust solution, that, at the worst, takes 1-month of data collection of ramp-up time. On the other hand, having to potentially collect 2 years' worth of data before being able to use the linear model does not seem very practical.

When further analyse the two time steps ahead forecast, it has been noticed, again, that the

single best fit is achieved by an RNN variant for a rolling-window length of 6 months (126 business days). The forecast quality for the short input interval of 1 month is also not bad, but is, indeed, just ever so slightly beaten by the linear model fit for a 2 years rolling-window length. Nevertheless, we cannot fail to notice that in all cases the RNN performs more robustly and more consistently.

Finally, for the 5-day ahead forecast, again it has the single best fit provided by the RNN with a rolling-window length of 22 days. The performance of the linear model for short rolling-window lengths is abysmal. However, when adding more data the linear model performance improves. It is almost on-par with the RNN results when considering the longest input intervals.

If consider other error measures like MAE and MAPE it has been also noticed, that the RNN solution consistently produces excellent results, even for the shortest rolling-window length. However, for MAE and MAPE the single best solution may in some cases also be produced by the linear model with the longest rolling-window. However, it has been noticed that in all cases the RNN solution has a robustly low error, while the linear model produces erratic results which does not inspire confidence in its robustness for the shorter intervals. Therefore, even when considering MAE and MAPE there is a good deal to be said in favour of the RNN model, because the result is among the best achievable results for all rolling-window lengths but only requires a small amount of training data.

It has been also noticed, that the RNN model keeps its generally good forecasting performance, even for longer forecast horizons. Indeed, we expect the effect that forecasts for longer horizons are less good than forecasts for shorter horizons. But, the linear model worsens dramatically, while the RNN model does not change much. As a general recommendation, it could be argued that, if a suitable implementation is available, the RNN seems like a good choice for practical applications,

because it produces excellent (often even best) results, for the shortest rolling window length. If, however, plenty of data is available, and/or a linear model is preferred for whatever reasons, then good results can also be expected in the linear model case, but a longer rolling-window interval should be used. As in Barunik and Krehlik (2016) for the energy market volatility, the hybrid model's quality is often between that of the RNN and linear version. This may seem disappointing at first sight. Indeed, the initial expectation would have been, that the RNN is able to make better use of the linear forecast. In theory, it is expected that the hybrid model to be at least as good as the basic RNN model. In the worst case, it can be argued that the hybrid RNN appears to ignore all the linear forecasts by setting the corresponding weights to zero in the training process. However, in practical applications, this is not quite so clear cut as the training process depends on many factors. As we see here, the linear forecast does not help in improving basic RNN forecast, although the performances are pretty close, see Table 3.2. Therefore, there does not seem to be any compelling reason for using a hybrid model in this specific case.

When using the log variant of the HAR model, Table 3.3, the results are less clear cut. Generally, we may find that the RNN still produces very good results, but they are on-par with the results of the linear model. Still, the feature remains, that for all forecast horizons and all error measures the RNN is the model that makes best use of a short dataset. With more data the linear model improves and, even beats the RNN for very long rolling windows, but the results are pretty close. Therefore, if data is scarce, the RNN model still provides good and robust results. However, when a very long dataset is available, then the linear model should be preferred.

Table 3.2: Comparison of Models I

	RMSE	MAE	MAPE		RMSE	MAE	MAPE		RMSE	MAE	MAPE
$t = 22, l = 1$				$t = 22, l = 2$				$t = 22, l = 5$			
HAR-RV-J	0.1963	0.0523	100.8305		0.2847	0.0705	152.0512		0.6694	0.1153	205.5808
RNN	<b>0.1192</b>	<b>0.0398</b>	<b>53.9350</b>		0.1329	0.0451	58.1954		<b>0.1329</b>	<b>0.0479</b>	<b>64.7888</b>
Hybrid	<b>0.1303</b>	<b>0.0483</b>	<b>55.9974</b>		0.1571	0.0568	61.1080		0.1564	0.0594	65.9130
$t = 63, l = 1$				$t = 63, l = 2$				$t = 63, l = 5$			
HAR-RV-J	0.1563	0.0448	64.8351		0.1712	0.0539	88.7721		0.2544	0.0714	99.4946
RNN	0.1214	0.0433	60.5210		0.1288	0.0476	63.9215		0.1367	0.0507	69.0026
Hybrid	0.1479	0.0646	63.4334		<b>0.1401</b>	<b>0.0573</b>	<b>65.3191</b>		<b>0.1381</b>	<b>0.0551</b>	<b>70.8009</b>
$t = 126, l = 1$				$t = 126, l = 2$				$t = 126, l = 5$			
HAR-RV-J	0.1363	0.0415	48.7744		0.1502	0.0485	58.2050		0.1933	0.0601	75.6028
RNN	0.1228	0.0445	64.1628		<b>0.1273</b>	<b>0.0479</b>	<b>66.7512</b>		0.1417	0.0532	71.6463
Hybrid	0.1515	0.0701	67.4231		0.1448	0.0635	70.1290		0.1482	0.0605	73.1933
$t = 252, l = 1$				$t = 252, l = 2$				$t = 252, l = 5$			
HAR-RV-J	0.1239	0.0349	46.9789		0.1410	0.0455	52.2581		0.1651	0.0529	88.1153
RNN	0.1238	0.0471	66.9135		0.1350	0.0501	70.5627		0.1421	0.0541	73.2343
Hybrid	0.1448	0.0700	70.6262		0.1700	0.0680	72.9267		0.1548	0.0656	74.5489
$t = 504, l = 1$				$t = 504, l = 2$				$t = 504, l = 5$			
HAR-RV-J	<b>0.1194</b>	<b>0.0382</b>	<b>45.7595</b>		<b>0.1316</b>	<b>0.0431</b>	<b>50.8488</b>		<b>0.1522</b>	<b>0.0497</b>	<b>61.2953</b>
RNN	0.1309	0.0508	72.4723		0.1364	0.0534	75.5360		0.1488	0.0572	80.2657
Hybrid	0.1438	0.0727	75.4535		0.1552	0.0747	76.3257		0.1635	0.0745	80.9149

**Note:** The table presents the comparison of forecasting performance of HAR-RV-J, RNN and hybrid models in which the daily realized volatility,  $RV_t$ , is used.  $t$  indicates the rolling window length, e.g.  $t = 22$  denotes the length of rolling window is 1 month.  $l = 1, 2, 5$  indicates one step forecast, two steps forecast and five steps forecast, respectively. The best results considering RMSE, MAE and MAPE for those models for the different lengths of rolling windows and steps of forecast are highlighted in bold.

Table 3.3: Comparison of Models: II

	RMSE	MAE	MAPE		RMSE	MAE	MAPE		RMSE	MAE	MAPE
$t = 22, l = 1$				$t = 22, l = 2$				$t = 22, l = 5$			
HAR-RV-J LOG	0.0593	0.0281	5.6179		0.0633	0.0331	6.8792		0.1012	0.0456	9.2539
RNN LOG	<b>0.0325</b>	<b>0.0253</b>	<b>5.1552</b>		<b>0.0338</b>	<b>0.0263</b>	<b>5.3496</b>		<b>0.0362</b>	<b>0.0281</b>	<b>5.7234</b>
Hybrid LOG	<b>0.0365</b>	<b>0.0287</b>	<b>5.8164</b>		<b>0.0370</b>	<b>0.0289</b>	<b>5.8884</b>		<b>0.0398</b>	<b>0.0312</b>	<b>6.3415</b>
$t = 63, l = 1$				$t = 63, l = 2$				$t = 63, l = 5$			
HAR-RV-J LOG	0.0345	0.0228	4.6061		0.0356	0.0261	5.2888		0.0537	0.0321	6.5048
RNN LOG	0.0381	0.0297	6.0166		0.0387	0.0302	6.1148		0.0400	0.0311	6.3093
Hybrid LOG	0.0418	0.0330	6.6341		0.0417	0.0327	6.5987		0.0423	0.0332	6.7356
$t = 126, l = 1$				$t = 126, l = 2$				$t = 126, l = 5$			
HAR-RV-J LOG	0.0329	0.0246	4.4305		0.0328	0.0246	5.0191		0.0390	0.0284	5.7771
RNN LOG	0.0415	0.0323	6.5449		0.0419	0.0326	6.5984		0.0430	0.0334	6.7578
Hybrid LOG	0.0446	0.0352	7.0505		0.0444	0.0349	7.0295		0.0454	0.0354	7.1593
$t = 252, l = 1$				$t = 252, l = 2$				$t = 252, l = 5$			
HAR-RV-J LOG	0.0276	0.0213	4.3524		0.0310	0.0239	4.8576		<b>0.0254</b>	<b>0.0269</b>	<b>5.4719</b>
RNN LOG	0.0457	0.0355	7.1384		0.0460	0.0358	7.1962		0.0466	0.0361	7.2589
Hybrid LOG	0.0490	0.0384	7.5772		0.0484	0.0380	7.5772		0.0494	0.0384	7.7117
$t = 504, l = 1$				$t = 504, l = 2$				$t = 504, l = 5$			
HAR-RV-J LOG	<b>0.0274</b>	<b>0.0212</b>	<b>4.3370</b>		<b>0.0303</b>	<b>0.0235</b>	<b>4.8048</b>		0.0342	0.0264	5.3852
RNN LOG	0.0490	0.0382	7.6719		0.0388	0.0498	0.0388		0.0502	0.0390	7.8337
Hybrid LOG	0.0526	0.0414	8.2160		0.0524	0.0411	8.1993		0.0529	0.0413	8.2777

**Note:** The table presents the comparison of forecasting performance of HAR-RV-J, RNN and hybrid models in which the logarithmic transformation of daily realized volatility,  $\log(RV_t)$ , is used.  $t$  indicates the rolling window length, e.g.  $t = 22$  denotes the length of rolling window is 1 month.  $l = 1, 2, 5$  indicates 1 step, 2 steps, and 5 steps forecast, respectively. The best results considering RMSE, MAE and MAPE for those models for the different lengths of rolling windows and steps of forecast are highlighted in bold.

### 3.5.2 Trading Efficiency

While modelling and forecasting RV is an interesting and instructive exercise in itself, the question arises, to which extent the results are useful in financial applications. Or, put in other words, can the models uncover an exploitable market inefficiency? For this reason it is useful to develop ideas as to how a forecast of RV could actually be traded on the financial markets. Two general strategies come to mind:

- An outright volatility trade.
- Trade the expected reaction of another asset with respect to volatility.

Since the advent of VIX futures in 2004 on the most important volatility index, a simple way for outright volatility trading is available. The VXX ETN that mimics a 30-day constant maturity future makes trading volatility even easier and puts it in reach of retail investors. The VXX ETN started trading in 2009. Our basic volatility trading strategy consists of very simple elements:

- If  $RV_{t+k} > RV_t$  then go (and stay) long volatility for  $k$  time steps,
- else go (and stay) short volatility for  $k$  time steps.

In the above,  $RV_{t+k}$  denotes the  $k$  time steps ahead RV forecast at time  $t$ , while  $RV_t$  simply denotes today's RV. In the case  $k > 1$ , the investing capital are equally split among the number of concurrent trades. In this case, it invests 50% of the capital on two potentially differing trades for 2-day ahead forecasts, and 20% of our capital for 5-day ahead forecasts. Note, that this may also lead to two trades cancelling each other out. That means, if we have a signal for a long volatility trade today (and a 2-day ahead forecast) and a signal for a short volatility trade the next day, then the net position would be zero, because the long and the short trade cancel each other out. In the



above strategy, even in the special case that  $RV_{t+k} = RV_t$  it goes short volatility, because going short volatility is a long term winning strategy (although with mind boggling drawdowns, therefore surely not advised for practical implementation). This is due to the observation that volatility tends to go up in short (but large) spikes, and then continuously fall again.

However, we may want to limit trading in this strategy and only trade when our forecast predicts a significant change in RV. If we assume a threshold  $p$  expressed as a percentage range, for example  $p = 0.01$  or  $p = 1\%$ , then we can modify the above strategy to:

- If  $RV_{t+k} > RV_t * (1 + p)$  then go (and stay) long volatility for  $k$  time steps,
- if  $RV_{t+k} < RV_t * (1 - p)$  then go (and stay) short volatility for  $k$  time steps,
- else stay out of the market (flat).

Again, investment capital is allocated equally, however still depending on the forecast horizon. When we speak of trading volatility, we have to be careful to note that RV is, in itself, not a tradable quantity. Also, the VIX index itself is not tradable. Strictly speaking, we can only trade volatility derivatives with a defined maturity. The threshold  $p$  has to be determined heuristically. However, it makes sense to set it in a way, that, at least, transaction costs are overcome if the forecast proves to be correct.

A second variant involves trading an asset, of which it is expected that it moves in close relation to volatility. Generally, asset returns and volatility are supposed to be negatively correlated. Indeed, in our sample period a quick analysis of the S&P500 Index and the VIX index reveals that in around 80% of the cases, where the S&P500 Index is up, the VIX index is down and the other way round. We would therefore expect that a trading strategy that uses a volatility forecast to trade the S&P500 index ETF would be a sensible approach. Therefore, the following base strategy has been

proposed:

- If  $RV_{t+k} > RV_t$  then go (and stay) short the base index ETF for  $k$  time steps,
- else go (and stay) long the base index ETF for  $k$  time steps.

Here, again, our default case for  $RV_{t+k} = RV_t$  is to go long the base index ETF, because, on average, and on a very long time frame, asset prices tend to go up. This argument is, of course, debatable, but it only concerns a very minor edge case.

In a similar way, we may want to implement a threshold in order to avoid overtrading and engaging in trades that potentially do not cover the transaction cost. Analogously this allows us to define the following trading strategy:

- If  $RV_{t+k} > RV_t * (1 + p)$  then go (and stay) short the base index ETF for  $k$  time steps,
- if  $RV_{t+k} < RV_t * (1 - p)$  then go (and stay) long the base index ETF for  $k$  time steps,
- else stay out of the market (flat).

The advantage of the direct volatility trade is that the asset traded (a volatility derivative) may correspond more closely, to what is actually forecast. On the other hand, the base index strategy can be implemented, even if there is no volatility derivative available. However, we have to be careful to first carry out an analysis whether the implied correlation between asset returns and changes in volatility is indeed present in the index that we want to trade. This significantly broadens the universe of assets within reach of a volatility forecasting model.

In the above strategies the threshold approach seems appealing to limit trades and whipsaw but introduces a new meta-parameter into the system. To avoid too much interference with the base system, it has been estimated the threshold that it removes the 10% smallest trades (by absolute

Table 3.4: Out-of-sample rolling window trading results for the Linear Model using a lookback of 252 days

Strategy	Ann. Ret.	Ann. Vol.	Sharpe
$h = 1$			
Index	5.68%	20.49%	0.28
Vol.	55.99%	108.58%	0.52
$h = 2$			
Index	5.18%	13.95%	0.37
Vol.	58.86%	73.54%	0.80
$h = 5$			
Index	5.57%	8.30%	0.67
Vol.	56.98%	42.60%	1.34

Table 3.5: Out-of-sample rolling window trading results for the Recurrent Neural Network (RNN) using a lookback of 22 days

Strategy	Ann. Ret.	Ann. Vol.	Sharpe
$h = 1$			
Index	0.36%	1.75%	0.20
Vol.	13.28%	14.61%	0.95
$h = 2$			
Index	5.79%	13.96%	0.41
Vol.	55.44%	73.56%	0.75
$h = 5$			
Index	5.72%	8.30%	0.69
Vol.	55.69%	42.63%	1.31

forecast difference) in the lookback interval. Of course, it could now be argued that removing the bottom  $x\%$  trades also represents a meta-parameter and that is true. However, it makes the meta-parameter adaptable and dependent on the actual history.

In the following, it presents trading results for the linear and RNN model. Only results for the lookbacks are presented which seem most suitable according to the statistical evaluation. Table 4.7 shows results for the linear model with a lookback of 252 days. For this lookback, performance seems to stabilize in the statistical evaluation. Table 3.5 presents trading results for the RNN model using the short lookback of 22 days in the rolling window forecast.

For each model type (linear and RNN), it presents trading results for trading the index and

the volatility derivative. As a risk adjusted measure of return, we opt for the Sharpe ratio. We are all very well aware of the inherent limitations of the Sharpe ratio, but in this study it is still used, as it is a good basis for comparison. And, despite, all the critics of the Sharpe ratio, mostly only artificially constructed return series may exhibit good Sharpe ratios but otherwise disappointing risk measures. For realistic return series, we may expect good Sharpe ratios to generally lead to also otherwise attractive capital curves.

As it is typical, the volatility derivative trading strategies show high returns. However, volatility is in itself a very volatile asset class, and the optically high returns are mitigated by correspondingly high volatility of the trade returns. Therefore, it is only advised allocating a small portion of the portfolio to a volatility strategy. It is not suitable as the only trading strategy for any but the most risk-loving portfolios.

As expected the results between the linear and RNN model do not differ much, as the chosen parameter sets have very similar statistical results. While the RNN model had overall better statistical error measures (and, definitely, was able to produce these good results with a modest amount of data) this does not translate to a clearly better trading strategy. For all forecast horizons  $h$ , the results are pretty similar when comparing linear and RNN model. There is no clearly dominant strategy.

We may want to put the trading results into the context of the overall Sharpe ratio for the S&P500 index, that is often used as benchmark. For the trading period of this study this is 0.29. It is therefore noticed that trading the index with either model improves slightly on this benchmark for forecast horizons  $h = 2$  and  $h = 5$ . For  $h = 1$  performance is not attractive. We, generally, get more attractive Sharpe ratios for trading a volatility derivative. This is expected as, in this

latter case, we are actually trading something very similar to what is forecast and not using an indirect correlation. Also, it has been noticed that Sharpe ratios tend to improve for the longer forecast horizon. Here, it uses the effect to our advantage, that a potentially wrong position might be corrected the next day by the correct forecast. It is a very crude way of diversification in time.

While Sharpe ratios of more than 1.3 for the longest horizon volatility strategies with either model seem attractive, it is necessary to notice, that we did not yet carry out an analysis for different time periods. The effect seems stable, but, as volatility gets more and more attention, it seems probably that any potential inefficiency will fade quickly.

As an overall observation, it concludes that the given inputs seem to be able to produce attractive Sharpe ratios for either an index or volatility derivative trade for a forecast horizon of 1 week,  $h = 5$ . This applies to both the linear and RNN models. As there does not seem to be any systematic bias towards one model type or the other the availability of data versus ease of implementation can be the basis for the choice of which strategy to adopt. In both cases 5-minute intraday data is necessary to operate the models. Both models train very quickly even on a standard laptop or personal computer. Therefore, training and evaluation time are of no concern, typically. If only a short amount of data is available the RNN model seems preferable. If a longer amount of data is available, then, the choice is pretty much up to the availability of an implementation of one or the other model.

An overall decision on which of the two basic model types, HAR or RNN should be chosen, cannot be taken without further knowledge of the context. In the end, given a sufficiently large lookback period, both models are able to achieve similar performance. This raises the question, why the RNN model could ever preferred to the simple HAR model. As already outlined, the RNN

model is able to realize satisfying (if not necessarily good) performance also in cases of a relatively low amount of training data. For fast changing environments or environments where data is expensive to obtain or for newly trading assets this can be considered an advantage. Also, when computational costs are considered, neither model represents a challenge for today's modern computing infrastructure. While estimating the HAR model is still several times faster than estimating an RNN, both estimations are easily done in less than one second, even on a standard personal laptop. In the context of this study only a single model estimation is needed at the close of each day. Whether we use the HAR model or the RNN model will not impact the daily operation of such a system

### 3.6 Conclusion

This chapter analyses the potential of a heterogeneous autoregressive model including jumps to forecast realized volatility (RV). This approach computed RV based on a 20-year history of 5-minutes intraday data for the S&P500 index. Our results show that the base HAR-RV-J model is indeed able to provide a satisfactory forecast of RV. This outlined not only by the statistical error measures, but also by an analysis of trading efficiency based on the SPY ETF, the VIX index and the VXX ETN. Using our approaches attractive Sharpe Ratios can be obtained that outperform a common benchmark.

Our analysis also includes a Recurrent Neural Network (RRN) that uses the same inputs as the linear base HAR-RV-J model for a comparison of performance. These inputs are daily, weekly, and monthly RVs, plus the jumps. Finally, a hybrid model is built that additionally feeds the linear forecast to an RNN.

The results of all three model types are of similar quality. However, it has been noticed that

the RNN models are able to achieve these results with a shorter input time frame. That means, that when historical data is scarce, we can rely on an RNN to still deliver robust performance. Additionally, the RNN errors are uniformly low, while the errors for the linear model only reduce, once an input time frame of 1 or 2 years is used. Finally, it has been observed that the results between HAR-RV-J and RNNs do not differ too much, and attractive Sharpe ratios are obtained for trading a volatility derivative. The present work is, in our view, just a starting point to analyse the trading efficiency of intraday RV models.

## Chapter 4

# An Intraday Risk Management Support System Based on Volatility

### 4.1 Introduction

Modelling and forecasting financial volatility is one the most challenging task for both financial practitioners and econometricians. An accurate volatility forecast would have to be linked to a corresponding trade in the underlying, an analysis of volatility can provide decision support for any regulatory body involved with trading and pricing financial assets. A long list of Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) and Stochastic Volatility (SV) type formulations have been drastically investigated in the literatures for estimating financial market volatility. These specific parametric models are not appropriately suited for applications when estimated directly with intraday data, which is now available for many financial assets. In response to more effectively exploit the information inherent in high-frequency data, Andersen and Bollerslev (1998) suggested a new measure, Realized Volatility which has been discussed in the previous chapter. Andersen et al. (2003) originally suggested the use of autoregressive fractionally integrated (ARFIMA) models for capturing the highly persistent property in volatility. In contrast to ARFIMA models, long-memory heterogeneous autoregressive (HAR) models which is proposed by Corsi (2009) have simpler structure and easier to estimate. Empirically this approach has ar-



guably emerged as the preferred specification for modelling and forecasting realized volatility when high frequency data are available.

Following the work of Liu et al. (2017) on realized volatility of large indices, the motivation of this study is to investigate the trading performance of HAR type models developed by Andersen et al. (2007a), our empirical analysis relies on 5-minute high-frequency data and corresponding realized volatility measures for the EURO STOXX 50 index, EURO STOXX 50 Volatility (VSTOXX) index, RUSSEL 2000 index, iShare RUSSEL 2000 index (ETF) and RUSSEL 2000 volatility index (RVX). Over time, numerous studies have been applied to the task of forecasting daily implied volatility on large stock indices, still the focus has largely remained with the S&P 500 index, therefore this study contributes to the literature by comprehensively evaluating multiple-step-ahead volatility forecast of the most popular models with different large indices, also including the implied volatility index. Additionally, this study not only evaluate the forecasting performance by statistical errors but also by financial criteria. Our results show that even after accounting for transaction costs the presented models are able to produce forecasts that are useful in the context of a very simple trading strategy. Such a strategy can be used on its own to diversify return streams. However, in the context of risk management support there is also a second alternative that it has been outlined: it is possible to use the forecast as a dynamic hedging tool. Indeed, indiscriminate hedging of a portfolio is, generally, an expensive proposition. With a realized volatility forecast we can limit the hedge to days when volatility is forecast to rise. This significantly reduces the cost of the hedge.

The rest of this chapter is structured as follows. Section 4.2 is a literature review of previous research on volatility forecasts, especially the application of HAR models. Afterwards, it presents the methodology. Specifically, how the different components of the HAR model (including jumps)

come together to form a simple, yet effective modelling approach to realized volatility. Section 4.4 outlines the intraday datasets which have used to estimate our models. Section 4.5, finally, presents results in three different ways. Firstly, it presents pure statistical error measures. Secondly, it analyses the economic significance of our results. Thirdly, some ways to improve hedging on equity portfolios have been suggested. The last section summarizes and discusses our results.

## 4.2 Literature Review

According to the literature there are diverse quantitative models and approaches have been developed and refined to address the importance of volatility forecasting on risk management, asset pricing and asset allocation (Christoffersen and Diebold (2000)). In the high frequency environment, Heterogeneous AutoRegressive (HAR) model was is widely used for the prediction of RV, which has a simple structure and also capture stylized facts of volatility such as long memory and multiscaling behaviour, Corsi (2009) showed a detailed theoretical results. It is difficult to draw clear conclusions from the existing literature of which model is the best as research designs vary considerably in terms of countries, assets classed, time periods, forecasting horizon and forecast evaluation methods, therefore Kourtis et al. (2016) overcome this difficulty by comparing some of the most popular volatility models within a common framework, their study suggested that at the daily horizon, the HAR model offered the most accurate prediction. More analysis of the basic framework of HAR model can be found at Ma et al. (2014); Bollerslev et al. (2016). Based on Corsi (2009)'s work, researchers have made much effort to improve this model for analysing and forecasting realized volatility, therefore, different implementation have been applied to HAR models (Bandi et al. (2013); Souček et al. (2013); Haugom et al. (2014); Sevi (2014); Huang et al. (2015); Tian et al. (2017); Cubadda et al. (2017)). Andersen et al. (2007a) proposed an alternative model

which allowing jumps or discontinuities in the estimation of RV, which referred as HAR-RV-J in this thesis. Their empirical studies show that incorporating the jumps to the HAR model increase the accuracy of forecasting performance. This has also been proven experimentally, see Andersen et al. (2007b); Corsi et al. (2010); Celik and Ergin (2014); Liu et al. (2016); Papavassiliou (2016). However, some researcher also argued that when jump component are zero or insignificant, the incorporation of a jump component in to HAR may lead to overfitting, see Rossi and Sekhposyan (2011). Additionally, Prokopczuk et al. (2016) characterized the dynamics of jumps and analysed their importance for volatility forecasting in four leading energy markets and investigated several HAR models which explicitly capture the dynamics of jumps, they establishes that explicitly modelling jumps does not significantly improve forecast accuracy. Busch et al. (2011) used implied volatility as an additional forecasting variable in HAR model and found that implied volatility contained incremental information about future volatility in foreign exchange, stock, and bond markets, for application of HAR on implied volatility, can also be found at Fernandes et al. (2014); Psaradellis and Sermpinis (2016). Recently, by taking the time-varying property of the HAR model's parameter and volatility of RV, Wang et al. (2016) used a dynamic model averaging approach and shows more accurate forecast than individual model in both statistical and economic sense.

In the following sections, the specific methodology is presented and empirical data is used to test the model.

## 4.3 Methodology

### 4.3.1 Realized Volatility

We consider an  $n$ -dimensional price process defined on a complete probability space,  $(\Omega, \mathcal{F}, P)$ , evolving in continuous time over the interval  $[0, T]$ , where  $T$  denotes a positive integer. Following closely the setup of Back (1991); Andersen et al. (2003, 2007a)'s work, we assume that the logarithm of asset price,  $p(T)$ , follows the continuous-time semimartingale jump diffusion model:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + \eta(t)dq(t), \quad 0 \leq t \leq T, \quad (4.3.1)$$

The mean  $\mu(t)$  is a continuous and locally bounded variation process, the instantaneous  $\sigma(t) > 0$  is a cadlag stochastic volatility process,  $W(t)$  is a driving standard Brownian motion, the counting process  $q(t)$  is normalized such that  $dq(t) = 1$  corresponding to a jump at time  $t$  and  $dq(t) = 0$  otherwise,  $\eta(t)$  refers the size of the corresponding discrete jumps in the logarithmic price process. The probability of a jump occurring in the time interval  $dt$  is  $P[dN_t = 1] = \lambda_t dt$  where  $\lambda(t)$  is the jump intensity, which is possibly time-varying, but does not allow infinite activity jumps processes. The leverage effect is addressed in model 4.3.1 through possible dependence between  $\sigma(t)$  and  $W(t)$ , see Barndorff-Nielsen et al (2006a); Barndorff-Nielsen and Shephard (2006b). The quadratic variation for the cumulative return process can be expressed as a sum of a continuous and a discontinuous component:

$$QV_t = \int_{t-1}^t \sigma^2(s)ds + \sum_{t-1 < \tau_i \leq t} \eta^2(\tau_i). \quad (4.3.2)$$

where  $0 \leq \tau_1 < \tau_2 < \dots$  are jump times, In 4.3.2, quadratic variation is decomposed as integrated volatility plus the sum of squared jumps through time  $\tau_t$ , see Andersen and Bollerslev (1998); Andersen et al. (2001, 2003, 2006); Barndorff-Nielsen and Shephard (2001, 2002a,b).

Suppose that  $M + 1$  evenly spaced intra-period observations on a trading day  $t$ , if  $p_{t,j}$  is the logarithmic price at time  $t_j$ , then for the continuously compounded intra-period returns are  $r_{t,j} = p_{t,j} - p_{t,j-1}$  for the  $j^{th}$  intraday interval of day  $t$ . The *realized volatility* for the day  $t$  is defined by summing the corresponding  $M$  high frequency intradaily squared returns:

$$RV_t = \sum_{j=1}^M r_{t,j}^2, \quad t = 1, \dots, T \quad (4.3.3)$$

By the theory of quadratic variation, see Back (1991); Andersen et al. (2003), the realized variation converges uniformly in probability to the increment of the quadratic variation process as the sampling frequency of the underlying returns go to infinity, that is

$$\lim_{M \rightarrow \infty} RV_t = QV_t = \int_{t-1}^t \sigma^2(s) ds + \sum_{t-1 < \tau_t \leq t} \eta^2(\tau_t), \quad (4.3.4)$$

Thus, in the absence of jumps the realized variation is consistent for the integrated volatility. However, in order to separate the continuous variation and jump components, Barndorff-Nielsen and Shephard (2004) proposed the *Bipower Variation* (BV), which is a continuous component of  $QV_t$ :

$$BV_t = \mu_1^{-2} \left( \frac{M}{M-2} \right) \sum_{j=2}^M |r_{t,j-1}| |r_{t,j}|. \quad (4.3.5)$$

where  $\mu_1 = \sqrt{\frac{2}{\pi}}$  is the first moment of the absolute value of a standard normal random variable. In theory, a higher value of  $M$  improved precision of the estimators, but in practice, this also makes them more susceptible to market microstructure noises, such as bid-ask bounces, stale prices, measurement errors, etc., see Busch et al. (2011).

Combining the results in Eqs. (4.3.4) and (4.3.5), the contribution to the quadratic variation process due to jumps in the underlying process can be estimated by:

$$RV_{t+1}(\delta) - BV_{t+1}(\delta) \sim \sum_{t < s \leq t+1} \kappa^2(s). \quad (4.3.6)$$

To prevent the right hand-side of Eq. (4.3.6) from becoming negative, Andersen et al. (2007a) imposed non-negativity truncation on the jump measurements:

$$J_{t+1}(\delta) = \max[RV_{t+1}(\delta) - BV_{t+1}(\delta), 0]. \quad (4.3.7)$$

### 4.3.2 Volatility Forecasting Model

HAR-RV model is introduced by Corsi (2009), its simple structure enables it to parsimoniously capture the long-memory property of realized volatility as it combining historical estimates of realized volatility computed over various non-overlapping horizons:

$$RV_{t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \varepsilon_{t+1}, \quad (4.3.8)$$

$t = 1, 2, \dots, T$ .  $RV_t$ ,  $RV_{t-5}$  and  $RV_{t-22}$  mark daily, weekly (5 business days) and monthly (22 business days) RV, respectively. Weekly and monthly RV is calculated as:  $RV_{t,t+h} = h^{-1}[RV_{t+1} + RV_{t+2} + \dots + RV_{t+h}]$ ,  $h = 1, 2, \dots$

Andersen et al. (2007a) proposed the new HAR-RV-J model, which seeks to capture the dynamics of jumps. Daily HAR-RV-J model is expressed as:

$$RV_{t,t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \beta_J J_t + \varepsilon_{t,t+1}. \quad (4.3.9)$$

Logarithmic and standard deviation form of HAR-RV-J model is given by:

$$(RV_{t,t+1})^{1/2} = \beta_0 + \beta_D (RV_t)^{1/2} + \beta_W (RV_{t-5,t})^{1/2} + \beta_M (RV_{t-22,t})^{1/2} + \beta_J (J_t)^{1/2} + \varepsilon_{t,t+1}, \quad (4.3.10)$$

and

$$\log(RV_{t,t+1}) = \beta_0 + \beta_D \log(RV_t) + \beta_W \log(RV_{t-5,t}) + \beta_M \log(RV_{t-22,t}) + \beta_J \log(J_t + 1) + \varepsilon_{t,t+1}. \quad (4.3.11)$$

## 4.4 Dataset

In contrast to previous work that modelled realized volatility on index data and used tradable assets to actually check for the economic significance, it will also use liquid ETF data to calibrate and test our models, thereby treating an ETF as a substitute for the index. We are well aware of the limitations that this approach incurs, notably the fact that ETFs will have costs associated to them which might bias the results and also include tracking errors. However, firstly, it only uses the top liquid ETFs, which, also, tend to have very low costs associated with them. Secondly, using ETFs is a straightforward way to test our approach on assets, where index data is not readily available. A case in point can be found with commodity ETFs or ETNs which might track an asset price through appropriate positions in the corresponding futures, thereby creating a constant maturity synthetic future.

Our dataset consists of five-minute interval data trimmed down to the earliest common date. This ensures comparability of results over all time series. It only includes liquid hours in our computation. Especially, the highly volatile first half hour of the trading day is removed. The last datapoint that goes into the computation is the close of the five minute bar *before* the actual close price. Therefore for the typical trading day the first datapoint will be at 10h00 and the last datapoint at 15h55. This leads to a total of, typically, 72 datapoints per day. Half-trading days are excluded from the computation. This concerns, mostly, the days around exchange holidays.

Finally, especially in the context of intraday-data, errors in the dataset have to be taken account. What should be considered an error is, of course, open to debate. Nevertheless, some cleaning method has to be come up with. As we are always looking at whole days we can, without incurring any lookahead bias, consider all datapoints in our cleaning procedure. Specifically, as the ultimate

Table 4.1: **Summary statistics for STOXX50E at 5-minute frequency**

	$RV_t$	$RV_t^{1/2}$	$\log(RV_t)$	$J_t$	$J_t^{1/2}$	$\log(J_t + 1)$
Mean	0.0796	0.2415	-3.1274	0.0504	0.1919	0.0466
St.dev	0.1383	0.1460	1.0588	0.0897	0.1164	0.0675
Skewness	8.1769	2.5581	-0.0714	9.0537	2.6415	5.7493
Kurtosis	107.8345	15.2369	5.4801	134.7332	16.5035	57.3278
Min	2.1505e-06	0.0015	-13.0498	1.9513e-06	0.0014	1.9513e-06
Max	2.5554	1.5986	0.9382	1.8375	1.3555	1.0429

Table 4.2: **Summary statistics for V2TX at 5-minute frequency**

	$RV_t$	$RV_t^{1/2}$	$\log(RV_t)$	$J_t$	$J_t^{1/2}$	$\log(J_t + 1)$
Mean	1.0415	0.9395	-0.2705	0.6579	0.7468	1.6579
St.dev	1.2047	0.3987	0.7443	0.7446	0.3166	0.7447
Skewness	7.2627	2.3216	0.3559	6.5541	2.2153	6.5541
Kurtosis	97.5384	14.3636	3.6035	78.2062	13.0376	78.2062
Min	0.0779	0.2791	-2.5524	0.0535	0.2312	1.0535
Max	24.9652	4.9965	3.2175	14.1203	3.7577	15.1203

method will look at five-minute returns it filters out any intraday returns that deviate more than four standard deviations from the mean. This large interval is designed to ensure that only true errors are filtered out. Such errors may happen, for example, when missing values are filled with zeroes or errors in the decimal point occur. The cleaning procedure replaces filtered values with the return mean for the day.

It has to be emphasized, again, that such a procedure is only valid, if the incoming data does not feed any online (live, just-in-time) system. In this case, it needs to use an online cleaning procedure that has to decide immediately (i. e., when a data point comes in) whether this datapoint is valid. In the case of an error, it needs to find a replacement value immediately. Such a procedure introduces more complexity than our off-line cleaning.

Table 4.3: **Summary statistics for RUT at 5-minute frequency**

	$RV_t$	$RV_t^{1/2}$	$\log(RV_t)$	$J_t$	$J_t^{1/2}$	$\log(J_t + 1)$
Mean	0.0686	0.2174	-3.3447	0.0433	0.1731	0.0394
St.dev	0.1601	0.1462	1.0266	0.1004	0.1154	0.0702
Skewness	11.3601	3.6702	0.5079	12.2978	3.6530	7.1679
Kurtosis	203.3252	26.3608	3.9711	255.1476	26.8485	81.4107
Min	0.0016	0.0400	-6.4392	0.0010	0.0318	0.0010
Max	4.3795	2.0928	1.4769	3.0783	1.7545	1.4057



Table 4.4: **Summary statistics for IWM at 5-minute frequency**

	$RV_t$	$RV_t^{1/2}$	$\log(RV_t)$	$J_t$	$J_t^{1/2}$	$\log(J_t + 1)$
Mean	0.1324	0.3095	-2.5954	0.0821	0.2442	0.0720
St.dev	0.2635	0.1912	0.9375	0.1650	0.1501	0.1053
Skewness	9.2937	3.2974	0.7630	10.4352	3.3426	5.4084
Kurtosis	137.1739	21.1086	3.9764	184.5929	22.3532	47.6530
Min	0.0064	0.0802	-5.0459	0.0038	0.0618	0.0038
Max	5.7934	2.4070	1.7567	4.1949	2.0481	1.6477

Table 4.5: **Summary statistics for RVX at 5-minute frequency**

	$RV_t$	$RV_t^{1/2}$	$\log(RV_t)$	$J_t$	$J_t^{1/2}$	$\log(J_t + 1)$
Mean	1.1854	1.0019	-0.1472	0.7504	0.7991	0.5015
St.dev	1.4144	0.4262	0.7664	0.8853	0.3344	0.3049
Skewness	10.5690	2.3517	0.1611	11.5448	2.3502	1.7962
Kurtosis	223.6934	17.4388	3.5992	266.1316	18.3848	9.4204
Min	0.0515	0.2269	-2.9664	0.0340	0.1843	0.0334
Max	37.5251	6.1258	3.6250	24.6604	4.9659	3.2449

## 4.5 Empirical Results

### 4.5.1 Forecast Evaluation

The result table shows several tendencies across the board. Firstly, errors tend to reduce when using a longer lookback period. This is mostly in-line with the results of Liu et al. (2017). However, previous studies showed a flattening out of the error improvement for a lookback period of around one year, while we still clear improvements when switching from a lookback period of one year (252 days) to two years (504 days). Similarly (and not surprisingly) errors tend to increase for longer forecast intervals. It uniformly obtains the best results for a one-day ahead forecast,  $l = 1$ . This result also carries over to an analysis of the economic significance in the following section.

While Liu et al. (2017) only analyse one asset, namely the S&P500 the present study analyses two different stock indices (ETFs) and their corresponding volatility as an asset class on their own. This allows for an intra-asset class comparison. It has been noticed clearly, that the two volatility forecasts for the Vstoxx (V2TX, Eurostoxx volatility) and the Russel (RVX, Russell 2000 volatility) show much higher errors, generally. However, this is explained by the fact that volatility

Table 4.6: **Comparison of Models**

	RMSE	MAE	MAPE		RMSE	MAE	MAPE		RMSE	MAE	MAPE
$t = 22, l = 1$				$t = 22, l = 2$				$t = 22, l = 5$			
STOXX50E	0.3448	0.0693	572.6532		0.4870	0.0868	597.8710		0.6191	0.1214	861.8646
V2TX	3.6374	0.8862	103.4957		7.8930	1.1138	129.1592		7.6551	1.4234	171.8143
Russell2000	0.1816	0.0466	77.1585		0.2931	0.0598	97.1462		0.7212	0.0980	145.4468
IWM	0.3521	0.0874	69.6220		0.4515	0.1080	90.1878		0.8474	0.1733	137.6451
RVX	4.8255	0.9163	86.5296		4.5828	1.0871	105.2760		5.8713	1.4170	152.5510
$t = 63, l = 1$				$t = 63, l = 2$				$t = 63, l = 5$			
STOXX50E	0.2149	0.0541	656.9858		0.2120	0.0600	772.0017		0.3570	0.0789	772.7852
V2TX	1.8199	0.6696	86.5409		1.7369	0.7105	95.9589		1.5600	0.8101	112.4515
Russell	0.1532	0.0404	67.1957		0.2325	0.0488	80.9854		0.3214	0.0654	85.7798
IWM	0.3002	0.0739	57.9539		0.3015	0.0837	69.3099		0.4273	0.1120	91.5067
RVX	2.1446	0.6930	75.4161		1.5555	0.7418	81.5747		1.8149	0.8652	100.3236
$t = 126, l = 1$				$t = 126, l = 2$				$t = 126, l = 5$			
STOXX50E	0.1772	0.0500	488.6619		0.1407	0.0522	463.5794		0.1739	0.0607	608.4748
V2TX	1.3483	0.6039	78.2367		1.2188	0.6352	86.4477		1.2609	0.6727	93.2666
Russell	0.1370	0.0371	65.1924		0.2030	0.0438	77.2291		0.2007	0.0527	95.4863
IWM	0.2281	0.0668	55.2839		0.2608	0.0747	64.8087		0.3386	0.0936	80.7807
RVX	1.6360	0.6406	72.0678		1.4352	0.6850	77.4747		1.5412	0.7551	91.0955
$t = 252, l = 1$				$t = 252, l = 2$				$t = 252, l = 5$			
STOXX50E	0.1428	0.0454	481.6909		0.1265	0.0474	484.5718		0.1367	0.0520	553.1643
V2TX	1.2421	0.5816	75.9609		1.1878	0.6129	83.8503		1.2161	0.6381	89.5973
Russell	0.1262	0.0352	64.8112		0.1686	0.0406	75.0148		0.1530	0.0449	89.0268
IWM	0.1996	0.0627	56.1966		0.2416	0.0709	65.0165		0.2603	0.0815	78.1806
RVX	1.3888	0.6116	69.3871		1.3756	0.6540	74.3515		1.4554	0.6974	84.8808
$t = 504, l = 1$				$t = 504, l = 2$				$t = 504, l = 5$			
STOXX50E	0.1252	0.0440	501.4259		0.1212	0.0463	531.7771		0.1275	0.0494	678.1119
V2TX	1.1395	0.5723	75.4532		1.1674	0.6037	82.8873		1.1971	0.6302	89.2114
Russell	0.1133	0.0338	65.4136		0.1372	0.0380	74.4394		0.1350	0.0413	87.4047
IWM	0.1826	0.0598	55.7916		0.2118	0.0665	63.5606		0.2369	0.0765	77.6579
RVX	1.3142	0.5974	67.2449		1.3356	0.6313	72.6705		1.3840	0.6655	81.1810

indices feature themselves a much higher volatility than the corresponding underlying stock indices.

Therefore, forecast errors will naturally lead to a much higher impact on the statistical errors. This, however, does not affect the economic significance which is similar for stock indices and volatility indices as it will show in the following section.

#### 4.5.2 Economic Significance

In the previous paragraphs it has been looking at statistical error measures. These might provide a first suggestion on whether the proposed method performs well or not. However, when we model financial markets, we are in a good position to assess the actual usefulness of new results in a, typically, better way than pure statistical measures normally allow. Indeed, if we assume that our

method is a useful decision support tool we should be able to use the forecasts in the context of a quantitative trading strategy. Such a strategy will, algorithmically and systematically, take the forecast output and convert it into an actual trading decision. The most basic way to do this is to find a mapping that will result in being long or short a certain asset. That is, we will buy (or sell short) an asset. Optionally, such a strategy can be augmented by a third state, that we will call *flat*. That is, the strategy is out of the market.

In the present case it forecasts the volatility of an underlying asset. This asset is a stock index or, again, a volatility index, whose volatility is forecast (volatility of volatility). To generate our decisions, the following heuristic based on empirical stylized facts of financial markets are used:

- When it forecasts the RV of a stock index we will assume that a rising forecast RV will lead to a decline in the stock index. On the other hand, we will also assume that a declining forecast RV will be associated with a rising stock index. This builds on the commonly documented observation that rising volatility is a sign of insecurity and often associated with falling asset prices.
- The position for a volatility index is reversed. That means, a rising forecast RV (rising volatility of volatility) will be assumed to lead to a rise in the volatility index itself.

This leads to the following basic trading strategy. It shows the rules for a stock index:

- if  $RV_{t+1} > RV_t$  then short the asset
- if  $RV_{t+1} < RV_t$  then go long the asset
- else stay out of the market.

These rules would be reversed for a volatility index. Furthermore, transaction costs are assumed

of 0.1% or 10bp per roundtrip. Such transaction costs are relatively low, but justified by the high liquidity of the assets that we are trading. Please note, that in the case of the Eurostoxx the trading simulation is actually carried out on Eurostoxx futures, in the case of the Vstoxx, Vstox futures are used. The Russel index is traded through the IWM ETF, and Russel volatility is traded through the corresponding futures.

Therefore, after transaction costs it obtains the results outlined in the following tables. It presents results for different combinations of lookbacks (training data) and look-aheads (how many days in advance do we forecast). We can see a clearly positive result for the one-day ahead forecast. For the two-day and five-day ahead forecasts the results are more mixed. Nevertheless, we shouldn't forget that the HAR-RV-J model was initially intended to be used for one-day forecasts. Insofar, we may expect the highest consistency for this forecast horizon.

Raw returns in themselves are not very useful in assessing the applicability of a trading strategy. Rather, we should look at risk adjusted returns. We take a very simple measure of risk-adjusted returns with the information ratio. We divide the annualized return by the annualized volatility. Ideally, we want the returns to be high and the volatility to be low. The information ratio is similar to the Sharpe ratio, which additionally takes current risk-free rates into account. However, these are essentially zero currently. Our results are in-line with Liu et al. (2017) which shows similar results for the S&P500 index. The present study analyses the usefulness of the model away from the asset that it has most been used for. Interestingly, even with the currently presented naive strategy, we manage to obtain acceptable risk-adjusted returns.

However, compared to studies on the S&P500 the consistency is slightly reduced for longer forecast horizons. This is to be expected, but should nevertheless give rise to the suggestion to use

this model mainly for one-day ahead forecasts and trading as initially intended. Two extensions of the basic strategy, naturally arise: Firstly, we may use a threshold, that would reduce trading to cases where the model suggests a significant change in RV. This, of course, introduces a new parameter and requires to define (and possibly optimize) the threshold. Secondly, we suggest to apply a dynamic leverage approach. Such an approach would analyse past strategy return volatility and scale the strategy exposure up and down accordingly.

The present chapter is in many ways an extension of the work in the previous chapter with a focus on intraday risk management. It makes sense to add the previously found modelling results on the S&P500 to the discussion. In many ways the promising robust results on the S&P500 can also be found in the new dataset. However, we might notice that, overall, the results are a bit less stable. Also, on average, the trading performance might be interpreted as slightly less robust. Nevertheless, we don't see any reason to reject the models as a useful addition to the quantitative analyst's toolbox because of that. Rather, a more in-depth analysis of the circumstances, under which some model might work better than another seems warranted. Combining this with additional results from the following chapter (energy commodity) seems promising and is left for further research.

### **4.5.3 Risk Management Support**

The previous section has outlined economic significance of the forecast using a trading strategy. While this may be the most direct way to assess the usefulness as a tool for tackling financial markets, another application is in the context of risk management decision support. Indeed, the owner of a portfolio of risky assets will have to assess risk on a continuous basis. Portfolio managers measure risk in different, but mostly similar, ways. One of the primary goals of successful risk management is to mitigate downside volatility. In the context of a long equity portfolio this entails

Table 4.7: Out-of-sample rolling window trading results for the HAR-RV-J model

Strategy	Ann. Ret.	Ann. Vol.	Info. Ratio	Ann. Ret.	Ann. Vol.	Info. Ratio	Ann. Ret.	Ann. Vol.	Info. Ratio
$t = 22, l = 1$				$t = 22, l = 2$			$t = 22, l = 5$		
STOXX50E	0.0385	0.2307	0.1669	-0.0301	0.2307	-0.1305	0.0471	0.2307	0.204
V2TX	0.4689	1.0277	0.4563	-0.3784	1.0278	-0.3682	-0.0241	1.0281	-0.0234
IWM	0.1329	0.2926	0.4541	-0.03	0.2927	-0.1025	0.0464	0.2927	0.1585
RVX	0.3691	0.9446	0.3908	0.6317	0.944	0.6692	0.2721	0.9447	0.2881
$t = 63, l = 1$				$t = 63, l = 2$			$t = 63, l = 5$		
STOXX50E	0.0387	0.2307	0.1676	-0.0344	0.2307	-0.149	0.0829	0.2306	0.3596
V2TX	0.5404	1.0276	0.526	0.2136	1.028	0.2078	-0.0192	1.0281	-0.0187
IWM	0.0849	0.2927	0.2901	0.0121	0.2927	0.0413	0.0087	0.2927	0.0296
RVX	0.2679	0.9447	0.2836	0.6708	0.9439	0.7107	0.5465	0.9443	0.5787
$t = 126, l = 1$				$t = 126, l = 2$			$t = 126, l = 5$		
STOXX50E	0.0483	0.2307	0.2093	-0.0419	0.2307	-0.1817	0.0525	0.2307	0.2274
V2TX	0.5785	1.0275	0.5631	0.327	1.0279	0.3181	0.2135	1.028	0.2077
IWM	0.0989	0.2926	0.3379	-0.0215	0.2927	-0.0736	-0.0124	0.2927	-0.0424
RVX	0.3674	0.9446	0.3889	0.5055	0.9443	0.5353	0.4194	0.9445	0.4441
$t = 252, l = 1$				$t = 252, l = 2$			$t = 252, l = 5$		
STOXX50E	0.0691	0.2307	0.2995	-0.035	0.2307	-0.1518	0.1395	0.2305	0.6052
V2TX	0.4737	1.0277	0.461	0.1968	1.028	0.1915	0.1611	1.0281	0.1567
IWM	0.0945	0.2927	0.3229	-0.0169	0.2927	-0.0576	-0.0168	0.2927	-0.0572
RVX	0.2547	0.9447	0.2696	0.3307	0.9446	0.3501	0.4375	0.9448	0.4632
$t = 504, l = 1$				$t = 504, l = 2$			$t = 504, l = 5$		
STOXX50E	0.0713	0.2307	0.3091	-0.0104	0.2307	-0.0452	0.1062	0.2306	0.4607
V2TX	0.5299	1.0276	0.5157	0.266	1.028	0.2588	0.1286	1.0281	0.1251
IWM	0.0906	0.2927	0.3096	-0.0277	0.2927	-0.0945	-0.0645	0.2927	-0.2202
RVX	0.2335	0.9448	0.2471	0.3487	0.9446	0.3692	0.3273	0.9447	0.3464

to control (or hedge) downside risk.

To achieve this aim it makes sense to consider the correlation between equity return and corresponding volatility. As is well-known on the financial markets daily returns of equities tend to be strongly negatively correlated. Or put in another way: on days where equities suffer large draw-downs, volatility will tend to spike up. On the other hand, on days where equities exhibit positive returns, volatility tends to go down. There is, however, an asymmetry in how this reaction will turn out. Volatility spikes are, generally, abrupt and short-lived. In calm periods, volatility will slowly fall.

We can use this tendency to hedge an equity portfolio by going long volatility. The long volatility position will generally lose a bit of its value every day (negative roll-yield, or contango of the volatility term structure). However, on days with large negative returns, volatility will spike up and therefore provide protection against sudden drops in the equity market. While this method is reasonably effective at hedging an equity portfolio it has one very significant drawback:

being long volatility is a losing proposition. For example, the most popular and most liquid long volatility ETF (VXX) has lost more than 99% of its value since its inception. We can think of this phenomenon as an insurance: being long volatility means we pay an insurance premium (to the opposite party of the trade, being short) to protect ourselves against tail risk. Like for every (well-designed) insurance we will, on average, lose money by paying the insurance premium. On the other hand, the seller of the insurance will tend to make money.

The whole picture would change, if we only had to pay the insurance premium in times that are perceived as risky. For this task a volatility forecast comes in handy. We would then only be long volatility when volatility is forecast to rise. For our analysis this involves only approximately 2% of the time under consideration. We therefore only pay insurance premium during times that are perceived as high-risk. Of course, we have to be aware that this is not as good (in terms of risk mitigation) as having a full-time insurance. Indeed, our forecasts are far from perfect and we still run the risk to be hit by an unforeseen event. However, preliminary studies show that this method has the potential to cut drawdowns by around 25 percentage points. This hedge is also easy to implement, as the only necessity is to buy the corresponding future. It is also not expensive, as, on average, the hedge will only be active a few business days per year.

A further application in the context of risk management support may also be found in the context of hedging specific assets against large transactions, as in Kurek (2016). Here, assuming information leakage, an appropriate model of realized volatility could help in safeguarding non-insider investors against block transactions.

## 4.6 Discussion and Outlook

This chapter analyses the forecastability of realized volatility for the base stock market indices Eurostoxx and Russell 2000. For the Eurostoxx we use index values and the corresponding volatility Vstoxx. For the Russell 2000 we use index values, the corresponding ETF IWM and the volatility index RVX. Using index data or ETFs as a base for estimating the model or evaluating the economic significance both provide individual challenges.

Index data provides an unbiased assessment of the current state of the market and is, arguably, harder to manipulate. Therefore, by calibrating our models on index data and by evaluating the economic significance on this same index data we use the *true* price. On the other hand, we have to be always aware of the fact that indices are, in themselves, not tradable assets. Only derivatives of the index data (in the broadest sense of the word) can be used to actually perform a trade, be it for the purpose of generating a profit or for hedging purposes. Considering our other asset class, ETF data always show tradable prices. This is of significance when we want to trade the assets. But, we have to keep in mind, that ETF data may be subject to stronger biases than index data, due to any of the following reasons. An obvious interference is the presence of costs. In our case these costs are low. Additionally, the price of the ETF may be biased, especially on an intraday basis, just because of changing demand and supply dynamics. This effect will be of very minor importance when considering only daily ETF data. At the close of trading the issuer or special market participants will establish a uniform price. However, during the day significant deviations from the corresponding index may occur. Calibrating a model on this data may lead to its own challenges.

To summarize, we show that the HAR model does a decent job at modelling realized volatility.



Results, even as measured as economic significance, are very satisfactory. Future work will involve extending the analysis to other assets outside the world of equities. Also, we will use different setups of our methods and add additional methods.

## Chapter 5

# Analysing and Modelling High Frequency Realized Volatility of Energy Commodities

### 5.1 Introduction

Energy prices have important effect on macro-economics and financial markets (Regnier (2007)). The price of energy commodity are more volatile than prices of non-energy products is likely due to the supply uncertainty (Susmel (1997); Regnier (2007); Lin and Zhu (2004); Wilson (1996)), therefore, modelling and forecasting the volatility of energy commodities is crucial for resource allocation, speculation, risk management, and real option valuation. The traditional GARCH-type models where the volatility is assumed to be latent are commonly used for conditional volatility prediction. However, jumps are generally not taken into account with these models. Recent developments in the econometrics literature have shown that it is possible to use non-parametric estimation of price volatility and jumps. In response to the increasing availability of high frequency data, Andersen and Bollerslev (1998) proposed the realized volatility (RV) which defined as the sum of squared intraday returns, this measure includes both jumps and non-jumps and allow the volatility to be treated observable. Later Barndorff-Nielsen and Shephard (2004) proposed the bipower variation (BV), which defined as the sum of the product of adjacent absolute intraday

returns, that is considered as a consistent estimation of non jump component volatility. Therefore, the discontinuous jumps can be constructed by the difference of RV and BV. Among the models of forecasting realized volatility, the parametric model termed Heterogeneous AutoRegressive model of Realized Volatility (HAR-RV) is commonly used as it can captures some of the volatility properties such as long memory and multiscaling in a very simple and parsimonious way (Corsi (2009)). In this chapter we used an alternative model which is an extension of HAR-RV model incorporating jumps or discontinuities, that is referred as HAR-RV-J (Andersen et al. (2007a))

However, the errors in prediction by using the parametric models are often argued by researchers. This is because the linear models are often based on certain distribution assumptions and the microstructure noise can rise by bid-ask bounce, asynchronous trading, and price discreteness. Artificial Neural Network model is considered as an alternative non-parametric approach as it is able to tolerant data with errors and find nonlinear associations between parameters. In this chapter, we test the ANN against the most widely used HAR-RV-J model, and also combine the HAR-RV-J forecast with ANN to build a hybrid model.

Our work contributes to the literature in three aspects. Firstly, we analyse energy commodities which do not have their own (implied) volatility futures to trade. This means, that a good forecast of implied volatility cannot directly be exploited. Rather, a good forecast would have to be linked to a corresponding trade in the underlying. It should therefore be of interest to any regulatory body involved with trading and pricing energy commodities, to have a good model of intraday realized volatility. Secondly, the studies on hybrid model which incorporates linear model and neural network to forecast energy commodities volatility are limited, also we target on high frequency intraday realized volatility that bring us closer to the true volatility process. Thirdly, to evaluate

the forecasting performance, we omit the traditional statistical accuracy criteria, but apply realistic volatility trading strategy which gives the economic point of view. To the best of our knowledge, this is the first attempt to comprehensively test this trading strategy based on the new approach of volatility forecasting by using high frequency data in energy literature.

The remainder of this chapter is structured as follows. We begin with a compact literature research. After this, we outline our methods and data. This includes presenting the HAR model including jumps, recurrent neural networks and the hybrid model. Also, we touch on data preprocessing and data cleaning. Finally, we present the results. The last section concludes and outlines potential for further research.

## 5.2 Literature Research

The existing literature mostly focused on finding the most accurate forecasting models by considering characteristics of volatility of energy commodities which include volatility spillover (Du et al. (2011); Lin and Tamvakis (2001); Mensi et al. (2013); Sadorsky (2012); Souček et al. (2013); Wang et al. (2008)), jumps (Ciarreta and Zarraga (2016); Qu et al. (2016); Chevallier (2012)), leverage effect (Charfeddine (2014); Doran and Ronn (2008)), volatility persistence (Fujihara and Mougoue (1997); Cunado et al. (2010)), etc.. Early research used generalized autoregressive conditional heteroskedastic (GARCH) model (Agnolucci (2009)) and its various extensions to model volatility in energy markets (Fong and See (2002); Hung et al. (2008); Kang et al. (2009); Sadorsky (2006)). The advent of high frequency data which contains more intraday trading information has allowed the volatility to be treated as an observable variable, the first attempt to apply the concept of RV on electricity data was carried out by Chan et al. (2008), later application can also be found by Haugom et al. (2011). The most widely used parametric model for the prediction of RV is

called Heterogeneous AutoRegressive (HAR) model, which was developed by Corsi (2009), captures volatility persistence by the sum of the heterogeneous components in the financial markets, this parsimonious model is considered in which volatilities are realized over different horizon. The simple structure and superior forecasting performance made the HAR model popular for the RV estimation, recently, different implementations have been applied to HAR model by incorporating other exogenous variables, for example, the jump component, (Andersen et al. (2007a); Corsi et al. (2010); Da Fonseca et al. (2016)). Sevi (2014) investigated the realized volatility forecast of crude oil by using intraday data, they extended the HAR model by considering various components, such as continuous and discontinuous part, leverage effects, and positive and negative returns, their results showed that these sophisticated models including jumps and other components have not improved better out-of-sample forecasts than the genuine model by Corsi (2009). Additionally, Prokopczuk et al. (2016) used high-frequency data on four prominent energy markets, and investigated the importance of jumps in energy market by employing various extended HAR models, their results indicated that explicitly modelling jumps does not significantly improve forecast accuracy in energy markets

Due to the development of the new techniques in artificial intelligence, non-parametric models have been used extensively on financial forecasting. Compared to parametric models, Artificial Neural Network (ANN) has the advantage of being able to detect non-linear relationship in the presence of noisy information (Jammazi and Aloui (2012); Panella et al. (2012); Papadimitriou et al. (2014)). In the literature of volatility forecasting and trading, ANN have been argued by several researchers, see Haigh et al. (2004); Dunis and Chen (2005); Sermpinis et al. (2013); Xiao et al. (2014); Hemanth Kumar and Patil (2015); Tung and Quek (2013). In the academic literature, the

applications of ANN in energy market have focused mainly on the forecasting prices (Fan et al. (2008); Xiong et al. (2008)), research using ANN to forecast volatility continued to be developed, among the few that do, the pioneering work has been done by Barunik and Krehlik (2016), who comprehensively evaluate multiple-step-ahead volatility forecast of energy market using HAR and ANN models, their results indicated that ANN yields both statistical and economic gains, and coupling ANN with high frequency data result in substantial reduction in the over-estimation tendency. Additionally, the so called Hybrid models which combined parametric models and ANN have been developed (Donaldson and Kamastra (1997); Bildirici and Ersin (2009); Park et al. (2014); Lahmiri (2015)). Whereas researchers mainly apply these hybrid models on stock market index (Kristjanpoller et al. (2014); Hajizadeh et al. (2012); Monfared and Enke (2014); Araujo et al. (2015)), exchange rate (Dunis and Huang (2002)), and non-energy commodity (Kristjanpoller and Minutolo (2015)).

Furthermore, a significant body of literature exists on the context of modelling volatility in the context of trading financial assets. This includes, more generally, for example, Haigh et al. (2007). On the other hand Herbert (1995) focus more specifically on a single energy commodity. As trading of energy commodities mostly occurs through futures (at least if we are trading financial assets and not the physical asset) we have to take maturity effects into account, see Serietis (1992). Another more recent example can also be found in Walls (1999) and Weiner (2002).

### 5.3 Method and Data

This section firstly provides a brief overview of quadratic variation theory, detailed theoretical explanation can be found in Barndorff-Nielsen and Shephard (2004). Then we introduce the competing models. Finally, we present our dataset.

### 5.3.1 The HAR Model

Let  $p(t)$  denote the logarithm of an asset price, and assume that the price process is governed by a continuous-time stochastic volatility jump model:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)q(t), \quad (5.3.1)$$

the mean  $\mu(t)$  is assumed continuous and locally bounded, the instantaneous volatility  $\sigma(t) > 0$  is càdlàg,  $W(t)$  is a standard Wiener process, and  $q(t)$  is a Poisson counting process, with  $q(t) = 1$  if there is a jump at time  $t$  and zero otherwise.  $\kappa(t)$  represents the corresponding size of jumps at time  $t$  if  $q(t) = 1$ , the intensity of the arrival process for jumps,  $\lambda(t)$ , is possibly time-varying, but does not allow infinite activity jump processes. The leverage effect is accommodated in (5.3.1) through possible dependence between  $\sigma(t)$  and  $W(t)$ . The quadratic variation of the cumulative return process is given by:

$$QV_t = \int_0^t \sigma^2(s)ds + \sum_{s=1}^{q(t)} \kappa^2(s). \quad (5.3.2)$$

In the above equation, quadratic variation is decomposed as integrated volatility plus the sum of squared jumps at time  $t$ .

Suppose that the price is observed at discrete times  $j = 1, 2, \dots, M$  within each day  $t = 1, 2, \dots$ , and let  $r_{t,j} = p_{t,j} - p_{t,j-1}$  be the  $j$ th continuously compounded intra-period return of day  $t$ . The realized volatility for period  $t$  is given by the sum of squared intra-period returns,

$$RV_t = \sum_{j=1}^M r_{t,j}^2, \quad t = 1, 2, \dots, T, \quad (5.3.3)$$

The theory of quadratic variation indicates that the realized volatility converges uniformly in probability to the quadratic variation as the sampling frequency increase, that is:  $RV_t \rightarrow QV_t$  for  $M \rightarrow \infty$ . In reality, a higher frequency leads inevitably to a larger microstructure noise, therefore,

five-minute sampling frequency is often chosen for active financial markets as a bias-variance tradeoff.

Barndorff-Nielsen and Shephard (2004) introduce a related concept known as bipower variation, defined as the sum of the product of adjacent absolute intraday returns:

$$BV_t = \mu_1^{-2} \left( \frac{M}{M-i} \right) \sum_{j=1}^{M-i} |r_{t,j}| |r_{t,j+i}| \rightarrow \int_{t-1}^t \sigma_s^2 ds, \quad M \rightarrow \infty \quad (5.3.4)$$

where  $\mu_1 = \sqrt{(2/\pi)}$  is the first moment of the absolute value of a standard normal random variable,  $i \geq 1$  is the lag length in the multiplication of absolute intraday returns.

The quadratic variation is the sum of continuous and discontinuous components, therefore, equation (5.3.2) can be expressed by:

$$RV_t - BV_t \rightarrow \sum_{j=1}^{q(t)} \kappa^2(s) \quad (5.3.5)$$

To prevent the right hand-side of equation (5.3.5) from becoming negative, Andersen et al. (2007a) imposed non-negativity truncation on the jump measurements:

$$J_t = \max[RV_t - BV_t, 0]. \quad (5.3.6)$$

Our benchmark econometric model is the HAR-RV-J, the simple structure of this model enables it to parsimoniously capture the long-memory behaviour of realized volatility. This is achieved by combining historical estimates of realized volatility computed over various non-overlapping horizons. The structure of our estimating model is as follows:

$$RV_{t,t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \beta_J J_t + \varepsilon_{t,t+1}. \quad (5.3.7)$$



Logarithmic and standard deviation form of HAR-RV-J model is given by:

$$(RV_{t,t+1})^{1/2} = \beta_0 + \beta_D(RV_t)^{1/2} + \beta_W(RV_{t-5,t})^{1/2} + \\ \beta_M(RV_{t-22,t})^{1/2} + \beta_J(J_t)^{1/2} + \varepsilon_{t,t+1},$$

and

$$\log(RV_{t,t+1}) = \beta_0 + \beta_D \log(RV_t) + \beta_W \log(RV_{t-5,t}) + \\ \beta_M \log(RV_{t-22,t}) + \beta_J \log(J_t + 1) + \varepsilon_{t,t+1}.$$

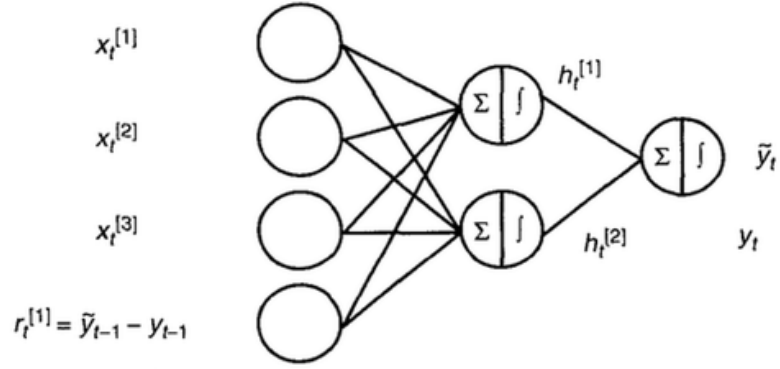
Each component in the HAR-RV model is computed over different horizon,  $RV_t$ ,  $RV_{t-5}$  and  $RV_{t-22}$  mark daily, weekly (5 business days) and monthly (22 business days) RV, respectively. Weekly and monthly RV is calculated as:  $RV_{t,t+h} = h^{-1}[RV_{t+1} + RV_{t+2} + \dots + RV_{t+h}]$ ,  $h = 1, 2, \dots$

### 5.3.2 Recurrent Neural Networks

In this chapter, we use *Recurrent Neural Network* (RNN) models which were introduced by Elman (1990), see also an application of RNNs in currency trading by Dunis and Huang (2002). Their only difference from multilayer neural network is that they include a loop back from one layer, either the output or the intermediate layer or the input layer. Figure 5.1 shows a single output RNN model with one hidden layer and two hidden nodes. The detailed discussion of ANN has been shown in Chapter 3.

In our empirical application, we choose the logistic function as squashing function, that is better adapted to the output domain of the numbers we analyse. However, the generated data is typically relatively small compared to unity. The typical order of magnitude is between  $10^{-5}$  and  $10^{-4}$ . In this case, a simple linear transformation is advised that makes better use of the available output range of the logistic function. To achieve this, we simply multiply all data with  $10^3$  for the linear

Figure 5.1: Single output Recurrent Neural Network (RNN) model with one hidden layer



HAR-RV-J model inputs. This simple linear transformation does not change the basic interpretation of the data but makes it easier to be learned by an ANN. The inputs to our RNN model are the same as to the linear model to allow a fair comparison. That is, we include the three RV and one jump inputs. In all our RNN models we determine the number of hidden neurons  $h$  in the following way:

$$h = 2 \times \sqrt{i \times o}, \quad (5.3.8)$$

where  $i$  and  $o$  refer to the size of the input and output layer, respectively. Note, that in the case of the input layer we do not take the bias neuron into account.

### 5.3.3 Hybrid Model

The hybrid model is also designed as an RNN. However, as an additional input we feed the *forecast* of the linear model to the RNN. We also keep our four basic inputs. Thus, the total number of inputs rises to five in the case of the hybrid model. All other model parameters are kept the same. Specifically, the number of hidden neurons is determined as above. Also, the model architecture stays identical.

The hybrid model shows increased complexity. We are combining the linear HAR model including jumps with the non-linear recurrent neural network. This recurrent network now uses five

inputs: the four basic inputs (three lags of realized volatility and the jump component) plus the output of the HAR model. It is, therefore, not very straightforward to set up. Such a model will not easily be available in an analyst's toolbox and will need to be built *by hand* in most cases.

The application of hybrid models is quite popular, see for example Araujo et al. (2015); Bildirici and Ersin (2009); Donaldson and Kamastra (1997); Hajizadeh et al. (2012); Monfared and Enke (2014); Park et al. (2014). Often, artificial neural networks are combined with models from the GARCH family. They are applied to the domain of forecasting stock market returns or stock market returns. We introduce models that combine a neural network with the HAR model.

#### 5.3.4 Data

We use five minute intraday data that spans from May 2007 to May 2017. This data is collected on the front-month future. When using future data it is important to take care of building a continuous future series out of the independent maturities. In the present case, the front month future is rolled over to the second month future, once the trading volume in the second month future has overtaken the trading volume in the first month future. This insures adequate liquidity in the asset that is modelled. Additionally, an additional analysis is provided excluding the data from the *crisis period* in 2008 and 2009. Therefore the *after crisis* results will be presented starting with data since 2010. The expectation is that excluding the very volatile periods in 2008 and 2009 should lead to results with better economic significance.

If we only consider returns on daily data, then rolling over the future poses a problem: naively doing so introduces an artificial jump when the future is rolled over. However, in the present case, this is of lesser importance. We only compute intraday five-minute returns. For any given day these returns will be computed off the *same* future. No artificial return jump is introduced. We

might, however, argue, that the future into which we roll over behaves somehow differently than the current front month future. Therefore, we would introduce an irregularity into the realized volatility time series. Alternatively, progressive weighting schemes could be explored. This introduces further complexity and is left to further research. It is questionable, whether such a process would indeed improve anything. We prefer to stick with an actually tradable asset like a future as opposed to a synthetically weighted portfolio of futures which would be much more cumbersome to trade. Nevertheless, it has to be outlined, that such constant maturity futures are synthetically implemented in Exchange Traded Funds. This might also be an alternative path to explore.

### **5.3.5 Data Preprocessing**

When computing a daily realized volatility out of intraday data the question arises, what exactly we should consider as a day. Most futures in the USA actually trade Monday till Friday, almost round the clock. Typically, a short trading break of fifteen minutes occurs in the afternoon. We have opted to stay with the methodology adopted in previous studies. We only take data from the liquid trading hours, which coincide with the liquid hours of the stock market. Specifically, we take five minute data from 10am to 4pm EST. This amounts to six hours or 72 five minute intervals. Actually, the trading day starts already at 9.30 am. However, as is usually done, we discard the first half hour as this is quite volatile. We don't want the first half hour volatility to bias our volatility computation for the whole day. Days that do not include 72 five minute data-points between 10am and 4pm are excluded. This takes care of removing half-holidays. Indeed, otherwise, we would have to normalize volatility. It is questionable and open to further research anyway, whether half-holidays would exhibit the same behaviour with respect to realized volatility, as typical full trading days.

Lastly, we have to take the possibility of errors in our intraday data into account. Due to the sheer amount of data, it is an almost hopeless task to filter out errors manually, just by eyeballing the numbers or the chart. Therefore we adopt the following data cleaning procedure. This procedure takes specifically into account, that the input to the realized volatility computation is returns and not raw prices. Therefore our data cleaning procedure acts on the raw prices:

- Sort the intraday returns by size
- Remove the smallest three and the largest three returns
- On the remaining returns, compute mean and standard deviation
- Replace all returns that are outside of five standard deviations from the mean by the mean return.

This cleaning procedure will keep most if not all legitimate returns, but filter out obviously wrong returns that may result from zero values in the intraday data. Note, that  $+\infty$  and  $-\infty$  are valid values in our programming environment and will get filtered out in the second step above. In the above procedure, we use the returns for an entire day. As we only compute realized volatility at the end of the day, this procedure is valid and does not introduce a look-ahead bias. We are not performing any kind of intraday online algorithm. Thus, taking the entire return set doesn't cause a problem.

## 5.4 Results

This section will focus on providing an *economic* interpretation of our results. Whereas often in forecast exercises we present statistical error measures, this often leads to neglecting the actual potential (or lack) of the method. For this reason we will omit any statistical error measures.

These are available upon demand and rather put the results in the context of a risk management strategy.

### 5.4.1 Crude Oil

It is useful, to first analyse, what our models are actually showing us. We will emphasize crude oil, here. Figure 5.2 illustrates the one day realized volatility. We notice the stylized fact of volatility clustering. This means that periods of high volatility tend to be followed by other periods of high volatility. And periods of low volatility will, generally, be followed by periods with low volatility again. Historically, we can identify several periods with high volatility, most notably in 2008, 2011 and around 2015. These high volatility periods in crude oil also coincide, approximately, with periods of high volatility on the broader financial markets.

More specifically, figure 5.3 illustrates the jump component. This component, again, becomes large when general volatility is high. Adding to this, figure 5.4 shows the corresponding (signed) forecast errors. Not surprisingly, periods of high volatility, lead to higher forecast error.

Therefore, we have to deal with the question, how these forecasts can actually help in improving a risky portfolio of crude oil. For this goal, we look at the information ratio as a measure of risk adjusted return. The information ratio divides the annualized return by the annualized volatility. We want this number to be large: a high return, paired with comparatively low volatility is desirable. This measure is related to the Sharpe ratio. In the case of the Sharpe ratio we would subtract the annualized risk free rates from the annualized return before dividing by volatility. However, firstly, what the risk free rate is, exactly, is often open to debate. And, secondly, the risk free rate has been essentially zero or even negative in the last few years. Therefore, just using the information ratio gives us a good idea of what the Sharpe ratio should be, without leaving anything open to

discussion.

Table 5.1 shows results of the economic significance for different lookbacks and lookaheads. This table should be interpreted as follows. The lookback outlines the trainset (in business days) on which the model is calibrated. For example, a lookback of 252 (business) days equals to a train set of one year, on which the model is fitted. This model is then used, with a lag of one day, to forecast volatility. It is then trained anew each day, leading to a rolling window forecast. The lookahead of 1 day, for example, means that we are forecasting 1 day ahead. It may seem like a large amount of work to retrain the model every day and arguments might be raised, whether this would actually be feasible in real-time. However, estimating a single model, takes much less than one second. And computing all numbers for the entire table 5.1 only takes a few minutes on a laptop. Therefore, all the computational needs of the model can be easily achieved without any specialized hardware or recourse to parallel computation. Indeed, introducing an artificial lag of one-day could even prove too much and our results are likely to be conservative. As model estimation (and evaluation) is almost instantaneous even a same-day forecast would be achievable. Comparatively table 5.1 shows the results excluding the initial crisis period. It can be noticed that, generally, results improve slightly. This can be interpreted as the model being better able to catch up with lower volatility periods. However, when performing this analysis and wondering whether even better results could be achieved, it should be kept in mind, that there are still volatile periods included in this sample. Specifically 2011 and 2016 have proven to be quite volatile. This cannot, of course, be known in advance.

When looking at the results of table 5.1 in more detail we should keep in mind that a pure buy and hold position in Crude Oil would have resulted in an information ratio of 0.2. This

means, always holding the front-month future and rolling it over, once the volume of the second-month future becomes larger than the volume of the front-month future. In contrast to this simple strategy, we use the volatility forecast in the following way. If volatility is forecast to rise, we short the asset. If volatility is forecast to go down, we buy the asset (go long). This exploits the general tendency of asset returns to be negatively correlated to volatility. If our forecast is valuable, then the risk adjusted return of such a strategy should be higher than just buy and hold. Of course, as the strategy may switch positions relatively often, we should account for transaction costs. Here, we assume that for each transaction we pay the bid-ask spread. Furthermore, we assume fixed transaction costs of 1 basis point per trade for the brokerage commissions and exchange fees. While this may seem low, such fees are actually achievable even for non-institutional traders at discount brokerages. For institutional traders with direct market access these fees would be almost negligible. The figures shown in the table represent therefore a true walk-forward test of what would have been achieved when following the recommendation of the forecast.

We see that, overall, the information ratio is always higher than buy-and-hold. Also, in no case, do we get a negative result. If we would be using the strategy not for risk management but rather for speculative purposes, we might also be interested in high annualized returns. Here we notice that the shortest lookback produces the overall best annualized returns on average. This is, of course, with the benefit of hindsight. We can't possibly know beforehand, which parameter set will produce the best results. The interpretation of this finding is, that dynamics seem to be shifting rapidly in the crude oil market. Therefore, short lookbacks allow for a fast adaptation of the system to these new dynamics. Longer lookbacks, on the other hand, produce lower (but arguably more stable) returns and information ratios.



Another finding is also of interest: the lookahead interval plays a minor role in producing good risk adjusted returns. While the information ratio tends to be lower for a lookahead of 5 days, this behaviour is not very strong. Overall, the information ratios stay comparatively constant for a given lookback over a range of lookahead intervals. This is in contrast to previous studies which show that the forecast efficiency diminishes markedly for multi-step forecasts. These previous studies analyse the equity market. It would be premature to draw any strong conclusions at the present state of research. However, one interpretation of this finding may be, that the dynamics of crude oil exhibit a more trending behaviour in the sense, that past behaviour is a better indicator of future performance. As the dynamics of the crude oil market may partly be driven by fundamentals which don't change significantly within a few days, we may observe such kind of behaviour.

The above pattern changes subtly, when we use a recurrent neural network to model realized volatility. Table 5.3 shows the result for different lookback intervals. To conserve space, we don't show results for different forecast horizons as the tendency remains the same: the information ratios stay relatively stable. However, we notice that the quality of the forecast as measured by its economic significance increases for a longer lookback (training) interval. This is interesting, because it shows a tendency of the network to better incorporate the past information given a longer training horizon. It is also a contrast to previous studies on equity markets. These studies found, that the economic significance of the forecast is essentially *not* sensitive to the lookback interval. The performance was equally satisfying for long and short intervals. Conversely, table 5.4 shows results excluding the initial crisis period.

Here, however, the behaviour is different: for short intervals the results are clearly not satisfying. They are sensibly below the average forecast quality of the basic HAR model. For the 126 days

and 252 days lookback the results are within the average of the HAR model, but not great. On the other hand for the long lookback interval of 504 days the resulting information ratio is clearly attractive and above what could be achieved with the basic HAR model. However, we shall always take such good results with a grain of salt. Specifically, the total forecast exercise is carried out on a dataset of only ten years. Essentially, from the end of the Global Financial Crisis upto now (May 2017) the equity markets have been predominantly in a bull market and the interest rate environment has been one of continuous quantitative easing. Our results may therefore be subject to data snooping: just by chance we found a specific model configuration (in this instance: using an RNN with two years of training data) that happened to work well on crude oil in this more or less steady market environment. The after crisis results show a similar behaviour as above. Slight, but not substantial, improvements.

Finally, table 5.5 highlights the results of using a hybrid model to forecast one-day ahead realized volatility. The results start with a very attractive information ratio for a lookback of 63 days. It seems, that for this short interval the hybrid model is able to combine the advantages of both the linear HAR model and the neural network to its advantage. Indeed, the HAR alone is also able to produce good results on a short lookback and feeding this result into a neural network seems to improve the performance further. However, for the longest lookback the performance deteriorates considerably and the information ratio becomes markedly negative. This should be a warning sign to use the hybrid model for long lookbacks. The decline in performance seems to be gradual instead of sudden. The longer the lookback, the worse the performance is. Again, table 5.6 highlights after crisis results.

We may attribute the above results to the general problem of overfitting. The least complex

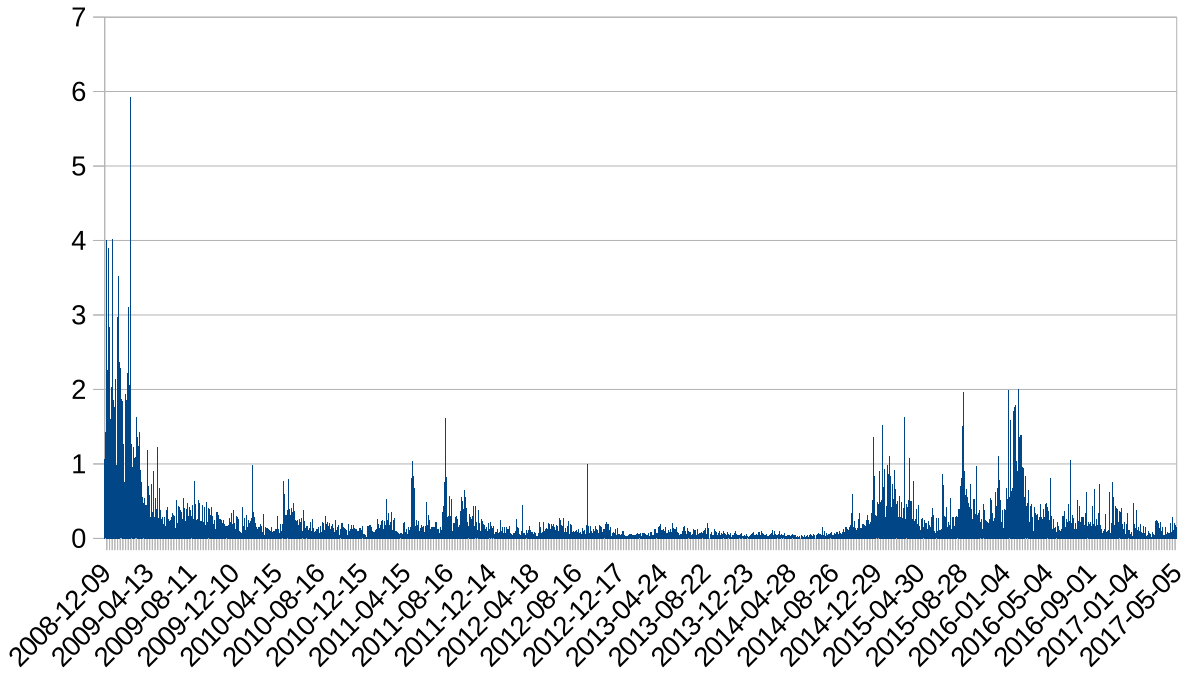


Figure 5.2: One day realized volatility over time.

model, the linear HAR model, produces average results, but does so consistently for all lookbacks. The non-linear pure recurrent neural network produces results that range from average to very good. And, finally, the most complex hybrid model, produces results that range from terrible to excellent. This may be an unfair treatment of the non-linear model: at least, they have domains that perform consistently well. For the pure recurrent network these are longer lookback intervals. For the hybrid model we should confine ourselves to shorter lookback intervals. Nevertheless, the behaviour among the non-linear models is actually not consistent. Before using any of the model, further research might be warranted. Also, excluding the crisis period does not improve the results by much.

#### 5.4.2 Natural Gas

After the extensive analysis of the results for Crude Oil we now shift our attention to another very popular and liquidly traded energy commodity: Natural Gas. Table 5.7 shows results for different

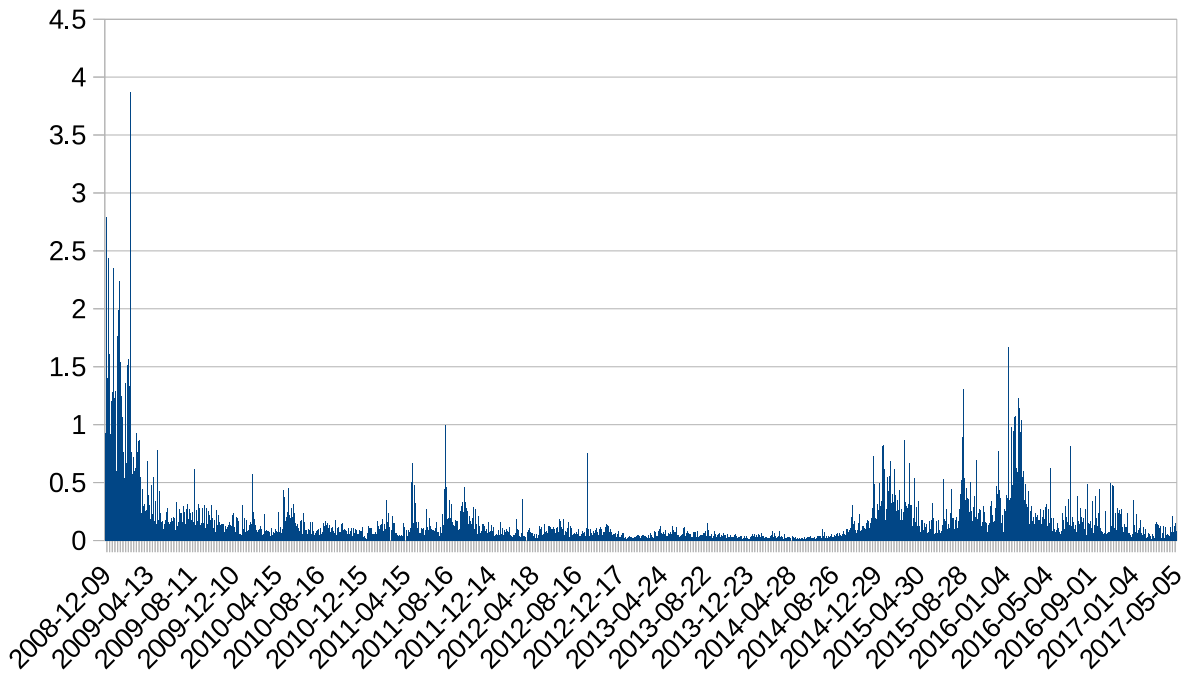


Figure 5.3: Jump component over time.

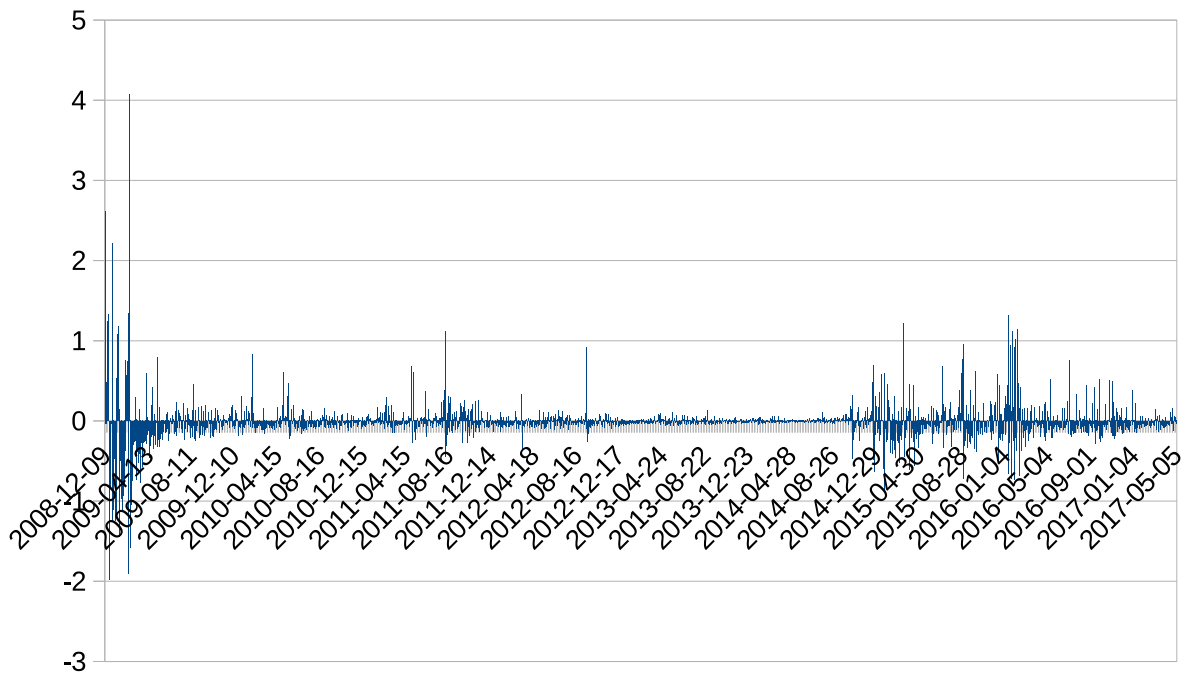


Figure 5.4: Forecast errors for a lookback of 252 days and a one day ahead forecast.

lookback	lookahead	ann. ret.	inf. ratio
63	1	0.2794	0.5408
63	2	0.4124	0.7989
63	3	0.3667	0.7101
63	4	0.3726	0.7216
63	5	0.3524	0.6825
126	1	0.2874	0.5563
126	2	0.283	0.5479
126	3	0.3694	0.7155
126	4	0.2346	0.454
126	5	0.27	0.5226
252	1	0.198	0.3831
252	2	0.2606	0.5045
252	3	0.2405	0.4655
252	4	0.2071	0.4008
252	5	0.1889	0.3656
504	1	0.2832	0.5482
504	2	0.2867	0.555
504	3	0.2498	0.4835
504	4	0.2249	0.4353
504	5	0.2615	0.5063

Table 5.1: Overview of economic significance for Crude Oil.

lookback	lookahead	ann. ret.	inf. ratio
63	1	0.2989	0.6174
63	2	0.4574	0.8271
63	3	0.3791	0.7397
63	4	0.4013	0.7589
63	5	0.3617	0.6724
126	1	0.2784	0.5457
126	2	0.3153	0.5683
126	3	0.3898	0.7415
126	4	0.2512	0.489
126	5	0.2974	0.5485
252	1	0.2015	0.3911
252	2	0.2751	0.5415
252	3	0.2814	0.4634
252	4	0.2275	0.4484
252	5	0.2018	0.3797
504	1	0.3014	0.5633
504	2	0.2915	0.5612
504	3	0.2659	0.4931
504	4	0.2441	0.4613
504	5	0.2723	0.5132

Table 5.2: Overview of economic significance for Crude Oil after crisis

lookback	ann. ret.	inf. ratio
63	0.12	0.2846
126	0.2118	0.5119
252	0.2096	0.5054
504	0.3186	1.041

Table 5.3: One step ahead forecast for Crude Oil using a Recurrent Neural Network.

lookback	ann. ret.	inf. ratio
63	0.1414	0.2944
126	0.2322	0.5231
252	0.2151	0.5341
504	0.3285	1.131

Table 5.4: One step ahead forecast for Crude Oil using a Recurrent Neural Network after crisis.

lookback	ann. ret.	inf. ratio
63	0.4921	1.1699
126	0.29	0.7013
252	0.2373	0.5721
504	-0.1972	-0.6436

Table 5.5: One step ahead forecast for Crude Oil using a Hybrid Model.

lookback	ann. ret.	inf. ratio
63	0.5234	1.2012
126	0.3153	0.7253
252	0.2389	0.5891
504	-0.1841	-0.6352

Table 5.6: One step ahead forecast for Crude Oil using a Hybrid Model after crisis.

lookbacks and lookaheads using the HAR model including jumps. Just glancing at the annualized returns and corresponding information ratios would give the impression, that the HAR model is able to produce forecast of excellent economic significance. There is also a remarkable tendency of improving forecasts for a longer lookahead.

However, we have to be very careful with the interpretation of such seemingly attractive results. It is true, without doubt, that the basic HAR model is able to nicely catch-up with the underlying volatility process of Natural Gas. But, over the last ten years, we notice that Natural Gas has exhibited a strong mean-reverting tendency. Therefore, any simple mean-reverting strategy would have produced attractive returns. While there is a fundamental reason for a slight mean-reversion of the Natural Gas reason due to seasonal demand, there is no specific reason, why such a mean-reverting tendency should persist forever. We might therefore expect a *regime change*. In this case, it is not clear, if the HAR model would be able to catch up with this regime change. We would therefore advise, to use the model only with extreme caution.

We also have another reason for caution when we look at table 5.8. This table reports one step ahead forecasts for natural gas, where we use a recurrent neural networks. It is the exact same setup as we used previously for crude oil. The results, here, are clearly disappointing, possibly indicating a high degree of overfitting. In contrast to the situation with Crude Oil, the hybrid model in table 5.9 also shows no improvement. To conserve space the after crisis results are omitted here as they present to substantial added information.

### 5.4.3 Summary of Results

In any case, the economic results are encouraging. Whether realized volatility forecasts are used for risk management purposes, for trading, or for policy decisions, there seems to be value over a

lookback	lookahead	ann. ret.	inf. ratio
63	1	0.3534	0.7495
63	2	0.5545	1.178
63	3	0.1224	0.2594
63	4	0.9093	1.9408
63	5	1.1651	2.4988
126	1	0.4317	0.9161
126	2	0.8469	1.8058
126	3	0.59	1.2538
126	4	0.9817	2.0979
126	5	0.9417	2.0109
252	1	0.7802	1.6619
252	2	0.9059	1.9335
252	3	0.7197	1.5319
252	4	0.733	1.5603
252	5	1.0129	2.1659
504	1	0.8915	1.9023
504	2	0.8445	1.8006
504	3	0.7801	1.6616
504	4	0.8915	1.9023
504	5	0.8915	1.9023

Table 5.7: Overview of economic significance for Natural Gas.

lookback	ann. ret.	inf. ratio
63	0.3512	0.7592
126	0.6891	1.485
252	0.4211	0.9126
504	0.5217	1.1083

Table 5.8: One step ahead forecast for Natural Gas using a Recurrent Neural Network.

lookback	ann. ret.	inf. ratio
63	0.386	0.8345
126	0.4223	0.9076
252	0.0126	0.0274
504	0.2041	0.4327

Table 5.9: One step ahead forecast for Natural Gas using a Hybrid Model.



naive strategy.

At the same time, clearly, more research is warranted. This especially concerns the use of non-linear models. While adequate model tuning would surely improve the economic results we have to be very careful in using non-linear models *out of the box*. The present analysis shows that excluding very volatile periods may improve model performance. However, this has also to be taken with care as it is not possible to know the volatile periods in advance. Therefore, we cannot, realistically exclude them. Of course, a volatility filter could be of interest here: for example, refraining from using the model, if historical realized volatility has been higher than average over the past  $n$  days. This introduces a new variable and should be applied with extreme care.

Nevertheless, we have a relatively clear assessment of the basic HAR model including jumps: this model seems, generally, to be able to produce robust results. This is irrespective of the interval that is used to train the model and irrespective of the forecast horizon ranging from one to five days. This is promising. Yet, as a rule, a walk-forward test on around ten years of data can only be a first hint at potential usability as a risk management and decision support tool. Within the last ten years we have been mostly stuck with a market regime of declining interest rates and rising asset prices.

## 5.5 Conclusions

This chapter presents an analysis of the modelling of realized volatility using futures data. To this end, we firstly use the comparatively simple HAR model including jumps. We also use more complex models, namely a recurrent neural network and, finally, a hybrid model. This hybrid model uses the HAR model forecast as an independent input to a neural network.

As it turns out all models are able to beat the naive benchmark of buy and hold. Most model

instances beat the benchmark by a wide margin. The linear HAR model including jumps shows the most consistent performance. The non-linear models are able to deliver outstanding performance in some cases, however the overall performance is less robust. This may be taken as a clue to explore further options with the non-linear models.

The first path to explore is to take an ensemble of models. As every model instance will be different due to random initialization of parameters (weights) each non-linear model instance will also produce a different output. Averaging over all these outputs might produce more stable forecasts. However, if for every timestep in the rolling-window approach one hundred models have to be computed, the overall time to analyse a time series will increase significantly. Nevertheless, the whole training process of a non-linear model only takes a few minutes. Therefore using an ensemble approach is a viable next step.

Further potential work includes more technical variations of the model building. Several meta-parameters used by our models are taken in an ad-hoc way, just replicating best practices of the literature. However, this does not imply that these meta-parameters are set in stone. Specifically, it would be instructive to change the actual frequency of the data for computing realized volatility. For example, evaluating the model on one minute data and on ten minute data. One minute data (or even higher frequency data) would presumably introduce noise in the estimation process, notably due to bid-ask bounce. However, the additional information contained in the higher frequency data might be worth it. On the other hand, lower frequency data would be more smooth, but we might lose important information. This tradeoff should be explored in more detail.

## Chapter 6

# Discussion and Limitations

Even if the obtained results were very satisfying, there is, nevertheless, always room for improvement, critique and further research. For our two main topics, NAV modelling and RV modelling the challenges were of different nature and shall be discussed subsequently.

### 6.1 Discussion of Modelling ETF Premium

#### 6.1.1 Problem Outline

Modelling ETF premium and discount and the accompanying potential feedback trading proved to be a worthwhile research quest. Especially, we could analyse to which extent feedback trading is actually present. This is a natural extension to previous work by other authors who already analyse premium and discount for more mainstream ETF. However, our analysis is distinct from the literature on closed-end fund premiums and discounts. In the case of closed-end funds there is a natural arbitrage opportunity that arises when a closed-end fund is traded on exchange and the price deviates from the supposedly well-known NAV.

The case in the present research is a slightly different one: the causes for premiums or discounts may be mainly two-fold. On the one hand, we may have an apparent deviation from NAV because of synchronicity issues: the ETF is trading, while the assets within the ETF are actually not trading. Here, we have a stale NAV, that might have been computed several hours ago. The price of the

ETF is influenced by supply and demand. And this supply and demand will be chiefly influenced by the market participants' perception on how the holdings of the ETF are likely to evolve. As a direct observation is not possible, each market participant will form her own model of the holdings' price evolution. In case of equity ETFs the price may just follow the US stock market as a leader, although there is really no compelling reason to do so. Or, market participants may use more or less sophisticated models to determine a likely price for the ETF. In any case, the price of the ETF is more of an educated guess than anything else.

### 6.1.2 Modelling Issues

A second reason for deviations from NAV can be found in liquidity issues. It might be rational for investors to over- or underprice the ETF compared to its NAV because of supposed better liquidity in the ETF or in the underlying assets. This could be a very valid concern and lead to a persistent discount or premium.

In the case of modelling ETF premium and discount it can be argued, that we didn't implement some rather obvious extensions of our research. We were mostly concerned with modelling the potential feedback trading, therefore describing its features. However, we didn't venture into analysing, whether a supposed anomaly could actually be exploited on the financial markets. Therefore, our work is limited in the sense that it (mostly) describes stylized facts of ETF premiums for emerging market ETFs. But there is also the question of *economic significance*. Or, put more bluntly: Is there a truly exploitable anomaly that could lead to abnormal returns by those who have a good grasp on the premium and discount mechanism of (emerging markets) ETFs? There seems to be a truly puzzling anomaly in the domain of closed-end funds which (even after transaction costs) offer abnormal returns to buying discounted funds and selling funds trading at a premium.

However, it should be explored, whether similar abnormal returns could also be obtained with the emerging markets ETFs.

The analysis of this question has to be left to future research. However, for the sake of a critical assessment, we can already highlight a few arguments that would speak against a truly exploitable market anomaly for emerging market ETFs. Firstly, as already outlined, premium or discounts might be rational and not likely to go away. Secondly, most ETFs we are analysing cannot be categorised as *very liquid*. In consequence, spreads might be too large to actually make a true arbitrage transaction attractive. The comparatively low daily turnover of these ETFs may also makes it difficult to execute large volume transaction. Additionally, if we want to engage into a hedged transaction we would look, for example, at buying the ETF and selling the basket of constituents (or vice versa, depending on the perceived discount/premium). Transacting in the ETF's holdings might prove difficult enough, to make such an operation financially unattractive. Spreads and outright transaction costs might be high. On top of that, certain emerging market countries may have short-selling restrictions, that effectively make selling the constituents impossible.

### **6.1.3 Financial Viability**

If we wanted to analyse the financial viability of modelling feedback trading on premium or discounted ETFs we might still engage in a third type of transaction. First, we will assume that the *emerging markets* as a whole will show some pronounced degree of correlation. Then, it might be tempting, to build a market neutral portfolio that would include all emerging markets ETFs. We can long the ETFs at a discount and short the ETFs at a premium. As we have a large number of ETFs in our analysis our portfolio would be relatively immune against external market shocks. Again, implementing this strategy will be subject to the same caveats as outlined above.

Another limitation is the availability of NAV data for emerging market ETFs. Generally, with the data sources accessible to us, we were mostly limited to NAV data from the iShares series of ETFs. This potentially introduces an *issuer bias* into the analysis. Indeed, with the current dataset, it is not meaningfully possible to check, whether there might be some systematic distortion in the recorded prices of iShares ETFs. There is no specific reason to assume that such a bias would be present, because it would be detrimental to the issuer. However, the academic literature has several examples of regular mispricings along the product line of a given issuer. These findings mostly relate to more complex products, like exotic derivatives. However, a comparison among different issuers would be relatively easy to do given access to the corresponding data and could answer this question. This is left to further research.

Additionally, the specific selection of ETFs was, again, dictated by the availability of NAV data. In this sense some country ETFs could just not be considered. It may be argued that the iShares product line is well-reputed and broad. Nevertheless, extending the analysis would have been easy with access to the data.

From a modelling perspective there is one broad line of thought that always crops up when using a parametric approach in contrast to a non-parametric approach: How do we know, that we have the right model? The answer is, obviously: We just cannot be sure. Each model based approach is subject to *model risk*. We should be aware of this and cannot rule out, that some other model might be a much better description of reality than what was used in this thesis. This critique could be mitigated by using a non-parametric approach, for example Artificial Neural Networks, Support Vector Machines or Genetic Programming. We address these concerns in our second research area, modelling realized volatility. For the ETF premium research area we leave a

non-parametric approach to future work.

## 6.2 Discussion of Modelling Realized Volatility

### 6.2.1 Problem Outline

Related to the modelling of Realized Volatility several limitations and, accordingly, possibilities of extensions of our work come to mind. Firstly, when just considering the pure HAR-RV-J model the inherent critique of model risk can be seen. HAR-RV-J is just one of many models that can be used to model RV. The literature also considers many more complex models, although HAR-RV-J is praised for its simplicity and accuracy. In the case of RV we try to counter this argument by also including a model-free approach by using an artificial neural network and also a hybrid model combining neural networks and the HAR-RV-J. The neural network does not imply any specific functional form, rather lets the data speak for itself. This, on the other hand, raises the concern of potential overfitting. However, the network is kept small and overfitting is controlled through early stopping.

### 6.2.2 Data Availability

Another point of critique may be raised when considering the actual data used to compute RV. In the present case we use five-minute data. This is seen as a good compromise between an adequate level of detail, not available at lower frequencies, and a reduction of noise, that would occur at higher frequencies. Of course, this is debatable. Arguably, qualitatively similar results could be obtained using ten-minute or three-minute data.

Also, this raises the interesting question whether there is some *optimal* frequency on which to compute RV and which parameters determine this frequency. This would imply actually determining meta-parameters of RV modelling, a task left for further research.

### 6.2.3 Application Domain

Our RV modelling applications are also limited by the specific application domains. On the one hand we model RV on very liquid, large indices and derived products. On the other hand we model RV on energy-related assets, putting RV modelling into a more risk-focused, regulatory perspective.

In all cases we explore the economic significance of the RV forecast in addition to computing the usual statistical error measures. However, more work could have been carried out in evaluating the risk assessment qualities of the presented RV forecasts.



## Chapter 7

# Outlook and Recommendations

### 7.1 Conclusions and Outlook

My thesis highlights potential market anomalies and inefficiencies using advanced methods and datasets. Specifically, the thesis is concerned with modelling and analysing feedback trading on ETFs and with modelling high-frequency intraday realized volatility.

However, the presented results should only be considered as one of many steps towards a more thorough treatment of the underlying problems. In the case of ETFs a host of different future research paths arises. These include, for example:

- Currently, the entire analysis is based on daily closing prices and net asset values. This data will be the most common, as it is readily available. On the other hand, little analysis has been carried out on intraday relationships and deviations of price and NAV. This is understandable, as, typically, the intraday NAV would have to be estimated manually out of the components of the ETF. Some providers and market participants (for example Interactive Brokers, among others) offer intraday updates of ETF NAVs included in their normal data feed. But, generally, this data has seen less attention. Yet, this might prove a very interesting area for research, as premiums and discounts intraday might be much more pronounced.
- Relatedly, the question arises, what the influence of pre-hours and after-hours trading is on

ETFs trading assets in a far-away timezone. Indeed, as the large Electronic Communication Networks (ECNs) offer almost round-the-clock off-exchange matching facilities for most US securities an overlap of trading hours could be observed. However, spreads outside the regular trading hours tend to be quite significant and could cancel out any supposedly available arbitrage opportunity. Also, data would have to be collected individually from the different ECNs, as the data is not necessarily consolidated into a single feed.

- Additionally, the current analysis is only concerned with aggregates. That is, I don't consider the individual components that make up the ETFs at all. Arguably, considering and modelling the components individually, and, ultimately, trying to forecast them, might lead to better overall results when evaluating feedback trading.

For the realized volatility research the potential for future analysis is also significant:

- Although I evaluated several asset classes already, including equities and commodities, the method is, in principle, applicable to any asset. Therefore, a very comprehensive analysis might allow to single out areas, where the analysis is especially suitable and also highlight assets, which do not lend themselves easily to such modelling. For example, a broad analysis could include the top 100 most liquid ETFs traded in the US, or the major future markets, or, potentially, also include bond markets. Finally, analysing all the constituents of a stock index would also prove instructive. Such an extended analysis necessitates ample data and still more automation in the different processing steps to be actually viable.
- Future research should also include a more detailed benchmarking of the applied HAR and neural network model. For example, as has been outlined in the literature research, some HAR variants also include a measure of implied market volatility (for example the VIX in the

case of the S&P500 index) as an additional model input. On the other hand, more traditional approaches are in existence to model volatility. Putting this into perspective, a comprehensive comparison of different model types would be useful. Ideally, this should be coupled with a broad analysis as outlined in the previous bullet point.

- On top of that, it might be worthwhile to try out a *big data* approach that includes still other, more exotic predictors. In this case, of course, this would deviate from the premise of building a purely autoregressive model. This tradeoff has to be weighted carefully: the three papers on this topic show, that the basic HAR model including jumps is already able to model volatility satisfactorily. On the other hand some predictors have been successful, especially when applied to intraday data. This includes, for example, the release of news. Even if the content of news is not known, just the fact that a news release is scheduled can lead to repeatable volatility patterns. If the content of the news can somehow be forecast, then even better assessments are possible.
- Also, modelling realized volatility around the close of each trading day seems a bit arbitrary. An intraday analysis and forecast of realized volatility would be both a novel application and potentially unearth new approaches in risk management. Indeed, out of my experience with the industry I noticed, that most risk management models are only updated and evaluated daily. However, markets are moving quickly and an intraday assessment of the probably path of volatility would surely be welcome.
- Related to the above future research idea more research is warranted as to which kind of intraday data should actually be used. Sticking to the literature standard, I use five minutes data. Yet, the options are many. The highest frequency available would be to use tick

data. While we have this data available, using tick data would require a good assessment of market noise. Using tick data for modelling realized volatility would necessitate, especially, a modelling of the bid-ask bounce.

Typically, one doesn't want the bid-ask bounce to interfere with our realized volatility computation. At the same time, most models assume equidistant spacing of data in time. This is not the case, obviously, for tick data. A suggestion might be, therefore, to use an *activity* or *volatility* based time scale. For example, realized volatility is computed every 100 ticks. At the other end of the spectrum, we have much lower frequency data that is less noisy but also less informative.

- At the same time the thesis has mostly been concerned with the quantitative finance aspects of the work. However, the work on realized volatility can and should also be seen in the context of decision support systems and from an information technology aspect. Due to my background in mathematics, obviously, we weren't able to carry out a usability and technology acceptance analysis of the models we developed. Yet, in the context of an interdisciplinary study, it would be useful to provide realized volatility forecasts to professional risk managers (and, perhaps, traders). This would allow the evaluation of the actual usefulness in combination with human interaction. For example, we could answer the question, whether human experts actually find value in our methods when applied on a daily basis. As I got a good amount of industry feedback on this project I tend to believe that the models developed have value. Yet, the feedback is by a comparatively limited number of people. This number of people would have to significantly extended to get a more truthful assessment.

Especially the presentation of the work on realized volatility has been well received at confer-

ences by academics and industry participants alike. I see good potential in developing this work further and bringing it to application in the context of risk management systems. Mostly, participants have commented positively on the simplicity, yet effectiveness, of the model. Also, the clarity of exposition has been mentioned as a good aspect of the work. I am very keen on pursuing research in these areas as the topic seems to be of high interest, currently, to the community.

I see the potential of actually improving today's risk management systems and provide a better understanding of short-term volatility. This in turn could lead to less hasty reactions on the financial markets and potentially avoid the *mini-crashes* we regularly encounter in different asset classes. Of course, nowadays, this is still more of a vision than a reality. Yet, it seems a worthwhile goal to pursue.

Summing up, my thesis contributes to the growing body of literature that deals with modelling and analysing potential market anomalies in the realm of high-frequency data.

## 7.2 Key Recommendations

The present section intends to give actionable key recommendations for the topics discussed in my thesis.

Related to ETF feedback trading we can notice that deviations from NAV are relatively frequent and may provide trading opportunities of economic significance. These findings could be exploited in the following ways. On the one hand, when purchasing ETF shares for long-term investments it might be useful to research those ETFs which trade at a discount to NAV, if some choice is available on the target investment goal. Conversely, the existence of premiums means that ETF investors might want to research those ETFs with the lowest average premium (for a given investment universe) to get a slightly better entry price.

More to the point of the actual research carried out this thesis finds that feedback trading may provide real opportunities with exploitable anomalies. Nevertheless, this recommendation has to be cautioned by the fact, that true opportunities with improved risk-adjusted returns will have to be researched thoroughly. Not every seemingly attractive deviation from NAV will result in an actionnable trade.

Related to the work carried out on realized intraday volatility we can take away, that the HAR model with jumps does a satisfying job of forecasting realized volatility. The inputs to the HAR model also lead to good forecasts when employing an RNN model. The ensuing double recommendation is to go for a pure HAR model, if enough lookback data (more than one year) is available. If only little data is available (one or three months), then an RNN model will be a good choice on average.

Additionally, it is recommended to use the above models in the context of risk management or systematic trading applications. Indeed, following the forecasts of the realized volatility models leads to improved risk-adjusted returns in terms of Sharpe ratios. Whether this will persist is up to discussion. In any case the recommendation is to always monitor the model performance in real time and to be aware of quickly changing market environments.

# Appendix A

## Abbreviations

ETFs: Exchange Traded Funds  
NAVs: Net Asset Values  
EMH: Efficient Market Hypothesis  
GARCH model: Generalized AutoRegressive Conditional Heteroskedasticity model  
GJR-GARCH model: Glosten-Jagannathan-Runkle GARCH model  
OU process: Ornstein-Uhlenbeck process  
SV model: Stochastic Volatility model  
RV model: Realized Volatility model  
HAR-RV: Heterogeneous AutoRegressive model of Realized Volatility  
HAR-RV-J: HAR-RV model with Jumps  
ANN: Artificial Neural Network  
RNN: Recurrent Neural Network  
ETNs: Exchange Traded Notes  
RGARCH model: Realized GARCH model  
BV: Bipower Variation  
RMSE: Root Mean Squared Error  
MAE: Mean Absolute Error  
MAPE: Mean Absolute Percent Error

# Bibliography

- Ackert, L., Tian, Y. Arbitrage, liquidity, and the valuation of exchange traded funds, *Financial Markets, Institutions and Instruments*, 2008, 17(5), 331-362.
- Aguirre, M.S., Saidi, R. Feedback trading in exchange-rate market: evidence from within and across economic blocks, *Journal of Economics and Finance*, 1999, 23(1), 1-14.
- Aït-Sahalia, Y. and Jacod, J., Testing for jumps in a discretely observed process, *The Annals of Statistics*, 37(1)(2009), 184–222.
- Alexander, C., Kapraun, J. and Korovilas, D., Trading and investing in volatility products, *Financial Markets, Institutions Instruments*, 24(4)(2015), 313–347.
- Agnolucci, P., Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models, *Energy Economics*, 31(2)(2009), 316-321.
- Andersen, T.G. and Bollerslev, T., Answering the skeptics: Yes, standard volatility models do provide accurate forecast, *International Economic Review*, 39(4)(1998), 885–905.
- Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P., The distribution of realized exchange rate volatility. *Journal of American Statistical Association*, 96(2001), 42–55.
- Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P., Modelling and forecasting realized volatility, *Econometrica*, 71(2)(2003), 579–625.



- Andersen, T.G., Bollerslev, T., Christoffersen, P.F. and Diebold, F.X., Volatility and correlation forecasting. *Handbook of Economic Forecasting*, 1(2006), 777–878.
- Andersen, T.G., Bollerslev, T. and Diebold, F.X., Roughing it up: Including jump components in the measurement, modelling, and forecasting of return volatility. *Review of Economics Statistics*, 89(4)(2007), 701–720.
- Andersen T.G., Bollerslev T. and Dobrev D., No-arbitrage semimartingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics*, 138(1)(2007b), 125–180.
- Andersen, T.G., Dobrev, D. and Schaumburg, E., Jump-robust volatility estimation using nearest neighbour truncation. *Journal of Econometrics*, 169(1)(2012), 75–93.
- Andreou, E. and Ghysels, E., Rolling-sample volatility estimators: Some new theoretical simulation, and empirical results. *Journal of Business & Economic Statistics*, 20(3)(2002), 363–376.
- Antoniou, A., Koutmos, G., Pericli, A. Index futures and positive feedback trading: Evidence from major stock exchanges, *Journal of Empirical Finance*, 2005, 12(2), 219-238.
- Araújo, R.d.A., Oliveira, A.L.I., Meira, S., A hybrid model for high-frequency stock market forecasting, *Expert System with Applications*, 42(8)(2015), 4081-4096.
- Back, K., Asset Pricing for General Processes, *Journal of Mathematical Economics*, 20(1991), 317–395.
- Bajgrowicz, P., Scaillet, O. and Treccani, A., Jumps in high-frequency data: Spurious detections, dynamics, and news. *Management Science*, 62(2016)(8), 2198–2217.

- Bandi, F.M., Russell, J.R., Yang, C., Realized volatility forecasting in the presence of time-varying noise, *Journal of Business and Economic Statistics*, 31 (2013) 331-345.
- Barber, B.M., Lee, Y., Liu, Y., Odean, T., Is the aggregate investor reluctant to realise losses? Evidence from Taiwan, *European Financial Management*, 2007, 13(3), 423-447.
- Barber, B.M., Lee, Y., Liu, Y., Odean, T., just how much do individuals lose by trading? *Review of financial studies*, 2009, 22(2), 609-632.
- Barberis, N., Shleifer, A., Vishny, R., A model of investor sentiment, *Journal of Financial Economics*, 1998, 49, 307-343.
- Barberis, N., Thaler, R., A survey of behavioural finance, In: Constantinides, G., Harris, M., Stulz, R. (Eds), *Handbook of the Economics of Finance*. North-Holland, Amsterdam, 2003
- Barndorff-Nielsen, E.E. and Shephard, N., Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics. *Journal of the Royal Statistical Society B*, 63(2001)(2), 167–241.
- Barndorff-Nielsen, E.E. and Shephard, N., Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of Royal Statistical Society B*, 64(2002a)(2), 253–280.
- Barndorff-Nielsen, E.E. and Shephard, N., Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17(2002)(5), 457–477.
- Barndorff-Nielsen, O. and Shephard, N., Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial. Econometrics*, 2(2004), 1–48.

- O.E.Barndorff-Nielsen, S.E. Graversen, J. Jacod, N. Shephard, Limit theorems for realized bipower variation in econometrics. *Econometric Theory*, 22 (2006) 677-719.
- O.E. Barndorff-Nielsen, N. Shephard, Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4 (2006) 1-30
- O.E. Barndorff-Nielsen, S. Kinnebrock, N. Shephard, Measuring downside risk-realized semivariance, T.Bollerslev, J.Russell, M.Watson(Eds.), *Volatility and time series econometrics: essays in honor of Robert F. Engle*, Oxford University Press (2010).
- Barunik, J. and Krehlik, T., Combining high frequency data with non-linear models for forecasting energy market volatility. *Expert Systems with Application*, 55(2016), 222–242.
- Barunik, J., Krehlik, T., Combining high frequency data with non-linear models for forecasting energy market volatility. *Expert Systems Applications*, 55 (2016a) 222–242.
- Barunik, J., Krehlik, T., Vacha, L., Modelling and forecasting exchange rate volatility in time-frequency domain. *European Journal of Operational Research*, 251 (2016b) 329-340.
- Bennett, J.R., Sias, R., Starks, L., Greener pastures and the impact of dynamic institutional preferences, *Review of Financial Studies*, 2003, 16, 1203-1238.
- Bildirici, M. and Ersin, O.O., Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange. *Expert Systems with Applications*, 36(2009)(4), 7355-7362.
- Blitz, D., Huij, J., Evaluating the performance of global emerging market equity exchange-traded funds, *Emerging Markets Review*, 2012, 13, 149-158.

- Bollerslev, T., Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(1986), 307–327.
- Bollerslev, T., Engle, R.F. and Nelson, D.B., ARCH models, in Daniel Mcfadden and Robert F. Engle, eds., *Handbook of Econometrics*, 4(1994), Elsevier: Amsterdam.
- Bollerslev, T., Patton, A.J., Quaedvlieg, R., Exploiting the errors: A simple approach for improved volatility forecasting, *Journal of Econometrics*, 192 (2016) 1-18.
- Bohl, M.T., Reitz, S., Do positive feedback traders act in Germany’s neuer market? *Quarterly Journal of Business and Economics*, 45(1/2), 2006, 3-14.
- Bond, S., Hawkins, M., Klemn, A., Stamp duty on shares and its effect on share prices, *FinanzArchiv: Public Finance Analysis*, 61(3), 2005, 275-297. Bohl, M.T., Siklos, P.L., Empirical Evidence on Feedback Trading in Mature and Emerging Stock Markets, *Applied Financial Economics*, 18-17, 2008, 1379-1389.
- Bormetti, G., Cazzola, V., Delpini, D., Option pricing under Ornstein-Uhlenbeck stochastic volatility: a linear model, *International Journal of Theoretical & Applied Finance*, 7, 2010, 1047-1063.
- Borovkova, S. and Mahakena, D., News, volatility and jumps: The case of natural gas futures. *Quantitative Finance*, 15(2015)(7), 1217–1242.
- Boudt, K. and Zhang, J., Jumps robust two time scale covariance estimation and realized volatility budgets. *Quantitative Finance*, 15(2015)(6), 1041–1054.

- Bowe, M., Domuta, D., Investor herding during financial crisis: a clinical study of the Jakarta stock exchange, *Pacific-Basin Finance Journal*, 12, 2004, 387-418
- Brennan, M.J., Cao, H., International portfolio investment flows, *Journal of Finance*, 52(5), 1997, 1851-1880.
- Busch, T., Christensen, J.B., Nielsen, M.O., The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock and bond markets, *Journal of Econometrics*, 160 (2011), 48-57.
- Carr, P and Wu, L (2003) What type of process underlies options? A simple robust test. *Journal of Finance*, 58(2003)(6), 2581–2610.
- Carr, P. and Lee, R., Volatility Derivatives. *Annual Review of Financial Economics*, 1(2009), 319–339.
- Celik, S. and Ergin, H., Volatility forecasting using high frequency data: Evidence from stock markets. *Economic Modelling*, 36(2014) 176–190.
- Charfeddine, L., True or spurious long memory in volatility: Further evidence on the energy futures markets, *Energy Policy*, 71(2014), 76-93.
- Chan, K.F., Gray, P., Van Campen, B., A new approach to characterizing and forecasting electricity price volatility, *International Journal of Forecasting*, 24(2008), 728-743.
- Charteris, A., Chau, F., Gavrilidis, K., Kallinterakis, V., Premiums, discounts and feedback trading: evidence from emerging markets' ETFs, *International Review of Financial Analysis*, 35, 2014, 80-89.

- Chau, F., Deesomsak R., Business cycle variation in positive feedback trading: Evidence from the G-7 Economies, *Journal of International Financial Markets, Institutions and Money*, 35, 2015, 147-159.
- Chau, F., Deesomsak, R., Lau, M., Investor sentiment and feedback trading: evidence from the exchange traded funds markets, *International Review of Financial Analysis*, 20, 2011, 292-305.
- Chau, F., Holmes, P., Paudyal, K., The impact of universal stock futures on feedback trading and volatility dynamics, *Journal of Business Accounting and Finance*, 35, 2008, 227-249.
- Chau, F., Kuo, J., Shi, Y. Arbitrage opportunities and feedback trading in emissions and energy markets. *Journal of Business Accounting and Finance*, 2015, 35, 227-249.
- Christensen, K., Oomen R.C.A. and Podolskij, M., Fact or friction: Jumps at ultra high frequency. *Journal of Financial Econometrics*, 114(2014)(3), 576–599.
- Chen, H.C., Ho, J.K., Lai, C.W., Morales-Camargo, E., Institutional trading behaviour in the ETF market, working paper, Anderson School of Management, University of New Mexico., 2012.
- Cherry, J. The limits of arbitrage: Evidence from exchange traded funds, Honours Dissertation, University of California-Berkely, California.
- Chevallier, J., Sévi, B., On the volatility-volume relationship in energy futures market using intraday data, *Energy Economics*, 34(6)(2012), 1896-1909.
- Christoffersen, P.F., Diebold, F.X., How relevant is volatility forecasting for financial risk management, *The Review of Economics and Statistics*, 82(1) (2000) 12-22.

- Choe, H., Kho, B.C., Stulz, R.M., Do foreign investors destabilize stock markets? The Korean experience in 1997, *Journal of Financial Economics*, 54, 1999, 227-264.
- Choi, N. Sias, R.W., Institutional industry herding, *Journal of Financial Economics*, 94, 2009, 469-491.
- Choi, N., Skiba, H., Institutional herding in international markets, *Journal of Banking and Finance*. 55, 2015, 246-259.
- Chou, R.K., Wang, G.H.K., Wang, Y.Y., Bjursell, J., The impacts of large trades by trader types on intraday futures prices: Evidence from the Taiwan futures exchange, *Pacific-Basin Financial Journal*, 19, 2011, 123-138.
- Ciarreta, A., Zarraga, A., Modelling realized volatility on the Spanish intra-day electricity market, *Energy Economics*, 58(2016), 152-163.
- Ciarreta, A., Muniain, P., Zarraga, A., Modelling realized volatility in Electricity spot prices: new insights and application to the Japanese electricity market, *ISER Discussion Paper No. 991*, (2017).
- Corsi, F., A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2009)(2), 174–196.
- Corsi, F., Threshold bipower variation and the impact of jumps on volatility forecasting, *Journal of Econometrics*, 159 (2010) 276-288.
- Cubadda, G., Guardabascio, B., Hecq, A., A vector heterogeneous autoregressive index model for realized volatility measures, *International Journal of Forecasting*, 33(2) (2017) 337-344.

- Cunado, J., Gil-Alana, L.A., De Gracia, F.P., Persistence in some energy futures markets, 30(5)(2010), 490-507.
- Culter, D.M., Poterba, J.M., Summers, L.H., Speculative dynamics and the role of feedback traders, American Economic Review, 80, 1990, 63-68.
- Da Fonseca, J., Ignatieva, K., Ziveyi, J., Explaining credit default swap spreads by means of realized jumps and volatility in the energy markets, Energy Economics, 56(2016), 215-228.
- De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., Positive feedback investment strategies and destabilizing rational speculation, Journal of Finance, 45(2), 1990, 379-395.
- Deville, L. Exchange traded funds: history, trading, research, in: Zopounidis, C., Doumpos, M. Pardalos, P. (Ed.), 2008, 1-37.
- Donaldson, G., Kamstra, M., An artificial neural network-GARCH model for international stock return volatility, Journal of Empirical Finance, 4(1)(1997), 17-46.
- Doran, J.S., Ronn, E.I., Computing the market price of volatility risk in the energy commodity markets, Journal of Banking and Finance, 32(12)(2008), 2541-2552.
- Dorn, D., Huberman, G., Sengmueller, P., Correlated trading and returns, Journal of Finance, 63, 2008, 885-920.
- Du, X.D, Yu, C.L., Hayes, D.J., Speculation and volatility spillover in the crude oil and agricultural commodity markets: A Bayesian analysis, Energy Economics, 33(3)(2011), 497-503.
- Dunis, C.L. and Huang, X., Forecasting and trading currency volatility: An application of recurrent neural regression and model combination. Journal of Forecasting, 21(5)(2002), 317-354.



- Dunis, C.L., Chen, Y.X., Alternative volatility models for risk management and trading: application to the EUR/USD and USD/JPY rates, *Trading and Regulation* 11 (2005) 126-156.
- Elman, J.L., Finding structures in time. *Cognitive Science*, 14(1990)(2), 179–211.
- Engle, R and Patton, S.J., What good is a volatility model? *Quantitative Finance*, 1(2001)(2), 237–245.
- Engle, R., Sarkar, D. Premiums-discounts and exchange traded funds, *Journal of Derivatives*, 2006, (Summer), 27-45.
- Fama, E.F., Efficient capital markets II, *Journal of Finance*, 46, 1991, 1575-1643.
- Fan, Y., Liang, Q., Wei, Y.M., A generalized pattern matching approach for multi-step prediction of crude oil price. *Energy Economics*, 30(3)(2008), 889-904.
- Fan J, Wang Y, Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of American Statistical Association*, 2007, 102, 1349–1362.
- Farmer, J.D., Market force, ecology and evolution, *Industrial and Corporate Change*, 11(5), 2002, 895-953.
- Farmer, J.D., Joshi, S., The price dynamics of common trading strategies, *Journal of Economic Behaviour and Organization*, 49(2), 2002, 149-171.
- Fassas, A. and Siriopoulos, C., The efficiency of VIX futures market – A panel data approach. *Journal of Alternative Investments*, 14(2012)(3), 55–65.
- Feng, L., Seaholes, M.S., Correlated trading and location, *Journal of Finance*, 59(5), 2004, 2117-2144.

- Fernandes, M., Medeiros, M.C., Scharth, M., Modelling and predicting the CBOE market volatility index, *Journal of Banking and Finance*, 40 (2014) 1-10.
- Fong, W.M., Yong, L.H.M., Chasing trends: recursive moving average trading rules and internet stocks, *Journal of Empirical Finance*, 12, 2005, 43-76.
- Fong, W.M., See, K.H., A Markov switching model of the conditional volatility of crude oil futures prices, *Energy Economics*, 24(1)(2002), 71-95.
- Froot, K., Teo, M., Style investing and institutional investors, *Journal of Financial and Quantitative Analysis*, 43(4), 2008, 883-906.
- Fujihara, R.A., Mougoue, M., Linear dependence, nonlinear dependence and petroleum futures market efficiency, *Journal of Futures Market*, 17(1)(1997), 75-99.
- Galariotis, E., Contrarian and momentum trading: a review of the literature, review of behavioural finance, 6(1), 2014, 63-82.
- Gastineau, G.L., Exchange traded funds: an introduction, *Journal of Portfolio Management*, 27(3), 2011, 88-96.
- Glosten, L.R., Jagannathan, R., Runkle, D.E., on the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, 48(5), 1993, 1779-1801.
- Griffin, J.E., Inference in infinite superposition of non-gaussian Ornstein Uhlenbeck process using bayesian nonparametric methods. *Journal of Financial Econometrics*, 3, 2010, 519-549.
- Grinblatt M., Titman, S., Wermers, R., Momentum investment strategies, Portfolio Performance, and herding: a study of mutual fund behaviour, *American Economic Review*, 85, 1995, 1088-1105.

- Hajizadeh, E., Seifi, A., Zarandi, M.F. and Turksen, I., A hybrid modelling approach for forecasting the volatility of S&P index return. *Expert Systems Application*, 39(2012)(1), 431–436.
- Haigh, M.S., Hranaiova, J., Overdahl, J.A., Hedge funds, volatility, and liquidity provision in energy futures markets, *Journal of Alternative Investments*, 9(4)(2007), 10-38.
- Hamid, S.a., Iqbal, Z., Using neural networks for forecasting volatility of S&P 500 Index futures prices, *Journal of Business Research*, 57 (2004) 1116-1125.
- Hansen, P.R. and Lunde, A., Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2006)(2), 127–161.
- Hansen, P.R., Huang, Z. and Shek, H.H., Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility. *Journal of Applied Econometrics*, 27(2012)(6), 877–906.
- Harper, J.T., Madura, J., Schnusenberg, O., Performance comparison between exchange traded funds and closed-end country funds, *Journal of International Financial Markets, Institutions, and Money*, 16(2), 2006, 104-122.
- Harvey, A., Ruiz, E., Shephard, N., Multivariate stochastic variance models. *Review of Economic Studies*, 61(1994)(2), 247–264 .
- Haugom, E., Westgaard, S., Solibakke, P.B., Lien, G., Realized volatility and the influence of market measures on predictability: Analysis of Nord Pool forward electricity data, *Energy Economics*, 33(6) (2011), 1206-1215.
- E. Haugom, H. Langeland, P. Molnár, S. Westgaard, Forecasting volatility of the US oil market, *Journal of Banking & Finance*, 47 (2014) 1-14

- Haykin, S., Neural networks: a comprehensive foundation. Prentice Hall Englewood Cliffs, 2007, NJ, USA.
- Hemanth Kumar, P., Patil, S.B., Estimation & forecasting of volatility using ARIMA, ARFIMA and Neural Network based techniques, Souvenir of the 2015 IEEE International Advance Computing Conference 10 (2015) 992-997.
- Herbert, J.H., Trading volume, maturity and natural gas futures price volatility, Energy Economics, 17(4)(1995), 293-299.
- Hirose, T., Kato, H.K., Bremer, M., Can margin traders predict future stock returns in Japan?, Pacific-basin Finance Journal, 17, 2009, 41-57.
- Hirshleifer, D., Teoh, S.H., Limited attention, information disclosure, and financial reporting, Journal of Accounting and Economics, 36, 2003, 337-386.
- Hirshleifer, D., Lim, S., Teoh, S.H., Limited investor attention and stock market misreactions to accounting information, Review of Asset Pricing Studies, 1(1), 2011, 35-73.
- Holmes, P.R., Kallinterakis, V., Feedback trading, in Volume 4 (Finance), 3<sup>rd</sup> Edition of Wiley Encyclopedia of Management, edited by K. Paudyal, John Wiley & Sons, Ltd: Oxford, UK.
- Huang, X. and Tauchen, G. (2005) The relative contribution of jumps to total price variance. Journal of Financial Econometrics, 3(2005)(4), 456-499.
- Huang, R.D., Masulis, R.W., Stoll, H.R., Energy shocks and financial markets, Journal of Futures Markets, 16(1)(1996), 1-27.

- Hung, J.C., Lee, M.C., Liu, H.C., Estimation of value-at-risk for energy commodities via fat-tailed GARCH models, *Energy Economics*, 30(3)(2008), 2326-2339.
- Hung, W., Lu, C.C., Lee, C.F., Mutual fund herding and its impact on stock returns: evidence from the taiwan stock market, *Pacific-Basin Finance Journal*, 18, 2010, 477-493.
- Z. Huang, H. Liu, T.Y. Wang, Modelling long memory volatility using realized measures of volatility: A realized HAR GARCH model, *Economic Modelling* 52 (2015) 812-821.
- Hyndman, R.J. and Koehler, A.B., Another look at measures of forecast accuracy, *International Journal of Forecasting*, 22(4)(2006), 679-688.
- Hyndman, R.J. and Athanasopoulos, G., *Forecasting: principles and practice*, OTexts.com (online open access), 2014.
- Jammazi, R. and Aloui, C., Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modelling. *Energy Economics*, 34(2012)(3), 828–841.
- Jares, T., Lavin, A., Japan and Hong-Kong exchange-traded funds (ETFs): discounts, returns and trading strategies. *Journal of Financial Services Research*, 2004, 25(1), 57-69.
- Jiang, G.J. and Oomen, R., Testing for jumps when asset prices are observed with noise - A swap variance approach. *Journal of Econometrics*, 144(2008)(2), 352–370
- Kallinterakis, V., Liu, F., Pantelous, A.A., On the impact of country ETFs' premiums and discounts over feedback trading, (2017) Working Paper.
- Kang, S.H., Kang, S.M., Yoon, S.M., Forecasting volatility of crude oil markets, *Energy Economics*, 31(1)(2009), 119-125.

- Kim, W., Wei, S.J., Offshore investment funds: monsters in emerging markets? *Journal of Development Economics*, 68(1), 2002a, 205-224.
- Kim, W., Wei, S.J., Foreign portfolio investors before and during a crisis, *Journal of International Economics*, 56(1), 2002b, 77-96.
- Kodres, L.E., The existence and impact of destabilizing positive feedback traders: Evidence from the S&P 500 index futures market, Working paper, Board of Governors of the Federal Reserve System.
- Kostovetsky, L. Index mutual funds and exchange-traded funds, *Journal of Portfolio Management*, 29(4), 2003, 80-92.
- Kotkatvuori-Ornberg, J., Measuring actual daily volatility from high frequency intraday returns of the S&P futures and index observations. *Expert Systems with Applications*, 43(2016), 213-222.
- Koutmos, G., Feedback trading and the autocorrelation pattern of stock returns: further empirical evidence, *Journal of International Money and Finance*, 16(4), 1997, 625-636.
- Koutmos, G., Pericli, A., Trigeorgis, L., Short-term dynamics in the Cyprus stock exchange, *European Journal of Finance*, 12(3), 2006, 205-216.
- Koutmos, G., Saidi, R., Positive feedback trading in emerging capital markets, *Applied Financial Economics*, 11, 2001, 291-297.
- Kourtis, A., Markellos, R.N., Symeonidis, L., An international comparison of implied, realized, and GARCH volatility forecast, *Journal of Futures Markets*, 36 (2016) 1164-1193.

- Kremer, S., Nautz, D., Causes and consequences of short-term institutional herding, *Journal of Banking and Finance*, 37, 2013, 1676-1686.
- Kristjanpoller, W., Fadic, A. and Minutolo, M., Volatility forecast using hybrid Neural Network models. *Expert Systems with Applications*, 41(2014)(5), 2437–2442.
- Kristjanpoller, W. and Minutolo, M., Gold price volatility: A forecasting approach using the Artificial Neural Network-GARCH model. *Expert Systems with Applications*, 42(2015)(20), 7245–7251.
- Kuo, W.S., Chung, S.L., Chang, C.Y., the impacts of individual and institutional trading on futures returns and volatility: evidence from emerging index futures markets, *Journal of Futures Markets*, 35(3), 2015, 222-244.
- B. Kurek, The information content of equity block trades on the Warsaw stock exchange: conventional and bootstrap approaches, *Journal of Forecasting*, 35(1) (2016) 43-53.
- Lahmiri, S., Boukadoum, M., An ensemble syetem based on Hybrid EGARCH-ANN with different distributional assumptions to predict S&P 500 intraday volatility, *Fluctuation and Noise Letters*, 14 (2015).
- Lakonishok, J., Shleifer, A., Vishny, R., The impact of institutional trading on stock prices, *Journal of Financial Economics*, 32, 1992, 23-43.
- Laopodis, N.T., Feedback trading and autocorrelation interactions in the foreign exchange market: further evidence, *Economic Modelling*, 22, 2005, 811-827.
- Lee, S.S., Jumps and information flow in financial markets. *Review of Financial Study*, 25(2012)(2), 439–479.

- Lee, S.S. and Hannig, J., Detecting jumps from Levy jump diffusion processes. *Journal of Financial Econometrics*, 96(2010)(2), 271–290.
- Lee, S.S. and Mykland, P.A., Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Study* 21(2008)(6), 2535–2563.
- Ljung, G.M. and Box, G.E.P., On a measure of lack of fit in time series models, *Biometrika*, 65(1987)(2), 297-303.
- Lin, S.X., Tamvakis, M.N., Spillover effects in energy futures markets, *Energy Economics*, 23(1)(2001), 43-56.
- Lin, S.C., Zhu, Z., Natural gas prices and the gas storage report: Public news and volatility in energy futures markets, *Journal of Futures Market*, 24(3)(2004), 283-313.
- Lin, Y.A., Swanson, P.E., Foreigners’ perceptions of US markets: do foreigners exhibit herding tendencies? *Journal of Economics and Business*, 60, 2008, 179-203.
- Liu, F., Pantelous, A.A., von Mettemheim, H., Forecasting and trading high frequency volatility on large indices. (2017) To appear in *Quantitative Finance*.
- Liu, Z., Ma, F., Wang, X., Xia, Z., Forecasting the realized volatility: the role of jumps, *Applied Economics Letters*, 23 (2016) 736-739.
- Lux, T. and Moreles-Arias, L., Relative forecasting performance of volatility models: Monte Carlo evidence. *Quantative Finance*, 13(2013), 1375–1394.
- Madura, J., Richie, N., Overreaction of exchange-traded funds during the bubble of 1998-2002, *Journal of Behavioral Finance*, 5, 2004, 91-104.



- Ma, F., Wei, Y., Chen, Y., Which is the better forecasting model? A comparison between HAR-RV and multifractality volatility, *Physica A: Statistical Mechanics and its Applications*, 405 (2014) 171-180.
- McAleer, M., Forecasting realized volatility with linear and nonlinear univariate models, *Journal of Economic Surveys*, 25 (2011) 6-18.
- Mensi, W., Beljid, M., Boubaker, A., Managi, S., Correlations and volatility spillovers across commodity and stock markets: Linking energies, food, and gold, *Economic Modelling*, 32(1)(2013), 15-22.
- Monfared, S.A., Enke, D., Volatility forecasting using a hybrid GJR-GARCH neural network model, *Procedia Computer Science*, 36(2014), 246-253.
- Nofsinger, J., Sias, R., Herding and feedback trading by institutional and individual investors, *Journal of Finance*, 45(6), 1999, 2263-2295.
- Osler, C.L., Stop-loss orders and price cascades in currency markets, *Journal of International Money and Finance*, 24 (2005) 219-241.
- Panella, M., Barcellona, F. and D'Ecclesia, R.L., Forecasting energy commodity prices using neural networks. *Advances in Decision Sciences*, 2012, Article ID 289810, 26 pages.
- Papavassiliou V.G., Allowing for jump measurement in volatility: A high-frequency financial data analysis of individual stocks. *Bulletin of Economic Research*, 68(2016)(2), 124-132.
- Papadimitriou, T., Gogas, P. and Stathakis, E., Forecasting energy markets using support vector machines. *Energy Economics*, 44(2014), 135–142.

- Park, H., Kim, N., Lee, J., Parametric models and non-parametric machine learning models for predicting option prices: Empirical comparison study over KOSPI 200 index options, *Expert Systems with Applications*, 41(11)(2014), 7653-7670.
- Prokopczuk, M., Symeonidis, L. and Wese Simen, C., Do jumps matter for volatility forecasting? evidence from markets. *Journal of Futures Markets*, 36(2015)(8), 758-792.
- Prokopczuk, M., Symeonidis, L., Wese Simen, C., Do jumps matter for volatility forecasting? Evidence from energy markets, *Journal of Futures Market*, 36 (2016) 758-792.
- Psaradellis, I., Sermpinis, G., Modelling and trading the U.S. implied volatility indices. Evidence from the VIX, VXN and VXD indices, *International Journal of Forecasting*, 32 (2016) 1268-1283.
- Qu, H., Chen, W., Niu, M.Y., Li, X.D., Forecasting realized volatility in electricity markets using logistic smooth transition heterogeneous autoregressive models, *Energy Economics*, 54(2016), 68-76.
- Regnier, E., Oil and energy price volatility, *Energy Economics*, 29(3)(2007), 405-427.
- Rossi, B., Sekhposyan, T., Understanding models' forecasting performance, *Journal of Econometrics*, 164 (2011) 158-172.
- Sadorsky, P., Modelling and forecasting petroleum futures volatility, *Energy Economics*, 28(4)(2006), 467-488.
- Sadorsky, P., Correlations and volatility spillovers between oil prices and the stock prices of clean energy and technology companies, *Energy Economics*, 32(1)(2012), 248-255.

- Schuppli, M., Bohl, M.T., do foreign institutional investors destabilize China's A-Share market, *Journal of International Financial Markets, Institutions and Money*, 20, 2010, 36-50.
- Sentana, E., Wadhwani, S., Feedback traders and stock return autocorrelations: Evidence from a century of daily data, *The Economic Journal*, 102(411), 1992, 415-426.
- Sevi, B., Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research*, 235(2014)(3), 643–659.
- Serletis, A., Maturity effects in energy futures, *Energy Economics*, 14(2)(1992), 150-157.
- Sermpinis, G., Dunis, C.L., Laws, J., Stasinakis, C., Forecasting and trading the EUR/USD exchange rate with stochastic Neural Network combination and time-varying leverage, *Decision Support System* 54 (2012) 316-329.
- Sermpinis, G., Jason, L., Dunis, c.L., Modelling and Trading the realised volatility of the FTSE100 Futures with Higher Order Neural Networks, *European Journal of Finance*, 19 (2013) 165-179.
- Shin, M., Zhong, M., Does realized volatility help bond yield density prediction? *International Journal of Forecasting*, 33(2) (2017) 373-389.
- Sias, R.W., Institutional Herding, *Review of Financial Studies*, 17, 2004, 165-206.
- Souček, M., Todorova, N., Realized volatility transmission between crude oil and equity futures market: A multivariate HAR approach, *Energy Economics*, 40 (2013) 586-597.
- Susmel, R., Thompson, A., Volatility, storage and convenience: Evidence from natural gas markets, *Journal of Futures Markets*, 17(1)(1997), 17-43.
- Taylor, S.J., *Modelling Financial Time Series*. Chichester, 1986, Wiley, UK.

- Tian, P.F., Yang, K., Chen, L.N., Realized volatility forecasting of agriculture commodity futures using the HAR model with time-varying sparsity, *International Journal of Forecasting*, 33(1) (2017) 132-52.
- Tung, W.L., Quek, C., Financial volatility trading using a self-organising neural-fuzzy semantic network and option straddle-based approach, *Expert Systems with Applications*, 38 (2011) 4668-4688.
- Voronkova, S., Bohl, M.T., Institutional traders' behaviour in an emerging stock market: Empirical evidence on Polish pension fund investors, *Journal of Business, Finance, Accounting*, 32, 2005, 1537-1560.
- Walls, W.D., Volatility, volume and maturity in electricity futures, *Applied Financial Economics*, 9(3)(1999), 283-287.
- Walter, A., Weber, M., Herding in the German mutual fund industry, *European Financial Management*, 12, 2006, 375-406.
- Y.D. Wang, F. Ma, Y. Wei, C.F. Wu, Forecasting realized volatility in a changing world: A dynamic model averaging approach, *Journal of Banking and Finance*, 64 (2016) 36-149.
- Wang, T., Wu, J., Yang, J., Realized volatility and correlation in energy futures markets, *Journal of Futures Markets*, 28(10)(2008), 993-1011.
- Watanabe, T., Margin requirements, positive feedback trading, and stock return autocorrelations: the case of Japan, *Applied Financial Economics*, 12, 2002, 395-403.
- Wermers, R., Mutual fund herding and the impact on stock prices, *Journal of Finance*, 54, 1999, 581-622.

- Weiner, R.J., Sheep in wolves' clothing? Speculators and price volatility in petroleum futures, *Quarterly Review of Economics and Finance*, 42(2)(2002), 391-400.
- Wilson B., Aggarwal, R., Inclan, C., Detecting volatility changes across the oil sector, *Journal of Futures Markets*, 16(3)(1996), 313-330.
- Wylie, S., Fund manager herding: a test of the accuracy of empirical results using U.K. data, *Journal of Business*, 78, 2005, 381-403.
- Xiao, Y., Xiao, J., Liu, J., Wang, S.Y., A multiscale modelling approach incorporating ARIMA and ANNs for financial market volatility, *Journal of Systems Science & Complexity*, 27 (2014) 225-236.
- Xiong, T., Bao, Y., Hu, Z., Beyond one-step-ahead forecasting: Evaluation of alternative multi-step-ahead forecasting models for crude oil prices. *Energy Economics*, 40(2008), 405-415.
- Zhang, L., Mykland, P. and Ait-Sahalia, Y., A tale of two time scales: Determining integrated volatility with noisy high frequency data. *Journal of American Statistical Association*, 100(2005), 1394-1411.
- Zhang, J.E., Shu, J. and Brenner, M. The new market for volatility trading, *Journal of Future Markets*, 30(2010)(9), 809-833.
- Zhou, B., High frequency data and volatility in foreign exchange rates. *Journal of Business & Economic Statistics*, 14(1996)(1), 45-52.