The impact of intraday periodicity and news announcements on high-frequency stock volatility

By

Yanying Guan

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Economics

Lancaster University

September 2020

Declaration

I hereby declare that this thesis is my own work and has not been submitted for the award of a higher degree elsewhere.

Yanying Guan

September 2020

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Professor Marwan Izzeldin and Professor Mike Tsionas, for their encouragement, support and guidance. I particularly thank Professor Marwan Izzeldin, who provided me with this PhD opportunity.

Contents

Abstract

High-frequency intraday financial data are commonly used in stock market volatility estimation and forecasting because they produce accurate results. However, little work to date has focused on the stylised facts of high-frequency returns, such as their tail properties, autocorrelations and leverage effects. One of the most discussed features of high-frequency returns is intraday periodicity, yet it is not well known how this feature operates in returns from data with different sampling schemes and frequencies. In addition, macroeconomic news announcements have been shown to have a large impact on first-moment and second-moment responses in financial markets. However, few existing models consider the effect of news on volatility estimation and forecasting, and those that do tend to treat it as a dummy variable, limiting its analytical power.

This thesis addresses these issues by reporting a study of the stylised facts of returns from S&P 500 stocks and the SPY index, and standardised returns from the latter, using various volatility measures in different financial regimes (i.e. before, during and after the 2008 financial crisis). It presents a comparison of the intraday patterns, jump frequencies, jump components and volatility forecasting of stock returns from calendar-time and business-time sampling schemes, as well as how these features are affected by intraday periodicity. It assesses the direct impact of macroeconomic news announcements on volatility estimation and forecasting for stock returns by incorporating significant news announcements as an index to identify the jumps caused by news in heterogeneous autoregressive (HAR) class models.

The results suggest that absolute intraday returns for high-frequency data exhibit autocorrelations and that aggregated returns display heavy tails. Standardising the returns of the SPY index using eleven different volatility measures produces distributions that are closer to a normal distribution. We find that various volatility measures are significantly correlated with trading volume, and hence that HAR-class models that include trading volume yield better volatility forecasting results than existing models. However, this effect may be limited to data from the relatively non-volatile pre-crisis and post-crisis periods. High-frequency returns based on business-time sampling have smaller jump frequencies, jump components and intraday periodicity patterns, than calendar-time data, which may be useful for volatility analysis. Intraday periodicity has a notable impact on jumps for both sampling schemes, however, and adjusting for intraday periodicity produces fewer jumps for all returns and smaller jump components for the majority. We also find that the forecasting results for less volatile data, such as healthcare stocks and data from the post-crisis period, improved after filtering for intraday periodicity. Finally, macroeconomic news announcements can affect jump components, and considering news outlets in HAR models can improve the forecasting results. The thesis thus contributes to our understanding of the factors affecting stock market volatility by providing evidence in support of including trading volume, efficient intraday periodicity estimators and news surprise in volatility estimation and forecasting models.

Chapter 1

In this chapter, we review the development of financial asset volatility estimation and forecasting methods in academic literature in light of the increasing availability of intraday data in recent decades. The most popular parametric and non-parametric models used for high-frequency asset return estimation and forecasting are the conditional volatility GARCH and HAR models respectively. The generalised autoregressive conditional heteroscedasticity (GARCH) model was introduced initially to estimate yearly and monthly asset returns and was then extended to fit intraday returns. The heteroscedastic autoregressive (HAR) model was introduced based on the notion that the non-parametric volatility measure known as realised variance (RV) has a long-term dependence and that lagged daily, weekly and monthly RV can provide useful information about current volatility.

Dramatic changes in intraday return volatility – known as jump components – have been used together with bi-power variation (BV) to separate the jump and continuous components from RV. This method has been incorporated into HAR-class models in recent years, which often yield better results than the traditional GARCH model. Since then, studies have integrated more advanced jump component measurements such as (corrected) threshold bi-power variation (TBV and CTBV) into HAR models, further improving their performance. Research on the stylised facts of intraday returns such as their long-memory properties and intraday periodicity has also contributed to improvements in estimation and forecasting for HAR and GARCH models. Finally, some studies have considered in their modelling a range of factors that are suspected to cause dramatic changes in asset returns, most notably announcements of macroeconomic news. It is clear that news announcements are likely to play a major role in stock market volatility, and so better integration of this phenomenon into volatility estimation and forecasting models using high-frequency data is a priority for future research in the field.

Chapter 2

This chapter investigates the stylised facts of high-frequency returns, together with eleven different volatility measures, in different financial regimes. We find autocorrelations for absolute intraday returns and volatility measures, and heavy tails for aggregated returns. Aggregated returns are not normally distributed, yet standardising returns for the SPY index using the set of volatility measures results in distributions that are significantly closer to normal. We also find a significant correlation between various volatility measures and trading volume, and thus that the inclusion of trading volume in HAR-class models produces better RV forecasting results, at least for the relatively non-volatile post-crisis period.

Chapter 3

In this chapter, we examine the volatility patterns for high-frequency returns, using business-time sampling and calendar-time sampling, along with the performances of different non-parametric intraday periodicity estimators for these two sampling schemes for stocks and the SPY index in different financial regimes. We also study the impact of intraday periodicity on jump frequency, jump components and volatility forecasting. The results provide empirical evidence that business-time sampling returns have fewer jumps, smaller jump components and less marked intraday periodicity patterns than calendar-time sampling returns. Filtering for intraday periodicity reduces jump frequency for both calendar-time and business-time returns and it reduces jump components for most returns. Finally, we find that HAR-class models perform better at RV forecasting for less volatile data, such as healthcare stocks and data from the post-crisis period, after filtering for intraday periodicity. We conclude that business-time sampling may be more useful for volatility analysis by virtue of its lower jump frequency, jump components and intraday periodicity, and that filtering for intraday periodicity is most effective for less volatile data.

Chapter 4

This chapter investigates the impact of different macroeconomic news announcements on the jump components of returns from twenty-one stocks with high trading volumes and low jump frequencies from 2000 to 2016. It also assesses the impact of news announcements on co-jumps between stocks. The results show that positive news surprises from the Consumer Price Index and Initial Jobless Claims sources, and negative news surprises from the University of Michigan Consumer Sentiment Index, have the biggest effects on stock jump components in the pre-crisis and post-crisis period respectively. We also find that treating the jump components that are caused by news separately in the HAR model can improve the model's forecasting performance. Co-jumps also have a significant effect on jump components, yet the number of news announcements that they capture is limited. Co-jumps are too few to have a significant impact on the HAR model when they are used as an index to separate news-related jumps from jumps caused by other factors. Finally, we do not find significant evidence that the effect of macroeconomic news announcements on stock jump components is related to their trading volumes or jump frequencies.

Introduction

Study overview

Financial market volatility (henceforth 'volatility') refers to the dispersion of financial returns for assets, security and market indices, which is widely used in risk management, portfolio allocation and asset pricing. Volatility can provide useful information for financial market participants, investment bankers, regulators and government agencies. In this thesis, we investigate how the stylised facts of high-frequency returns and the consideration of announcements of macroeconomic news can improve volatility estimation and forecasting.

High-frequency intraday data has in recent decades become available for financial analysis and has emerged as a key feature of stock volatility. However, the stylised facts of high-frequency returns, such as autocorrelation and leverage effects, are rarely discussed in previous literature. In Chapter 2 of this thesis, we examine the stylised facts (distribution properties, autocorrelation and leverage effects) of high-frequency intraday returns in different financial regimes and after being standardised using eleven different volatility measures.

One of the most widely discussed stylised facts of stock returns is intraday periodicity. The most recent and popular non-parametric intraday periodicity estimation methods are the Weighted Standard Deviation estimator (WSD; Boudt et al. 2011b) and the Shortest Half scale estimator (Short-H; Rousseeuw & Leroy 1988). They share some methodological similarities with the earlier Standard Deviation periodicity estimator (SD; Taylor & Xu 1997), which uses the standard deviation of standardised returns from the same local window. The Short-H and WSD estimators, however, consider order statistics and weights in the estimations, and are thus more robust to the presence of jumps; this represents an advantage compared to the SD estimator. Boudt et al. (2011b) use simulation results to demonstrate that the WSD estimator is more robust to jumps than Short-H; however, they do not provide evidence from empirical data. In Chapter 3, we discuss the impact of the SD, WSD and Short-H estimators on jump frequency and volatility components in different financial regimes, using intraday returns measured using both business-time and calendar-time sampling.

The increasing availability of high-frequency data has enabled researchers to explore fine-grained patterns in financial data. A major innovation is Barndorff-Nielsen and Shephard's (2004) development of a non-parametric method which can separate the volatility caused by large changes in returns in one day from the total volatility. The high volatility caused by large changes in returns within one day is known as the jump component, and the remaining variation in returns in that day is known as the continuous component. Barndorff-Nielsen and Shephard's (2004) volatility measures are bi-power variation (BV), which estimates the volatility of the continuous components, and realised variance (RV), which estimates the volatility of all returns (both jump and continuous components). Therefore, separating BV from RV leaves only the volatility caused by the large discontinuous returns (the jump components). This is further developed by Corsi et al.'s (2010) threshold bi-power variation (TBV) measure, which helps mitigate estimation bias, and their corrected threshold bi-power variation measure (CTBV), which helps account for the over-estimations produced by TBV. The stylised facts of returns standardised using these four volatility measures, along with seven other bias-corrected volatility measures, are studied in Chapter 2, together with their correlations with trading volume. In Chapter 3, we use BV, TBV and CTBV in our estimation of jump components to examine how they are affected by intraday periodicity. The relative performances of these three non-parametric methods are also compared in Chapter 4 as part of our study of the impact of macroeconomic news announcements on volatility.

Previous work typically uses two models that incorporate intraday asset returns for volatility estimation and forecasting – these are the parametric (Generalised) Autoregressive Conditional Heteroscedasticity (GARCH) model (Engle, 1982) and the non-parametric Heterogeneous Autoregressive (HAR) model (Corsi, 2009). The GARCH model was initially used for modelling monthly or yearly rather than intraday data (e.g. Akgiray, 1989), as only low-frequency data were widely available until the early 1990s. Since then, many scholars have adapted the GARCH model to fit intraday financial data (e.g. Andersen & Bollerslev 1997, 1998; Walsh & Tsou 1998; Andersen et al, 1999, 2003; Engle & Sokalska, 2012). Heterogeneous Autoregressive (HAR)-family models (Corsi, 2009) are non-parametric volatility estimation and forecasting models based on the non-parametric RV and BV volatility measures from Barndorff-Nielsen and Shephard (2004). The original HAR model uses lagged daily, weekly and monthly RV, while the Heterogeneous Autoregressive model with Jumps (HAR-J) variant uses both RV and BV. By considering the continuous and jump components separately in the model via BV, Andersen et al. (2007a) find that the HAR-J model produces better results than the original. Corsi et al. (2010) later apply TBV and CBTV to the HAR-J model, which yields further improvements. In Chapter 2 of this thesis, we adapt the HAR model by considering trading volume as a term in light of its positive relationship with eleven different volatility measures. In Chapter 3, we compare the forecasting performances of HAR-class models before and after filtering for intraday periodicity in order to investigate the impact of intraday periodicity on volatility forecasting for both calendar-time sampling and business-time sampling data. In Chapter 4, we further extend HAR-family models by differentiating between the jump components caused by macroeconomic news announcements and those caused by other factors.

The data we use in this thesis are 30-second, 60-second, 150-second and 300-second (5-minute) calendar-time sampling returns (Chapters 2, 3 and 4) and 300-second business-time sampling returns (Chapter 3) for S&P 500 stocks and the SPY index. The data are from 2000 to 2016, which can be divided into three financial regimes: the pre-financial crisis period (2000-2007), the crisis period (January 2008 to June 2009) and the post-crisis period (July 2009 to December 2016). We consider the whole data set in our analysis of the stylised facts of high-frequency returns (Chapter 2), our comparison of volatility and intraday periodicity patterns between different sampling schemes (Chapter 3) and the effect of news announcements on stock market volatility forecasting (Chapter 4), as well as data broken down by different regimes. Of particular interest in these analyses are stocks from the information technology (IT) sector, which have undergone a dramatic increase in volumes and prices over the last three decades due to rapid technological developments. IT stocks' high sensitivity to market forces thus make them excellent candidates for the study of returns' stylised facts and the impact of intraday periodicity and announcements of macroeconomic news on financial return volatility.

The aim of this thesis, therefore, is to help progress our understanding of the stylised facts of high-frequency stock returns and the factors affecting stock market volatility, namely the use of different sampling schemes, the impact of trading volumes, intraday periodicity, news announcements and jumps, so that more effective ways of improving models for forecasting stock volatility can be developed.

Thesis outline

In Chapter 1, we discuss the development in the literature of two families of intraday return volatility models: the parametric GARCH family and the non-parametric HAR family. In particular, we focus on the importance of further study of the stylised facts of high-frequency data and how to extend HAR-family models by incorporating different intraday periodicity estimators and macroeconomic news announcements.

Chapter 2 discusses the stylised facts of high-frequency intraday returns and aggregated high-frequency returns standardised using eleven different volatility measures. We also study the correlations between various volatility measures and trading volume in different financial regimes. Finally, the performance of a new extension to the HAR model that incorporates trading volume is compared with other HAR-family models in different financial regimes.

Chapter 3 presents a discussion of different volatility patterns for high-frequency data using business-time sampling and calendar-time sampling. We also investigate the impact of three non-parametric intraday periodicity estimators – SD, WSD and Short-H – on jump frequency, jump components and volatility forecasting in different financial regimes.

Chapter 4 studies the impact of macroeconomic news announcements on volatility measures and on HAR-class models. We first discuss how news affects the jump components estimated via BV, TBV and CTBV during different financial regimes. We then analyse the relationship between the co-jumps caused by news announcements and the daily jump components of various stocks. Finally, we examine the performance of HAR-class models that incorporate the effect of news announcements.

This is followed by a short concluding chapter which summarises the findings and offers some suggestions for future research.

Chapter 1 – The effect of jumps, intraday periodicity and news announcements on financial market volatility: A review

1.1 Overview

The first two sections of this chapter review the development of methods for estimating the important features of intraday returns, namely intraday periodicity and jumps. Section 1.3 discusses how volatility estimation and forecasting models can be improved by using high-frequency intraday data. We particularly focus on the development of parametric GARCH-family models and non-parametric HAR-family models. We then discuss the importance of considering news announcements' effects on the volatility of financial markets, as shown in previous literature. The final section of this chapter sets forth how this thesis makes an original contribution to the field by addressing some of the gaps in our understanding identified in the preceding sections.

1.2 Stylised facts of intraday data

1.2.1 Autocorrelation and distributional properties for intraday data

High-frequency intraday data can be used to study market microstructures (e.g. Goodhart & O'Hara, 1997; O'Hara, 2015) and can be employed in volatility estimation and forecasting (e.g. Andersen, 1997; Corsi et al., 2010). High-frequency data have unique stylised facts compared to low-frequency data, such as autocorrelations and leverage effects, but these differences are considered by few previous studies. Guillaume et al. (1997) find evidence that distributions of price changes in high-frequency foreign exchange rates are symmetric and fat-tailed with decreasing leptokurtosis. Shakeel and Srivastava (in press) investigate the autocorrelation and distributional properties for tick-by-tick S&P CNX NIFTY futures index data and find that high-frequency returns are positively skewed and have leptokurtic distribution. They also find that intraday returns display a slow decay in autocorrelations, with significant correlations at the first two lags. This thesis contributes to our knowledge of the properties of high-frequency data by investigating the autocorrelations, tail properties and leverage effects of intraday returns in different financial regimes in Chapter 2. The chapter also examines changes in the stylised facts of intraday returns when they are standardised using different volatility measures.

1.2.2 Intraday periodicity

Intraday periodicity refers to the systematic patterns of intraday return volatility over the course of a trading day, which are mainly caused by variation in trading volumes and bid-ask spreads. The intraday periodicity components that are observed in high-frequency data are mainly induced by regular trading patterns, such as the openings and closings of markets (Andersen & Bollerslev, 1997; Erdemlioglu et al., 2015). Andersen et al. (1998, 2003, 2007a) find a significant increase in market volatility immediately after macroeconomic news is announced. Bollerslev et al. (2008) report that the peak of the intraday pattern of realised variation tends to occur at 10am EST, which is when news announcements are usually scheduled to be released. Previous studies document that the presence of periodicity in financial data has an impact on intraday volatility estimation and forecasting (Andersen & Bollerslev, 1998), co-volatility estimation of multivariate price processes (Boudt et al., 2011a) and intraday jump detection (Boudt et al., 2011b). In Chapter 3 of this thesis, we study the impact of periodicity on jumps and volatility forecasting using data from different sampling schemes.

1.2.3 Periodicity estimation

Estimation methods for intraday periodicity in previous literature can be divided into parametric and non-parametric methods. One of the best-known parametric intraday periodicity estimators is the Flexible Fourier form (Andersen & Bollerslev, 1997). The authors use this estimator to filter intraday periodicity for exchange returns and equity returns. By fitting these filtered returns in the autoregressive conditional heteroskedasticity (GARCH) model, they find that the intraday periodicity-adjusted Deutsche Mark/US Dollar exchange rate and Standard and Poor's 500 (S&P 500) returns fit the GARCH model better. This highlights the fact that accurately measuring intraday periodicity can improve the efficiency of volatility estimation. It also shows the complexity of intraday volatility, which demands robust measurement methods.

Non-parametric periodicity estimators are a product of the scale estimate of the standardised returns that have the same periodicity factor – that is, those that share similar daily patterns. Taylor and Xu (1997) developed the SD periodicity estimator, which is based on the standard deviation (SD) of standardised returns. However, if jumps are present, SD produces inaccurate results (Boudt et al., 2011b). The Shortest Half (Short-H) estimator (Rousseeuw & Leroy, 1988) attempts to deal with jumps by taking into account order statistics. This method is more robust than SD even when jumps are present, and in their comparison of a variety of scale estimators, Martin and Zamar (1993) find that Short-H produces the smallest possible maximum bias. However, although Shortest Half is a relatively robust estimator of periodicity when jumps are present, it has only 37% efficiency when they are absent. In order to mitigate this bias caused by jumps, later work introduced more robust scale estimators. Boudt et al.'s (2011b) Weighted Standard Deviation (WSD) is one such method, which includes a weight function in the estimation based on the standardised return and the Shortest Half estimate. Boudt et al. (2011b) argue that the WSD is not only highly robust when estimating periodicity of returns in the presence of jumps, but also in the absence of jumps.

This thesis contributes to the literature on periodicity estimation by comparing the SD, WSD and Short-H estimators when filtering for intraday periodicity in intraday stock returns using calendar-time and business-time sampling data in Chapter 3. It also investigates the impact of intraday periodicity on volatility forecasting and jumps in intraday returns.

1.3 The relationship between intraday periodicity, volatility and jumps

1.3.1 Non-parametric volatility measures

Early studies (e.g. French et al. 1987; Schwert 1989, 1990) use standard deviations or variances as model-free volatility measures because they summarise the probability of extreme values of returns. Large standard deviations or variances indicate a high probability of extreme values. Schwert (1989, 1990) measures monthly volatility using the variance of monthly returns and by summing the square of daily returns. French et al.'s (1987) alternative method of measuring monthly volatility involves calculating the product of the sum of squared daily returns and the sum of the product of adjacent returns. They suggest that monthly volatility are more accurate than other volatility measures such as twelve-month rolling estimates at estimating risk premiums (Officer, 1973).

The availability of high-frequency data makes it possible to observe price movements at very fine intervals of time. This in turn has led to the development of daily measures that are calculated from high-frequency data. For example, Barndorff-Nielsen and Shephard's (2002) realised daily variance (RV) method uses continuously recorded transaction prices. Because the realised volatility in multivariate context is a product of the realised variance and covariance of intraday high-frequency returns (Zhao and Li, 2010) and in univariate context is the square root of realised variance, this measure is able to adequately characterise the distributional properties of stock return volatility. In addition, these model-free volatility measures are also easy to implement in high dimensional contexts because they are better at characterising the distributional and dynamic properties of correlations compared to the traditional multivariate Autoregressive Conditional Heteroscedasticity (ARCH) models (Engle, 1982).

Moreover, many asset pricing models (such as the option pricing model) are based on continuous-time models. Therefore, GARCH models, which are based on discrete-time formulation, are unable to construct good volatility measures for these pricing models. Hull and White (1987) find that the distribution of integrated volatility of an underlying asset is the determinant of the option price. Realised volatility converges to the integrated volatility when the sampling frequency for the intraday returns is close to zero (Barndorff-Nielsen & Shephard, 2002). Therefore, realised volatility is commonly used to estimate integrated volatility in asset pricing models.

1.3.2 Estimation of jumps

Since the increase in availability of intraday data in financial analysis, more studies find the evidence of the presence of jumps and find the importance of jumps as they are larger and more visible in the high-frequency intraday data. Given that jumps are among the main factors that contribute to the observed excess kurtosis in unconditional distributions of prices or returns, identifying jumps is central in risk pricing and estimation. Previous work highlights the importance of jumps in volatility. For example, Carr and Wu (2003) find strong evidence for the presence of jump components and continuous components in the S&P 500 index. They find that the decline of jump components in volatility can increase the slope of the OTM S&P 500 index option plot, therefore affecting the pricing for the

S&P 500 index option. Eraker et al. (2003) also find strong evidence for the presence of jumps in both volatility and returns for option data; they argue that including jumps in volatility estimation affects derivative pricing because the misspecification problems in the stochastic volatility (SV) models are removed when considering jumps. In this section, we discuss how to incorporate jumps in volatility estimation and forecasting for different financial assets.

Two early studies are Merton (1976) and Hull and White (1987), who show that it is important to consider the impact of jumps in the pricing of options. Since the millennium, the number of studies that find evidence that jumps play an important role in asset volatility has increased dramatically, especially those that use high-frequency data. This is because the increased availability of high-frequency intraday financial data since the 1990s has allowed many scholars to examine financial market volatility to an intraday level, thus enabling them to find more visible evidence of the presence of jumps in intraday volatility. Much of this work has demonstrated the importance of incorporating jumps into intraday volatility models when estimating and forecasting different assets. For example, Andersen et al. (2002) and Chernov et al. (2003) agree that financial volatility models (e.g. stochastic volatility models) provide satisfactory estimation of option prices if the models allow for both time-varying volatility and jump effects. Duffie et al. (2000) propose the Stochastic Volatility Model with Simultaneous and Correlated Jumps in Returns and Volatility (SVJJ) model, which allows for finite jumps in both volatility and prices. They find that the over-pricing problems present in Bakshi et al.'s (1997) Stochastic Volatility Model with Jumps in returns (SV-J) are alleviated when the model considers jumps in volatility for out-of-the-money (OTM) calls.

By further investigating the features of jumps, many studies (e.g. Barndorff-Nielsen & Shephard, 2004, 2006) suggest that separating the variation of the financial assets into the parts that are and are not caused by jumps will be useful for volatility analysis. Before the era of high-frequency data, estimation of jump components was a rather tedious task because smaller-sized jumps in low-frequency financial data (e.g. weekly and monthly data) are harder to detect. High-frequency data makes it possible to decompose the total daily return variation into its continuous and discontinuous components.

Realised variance (RV), which is the sum of squares of the intraday returns of a given day, contains both continuous and discontinuous components. Three methods of splitting the components of RV into its two components (continuous and jump components) are typically used. The first method provides an estimate of a realised measure known as Bi-power Variation (BV; Barndorff-Nielsen & Shephard, 2004, 2006). Bi-power variation is estimated based on the sum of the same day absolute intraday returns at time *t* multiplied by the absolute returns collected at the previous time. The number of intraday returns (the number of *t*s) depends on the chosen sampling frequency. Corsi et al. (2010) argue that BV is biased and tends to over-estimate the continuous components, prompting them to propose the so-called Threshold Bi-power Variation (TBV) method, which adds a threshold to BV in order reduce the bias that it causes. Aït-Sahalia and Jacod (2012) present an alternative framework to measure the presence of the relative components of the quadratic variation process, which involves defining the jump components using a discretely sampled semi-martingale beyond its volatility.

Aït-Sahalia and Xiu (2016) find that co-movements in asset volatility can be separated into continuous and jump co-movements. They suggest that separating the continuous and discontinuous components of assets is important in portfolio optimisation settings, especially during financial crises. During a financial crisis, more macroeconomic news is released into the market, and so the correlations between asset classes increase. The increase in co-movement of assets can be explained by the increase in co-movement of either the continuous part or the jump part of those assets. Knowing the source of increased correlations between assets is critical in making optimal hedging decisions, thus it is important to separate the continuous components from the jump components to find out how unusual shocks affect the co-movements of asset classes. Aït-Sahalia and Xiu (2016) find that co-movements in the continuous components contribute more to the correlation between two asset returns than co-movements in the jump components.

To conclude, recent studies have demonstrated the importance of jumps in volatility estimation based on high-frequency data. Scholars have suggested methods to improve volatility estimation and forecasting by introducing more advanced methods of detecting and estimating the variations in jumps and taking them into account in volatility models.

1.3.3 Testing for jumps

In order to examine the extent to which jumps have a notable effect on financial market volatility, tests are used to detect the frequency and nature of jumps. Barndorff-Nielsen and Shephard (2006) and Jiang and Oomen (2008) introduce two early non-parametric jump tests (referred to as BN-S and JO respectively) whose results are robust to the presence of leverage effects and to infinite jump activity and macroeconomic noise respectively. Barndorff-Nielsen and Shephard (2006) test the presence of jumps in a time interval by differentiating between realised variance and the continuous components, the latter of which are measured using bi-power variation. Jiang and Oomen (2008) adapt BN-S by using swap variance instead of bi-power variation. Barnorff-Nielsen and Shephard's method (2004, 2006) is able to detect the presence of jumps in exchange rates, while Andersen et al. (2007a) use it on fixed income and equity data and conclude that there are many jumps in the data.

However, Corsi et al. (2010) find that the BN-S method with bi-power variation tends to overestimate the jump components. They introduce a threshold into the non-parametric measures and produce threshold multi-power variation. Although the BN-S test with threshold multi-power variation performs better than multi-power variation in detecting jump components, Theodosiou and Zikes' (2011) simulation study finds that the choice of threshold affects the results of the jump test, and that its performance is dependent on the trade-off between size and power. A test with a low threshold is able to detect jumps more precisely, yet it may also increase the probability of detecting the presence of infinite jumps from a no-jump series. In addition, because both the JO test and BN-S test are based on integrated quantities, they are unable to detect the number of jumps accurately. They are also unable to test the size and time of jumps. Lee and Mykland (2007) introduce another jump detection test which is based on the returns scaled by the estimate of a local volatility measure known as contiguous intraday returns. The authors find that this new jump test out-performs the BN-S and JO tests in simulation studies by producing less biased detection rates. Lee and Mykland's (2007) jump test does not suffer from the same problem that the BN-S and JO jump tests do as it is not based on the integrated quantities, which make the tests unable to distinguish the difference between two small jumps and one large jump in volatility.

Dumitru and Urga (2012) compare various jump testing procedures and conclude that no single test is uniformly powerful at a specific sampling frequency. As such, they suggest a combination of tests to be employed at different sampling frequencies. However, jumps appear to change in nature with the sampling frequency, which makes interpreting average jump statistics over sampling frequencies unreliable. Aït-Sahalia and Jacod (2012) provide an estimate of the so-called beta index of jump activity, which mostly reflects the concentration of small jumps. At high sampling frequencies they find that jumps are characterised by infinite levels of activity and finite levels of activity at low frequencies. Thus, averaging frequencies in jump tests might not be the ideal approach to follow given the changing nature of jumps.

1.3.4 The impact of intraday periodicity on jumps

According to Andersen and Bollerslev (1997), it is important to consider jumps and intraday periodicity in non-parametric volatility forecasting. However, they consider the impacts of jumps and intraday periodicity on return volatility separately, while later work has argued that intraday periodicity and jumps affect each other when estimating and forecasting volatility. For example, Aït-Sahalia and Xiu (2016) find that jumps and co-jumps between assets have intraday patterns. They report that a large proportion of jumps can be observed in predicted time, because those jumps are the results of the surprise in scheduled news announcements. Some scheduled news announced before the openings of the US market tends to have a larger impact on European trading markets and causes larger jumps. Scheduled news announcements from the US Energy Information Administration also tends to produce large jumps on Wednesday at 10.30 am EST during US trading time. Boudt et al. (2011b) show that allowing for intraday periodicity can reduce bias in detecting jumps, while neglecting the presence of jumps can cause bias. Additionally, Erdemlioglu et al. (2015) find that the presence of intraday periodicity has an impact on the truncation mechanism in truncated power variation, which is commonly used in detecting the jump components in returns' volatility. Therefore, in order to improve the jump detection, they filter out the intraday periodicity from returns by using the periodicity component estimator known as Weighted Standard Deviation (WSD). They find that the WSD-filtered jump test performs better than other conventional tests at detecting jumps.

The most efficient intraday periodicity estimators introduced in recent decades, including the WSD estimator (Boudt et al., 2011b) are discussed in Section 1.2.2. However, they are mainly used to show the improved intraday periodicity pattern estimations for different assets. In this thesis, we incorporate these efficient intraday periodicity estimators into a volatility estimation and forecasting model and investigate how they affect volatility estimation and forecasting of stock returns.

1.4 Volatility forecasting models

1.4.1 Early forecasting models

Forecasting models were initially developed to estimate and forecast low-frequency data such as daily, weekly and monthly data. The extreme value volatility estimator was an early method for forecasting used and advocated by Taylor (1987) and Wiggins (1992). Taylor (1987) finds that the weighted average composite forecast performs best when forecasting 1- to 20-day Deutsche Mark/US dollar future volatility using high, low and closing prices. Alford and Boatsman (1995) use weekly and monthly data to improve five-years-ahead volatility forecasting by using the historical volatility based on the standard deviation of past returns in a fixed interval, which is called the HIS method. In order to improve the forecast returns, they used an HIS adjusted with the 'Shrinkage' forecasting method. This method adjusts the historical volatility based on the volatility estimated comparable firms from the same industrial sector. Figlewski (1997) finds that the volatility mean reversion is difficult to adjust when using daily data to forecast long-term volatility. They hence argue for the use of long horizontal historical data for forecasting using monthly data. This early work showed the potential of using volatility for forecasting using low-frequency data but was limited by the unavailability of intraday data, which only emerged in the late 1980s and early 1990s.

1.4.2 GARCH model

As one of the most popular conditional volatility models since the late 1980s, the ARCH model was first introduced by Engle (1982) to estimate UK economic inflation data. The advantage of this model compared to previous methods (e.g. Taylor, 1987; Wiggins, 1992) is that it allows the weights of past error variance to be estimated based on the data, but not under the assumption that the square residuals for every day in the past provides equal information for the conditional variance in the future. The ARCH model assumes that the square residuals follow an autoregressive (AR) pattern because more recent days are more relevant to future variance. This model can then be used to estimate and forecast the conditional variance of financial asset returns. Taylor (1986) uses ARCH family models in volatility estimation for various financial data, including 23 spot prices and 17 future prices. The average forecasting results for the conditional variation for those spot prices and future prices are best when using two modified ARCH processes versus the benchmark forecast that uses the natural estimate of the time series volatility. In the same year, the generalised ARCH (GARCH) model was proposed by Bollerslev (1986) for conditional variance forecasting and economic analysis. The weights for the past square residuals in this model are not only estimated based on the data, but also the weights can never go to zero, as the error variances are assumed to follow an autoregressive moving average (ARMA) model. Bollerslev (1986) finds that the GARCH (1,1) model has a better lag structure and performs better at describing inflation rates than the ARCH (8) model (Engle & Kraft 1983). Later, Akigray (1989) uses both the ARCH (2) and GARCH (1,1) to forecast the conditional variance of value-weighted and equal-weighted indices from the US Center for Research in Security Prices. He finds that the GARCH model out-performs the Exponentially

Weighted Moving Average (EWMA), HIS and ARCH models in monthly conditional variance forecasting. Akigray (1989) also suggests that the GARCH model can provide useful information for understanding the relationship between asset returns and volatility. This study is the first time that the GARCH model was used in financial volatility forecasting, and since then it has become one of the most popular volatility estimators and forecasting models for financial asset returns up to the present day.

1.4.3 Extensions for GARCH models

There are different extensions for GARCH models that capture different features of asset volatility. One of the early extensions for GARCH models is the integrated GARCH (IGARCH; Engle & Bollerslev 1986). The authors argue that if the sum of the persistent parameters for the first lag of the squared residuals and the first lag of the conditional variance is equal to 1 in the GARCH (1,1) model, then that indicates that the conditional variance has a unit root. This also means that the current shock will persistently affect the conditional variance forecast. Therefore, Engle & Bollerslev (1986) introduce the IGARCH model, which adds a restriction to the GARCH model in order to make the GARCH process a unit root process.

Choudhry (1995) also uses the IGARCH model to test the monthly stock returns from five European countries from 1919 to 1936 and finds that the persistence measurements for a few stocks in certain periods are significantly less than 1 (i.e. not persistent). Therefore, he concludes that the shocks have a persistent impact for stock volatility for most stocks in most of the periods he studied. Baillie et al. (1996) further extend the IGARCH model to the Fractional Integrated Generalised Auto-Regressive Conditionally Heteroscedastic (FIGARCH) model because they find that the impact of shocks is not permanent on the GARCH process, as assumed in the IGARCH model. Instead, the impact of shocks on conditional variance will die out in long-term forecasting. Therefore, they modify the IGARCH model by considering the features of the fractional order of integration in the mean equation in order to capture the slow hyperbolic rate of decay for the lagged squared residuals for the GARCH process. They use fractionally integrated GARCH (FIGARCH), IGARCH and GARCH models to fit the daily Deutsche Mark/US dollar spot exchange rates and find that the FIGARCH model out-performs the other two models at describing the data.

Another two extensions for the GARCH model are the GARCH-in-mean (GARCH-M) model and the exponential generalised autoregressive conditional heteroscedastic (EGARCH) model, which are introduced in Engle et al. (1987) and Nelson (1991) respectively. The GARCH-M model extends the GARCH model by adding a heteroscedasticity term into the mean equation in order to assume that the series is linearly affected by conditional variance. The parameter of the conditional variance term in the mean equation is called the risk premium parameter, which reflects how much the returns of an asset are expected to exceed the returns of a risk-free asset. The GARCH-M model fixes the problem encountered by the original GARCH model that the latter is unable to capture the linear impact of the risk as measured by conditional variance on the time series.

The EGARCH model extends the GARCH model by assuming the lagged squared innovation term as a function of standard normal variable or a variable from a generalised error distribution. Additionally, logarithm transformations for conditional variance are considered to overcome some shortfalls of the GARCH model, such as the fact that its estimated coefficients always violate its constraints imposed in the parameters. Also, it is difficult to interpret the persistence of shocks on future volatility in the GARCH model. Nelson (1991) applies the EGARCH model to CRSP Value-Weighted Market Index data from 1962 to 1987 and finds that the EGARCH model fits the conditional variance data very well. Both the GARCH-M and EGARCH model are mainly used in asset pricing contexts to find out the relationship between series with risk premia.

Bera and Higgins (1992) propose the Nonlinear Asymmetric GARCH (1,1) (NGARCH) model, which captures the leverage effect of the returns on future volatility. The model imposes a parameter which makes sure the impact of the negative returns is larger on future volatility than positive returns. In addition, since the standard GARCH model is unable to describe the asymmetric effects of negative and positive shock conditional volatility, other GARCH extensions since the early 1990s have attempted to account for these the asymmetric effects of shocks. The most notable extensions include the Glosten-Jagannathan-Runkle GARCH (GJR-GARCH), Threshold GARCH (TGARCH), Quadratic GARCH (QGARCH) and Family GARCH (FGARCH) models, suggested by Glosten et al. (1993), Zakoïan (1994), Sentana (1995) and Hentschel (1995) respectively.

The asymmetric effects of shocks are taken into account by the GJR-GARCH model by adding an extra term if the lagged innovation is smaller than zero. This term is the lagged squared residual multiplied by a dummy variable with a value of 1. Glosten et al. (1993) find that the negative monthly excess returns have a larger impact on the volatility of the CRSP value-weighted stock index portfolio than the positive monthly excess returns, as the asymmetric parameters are statistically significant. The structure of the Threshold GARCH

(TGARCH) model is slightly different to the other GARCH models as it has the conditional standard deviation equation instead of conditional variance equation in the model. The TGARCH model has the equation of the current conditional standard deviation on (i) the lagged conditional standard deviation, (ii) the lagged positive innovation and (iii) the lagged negative innovation. Zakoïan (1994) finds that the negative lagged residuals have a significantly larger impact on the current conditional volatility than the positive lagged residual for daily French ACA stock index data from January 1976 to July 1990. The Quadratic GARCH (QGARCH) model incorporates the lagged innovation in the conditional variance equation to capture the asymmetric effects of the shocks. The QGARCH model appeared to perform efficiently when used to capture the conditional variance and risk premia for daily US and monthly UK stock returns in Sentana (1995). The family GARCH model introduced by Hentschel (1995) nests a variety of asymmetric GARCH models including EGARCH, NGARCH, TGARCH, GJR-GARCH and absolute value of GARCH(AVGARCH), and a variety of symmetric GARCH models including Bollerslev's (1986) GARCH model and GARCH-M family models. He also finds evidence that negative shocks have more of an impact on conditional variance than positive shocks when the models are applied to daily US stock returns from 1926 to 1990 in the omnibus model he proposes.

1.4.4 High-frequency data used in GARCH models

High-frequency data were first estimated and forecasted in GARCH models in Andersen and Bollerslev (1997). They estimate the 5-minute returns for DM/$ and S&P 500 equities, yet distortions are found from the model when they apply it to high-frequency data. They also find that the distortions from the GARCH model can be eliminated if the intraday returns for the DM/$ exchange rate and S&P 500 equities are filtered or standardised by the intraday periodicity estimated from the Flexible Fourier form. The 5-minute returns for the DM/$ and ¥/$ exchange rates and the hourly returns for Australian indices have been tested using GARCH models by Andersen and Bollerslev (1998) and Walsh and Tsou (1998). Andersen and Bollerslev (1998) find that the GARCH (1,1) model better fits the higher-frequency data as the R^2 for the 5-minute intraday data is higher than that for the daily data. However, Walsh and Tsou (1998) find that the estimating hourly conditional variance for Australian indices with a large number of stocks using GARCH (1,1) is challenging because the large diversity of stocks in the index causes more apparent non-synchronous trading problems in higher-frequency data.

Many scholars have used GARCH-family models to fit high-frequency intraday returns for different types of financial data since the late 1990s. Andersen and Bollerslev (1999) use GARCH (1,1) to test 5-minute DM/$ Reuters quote data and find that the forecasted conditional variance improves for longer horizons when using high-frequency data. Andersen et al. (2003) use the GARCH-family models GARCH and the FIEGARCH model (Bollerslev & Mikkelsen, 1996) to estimate 30-minute tick data for ¥/$ and DM/$ Reuters FXFX quotes. They find that the long-memory Gaussian vector autoregression for the realised logarithmic volatilities (VAR-RV) model is better at forecasting than the GARCH model. However, they argue that the improvements gained by using VAR-RV do not render the GARCH model obsolete because it is the use of high-frequency data and its volatility measure RV in the VAR-RV model that contributes the most to improvements in forecasting rather than the model itself.

Another important extension of GARCH models takes into account the intraday patterns in the data. Andersen and Bollerslev (1997) find that the GARCH (1,1) model performs well at capturing the intraday volatility dynamics of exchange rates when their returns are filtered by intraday periodicity. They also suggest that intraday periodicity-adjusted returns give a cleaner picture of asset return volatilities.

Engle and Sokalska (2012) propose the Multiplicative Component GARCH (MC-GARCH) model, which considers intraday periodicity estimated via a non-parametric method. Intraday periodicity in the MC-GARCH model is estimated by using the average value of the squared intraday returns standardised by daily volatility. The authors find that this new model performs better at forecasting for less liquid stocks. Additional extensions of GARCH models which combine volatility measures from high-frequency data will be reviewed at the end of the next section.

1.4.5 HAR model and its extensions

The increasing availability of high-frequency data has resulted in the development of volatility measures such as realised variance (RV), which describes the unconditional volatility for asset returns. Estimation and forecasting models have been developed that use several non-parametric volatility measures. The non-parametric volatility measure RV is added to different parametric volatility estimation and forecasting models such as GARCH, SV and vector autoregression (VAR) models in order to improve their performance at volatility estimation and forecasting. Andersen et al. (2003) include RV in the VAR models by using 30-minute tick data for ¥/$ and DM/$ Reuters FXFX quotes, and they find that

VAR-RV model outperform the traditional GARCH model. Koopman et al. (2005) introduces the GARCH-RV, SV-RV and ARFIMA-RV models, which add the RV volatility measure to the GARCH, SV and autoregressive fractionally integrated moving average (ARFIMA) models. They test these models using 5-minute returns from the S&P 500 index and they find that not only the GARCH-RV model, but also the SV-RV and ARFIMA-RV models are superior than traditional volatility models. They conclude that RV, which is estimated from high-frequency intraday data, plays an important role in volatility forecasting, which supports the conclusions from Andersen et al. (2003).

Based on RV's effectiveness at capturing asset volatility, more research has been done to develop volatility estimation and forecasting models using non-parametric volatility measures including RV. The first volatility model based on RV is the Heterogeneous Auto-Regressive (HAR) model (Corsi. 2009), which uses the lagged RV, lagged weekly RV and lagged monthly RV to describe current volatility. They use the HAR model to test the tick-by-tick series for USD/CHF, S&P 500 Futures and 30-year US Treasury Bond Futures and find that this model is more accurate at forecasting than AR models and yields similar results to the ARFIMA model. However, they suggest that the advantage of the HAR model is that it is much easier to estimate than the ARFIMA model. Ma et al. (2014) compare the performance of different volatility RV models using Model Confidence Set (MCS) tests and find that the HAR-RV model outperforms the ARFIMA-RV model as well as its variants based on realised bi-power variation (ARFIMA-RBV) and multifractal volatility (ARFIMA-MFV) in all loss functions.

There are many extensions of HAR-RV models which aim to improve its forecasting ability for different financial data. Andersen et al. (2007a) find that the HAR model can be further improved by considering the jump components and continuous components separately based on Barndorff-Nielsen and Shephard's (2006) BV volatility measure. The new extensions HAR-RV-J and HAR-RV-CJ models are better at forecasting than the HAR-RV model from Corsi (2009) using tick-by-tick data for the DM/$ exchange rate, the S&P 500 market index, and the 30-year US Treasury bond yield (Andersen et al., 2007a). Andersen et al. (2011) and Corsi et al. (2010) propose extensions to the HAR family by considering the overnight variance and jump components respectively using more accurate non-parametric estimators. Andersen et al. (2011) find that over 16% of the variation in the S&P 500 and US daily variation data are caused by changes in stock prices from the closing prices of the previous day to the opening prices of the current day. Therefore, they suggest a HAR-RV-CNJ model which includes the overnight return variability in the HAR model. By using the new model to test the five-minute S&P 500 futures (SP) and 30-year US treasury bond futures, they find that the HAR-CNJ models outperforms the HAR-RV and GARCH-class models at both in-sample and out-of-sample forecasting. Wang and Xu (2015) further extend the concept not only to account for the overnight returns, but also to include the lunch-break returns, trading volumes and the leverage effects in the HAR model to fit the 5-minute data for the Shanghai Stock Exchange Composite Index (SHCI) and the Shenzhen Composite Index (SZCI). Both the DM test and the R^2 values from the Mincer-Zarnowitz regression are better for the new model than for HAR-RV, which is attributed to the significant impact of negative lunch-break returns and negative overnight returns.

Corsi et al. (2010) find evidence that BV underestimates jump components, so they introduce threshold bi-power variation (TBV) to reduce the bias. In order to avoid the possibility that TBV overestimates the jump components, they also put forward a corrected

TBV (CTBV) estimator to improve the accuracy of jump components further. Based on the new advanced jump component estimators, Corsi et al. (2010) improve the HAR-RV-J and HAR-RV-CJ models, resulting in the HAR-TJ and HAR-CTJ models. By using the new models and a standard HAR-RV model to test 5-minute 30-year US Treasury Bond futures data, they find both new models are better at forecasting than the standard HAR-RV model. Duong and Swanson (2015) also improve HAR models by considering asymmetric information provided by the jump components (HAR-RV-C-APJ) or truncated large jumps (HAR-RV-C-UDJ). They use 5-minute S&P 500 futures data to test these models and find that the HAR-RV-C-APJ model shows more obvious improvements than the HAR-RV-C-UDJ model, with respective 8% and 7.5% increases in R^2 compared to the HAR-RV-C model at forecasting horizons of 1 and 5.

Pu et al. (2016) argue that past jumps caused by negative price changes may have a different impact on current volatility from past jumps caused by positive price changes. They hence attempt to account for the realised semi-variance (RS) estimators introduced by Barndorff-Nielsen et al. (2010) as an additional explanatory variable in a HAR model on 1-minute high-frequency data from the Shanghai Stock Exchange Composite (SSEC) Index. The RS estimator calculates the jump variation by subtracting the variation from negative price changes (RS^-) from the variation from positive price changes (RS^+). From the forecasting results and mode confidence set (MCS) tests, they find that the models with the RS estimators are most accurate at forecasting. This is especially so for the HAR-RV-TJ-SJV-D, which adds the polarity of the jump variations to the HAR-RV-TJ model.

Market transactions caused by different traders are heterogeneous and this is the main reason for market volatility (Müller et al. 1993). Dong and Feng (2018) use daily, weekly and monthly RV to capture these heterogeneous trading characteristics, which are the short-term, medium-term and long-term market transactions respectively. By applying the new model to 1-minute data from the CSI 300 index, the authors find that the expected increase on long-term speculative behaviour has a negative impact on the market and triggers market volatility. Gong and Lin (2018) following Bandi and Russell (2008), suggest that RV shows bias when estimating volatility because of microstructure noise. Therefore, they use realised range-based variance (RRV; Christensen & Podolskij, 2007; Martens & van Dijk, 2007) to replace RV in the HAR model, the former of which is claimed to be five times better than RV (Christensen & Podolskij, 2007). Gong and Lin (2018) dub their RRV-based HAR models HAR-RRV and HAR-RRV-SC, which respectively exclude and include structural change. The forecasting and DM test results based on 5-minute S&P 500 index data show that the model that takes structural changes into account performs better than one that does not. This highlights the importance of considering structural changes when forecasting volatility. However, there is no direct evidence from the study for the superiority of HAR models that use RRV rather than RV in forecasting.

Peng et al. (2018) test the impact of 5-minute index data from G7 countries on volatility forecasting via RV estimated from the Shanghai Stock Exchange Composite Index (SSEC). Their method takes into account the RV estimated from the S&P 500, FTSE 100, Nikkei 225, DAX, CAC 40, FTSEMIB and S&P/TSX composite indices as news variables in a HAR-RV model, both individually and together. From the in-sample and out-of-sample forecasting results, they find that the indices from the Japanese and US markets have a positive impact on future volatility for the Chinese market. They also find that the model which incorporates all of the G7 stock markets provides more accurate forecasting for one-day volatilities of the Chinese stock market than the standard HAR-RV model. Therefore, the information from G7 stock markets has an impact on Chinese stock market forecasting.

Unlike the traditional way of separating the continuous and jump component using continuous component estimators such as BV or TBV, Gong and Li (2018) attempt to improve the HAR-RV-CJ model by applying the ensemble EMD method to separate the RV into several intrinsic mode function components and one non-oscillatory trend component. They also add the impact of high-frequency, low-frequency and trend volatility as well as leverage effects in the models, which result in the HAR-RV-HLT and LHAR-RV-HLT models. They conclude that the 5-minute S&P 500 data favour these models compared to other HAR models including the traditional HAR-RV, HAR-RV-J and HAR-RV-CJ models regarding one-month future volatility forecasting.

Bollerslev et al.'s (2016) full HARQ (HARQ-F) model includes the lagged realised quarticity (RQ) and lagged weekly and monthly RQ in the HAR model in order to correct the heteroscedastic measurement errors in the model. This is because take the RQs into account can lead to faster mean reversion for the model when measurement errors are large. The authors claim that the weekly and monthly lags for RQ in the HARQ-F model do not play a large role in correcting bias compared to the lagged RQ; therefore, they propose the HARQ model, which only considers lagged RQ. The authors also propose the CHARQ model, HARQ-J model and SHARQ models, which account for the following elements: the continuous components only; the continuous and jump components; and the RV estimated from the negative and positive returns separately. They apply these extended HAR models to tick-by-tick S&P 500 data and find that the data favour the HARQ model the most among all the new models. They also find that the HARQ model is better at estimating risk premia and volatility forecasting than the HAR and HAR-J models.

Bekierman and Manner (2018) suggest that the realised quarticity (RQ) is a noisy estimator of Integrated Quarticity (IQ), which may cause bias in the HARQ model. In order to avoid this bias, they introduce a state-space HAR model (HARS) which includes a state equation in the model. This makes the autoregressive parameter a time-varying parameter driven by a latent Gaussian process. Also, the RQ is included in the state equation for the HARSQ model, which combines the state equation with the HARQ model. They also propose HARL and HARSL models, which replace the RV from the HAR and HARS models with the logarithm of RV. They compare the forecasting results of these two models based on 40 stocks' 5-minute returns from the S&P 500 index with those from the HAR, HARQ, HARS and HARSQ models. They find evidence that the HARL and HARSL models provide more accurate forecasting than the other models.

Some work has also attempted to combine GARCH and HAR models. Corsi et al. (2008) extend the HAR model by assuming that the error term follows the GARCH process. They find that their 5-minute S&P 500 index data favour the HAR-GARCH model, especially the model with a Normal-Inverse Gaussian (NIG)-distributed innovation on in-sample estimation. They argue that it is important for the model to incorporate the GARCH specification as it is able to accommodate fat-tailed and/or skewed distributions. Models combining GARCH and HAR are also used in Vale-at-Risk forecasting. Będowska-Sójka (2015) forecasts the Value-at-Risk (VaR) for 5-minute EUR/PLN exchange rate data using hybrid models, which combine different methodologies including HAR-class and GARCH-class models. The models she uses include HAR-RV, HAR-RV-J, GARCH, EGARCH and

FIGARCH. She finds that the combination of HAR- and GACRH-class models is better at forecasting VaR when the accuracy test results are not considered, because they offer firm loss functions. However, by comparing the accuracy test results, the various GARCH-class models are generally better at VaR forecasting.

It is clear from previous literature that there are numerous ways of extending HAR-family models by considering different factors, such as improving the methods of estimating the jump and continuous components, and combined the model with parametric volatility methods. However, despite the wealth of scholarly work described in this section, few studies have considered extending the HAR model by incorporating macroeconomic news announcements. We therefore help fill this gap in Chapter 4 by extending non-parametric volatility models to include information given by news announcements and examining how they affect the forecasting performance of HAR-family models. In the next section, we discuss the importance of the impact of news announcements and why they should be incorporated into HAR models.

1.5 The impact of news announcements on asset return volatility

1.5.1 Market responses towards news announcements

The announcement of macroeconomic news often results in immediate changes in financial asset prices and volatility (Andersen & Bollerslev, 1998; Balduzzi et al., 2001; Huang, 2018). Changes in asset prices are defined as first-moment market responses, and changes in their volatility are known as second-moment market responses. The changing patterns of financial asset prices and volatility in response to macroeconomic news announcements is important, especially in estimation and forecasting (Andersen & Bollerslev, 1998; Balduzzi et al., 2001; Andersen et al., 2007a). In this section, we review previous studies of first- and second-moment market responses of financial assets to macroeconomic news announcements and how taking these announcements into account can prove advantageous for volatility estimation and forecasting.

1.5.1.1 First-moment market responses

Evidence from early studies of the impact of news announcements on market responses is limited as they typically analyse monthly or weekly data. In the late 1980s and early 1990s, information regarding market responses to news became clearer as intraday data became available. Early work mainly focuses on changes in asset prices (first-moment market responses). For example, Jain (1988) finds that stock price adjustments in response to news announcements can take place within an hour. Ederington and Lee (1993) find that the market's first-moment responses for Treasury bond (T-bond), Euro-dollar, and Deutsche Mark futures to the announcement of news can be completed within a minute, while bigger changes in volatility may last longer (up to 15-30 minutes). Recent work on first-moment responses by Bollerslev et al. (2018) using high-frequency S&P 500 and US Treasury bonds data finds evidence that large price jumps are influenced by news announcements from the US Federal Open Market Committee (FOMC).

1.5.1.2 Second-moment market responses

Early research on second-moment market responses to macroeconomic news announcements mainly focuses on the observable patterns of asset volatility caused by the release of news. One such study is Ederington and Lee (1993), who find that treasury bond futures have a high volatility between 8.30 am and 8.35 am after the release of monthly economic news announcements. Bollerslev et al. (2000) also find two spikes in volatility at 8.30 am and 10.00 am on treasury bond futures provoked by regular scheduled news announcements. Meanwhile, Balduzzi et al. (2001) examine the differences between first-moment and second-moment market responses to news. They find that there are three phases of market responses to news announcements on bond markets. Price adjustment typically happens immediately after a news announcement and before any changes in volume and volatility because they are driven by public information. Volatility and volume then increase in the second phase for up to 15 minutes as they are partly driven by informed trading. In the third phase, liquidity trading tends to drive volume and volatility back to normal.

In addition, studies find evidence that the announcement of macroeconomic news has an impact on overseas markets. Wongswan (2006) finds that the announcement of American and Japanese news can result in a 30-minute change of volatility in Korean and Thai equity markets. However, Kleinnijenhuis et al. (2013) find that the announcement of negative financial news in the Netherlands not only causes short-lived changes in market responses because of the actions taken by Dutch investors, but it can also cause panic in global players. This can lead these traders to conduct massive sales, which can affect the market for a whole month.

1.5.1.3 Regression analysis of news on market responses

Given the importance of macroeconomic news announcements on second-moment market responses, statistical methods have been suggested to find out the quantitative relationship between the two variables. Andersen and Bollerslev (1998) find that regularly scheduled macroeconomic news announcements can produce daily patterns in second-moment responses, suggesting that news can affect returns' intraday patterns. They therefore include the impact of news announcements as a dummy variable in the intraday periodicity estimation, which contributes to the volatility of the returns. By applying the model to the 5-minute Deutsche Mark/US Dollar spot exchange rate, they find clear evidence that news announcements can affect exchange rate volatility.

Instead of using a dummy variable to model the arrival of macroeconomic news announcements, Balduzzi et al. (2001) introduce a very useful measure, the *z*-type standardised measure, in order to transform the released value of news into a variable called standardised news surprise (see Section 4.3.1 for further methodological details). They use a linear regression to test the impact of standardised news surprise on returns 30 minutes after the announcement of news and find that at least ten news announcements significantly affect the prices of T-bill, two-year notes, ten-year notes and 30-year bonds. This method allows researchers to test the impact of news on volatility components more rigorously than with a dummy variable or via impressionistic observation of graphs.

Balduzzi et al.'s (2001) method of calculating news surprise is used in many later studies as well. For example, Wongswan (2006) tests the impact of macroeconomic news from developed countries such as the USA and Japan on the equity markets of emerging economies such as South Korea and Thailand. He does so by running regressions on the estimated pattern effects of news announcements from the USA and Japan on intraday volatility and intraday trading volumes for the Korean and Thai equity markets. He accounts for macroeconomic news effects using both the dummy variable method from Andersen and Bollerslev (1998) as well as standardised news surprise (Balduzzi et al., 2001), together with the dispersion of expectations for news announcements to describe the pattern effects. Wongswan (2006) finds that US and Japanese news items significantly affect the volatility of Korean and Thai equity markets 30 minutes after they are announced, in contrast to domestic news, which does not appear to have a significant effect. He suggests that this may be because the domestic information was leaked before the official announcements.

In a similar vein, Andersen et al. (2007b) run a regression of returns from nine futures markets from the USA, UK and Germany on the lags of returns and lags of different news surprises using the method from Balduzzi et al. (2001). They find clear evidence that the news contributes to large changes in returns for those futures. Also, Huang (2018) tests the impact of US news announcements on American equity and bond markets using 5-minute returns for S&P 500 futures and 30-year US Treasury bond futures. He uses the standardised measure from Balduzzi et al. (2001) and calculates standardised news surprises based on the released values of news and disagreement about the news, estimated by taking the standard deviation of the survey forecasts. He regresses the transformed jump components and continuous components on standardised news surprise and disagreement, which reflects how agents feel about their forecasts of upcoming news. The regression results indicate that both news surprises and the disagreement about the news affects second-moment market responses.

Instead of using the standard news surprises (Balduzzi et al., 2001), later studies introduce other methods to test the impact of macroeconomic news announcements on financial market volatility (second-moment market responses). Lee and Mykland (2007) test the impact of market-level news and company-level news on three stocks and the S&P 500 Index by observing the relationship between news arrivals and the frequency and size of jumps. Their results show that the jumps for the Walmart (WMT), IBM (IBM) and General Electric (GE) stocks are more influenced by company-related scheduled and unscheduled news such as reports of earnings, while the S&P 500 as a whole is affected more by market-level news such as Federal Open Market Committee (FOMC) reports. Lee (2011) investigates how market-level and company news affects the prediction of jumps using her own jump prediction test. Her analysis of 23 stocks from the US market shows that macroeconomic news announcements such as initial jobless claims significantly affects the occurrence of jumps, especially in the short horizon of 30 minutes.

It is clear from the literature discussed in this section that macroeconomic news announcements have an impact on financial markets, especially in the second-moment market responses, as shown in a range of financial assets. Some news, such as that from the FOMC, is particularly influential on market volatility. In this analysis in Chapter 4, we consider the more rigorous news surprise value (Balduzzi et al., 2001) for the release of news rather than incorporating them as dummy variables to examine the impact of news on financial assets. In addition, our analysis takes into account a wide range of news outlets, some of which are not considered in previous work. We also contribute to the literature by incorporating the information from the significant news announcements in a non-parametric volatility forecasting model (HAR model).

1.5.2 News announcements and co-jumps

News announcements may produce similar changes across a range of financial asset returns at the same time, which are known as co-jumps. Most previous studies of co-jumps are done using logit or probit models. Dungey et al. (2009) examine the impact of macroeconomic news announcements on US bond markets using a panel logit model and find that co-jumps of bonds of different maturities are strongly affected by news about US interest-rate term structure. Lahaye et al. (2011) use a probit model to test the relationship between co-jumps and news announcements for USD exchange rates, US Treasury bond futures, and US equity futures. From the regression results, they find that news announcements significantly affect the co-jumps between different financial returns. By calculating the conditional probably of news on co-jumps, they find that macroeconomic news announcements more strongly influence co-jumps for the equity and bond markets than those for the exchange market. The impact of news announcements on co-jumps for the US Treasury market are examined using a panel logit model by Dungey and Hvozdyk (2012). Their model regresses on a joint jump day or a conflicting day with news announcements as dummy variables, the estimation results indicating a significant positive effect of news announcements on the probability of co-jumps. The authors also find that announcements of non-farm payrolls, consumer price index (CPI), gross domestic product (GDP) and retail sales have more of an impact than other news on the probability of co-jumps. Chatrath et al. (2014) run a probit regression on the negative and positive co-jumps for news surprises for four currencies – pound sterling, euro, Japanese yen and Swiss franc. They find that macroeconomic news has a significant impact on co-jumps. They also find that positive and negative news surprises increase the probability of negative and positive co-jumps respectively.

Instead of investigating the impact of news on the co-movements between financial asset returns, some literature (e.g. Maio et al., 2014; Gilder et al., 2014) examines the impact of news announcements on systematic co-jumps for portfolios and provide evidence for a significant effect of macroeconomic news on systematic co-jumps. One of the most recent is Chan et al. (2017), who examine the impact of macroeconomic news announcements on co-jumps of book-to-market (B/M) portfolios, based on their B/M price ratio and market capitalisation. They do this using probit and tobit regression of the probability and magnitude of systematic co-jumps on the standardised announcement surprises. They find a significant effect for various types of news, including announcements regarding the Federal Funds target rate, nonfarm payroll statistics, the unemployment rate, the producer price index and the Institute for Supply Management index, on systematic co-jumps.

The impact of news surprises on asset returns are discussed in previous literature, and it is clear that news announcements play an important role on the volatility of asset returns. However, not many volatility models from previous work consider the impact of news announcements. Andersen and Bollerslev (1997) is one of the few studies that includes the impact of news announcements in a volatility model, but the authors only treat news as a dummy variable of intraday periodicity to explain volatility with a GARCH model. To our knowledge, there are also no studies that consider news announcements to be an important factor for volatility forecasting using non-parametric HAR-class models. Therefore, this thesis helps fill the gap in the literature by accounting for news announcements in non-parametric HAR-class volatility models and examining how news may contribute to volatility forecasting using such models.

1.6 This thesis's original contribution

It is clear from the literature discussed in this chapter that high-frequency stock market data have been studied extensively, yet the stylised facts of such data, such as their leverage effects and tail properties, are less well understood. This thesis directly investigates the stylised facts of high-frequency data using intraday asset returns, while also examining how these facts vary when intraday returns are standardised using different volatility measures.

In addition, much of the literature discussed in this chapter has highlighted the importance of intraday periodicity patterns in intraday returns (e.g. Andersen & Bollerslev, 1997, 1998). Efficient non-parametric intraday periodicity estimation methods have been developed, such as standard deviation (SD), weighted standard deviation (WSD; Boudt et al., 2011b) and Shortest Half (Short-H; Rousseeuw & Leroy, 1988). However, few studies have compared the performance of different estimators on data using different sampling schemes. We therefore compare the performance of the SD, WSD and Shortest Half intraday periodicity estimators on high-frequency stock returns using business-time sampling and calendar-time sampling in different financial regimes, which allows us to assess their impact on volatility forecasting.

Macroeconomic news announcements are another factor which evidence suggests significantly contributes to the volatility of asset returns. However, it has not been considered in volatility models using rigorous methods such as standardised news surprise (Balduzzi et al., 2001), especially as part of non-parametric HAR-class models. Therefore, we extend HAR-class models by considering the impact of macroeconomic news on jump components and how such models can be improved by incorporating news announcements using news surprise. This will help deepen our understanding of the factors influencing market volatility and how best to forecast volatility patterns using high-frequency data.

Chapter 2 – The stylised facts of high-frequency returns and volatility measures

2.1 Introduction

Modelling intraday volatility has been a popular topic over the last two decades, as high-frequency time-series data are widely used in financial analysis. Non-parametric Heterogeneous Auto-Regressive (HAR) models (Corsi, 2009) are typically used to forecast realised variation (RV). Non-parametric measures, such as bi-power variation (BV) and threshold bi-power variation (TBV), are included in HAR-family models to generate continuous and jump components in order to describe and forecast the realised variance of financial returns.

Understanding the statistical properties of stocks is essential for stock volatility estimation and forecasting. The stylised facts of financial assets can be defined as the consistent statistical findings for financial time series within a particular market or time period. Stylised facts are helpful for us to understand the quantitative properties of financial time series, such as their distribution properties and tail properties, which can inform the selection of appropriate methods of asset analysis and potential improvements in statistical estimations.

In this chapter, we discuss stylised facts such as linear dependence, leverage effects and tail properties (for aggregated returns) for stocks and stock market indices across different sampling frequencies in different financial conditions (pre-crisis, crisis or post-crisis periods). We also study the linear dependence and long-memory properties of many different volatility measures, which have not been compared in previous literature. In addition, we examine how volatility measures correlate with trading volume and how they standardise returns across different financial regimes. We then compare the performances of different HAR-class models that consider trading volume across different regimes.

The remainder of the chapter is structured as follows. In Section 2.2, we introduce the data used in the analysis. This is complemented in Section 2.3 by an overview of the methods for different measures for estimating volatility, together with a description of HAR volatility models. Section 2.4 presents the empirical results of several stylised facts of stocks and the SPY index, including autocorrelation, tail properties and standardised returns. In Section 2.5, the volatility measures' long-memory properties and correlations with trading volume, as well as the leverage effects of the returns and standardised returns, are discussed. Section 2.6 provides an analysis the volatility forecasting. Section 2.7 compares the above properties across different regimes. Section 2.8 concludes the main findings of this chapter.

2.2 Data

The data used in this chapter are high-frequency stock returns from the NASDAQ index from 2000 to 2016. There are four stocks from two different industrial sectors: PFE and JNJ from the healthcare (HC) sector; and AAPL and MSFT from the information technology (IT) sector. We also include the SPDR S&P 500 exchange-traded fund (SPY) in this chapter, as it helps provide a picture for the stock market as a whole. We include four sampling frequencies in our analysis, which includes 30-second, 60-second, 150-

second and 300-second data. We also examine the stylised facts of volatility measures in different regimes by separating the 2000-2016 data into three parts: pre-crisis (01/01/2000 to 30/12/2007), crisis (01/01/2008 to 30/06/2009) and post-crisis (01/07/2009 to 30/12/2016) periods.

[Insert Figures 2.1 to 2.5 here]

Figures 2.1 to 2.5 show that the stocks and the SPY index show the most significant changes in prices from 2008 to 2009, which also causes the biggest change in volatility across all frequencies for both calendar-time and business-time sampling data. This result is caused by the global financial crisis of 2008 to 2009. There are also dramatic changes in prices from 2000 to 2002, which may be part of the bear market that began in 2000. This bear market affected the IT companies AAPL and MSFT to a greater degree than the healthcare stocks because of the burst of the tech bubble in 2000. In addition, we can see that these big shocks in the stock market have a relatively small impact on the volatility of SPY compared to individual stocks. For example, the burst of the tech bubble caused the RV of each stock to exceed 10 for 30-second returns at the end of 2002, but the RV of SPY only reached 5 at the same time.

Figures 2.1 to 2.5 also show that the volatility of stocks and SPY has the highest peak during the financial crisis period (2008 to 2009), with a dramatic increase in trading volume at the same time for most stocks and SPY (with the exception of MSFT) as investors' fears over the financial crisis led to increased trading. However, during the small peak in volatility from 2000 to 2002 caused by the burst of the tech bubble, the amount of trading volume does not rise significantly for stocks or SPY. This shows that the volatility of stocks and SPY are not only affected by trading volume, but also affected by the types of trade orders. The trading volumes from 2000 to 2002 are not particularly high, yet the large number of sell orders with relatively few buy orders resulted in high volatility after the burst of the tech bubble.

[Insert Table 2.1 here]

Table 2.1 shows the descriptive statistics for returns and conditional variances for 300-second stocks and the SPY index. The conditional variances are estimated from the GARCH (1,1) model with normally distributed error terms. The standard deviations for returns and the mean conditional variance are higher in the crisis period, followed by the pre-crisis period, and the average returns for stocks and SPY are much more negative during the crisis. The findings show the impact of the financial crisis and the burst of the tech bubble, which are in line with the results in Figures 2.1 and 2.5. In addition, the stocks from the IT sector are more volatile in all periods than stocks from the HC sector. This is because the companies in the IT sector grew dramatically due to the development of cloud computing, mobile computing and big data and they have capacity to alter their operations and innovations frequently. News announcements related to these companies, especially their quarterly earnings reports, are watched closely by investors and often result in fluctuations in investor sentiment. Companies from the HC sector are less volatile as they are less sensitive to economic cycles and are typically regarded as defensive stocks when market is going down.

2.3 Methodology

2.3.1 Volatility estimation measures

In this chapter, we use different volatility measures to estimate quadratic variation (QV), which describes the variation in financial time series. Andersen and Bollerslev (1998) use realised variance (RV) to estimate QV, as shown in equation (2.1). The RV of a given trading day can be calculated using equation 2.1.

$$\mathrm{RV} = \sum_{\mathrm{i}=1}^{\mathrm{M}} (\Delta_{\mathrm{i}}^{\mathrm{n}} \mathrm{X})^2 \underset{M \to \infty}{\overset{plim}{\to}} \mathrm{QV}_{\mathrm{t}}, \quad \Delta_{\mathrm{n}}(\equiv 1/\mathrm{M}) \to 0 \tag{2.1}$$

Where $\Delta_{\mathrm{i}}^{\mathrm{n}}\mathrm{X} = \mathrm{X}(\mathrm{t+i}\Delta) - \mathrm{X}(\mathrm{t+(i-1)}\Delta)$, which are the equally spaced intraday returns from a financial time series. $(\Delta_{\mathrm{i}}^{\mathrm{n}}\mathrm{X})^2$ is the *i*th squared returns of the trading day at stage *n*. RV_{t} converges in probability to the QV_{t}. M is the number of sampled observations per trading day.

Hansen and Lunde (2004) use bias-corrected realised variance (RV_{AC}), which can be calculated using equation (2.2).

$$RV_{AC} = \sum_{\mathrm{i}=2}^{\mathrm{M}} (\Delta_{\mathrm{i}}^{\mathrm{n}}\mathrm{X})^2 + 2\sum_{\mathrm{i}=2}^{\mathrm{qM}} \frac{M}{M-h} \sum_{\mathrm{i}=2}^{\mathrm{M-h}} (\Delta_{\mathrm{i}}^{\mathrm{n}}\mathrm{X})(\Delta_{\mathrm{i+h}}^{\mathrm{n}}\mathrm{X}) \tag{2.2}$$

$\Delta_{\mathrm{i+h}}^{\mathrm{n}}\mathrm{X}$ refers to the $(i+h)$th return of a given trading day. This approach eliminates microeconomic noise in realised variation estimation by correcting for the first *qM* autocorrelations. Huang and Tauchen (2005) argue that the presence of microeconomic noise seriously contaminates jump detections, so eliminating the noise may improve jump test accuracy.

Barndoff-Nielsen and Shephard (2004) use bi-power variation (BV), which is robust to jumps and can be written as:

$$BV = \mu_1^{-2} \frac{M}{M-1} \sum_{i=2}^{M} |\Delta_i^n X|^r \, |\Delta_{i-1}^n X|^s \underset{M \to \infty}{\overset{plim}{\rightarrow}} \int_0^t \sigma_s^2 \, ds, \quad r=s=1 \tag{2.3}$$

Where $\mu_1 = E(\mu) = \sqrt{2}/\Gamma(\frac{1}{2}) \approx 0.7978$, and *r* and *s* are the powers of the absolute return and its first lag respectively. $\int_0^t \sigma_s^2 \, ds$ corresponds to the integrated variance (IV). They also employ tri-power variation (TPV) and quad-power variation (QPV) by adding one or two extra adjacent returns in the estimation respectively, in order to make the estimators more robust to microeconomic noise. These are shown in equations (2.4) and (2.5) respectively.

$$TPV = \mu_{3/2}^{-3} \frac{M}{M-2} \sum_{i=3}^{M} |\Delta_i^n X|^r \, |\Delta_{i-1}^n X|^s |\Delta_{i-2}^n X|^q \underset{M \to \infty}{\overset{plim}{\rightarrow}} \int_0^t \sigma_s^2 \, ds, \quad r=q=s=2/3. \tag{2.4}$$

$$QPV = \mu_{1/2}^{-4} \frac{M}{M-3} \sum_{i=4}^{M} |\Delta_i^n X|^r \, |\Delta_{i-1}^n X|^s |\Delta_{i-2}^n X|^q |\Delta_{i-3}^n X|^u \underset{M \to \infty}{\overset{plim}{\rightarrow}} \int_0^t \sigma_s^2 \, ds, \quad r=t=q=u=1/2 \tag{2.5}$$

Where $\mu_1 = E(\mu) = \sqrt{2}/\Gamma(\frac{1}{2}) \approx 0.7978$ and *q* and *u* are the powers of the absolute values of the 2nd and 3rd lags of a given return. TPV keeps the rule that the sum of the exponents must equal 2. The QPV estimator is more robust to microeconomic noise but less efficient than TPV.

Another estimator is skipped bi-power variation (SBV; Huang & Tauchen, 2005). SBV is a BV estimator that is robust to microeconomic noise by adding a more distant adjacent return term in the estimation. SBV also adopts the rule that the sum of the exponents must equal 2.

$$\mathrm{SBV}=\mu_1^{-2}\frac{M}{M-2}\sum_{i=3}^{M}|\Delta_i^n X|^r\,|\Delta_{i-1}^n X|^s|\Delta_{i-2}^n X|^q \underset{M\to\infty}{\overset{plim}{\rightarrow}} \int_0^t \sigma_s^2\,ds,\ \ r=q=1,\ s=0 \tag{2.6}$$

Threshold bi-power variation (TBV; Corsi et al., 2010), adds a threshold parameter to make the estimator more robust to large jumps. TBV can be written as:

$$\mathrm{TBV}=\mu_1^{-2}\frac{M}{M-1}\sum_{i=2}^{M}|\Delta_i^n X|^r\,|\Delta_{i-1}^n X|^s 1_{\{|\Delta_{i-1}^n X|^2\le\vartheta_i\}}1_{\{|\Delta_{i-1}^n X|^2\le\vartheta_{i-1}\}} \underset{M\to\infty}{\overset{plim}{\rightarrow}} \int_0^t \sigma_s^2\,ds \tag{2.7}$$

r=t=1, s=0. The threshold parameter $\vartheta=c_\vartheta^2\cdot\widehat{V}_t$ is estimated with c_∂=3 and $\widehat{V}$ is an auxiliary estimator of σ^2.

Andersen et al. (2012) introduce the minRV and medRV estimators, which take the minimum and the median over the adjacent returns respectively in order to eliminate the jumps in the estimators.

$$\mathrm{minRV}=\frac{\pi}{\pi-2}\frac{M}{M-1}\sum_{i=2}^{M}\min\left(|\Delta_i^n X|,|\Delta_{i-1}^n X|\right)^2 \underset{M\to\infty}{\overset{plim}{\rightarrow}} \int_0^t \sigma_s^2\,ds \tag{2.8}$$

$$\mathrm{MedRV}=\frac{\pi}{6-4\sqrt{3}+\pi}\frac{M}{M-2}\sum_{i=3}^{M}\min\left(|\Delta_i^n X|,|\Delta_{i-1}^n X|,|\Delta_{i-2}^n X|\right)^2 \underset{M\to\infty}{\overset{plim}{\rightarrow}} \int_0^t \sigma_s^2\,ds \tag{2.9}$$

2.3.2 HAR-class models

Corsi et al. (2010) introduced the HAR-J and HAR-TJ models, which are shown in equations (2.10) and (2.11) respectively. The HAR-J model can be written as:

$$\mathrm{RV}_{t:t+h-1}=\beta_0+\beta_d\widehat{C}_{t-1}+\beta_w\widehat{C}_{t-5:t-1}+\beta_m\widehat{C}_{t-22:t-1}+\beta_j\widehat{J}_{t-1}+\varepsilon_t \tag{2.10}$$

$\mathrm{RV}_{t_1:t_2} = \frac{1}{t_2-t_1+1}\sum_{t=t_1}^{t_2} RV_t$, with $t_1 \leq t_2$. The error term is an independent and identically distributed (i.i.d.) random variable with mean 0 and variance σ^2. The jump and continuous components in equation (2.10) can be expressed as $\hat{\mathrm{J}}_{\mathrm{t}} = \mathrm{I}_{\{\mathrm{z_t} > \phi_\alpha\}} \cdot \max\left[(\mathrm{RV_t} - \mathrm{BV_t}), 0\right]$ and $\hat{\mathrm{C}}_{\mathrm{t}} = \mathrm{RV_t} - \hat{\mathrm{J}}_{\mathrm{t}}$ respectively. The HAR-TJ model can be expressed as:

$$\mathrm{RV}_{\mathrm{t:t+h-1}} = \beta_0 + \beta_{\mathrm{d}}\widehat{\mathrm{TC}}_{\mathrm{t-1}} + \beta_{\mathrm{w}}\widehat{\mathrm{TC}}_{\mathrm{t-5:t-1}} + \beta_{\mathrm{m}}\widehat{\mathrm{TC}}_{\mathrm{t-22:t-1}} + \beta_{\mathrm{j}}\widehat{\mathrm{TJ}}_{\mathrm{t-1}} + \varepsilon_{\mathrm{t}} \quad (2.11)$$

with the jump and continuous components $\widehat{\mathrm{TJ}}_{\mathrm{t}} = \mathrm{I}_{\{\mathrm{z_t} > \phi_\alpha\}} \cdot \max\left[(\mathrm{RV_t} - \mathrm{TBV_t}), 0\right]$ and $\widehat{\mathrm{TC}}_{\mathrm{t}} = \mathrm{RV_t} - \widehat{\mathrm{TJ}}_{\mathrm{t}}$ in equation (2.11). The error term ε_{t} is an i.i.d. random variable with mean 0 and variance σ^2.

In this thesis, we present new models, HAR-J-Vol and HAR-J-Vol, which consider trading volume. This is because volatility measures are significantly correlated with trading volume, as discussed in Sections 2.5.3 and 2.7.3. In addition, we use LASSO regression to examine the relationship between trading volume lags and realised variance. We find that the first lags of trading volume have a significant impact on RV for most stocks, and so they are considered in the HAR-J-Vol and HAR-TJ-Vol models, as shown in equations (2.12) and (2.13). Higher lags are not significant for most stocks and hence are not included.

The HAR-J-Vol and HAR-TJ-Vol models can be written as:

$$\mathrm{RV}_{\mathrm{t:t+h-1}} = \beta_0 + \beta_{\mathrm{d}}\hat{\mathrm{C}}_{\mathrm{t-1}} + \beta_{\mathrm{w}}\hat{\mathrm{C}}_{\mathrm{t-5:t-1}} + \beta_{\mathrm{m}}\hat{\mathrm{C}}_{\mathrm{t-22:t-1}} + \beta_{\mathrm{j}}\hat{\mathrm{J}}_{\mathrm{t-1}} + \beta_{\mathrm{v}}\mathrm{Vol}_{\mathrm{t-1}} + \varepsilon_{\mathrm{t}} \quad (2.12)$$

$$\mathrm{RV}_{\mathrm{t:t+h-1}} = \beta_0 + \beta_{\mathrm{d}}\widehat{\mathrm{TC}}_{\mathrm{t-1}} + \beta_{\mathrm{w}}\widehat{\mathrm{TC}}_{\mathrm{t-5:t-1}} + \beta_{\mathrm{m}}\widehat{\mathrm{TC}}_{\mathrm{t-22:t-1}} + \beta_{\mathrm{j}}\widehat{\mathrm{TJ}}_{\mathrm{t-1}} + \beta_{\mathrm{v}}\mathrm{Vol}_{\mathrm{t-1}} + \varepsilon_{\mathrm{t}} \quad (2.13)$$

Where the error terms are i.i.d. random variables with mean 0 and variance σ^2.

2.4 Stylised facts of intraday returns

In this section, we investigate the stylised facts of intraday returns from stock market assets using the SPY index for the whole data set from 2000 to 2016, as SPY reflects the average movements of the stocks in the market for the whole period.

2.4.1 Autocorrelations

Figure 2.6 shows that there are no autocorrelations for intraday returns across different frequencies. Figure 2.7 shows a slow decay in autocorrelations for SPY's absolute intraday returns, which suggests that they are likely to have long-term dependence.

[Insert Figures 2.6 to 2.8 here]

The autocorrelation plot for the volatility measures in Figure 2.8 shows that they have positive partial autocorrelations for the majority of the lags, suggesting that high-volatility events tend to follow one other in rapid succession. This can be easily observed from the changes in returns during the financial crisis period (2008-2009) in Figure 2.1, as many high-volatility events occurred during this period. In addition, the partial autocorrelation results in Figure 2.8 show that the RV and BV volatility measures are AR (9) processes because the partial autocorrelations cut off after the ninth lag. This indicates that stock volatility can have a long-term impact (up to nine days). Also, the CTBV partial autocorrelations have large correlations for lags 11, 15 and 20 in Figure 2.8. This suggests that long-range dependence also differs between different volatility measures calculated from the same set of intraday returns.

2.4.2 Unconditional and conditional heavy tails

Figure 2.9 shows quantile-quantile (Q-Q) plots for SPY returns.

[Insert Figure 2.9 here]

It is obvious that the returns have a heavy tail because they have a large number of outliers. The Q-Q plots are all S-shaped, with small values on the left side of the x-axis quantiles and large values on the right side compared to the theoretical quantile.

[Insert Figure 2.10 here]

The Q-Q plot for the residuals of the GARCH (1,1) model is shown in Figure 2.10. This model corrects the volatility clustering for daily returns. Comparing these two figures shows that SPY's conditional heavy tails are smaller than its unconditional heavy tails.

2.4.3 Standardised returns

By comparing the descriptive statistics for SPY returns in Table 2.1 with those for returns standardised by different volatility measures shown in Table 2.2, we can see that standardising returns using volatility measures does not have a dramatic impact on the mean of the returns.

[Insert Table 2.2 here]

However, the standardisation does produce a fall in the absolute values of the minimum and maximum values, as well as the skewness and kurtosis of their distributions. In addition, the results of the Jarque-Bera (JB) test for normality for the standardised returns are generally closer to a normal distribution across different frequencies. Volatility measures from different sampling frequencies have different effects on standardising returns for SPY, as shown in Table 2.2. SPY returns are more likely to be normally distributed when standardising volatility measures using 30-second sampling frequencies than for lower frequencies. For example, we fail to reject the null hypothesis at the 5% significance level after performing the JB test on SPY's daily returns after standardising all the volatility measures estimated using 30-second intraday returns. This is likely because the volatility measures estimated using higher-frequency data (e.g. 30-second data) can more easily capture the large volatility caused by large changes in stock returns. Therefore the extreme values in the return distributions can be standardised more easily with 30-second data. However, when the estimation uses 300-second intraday returns, this effect only holds for QPV, minRV and TBV.

2.5 Leverage effects and long-memory properties

In the previous section, we discussed some of the stylised facts (e.g. autocorrelation and heavy tails) of calendar-time intraday returns using the graphs in Figures 2.6 to 2.10. The use of graphs for analyses necessitated restricting the discussion to SPY as representative of the whole market. In this section, however, we investigate volatility forecasting using results from stocks, as well as SPY, for the whole data set (2000-2016). We examine stocks' leverage effects, long-memory properties and the correlations between trading volume and volatility measures, which are important considerations in volatility estimation and forecasting. The stocks considered in this section are AAPL and MSFT (IT sector), and JNJ and PFE (healthcare sector). These were chosen because IT is the most volatile sector due to the rise of cloud computing, big data and mobile computing, while healthcare is relatively stable.

2.5.1 Leverage effects

Table 2.3 shows the leverage effects estimated from the EGARCH (1,1) model for the stocks and the SPY index. We choose the minRV volatility measure because it is one of the most effective at standardising returns, as discussed in Section 2.4.3. The EGARCH model is used because previous literature finds that negative shocks have a bigger effect on the future volatility of stocks than positive shocks (e.g. Chou, 1988; Baillie & De Gennaro, 1990; Tiwar et al., 2019). It is clear from Table 2.3 that the stocks and SPY have leverage effects with significant negative coefficients.

[Insert Table 2.3 here]

The leverage effects for SPY are much higher than those for individual stocks, with an absolute value of -0.115 for SPY compared to a range of -0.069 to -0.027 for stocks. Also, the leverage effects fall after standardising returns when using the minRV volatility measure, especially for lower-frequency stock returns. For example, the leverage effect coefficients for 60-second, 150-second and 300-second standardised intraday stock returns are not significant, as shown in Table 2.3. By observing the 30-second standardised stock returns, we can see that the leverage effects for AAPL and MSFT are not significant, while those for JNJ and PFE are significant at the 5% level but with smaller absolute values for the estimated parameters. In addition, the estimated leverage effect parameters for SPY's standardised returns are either non-significant (150 seconds), or significant but with much smaller absolute values (30 seconds, 60 seconds and 300 seconds).

2.5.2 Long-memory properties

The long-memory properties for the intraday returns for SPY and stocks can be estimated using an autoregressive fractionally integrated moving average (ARFIMA) model, which is shown in Table 2.4.

[Insert Table 2.4 here]

The long-memory results in Table 2.4 show that the long-memory properties for TRV and TBV are higher on average than other volatility measures, while those for RV_{ac} are lower than others on average. In addition, the table shows that the volatility measures estimated from highly volatile stocks (i.e. AAPL and MSFT) have higher average long-memory properties than those estimated from less volatile stocks (i.e. JNJ and PFE). This is because the stocks from the IT sector tend to have stronger volatility clustering after the burst of the tech bubble (2000 to 2002) and during the financial crisis (2008 to 2009), as the large changes in prices for stocks cluster together, as shown in Figures 2.1 to 2.5. Therefore, this high persistence in price change magnitudes for IT stocks results in larger long-memory properties for their volatility measures.

2.5.3 Trading volume and volatility correlation

Table 2.5 shows the correlation between volatility measures and trading volume for stocks and SPY using correlation coefficients. Trading volume is the aggregated number of shares traded during a given day. The correlation coefficients between the same-day volatility measures and trading volume (i.e. both at time *t)* are calculated and reported in Table 2.5. The results in show that, on average, two of the bias-corrected volatility measures, namely Quad-Power Variation (QPV) and realised variance (RV_{AC}), have stronger correlations with trading volume (0.385 and 0.383 respectively) than the other measures, including uncorrected measures such as RV.

[Insert Table 2.5 here]

Also, the bias-corrected Skipped Bi-power Variation (SBV), Threshold Bi-power Variation (TBV) and Corrected Threshold Bi-power Variation (CTBV) measures shows a higher correlation with trading volume than their uncorrected equivalent, BV (average correlations of 0.364, 0.371 and 0.378 versus 0.351 respectively). However, among the bias-corrected measures, QPV is more correlated with trading volume than SBV, suggesting that both the jump and continuous components in realised variance are closely correlated with trading volume. These results demonstrate the value of examining the impact of the lag of trading volume on realised variance estimation and forecasting.

2.6 Volatility forecasting

The long-memory results in the previous section show that long-memory properties are present in the volatility measures estimated from intraday returns, and that volatility measures are significantly correlated with trading volume. It is therefore a worthy endeavour to study the impact of trading volume on forecasting realised variances and compare the performance of the trading volume model with other HAR-family models. We use lasso regressions to examine the impact of 100 lags of trading volume on realised variance, shown in Figure 2.11. We first chose the best tuning parameter using 10-fold cross-validation in the lasso regression. Then the best tuning parameters were used in the lasso regression with 10,000 iterations to select the relevant lags for predicting RV_{t+1}. Figure 2.11 shows that the first lag for trading volume is an important predictor for RV for most stocks across different frequencies, hence why it was chosen for inclusion in our HAR forecasting model.

[Insert Figure 2.11 here]

Table 2.6 reports the regression results for HAR family models using price data from the SPY ETF that tracks the S&P 500 index. The table also includes the MAE ratio, which compares the forecasting performance of the HAR-J-Vol and HAR-TJ-Vol models versus the HAR-J and HAR-TJ models. The Diebold and Mariano (DM) test results also reported in Table 2.6 show whether the inclusion of the first lag of trading volume in the HAR-J and HAR-TJ models yields significant improvements in forecasting. The alternative hypothesis for the DM test is that the HAR family models with the first lag of trading volumes perform better than those without it.

[Insert Table 2.6 here]

The regression results for HAR family models in Table 2.6 show that the estimated coefficient of the first lag of trading volume is significant. Although the estimated coefficients are small, they improve the value of R-squared for the HAR-J and HAR-TJ models.

Table 2.6 also shows that considering the first lag of volume improves the forecasting performance for HAR-J across all stocks and frequencies. This result holds for the majority of stocks for the HAR-TJ model. The DM test results show that these improvements are significant at the 5% level for more than half of the cases. The results suggest that the HAR-J-Vol model is better at forecasting than the HAR-J model, as its MAE ratio is smaller than the latter's across different stocks and frequencies. This improvement is significant at the 5% significance level for the PFE and JNJ stocks across all frequencies, and for AAPL and MSFT at certain sampling frequencies. The DM results show that the improvement in forecasting is significant for SPY using 30-second and 60-second sampling frequencies at the 10% significance level. The MAE ratios for the HAR-TJ-Vol model versus the HAR-TJ model are less than 1 for SPY, JNJ and MSFT across all frequencies and for AAPL and PFE at certain frequencies. The DM test results show that the gains of the HAR-TJ-Vol model are significant for JNJ and MSFT at the 1% significance level across all frequencies, while the improvements for SPY, AAPL and PFE are significant at the 1% and/or 5% significance levels for some of the frequencies. In summary, then, it is clear that adding the first lag of trading volume to HAR models can significantly improve stock volatility forecasting in the majority of cases.

2.7 Stylised facts of stock returns in different regimes

Previous sections discussed the stylised facts of the high-frequency aggregated SPY return and realised volatility measures, followed by the impact of these measures on standardising the return. In addition, we analysed the correlation between trading volume and volatility measures and how they affect unconditional volatility forecasting for stocks and SPY. In this section, we discuss the impact of volatility measures affecting the high frequency SPY returns in the pre-crisis, crisis and post-crisis periods. We also examine the correlation between volatility measures and trading volume, and the impact of the latter on volatility forecasting for stocks and SPY in different regimes.

2.7.1 Standardised SPY returns

Tables 2.7 to 2.9 show the descriptive statistics for SPY returns standardised using different realised measures in the pre-crisis (Table 2.7), crisis (Table 2.8) and post-crisis (Table 2.9) periods.

[Insert Tables 2.7 to 2.9 here]

The changes in maximum and minimum values of standardised returns, as well as the skewness and kurtosis of returns' distributions, all decrease across different regimes, which is in line with the changes in the returns of the whole data set (2000-2016) discussed in Section 2.4.3. In addition, the JB test results also show that standardising returns with volatility measures results in more normal distributions across different regimes, which is in line with the results for the 2000-2016 returns. Comparing the JB tests for the three periods, we find standardising returns is more effective in the most volatile period (2008 to 2009), as the returns all fail to reject the null hypothesis during this period (i.e. the returns follow a normal distribution). For the less volatile periods, the 300-second returns reject the null hypothesis for some of the volatility measures: three for the pre-crisis period (RV, SBV and RV_{ac}) and only one for the post-crisis period (RV_{ac}). This is because the returns for SPY display volatility clustering during the financial crisis, which results in much bigger changes in prices during this period than before or after the crisis. These big changes in prices can easily be captured by volatility measures. Therefore, standardising returns with volatility measures can help eliminate the impact of large changes in returns and can yield return distributions that are closer to a normal distribution.

2.7.2 Leverage effects

Tables 2.10 to 2.12 show the leverage effects for returns and standardised returns for stocks and SPY in the pre-crisis, crisis and post-crisis periods.

[Insert Tables 2.10 to 2.12 here]

The results in Tables 2.10 and 2.12 show that leverage effects are present in the 2000-2016 returns (as shown in Section 2.5.1), but they are also significant in returns for most stocks in the pre-crisis, crisis and post-crisis periods. (The exceptions to this are AAPL before the crisis and PFE during it.) When standardising returns using volatility measures, the leverage effects either fall dramatically or became non-significant across different regimes, which is in line with the returns for the data set as a whole. For the assets that do exhibit leverage effects during the crisis (that is, all except PFE), the effects are much larger during the crisis than in the pre-crisis and post-crisis periods, as the estimated parameters are all below -1.3.

2.7.3 Trading volume and volatility correlation

The results from Tables 2.13 to 2.15 show that the bias-corrected BV measures (SBV, TBV and CTBV) have higher correlations with trading volume than BV across different regimes, which is in line with the 2000-2016 data shown in Section 2.5.3.

[Insert Tables 2.13 to 2.15 here]

By comparing the correlation results in different regimes, we find that the volatility measures have, on average, the highest correlations with trading volume during the financial crisis. This is because investors tend to trade more frequently in response to fears of a crisis, as shown in Section 2.2. The correlation between the volatility measures and trading volume are particularly low on average in the pre-crisis period. Some volatility measures are even negatively or not significantly correlated with trading volume for the AAPL and PFE stocks and the SPY index. This indicates that although the burst of the tech bubble (2000-2002) affected the volatility of stocks dramatically, this large volatility was not necessarily caused by high trading volume; rather, the cause may be the large amount of sell orders compared to buy orders (see Section 2.2). Therefore, the correlation between volatility measures and trading volume may not always be significant or positive.

2.7.4 Volatility forecasting

Tables 2.16 to 2.18 show the regression results for HAR family models and their one-day ahead forecasting results using the pre-crisis (2000-2007), crisis (2007-2009) and post-crisis (2010-2016) data respectively. The results show that the coefficients for the first lag of trading volume are significant for both the HAR-J-Vol and HAR-TJ-Vol models across all frequencies for SPY for the post-crisis period in Table 2.18. However, the coefficients for trading volume are not significant for the pre-crisis and crisis periods. Considering trading volume also improves the R-squared values for both HAR models in the post-crisis period, though changes in R-squared are negligible for the pre-crisis period and mixed for the crisis period.

This suggests that the impact of trading volume on RV in regression varies for different financial regimes, which is reflected in the forecasting results. The HAR-J-Vol and HAR-TJ-Vol models perform better at forecasting RV, as their MAE ratios are less than 1 for all stocks in the post-crisis period (Table 2.18) across different frequencies with only one exception (SPY at 300 sec for the HAR-J-Vol versus the HAR-J model). The biggest improvement is for AAPL as its MAE ratios are all less than 0.8, with some of them close to 0.5 (e.g. 0.511 for HAR-J-Vol versus HAR-J using 30-second sampling). The improvements are significant for all stocks at the 5% significance level across all frequencies with only two exceptions. In line with the results for the full data set (2000-2016) discussed in Section 2.6, the DM test results show fewer significant improvements for SPY across different sampling frequencies in the post-crisis period compared to stocks.

The forecasting results in Table 2.16 show that the MAEs are smaller for the models that consider the first lag of trading volume for most cases in the pre-crisis period. However, most of the improvements are not significant for the HAR-J-Vol model, while the significant improvements for the HAR-TJ-Vol model can be found for stocks AAPL and PFE at the 1% significance level, but not for JNJ, MSFT and SPY. Table 2.17 shows that including the first lag of trading volume does not produce any obvious improvements for most stocks and frequencies in the crisis period. In sum, the findings suggest that trading volume has a more significant impact on estimating and forecasting volatility in the less volatile periods such as the post-crisis period.

2.8 Conclusion

In this chapter, we have examined the stylised facts of high-frequency returns and volatility measures from stock markets. We find that the intraday returns for the SPY index do not display autocorrelations across different frequencies, yet the opposite is true for their absolute values. By investigating the stylised facts of high-frequency aggregated returns, we find that high-frequency daily returns have both unconditional and conditional heavy tails, the latter of which are smaller than the former. In addition, we find that the volatility measures of intraday returns have autocorrelations and long-memory properties, the latter of which are tested using eleven measures, some of which have not been considered in previous literature. The bias-corrected volatility measures based on RV and BV have higher long-memory properties than the original uncorrected measures. In addition, long-memory properties are higher for RV (which captures both the continuous and jump components) than for BV (which only captures the continuous components). This shows that both jump components and continuous components have long-memory properties. We also find that long-memory properties are higher for volatility measures estimated from more volatile stocks such as MSFT and AAPL.

Our examination of the correlation between trading volume and eleven volatility measures reveals that correlations are present for both the whole data set and the data from the crisis and post-crisis periods. The highest correlation between trading volume and volatility measures are in during the crisis. However, the negative and non-significant correlations between trading volume and volatility for AAPL, PFE and the SPY index during the burst of the tech bubble highlight the fact that high volatility is not always caused by trading volume only, but may also be affected by the type of trading orders.

By standardising SPY returns using volatility measures, we find that the standardised returns are closer to normal distributions. The impact of volatility measures on returns is more obvious for returns during the financial crisis than for the pre- and post-crisis periods. We also find evidence that intraday returns have leverage effects and that standardising returns using the minRV volatility measure can help eliminate or decrease the leverage effects for the data from 2000-2016 and from different regimes. In addition, the leverage effects for most asset returns (except PFE) tend to be higher during the crisis than before or after it.

In this chapter, we also examined volatility forecasting for stocks and SPY across different sampling frequencies using HAR-family models. The forecasting results show that the HAR-J-Vol and HAR-TJ-Vol models performs better than HAR-J and HAR-TJ models across the whole period under study (2000-2016) as well as the post-crisis period. This indicates that trading volume may be helpful for volatility forecasting. This improvement of forecasting with using the first lag of volumes results in more number of significant results in the post-crisis periods, suggesting the impact of past trading volumes on stock volatility are reduced in more volatile period. Future work may wish to explore and assess how various volatility measures, such as the eleven studied in this thesis, can be used in volatility estimation and forecasting. This particularly applies to parametric methods, which would represent a step beyond the non-parametric methods used to date.

Appendix

Figures

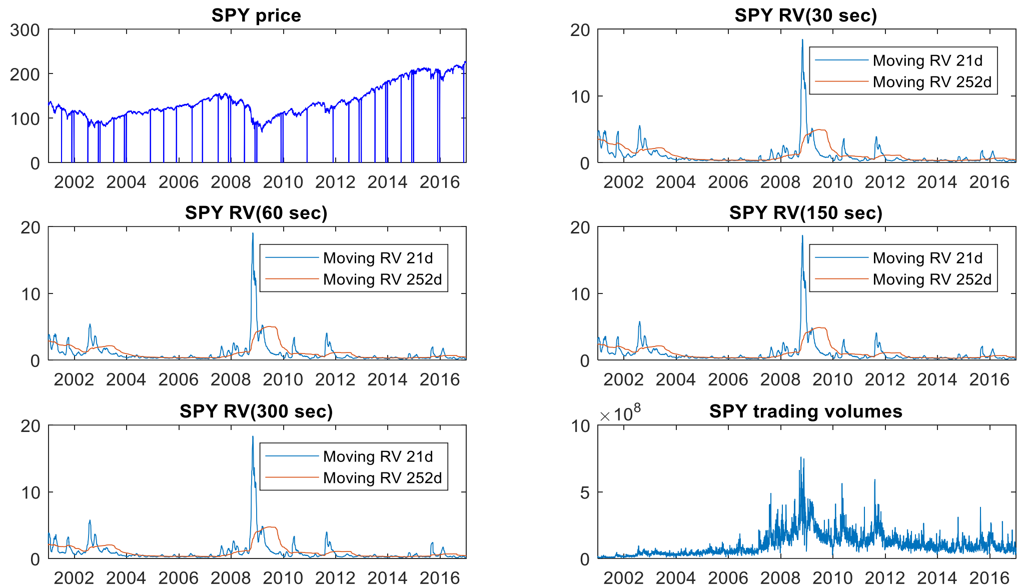

Figure 2.1: Price, volatility and trading volume for the SPY index. Volatility is measured using 21-day and 252-day moving average realised variance (RV), estimated with 30-second, 60-second, 150-second and 300-second returns.

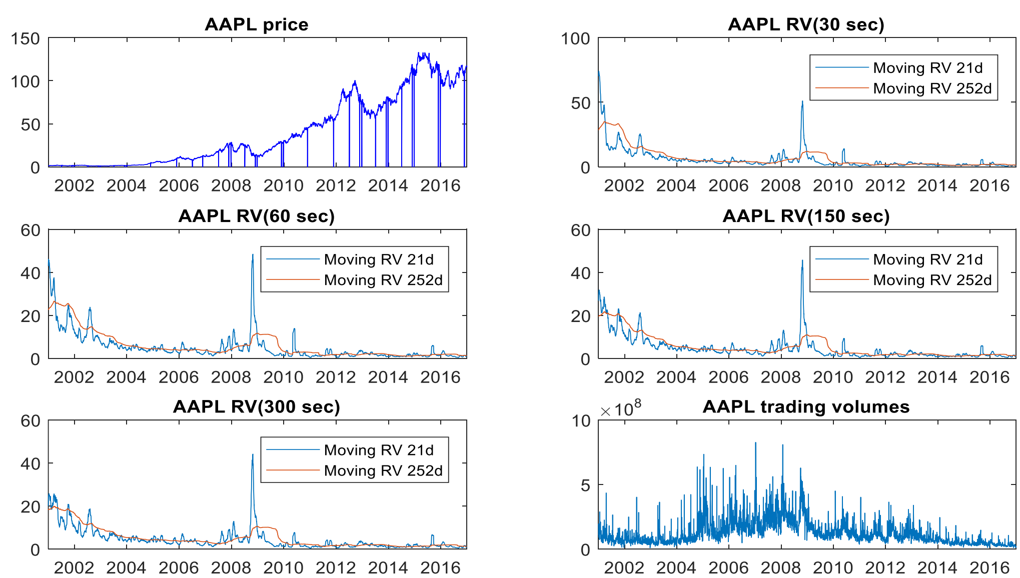

Figure 2.2: Price, volatility and trading volume for the AAPL stock (IT sector). Volatility is measured using 21-day and 252-day moving average realised variance (RV), estimated with 30-second, 60-second, 150-second and 300-second returns.

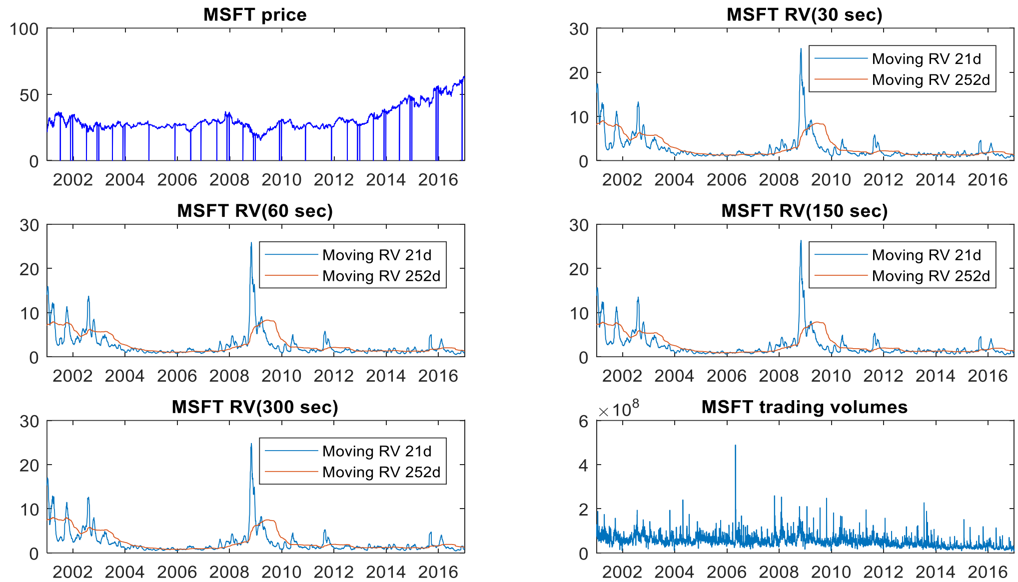

Figure 2.3: Price, volatility and trading volume for the MSFT stock (IT sector). Volatility is measured using 21-day and 252-day moving average realised variance (RV), estimated with 30-second, 60-second, 150-second and 300-second returns.

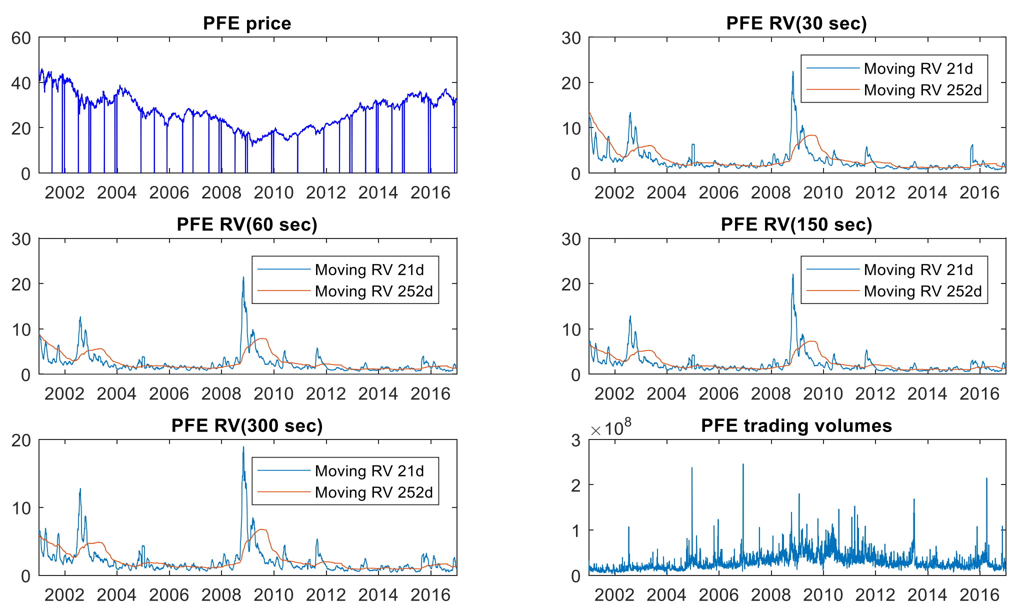

Figure 2.4: Price, volatility and trading volume for the PFE stock (healthcare sector). Volatility is measured using 21-day and 252-day moving average realised variance (RV), estimated with 30-second, 60-second, 150-second and 300-second returns.

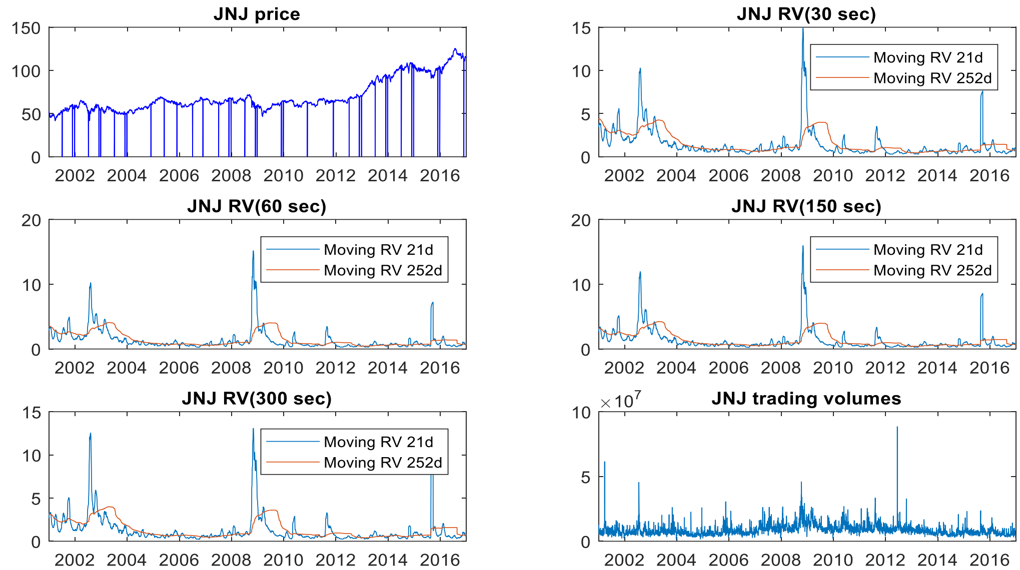

Figure 2.5: Price, volatility and trading volume for the JNJ stock (healthcare sector). Volatility is measured using 21-day and 252-day moving average realised variance (RV), estimated with 30-second, 60-second, 150-second and 300-second returns.

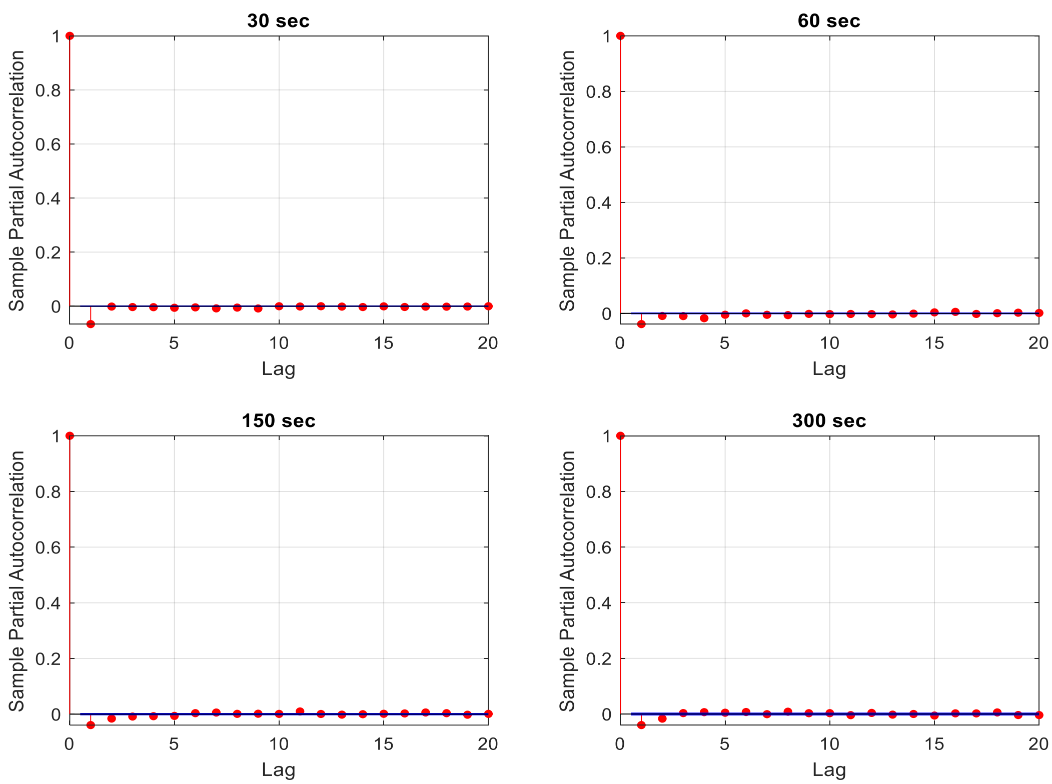

Figure 2.6: Partial autocorrelations for intraday returns for the SPY index, using 30-second, 60-second, 150-second and 300-second returns.

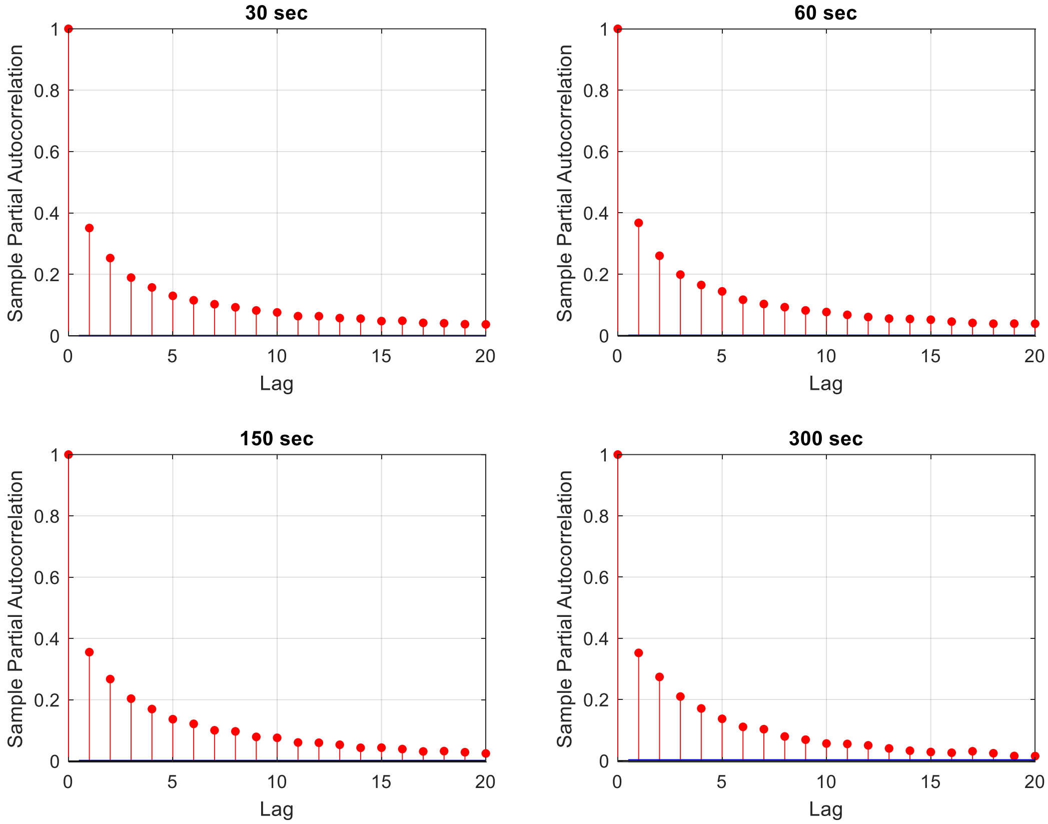

Figure 2.7: Partial autocorrelations for absolute intraday returns for the SPY index, using 30-second, 60-second, 150-second and 300-second returns.

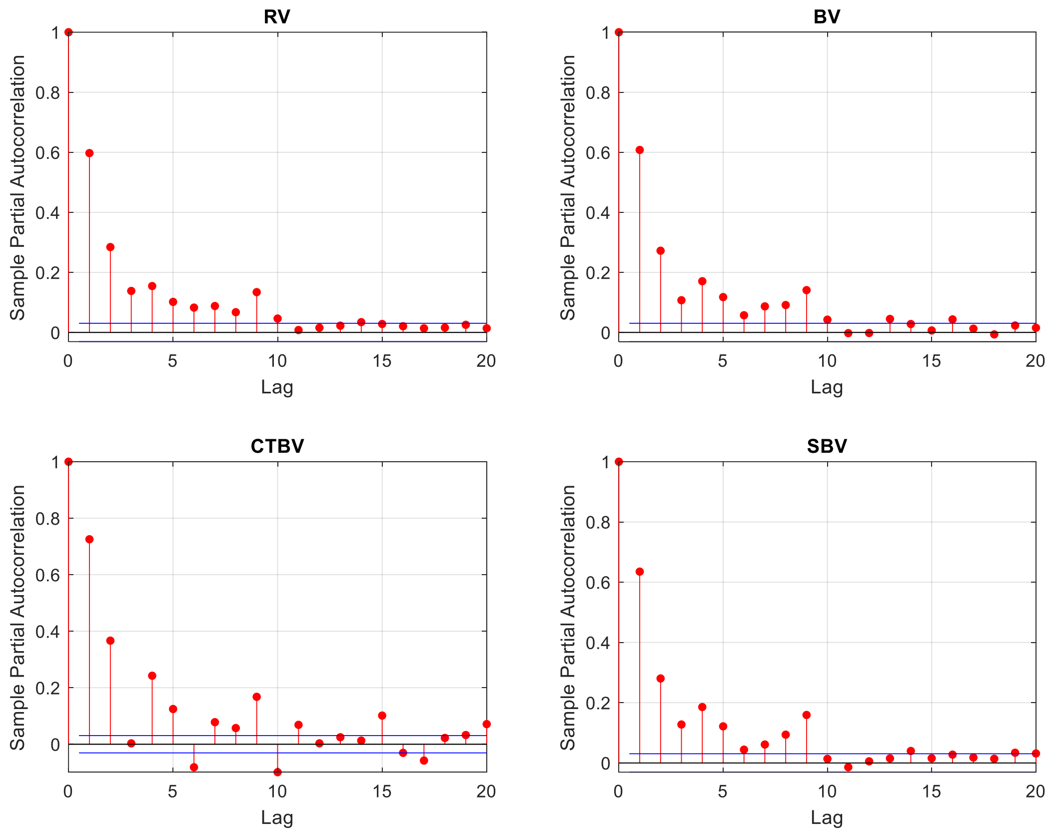

Figure 2.8: Partial autocorrelations for 300-second returns for the SPY index using RV, BV, CTBV and SBV.

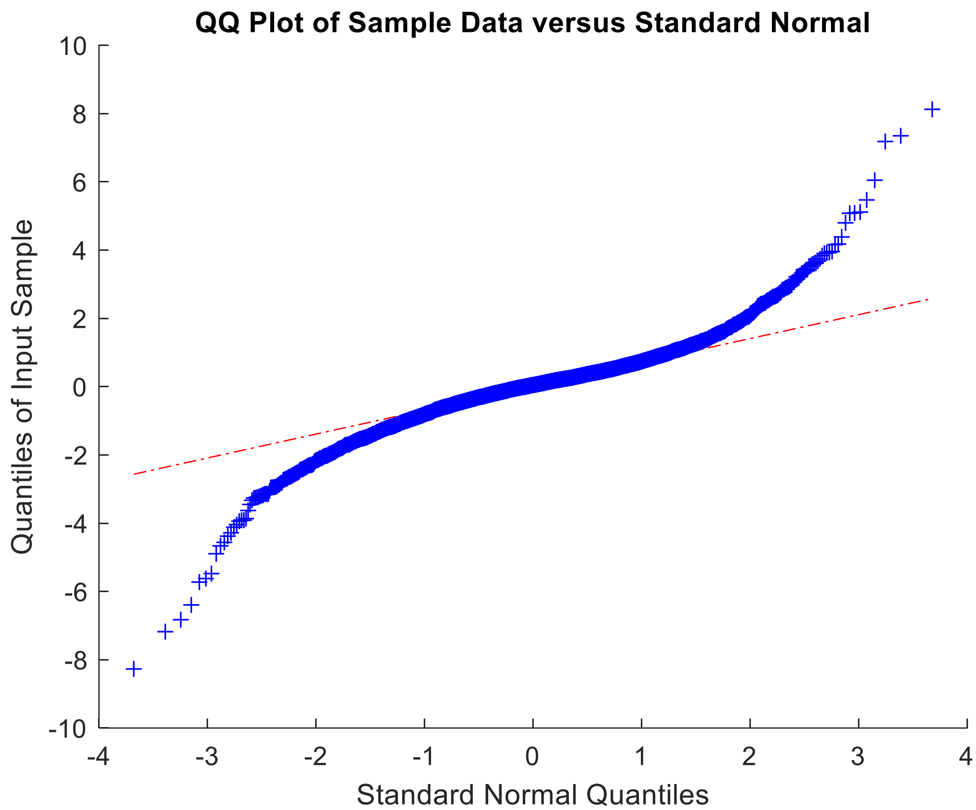

Figure 2.9: Quantile-quantile (Q-Q) plot for the daily returns for the SPY index.

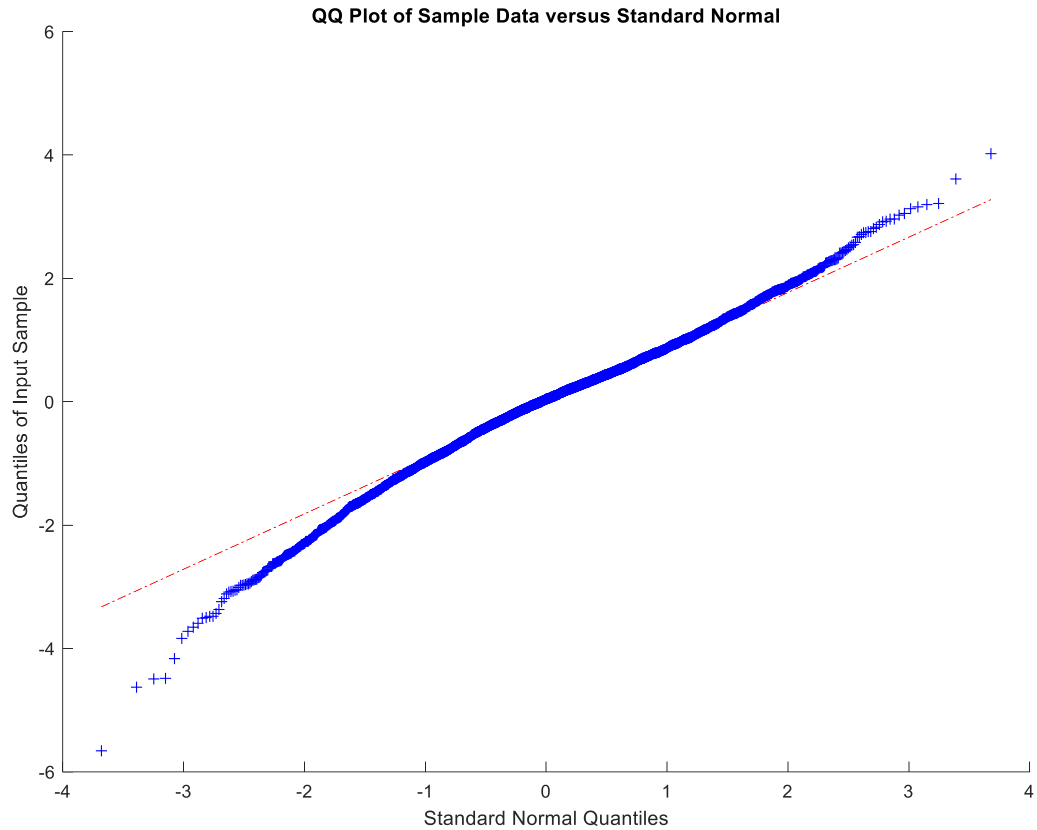

Figure 2.10: Quantile-quantile (Q-Q) plot for the daily returns for the SPY index after correcting for volatility clustering using the GARCH (1,1) model.

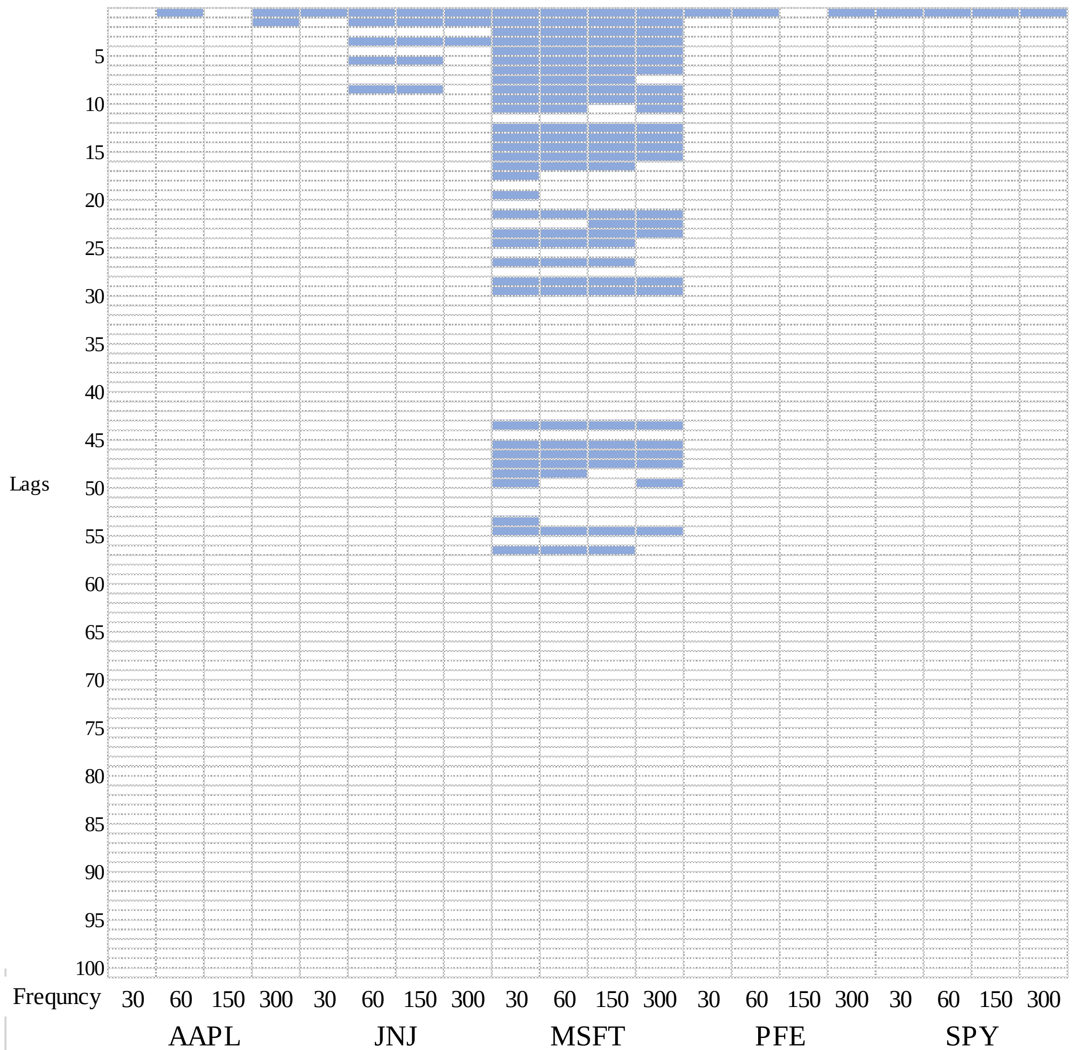

Figure 2.11: Lags for trading volumes selected by LASSO regression for RV predictions using the best tuning parameters. Each blue cell represents one lag selected by the regression model.

Tables

Table 2.1 Descriptive statistics for daily returns and conditional variance

	AAPL		JNJ		PFE		MSFT		SPY	
	Return	***GARCH***	***Return***	***GARCH***	***Return***	***GARCH***	***Return***	***GARCH***	***Return***	***GARCH***
					2000-2016					
Mean	-0.012	4.922	0.020	1.086	-0.011	1.879	0.009	2.440	0.000	1.033
STD. DEV.	2.205	4.767	1.035	1.328	1.379	1.492	1.571	2.585	1.024	1.568
Median	0.015	3.326	0.017	0.691	0.000	1.315	0.000	1.440	0.048	0.576
Max	12.540	42.173	7.921	16.290	6.925	10.623	11.050	20.098	8.124	20.936
Min	-12.201	0.620	-7.486	0.209	-6.812	0.428	-7.753	0.406	-8.270	0.149
Skewness	0.048	2.540	0.057	4.729	0.113	2.121	0.207	2.933	-0.077	6.049
Kurtosis	6.256	11.861	9.047	33.559	5.719	8.336	7.081	13.985	10.326	51.016
JB Test	1891	18590	6519	182360	1327	8281	2998	27637	9569	436940
p-value	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
					Pre-crisis					
Mean	0.009	6.911	0.033	1.361	-0.022	2.292	-0.017	2.997	-0.020	0.961
STD. DEV.	2.656	3.849	1.160	1.237	1.527	1.414	1.726	2.966	0.983	0.809
Median	0.000	5.442	0.031	0.955	-0.068	1.712	-0.037	1.798	0.024	0.656
Max	12.285	21.536	7.921	8.598	6.925	6.976	11.050	21.700	8.124	7.029
Min	-11.405	2.441	-7.486	0.218	-6.192	0.686	-7.753	0.361	-4.894	0.229
Skewness	0.117	1.391	0.125	2.593	0.292	1.023	0.268	1.877	0.251	2.597
Kurtosis	4.550	4.335	6.694	11.229	4.935	2.979	6.080	7.533	6.978	13.543
JB Test	206	798	1148	7923	342	350	819	2901	1347	11567
p-value	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
					Crisis					
Mean	-0.035	8.618	-0.061	2.123	-0.114	4.168	-0.085	5.755	-0.036	4.025
STD. DEV.	2.955	7.097	1.482	2.773	2.007	2.727	2.406	4.056	2.001	4.135
Median	0.147	6.365	-0.039	1.115	-0.281	3.488	-0.267	4.404	-0.015	2.417
Max	12.540	49.416	7.487	18.382	6.667	12.998	10.890	21.156	7.349	24.857
Min	-12.201	2.610	-7.450	0.516	-6.812	1.045	-7.565	2.131	-8.270	0.663
Skewness	-0.257	3.028	0.054	3.244	-0.058	1.422	0.357	1.921	-0.108	2.318
Kurtosis	5.206	13.683	8.807	14.576	4.224	4.273	4.951	6.166	5.332	8.535
JB Test	81	2369	530	2766	24	153	68	389	86	819
p-value	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
					Post-crisis					
Mean	-0.030	1.779	0.022	0.569	0.022	1.025	0.055	1.263	0.028	0.551
STD. DEV.	1.329	0.900	0.745	0.360	1.011	0.428	1.122	0.386	0.741	0.473
Median	0.014	1.577	0.015	0.467	0.031	0.886	0.021	1.166	0.069	0.407
Max	8.348	12.712	5.023	4.044	4.313	3.773	4.727	4.388	3.599	5.591
Min	-6.845	0.729	-5.280	0.234	-5.734	0.461	-6.067	0.747	-4.281	0.168
Skewness	-0.050	3.914	-0.129	3.923	-0.107	1.823	-0.172	2.309	-0.417	4.079
Kurtosis	4.786	33.434	6.997	27.373	4.505	7.552	4.863	12.077	6.093	27.812
JB Test	252	77767	1264	51630	182	2678	283	8167	808	53722
p-value	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Note: This table reports the descriptive statistics and the results of the Jarque-Bera test for normality for returns and conditional variance for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index in different financial regimes.

Table 2.2 Descriptive statistics for SPY daily returns standardised by volatility measures

	$\frac{r_t}{\sqrt{RV_t}}$	$\frac{r_t}{\sqrt{BV_t}}$	$\frac{r_t}{\sqrt{TRV_t}}$	$\frac{r_t}{\sqrt{QPV_t}}$	$\frac{r_t}{\sqrt{TPV_t}}$	$\frac{r_t}{\sqrt{minRV_t}}$	$\frac{r_t}{\sqrt{medRV_t}}$	$\frac{r_t}{\sqrt{SBV_t}}$	$\frac{r_t}{\sqrt{TBV_t}}$	$\frac{r_t}{\sqrt{CTBV_t}}$	$\frac{r_t}{\sqrt{RVac_t}}$
					30 seconds						
Mean	0.055	0.059	0.066	0.070	0.064	0.060	0.060	0.061	0.069	0.062	0.062
STD. DEV.	0.957	1.000	1.020	1.106	1.048	1.002	0.997	1.010	1.079	1.009	1.009
Median	0.074	0.076	0.076	0.087	0.082	0.076	0.075	0.077	0.081	0.075	0.077
Max	3.305	3.385	3.611	3.687	3.509	3.425	3.384	3.495	3.822	3.487	3.502
Min	-3.076	-3.270	-3.515	-3.802	-3.414	-3.298	-3.204	-3.330	-3.675	-3.276	-3.467
Skewness	-0.014	-0.003	0.035	0.016	0.005	0.001	0.001	0.003	0.031	0.006	0.008
Kurtosis	2.939	2.915	2.977	2.907	2.901	2.933	2.929	2.924	3.000	2.950	2.881
JB test	0.811	1.294	0.984	1.724	1.782	0.810	0.895	1.046	0.699	0.479	2.579
p-value	0.500	0.500	0.500	0.411	0.400	0.500	0.500	0.500	0.500	0.500	0.273
					60 seconds						
Mean	0.057	0.062	0.070	0.071	0.067	0.063	0.063	0.064	0.072	0.064	0.066
STD. DEV.	0.972	1.005	1.032	1.074	1.037	1.015	1.014	1.016	1.080	1.016	1.011
Median	0.074	0.077	0.080	0.084	0.082	0.079	0.077	0.080	0.083	0.077	0.080
Max	3.322	3.410	3.468	3.582	3.508	3.476	3.605	3.534	3.771	3.437	3.300
Min	-3.113	-3.370	-3.670	-4.103	-3.762	-3.526	-3.430	-3.443	-4.028	-3.425	-3.018
Skewness	-0.005	0.009	0.043	0.028	0.021	0.016	0.018	0.014	0.037	0.017	0.019
Kurtosis	2.894	2.900	2.948	2.924	2.912	2.932	2.935	2.915	2.973	2.915	2.788
JB test	2.044	1.855	1.815	1.588	1.682	1.006	1.002	1.413	1.116	1.511	8.261
p-value	0.354	0.386	0.394	0.440	0.420	0.500	0.500	0.487	0.500	0.459	0.017**
					150 seconds						
Mean	0.063	0.069	0.077	0.078	0.074	0.073	0.072	0.071	0.083	0.073	0.067
STD. DEV.	0.996	1.037	1.062	1.089	1.063	1.064	1.056	1.041	1.127	1.052	1.035
Median	0.077	0.078	0.080	0.080	0.079	0.079	0.080	0.078	0.084	0.078	0.087
Max	3.256	3.648	3.551	3.906	3.823	3.977	3.786	3.592	3.916	3.707	3.137
Min	-3.135	-3.164	-3.194	-3.182	-3.198	-3.255	-3.203	-3.160	-3.224	-3.175	-2.946
Skewness	0.005	0.039	0.071	0.066	0.055	0.057	0.042	0.034	0.091	0.052	0.008
Kurtosis	2.814	2.866	2.895	2.909	2.886	2.922	2.900	2.854	2.962	2.871	2.642
JB test	6.198	4.291	5.584	4.602	4.485	3.410	3.061	4.599	6.179	4.921	22.891
p-value	0.045**	0.115	0.061	0.098	0.104	0.178	0.213	0.098	0.045**	0.084	0.001***
					300 seconds						
Mean	0.065	0.071	0.078	0.080	0.075	0.075	0.074	0.074	0.085	0.074	0.064
STD. DEV.	1.015	1.059	1.082	1.107	1.083	1.095	1.081	1.059	1.162	1.079	1.054
Median	0.082	0.084	0.084	0.085	0.084	0.084	0.084	0.083	0.086	0.085	0.088
Max	3.227	3.270	3.837	3.555	3.382	3.699	3.525	3.185	4.501	3.416	3.717
Min	-2.960	-3.294	-3.894	-3.404	-3.381	-3.423	-3.220	-3.082	-3.487	-3.315	-2.981
Skewness	0.004	0.021	0.060	0.042	0.033	0.039	0.042	0.037	0.088	0.031	-0.020
Kurtosis	2.718	2.785	2.852	2.844	2.809	2.883	2.798	2.754	2.982	2.796	2.607
JB test	14.184	8.521	6.482	5.587	7.303	3.499	8.501	11.759	5.619	8.134	27.885
p-value	0.001***	0.015**	0.039**	0.061	0.027**	0.171	0.015**	0.004***	0.060	0.018**	0.001***

Note: This table reports the descriptive statistics and the results of the Jarque-Bera test for normality for eleven different volatility measures for the SPY index using 30-second, 60-second, 150- second and 300-second returns. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Table 2.3 Leverage effects estimated from the EGARCH (1,1) model (2000-2016)

	Return (r_t)	$\frac{r_t}{\sqrt{minRV_t}}$			
		30 sec	*60 sec*	*150 sec*	*300 sec*
		AAPL (IT)			
Leverage	-0.027	0.004	0.008	0.004	0.019
S.E.	0.006	0.003	0.023	0.003	0.023
T-statistics	-4.323	1.309	0.350	1.321	0.809
p-value	0.000	0.190	0.726	0.186	0.418
		JNJ (HC)			
Leverage	-0.069	-0.034	-0.002	-0.006	-0.003
S.E.	0.008	0.012	0.008	0.007	0.007
T-statistics	-9.083	-2.742	-0.261	-0.786	-0.376
p-value	0.000	0.006	0.794	0.432	0.707
		PFE (HC)			
Leverage	-0.032	-0.008	-0.003	0.002	-0.015
S.E.	0.005	0.004	0.020	0.021	0.021
T-statistics	-6.938	-2.324	-0.129	0.093	-0.728
p-value	0.000	0.020	0.898	0.926	0.466
		MSFT (IT)			
Leverage	-0.035	0.003	0.015	0.011	-0.027
S.E.	0.006	0.003	0.015	0.014	0.022
T-statistics	-5.507	0.989	0.994	0.755	-1.232
p-value	0.000	0.323	0.320	0.450	0.218
		SPY			
Leverage	-0.115	-0.064	-0.062	-0.041	-0.046
S.E.	0.007	0.023	0.022	0.022	0.022
T-statistics	-15.600	-2.840	-2.781	-1.850	-2.053
p-value	0.000	0.005	0.005	0.064	0.040
Mean	-0.056	-0.020	-0.009	-0.006	-0.015

Note. This table shows the leverage effects estimated from the EGARCH (1,1) model for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the whole data set (2000-2016) using 30-second, 60-second, 150-second and 300-second returns.

Table 2.4 Long-memory properties for stocks' and SPY's volatility measures (2000-2016)

Sampling Frequency (Seconds)		RV_t	BV_t	TRV_t	QPV_t	TPV_t	$minRV_t$	$medRV_t$	SBV_t	TBV_t	$CTBV_t$	$RVac_t$	Mean
							AAPL (IT)						
30	***d***	0.474	0.429	0.499	0.363	0.386	0.468	0.480	0.433	0.498	0.489	0.341	0.442
	S.E.	0.040	0.049	0.001	0.053	0.051	0.047	0.042	0.045	0.004	0.035	0.051	0.038
60	***d***	0.432	0.420	0.490	0.404	0.405	0.455	0.455	0.426	0.484	0.471	0.372	0.438
	S.E.	0.041	0.038	0.026	0.041	0.039	0.036	0.037	0.040	0.042	0.046	0.041	0.039
150	***d***	0.427	0.412	0.467	0.399	0.403	0.412	0.427	0.425	0.477	0.451	0.372	0.425
	S.E.	0.035	0.034	0.041	0.037	0.036	0.034	0.036	0.037	0.045	0.041	0.034	0.037
300	***d***	0.396	0.387	0.414	0.387	0.383	0.372	0.384	0.399	0.459	0.425	0.361	0.397
	S.E.	0.036	0.039	0.039	0.038	0.037	0.040	0.039	0.034	0.052	0.042	0.036	0.039
							JNJ HC)						
30	***d***	0.289	0.301	0.422	0.391	0.368	0.306	0.307	0.397	0.391	0.396	0.333	0.355
	S.E.	0.060	0.062	0.086	0.094	0.072	0.060	0.060	0.076	0.089	0.094	0.076	0.075
60	***d***	0.277	0.318	0.417	0.351	0.351	0.350	0.326	0.334	0.463	0.415	0.310	0.356
	S.E.	0.054	0.058	0.087	0.070	0.064	0.059	0.059	0.068	0.086	0.087	0.061	0.068
150	***d***	0.261	0.257	0.404	0.334	0.304	0.234	0.224	0.339	0.412	0.427	0.373	0.324
	S.E.	0.062	0.062	0.081	0.061	0.058	0.062	0.055	0.062	0.070	0.087	0.047	0.064
300	***d***	0.232	0.248	0.399	0.364	0.373	0.181	0.257	0.385	0.420	0.417	0.396	0.334
	S.E.	0.058	0.058	0.063	0.055	0.050	0.051	0.056	0.054	0.065	0.059	0.046	0.056
							PFE (HC)						
30	***d***	0.359	0.359	0.469	0.353	0.359	0.376	0.373	0.399	0.459	0.435	0.332	0.388
	S.E.	0.042	0.051	0.076	0.068	0.060	0.051	0.049	0.057	0.095	0.081	0.052	0.062
60	***d***	0.381	0.394	0.448	0.382	0.385	0.406	0.401	0.388	0.436	0.412	0.351	0.399
	S.E.	0.057	0.077	0.084	0.099	0.092	0.071	0.070	0.082	0.104	0.092	0.071	0.082
150	***d***	0.349	0.339	0.401	0.335	0.332	0.316	0.324	0.350	0.479	0.369	0.359	0.359
	S.E.	0.071	0.077	0.075	0.082	0.078	0.069	0.074	0.081	0.065	0.086	0.061	0.074
300	***d***	0.368	0.377	0.404	0.368	0.376	0.373	0.376	0.390	0.447	0.420	0.331	0.385
	S.E.	0.046	0.062	0.069	0.059	0.065	0.072	0.063	0.062	0.108	0.090	0.034	0.066
							MSFT (IT)						
30	***d***	0.487	0.481	0.496	0.472	0.479	0.484	0.483	0.489	0.493	0.487	0.433	0.480
	S.E.	0.055	0.063	0.013	0.072	0.066	0.056	0.061	0.050	0.037	0.058	0.076	0.055
60	***d***	0.458	0.447	0.481	0.434	0.441	0.444	0.447	0.457	0.462	0.456	0.440	0.452
	S.E.	0.068	0.071	0.067	0.079	0.077	0.068	0.071	0.076	0.080	0.076	0.056	0.072
150	***d***	0.456	0.449	0.479	0.436	0.442	0.444	0.452	0.452	0.488	0.462	0.416	0.452
	S.E.	0.063	0.061	0.052	0.065	0.064	0.053	0.049	0.063	0.036	0.062	0.041	0.055
300	***d***	0.435	0.437	0.475	0.429	0.433	0.438	0.450	0.447	0.481	0.455	0.380	0.442
	S.E.	0.043	0.049	0.042	0.047	0.050	0.048	0.044	0.044	0.039	0.048	0.035	0.044
							SPY						
30	***d***	0.405	0.389	0.452	0.395	0.389	0.370	0.392	0.406	0.429	0.407	0.386	0.402
	S.E.	0.087	0.085	0.120	0.093	0.087	0.080	0.086	0.091	0.116	0.093	0.114	0.095
60	***d***	0.406	0.400	0.437	0.381	0.387	0.400	0.406	0.390	0.418	0.405	0.402	0.403
	S.E.	0.113	0.114	0.124	0.115	0.116	0.105	0.114	0.116	0.128	0.118	0.110	0.116
150	***d***	0.399	0.378	0.444	0.372	0.374	0.356	0.371	0.396	0.472	0.409	0.378	0.395
	S.E.	0.108	0.111	0.118	0.122	0.119	0.102	0.110	0.124	0.114	0.120	0.086	0.112
300	***d***	0.403	0.417	0.469	0.429	0.429	0.406	0.409	0.421	0.496	0.461	0.394	0.430
	S.E.	0.074	0.074	0.084	0.087	0.082	0.061	0.072	0.093	0.015	0.088	0.070	0.073

Note: This table shows the leverage effect coefficients (*d*) and their standard errors (*SE*) estimated from the ARFIMA model for volatility measures for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the whole data set (2000-2016) using 30-second, 60-second, 150-second and 300-second returns.

Table 2.5 Correlations between trading volume and volatility measures (2000-2016)

Sampling Frequency (Seconds)		RV_t	BV_t	TRV_t	QPV_t	TPV_t	$minRV_t$	$medRV_t$	SBV_t	TBV_t	$CTBV_t$	$RVac_t$
							AAPL (IT)					
30	***Estimate***	0.203	0.243	0.201	0.355	0.305	0.193	0.188	0.260	0.316	0.300	0.274
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.251	0.279	0.249	0.354	0.318	0.248	0.245	0.289	0.334	0.325	0.294
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.284	0.291	0.277	0.329	0.309	0.271	0.285	0.306	0.334	0.332	0.321
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.301	0.316	0.298	0.343	0.331	0.305	0.306	0.320	0.348	0.352	0.322
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							JNJ (HC)					
30	***Estimate***	0.325	0.351	0.361	0.438	0.413	0.336	0.331	0.398	0.370	0.376	0.403
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.344	0.381	0.388	0.423	0.412	0.389	0.376	0.392	0.398	0.406	0.387
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.332	0.326	0.392	0.394	0.372	0.301	0.298	0.394	0.389	0.402	0.415
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.303	0.317	0.398	0.409	0.408	0.247	0.326	0.408	0.383	0.402	0.424
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							PFE (HC)					
30	***Estimate***	0.139	0.168	0.092	0.275	0.224	0.098	0.095	0.167	0.155	0.161	0.313
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.214	0.244	0.186	0.303	0.275	0.193	0.196	0.244	0.228	0.236	0.295
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.260	0.276	0.250	0.292	0.287	0.258	0.258	0.272	0.239	0.264	0.281
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.282	0.294	0.264	0.316	0.305	0.273	0.269	0.288	0.256	0.270	0.301
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							MSFT (IT)					
30	***Estimate***	0.526	0.523	0.518	0.494	0.511	0.523	0.523	0.519	0.557	0.565	0.507
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.520	0.510	0.513	0.485	0.499	0.509	0.513	0.510	0.556	0.563	0.518
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.520	0.517	0.514	0.501	0.508	0.517	0.517	0.511	0.560	0.569	0.519
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.518	0.513	0.511	0.502	0.507	0.504	0.510	0.512	0.553	0.563	0.501
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							SPY					
30	***Estimate***	0.406	0.420	0.419	0.468	0.445	0.394	0.397	0.430	0.390	0.396	0.495
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.461	0.464	0.463	0.472	0.466	0.455	0.451	0.455	0.453	0.462	0.488
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.474	0.454	0.480	0.450	0.450	0.434	0.448	0.464	0.468	0.479	0.482
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.485	0.487	0.499	0.492	0.492	0.482	0.488	0.495	0.501	0.510	0.502
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mean	***Estimate***	0.340	0.351	0.346	**0.385**	0.373	0.330	0.334	0.364	0.371	**0.378**	**0.383**

Note: This table shows the correlation coefficients (estimates) and their significance levels (p-values) for different volatility measures and trading volumes for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the whole data set (2000-2016) using 30-second, 60-second, 150-second and 300-second sampling data.

Table 2.6: Estimation and out-of-sample forecasting losses (mean absolute errors) using the whole data set (2000-2016) (forecast horizon h=1)

		HAR-J				HAR-TJ			
		30	60	150	300	30	60	150	300
	β_0	0.125***	0.125***	0.119***	0.103***	0.193***	0.179***	0.167***	0.130***
	s.e.	0.032	0.031	0.031	0.030	0.031	0.030	0.029	0.029
	β_d	0.226***	0.227***	0.208***	0.246***	0.279***	0.232***	0.414***	0.478***
	s.e.	0.019	0.019	0.019	0.019	0.022	0.020	0.024	0.025
	β_w	0.468***	0.479***	0.493***	0.432***	0.487***	0.512***	0.482***	0.413***
	s.e.	0.033	0.032	0.032	0.032	0.035	0.033	0.036	0.037
	β_m	0.228***	0.211***	0.205***	0.242***	0.213***	0.196***	0.114***	0.138***
	s.e.	0.029	0.028	0.029	0.029	0.030	0.029	0.030	0.031
	β_j	0.049	-0.591***	0.116	0.083	-0.078	0.008	-0.303***	-0.125***
	s.e.	0.245	0.200	0.238	0.127	0.058	0.063	0.039	0.029
	$\bar{R}^2$	0.545	0.541	0.518	0.509	0.552	0.543	0.553	0.548
		HAR-J-Vol				HAR-TJ-Vol			
SPY	β_0	-0.010	-0.016	-0.033	-0.042	0.061	0.025	0.018	0.029
	s.e.	0.040	0.039	0.039	0.039	0.040	0.039	0.037	0.037
	β_d	0.205***	0.203***	0.182***	0.220***	0.257***	0.205***	0.382***	0.453***
	s.e.	0.019	0.020	0.019	0.019	0.022	0.020	0.025	0.025
	β_w	0.460***	0.471***	0.483***	0.422***	0.479***	0.503***	0.475***	0.410***
	s.e.	0.033	0.032	0.032	0.032	0.035	0.033	0.036	0.037
	β_m	0.229***	0.206***	0.197***	0.231***	0.214***	0.191***	0.107***	0.131***
	s.e.	0.029	0.028	0.028	0.029	0.030	0.029	0.030	0.031
	β_j	-0.063	-0.571***	-0.022	0.072	-0.085	-0.022	-0.321***	-0.132***
	s.e.	0.245	0.199	0.237	0.126	0.058	0.063	0.039	0.029
	β_v	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.001***
	s.e.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	$\bar{R}^2$	**0.548**	**0.545**	**0.522**	**0.514**	**0.554**	**0.547**	**0.557**	**0.550**
	MAE ratio	0.906*	0.911*	0.923	0.942	0.893***	0.888***	0.923	0.956
MAE ratio (stocks)									
AAPL		0.834***	0.996	0.995	0.941***	1.537	1.251	1.055	0.986***
PFE		0.969***	0.983**	0.972**	0.992***	1.034	1.010	0.989	0.994**
JNJ		0.756**	0.735**	0.712**	0.730**	0.876***	0.876***	0.790***	0.830***
MSFT		0.941	0.879**	0.899**	0.894*	0.885***	0.854***	0.875***	0.906***

Note. This table reports the estimated coefficients and adjusted R-squared for the HAR-J, HAR-TJ, HAR-J-Vol and HAR-TJ-Vol models. Bold numbers indicate the adjusted R-squared values that are higher for models that include the lag of trading volume in the regression compared to the equivalent models without the lag of trading volume. The MAE ratio panel reports the ratio of the losses from the HAR-J-Vol, HAR-TJ-Vol versus the HAR-J and HAR-TJ models respectively. Ratios smaller than 1 indicate that the HAR-J-Vol and HAR-TJ-Vol models outperform the models without trading volume lag. *, ** and *** highlight the models with trading volume whose losses are significantly lower than the equivalent original model based on the Diebold and Mariano test at the 10%, 5% and 1% levels, respectively.

Table 2.7 Descriptive statistics for SPY's standardised returns (pre-crisis)

	$\frac{r_t}{\sqrt{RV_t}}$	$\frac{r_t}{\sqrt{BV_t}}$	$\frac{r_t}{\sqrt{TRV_t}}$	$\frac{r_t}{\sqrt{QPV_t}}$	$\frac{r_t}{\sqrt{TPV_t}}$	$\frac{r_t}{\sqrt{minRV_t}}$	$\frac{r_t}{\sqrt{medRV_t}}$	$\frac{r_t}{\sqrt{SBV_t}}$	$\frac{r_t}{\sqrt{TBV_t}}$	$\frac{r_t}{\sqrt{CTBV_t}}$	$\frac{r_t}{\sqrt{RVac_t}}$
						30 seconds					
Mean	0.008	0.009	0.013	0.015	0.011	0.008	0.009	0.011	0.015	0.011	0.011
STD. DEV.	0.928	0.972	0.981	1.109	1.033	0.965	0.960	0.985	1.045	0.977	1.007
Median	0.024	0.026	0.027	0.031	0.028	0.025	0.025	0.026	0.027	0.026	0.029
Max	3.171	3.315	3.412	3.677	3.454	3.370	3.367	3.343	3.684	3.341	3.362
Min	-2.851	-2.999	-2.946	-3.235	-3.077	-3.005	-2.953	-2.919	-3.151	-3.041	-3.260
Skewness	0.017	0.024	0.058	0.050	0.036	0.025	0.026	0.037	0.063	0.033	0.039
Kurtosis	2.952	2.911	3.001	2.877	2.880	2.946	2.955	2.921	3.050	2.976	2.863
JB test	0.287	0.855	1.124	2.101	1.632	0.447	0.390	0.971	1.527	0.412	2.082
p-values	0.500	0.500	0.500	0.341	0.427	0.500	0.500	0.500	0.451	0.500	0.344
						60 seconds					
Mean	0.009	0.012	0.016	0.017	0.014	0.012	0.013	0.013	0.018	0.013	0.016
STD. DEV.	0.958	0.992	1.011	1.078	1.032	0.997	0.993	1.003	1.065	1.001	1.004
Median	0.026	0.025	0.028	0.029	0.027	0.026	0.026	0.027	0.029	0.027	0.031
Max	3.095	3.196	3.405	3.476	3.383	3.229	3.312	3.367	3.652	3.276	2.962
Min	-3.015	-3.120	-3.079	-3.299	-3.208	-3.136	-3.039	-3.090	-3.204	-3.140	-2.962
Skewness	0.027	0.045	0.074	0.061	0.055	0.056	0.054	0.046	0.073	0.049	0.055
Kurtosis	2.885	2.885	2.966	2.896	2.892	2.925	2.920	2.907	2.964	2.907	2.759
JB test	1.343	1.796	1.911	2.146	1.996	1.512	1.523	1.427	1.885	1.522	5.860
p-values	0.500	0.394	0.373	0.334	0.359	0.455	0.452	0.478	0.378	0.452	0.052
						150 seconds					
Mean	0.012	0.016	0.020	0.022	0.019	0.018	0.017	0.017	0.024	0.019	0.017
STD. DEV.	0.981	1.018	1.040	1.074	1.044	1.040	1.034	1.024	1.106	1.032	1.037
Median	0.029	0.032	0.031	0.032	0.032	0.032	0.032	0.031	0.034	0.032	0.032
Max	2.976	3.333	3.534	3.582	3.501	3.552	3.340	3.262	3.796	3.337	2.998
Min	-2.911	-2.868	-3.043	-3.043	-2.970	-2.858	-2.923	-3.029	-3.128	-2.879	-2.723
Skewness	0.036	0.069	0.087	0.091	0.082	0.084	0.066	0.058	0.119	0.081	0.067
Kurtosis	2.783	2.833	2.895	2.886	2.851	2.888	2.877	2.832	2.972	2.842	2.584
JB test	4.389	3.918	3.461	3.880	4.112	3.434	2.752	3.479	4.789	4.307	15.959
p-values	0.107	0.136	0.171	0.138	0.123	0.173	0.245	0.169	0.088	0.112	**0.001**
						300 seconds					
Mean	0.014	0.017	0.020	0.021	0.019	0.018	0.018	0.019	0.025	0.019	0.012
STD. DEV.	1.009	1.052	1.071	1.102	1.076	1.086	1.071	1.051	1.157	1.072	1.063
Median	0.031	0.032	0.032	0.032	0.032	0.032	0.032	0.031	0.035	0.032	0.036
Max	3.227	3.270	3.837	3.476	3.366	3.616	3.525	3.162	4.501	3.416	2.953
Min	-2.860	-3.294	-2.938	-3.404	-3.381	-3.423	-3.220	-3.082	-3.315	-3.315	-2.981
Skewness	0.053	0.051	0.083	0.054	0.053	0.054	0.064	0.065	0.117	0.061	0.023
Kurtosis	2.688	2.761	2.818	2.821	2.786	2.856	2.777	2.727	3.036	2.777	2.566
JB test	9.083	5.624	5.076	3.654	4.795	2.723	5.533	7.646	4.705	5.418	15.954
p-values	**0.013**	0.059	0.077	0.155	0.088	0.249	0.061	**0.023**	0.092	0.065	**0.001**

Note: This table reports the descriptive statistics and the results of the Jarque-Bera test for normality for eleven different volatility measures for the SPY index using 30-second, 60-second, 150- second and 300-second returns for the pre-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Table 2.8 Descriptive statistics for SPY's standardised returns (crisis)

	$\frac{r_t}{\sqrt{RV_t}}$	$\frac{r_t}{\sqrt{BV_t}}$	$\frac{r_t}{\sqrt{TRV_t}}$	$\frac{r_t}{\sqrt{QPV_t}}$	$\frac{r_t}{\sqrt{TPV_t}}$	$\frac{r_t}{\sqrt{minRV_t}}$	$\frac{r_t}{\sqrt{medRV_t}}$	$\frac{r_t}{\sqrt{SBV_t}}$	$\frac{r_t}{\sqrt{TBV_t}}$	$\frac{r_t}{\sqrt{CTBV_t}}$	$\frac{r_t}{\sqrt{RVac_t}}$
						30 seconds					
Mean	-0.021	-0.018	-0.015	-0.017	-0.018	-0.017	-0.018	-0.018	-0.016	-0.017	-0.024
STD. DEV.	1.049	1.085	1.101	1.149	1.115	1.098	1.089	1.091	1.158	1.094	1.073
Median	-0.010	-0.010	-0.010	-0.010	-0.010	-0.010	-0.010	-0.010	-0.010	-0.010	-0.009
Max	2.733	2.866	2.968	2.957	2.914	2.938	2.853	2.826	3.011	2.868	2.678
Min	-2.423	-2.406	-2.473	-2.486	-2.426	-2.399	-2.417	-2.434	-2.515	-2.420	-2.601
Skewness	0.065	0.079	0.104	0.095	0.087	0.077	0.076	0.083	0.091	0.082	0.072
Kurtosis	2.622	2.624	2.661	2.666	2.644	2.622	2.614	2.642	2.639	2.624	2.638
JB test	2.509	2.615	2.491	2.331	2.463	2.609	2.710	2.451	2.571	2.651	2.385
p-values	0.253	0.237	0.256	0.281	0.260	0.238	0.225	0.262	0.243	0.233	0.272
						60 seconds					
Mean	-0.023	-0.019	-0.016	-0.017	-0.018	-0.020	-0.021	-0.021	-0.016	-0.018	-0.025
STD. DEV.	1.054	1.078	1.104	1.118	1.098	1.092	1.094	1.089	1.144	1.087	1.091
Median	-0.009	-0.009	-0.010	-0.009	-0.009	-0.009	-0.009	-0.009	-0.010	-0.009	-0.010
Max	2.685	2.800	3.027	2.973	2.924	2.892	2.890	2.890	3.315	2.848	2.768
Min	-2.523	-2.570	-2.523	-2.678	-2.654	-2.617	-2.678	-2.655	-2.616	-2.572	-2.666
Skewness	0.071	0.089	0.113	0.103	0.097	0.091	0.092	0.080	0.118	0.097	0.081
Kurtosis	2.629	2.639	2.663	2.669	2.663	2.652	2.661	2.657	2.677	2.646	2.661
JB test	2.486	2.548	2.575	2.382	2.373	2.432	2.347	2.249	2.513	2.559	2.218
p-values	0.256	0.247	0.243	0.273	0.274	0.265	0.278	0.294	0.252	0.245	0.299
						150 seconds					
Mean	-0.026	-0.022	-0.013	-0.018	-0.020	-0.016	-0.017	-0.018	-0.008	-0.018	-0.032
STD. DEV.	1.076	1.117	1.129	1.156	1.138	1.153	1.140	1.117	1.193	1.131	1.101
Median	-0.010	-0.010	-0.010	-0.010	-0.010	-0.011	-0.010	-0.010	-0.011	-0.010	-0.009
Max	2.682	2.843	3.056	2.988	2.929	3.151	2.862	2.768	3.569	2.865	2.578
Min	-2.632	-2.732	-2.737	-2.977	-2.831	-2.821	-2.827	-2.809	-2.847	-2.763	-2.575
Skewness	0.067	0.090	0.132	0.093	0.096	0.124	0.095	0.084	0.170	0.109	0.016
Kurtosis	2.625	2.679	2.699	2.729	2.708	2.762	2.703	2.675	2.801	2.696	2.546
JB test	2.491	2.125	2.520	1.703	1.914	1.850	1.952	2.112	2.443	2.202	3.256
p-values	0.256	0.314	0.251	0.390	0.350	0.362	0.343	0.316	0.263	0.301	0.167
						300 seconds					
Mean	-0.026	-0.027	-0.023	-0.023	-0.024	-0.028	-0.025	-0.019	-0.022	-0.028	-0.036
STD. DEV.	1.091	1.135	1.142	1.171	1.155	1.174	1.164	1.135	1.218	1.151	1.114
Median	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.011	-0.010	-0.010
Max	2.598	2.776	3.218	3.140	2.903	2.879	2.959	2.874	3.613	2.879	2.733
Min	-2.604	-2.871	-2.604	-3.008	-2.972	-2.777	-2.650	-2.620	-2.889	-2.889	-2.519
Skewness	0.047	0.033	0.084	0.054	0.052	0.043	0.069	0.097	0.082	0.035	-0.015
Kurtosis	2.601	2.639	2.707	2.736	2.694	2.688	2.710	2.697	2.781	2.655	2.434
JB test	2.644	2.117	1.797	1.279	1.643	1.645	1.616	2.037	1.182	1.954	5.056
p-values	0.234	0.315	0.372	0.500	0.403	0.402	0.408	0.329	0.500	0.343	0.069

Note: This table reports the descriptive statistics and the results of the Jarque-Bera test for normality for eleven different volatility measures for the SPY index using 30-second, 60-second, 150- second and 300-second returns for the crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Table 2.9 Descriptive statistics for SPY's standardised returns (post-crisis)

	$\frac{r_t}{\sqrt{RV_t}}$	$\frac{r_t}{\sqrt{BV_t}}$	$\frac{r_t}{\sqrt{TRV_t}}$	$\frac{r_t}{\sqrt{QPV_t}}$	$\frac{r_t}{\sqrt{TPV_t}}$	$\frac{r_t}{\sqrt{minRV_t}}$	$\frac{r_t}{\sqrt{medRV_t}}$	$\frac{r_t}{\sqrt{SBV_t}}$	$\frac{r_t}{\sqrt{TBV_t}}$	$\frac{r_t}{\sqrt{CTBV_t}}$	$\frac{r_t}{\sqrt{RVac_t}}$
						30 seconds					
Mean	0.120	0.128	0.139	0.146	0.137	0.129	0.129	0.131	0.144	0.132	0.133
STD. DEV.	0.965	1.007	1.040	1.089	1.046	1.018	1.012	1.016	1.094	1.021	0.993
Median	0.121	0.128	0.132	0.137	0.134	0.129	0.129	0.128	0.139	0.130	0.127
Max	3.305	3.385	3.611	3.687	3.509	3.425	3.384	3.495	3.822	3.487	3.502
Min	-3.076	-3.270	-3.515	-3.802	-3.414	-3.298	-3.204	-3.330	-3.675	-3.276	-3.467
Skewness	-0.067	-0.053	-0.013	-0.029	-0.043	-0.049	-0.049	-0.051	-0.020	-0.044	-0.028
Kurtosis	3.006	2.995	3.022	3.021	3.000	2.988	2.976	3.005	3.044	3.005	2.976
JB test	1.417	0.886	0.093	0.303	0.575	0.765	0.813	0.820	0.274	0.623	0.293
p-value	0.481	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
						60 seconds					
Mean	0.124	0.132	0.143	0.147	0.139	0.134	0.134	0.135	0.146	0.135	0.138
STD. DEV.	0.966	0.999	1.036	1.056	1.027	1.014	1.014	1.010	1.079	1.012	0.999
Median	0.123	0.125	0.133	0.134	0.130	0.127	0.125	0.127	0.137	0.128	0.133
Max	3.322	3.410	3.468	3.582	3.508	3.476	3.605	3.534	3.771	3.437	3.300
Min	-3.113	-3.370	-3.670	-4.103	-3.762	-3.526	-3.430	-3.443	-4.028	-3.425	-3.018
Skewness	-0.052	-0.042	-0.004	-0.013	-0.027	-0.039	-0.034	-0.030	-0.016	-0.031	-0.024
Kurtosis	2.981	2.998	3.011	3.039	3.015	3.026	3.030	3.003	3.074	3.004	2.861
JB test	0.871	0.561	0.016	0.174	0.242	0.546	0.439	0.293	0.516	0.312	1.709
p-value	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.410
						150 seconds					
Mean	0.134	0.144	0.156	0.158	0.151	0.149	0.147	0.146	0.164	0.148	0.140
STD. DEV.	0.992	1.037	1.068	1.086	1.063	1.067	1.057	1.040	1.130	1.053	1.016
Median	0.131	0.133	0.139	0.139	0.136	0.136	0.136	0.137	0.142	0.135	0.138
Max	3.256	3.648	3.551	3.906	3.823	3.977	3.786	3.592	3.916	3.707	3.137
Min	-3.135	-3.164	-3.194	-3.182	-3.198	-3.255	-3.203	-3.160	-3.224	-3.175	-2.946
Skewness	-0.038	-0.001	0.040	0.038	0.020	0.015	0.006	0.000	0.045	0.012	-0.043
Kurtosis	2.900	2.950	2.946	2.981	2.968	2.998	2.973	2.925	3.006	2.949	2.744
JB test	1.237	0.196	0.728	0.475	0.207	0.067	0.069	0.444	0.639	0.247	5.732
p-value	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.056
						300 seconds					
Mean	0.138	0.148	0.159	0.162	0.155	0.155	0.153	0.151	0.171	0.152	0.141
STD. DEV.	1.001	1.047	1.076	1.095	1.070	1.083	1.070	1.047	1.150	1.066	1.028
Median	0.132	0.136	0.140	0.143	0.143	0.140	0.136	0.136	0.146	0.139	0.140
Max	2.967	3.210	3.761	3.555	3.382	3.699	3.435	3.185	4.316	3.235	3.717
Min	-2.960	-3.048	-3.894	-3.077	-2.991	-3.395	-2.985	-3.080	-3.487	-3.051	-2.706
Skewness	-0.047	-0.003	0.037	0.038	0.018	0.032	0.023	0.002	0.069	0.010	-0.052
Kurtosis	2.793	2.849	2.929	2.892	2.861	2.956	2.840	2.803	2.981	2.852	2.700
JB test	4.077	1.805	0.838	1.367	1.633	0.480	2.179	3.063	1.517	1.749	7.922
p-value	0.125	0.392	0.500	0.498	0.426	0.500	0.328	0.209	0.453	0.402	**0.021**

Note: This table reports the descriptive statistics and the results of the Jarque-Bera test for normality for eleven different volatility measures for the SPY index using 30-second, 60-second, 150- second and 300-second returns for the post-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Table 2.10 Leverage effects estimated from the EGARCH (1,1) model (pre-crisis)

	Return (r_t)	$\frac{r_t}{\sqrt{minRV_t}}$			
		30 sec	***60 sec***	***150 sec***	***300 sec***
		AAPL (IT)			
Leverage	-0.010	0.011	0.043	0.041	0.051
S.E.	0.010	0.007	0.034	0.033	0.034
T-statistics	-1.054	1.665	1.260	1.244	1.482
p-value	0.292	0.096	0.208	0.214	0.138
		JNJ (HC)			
Leverage	-0.063	-0.052	-0.052	-0.024	-0.020
S.E.	0.011	0.020	0.022	0.013	0.013
T-statistics	-5.631	-2.581	-2.371	-1.871	-1.553
p-value	0.000	0.010	0.018	0.061	0.121
		PFE (HC)			
Leverage	-0.021	-0.013	-0.011	0.034	-0.023
S.E.	0.004	0.005	0.005	0.025	0.031
T-statistics	-4.857	-2.671	-2.096	1.369	-0.737
p-value	0.000	0.008	0.036	0.171	0.461
		MSFT (IT)			
Leverage	-0.044	0.005	0.025	0.005	0.004
S.E.	0.010	0.005	0.026	0.019	0.004
T-statistics	-4.303	0.951	0.962	0.277	1.044
p-value	0.000	0.342	0.336	0.782	0.296
		SPY			
Leverage	-0.096	-0.103	-0.094	-0.086	-0.077
S.E.	0.009	0.033	0.034	0.032	0.033
T-statistics	-10.200	-3.096	-2.785	-2.645	-2.302
p-value	0.000	0.002	0.005	0.008	0.021
Mean	-0.047	-0.030	-0.018	-0.006	-0.013

Note: This table shows the leverage effects estimated from the EGARCH (1,1) model for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the pre-crisis period using 30-second, 60-second, 150-second and 300-second returns.

Table 2.11 Leverage effects estimated from the EGARCH (1,1) model (crisis)

	Return (r_t)	$\frac{r_t}{\sqrt{minRV_t}}$			
		30 sec	*60 sec*	*150 sec*	*300 sec*
		AAPL (IT)			
Leverage	-0.143	0.100	0.011	0.095	0.064
STD. DEV.	0.030	0.072	0.020	0.073	0.068
T-statistics	-4.778	1.398	0.541	1.298	0.952
p-value	0.000	0.162	0.589	0.194	0.341
		JNJ (HC)			
Leverage	-0.178	-0.124	-0.096	-0.091	-0.088
STD. DEV.	0.034	0.066	0.066	0.071	0.065
T-statistics	-5.242	-1.866	-1.454	-1.277	-1.359
p-value	0.000	0.062	0.146	0.202	0.174
		PFE (HC)			
Leverage	-0.040	0.196	0.185	0.167	0.150
STD. DEV.	0.030	0.067	0.068	0.069	0.068
T-statistics	-1.321	2.923	2.699	2.423	2.204
p-value	**0.187**	0.003	0.007	0.015	0.028
		MSFT (IT)			
Leverage	-0.139	-	-0.014	0.092	0.024
STD. DEV.	0.030	-	0.067	0.081	0.078
T-statistics	-4.696	-	-0.206	1.137	0.305
p-value	0.000	-	0.837	0.255	0.760
		SPY			
Leverage	-0.148	0.023	0.045	0.015	0.016
STD. DEV.	0.031	0.076	0.073	0.071	0.076
T-statistics	-4.837	0.308	0.612	0.208	0.206
p-value	0.000	0.758	0.541	0.835	0.837
Mean	-0.130	0.049	0.026	0.056	0.033

Note: This table shows the leverage effects estimated from the EGARCH (1,1) model for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the crisis period using 30-second, 60-second, 150-second and 300-second returns.

Table 2.12 Leverage effects estimated from the EGARCH (1,1) model (post-crisis)

	Return (r_t)	$\frac{r_t}{\sqrt{minRV_t}}$			
		30 sec	***60 sec***	***150 sec***	***300 sec***
		AAPL (IT)			
Leverage	-0.111	-0.015	-0.020	-0.026	-0.018
STD. DEV.	0.019	0.029	0.010	0.012	0.011
T-statistics	-5.939	-0.513	-1.972	-2.171	-1.699
p-value	0.000	0.608	0.049	0.030	0.089
		JNJ (HC)			
Leverage	-0.063	-0.013	-0.012	-0.010	-0.027
STD. DEV.	0.012	0.017	0.018	0.012	0.018
T-statistics	-5.103	-0.754	-0.668	-0.802	-1.485
p-value	0.000	0.451	0.504	0.422	0.137
		PFE (HC)			
Leverage	-0.065	-0.037	-0.038	-0.025	-0.043
STD. DEV.	0.012	0.029	0.028	0.029	0.031
T-statistics	-5.229	-1.271	-1.342	-0.860	-1.405
p-value	0.000	0.204	0.180	0.390	0.160
		MSFT (IT)			
Leverage	-0.038	-0.057	-0.052	-0.046	-0.059
STD. DEV.	0.014	0.032	0.032	0.033	0.034
T-statistics	-2.648	-1.779	-1.599	-1.363	-1.728
p-value	0.008	0.075	0.110	0.173	0.084
		SPY			
Leverage	-0.146	-0.067	-0.065	-0.061	-0.062
STD. DEV.	0.014	0.022	0.022	0.034	0.034
T-statistics	-10.270	-2.988	-2.941	-1.775	-1.826
p-value	0.000	0.003	0.003	0.076	0.068
Mean	-0.085	-0.038	-0.037	-0.034	-0.042

Note: This table shows the leverage effects estimated from the EGARCH (1,1) model for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the post-crisis period using 30-second, 60-second, 150-second and 300-second returns.

Table 2.13 Correlations between trading volume and volatility measures (pre-crisis)

Sampling Frequency (Seconds)		RV_t	BV_t	TRV_t	QPV_t	TPV_t	$minRV_t$	$medRV_t$	SBV_t	TBV_t	$CTBV_t$	$RVac_t$
							AAPL (IT)					
30	***Estimate***	-0.045	-0.016	-0.032	0.086	0.028	-0.022	-0.031	-0.008	-0.013	-0.039	-0.013
	p-value	0.042	0.467	0.146	0.000	0.202	0.314	0.166	0.728	0.562	0.082	0.575
60	***Estimate***	-0.032	-0.014	-0.033	0.045	0.013	-0.018	-0.021	-0.003	-0.012	-0.028	0.029
	p-value	0.149	0.530	0.139	0.044	0.555	0.422	0.337	0.905	0.605	0.210	0.187
150	***Estimate***	-0.004	0.007	-0.011	0.042	0.025	0.000	0.007	0.022	-0.008	-0.013	0.070
	p-value	0.873	0.753	0.621	0.059	0.265	0.994	0.755	0.332	0.709	0.563	0.002
300	***Estimate***	0.030	0.052	0.022	0.080	0.069	0.048	0.041	0.051	0.012	0.021	0.093
	p-value	0.172	0.020	0.316	0.000	0.002	0.030	0.063	0.022	0.578	0.355	0.000
							JNJ (HC)					
30	***Estimate***	0.279	0.289	0.238	0.328	0.310	0.260	0.255	0.286	0.201	0.215	0.325
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.300	0.304	0.258	0.319	0.312	0.288	0.287	0.299	0.225	0.243	0.353
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.334	0.325	0.297	0.321	0.325	0.314	0.314	0.321	0.246	0.269	0.370
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.361	0.349	0.333	0.346	0.345	0.340	0.346	0.342	0.257	0.293	0.375
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							PFE (HC)					
30	***Estimate***	-0.018	-0.001	-0.168	0.093	0.035	-0.041	-0.043	-0.047	-0.164	-0.154	0.161
	p-value	0.422	0.971	0.000	0.000	0.116	0.064	0.052	0.037	0.000	0.000	0.000
60	***Estimate***	-0.015	-0.007	-0.128	0.040	0.007	-0.059	-0.041	-0.014	-0.117	-0.097	0.123
	p-value	0.505	0.770	0.000	0.072	0.770	0.008	0.066	0.525	0.000	0.000	0.000
150	***Estimate***	0.072	0.112	0.009	0.131	0.132	0.109	0.094	0.081	-0.091	-0.033	0.127
	p-value	0.001	0.000	0.688	0.000	0.000	0.000	0.000	0.000	0.000	0.144	0.000
300	***Estimate***	0.119	0.125	0.053	0.162	0.138	0.069	0.050	0.097	-0.048	-0.019	0.196
	p-value	0.000	0.000	0.018	0.000	0.000	0.002	0.026	0.000	0.032	0.407	0.000
							MSFT (IT)					
30	***Estimate***	0.424	0.426	0.409	0.414	0.424	0.414	0.418	0.424	0.410	0.417	0.430
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.428	0.427	0.418	0.412	0.421	0.421	0.423	0.423	0.418	0.425	0.432
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.434	0.430	0.422	0.423	0.425	0.428	0.429	0.429	0.425	0.433	0.442
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.436	0.433	0.424	0.424	0.429	0.420	0.430	0.431	0.422	0.440	0.422
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							SPY					
30	***Estimate***	-0.139	-0.105	-0.163	-0.055	-0.079	-0.111	-0.128	-0.117	-0.184	-0.173	-0.052
	p-value	0.000	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.020
60	***Estimate***	-0.107	-0.099	-0.121	-0.068	-0.084	-0.107	-0.114	-0.101	-0.132	-0.126	-0.037
	p-value	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.098
150	***Estimate***	-0.068	-0.071	-0.077	-0.064	-0.068	-0.075	-0.075	-0.071	-0.093	-0.089	0.000
	p-value	0.002	0.001	0.001	0.004	0.002	0.001	0.001	0.002	0.000	0.000	0.989
300	***Estimate***	-0.029	-0.028	-0.027	-0.022	-0.025	-0.030	-0.029	-0.025	-0.046	-0.040	0.020
	p-value	0.199	0.204	0.218	0.319	0.267	0.185	0.194	0.256	0.040	0.069	0.368
Mean	***Estimate***	0.131	0.140	0.101	0.165	0.151	0.126	0.124	0.134	0.081	0.093	0.184

Note: This table shows the correlation coefficients (estimates) and their significance levels (p-values) for different volatility measures and trading volumes for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the pre-crisis period using 30-second, 60-second, 150-second and 300-second sampling data.

Table 2.14 Correlations between trading volume and volatility measures (crisis)

Sampling Frequency (Seconds)		RV_t	BV_t	TRV_t	QPV_t	TPV_t	$minRV_t$	$medRV_t$	SBV_t	TBV_t	$CTBV_t$	$RVac_t$
							AAPL (IT)					
30	***Estimate***	0.669	0.675	0.701	0.680	0.678	0.676	0.678	0.676	0.788	0.780	0.667
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.677	0.693	0.707	0.700	0.699	0.694	0.697	0.699	0.795	0.791	0.706
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.678	0.673	0.718	0.686	0.686	0.672	0.674	0.698	0.772	0.781	0.688
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.681	0.659	0.684	0.677	0.664	0.633	0.652	0.690	0.771	0.778	0.667
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							JNJ (HC)					
30	***Estimate***	0.656	0.656	0.652	0.655	0.658	0.650	0.652	0.659	0.732	0.749	0.616
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.628	0.662	0.667	0.655	0.662	0.674	0.652	0.634	0.747	0.756	0.610
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.633	0.642	0.669	0.655	0.654	0.628	0.601	0.644	0.740	0.761	0.685
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.682	0.682	0.659	0.657	0.669	0.676	0.674	0.660	0.720	0.737	0.675
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							PFE (HC)					
30	***Estimate***	0.549	0.547	0.562	0.527	0.539	0.547	0.550	0.559	0.594	0.605	0.559
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.549	0.550	0.561	0.532	0.540	0.550	0.557	0.545	0.609	0.617	0.552
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.527	0.522	0.557	0.505	0.512	0.510	0.510	0.522	0.598	0.613	0.522
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.547	0.550	0.542	0.551	0.552	0.538	0.545	0.551	0.572	0.598	0.508
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							MSFT (IT)					
30	***Estimate***	0.591	0.593	0.583	0.595	0.594	0.590	0.587	0.590	0.592	0.600	0.596
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.600	0.593	0.592	0.586	0.590	0.590	0.592	0.593	0.606	0.610	0.612
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.601	0.605	0.602	0.602	0.604	0.603	0.604	0.601	0.611	0.619	0.595
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.602	0.601	0.601	0.598	0.600	0.596	0.603	0.606	0.603	0.612	0.580
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							SPY					
30	***Estimate***	0.789	0.784	0.788	0.777	0.782	0.780	0.784	0.785	0.888	0.890	0.760
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.773	0.766	0.785	0.748	0.752	0.768	0.767	0.751	0.887	0.890	0.777
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.775	0.744	0.792	0.727	0.731	0.720	0.737	0.752	0.887	0.890	0.771
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.798	0.800	0.810	0.793	0.799	0.787	0.798	0.795	0.884	0.891	0.789
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mean	***Estimate***	0.619	0.619	**0.630**	0.615	0.617	0.613	0.615	0.620	**0.685**	**0.694**	0.616

Note: This table shows the correlation coefficients (estimates) and their *p*-values for different volatility measures and trading volumes for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the crisis period using 30-second, 60-second, 150-second and 300-second sampling data. Top three average estimate values are in bold.

Table 2.15 Correlations between trading volume and volatility measures (post-crisis)

Sampling Frequency (Seconds)		RV_t	BV_t	TRV_t	QPV_t	TPV_t	$minRV_t$	$medRV_t$	SBV_t	TBV_t	$CTBV_t$	$RVac_t$
							AAPL (IT)					
30	***Estimate***	0.325	0.282	0.541	0.276	0.273	0.274	0.289	0.307	0.635	0.623	0.234
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.332	0.356	0.535	0.361	0.344	0.414	0.356	0.337	0.628	0.632	0.283
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.438	0.393	0.515	0.379	0.377	0.361	0.413	0.417	0.629	0.632	0.379
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.458	0.457	0.498	0.448	0.449	0.425	0.459	0.450	0.611	0.617	0.399
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							JNJ (HC)					
30	***Estimate***	0.202	0.218	0.466	0.413	0.321	0.220	0.221	0.365	0.472	0.484	0.377
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.214	0.254	0.480	0.327	0.308	0.284	0.272	0.314	0.501	0.503	0.279
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.190	0.182	0.481	0.293	0.237	0.162	0.167	0.317	0.514	0.513	0.337
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.176	0.191	0.497	0.434	0.404	0.152	0.200	0.449	0.507	0.519	0.433
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							PFE (HC)					
30	***Estimate***	0.410	0.431	0.660	0.473	0.473	0.474	0.489	0.539	0.673	0.687	0.503
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.566	0.589	0.652	0.546	0.566	0.605	0.584	0.556	0.659	0.670	0.490
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.599	0.598	0.644	0.581	0.589	0.588	0.598	0.597	0.646	0.666	0.575
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.567	0.548	0.622	0.562	0.559	0.523	0.532	0.590	0.633	0.657	0.579
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							MSFT (IT)					
30	***Estimate***	0.548	0.533	0.550	0.494	0.514	0.528	0.528	0.527	0.578	0.599	0.496
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.505	0.475	0.505	0.457	0.471	0.463	0.483	0.496	0.564	0.588	0.508
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.517	0.506	0.511	0.469	0.483	0.505	0.499	0.479	0.572	0.597	0.499
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.498	0.503	0.500	0.486	0.488	0.507	0.482	0.472	0.551	0.573	0.502
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							SPY					
30	***Estimate***	0.725	0.664	0.817	0.638	0.647	0.635	0.694	0.699	0.877	0.878	0.736
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60	***Estimate***	0.755	0.740	0.801	0.711	0.729	0.741	0.751	0.741	0.878	0.884	0.748
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	***Estimate***	0.756	0.762	0.785	0.746	0.745	0.748	0.751	0.733	0.876	0.883	0.691
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	***Estimate***	0.688	0.656	0.731	0.704	0.677	0.596	0.641	0.719	0.870	0.879	0.726
	p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mean	***Estimate***	0.451	0.445	**0.561**	0.467	0.460	0.438	0.448	0.481	**0.613**	**0.623**	0.466

Note: This table shows the correlation coefficients (estimates) and their *p*-values for different volatility measures and trading volumes for four stocks (AAPL, JNJ, PFE and MSFT) and the SPY index for the post-crisis period using 30-second, 60-second, 150-second and 300-second sampling data. Top three average estimate values are in bold.

Table 2.16: Estimation and out-of-sample forecast losses (mean absolute errors) for the pre-crisis period (forecast horizon h=1)

		HAR-J				HAR-TJ			
		30	60	150	300	30	60	150	300
	β_0	0.131***	0.099***	0.104***	0.135***	0.166***	0.134***	0.132***	0.131***
	s.e.	0.046	0.031	0.035	0.042	0.044	0.029	0.032	0.039
	β_d	0.209***	0.466***	0.277***	0.257***	0.690***	0.676***	0.551***	0.558***
	s.e.	0.028	0.030	0.028	0.029	0.052	0.035	0.037	0.040
	β_w	0.321***	0.250***	0.358***	0.258***	0.143***	0.125***	0.173***	0.082
	s.e.	0.053	0.048	0.048	0.052	0.071	0.053	0.056	0.064
	β_m	0.398***	0.230***	0.275***	0.374***	0.226***	0.207***	0.259***	0.364***
	s.e.	0.049	0.042	0.043	0.051	0.059	0.043	0.047	0.057
	β_j	-0.424	-0.790***	0.349	-0.255*	-0.468***	-0.223***	0.036	-0.006
	s.e.	0.295	0.133	0.293	0.138	0.063	0.047	0.035	0.032
	$\bar{R}^2$	0.480	0.625	0.517	0.374	0.526	0.655	0.555	0.425
		HAR-J-Vol				HAR-TJ-Vol			
SPY	β_0	0.155*	0.167***	0.143***	0.182***	0.223***	0.202***	0.185***	0.232***
	s.e.	0.085	0.055	0.059	0.071	0.081	0.052	0.056	0.066
	β_d	0.210***	0.470***	0.279***	0.259***	0.694***	0.681***	0.555***	0.565***
	s.e.	0.028	0.031	0.029	0.029	0.052	0.036	0.037	0.040
	β_w	0.322***	0.252***	0.360***	0.259***	0.146***	0.128***	0.175***	0.084
	s.e.	0.053	0.048	0.048	0.052	0.071	0.053	0.056	0.064
	β_m	0.392***	0.210***	0.264***	0.362***	0.209***	0.186***	0.243***	0.334***
	s.e.	0.052	0.044	0.045	0.053	0.062	0.046	0.049	0.059
	β_j	-0.423	-0.800***	0.352	-0.255***	-0.470***	-0.223***	0.037	-0.006
	s.e.	0.295	0.133	0.293	0.138	0.063	0.047	0.035	0.032
	β_v	-0.001	-0.002	-0.001	-0.001	-0.001	-0.002	-0.001	-0.002*
	s.e.	0.002	0.001	0.001	0.001	0.002	0.001	0.001	0.001
	$\bar{R}^2$	0.480	0.625	0.517	0.374	0.526	0.655	0.555	**0.426**
	MAE ratio	1.012	0.972	0.994	0.997	*0.943*	*0.927*	*0.930*	*0.906*
MAE ratio (Stocks)									
AAPL		0.971	0.982	0.982*	0.990***	0.920***	0.950***	0.961***	0.978***
PFE		0.888*	0.893	0.928	0.930	0.884***	0.842***	0.882***	0.832***
JNJ		0.995**	1.013	1.055	1.031	1.011	1.028	1.054	1.038
MSFT		0.997	0.991	0.990	0.988	1.001	0.991	0.983	0.986

Note. This table reports the estimated coefficients and adjusted R-squared for the HAR-J, HAR-TJ, HAR-J-Vol and HAR-TJ-Vol models. Bold numbers indicate the adjusted R-squared values that are higher for models that include the lag of trading volume in the regression compared to the equivalent models without the lag of trading volume. The MAE ratio panel reports the ratio of the losses from the HAR-J-Vol, HAR-TJ-Vol versus the HAR-J and HAR-TJ models respectively. Ratios smaller than 1 indicate that the HAR-J-Vol and HAR-TJ-Vol models outperform the models without trading volume lag. *, ** and *** highlight the models with trading volume whose losses are significantly lower than the equivalent original model based on the Diebold and Mariano test at the 10%, 5% and 1% levels, respectively.

Table 2.17: Estimation and out-of-sample forecast losses (mean absolute errors) for the crisis period (forecast horizon h=1)

			HAR-J				HAR-TJ		
		30	60	150	300	30	60	150	300
	β_0	0.297**	0.282**	0.285**	0.272**	0.300**	0.263**	0.298**	0.203*
	s.e.	0.122	0.120	0.124	0.113	0.120	0.116	0.120	0.117
	β_d	0.665***	0.651***	0.584***	0.382***	0.691***	0.750***	0.592***	0.381***
	s.e.	0.085	0.087	0.090	0.098	0.100	0.116	0.092	0.129
	β_w	0.109	0.144	0.219*	0.411***	0.125	0.143	0.217	0.556***
	s.e.	0.120	0.123	0.131	0.136	0.133	0.147	0.144	0.184
	β_m	0.017	-0.004	-0.024	-0.046	0.003	-0.011	-0.023	-0.066
	s.e.	0.103	0.100	0.104	0.097	0.113	0.109	0.110	0.115
	β_j	-0.071	0.000	0.296	1.255***	0.480	0.310	0.514**	0.621***
	s.e.	0.804	0.768	0.399	0.245	0.365	0.229	0.207	0.116
	$\bar{R}^2$	0.528	0.532	0.482	0.500	0.526	0.540	0.480	0.486
SPY			HAR-J-Vol				HAR-TJ-Vol		
	β_0	0.154	0.189	0.180	-0.041	0.171	0.230	0.174	-0.022
	s.e.	0.214	0.217	0.225	0.195	0.214	0.216	0.221	0.201
	β_d	0.614***	0.616***	0.547***	0.225*	0.644***	0.734***	0.548***	0.244
	s.e.	0.105	0.111	0.112	0.126	0.119	0.144	0.113	0.163
	β_w	0.090	0.133	0.201	0.364***	0.106	0.140	0.194	0.522***
	s.e.	0.122	0.126	0.135	0.137	0.135	0.149	0.148	0.185
	β_m	0.046	0.013	-0.007	0.014	0.031	-0.005	-0.001	-0.013
	s.e.	0.109	0.106	0.109	0.101	0.119	0.115	0.115	0.121
	β_j	-0.145	-0.001	0.292	1.230***	0.444	0.306	0.476**	0.564***
	s.e.	0.810	0.770	0.400	0.243	0.369	0.231	0.215	0.123
	β_v	0.001	0.001	0.001	0.002*	0.001	0.000	0.001	0.002
	s.e.	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	$\bar{R}^2$	0.527	0.530	0.480	**0.510**	0.525	0.537	0.478	**0.489**
	MAE ratio	0.999	0.997	0.990	1.019	0.974	0.974	1.014	0.995
MAE ratio (Stocks)									
AAPL		0.969	0.992	0.957	0.957	0.942	0.958	0.958	0.947
PFE		1.046	1.041	1.038	1.017	1.027	1.019	1.030	1.014
JNJ		1.026	1.029	1.012	0.992	0.988	1.071	0.957	0.938
MSFT		0.975	0.986	1.012	1.001	0.965	0.965	0.993	0.970

Note. This table reports the estimated coefficients and adjusted R-squared for the HAR-J, HAR-TJ, HAR-J-Vol and HAR-TJ-Vol models. Bold numbers indicate the adjusted R-squared values that are higher for models that include the lag of trading volume in the regression compared to the equivalent models without the lag of trading volume. The MAE ratio panel reports the ratio of the losses from the HAR-J-Vol, HAR-TJ-Vol versus the HAR-J and HAR-TJ models respectively. Ratios smaller than 1 indicate that the HAR-J-Vol and HAR-TJ-Vol models outperform the models without trading volume lag. *, ** and *** highlight the models with trading volume whose losses are significantly lower than the equivalent original model based on the Diebold and Mariano test at the 10%, 5% and 1% levels, respectively.

Table 2.18: Estimation and out-of-sample forecast losses (mean absolute errors) for the post-crisis period (forecast horizon h=1)

			HAR-J				HAR-TJ		
		30	60	150	300	30	60	150	300
	β_0	0.132***	0.115***	0.111***	0.143***	0.134***	0.126***	0.100***	0.137***
	s.e.	0.030	0.027	0.027	0.033	0.031	0.026	0.027	0.032
	β_d	0.451***	0.509***	0.506***	0.396***	0.457***	0.646***	0.654***	0.754***
	s.e.	0.027	0.027	0.027	0.028	0.028	0.036	0.044	0.046
	β_w	0.099**	0.084**	0.095**	0.120**	0.079*	0.037	0.097	-0.032
	s.e.	0.046	0.043	0.043	0.047	0.047	0.051	0.059	0.063
	β_m	0.277***	0.245***	0.237***	0.261***	0.301***	0.221***	0.190***	0.177***
	s.e.	0.050	0.046	0.047	0.055	0.055	0.050	0.056	0.062
	β_j	-0.120	-0.144	-0.064	-0.084	0.514**	0.153***	0.266***	0.028
	s.e.	0.431	0.395	0.294	0.408	0.201	0.059	0.049	0.042
	$\bar{R}^2$	0.360	0.416	0.408	0.289	0.357	0.432	0.422	0.339
			HAR-J-Vol				HAR-TJ-Vol		
SPY	β_0	-0.264***	-0.233***	-0.250***	-0.345***	-0.291***	-0.177***	-0.212***	-0.198***
	s.e.	0.050	0.047	0.047	0.056	0.050	0.048	0.048	0.058
	β_d	0.300***	0.359***	0.350***	0.248***	0.306***	0.479***	0.454***	0.561***
	s.e.	0.031	0.031	0.031	0.030	0.031	0.042	0.050	0.053
	β_w	0.028	0.030	0.043	0.034	0.010	0.017	0.078	-0.045
	s.e.	0.045	0.043	0.043	0.046	0.046	0.050	0.058	0.063
	β_m	0.167***	0.160***	0.139***	0.111***	0.207***	0.148***	0.107***	0.088
	s.e.	0.050	0.046	0.047	0.055	0.054	0.050	0.056	0.063
	β_j	-0.925**	-0.643*	-0.358	-0.539	-0.553**	0.104*	0.207***	0.008
	s.e.	0.426	0.390	0.289	0.397	0.219	0.058	0.049	0.042
	β_v	0.005***	0.004***	0.004***	0.006***	0.005***	0.004***	0.004***	0.004***
	s.e.	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.001
	$\bar{R}^2$	**0.395**	**0.442**	**0.437**	**0.333**	**0.398**	**0.451**	**0.442**	**0.357**
	MAE ratio	0.897	0.928	0.950	1.015	0.791**	0.793	0.850	0.845
MAE ratio (Stocks)									
AAPL		0.511***	0.530***	0.695***	0.773***	0.576***	0.599***	0.658***	0.734***
PFE		0.838**	0.919*	0.929**	0.895***	0.950***	0.920***	0.919***	0.927***
JNJ		0.666**	0.678**	0.650**	0.643**	0.900*	0.909***	0.923***	0.856***
MSFT		0.984**	0.932***	0.962**	0.944**	0.948***	0.944***	0.961***	0.977***

Note. This table reports the estimated coefficients and adjusted R-squared for the HAR-J, HAR-TJ, HAR-J-Vol and HAR-TJ-Vol models. Bold numbers indicate the adjusted R-squared values that are higher for models that include the lag of trading volume in the regression compared to the equivalent models without the lag of trading volume. The MAE ratio panel reports the ratio of the losses from the HAR-J-Vol, HAR-TJ-Vol versus the HAR-J and HAR-TJ models respectively. Ratios smaller than 1 indicate that the HAR-J-Vol and HAR-TJ-Vol models outperform the models without trading volume lag. *, ** and *** highlight the models with trading volume whose losses are significantly lower than the equivalent original model based on the Diebold and Mariano test at the 10%, 5% and 1% levels, respectively.

Chapter 3 – The impact of intraday periodicity on stock volatility components and forecasting using different sampling schemes

3.1 Introduction

The importance of intraday periodicity in estimating and forecasting volatility is widely recognised. Andersen and Bollerslev (1998) and Andersen et al. (2003, 2007a) find that the announcement of news has a significant impact on financial market prices, while Bollerslev et al. (2008) suggest that the peak of realised variation is caused by scheduled announcements of news. In addition, intraday periodicity has an impact on detecting jumps and co-jumps (Boudt et al., 2011b; Aït-Sahalia & Xiu, 2016). Boudt et al. (2011b) find that detecting jump components without considering intraday periodicity is likely to produce biased results, and Aït-Sahalia and Xiu (2016) suggest that jumps and co-jumps have intraday patterns and that a large proportion of jumps can be predicted. In this chapter, we examine the impact of intraday periodicity on the frequency of intraday jumps, the proportion of volatility components and on volatility forecasting. We do this by studying the effects of intraday periodicity on the volatility of stocks and the SPY index in different financial regimes (the pre-crisis, crisis and post-crisis periods) and using different sampling schemes (calendar-time and business-time sampling schemes).

Non-parametric volatility measures, which are generated from high-frequency data, are able to separate the continuous and jump components of volatility. Filtering by intraday periodicity within these measures makes it possible to examine the volatility components in the absence or presence of intraday periodicity. The non-parametric volatility measures realised volatility (RV), realised bi-power variation (BV), threshold bi-power variation (TBV) and corrected threshold bi-power variation (CTBV) are constructed from the intraday returns of four stocks and the SPY index, which are sampled at 30, 60, 150 and 300 seconds based on calendar-time sampling schemes. The 300-second data are compared with equivalent data using business-time sampling, which has yet to be studied in previous literature. Non-parametric volatility measures play an important role in HAR class volatility forecasting models, thus allowing us to consider the impact of intraday periodicity on volatility forecasting.

This chapter contains six further sections. The data are described in Section 3.2 and the methodology is explained Section 3.3. The empirical results are discussed in Sections 3.4 to 3.6, beginning with the impact of intraday periodicity on the stylised facts of returns and volatility measures. Sections 3.5 and 3.6 present the effect of intraday periodicity on jump frequency, the proportion of volatility components and volatility forecasting in different regimes based on data from calendar-time and business-time sampling schemes respectively. Section 3.7 provides a conclusion and suggestions for future research.

3.2 Data

The empirical analysis in this research is carried out using high frequency stocks. We consider two active stocks from the highly volatile IT sector, MSFT and AAPL, as stocks in this sector have grown dramatically in recent years due to the rise of cloud computing, big data and mobile computing. We also include two stocks from a less active sector, namely healthcare (JNJ and PFE), because companies from this sector have a stable demand and are less sensitive to economic cycles. The SPY index is also considered in our analysis as it reflects the general trends of the stock market. The sampling period ranges from 2000 to 2016, which can be separated into three different regimes: the period before the 2008 financial crisis (01/01/2000 to 30/12/2007), the period during the crisis (01/01/2008 to 30/06/2009) and the period after the crisis (01/07/2009 to 30/12/2016). The trading day is 9:30am-4:00pm. We sample prices at tick level down to 30, 60, 150 and 300 seconds.

3.3 Methodology

High-frequency data are usually described using the Brownian Semi-Martingale with Finite Activity Jumps (BSMFAJ) model (Lee & Mykland, 2008; Boudt et al, 2011a; Erdemlioglu et al., 2015). In the BSMFAJ model, the logarithmic price process $p(t)$ in continuous time is governed by a semi-martingale with jumps. This can be defined as

$$dp(s) = u(s)ds + \sigma(s)dw(s) + \kappa(s)dq(s) \tag{3.1}$$

where the $u(s)ds$ and $\sigma(s)ds$ are the mean and standard deviation of a conditional random normal process and $w(t)$ is a standard Brownian motion. The finite activity counting process $dq(s)$ is used for capturing jumps occurrence, while $\kappa(s)$ represents the size of the jumps. According to Boudt et al. (2011b), the discrete model followed by BSMFAJ for representing data can be defined as:

$$r_i = f_i s_i u_i + a_i \tag{3.2}$$

where the return r_i is a normal random variable with mean u_i and standard deviation σ_i. The occurrence of jumps is governed by a_i. The standard deviation σ_i=$f_i s_i$, where s_i is an average local factor and f_i is a deterministic component. The mean of the average local volatility, which is the squared periodicity factor, is constant over a local window. The robust estimate of the average volatility of the r_js are in the same local window as r_i, which can be defined based on realised bi-power variation as:

$$\hat{s}_i = \sqrt{\frac{\pi}{2}\frac{1}{[\frac{\lambda}{\Delta}]}\sum_{l=j+2}^{j+[\frac{\lambda}{\Delta}]} |r_l||r_{l-1}|} \tag{3.3}$$

where $r_{j+1}, \ldots, r_{j+[\frac{\lambda}{\Delta}]}$ are $[\frac{\lambda}{\Delta}]$ returns, which are in the same local window as r_i, and λ is the length of the local window.

3.3.1 Intraday periodicity estimation

The SD non-parametric intraday periodicity estimator was first introduced by Taylor and Xu (1997). It can be defined as

$$\widehat{f_i^{SD}} = \frac{SD_i}{\sqrt{\frac{1}{\lceil\frac{\lambda}{\Delta}\rceil}\sum_{j\in N_i} SD_j^2}} \tag{3.4}$$

with $SD_i = \sqrt{\frac{1}{n_i}\sum_{j=1}^{n_i} \bar{r}_{j,i}^2}$. $\bar{r}_{i,1}, \bar{r}_{i,2}, \dots, \bar{r}_{i,j}$ are standardised returns that have the same periodicity factor as $\bar{r}_i$. However, the SD intraday periodicity estimator produces biased results when jumps are present, and so the Shortest Half estimator was proposed by Rousseeuw and Leroy (1988). The Shortest Half estimator is based on the length of all 'halves' of the order statistics, which can be defined as:

$$ShortH_i = 0.741 \cdot \min\{\bar{r}_{(n_i),1} - \bar{r}_{(1),i}, \dots, \bar{r}_{(n_i),1} - \bar{r}_{(n_i-h_i+1),1}\} \tag{3.5}$$

The 'halves' consist of $h_i = \left[\frac{n_i}{2}\right] + 1$ and the order statistics satisfy the condition that $\bar{r}_{(1),i} \leq \bar{r}_{(2),i} \leq \bar{r}_{(n_i),i}$.

The Shortest Half estimator can be defined as

$$\hat{f}_i^{ShortH} = \frac{ShortH_i}{\sqrt{\frac{1}{\lceil\frac{\lambda}{\Delta}\rceil}\sum_{j\in N_i} ShortH_j^2}} \tag{3.6}$$

However, Rousseeuw and Leroy (1988) find that the efficiency of the Shortest Half estimator is lacking. A more efficient estimator, which also robust to jumps, is obtained using weighted standard deviation (WSD). This estimator can be defined as

$$\widehat{f_i^{WSD}} = \frac{WSD_i}{\sqrt{\frac{1}{\lceil\frac{\lambda}{\Delta}\rceil}\sum_{j\in N_i} WSD_j^2}} \tag{3.7}$$

and $$WSD_i = \sqrt{1.08 \cdot \frac{\sum_{i=1}^{n_i} w_{l,j}\bar{r}_{l,j}^2}{\sum_{l=1}^{n_j} w_{l,j}}}$$

The $w_{l,j}$ is the weight of the standardised return $\bar{r}_{l,j}$ and can be given by $w_{l,j} = w(\bar{r}_{l,j}/\hat{f}_j^{ShortH})$. N_i denotes the number of observations per trading day.

3.3.2 Volatility estimation

There are three different methods of evaluating the continuous components of the quadratic variation process. The first method is attributed to Barndorff-Nielsen and Shephard (2004, 2006) which provides an estimate of a realised metric known as bi-power variation. Corsi et al. (2010) argue that bi-power variation (BV) is biased and tends to over-estimate the continuous component, and so they propose threshold bi-power variation (TBV). Aït-Sahalia and Jacod (2012) present an alternative apparatus to measure the presence of the relative components of the quadratic variation process. Below we outline this method and provide some illustrative examples to highlight their differences.

An early way of detecting jumps in the high frequency series was introduced by Barndorff-Nielsen and Shephard (2004, 2006). Their basic methodology consists of constructing a measure of variance that is robust to jumps; the difference between this measure and realised variance can be used to detect the present of jumps. Realised variance is the sum of the squared intraday high frequency returns, which can be defined as

$$RV_t = \sum_{j=1}^{N} r_{t,j}^2 \rightarrow \int_0^t \sigma_\mu^2 d\mu + \sum_{j=1}^{N_i} J_j^2 \, , j = 1, \dots, N \tag{3.8}$$

Where N shows the equally spaced time points while $r_{t,j}$=$p_{t,j} - p_{t,j-1}, t = 1, \dots, T$ is the jth intraday return. Barndorff-Nielsen and Shephard (2006) also propose realised bi-power variation, which is defined as

$$BV_t = \frac{\pi}{2}\sum_{j=2}^{N}\left|r_{t,j-1}\right|\left|r_{t,j}\right| \rightarrow \int_0^t \sigma_\mu^2 d\mu, j = 1, \dots, N \tag{3.9}$$

RV and BV converge to the quadratic variation and integrated variation respectively at a rate of $\sqrt{N}$ as $N\rightarrow \infty$ when they are under weak regularity conditions. Since BV constitutes the continuous part of the quadratic variation, and RV constitutes both the continuous and discontinuous parts, differentiating between BV and RV can separate the quadratic variation into continuous and jump components and thus the impact of jumps can be measured.

Corsi et al. (2010) suggest that estimation bias is caused by the presence of jumps when volatility measures are used in forecasting. The authors argue that jumps are underestimated in previous studies since bi-power variation is biased, particularly when continuous jumps are present in multi-power variation. If $\left|r_{t,i}\right|$ contains a jump and δ is finite, $|r_{t,j-1}|$ and $\left|r_{t,j+1}\right|$, which are multiplied in bi-power variation, do not disappear. Therefore, bi-power variation produces positive bias, which increases as $\left|r_{t,j}\right|$ increases.

Threshold realised variance (Mancini 2009) is defined as:

$$TRV_t = \sum_{j=1}^{N}\left|r_{t,j}\right|^2 I_{\left\{\left|r_{t,j}\right|^2 \leq \Theta(\delta)\right\}} \tag{3.10}$$

Where the $\Theta(\delta)$ and $I_{\{\cdot\}}$ are the threshold and indicator functions respectively, and the threshold function has to satisfy

$$\lim_{\delta\rightarrow 0} \Theta(\delta) = 0$$

$$\lim_{\delta\rightarrow 0} \frac{\delta \log\frac{1}{\delta}}{\Theta(\delta)} = 0$$

Let $\gamma_1, \ldots, \gamma_M > 0$. Realised threshold multi-power variation (TMPV) can be defined as

$$TMPV^{[\gamma_1,\ldots,\gamma_M]} = \delta^{1-\frac{1}{2}(\gamma_1+\cdots+\gamma_M)} \sum_{j=M}^{N} \Pi_{k=1}^{M} |r_{t,j-k+1}|^{\gamma_k} I_{\{|r_{t,j-k+1}|^2 \le \nu_{j-k+1}\}} \tag{3.11}$$

where ν_{j-k+1} is a strictly positive random threshold function. *M* is the number of absolute returns raised to a non-negative power. *N* denotes equally spaced time points within a day.

A special simple case of TMPV is threshold bi-power variation (TBV) with $\gamma_1 = \gamma_2$=1, given by

$$TBV = \frac{\pi}{2} TMPV^{[1,1]} = \frac{\pi}{2} \sum_{j=2}^{N} |r_{t,j-1}||r_{t,j}| I_{\{|r_{t,j-1}|^2 \le \nu_{j-1}\}} I_{\{|r_{t,j}|^2 \le \nu_j\}} \tag{3.12}$$

TBV will correct the bias in bi-power variation since the indicator function vanishes when jumps are present (Corsi et al., 2010). The jump test used in Corsi et al.'s study is defined as z=$\delta^{-\frac{1}{2}} \frac{(RV_t - BPV_t) \cdot RV_t^{-1}}{\sqrt{\bar{\vartheta} \max\{1, \frac{TriPV_t}{(BPV_t)^2}\}}}$ with $\bar{\vartheta} = \frac{\pi^2}{4} + \pi - 5$, where TriPV is one of the fourth-power counterparts of the estimates of BV. Corrected threshold bi-power variation (CTBV) corrects the problem caused by TBV when δ is finite.

3.3.3 HAR-class models

Andersen et al. (2007a) find that realised volatility is strongly temporally dependent, and so they introduce the ABD model to describe RV's slow-decaying autocorrelations. However, the ABD model does not the separate the jump components and continuous components of quadratic variation. Corsi et al. (2008, 2010) and Corsi (2009) extend the

ABD model to the HAR-J, HAR-TJ, HAR-CJ and HAR-TCJ models, which can separate the jump parts and the continuous parts when forecasting realised volatility. The basic HAR model introduced by Corsi (2009) is defined as:

$$\mathrm{RV}_{t:t+h-1}=\beta_0+\beta_d\mathrm{RV}_{t-1}+\beta_w\mathrm{RV}_{t-5:t-1}+\beta_m\mathrm{RV}_{t-22:t-1}+\varepsilon_t$$

$$\mathrm{RV}_{t_1:t_2}=\frac{1}{t_2-t_1+1}\sum_{t=t_1}^{t_2}RV_t,\text{with } t_1\leq t_2 \tag{3.13}$$

RV_{t-1}, $\mathrm{RV}_{t-5:t-1}$ and $\mathrm{RV}_{t-22:t-1}$ are the daily, weekly and monthly lags of realised volatility, which are used for capturing the long-memory dynamic dependence of RV. The error term is an independent and identically distributed (i.i.d.) random variable with mean 0 and variance σ^2. Based on the basic HAR model, the HAR-J and HAR-CJ models introduce a way to separate the continuous and jump parts in volatility forecasting by using threshold bi-power variation. The HAR-J model can be presented as

$$\mathrm{RV}_{t:t+h-1}=\beta_0+\beta_d\mathrm{RV}_{t-1}+\beta_w\mathrm{RV}_{t-5:t-1}+\beta_m\mathrm{RV}_{t-22:t-1}+\beta_j\hat{J}_{t-1}+\varepsilon_t \tag{3.14}$$

and the HAR-CJ model is defined as

$$\mathrm{RV}_{t:t+h-1}=\beta_0+\beta_d\hat{C}_{t-1}+\beta_w\hat{C}_{t-5:t-1}+\beta_m\hat{C}_{t-22:t-1}+\beta_j\hat{J}_{t-1}+\varepsilon_t \tag{3.15}$$

The jump parts in (3.14) and (3.15) can be expressed as $\hat{J}_t=\mathrm{I}_{\{z_t>\phi_\alpha\}}\cdot\max[(\mathrm{RV}_t-\mathrm{BV}_t),0]$ and the continuous parts as $\hat{C}_t=RV_t-\hat{J}_t$. $\beta_d,\beta_w, and\ \beta_m$ are the parameters for the daily, weekly and monthly lags of the continuous components of quadratic variation, while β_j is the parameter for the daily lag of the jump components. However, Corsi et al. (2010) show that bi-power variation underestimates the jump components, which leads to biased forecasting results using HAR-CJ. Therefore, they introduce the HAR-TJ and HAR-

TCJ models, which use threshold bi-power variation to separate the jump components from the quadratic variation. The HAR-TJ and HAR-TCJ models can be read as

$$\mathrm{RV}_{t:t+h-1}=\beta_0+\beta_d \mathrm{RV}_{t-1}+\beta_w \mathrm{RV}_{t-5:t-1}+\beta_m \mathrm{RV}_{t-22:t-1}+\beta_j \widehat{\mathrm{TJ}}_{t-1}+\varepsilon_t \tag{3.16}$$

$$\mathrm{RV}_{t:t+h-1}=\beta_0+\beta_d \widehat{TC}_{t-1}+\beta_w \widehat{TC}_{t-5:t-1}+\beta_m \widehat{TC}_{t-22:t-1}+\beta_j \widehat{TJ}_{t-1}+\varepsilon_t \tag{3.17}$$

The jump part $\widehat{TJ}_t$ and continuous part $\widehat{TC}_t$ can be expressed as $\widehat{TJ}_t = I_{\{z_t>\Phi_\alpha\}} \cdot \max\,[(RV_t - TBV_t), 0]$ *and* $\widehat{TC}_t = RV_t - \widehat{TJ}_t$ respectively, where Φα is the cumulative distribution function of the normal distribution at confidence level α and *x*+ = max(*x*, 0). In addition, $\widehat{TC}_{t-1}$, $\widehat{TC}_{t-5:t-1}$ and $\widehat{TC}_{t-22:t-1}$ are the daily, weekly and monthly lags of the continuous components. The error terms in these HAR-family models are i.i.d. random variables with mean 0 and variance σ^2.

3.3.4 Volatility estimation models with leverage effects

We use the Glosten-Jagannathan-Runkle GARCH model (GJR) model to estimate the conditional variance of intraday returns. This is because we find leverage effects in intraday returns, which can be captured in the GJR model. The GJR (1,1) with a Gaussian error term distribution can be written as:

$$r_{t,n}=\mu+\varepsilon_{t,n} \text{ where } \varepsilon_t = \sigma_t z_t \tag{3.18}$$

$$\sigma_{t,n}^2 = \gamma_1 \sigma_{t,n-1}^2 + \alpha_1 \varepsilon_{t,n-1}^2 + \xi_1 I[\varepsilon_{t,n-1} < 0]\varepsilon_{t,n-1}^2, \text{ where } n=1,\dots,M,\ t=1,\dots,T \tag{3.19}$$

The intraday returns $r_{t,n}$ that we use here refer to the *n*th return in day *t*. In our analysis, we use the intraday periodicity-filtered and unfiltered returns, with the SD, WSD and Short-H estimators as intraday periodicity filters, in order to compare how different filters can affect estimation in GJR models.

3.4 Impact of intraday periodicity on stylised facts of returns and volatility measures

This section first examines the impact of intraday periodicity on the autocorrelations for SPY returns and its realised measures, together with its unconditional and conditional tail properties and leverage effects for returns. The returns and volatility generated from business-time and calendar-time sampling schemes for the SPY index are used in this section, as this index can reflect the average movements of the stocks in the market.

3.4.1 Autocorrelations

[Insert Figures 3.1 to 3.3 here]

Figures 3.1 and 3.2 show that filtering by intraday periodicity does not have a significant impact on the autocorrelations for SPY returns. Comparing Figures 3.1 and 3.3, we can also see that the autocorrelations are absent for both calendar-time sampling and business-time sampling returns.

[Insert Figures 3.4 to 3.6 here]

Figures 3.4 and 3.5 show that both the intraday periodicity-filtered and unfiltered absolute returns have a slow decay in autocorrelations, and that the partial autocorrelation values are not notably different between the two types of returns. In addition, the partial autocorrelation values for lags 15 to 20 for 300-second business-time absolute intraday returns are greater than zero, while the equivalent values for the calendar-time absolute intraday returns in Figure 3.6 are comparatively much closer to zero. This indicates that the autocorrelations for business-time sampling intraday returns decay more slowly than those for calendar-time sampling intraday returns.

[Insert Figures 3.7 to 3.8 here]

The partial autocorrelation results in Figures 3.7 and 3.8 show that the RV and BV realised measures are AR (9) processes, because the partial autocorrelations cut off after the ninth lag. This shows that a given day's volatility can have a significant impact on its volatility up to nine days later, implying that past stock volatility can be useful in forecasting future volatility. However, the intraday periodicity-filtered and business-time RV and BV show larger correlations at lags 15 and 20. This indicates that the realised measures have different long-range dependencies when using different sampling schemes or when filtering by intraday periodicity.

3.4.2 Conditional and unconditional heavy tails

Figures 3.9 to 3.10 show quantile-quantile (Q-Q) plots for intraday periodicity-filtered and unfiltered returns.

[Insert Figures 3.9 to 3.10 here]

It is evident that the returns have a heavy tail, whether filtering by intraday periodicity or not. Filtering by intraday periodicity does not have a significant impact on the heavy tails for SPY's daily returns.

[Insert Figures 3.11 to 3.12 here]

The Q-Q plot for the residuals of the GARCH (1,1) model in Figures 3.11 to 3.12 shows that intraday periodicity does not have a significant impact on SPY's conditional heavy tails either.

3.4.3 Leverage effects

Table 3.1 shows the leverage effects estimated from the GJR (1,1) model for intraday periodicity-filtered and unfiltered returns for the AAPL, MSFT, JNJ and PFE stocks and SPY.

[Insert Table 3.1 here]

From the GRJ (1,1) estimation results shown in Table 3.1, we can see that the leverage effects for intraday returns are higher for lower-frequency data (with one exception, discussed below). For example, the leverage effect parameters in the GJR (1,1) models are the highest for 300-second intraday returns for both MSFT (0.01591) and SPY (0.03573). The only exception is that the leverage effects for 150-second PFE is 0.01535, which is larger than the equivalent 300-second data (0.01517). However, the difference between the estimated coefficients using these two sampling frequencies is very small at 0.00018.

It is clear that the leverage effects fall after adjusting for intraday periodicity for most frequencies. For example, the leverage effect coefficients for 300-second intraday returns for MSFT and SPY fall from 0.01591 and 0.03573 to 0.00978 and 0.02725 respectively.

Table 3.1 also shows that adjusting for intraday periodicity does not have a significant impact on the volatility persistence of the intraday returns. However, the impact on a given day's volatility of past volatility (γ_1) and past innovations (α_1) are stronger and weaker respectively after adjusting for intraday periodicity.

3.5 Impact of intraday periodicity on volatility from calendar-time sampling data

In this section, we test the impact of intraday periodicity on the number of intraday jumps (based on the LM jump test) and the volatility components (the jump and continuous components) estimated via BV, TBV and CTBV. Threshold bi-power variation (TBV) is used in our analysis as Corsi et al. (2010) suggests that it performs better than bi-power variation (BV) at separating the jump components from the quadratic variation. Here, the threshold $\hat{v}_t$ in equation (5.13) is defined as $\hat{v}_t = c_v^2 \hat{V}_t$, where $c_v = 3$ and the $\hat{V}_t$ is local spot variance, which is generated from return data.

3.5.1 Jumps and jump components

We first examine the impact of intraday periodicity on stock volatility, followed by the jump components estimated using different volatility measures (BV and TBV) from financial assets. Table 3.2 shows the realised volatility and conditional variance of stocks and SPY from 2000 to 2016, with and without filtering for intraday periodicity.

[Insert Table 3.2 here]

On average, the volatility of the two stocks from the IT sector is higher than that of the stocks from the healthcare sector and the SPY index. One of the main reasons for this result is that the burst of the tech bubble in 2000 had a bigger effect on the IT sector compared to other sectors. Andersen et al. (2010) report that the price of the index for the IT sector in the S&P 500 decreased by 50% in 2000, and by August 2002 had fallen by 80% from its highest point. In addition, the lower unconditional and conditional volatility for SPY shows that this index is less sensitive to market shocks than individual stocks.

It is also apparent from Table 3.2 that regardless of the estimation method used for intraday periodicity filtering, the changes in both RV and conditional volatility are all under 15%**.** Filtering by intraday periodicity results in a rise in RV for all stocks across different frequencies, but for conditional variance, intraday periodicity filtering has mixed effects. In addition, the impact of the WSD estimator on RV and conditional variance is closer to that of the SD estimator across different frequencies than the Shortest Half estimator. This is because the Shortest Half estimator is estimated using order statistics, which are not used in the SD and WSD estimators.

[Insert Figure 3.13 here]

The different patterns of intraday periodicity estimated using different methods are very clear in Figure 3.13. The graph shows that the Shortest Half estimator is slightly higher than the SD and WSD estimators at the beginning of day, but it falls below them in the middle of the day. These differences produce the differing effects of filtering on RV and conditional variance as seen in Table 3.1. In addition, we can also see that all three estimators are less smooth when they are estimated from higher frequency data, showing that they are more likely to be affected by jumps using higher frequency estimation. Also, the SD estimator tends to have larger spikes than WSD and Shortest Half, which suggests that SD is most dramatically affected by the presence of jumps.

Table 3.3 reports the number of intraday jumps estimated from the Lee and Mykland (2007) jump test for intraday periodicity-filtered and unfiltered returns. The LM test statistic is defined as:

$$T_{LM,t} = \frac{(\max(\tilde{T}_{LM,t_i}) - C_M}{S_M} \text{ where } \tilde{T}_{LM,t_i} = \frac{|r_{t_i}|}{\sqrt{\hat{V}_{t_i}}} \tag{3.20}$$

Where $C_M = \frac{(2logM)^{1/2}}{0.8}$ - $\frac{log\pi + \log(logM)}{1.6(2log\pi)^{1/2}}$ and $S_M = \frac{1}{0.6(2log\pi)^{1/2}}$. r_{t_i} is the *i*th return at day *t* and where *M* is the number of sampled observations per trading day. $\hat{V}_{t_i}$ denotes the local variance estimate. The null hypothesis for the test is that no jumps are present in the price series.

[Insert Table 3.3 here]

Table 3.3 shows that the number of estimated intraday jumps is higher when using higher frequency sampling data. However, we can also see from the table that although the intraday periodicity patterns are different, the number of intraday jumps for each stock and for SPY are all lower after filtering for intraday periodicity. In addition, large shocks have a smaller impact on volatility for SPY than for individual stocks (i.e. it has low volatility, as shown in Table 3.1), therefore the number of intraday jumps is lower for SPY than for stocks. We also find that intraday periodicity has less of an impact on reducing jump frequency for SPY than for stocks.

[Insert Table 3.4 here]

Intraday periodicity has a similar impact on daily jump components, as shown in Table 3.4. The continuous components estimated using BV, TBV and CTBV are all larger after filtering for intraday periodicity for 150-second and 300-second realised measures, which means that the jump components are smaller. However, the intraday periodicity estimators give mixed results for the continuous components for 30 seconds and 60 seconds. This is probably caused by the fact that these three estimators are more sensitive to the presence of jumps when using higher frequency data.

SPY in Table 3.1 has a lower RV and daily conditional variance compared to the individual stocks, both before and after filtering for intraday periodicity. It also has a higher proportion of continuous components than for each individual stock before filtering in Table 3.4. This is because the volatility of SPY is affected by the collective performance of many companies, rather than just one, and hence reflects general trends in the whole financial market. While individual stocks may have extreme values caused by major events related to specific companies, SPY has less extreme fluctuations and therefore has smaller jump components and RV.

[Insert Figure 3.14 here]

Figure 3.14 shows the comparison of intraday periodicity between SPY and stocks using 300-second sampling frequency data. The WSD estimator is used to estimate intraday periodicity as Boudt et al. (2011b) show that it is more robust to the present of jumps than the ShortH estimator. The difference in intraday periodicity between SPY and stocks is clear, as the results for stocks have a strong L-shaped pattern, while the result for SPY is more U-shaped. In other words, the intraday periodicity for stocks all start at or close to 3 at the beginning of the trading day, while that for SPY starts at a value close to 1.5, with similar end-points for both. Therefore, SPY's intraday periodicity shows a less dramatic change over the course of the day, which results in a smaller impact on jumps and volatility compared to stocks.

3.5.2 Volatility forecasting

Table 3.5 shows the impact of intraday periodicity on volatility forecasting using HAR-family models.

[Insert Table 3.5 here]

The results in Table 3.5 show that filtering by intraday periodicity improves the performance of HAR-family models for the JNJ, PFE, MSFT stocks and the SPY index. This indicates that the decrease in jump frequency caused by filtering discussed in Section 3.5.1 produces better forecasting results. In addition, the impact of intraday periodicity on

SPY volatility forecasting is not as significant as that on these three stocks, as the changes in MSE for volatility forecasting for SPY are minor and are restricted to the SD and WSD estimators. This is because intraday periodicity has less of an effect on the jumps for SPY than for stocks, as shown in Section 3.5.1, therefore resulting in a weaker impact on volatility forecasting for stocks and SPY. We also find that filtering by intraday periodicity fails to improve forecasting results for AAPL. Because AAPL has more than double the RV of other stocks and SPY, as shown in the previous section, it poses difficulties for intraday periodicity estimators to capture the intraday periodicity patterns for this highly volatile stock, and therefore produces biased forecasting results.

3.5.3 Volatility in different regimes

Table 3.6 shows the number of intraday jumps for different stocks and SPY in different regimes.

[Insert Table 3.6 here]

The results show that the stocks have more jumps for higher frequency intraday returns and that filtering reduces the number of jumps in each regime, which is in line with the results for the data set as a whole, discussed in Section 3.5.1.

[Insert Tables 3.7 to 3.9 here]

Tables 3.7 to 3.9 show the impact of intraday periodicity on the continuous components for stocks and SPY in the pre-crisis, crisis and post-crisis periods respectively. The results show that filtering by intraday periodicity increases the continuous components for stocks and SPY across different frequencies in the post-crisis period when using the SD and WSD estimators. However, intraday periodicity filtering raises the continuous components (hence producing falls in the jump components) when using low-frequency (150 and 300 second) data and has a mixed impact on high-frequency (30 and 60 second) volatility components in the pre-crisis and crisis periods. This may be because the stock returns are highly volatile between 2000 and 2002 in the pre-crisis period, and between 2008 and mid-2009 in the crisis period, caused by the burst of the tech bubble and the financial crisis respectively. These volatile periods cause difficulties for intraday periodicity estimators when capturing high frequency (30 second and 60 second) return patterns with higher volatility and many intraday jumps (as discussed in Section 3.5.1). This has an effect on the performance of intraday periodicity-filtered volatility measures, and therefore affects the estimation of the volatility components.

[Insert Tables 3.10 to 3.12 here]

Tables 3.10 to 3.12 show the impact of intraday periodicity on volatility forecasting across different regimes. The MSE results show that the performance of HAR-class models at volatility forecasting is worse during the crisis than in the pre-crisis and post-crisis periods. Periods of high volatility produce difficulties in volatility forecasting. In addition, the MSE results for forecasting are higher for AAPL than other stocks in both the pre-crisis and post-crisis periods. APPL exhibits higher volatility after the burst of the tech bubble (2000-2002) and during the financial crisis compared to other stocks and SPY, and is thus the hardest stock for volatility forecasting.

The MSEs in the forecasting results also show that filtering by intraday periodicity fails to improve the performance of HAR-family models for AAPL due to its high volatility, which is in line with the 2000-2016 results presented in Section 3.5.2. Filtering by intraday periodicity has a positive impact on volatility forecasting across different frequencies for the less volatile stocks JNJ and PFE in the crisis and post-crisis periods. For the volatile stocks MSFT and SPY, however, the improvements are present overall yet are not consistent across different frequencies. On the other hand, filtering by intraday periodicity has a negative impact on volatility forecasting for SPY and all stocks except PFE in the pre-crisis period. This is because the high volatility caused by the burst of the tech bubble from 2000 to 2002 may create difficulties for intraday periodicity estimators to capture the intraday periodicity patterns, leading to potentially biased intraday periodicity-filtered volatility measures in forecasting for highly volatile data.

3.6 Impact of intraday periodicity on volatility from business-time sampling data

3.6.1 Jumps and jump components

This section examines the impact of intraday periodicity on assets' jump frequencies, volatility components and volatility forecasting based on returns generated from business-time sampling with 78 returns per trading day, which is equivalent to 300-second calendar-time sampling. First, we compare the differences between stocks estimated using calendar-time sampling (Figure 3.14) and business-time sampling (Figure 3.15), both with the WSD intraday periodicity estimator.

[Insert Figure 3.15 here]

The intraday periodicity patterns estimated using business-time sampling have smaller values at both the beginning and end of the day compared to the calendar-time sampling returns. The values in Figure 3.15 are close to or smaller than 2.5 and 1 at the beginning and end of the day respectively, while the values for calendar-time sampling in Figure 3.14 are close to 3 and 1.5 respectively. This shows that business-time sampling can better reduce the impact of intraday periodicity compared to calendar-time sampling.

Table 3.13 shows the average unconditional and conditional volatility for intraday periodicity-filtered and unfiltered returns using intraday periodicity estimated via business-time sampling.

[Insert Table 3.13 here]

Unconditional volatility and conditional volatility are highest for the two stocks from the IT sector (AAPL and MSFT) for business-time returns, which is in line with the results for calendar-time returns. In addition, filtering by intraday periodicity produces higher RV for stocks and SPY.

[Insert Table 3.14 here]

The effect of intraday periodicity on the number of intraday jumps for business-time sampling data is shown in Table 3.14. The results show that business-time sampling returns have much fewer intraday jumps than calendar-time sampling returns for each stock in different financial regimes (pre-crisis, crisis and post-crisis periods). The drop in the number of intraday jumps when using business-time data is clearest for AAPL and SPY, which drop from 704 and 300 with calendar-time sampling to 103 and 62 with business-time sampling. Filtering by intraday periodicity reduces the number of intraday jumps across different periods, which reflects the results from the calendar-time sampling data.

[Insert Table 3.15 here]

Table 3.15 shows the impact of intraday periodicity on jump components for business-time sampling returns. By comparing the results in Table 3.15 with those discussed in Section 3.5.1, we can see that the continuous components estimated using business-time sampling are higher than those using calendar-time sampling for each stock across different financial conditions. This means that business-time sampling has a lower proportion of jump components, which gives it an advantage over calendar-time sampling. In addition, filtering by intraday periodicity increases stocks' continuous components (hence reducing their jump components) in the pre-crisis, crisis and post-crisis periods across most frequencies using different intraday periodicity estimators. The intraday periodicity-filtered volatility measures produce very large continuous components with values all above 96%, 83% and 88% using the BV, TBV and CTBV volatility measures respectively for all time periods. This suggests that filtering by intraday periodicity has a more consistent impact on business-time returns and is more efficient at reducing the jump components in this type of data compared to calendar-time data.

3.6.2 Volatility forecasting

Table 3.16 shows the impact of intraday periodicity on volatility forecasting for business-time sampling data in different financial regimes.

[Insert Table 3.16 here]

The MSE results in Table 3.16 show that the HAR-family models are better at forecasting RV in less volatile periods (the pre-crisis and post-crisis periods) or when using the whole sample from 2000-2016. The HAR-family models have difficulties in forecasting volatility during the financial crisis, as stocks and SPY are very volatile during this period. The HAR-TJ model produces similar MSE results to the other models in the pre-crisis and post-crisis periods, but it outperforms them for the most volatile crisis period, which indicates that threshold bi-power variation is superior to bi-power variation for data obtained from business-time sampling in volatile periods. Filtering by intraday periodicity decreases the MSEs for the stocks from the less active healthcare sector (JNJ and PFE) in the pre-crisis and post-crisis periods, while it produces mixed results for the volatile IT stocks AAPL and MSFT. This shows that filtering by intraday periodicity can reduce the number of jumps and jump components (thus increasing the continuous components), but that these improvements in forecasting may be limited to low-volatility data. These results are in line with the calendar-time sampling data results in Section 3.5.

3.7 Conclusion

This chapter has analysed the impact of intraday periodicity on jumps and volatility forecasting in different financial regimes (before, during and after the 2008 financial crisis) using different sampling schemes (calendar-time sampling and business-time sampling). We first examined the impact of intraday periodicity on the stylised facts of the SPY index, including the autocorrelation of returns, absolute returns, volatility measures, and the unconditional and conditional tail properties and leverage effects for daily returns. Second, we compared the usefulness of calendar-time sampling and business-time sampling data by examining the impact of filtering by intraday periodicity on data using both sampling schemes. We did this by studying the effect of intraday periodicity filtering on assets' conditional and unconditional volatility, jump frequency, volatility components (estimated using BV, TBV and CTBV) and volatility forecasting using HAR-class models in different regimes for different sampling schemes.

The results show that filtering by intraday periodicity lowers jump frequency and leverage effects of daily returns for stocks and SPY in different regimes. However, it only reduces the jump components and improves forecasting for less volatile data, such as low volatility stocks or the less volatile post-crisis period. Highly volatile data such as IT stocks and crisis data present difficulties for intraday periodicity estimators in capturing intraday patterns.

The comparison of business-time sampling and calendar-time sampling data revealed that the absolute returns from business-time sampling have autocorrelations that decay more slowly and have lower volatility persistence in a GJR model than calendar-time sampling data. In addition, intraday periodicity patterns are weaker in business-time sampling data than calendar-time sampling data. Also, volatility for business-time sampling has fewer jumps and a smaller proportion of jumps than for calendar-time sampling data. Filtering by intraday periodicity consistently reduces the number of jumps for all stocks for data using both sampling schemes, but business-time sampling shows more consistent jump component reductions than calendar-time data using the SD and WSD estimators. Finally, filtering by intraday periodicity improves RV forecasting for less volatile business-time data such as healthcare stocks and data from the post-crisis period, which is in line with the results from the calendar-time sampling data.

In sum, the findings suggest that business-time sampling data may be more useful for volatility analysis than calendar-time data. Future work could conduct further comparisons of business-time and calendar-time sampling schemes by replicating the present study with new data. As stock markets currently undergo notable volatility as a result of the Covid-19 pandemic, future researchers will have a new set of volatile stock market data with which to make useful comparisons between different sampling schemes for stock market volatility analysis.

Appendix

Figures

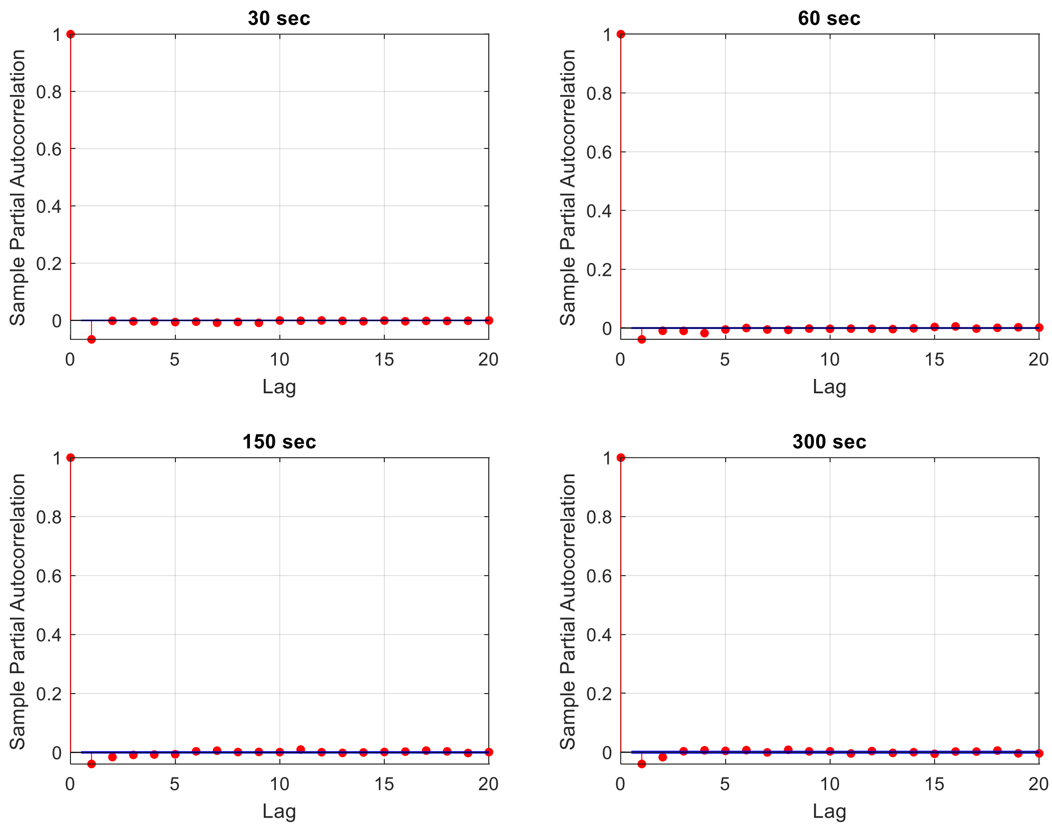

Figure 3.1: Partial autocorrelations for intraday returns for the SPY index, using 30-second, 60-second, 150-second and 300-second returns.

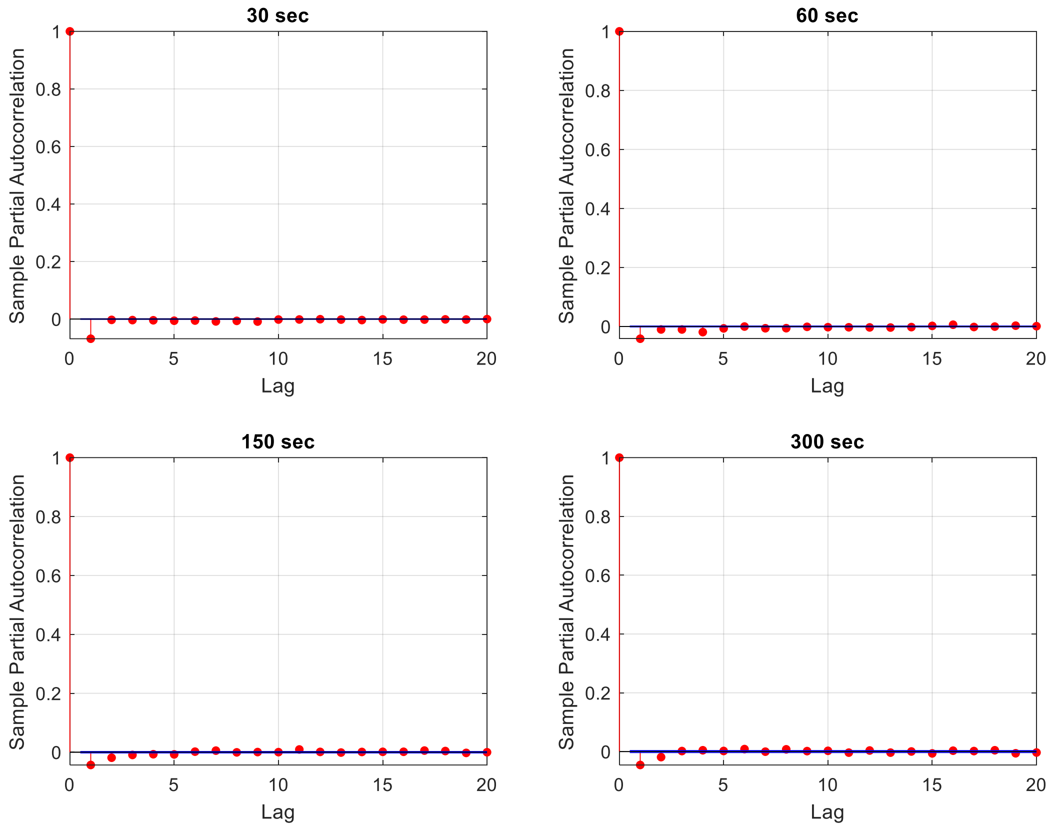

Figure 3.2: Partial autocorrelations for intraday periodicity-adjusted intraday returns for the SPY index using the WSD intraday periodicity estimator. 30-second, 60-second, 150-second and 300-second returns are shown.

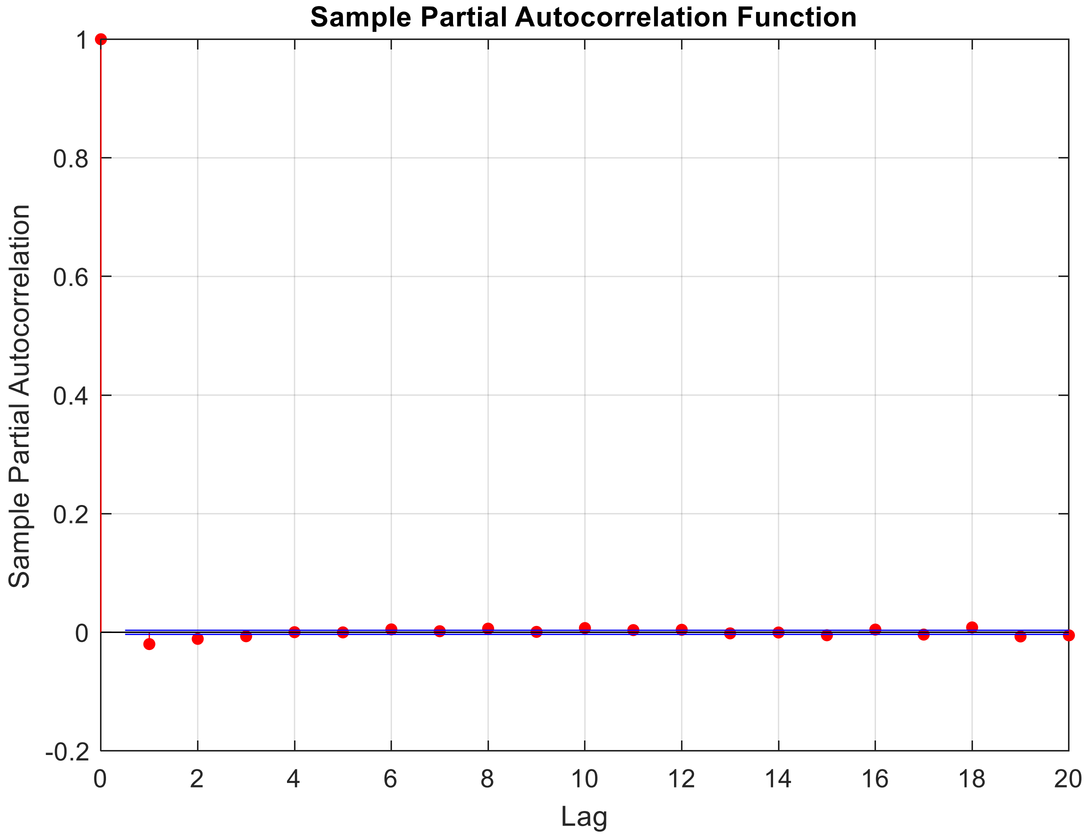

Figure 3.3: Partial autocorrelations for 300-second business-time sampling returns for the SPY index.

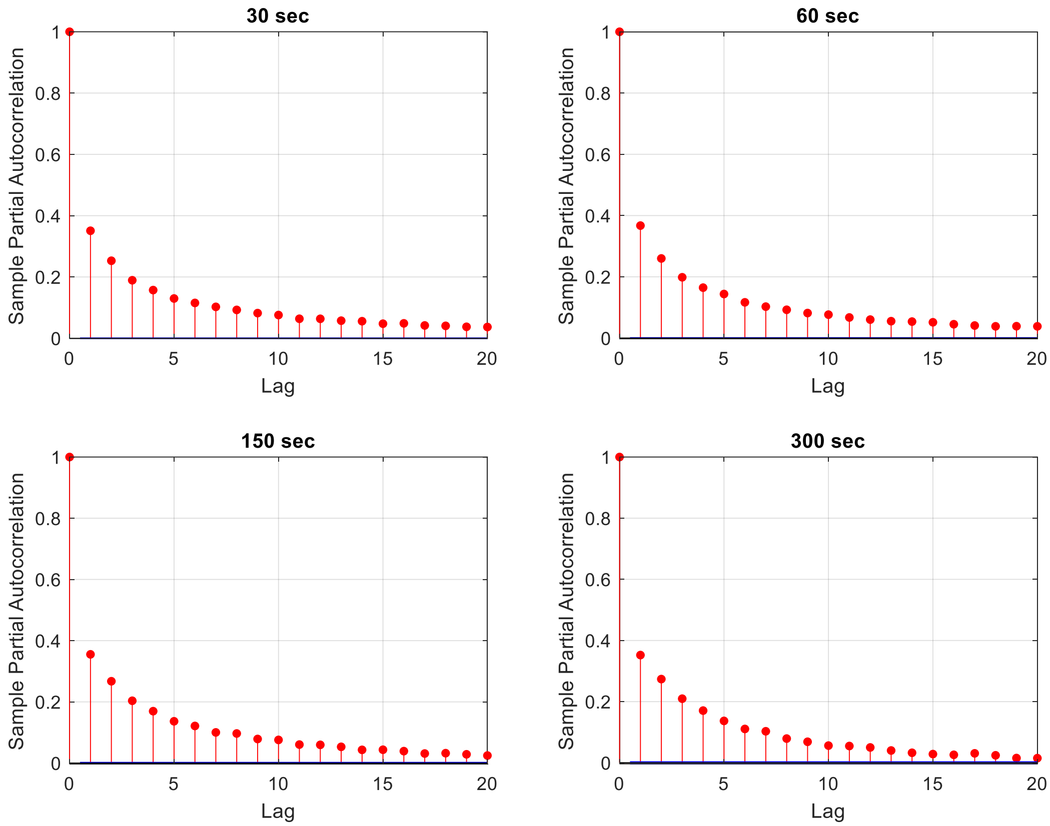

Figure 3.4: Partial autocorrelations for absolute intraday returns for the SPY index using 30-second, 60-second, 150-second and 300-second returns.

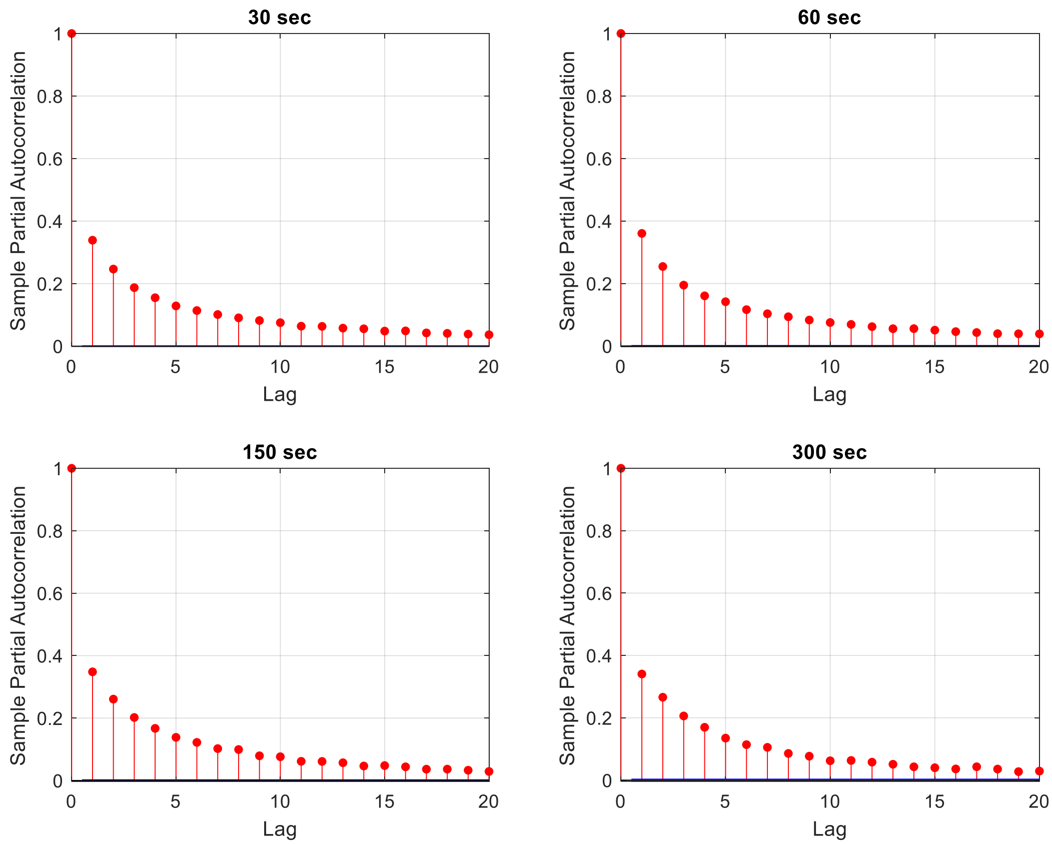

Figure 3.5: Partial autocorrelations for absolute intraday periodicity-adjusted returns for the SPY index using 30-second, 60-second, 150-second and 300-second returns. The filtered returns are estimated using the WSD intraday periodicity estimator.

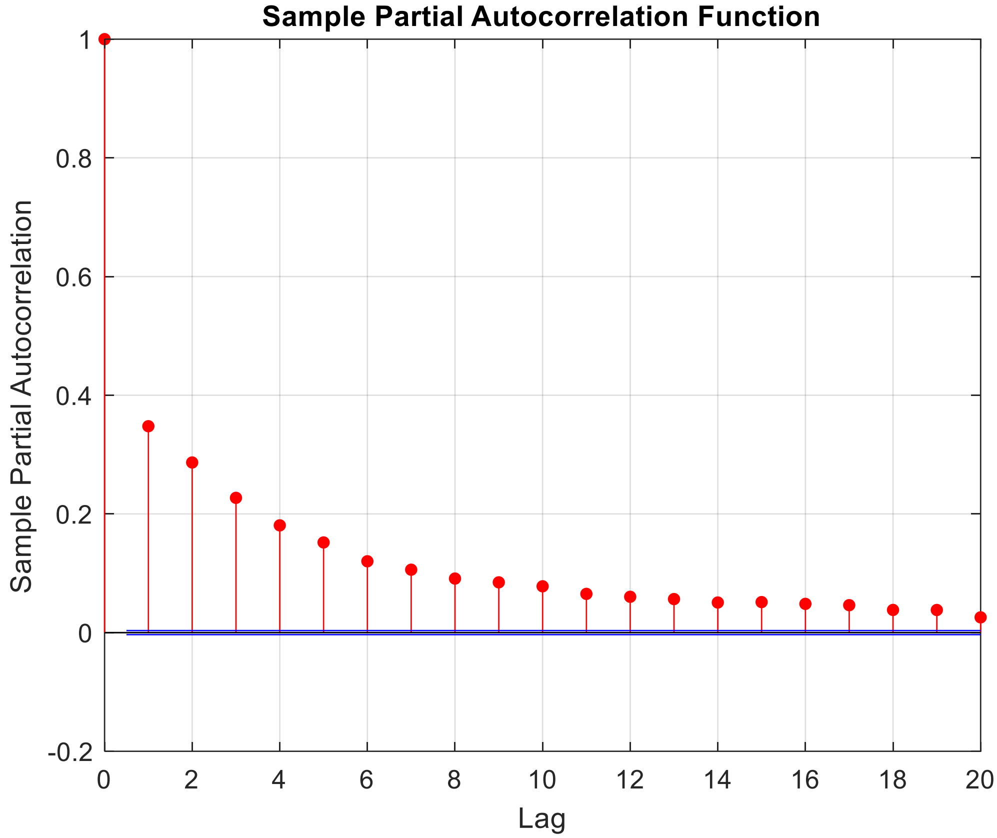

Figure 3.6: Partial autocorrelations for absolute 300-second business-time sampling returns for the SPY index.

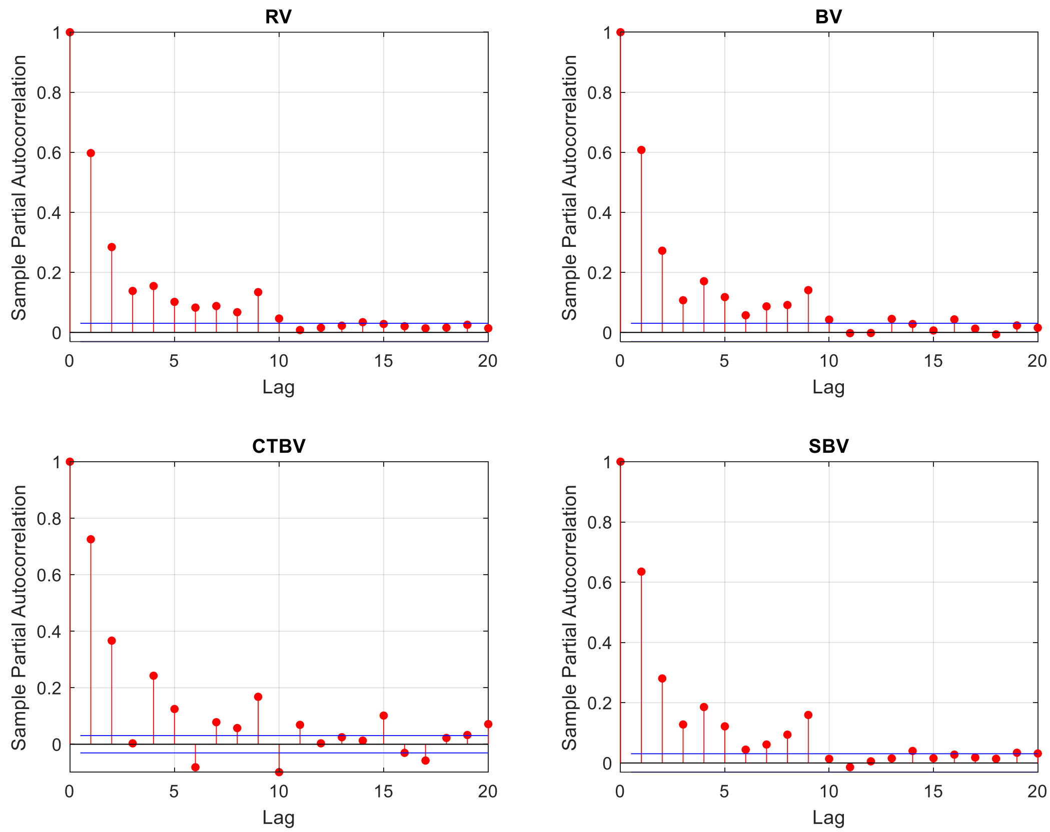

Figure 3.7: Partial autocorrelations for RV, BV, CTBV and SBV for 300-second calendar-time sampling returns for SPY index.

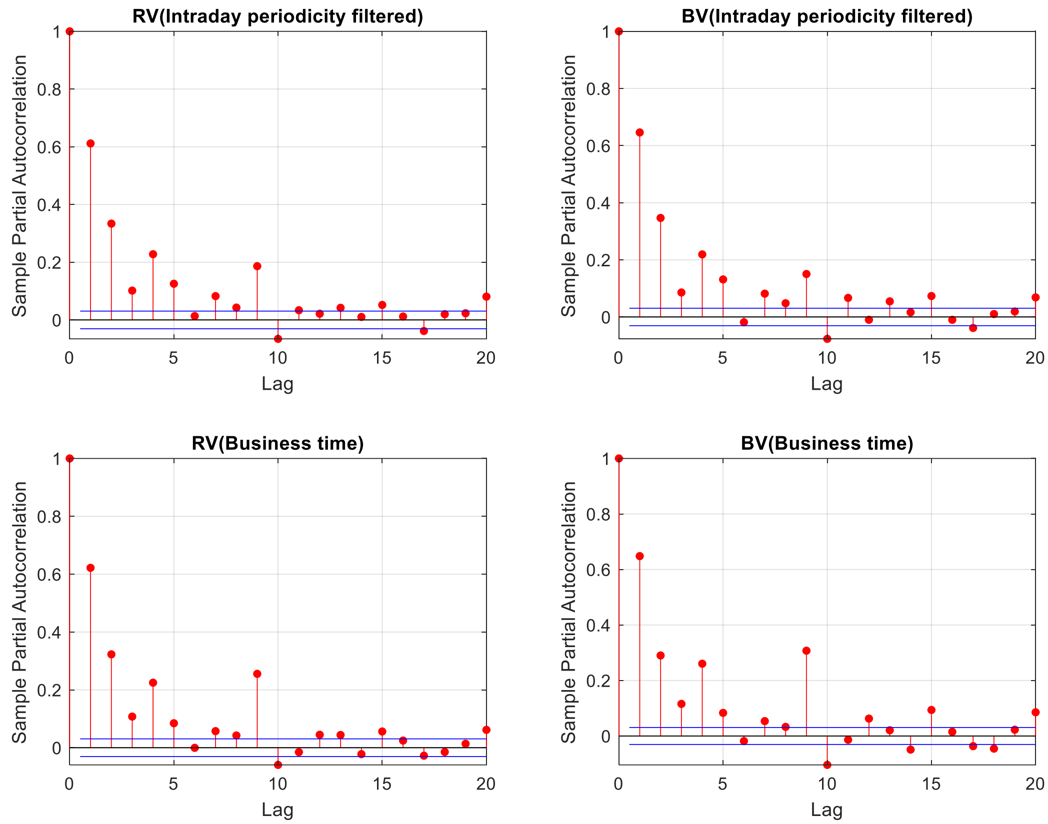

Figure 3.8: Partial autocorrelations for RV and BV estimated using intraday periodicity-filtered returns (top panels) and business-time returns (bottom panels), The filtered returns are estimated using the WSD intraday periodicity estimator.

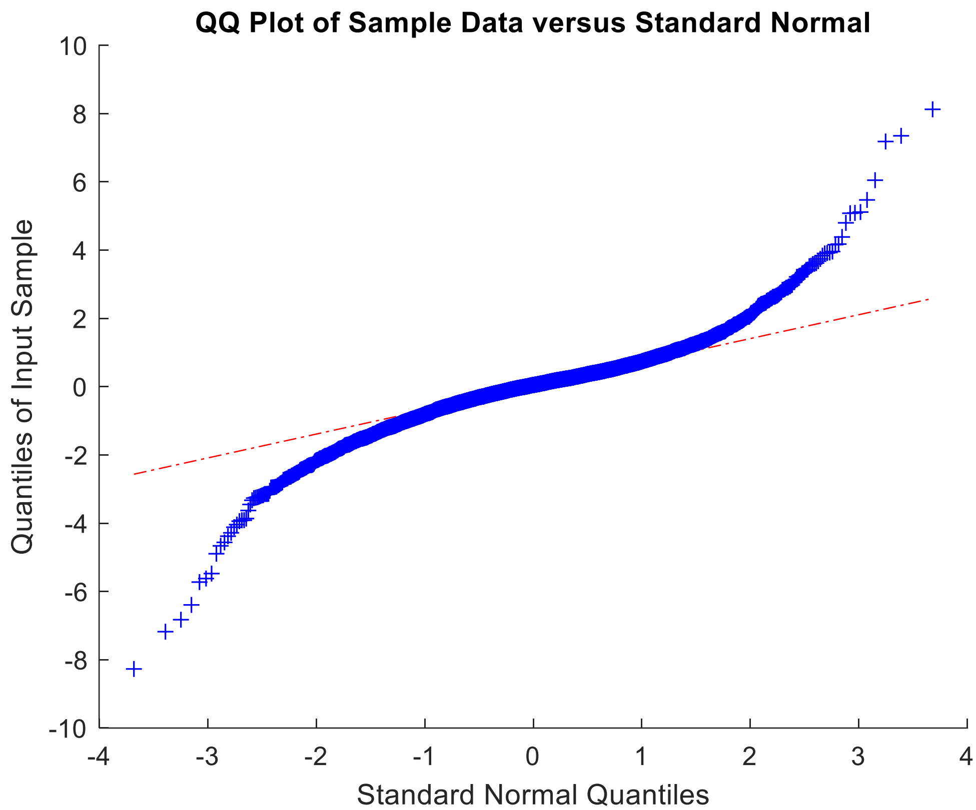

Figure 3.9: Quantile-quantile (Q-Q) plot for high-frequency aggregated returns for the SPY index.

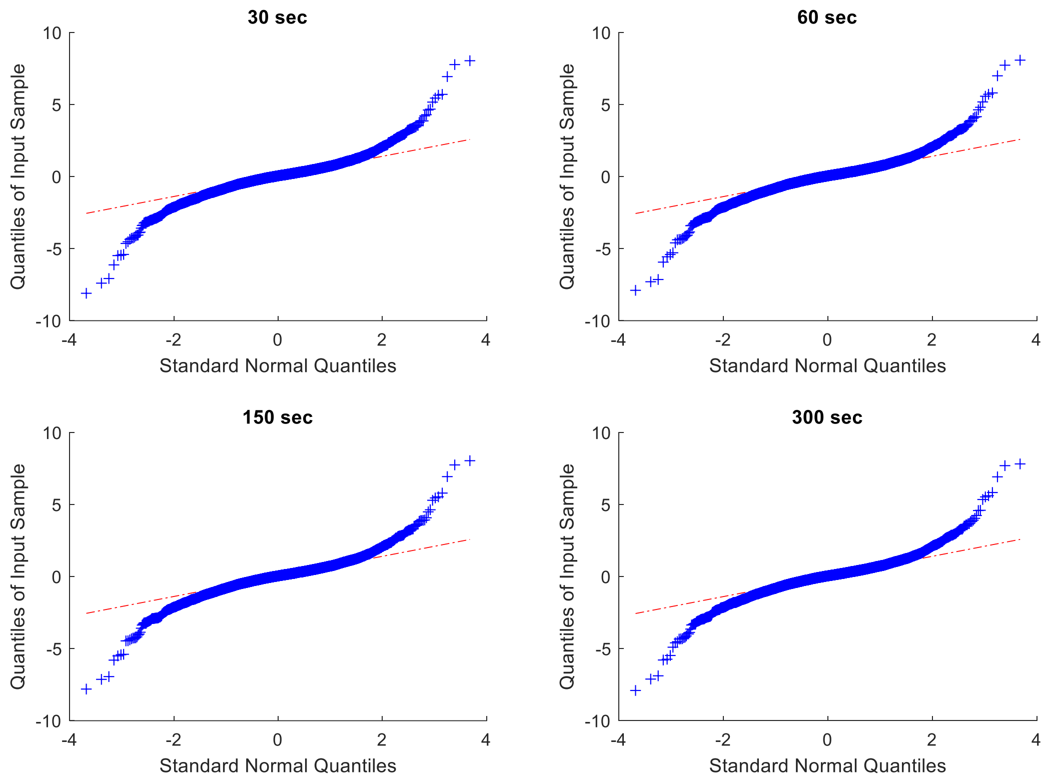

Figure 3.10: Quantile-quantile (Q-Q) plots for intraday periodicity-adjusted high-frequency aggregated returns for the SPY index using 30-second, 60-second, 150-second and 300-second returns, filtered using the WSD intraday periodicity estimator.

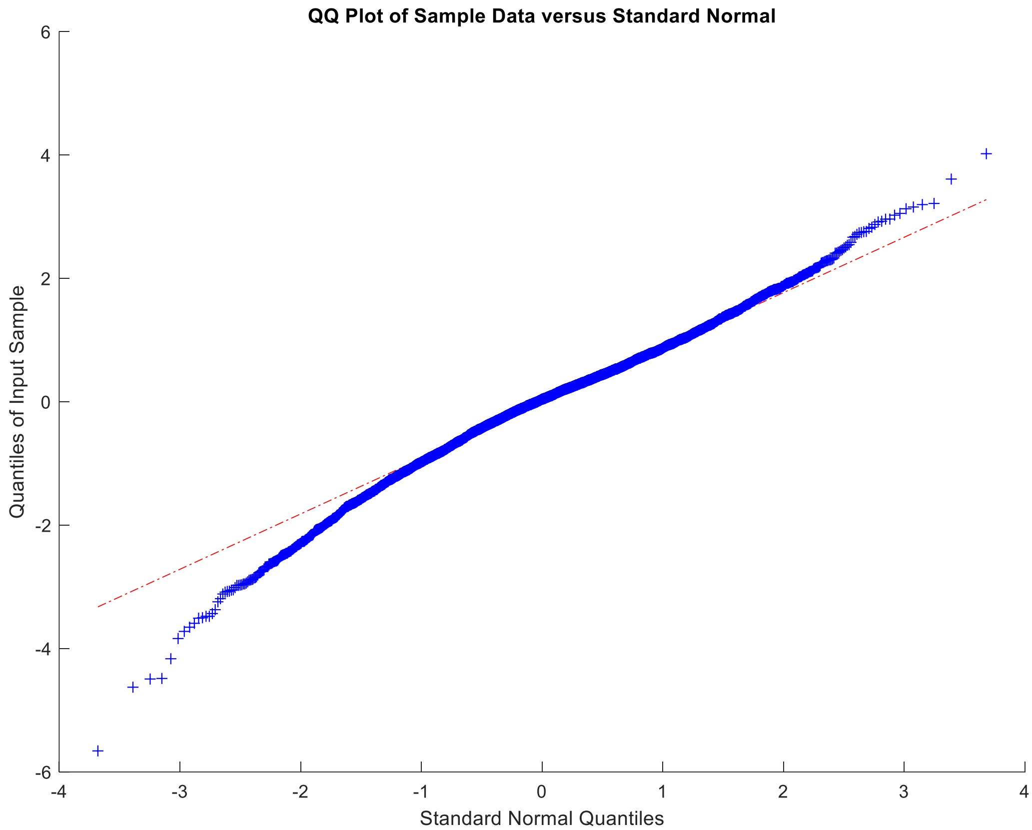

Figure 3.11: Quantile-quantile (Q-Q) plot for high-frequency aggregated returns for the SPY index after correcting for volatility clustering using the GARCH (1,1) model.

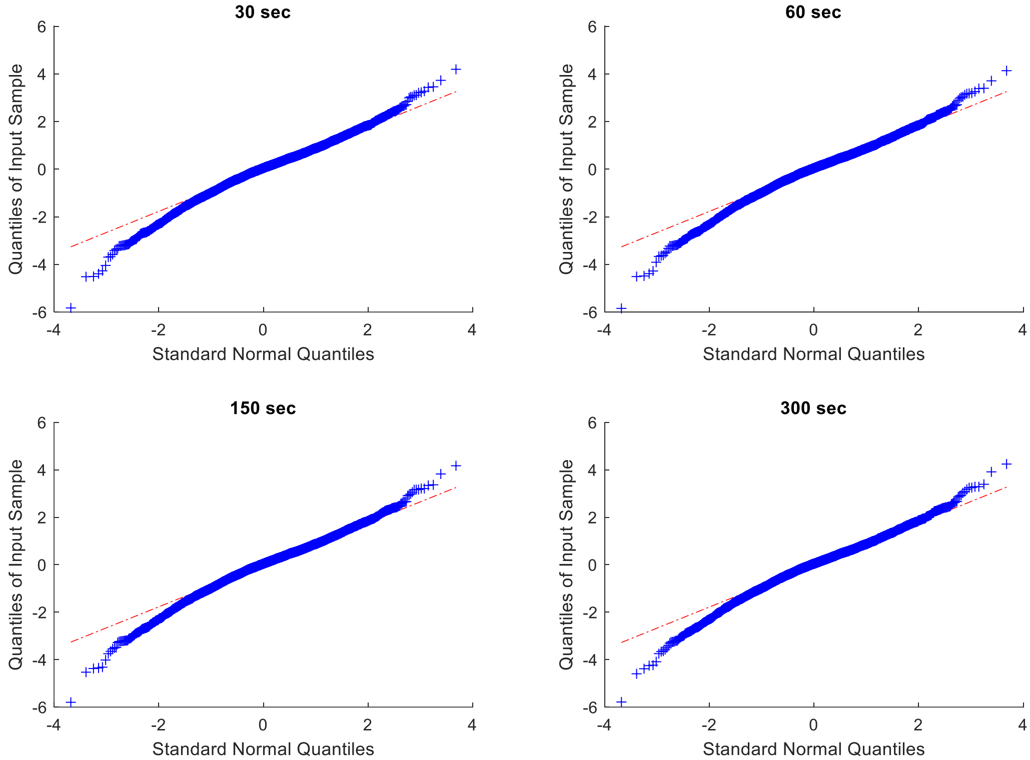

Figure 3.12: Quantile-quantile (Q-Q) plots for intraday periodicity-filtered returns for the SPY index using 30-second, 60-second, 150-second and 300-second returns, after correcting for volatility clustering using the GARCH (1,1) model. The returns are filtered using the WSD intraday periodicity estimator.

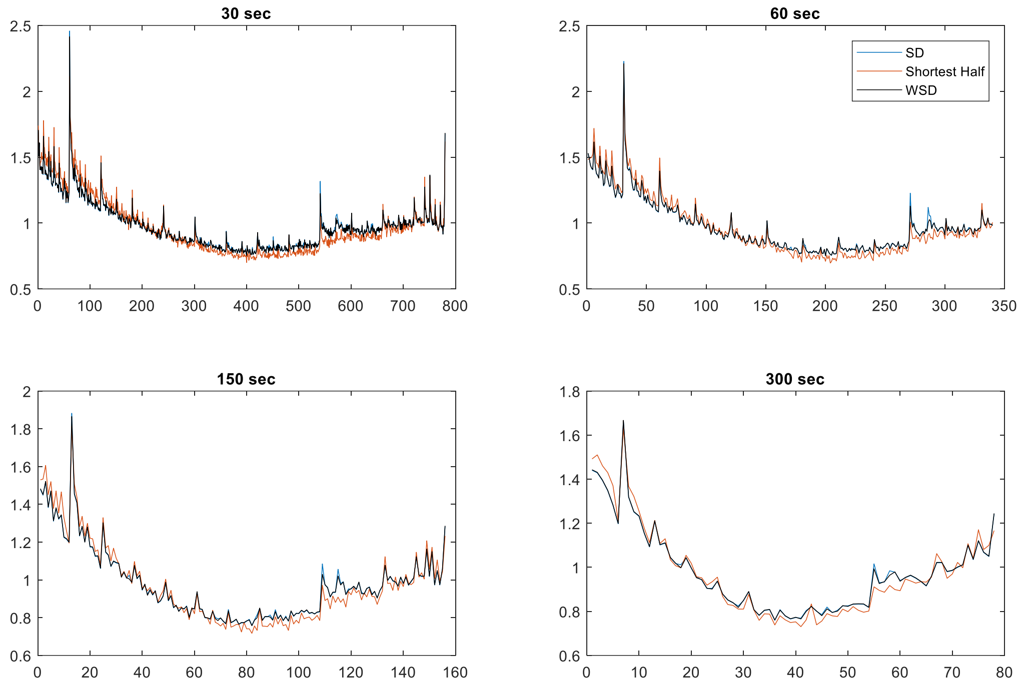

Figure 3.13: Intraday periodicity-filtered returns for the SPY index with 30-second, 60-second, 150-second and 300-second returns using the SD, Shortest Half, WSD intraday periodicity estimators.

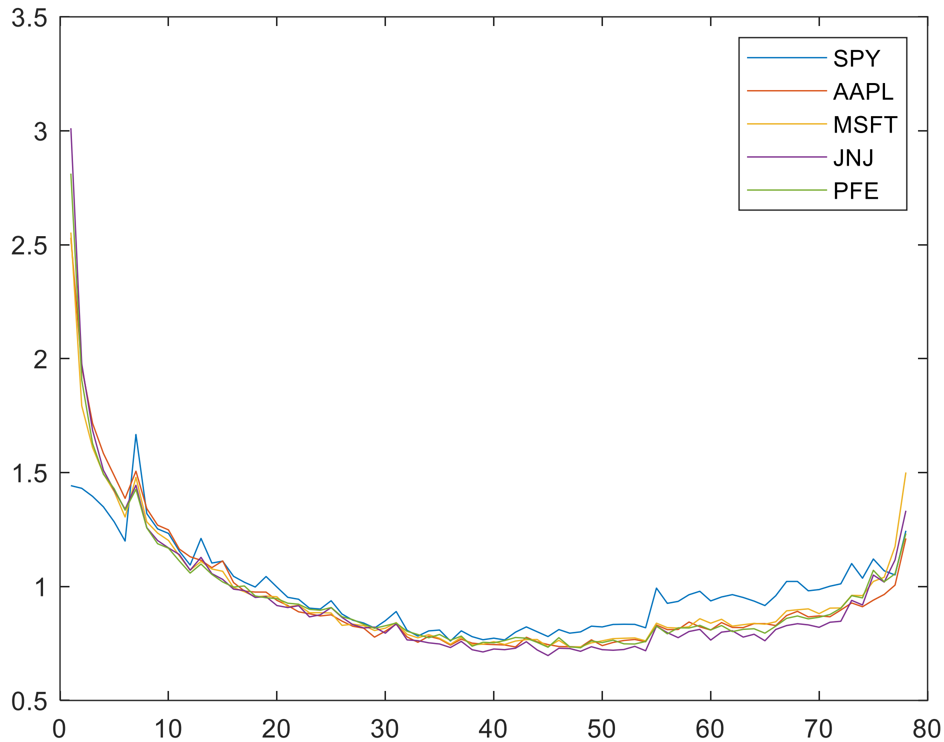

Figure 3.14: 300-second intraday periodicity-filtered returns for four stocks (AAPL, MSFT, JNJ and PFE) and the SPY index using the WSD estimator.

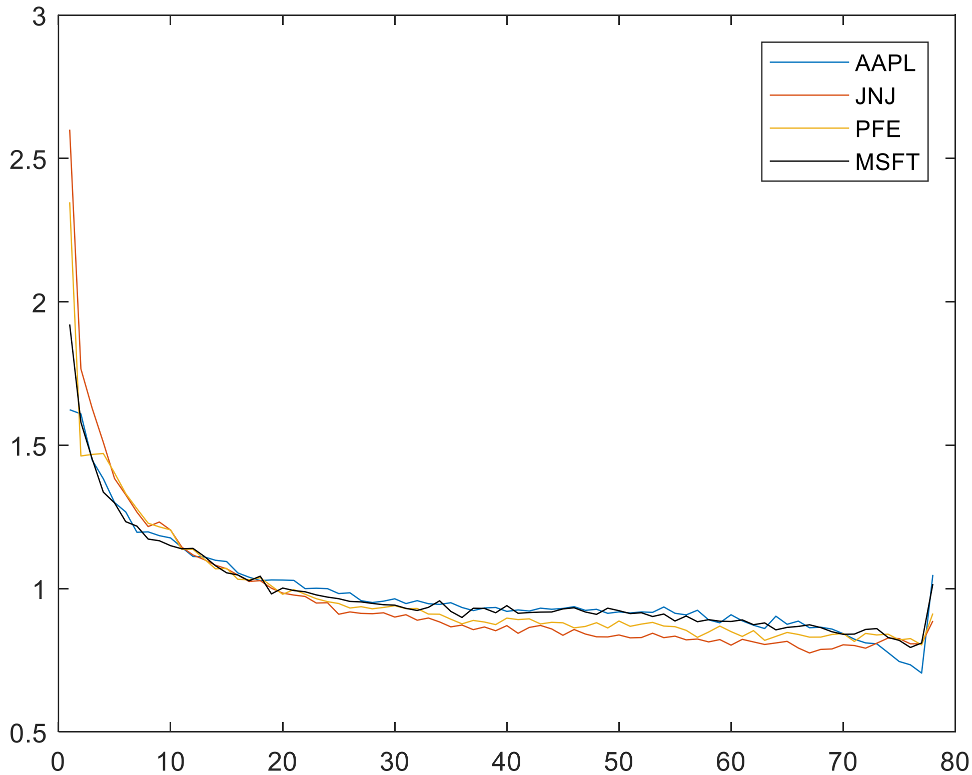

Figure 3.15: WSD Intraday periodicity estimator using 300-second business-time sampling returns for stocks.

Tables

Table 3.1 GJR (1,1) model results for SPY and MSFT

$r_{t,n}$=μ+$\varepsilon_{t,n}$ where $\varepsilon_t = \sigma_t z_t$

$\sigma^2_{t,n} = \beta_0 + \gamma_1\sigma^2_{t,n-1} + \alpha_1\varepsilon^2_{t,n-1} + \xi_1 I[\varepsilon_{t,n-1} < 0]\varepsilon^2_{t,n-1}$

	AAPL	PFE	JNJ	MSFT	SPY	AAPL	PFE	JNJ	MSFT	SPY
					Intraday periodicity unfiltered					
		Calendar time sampling (30 sec)					**Calendar time sampling (150 sec)**			
β_0	0.00001***	0.00001***	0.00001***	0.00002***	0.00000***	0.00018***	0.00021***	0.00016***	0.00019***	0.00002***
s.e.	4.5E-09	1.8E-09	1.3E-09	2.1E-08	6.6E-09	4.9E-07	2.9E-07	2.5E-07	6.4E-07	8.3E-08
γ_1	0.94037***	0.95448***	0.93168***	0.94231***	0.96164***	0.87996***	0.87132***	0.83917***	0.87277***	0.91878***
s.e.	5.6E-05	4.0E-06	1.4E-05	9.4E-05	3.4E-05	2.4E-04	1.2E-04	2.5E-04	3.9E-04	2.0E-04
α_1	0.05482***	0.03944***	0.06490***	0.05171***	0.03368***	0.10708***	0.11579***	0.14962***	0.11536***	0.06750***
s.e.	9.6E-05	4.5E-05	4.5E-05	1.2E-04	6.9E-05	4.0E-04	2.7E-04	4.8E-04	5.3E-04	2.4E-04
ξ_1	0.00962***	0.00714***	0.00333***	0.00446***	0.00936***	0.02592***	0.01535***	0.00998***	0.01250***	0.02744***
s.e.	1.2E-04	8.0E-05	7.0E-05	1.6E-04	1.2E-04	4.9E-04	4.9E-04	4.8E-04	6.7E-04	3.7E-04
		Calendar time sampling (300 sec)					**Business time sampling**			
β_0	0.00073***	0.00079***	0.00055***	0.00071***	0.00007***	0.00014***	0.00035***	0.00035***	0.00018***	0.00001***
s.e.	3.5E-06	2.3E-06	1.4E-06	4.0E-06	4.6E-07	2.1E-06	1.8E-06	1.8E-06	2.8E-06	4.2E-07
γ_1	0.83670***	0.82110***	0.77206***	0.82154***	0.90043***	0.93201***	0.91171***	0.86424***	0.93243***	0.95566***
s.e.	5.1E-04	2.3E-04	5.7E-04	7.2E-04	2.1E-04	4.0E-04	3.1E-04	5.5E-04	5.0E-04	3.8E-04
α_1	0.14391***	0.15997***	0.21274***	0.16023***	0.08102***	0.05770***	0.07065***	0.11515***	0.05745***	0.03190***
s.e.	8.2E-04	5.0E-04	1.0E-03	8.5E-04	3.9E-04	5.6E-04	4.7E-04	7.3E-04	5.8E-04	4.9E-04
ξ_1	0.03878***	0.01517***	0.01471***	0.01591***	0.03573***	0.02043***	0.01647***	0.00701***	0.00884***	0.02444***
s.e.	1.0E-03	9.6E-04	1.0E-03	1.2E-03	7.4E-04	8.7E-04	5.2E-04	8.5E-04	7.7E-04	7.1E-04
					Intraday periodicity filtered					
		Calendar time sampling (30 sec)					**Calendar time sampling (150 sec)**			
β_0	0.00000***	0.00001***	0.00000***	0.00001***	0.00000***	0.00004***	0.00004***	0.00002***	0.00004***	0.00001***
s.e.	1.2E-08	1.1E-09	6.0E-09	1.4E-08	9.2E-09	4.3E-07	4.4E-07	1.7E-07	4.6E-07	4.9E-08
γ_1	0.96039***	0.96933***	0.95876***	0.96353***	0.97074***	0.94225***	0.95184***	0.94869***	0.95577***	0.95604***
s.e.	6.1E-05	8.5E-06	3.2E-05	8.0E-05	2.4E-05	1.3E-04	7.1E-05	5.0E-05	6.6E-05	1.4E-04
α_1	0.03512***	0.02613***	0.03928***	0.03126***	0.02423***	0.04895***	0.04072***	0.04715***	0.03918***	0.03270***
s.e.	9.5E-05	5.7E-05	7.0E-05	1.0E-04	5.9E-05	2.5E-04	2.1E-04	1.8E-04	2.6E-04	1.6E-04
ξ_1	**0.00897*** **	**0.00483*** **	**0.00148*** **	**0.00433*** **	0.01007***	**0.01761*** **	**0.01064*** **	**0.00530*** **	**0.00690*** **	**0.02207*** **
s.e.	1.3E-04	1.0E-04	1.1E-04	1.5E-04	1.1E-04	3.6E-04	3.6E-04	2.6E-04	4.0E-04	2.6E-04
		Calendar time sampling (300 sec)					**Business time sampling**			
β_0	0.00013***	0.00010***	0.00006***	0.00010***	0.00002***	0.00083***	0.00008***	0.00005***	0.00008***	0.00001***
s.e.	2.1E-06	2.0E-06	8.7E+00	1.8E-06	2.0E-07	4.8E-05	2.2E-06	1.1E-06	2.5E-06	4.4E-07
γ_1	0.93282***	0.94433***	0.93818***	0.94618***	0.95253***	0.93907***	0.94537***	0.93718***	0.94751***	0.95556***
s.e.	2.9E-04	2.4E-04	2.0E-04	3.0E-04	1.8E-04	1.2E-03	3.6E-04	4.4E-04	5.1E-04	4.0E-04
α_1	0.05651***	0.04621***	0.05557***	0.04667***	0.03340***	0.05328***	0.04518***	0.05770***	0.04556***	0.03196***
s.e.	4.0E-04	4.1E-04	4.2E-04	4.9E-04	2.1E-04	1.5E-03	5.7E-04	5.8E-04	6.0E-04	5.1E-04
ξ_1	**0.02134*** **	**0.01437*** **	**0.00864*** **	**0.00978*** **	**0.02725*** **	**0.00625*** **	**0.01527*** **	**0.00612*** **	0.00996***	0.02467***
s.e.	5.5E-04	5.2E-04	5.1E-04	6.6E-04	4.6E-04	2.0E-03	5.7E-04	7.3E-04	7.0E-04	7.1E-04

Note: This table shows the coefficients (Estimate) and their standard errors (SE) estimated using the GJR (1,1) model for the SPY index and for the MSFT stock, for calendar-time and business-time sampling schemes, using 30-second, 60-second, 150-second and 300-second returns. The calendar-time results are shown for both unfiltered data and data filtered using the WSD estimator. The coefficients and their standard errors (SE) for GARCH lag (γ_1), ARCH lag (α_1) and leverage effect (ξ_1) are shown in the table. The bold leverage effect coefficients (ξ1) indicate that the leverage effects are highest when using business-time data.

Table 3.2 Average realised variance (RV) and conditional volatility for stocks and SPY (2000-2016)

Sampling Frequency (Seconds)	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}
	Realised Variance (RV)				*GARCH*			
	AAPL (IT)							
30	6.823	7.352	7.365	7.795	4.922	4.821	4.828	5.076
60	6.156	6.604	6.615	6.975		4.863	4.871	5.111
150	5.579	5.912	5.922	6.256		4.878	4.885	5.135
300	5.292	5.543	5.549	5.880		4.873	4.879	5.156
	JNJ (HC)							
30	1.603	1.664	1.640	1.639	1.086	1.086	1.069	1.048
60	1.478	1.530	1.521	1.484		1.089	1.081	1.054
150	1.432	1.476	1.469	1.465		1.100	1.095	1.089
300	1.385	1.418	1.415	1.427		1.102	1.099	1.107
	PFE (HC)							
30	3.369	3.408	3.381	3.407	1.879	1.758	1.744	1.750
60	2.809	2.840	2.827	2.895		1.792	1.784	1.805
150	2.475	2.509	2.501	2.504		1.830	1.823	1.822
300	2.332	2.361	2.357	2.389		1.854	1.850	1.873
	MSFT (IT)							
30	3.065	3.140	3.142	3.191	2.440	2.291	2.292	2.315
60	2.887	2.969	2.972	2.978		2.316	2.317	2.316
150	2.746	2.833	2.834	2.864		2.348	2.348	2.369
300	2.679	2.751	2.752	2.847		2.365	2.366	2.439
	SPY							
30	1.213	1.233	1.234	1.287	1.033	0.991	0.992	1.023
60	1.143	1.166	1.166	1.204		0.994	0.995	1.018
150	1.081	1.099	1.100	1.128		0.993	0.993	1.014
300	1.037	1.051	1.052	1.073		0.995	0.996	1.014

Note. This table shows the realised volatility (RV) and conditional variance (GARCH(1,1)) of four stocks (AAPL, JNJ, PFE and MSFT) and SPY from 2000 to 2016, with and without filtering for intraday periodicity, using 30-second, 60-second, 150-second and 300-second returns. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The r column for GARCH only has one value per stock because the aggregated daily returns used for the GARCH (1,1) are same across different frequencies, which result in the same conditional variance.

Table 3.3 Number of intraday jumps estimated using the Lee-Mykland intraday jump test (2000-2016)

Sampling Frequency (Seconds)	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}
	AAPL (IT)			
30	4692	2231	2293	2239
60	3173	1241	1250	1240
150	1486	475	490	476
300	704	292	304	293
	JNJ (HC)			
30	6134	2827	3336	2845
60	3929	1357	1451	1354
150	1747	411	429	411
300	869	226	227	226
	PFE (HC)			
30	3982	1880	2070	1897
60	2850	1052	1124	1050
150	1390	366	364	364
300	709	195	203	195
	MSFT (IT)			
30	3604	1644	1682	1640
60	2613	949	979	952
150	1246	315	315	314
300	554	162	168	163
	SPY			
30	2484	1983	2062	1987
60	1555	1113	1194	1122
150	641	460	497	464
300	300	242	254	245

Note: This table shows that the number of estimated intraday jumps using the Lee-Mykland (2008) jump test for four stocks (AAPL, JNJ, PFE and MSFT) and SPY before and after filtering for intraday periodicity, using 30-second, 60-second, 150-second and 300-second returns. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively.

Table 3.4 Contribution of continuous components to QV (%) for stocks and SPY (2000-2016)

Sampling Frequency (Seconds)	*BV*				*TBV*				*CTBV*			
	r	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**	**r**	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**	**r**	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**
						AAPL (IT)						
30	91.011	87.902	87.675	87.898	75.195	73.490	72.995	73.489	79.453	76.250	75.869	76.236
60	94.421	93.850	93.742	93.856	78.724	**79.452**	**79.064**	**79.429**	86.189	85.810	85.226	85.723
150	96.787	**97.629**	**97.536**	**97.628**	80.200	**83.847**	**83.370**	**83.799**	90.681	**93.381**	**93.081**	**93.345**
300	96.620	**98.081**	**97.966**	**98.073**	77.619	**84.930**	**84.498**	**84.853**	90.339	**94.238**	**93.903**	**94.293**
						JNJ (HC)						
30	87.465	86.713	86.195	86.694	65.624	**68.714**	**67.721**	**68.668**	68.872	**71.279**	**70.001**	**71.240**
60	91.296	**92.899**	**92.672**	**92.878**	69.824	**75.675**	**75.192**	**75.554**	76.795	**82.452**	**81.924**	**82.361**
150	94.838	**97.435**	**97.324**	**97.420**	72.785	**82.208**	**81.894**	**82.271**	83.583	**91.730**	**91.390**	**91.711**
300	95.132	**98.370**	**98.299**	**98.365**	72.067	**84.323**	**83.825**	**84.222**	84.104	**94.237**	**94.111**	**94.264**
						PFE (HC)						
30	86.052	**86.668**	**86.246**	**86.631**	69.359	**73.002**	**72.282**	**72.957**	71.243	**74.429**	**73.859**	**74.340**
60	89.950	**91.733**	**91.334**	**91.689**	72.619	**77.424**	**77.002**	**77.434**	79.166	**83.466**	**82.782**	**83.393**
150	93.969	**96.589**	**96.474**	**96.563**	75.168	**83.359**	**82.983**	**83.280**	86.224	**92.350**	**92.059**	**92.318**
300	93.951	**97.583**	**97.511**	**97.581**	72.974	**84.263**	**83.912**	**84.223**	85.647	**93.072**	**93.266**	**93.066**
						MSFT (IT)						
30	91.843	91.754	91.629	91.759	77.574	**79.336**	**78.962**	**79.325**	81.907	**83.231**	**82.866**	**83.196**
60	95.430	**95.889**	**95.779**	**95.872**	81.564	**83.846**	**83.439**	**83.831**	89.030	**90.722**	**90.356**	**90.719**
150	96.725	**97.997**	**97.972**	**97.989**	82.626	**87.622**	**87.578**	**87.666**	91.884	**95.510**	**95.359**	**95.493**
300	96.296	**98.493**	**98.396**	**98.490**	80.569	**88.486**	**88.284**	**88.490**	91.478	**96.396**	**96.334**	**96.385**
						SPY						
30	97.193	93.328	93.124	93.325	87.215	82.567	82.157	82.506	92.740	87.838	87.426	87.834
60	98.242	97.692	97.594	97.681	89.924	89.140	88.649	89.061	95.342	**94.491**	**94.191**	**94.463**
150	98.099	**98.532**	**98.445**	**98.519**	87.805	**88.866**	**88.638**	**88.775**	94.270	**94.960**	**94.696**	**94.898**
300	97.839	**98.371**	**98.322**	**98.371**	86.796	**88.173**	**87.652**	**88.149**	93.562	**93.985**	**93.778**	**93.959**

Note: The table reports the estimated percentage of the contribution of the continuous components to QV before and after filtering for intraday periodicity using different realised measures (BV, TBV and CTBV) and sampling frequencies (30-second, 60-second, 150-second and 300-second returns) for four stocks (AAPL, JNJ, PFE and MSFT) and SPY for the whole data set (2000-2016). The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The values in bold text indicate cases where the percentage of continuous components is larger after filtering for intraday periodicity.

Table 3.5 MSE results for HAR-family models (2000-2016) (forecast horizon h=1)

Sampling Frequency (Seconds)	r				r/f_{SD}				r/f_{Short}				r/f_{WSD}			
	HAR-RV	*HAR-Vol*	*HAR-J*	*HAR-TJ*	*HAR-RV*	*HAR-Vol*	*HAR-J*	*HAR-TJ*	*HAR-RV*	*HAR-Vol*	*HAR-J*	*HAR-TJ*	*HAR-RV*	*HAR-Vol*	*HAR-J*	*HAR-TJ*
	AAPL (IT)															
30	0.345	**0.344**	0.418	0.363	**0.412**	0.435	0.507	0.465	**0.475**	0.500	0.595	0.541	**0.415**	0.438	0.518	0.471
60	0.401	**0.377**	0.391	0.443	0.481	**0.476**	0.515	0.547	0.548	**0.532**	0.587	0.637	0.486	**0.477**	0.520	0.554
150	0.507	**0.489**	0.508	0.543	0.550	**0.542**	0.550	0.587	0.621	**0.609**	0.621	0.668	0.552	**0.543**	0.551	0.588
300	0.693	**0.648**	0.697	0.708	0.724	**0.675**	0.729	0.718	0.807	**0.747**	0.813	0.798	0.725	**0.676**	0.730	0.706
	JNJ (HC)															
30	0.594	0.582	0.616	**0.580**	0.199	**0.196**	0.206	0.218	0.178	**0.175**	0.185	0.192	0.194	**0.191**	0.200	0.211
60	0.368	**0.359**	0.372	**0.359**	0.116	**0.112**	0.124	0.132	0.112	**0.109**	0.119	0.128	0.115	**0.111**	0.123	0.130
150	0.366	0.350	0.366	**0.340**	0.119	**0.114**	0.120	0.134	0.116	**0.111**	0.117	0.127	0.118	**0.113**	0.119	0.131
300	0.405	0.383	0.410	**0.372**	0.131	**0.126**	0.131	0.136	0.129	**0.123**	0.128	0.135	0.131	**0.125**	0.130	0.135
	PFE (HC)															
30	0.859	**0.815**	0.899	0.892	0.483	**0.474**	0.489	0.544	0.477	**0.468**	0.486	0.531	0.477	**0.467**	0.483	0.537
60	0.969	**0.937**	0.965	0.989	0.505	**0.495**	0.511	0.539	0.497	**0.487**	0.503	0.528	0.501	**0.491**	0.507	0.536
150	1.461	1.422	1.464	**1.375**	0.646	**0.632**	0.653	0.730	0.638	**0.625**	0.646	0.689	0.642	**0.629**	0.649	0.717
300	1.443	1.429	1.436	**1.352**	0.775	**0.763**	0.780	0.816	0.762	**0.748**	0.768	0.777	0.771	**0.758**	0.775	0.811
	MSFT (IT)															
30	0.339	**0.331**	0.333	0.375	0.283	0.274	**0.271**	0.337	0.289	0.281	**0.279**	0.347	0.283	**0.275**	0.272	0.338
60	0.390	**0.378**	0.394	0.414	0.337	**0.325**	0.349	0.378	0.338	**0.325**	0.351	0.383	0.339	**0.326**	0.351	0.379
150	0.431	**0.422**	0.435	0.488	0.413	**0.402**	0.418	0.445	0.421	**0.410**	0.426	0.455	0.413	**0.403**	0.418	0.447
300	0.589	0.582	**0.570**	0.610	0.522	**0.512**	**0.512**	0.528	0.547	**0.536**	0.539	0.546	0.523	**0.513**	0.514	0.529
	SPY															
30	0.038	**0.035**	0.040	0.049	0.037	**0.035**	0.037	0.048	0.042	**0.038**	0.041	0.053	0.038	**0.035**	0.037	0.048
60	0.041	**0.039**	0.044	0.050	0.041	**0.039**	0.044	0.049	0.044	**0.042**	0.047	0.053	0.041	**0.039**	0.044	0.050
150	0.048	**0.046**	0.049	0.053	0.043	**0.042**	0.044	0.048	0.045	**0.044**	0.046	0.051	0.043	**0.042**	0.044	0.049
300	**0.045**	**0.045**	0.046	0.047	0.044	**0.043**	0.045	0.045	0.047	**0.045**	0.047	0.048	0.044	**0.043**	0.045	0.045

Note: This table shows the mean squared errors (MSE) for forecasting filtered and unfiltered RVs using HAR-class models (HAR-RV, HAR-Vol, HAR-J and HAR-TJ models) for four stocks and SPY. RVs are sampled using 30-second, 60-second, 150-second and 300-second returns. The panels for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The bold MSEs indicate the best HAR class model (i.e. the lowest MSE) for that particular stock and sampling frequency.

Table 3.6 Number of intraday jumps estimated using the Lee-Mykland intraday jump test for different financial regimes

Sampling Frequency (Seconds)	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}
	Pre-crisis				*Crisis*				*Post-crisis*			
					AAPL (IT)							
30	2395	1454	1529	1458	183	85	89	88	2114	692	675	693
60	1440	732	731	727	149	59	60	60	1584	450	459	453
150	630	228	238	227	85	35	38	35	771	212	214	214
300	284	133	136	133	35	18	19	18	385	141	149	142
					JNJ (HC)							
30	2604	1738	1867	1745	549	203	267	206	2980	886	1202	894
60	1471	765	810	763	390	112	127	110	2067	480	514	481
150	590	217	231	215	216	39	40	37	941	155	158	159
300	336	117	117	117	98	15	14	15	435	94	96	94
					PFE (HC)							
30	1778	1145	1204	1158	284	108	126	109	1920	627	740	630
60	1115	620	659	620	230	77	77	76	1505	355	388	354
150	543	234	230	232	111	19	22	20	736	113	112	112
300	300	120	128	120	54	13	13	13	355	62	62	62
					MSFT (IT)							
30	1291	821	850	821	194	95	95	95	2119	728	737	724
60	843	468	477	470	168	59	58	58	1602	422	444	424
150	365	158	162	158	83	17	16	16	798	140	137	140
300	159	77	85	78	38	11	11	11	357	74	72	74
					SPY							
30	995	859	899	859	162	140	146	140	1327	984	1017	988
60	574	478	522	482	106	84	92	85	875	551	580	555
150	251	198	217	200	41	26	31	26	349	236	249	238
300	128	104	109	105	20	14	16	14	151	124	128	126

Note: This table shows the number of estimated intraday jumps using the Lee-Mykland (2008) jump test for four stocks (AAPL, JNJ, PFE and MSFT) and SPY before and after filtering for intraday periodicity, using 30-second, 60-second, 150-second and 300-second returns. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively.

Table 3.7 Contribution of continuous components to QV (%) for stocks and SPY (pre-crisis)

Sampling Frequency (Seconds)	*BV*				*TBV*				*CTBV*			
	r	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**	**r**	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**	**r**	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**
						AAPL (IT)						
30	88.558	84.623	84.366	84.619	72.364	69.905	69.382	69.909	75.632	71.486	71.072	71.506
60	92.869	91.976	91.860	91.983	76.491	**76.537**	76.116	**76.535**	84.038	82.825	82.271	82.731
150	96.336	**97.199**	**97.062**	**97.199**	79.596	**82.593**	**82.073**	**82.560**	90.967	**93.249**	**92.915**	**93.207**
300	96.251	**97.747**	**97.598**	**97.741**	76.940	**84.098**	**83.824**	**84.075**	89.993	**94.431**	**93.959**	**94.416**
						JNJ (HC)						
30	83.586	82.102	81.743	82.089	63.316	63.203	62.528	63.193	64.590	63.883	63.170	63.868
60	88.686	**89.965**	**89.770**	**89.960**	68.682	**71.412**	**71.098**	**71.328**	74.742	**76.904**	**76.487**	**76.787**
150	94.380	**96.593**	**96.522**	**96.570**	75.315	**80.978**	**80.614**	**80.988**	87.267	**91.059**	**90.793**	**91.094**
300	95.547	**98.274**	**98.274**	**98.268**	75.378	**83.971**	**83.574**	**83.825**	88.466	**95.303**	**95.268**	**95.299**
						PFE (HC)						
30	83.778	**84.463**	**84.110**	**84.417**	66.972	**70.280**	**69.685**	**70.224**	68.149	**70.980**	**70.304**	**70.887**
60	87.822	**89.728**	**89.354**	**89.688**	70.156	**74.431**	**74.022**	**74.426**	75.831	**79.605**	**79.142**	**79.532**
150	93.355	**95.581**	**95.467**	**95.547**	74.272	**80.930**	**80.597**	**80.887**	84.752	**90.759**	**90.483**	**90.741**
300	93.392	**96.942**	**96.899**	**96.949**	72.008	**81.701**	**81.400**	**81.670**	85.152	**91.967**	**91.915**	**91.975**
						MSFT (IT)						
30	90.917	90.540	90.427	90.548	77.310	**78.115**	**77.766**	**78.103**	80.783	**81.405**	**81.100**	**81.386**
60	94.981	**94.996**	94.878	94.973	81.675	**82.227**	**81.852**	**82.168**	88.716	**88.954**	**88.586**	**88.955**
150	96.829	**97.603**	**97.593**	**97.594**	83.547	**86.414**	**86.268**	**86.447**	92.352	**94.676**	**94.494**	**94.677**
300	96.472	**98.295**	**98.169**	**98.287**	82.311	**88.129**	**87.550**	**88.109**	92.566	**95.643**	**95.557**	**95.604**
						SPY						
30	97.639	90.159	89.901	90.148	87.864	79.615	79.247	79.580	93.859	84.345	83.948	84.331
60	98.048	96.768	96.641	96.752	89.869	88.305	87.861	88.201	95.120	93.453	93.144	93.418
150	98.306	**98.673**	**98.589**	**98.659**	88.860	**89.602**	**89.244**	**89.575**	94.271	**94.929**	**94.625**	**94.830**
300	97.888	**98.296**	**98.244**	**98.290**	86.880	**88.092**	**87.807**	**88.028**	92.983	**93.018**	**92.771**	**92.955**

Note: The table reports the estimated percentage of the contribution of the continuous components to QV before and after filtering for intraday periodicity using different realised measures (BV, TBV and CTBV) and sampling frequencies (30-second, 60-second, 150-second and 300-second returns) for four stocks (AAPL, JNJ, PFE and MSFT) and SPY for the pre-crisis period. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The values in bold text indicate cases where the percentage of continuous components is larger after filtering for intraday periodicity.

Table 3.8 Contribution of continuous components to QV (%) for stocks and SPY (crisis)

Sampling Frequency (Seconds)	*BV*				*TBV*				*CTBV*			
	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}
					AAPL (IT)							
30	98.936	**99.197**	**99.194**	**99.198**	87.526	**88.867**	**88.686**	**88.938**	94.091	**95.447**	**95.557**	**95.367**
60	98.729	**99.661**	**99.564**	**99.665**	89.224	**91.552**	**91.743**	**91.509**	94.208	**97.565**	**96.990**	**97.555**
150	98.803	**99.069**	**99.136**	**99.065**	85.745	**89.280**	**88.983**	**89.256**	90.987	**94.784**	**94.574**	**94.749**
300	98.335	**99.230**	**99.188**	**99.205**	83.698	**87.568**	**86.362**	**87.229**	92.926	**93.393**	**93.628**	**93.858**
					JNJ (HC)							
30	93.484	**96.146**	**95.307**	**96.111**	72.826	**80.046**	**78.555**	**79.926**	79.708	**87.455**	**84.652**	**87.388**
60	93.818	**97.486**	**97.224**	**97.395**	74.807	**83.137**	**82.562**	**82.913**	81.810	**91.458**	**91.042**	**91.478**
150	95.582	**98.957**	**98.784**	**98.944**	73.283	**84.715**	**84.699**	**85.112**	81.924	**94.619**	**94.180**	**94.564**
300	94.652	**98.363**	**98.070**	**98.365**	76.727	**86.277**	**85.542**	**86.243**	86.288	**94.202**	**94.002**	**94.381**
					PFE (HC)							
30	89.133	**91.253**	**90.494**	**91.224**	74.586	**79.031**	**78.026**	**79.084**	76.910	**81.785**	**82.231**	**81.642**
60	93.018	**95.502**	**95.037**	**95.471**	77.659	**83.994**	**83.835**	**84.098**	84.719	**91.357**	**90.184**	**91.241**
150	94.747	**98.332**	**98.236**	**98.319**	77.202	**87.736**	**87.268**	**87.759**	90.373	**96.122**	**95.899**	**96.102**
300	93.514	**97.960**	**97.831**	**97.954**	76.802	**88.586**	**88.303**	**88.548**	87.541	**95.138**	**94.957**	**95.130**
					MSFT (IT)							
30	94.141	**94.624**	**94.533**	**94.636**	81.007	**83.209**	**82.765**	**83.211**	87.266	**89.016**	**88.672**	**88.902**
60	96.974	**97.519**	**97.452**	**97.517**	85.109	**88.027**	**87.399**	**88.047**	92.483	**94.564**	**94.161**	**94.570**
150	97.363	**98.673**	**98.649**	**98.663**	85.888	**91.752**	**91.808**	**91.693**	94.570	**97.441**	**97.332**	**97.351**
300	97.359	**99.230**	**99.145**	**99.229**	84.784	**90.538**	**90.654**	**90.542**	94.272	**99.121**	**99.009**	**99.120**
					SPY							
30	97.747	97.664	97.573	97.674	89.545	88.534	88.208	88.474	94.564	94.089	93.771	94.090
60	99.113	98.984	98.926	98.977	93.286	92.790	92.233	92.716	97.226	96.520	96.313	96.525
150	98.477	**98.695**	**98.589**	**98.680**	89.481	**90.161**	**90.365**	**89.926**	96.630	**96.914**	**96.809**	**96.922**
300	97.858	**98.375**	**98.313**	**98.386**	88.292	**89.289**	**88.658**	**89.287**	94.720	**95.319**	**95.233**	**95.351**

Note: The table reports the estimated percentage of the contribution of the continuous components to QV before and after filtering for intraday periodicity using different realised measures (BV, TBV and CTBV) and sampling frequencies (30-second, 60-second, 150-second and 300-second returns) for four stocks (AAPL, JNJ, PFE and MSFT) and SPY for the crisis period. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The values in bold text indicate cases where the percentage of continuous components is larger after filtering for intraday periodicity.

Table 3.9 Contribution of continuous components to QV (%) for stocks and SPY (post-crisis)

Sampling Frequency (Seconds)	*BV*				*TBV*				*CTBV*			
	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}
						AAPL (IT)						
30	98.220	**98.440**	**98.318**	**98.429**	80.169	**81.715**	**81.202**	**81.600**	88.270	**88.529**	88.019	**88.349**
60	98.647	**98.930**	**98.865**	**98.924**	80.330	**83.983**	**83.160**	**83.850**	89.916	**91.159**	**90.369**	**91.012**
150	97.104	**98.474**	**98.455**	**98.477**	77.650	**84.932**	**84.488**	**84.771**	88.821	**92.557**	**92.343**	**92.555**
300	96.753	**98.598**	**98.579**	**98.595**	74.839	**86.468**	**86.050**	**86.399**	89.451	**94.142**	**93.906**	**94.121**
						JNJ (HC)						
30	93.165	**94.035**	93.127	**93.990**	66.355	**77.417**	**75.637**	**77.295**	72.081	**82.168**	**80.085**	**82.048**
60	95.630	**97.671**	**97.311**	**97.658**	68.771	**81.938**	**80.938**	**81.789**	77.902	**91.199**	**90.120**	**91.083**
150	95.358	**98.572**	**98.390**	**98.577**	66.397	**83.628**	**83.197**	**83.561**	76.097	**91.223**	**90.739**	**91.075**
300	94.522	**98.640**	**98.552**	**98.633**	61.411	**83.742**	**83.142**	**83.710**	72.750	**91.333**	**91.022**	**91.327**
						PFE (HC)						
30	89.643	89.411	89.022	89.394	71.691	**76.116**	**75.229**	**76.028**	75.066	**78.584**	**77.518**	**78.546**
60	92.392	**93.678**	**93.286**	**93.617**	74.331	**79.611**	**78.975**	**79.592**	82.434	**86.732**	**85.911**	**86.698**
150	94.676	**97.519**	**97.387**	**97.499**	75.510	**85.460**	**85.055**	**85.221**	86.168	**93.022**	**92.645**	**92.948**
300	95.412	**98.680**	**98.592**	**98.660**	72.148	**86.543**	**86.035**	**86.484**	85.272	**93.903**	**94.911**	**93.869**
						MSFT (IT)						
30	92.169	**92.458**	**92.273**	**92.453**	75.460	**79.204**	**78.821**	**79.185**	80.261	**83.038**	**82.498**	**83.033**
60	95.184	**96.706**	**96.576**	**96.691**	78.449	**84.255**	**83.959**	**84.320**	86.937	**91.801**	**91.467**	**91.778**
150	95.963	**98.400**	**98.337**	**98.398**	77.830	**87.026**	**87.153**	**87.193**	88.602	**95.920**	**95.811**	**95.922**
300	95.039	**98.363**	**98.329**	**98.375**	73.168	**87.626**	**88.121**	**87.693**	86.725	**95.972**	**96.016**	**96.024**
						SPY						
30	95.411	**95.528**	95.308	**95.529**	82.675	82.056	81.404	81.928	87.719	**88.307**	87.716	**88.322**
60	97.525	**97.989**	**97.900**	**97.982**	85.710	**86.080**	85.559	**86.051**	93.380	**94.049**	**93.640**	**93.992**
150	97.190	**98.018**	**97.948**	**98.013**	83.501	**85.584**	**85.031**	**85.553**	91.205	**92.388**	**91.978**	**92.307**
300	97.722	**98.514**	**98.489**	**98.513**	84.723	**86.858**	**86.011**	**86.885**	93.185	**94.130**	**93.842**	**94.104**

Note: The table reports the estimated percentage of the contribution of the continuous components to QV before and after filtering for intraday periodicity using different realised measures (BV, TBV and CTBV) and sampling frequencies (30-second, 60-second, 150-second and 300-second returns) for four stocks (AAPL, JNJ, PFE and MSFT) and SPY for the post-crisis period. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The values in bold text indicate cases where the percentage of continuous components is larger after filtering for intraday periodicity.

Table 3.10 MSE for volatility forecasting using HAR-class models for the pre-crisis period (forecast horizon h=1))

Sampling Frequency (Seconds)	r				r/f_{SD}				r/f_{Short}				r/f_{WSD}			
	HAR-RV	**HAR-Vol**	**HAR-J**	**HAR-TJ**	**HAR-RV**	**HAR-Vol**	**HAR-J**	**HAR-TJ**	**HAR-RV**	**HAR-Vol**	**HAR-J**	**HAR-TJ**	**HAR-RV**	**HAR-Vol**	**HAR-J**	**HAR-TJ**
								AAPL (IT)								
30	25.412	25.244	24.931	26.644	41.605	41.258	40.046	43.115	52.901	52.489	50.859	55.163	42.349	42.001	40.769	43.922
60	32.222	32.140	32.282	32.924	53.341	53.180	53.065	55.241	67.618	67.446	67.175	69.973	54.440	54.281	54.138	56.589
150	19.439	19.380	19.590	20.015	32.488	32.369	32.596	33.473	40.070	39.932	40.202	41.135	33.256	33.137	33.367	34.273
300	22.514	22.459	22.654	22.879	37.106	37.021	37.243	36.717	46.088	45.992	46.298	45.724	37.587	37.501	37.725	37.157
								JNJ (HC)								
30	0.203	0.204	0.208	0.208	0.181	0.183	0.195	0.207	0.183	0.185	0.195	0.208	0.176	0.178	0.190	0.202
60	0.183	0.184	0.180	0.178	0.169	0.174	0.180	0.182	0.161	0.166	0.170	0.175	0.167	0.173	0.178	0.182
150	0.190	0.197	0.186	0.192	0.181	0.198	0.179	0.190	0.182	0.198	0.180	0.192	0.180	0.196	0.177	0.190
300	0.229	0.234	0.237	0.239	0.156	0.182	0.156	0.165	0.160	0.187	0.160	0.171	0.155	0.181	0.156	0.164
								PFE (HC0								
30	0.417	0.449	0.429	0.423	0.330	0.372	0.334	0.401	0.344	0.382	0.342	0.416	0.329	0.370	0.332	0.401
60	0.370	0.397	0.360	0.343	0.349	0.371	0.347	0.365	0.383	0.401	0.383	0.402	0.348	0.370	0.347	0.364
150	0.377	0.403	0.379	0.380	0.385	0.411	0.392	0.400	0.393	0.418	0.401	0.405	0.383	0.409	0.390	0.396
300	0.404	0.446	0.403	0.440	0.466	0.505	0.474	0.518	0.486	0.522	0.493	0.526	0.464	0.502	0.471	0.514
								MSFT (IT)								
30	0.956	0.958	0.978	0.959	1.297	1.298	1.309	1.331	1.436	1.437	1.460	1.470	1.307	1.308	1.321	1.344
60	1.097	1.099	1.090	1.075	1.480	1.486	1.482	1.536	1.570	1.576	1.557	1.629	1.506	1.511	1.507	1.560
150	1.083	1.084	1.101	1.085	1.708	1.715	1.713	1.723	1.863	1.870	1.869	1.840	1.727	1.735	1.733	1.739
300	0.820	0.824	0.785	0.809	1.091	1.097	1.081	1.136	1.233	1.239	1.226	1.277	1.094	1.100	1.085	1.140
								SPY								
30	0.289	0.290	0.287	0.249	0.330	0.335	0.329	0.313	0.384	0.392	0.383	0.368	0.332	0.338	0.332	0.316
60	0.289	0.295	0.281	0.288	0.339	0.347	0.334	0.339	0.399	0.408	0.391	0.404	0.348	0.356	0.340	0.347
150	0.280	0.284	0.277	0.268	0.341	0.345	0.339	0.339	0.385	0.390	0.382	0.371	0.344	0.348	0.342	0.341
300	0.347	0.351	0.341	0.340	0.446	0.447	0.445	0.435	0.488	0.488	0.486	0.475	0.450	0.451	0.448	0.439
Mean	5.356	**5.346**	5.349	5.497	8.689	8.665	**8.612**	8.906	10.836	10.806	**10.731**	11.116	8.847	8.822	**8.767**	9.079

Note: This table shows the mean squared errors (MSE) for forecasting filtered and unfiltered RVs using HAR-class models (HAR-RV, HAR-Vol, HAR-J and HAR-TJ models) for four stocks and SPY for the pre-crisis period. RVs are sampled using 30-second, 60-second, 150-second and 300second returns. The panels for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The bold mean MSEs indicate the best HAR class model (i.e. the lowest MSE) for each estimator.

Table 3.11 MSE for volatility forecasting using HAR-class models for the crisis period (forecast horizon h=1)

Sampling Frequency (Seconds)	r				r/f_{SD}				r/f_{Short}				r/f_{WSD}			
	HAR-RV	**HAR-Vol**	**HAR-J**	**HAR-TJ**	**HAR-RV**	**HAR-Vol**	**HAR-J**	**HAR-TJ**	**HAR-RV**	**HAR-Vol**	**HAR-J**	**HAR-TJ**	**HAR-RV**	**HAR-Vol**	**HAR-J**	**HAR-TJ**
	AAPL (IT)															
30	214.990	217.930	223.160	242.720	206.750	209.760	205.180	240.620	232.940	236.160	231.840	289.450	208.200	211.190	206.630	239.530
60	200.140	202.550	196.050	228.960	214.870	218.350	214.720	217.190	243.100	246.910	245.860	247.110	215.970	219.460	215.810	218.380
150	210.260	212.250	210.670	396.010	283.510	288.550	289.010	360.030	330.030	336.510	337.100	417.170	285.230	290.340	290.810	361.460
300	225.020	228.910	228.350	464.690	378.260	386.730	386.750	617.420	453.600	464.230	465.270	727.340	378.910	387.410	387.430	610.270
	JNJ (HC)															
30	140.110	189.780	190.470	80.699	76.142	92.745	78.535	111.630	62.328	72.294	63.853	109.160	73.630	89.667	75.900	108.210
60	152.300	176.670	100.410	92.168	66.273	70.853	52.154	71.673	60.275	64.135	51.532	67.391	64.928	69.340	51.497	71.147
150	370.180	461.180	381.440	399.050	115.100	137.310	129.080	105.000	105.010	124.500	109.450	98.642	113.410	135.340	125.750	104.910
300	34.924	36.706	36.743	63.502	42.143	43.997	42.894	53.780	42.748	44.562	43.821	53.639	41.932	43.800	42.685	53.730
	PFE (HC)															
30	92.671	91.432	100.680	109.970	64.917	65.115	74.631	56.294	62.782	63.150	72.046	56.385	63.797	63.988	73.336	55.195
60	90.769	90.616	92.076	117.170	55.661	56.310	56.985	81.338	55.581	56.249	57.004	63.276	55.018	55.660	56.336	80.085
150	152.680	152.150	162.620	79.129	73.019	73.695	72.222	58.376	71.044	71.652	70.361	58.165	72.304	72.973	71.550	57.796
300	50.583	51.072	50.901	72.843	49.010	49.376	49.067	63.123	49.883	50.273	49.967	63.366	48.849	49.212	48.903	62.962
	MSFT (IT)															
30	60.574	61.021	67.164	65.587	53.946	54.285	52.888	56.491	53.975	54.342	52.893	56.532	53.925	54.258	52.871	56.395
60	71.233	72.457	72.900	81.730	62.013	62.406	63.272	66.956	62.826	63.174	64.045	72.051	62.049	62.440	63.299	66.952
150	68.301	68.617	70.822	62.770	62.555	62.973	62.239	69.572	63.772	64.220	63.437	70.950	62.478	62.899	62.124	69.514
300	53.120	53.199	54.137	53.200	58.791	58.875	59.177	62.156	64.147	64.222	64.647	68.074	58.784	58.870	59.173	62.157
	SPY															
30	54.953	59.644	58.038	78.839	48.405	50.100	50.423	67.199	50.799	51.191	52.992	67.608	48.316	49.934	50.433	66.479
60	75.664	82.272	76.036	84.630	59.592	59.951	60.075	65.251	60.294	59.690	60.707	65.664	59.295	59.599	59.781	64.985
150	70.049	74.387	74.957	58.919	53.794	53.599	56.576	59.461	54.194	53.505	56.898	71.948	53.720	53.545	56.522	59.241
300	44.317	44.259	45.731	40.457	39.467	37.721	39.799	39.307	40.722	38.792	41.373	41.043	39.453	37.715	39.776	39.500
Mean	**121.642**	131.355	124.668	143.652	**103.211**	106.635	104.784	126.143	**111.003**	113.988	112.755	138.248	**103.010**	106.382	104.531	125.445

Note: This table shows the mean squared errors (MSE) for forecasting filtered and unfiltered RVs using HAR-class models (HAR-RV, HAR-Vol, HAR-J and HAR-TJ models) for four stocks and SPY for the crisis period. RVs are sampled using 30-second, 60-second, 150-second and 300-second returns. The panels for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The bold mean MSEs indicate the best HAR class model (i.e. the lowest MSE) for each estimator.

Table 3.12 MSE for volatility forecasting using HAR-class models for the post-crisis period (forecast horizon h=1)

Sampling Frequency (Seconds)	r				r/f_{SD}				r/f_{Short}				r/f_{WSD}			
	HAR-RV	HAR-Vol	HAR-J	HAR-TJ	HAR-RV	HAR-Vol	HAR-J	HAR-TJ	HAR-RV	HAR-Vol	HAR-J	HAR-TJ	HAR-RV	HAR-Vol	HAR-J	HAR-TJ
AAPL (IT)																
30	0.522	0.293	0.528	0.371	0.687	0.360	0.688	0.475	0.801	0.416	0.799	0.547	0.697	0.363	0.695	0.479
60	0.547	0.321	0.547	0.389	0.715	0.395	0.868	0.520	0.844	0.448	1.041	0.592	0.730	0.398	0.890	0.526
150	0.507	0.401	0.516	0.477	0.594	0.432	0.602	0.526	0.676	0.487	0.689	0.592	0.596	0.433	0.604	0.525
300	0.646	0.565	0.644	0.621	0.671	0.562	0.664	0.635	0.750	0.620	0.741	0.699	0.672	0.562	0.665	0.635
JNJ (HC)																
30	0.583	0.559	0.586	0.590	0.192	0.187	0.194	0.190	0.171	0.165	0.171	0.171	0.187	0.181	0.188	0.186
60	0.366	0.339	0.378	0.368	0.111	0.105	0.111	0.116	0.108	0.102	0.108	0.113	0.111	0.105	0.111	0.115
150	0.364	0.330	0.365	0.379	0.116	0.106	0.116	0.148	0.112	0.103	0.112	0.134	0.115	0.105	0.115	0.141
300	0.408	0.364	0.407	0.374	0.130	0.117	0.130	0.137	0.129	0.114	0.128	0.134	0.130	0.116	0.130	0.135
PFE (HC)																
30	0.857	0.752	0.866	0.814	0.448	0.423	0.449	0.474	0.444	0.420	0.445	0.464	0.442	0.417	0.444	0.467
60	0.953	0.902	0.946	0.980	0.488	0.452	0.493	0.519	0.481	0.446	0.482	0.506	0.484	0.449	0.489	0.515
150	1.446	1.389	1.471	1.375	0.630	0.571	0.632	0.667	0.626	0.563	0.628	0.630	0.627	0.569	0.629	0.659
300	1.409	1.342	1.472	1.309	0.796	0.681	0.797	0.774	0.785	0.668	0.788	0.724	0.791	0.677	0.793	0.769
MSFT (IT)																
30	0.332	0.329	0.332	0.362	0.271	0.272	0.281	0.282	0.278	0.279	0.287	0.289	0.271	0.272	0.282	0.283
60	0.384	0.372	0.384	0.396	0.321	0.320	0.324	0.327	0.321	0.320	0.326	0.326	0.322	0.321	0.325	0.328
150	0.418	0.411	0.430	0.410	0.395	0.392	0.393	0.393	0.402	0.399	0.408	0.401	0.395	0.392	0.393	0.393
300	0.577	0.568	0.601	0.603	0.497	0.493	0.555	0.498	0.520	0.517	0.576	0.524	0.498	0.494	0.556	0.500
SPY																
30	0.031	0.035	0.033	0.032	0.032	0.036	0.036	0.034	0.037	0.040	0.041	0.038	0.033	0.036	0.036	0.034
60	0.034	0.038	0.035	0.035	0.035	0.037	0.036	0.036	0.038	0.041	0.039	0.039	0.035	0.037	0.036	0.036
150	0.041	0.047	0.042	0.041	0.037	0.042	0.038	0.038	0.040	0.044	0.041	0.040	0.037	0.042	0.038	0.038
300	0.042	0.054	0.043	0.042	0.039	0.047	0.040	0.039	0.041	0.049	0.042	0.042	0.039	0.047	0.040	0.039
Mean	0.523	**0.471**	0.531	0.498	0.360	**0.302**	0.372	0.341	0.380	**0.312**	0.395	0.350	0.361	**0.301**	0.373	0.340

Note: This table shows the mean squared errors (MSE) for forecasting filtered and unfiltered RVs using HAR-class models (HAR-RV, HAR-Vol, HAR-J and HAR-TJ models) for four stocks and SPY for the post-crisis period. RVs are sampled using 30-second, 60-second, 150-second and 300-second returns. The panels for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The bold mean MSEs indicate the best HAR class model (i.e. the lowest MSE) for each estimator.

Table 3.13 Average realised variance (RV) and conditional variance for business-time sampling data (2000-2016)

	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}
	Realised Variance (RV)				***GARCH***			
AAPL (IT)	5.290	5.469	5.470	5.576	4.912	4.866	4.867	4.981
JNJ (HC)	1.373	1.445	1.436	1.384	1.089	1.121	1.113	1.065
PFE (HC)	2.416	2.444	2.422	2.359	1.879	1.878	1.860	1.803
MSFT (IT)	2.689	2.783	2.783	2.756	2.442	2.453	2.453	2.437
SPY	1.041	1.049	1.049	1.053	1.033	1.057	1.057	1.066

Note: This table shows the realised volatility (RV) and conditional variance (GARCH(1,1)) for four stocks (AAPL, JNJ, PFE and MSFT) and SPY using 300-second business-time sampling data for the whole data set (2000-2016), with and without filtering for intraday periodicity. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively.

Table 3.14 Number of intraday jumps estimated using the Lee-Mykland intraday jump test for business-time sampling data

	r	r/f_{SD}	r/f_{Short}	r/f_{WSD}
		2000-2016		
AAPL (IT)	103	51	53	51
JNJ (HC)	608	132	191	134
PFE (HC)	430	131	190	134
MSFT (IT)	182	43	49	43
SPY	62	56	57	54
		Pre-crisis		
AAPL (IT)	46	25	26	25
JNJ (HC)	225	71	97	73
PFE (HC)	226	98	124	99
MSFT (IT)	42	24	26	24
SPY	33	30	33	30
		Crisis		
AAPL (IT)	7	3	3	3
JNJ (HC)	77	7	18	7
PFE (HC)	31	6	6	5
MSFT (IT)	10	1	1	1
SPY	2	2	1	0
		Post-crisis		
AAPL (IT)	50	23	24	23
JNJ (HC)	306	54	76	54
PFE (HC)	173	27	60	30
MSFT (IT)	130	18	22	18
SPY	27	24	23	24

Note: This table shows the number of estimated intraday jumps using the Lee-Mykland (2008) jump test for four stocks (AAPL, JNJ, PFE and MSFT) and SPY before and after filtering for intraday periodicity for the whole data set (2000-2016) and in different financial regimes. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively.

Table 3.15 Contribution of continuous components to QV (%) for stocks and SPY using business-time sampling data for different financial regimes

	BV				*TBV*				*CTBV*			
	r	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**	**r**	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**	**r**	**r/f_{SD}**	**r/f_{Short}**	**r/f_{WSD}**
2000-2016												
AAPL	98.769	**99.090**	**98.639**	**99.090**	92.671	**94.164**	**93.027**	**94.134**	96.727	**97.291**	96.714	**97.278**
JNJ	96.310	**98.803**	**98.300**	**98.793**	79.102	**88.804**	**87.591**	**88.759**	90.818	**96.935**	**96.023**	**96.923**
PFE	94.452	**98.024**	**97.080**	**97.941**	79.329	**88.336**	**86.400**	**88.246**	88.224	**94.667**	**93.755**	**94.560**
MSFT	98.689	**99.232**	**99.136**	**99.231**	92.311	**95.479**	**95.173**	**95.481**	97.275	**98.430**	**98.257**	**98.430**
SPY	99.534	**99.577**	99.524	**99.573**	95.417	**95.583**	**95.647**	**95.576**	98.004	**98.040**	97.951	**98.021**
Pre-crisis												
AAPL	98.781	**99.113**	**98.886**	**99.113**	93.357	**94.603**	**93.538**	**94.580**	97.577	**98.298**	**97.933**	**98.298**
JNJ	96.143	**98.393**	**97.838**	**98.381**	79.404	**87.181**	**86.000**	**87.125**	91.676	**95.633**	**94.689**	**95.618**
PFE	93.004	**97.370**	**96.194**	**97.319**	75.937	**85.066**	**83.457**	**84.990**	86.182	**93.818**	**92.280**	**93.740**
MSFT	98.852	**99.016**	**99.065**	**99.013**	94.054	**95.105**	**94.983**	**95.113**	97.730	**98.095**	**98.050**	**98.094**
SPY	99.194	**99.296**	**99.208**	**99.289**	95.206	**95.466**	**95.357**	**95.449**	97.520	**97.525**	**97.403**	97.503
Crisis												
AAPL	99.577	**99.773**	**99.696**	**99.772**	95.167	**97.544**	**96.838**	**97.543**	99.150	**99.740**	**99.456**	**99.739**
JNJ	96.368	**99.539**	**99.143**	**99.538**	80.591	**92.478**	**90.990**	**92.460**	91.996	**99.386**	**98.481**	**99.384**
PFE	95.626	**99.258**	**98.312**	**99.107**	86.293	**93.754**	**91.251**	**93.688**	94.070	**96.778**	**97.799**	**96.575**
MSFT	98.838	**99.551**	**99.137**	**99.550**	92.305	**96.098**	**95.720**	**96.076**	97.728	**98.765**	**98.410**	**98.765**
SPY	99.825	**99.804**	99.802	99.804	96.338	**96.133**	**96.555**	**96.136**	99.760	**99.769**	99.659	**99.770**
Post-crisis												
AAPL	97.926	**98.294**	96.401	**98.296**	87.030	**88.657**	86.782	**88.575**	90.395	89.989	88.147	89.906
JNJ	96.686	**99.352**	**98.884**	**99.339**	77.165	**90.333**	**89.176**	**90.293**	87.714	**98.576**	**97.683**	**98.559**
PFE	96.571	**98.489**	**98.016**	**98.387**	81.169	**91.234**	**88.914**	**91.091**	88.106	**94.871**	**93.753**	**94.774**
MSFT	98.190	**99.503**	**99.316**	**99.503**	88.306	**95.890**	**95.177**	**95.897**	95.853	**98.989**	**98.650**	**98.990**
SPY	99.785	**99.797**	99.743	**99.797**	94.613	**95.072**	**94.980**	**95.069**	96.618	**96.703**	**96.701**	**96.664**

Note: The table reports the estimated percentage of the contribution of the continuous components to QV before and after filtering for intraday periodicity using different realised measures (BV, TBV and CTBV) for four stocks (AAPL, JNJ, PFE and MSFT) and SPY for the whole data set (2000-2016) and for different financial regimes. The columns for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The values in bold text indicate cases where the percentage of continuous components is larger after filtering for intraday periodicity.

Table 3.16 MSE results for HAR-class models with business-time sampling data (forecast horizon h=1)

	r				r/f_{SD}				r/f_{Short}				r/f_{WSD}			
HAR-	**RV**	**Vol**	**J**	**TJ**	**RV**	**Vol**	**J**	**TJ**	**RV**	**Vol**	**J**	**TJ**	**RV**	**Vol**	**J**	**TJ**
								2000-2016								
AAPL	0.545	0.455	0.545	0.456	0.588	0.484	0.592	0.494	0.614	0.500	0.618	0.506	0.591	0.486	0.595	0.496
JNJ	0.311	0.307	0.314	0.318	0.157	0.152	0.159	0.166	0.157	0.153	0.158	0.155	0.156	0.151	0.157	0.164
PFE	1.316	1.316	1.331	1.482	1.224	1.225	1.232	1.352	1.195	1.196	1.206	1.232	1.201	1.202	1.211	1.327
MSFT	0.471	0.471	0.475	0.501	0.440	0.440	0.441	0.445	0.429	0.429	0.429	0.432	0.440	0.440	0.441	0.445
SPY	0.045	0.045	0.045	0.044	0.046	0.046	0.046	0.044	0.046	0.046	0.046	0.045	0.046	0.046	0.046	0.044
								Pre-crisis								
AAPL	28.181	28.082	28.224	28.168	35.580	35.465	35.637	35.196	37.176	37.055	37.211	37.229	35.613	35.498	35.670	35.248
JNJ	0.206	0.205	0.200	0.205	0.177	0.174	0.181	0.207	0.168	0.166	0.170	0.194	0.175	0.172	0.179	0.204
PFE	0.478	1.477	0.495	0.439	0.391	0.684	0.393	0.411	0.369	0.651	0.372	0.384	0.383	0.668	0.385	0.402
MSFT	1.197	1.662	1.195	1.211	1.453	1.978	1.456	1.474	1.417	1.902	1.420	1.431	1.452	1.981	1.455	1.473
SPY	0.321	0.321	0.323	0.330	0.340	0.341	0.344	0.348	0.344	0.344	0.348	0.353	0.341	0.341	0.344	0.349
								Crisis								
AAPL	170.670	171.370	174.380	204.980	199.320	200.020	203.550	217.440	205.830	206.560	210.100	223.300	199.340	200.040	203.590	217.560
JNJ	100.260	100.150	104.430	59.004	95.079	95.091	95.925	82.769	82.646	82.667	83.755	72.636	93.285	93.297	94.123	81.435
PFE	127.840	129.210	133.900	60.261	155.360	156.690	155.270	76.304	154.330	155.560	154.590	684.580	151.640	152.940	150.940	75.260
MSFT	77.754	77.472	78.601	60.536	65.593	65.238	66.115	71.641	64.282	63.974	64.539	70.149	65.578	65.224	66.090	71.568
SPY	87.175	87.163	87.172	47.248	91.195	91.181	91.210	47.089	92.441	92.428	92.456	47.940	91.071	91.057	91.086	47.062
								Post-crisis								
AAPL	0.743	0.717	0.750	0.435	0.850	0.821	0.856	0.473	0.896	0.864	0.900	0.497	0.855	0.827	0.861	0.474
JNJ	0.293	0.293	0.293	0.314	0.150	0.150	0.150	0.157	0.148	0.148	0.147	0.150	0.148	0.149	0.149	0.156
PFE	1.309	1.309	1.313	1.478	1.211	1.211	1.203	1.358	1.180	1.180	1.180	1.217	1.188	1.188	1.183	1.337
MSFT	0.462	0.462	0.465	0.528	0.424	0.424	0.423	0.439	0.415	0.414	0.414	0.431	0.424	0.424	0.423	0.439
SPY	0.043	0.043	0.043	0.041	0.044	0.044	0.044	0.042	0.045	0.045	0.044	0.043	0.044	0.045	0.044	0.042
Mean	**29.981**	30.127	30.725	23.399	32.481	32.593	32.761	**26.892**	**32.206**	32.314	32.505	57.145	32.199	32.309	32.449	**26.774**

Note: This table shows the mean squared errors (MSE) for forecasting filtered and unfiltered RVs using HAR-class models (HAR-RV, HAR-Vol, HAR-J and HAR-TJ models) for four stocks (AAPL, PFE, JNJ, MSFT) and SPY with business-time sampling data for the whole data set (2000-2016) and for different financial regimes. RVs are sampled using 30-second, 60-second, 150-second and 300-second returns. The panels for r, r/f_{SD}, r/f_{Short} and r/f_{WSD} show the results for the unfiltered returns and the returns filtered using the SD, Shortest Half and WSD estimators respectively. The bold mean MSEs indicate the best HAR class model (i.e. the lowest MSE) for each estimator.

Chapter 4 – The impact of co-jumps and news announcements on stock volatility

4.1 Introduction

This chapter discusses how co-jumps and news affect the continuous and jump components of stocks and the realised volatility of forecasting returns before and after the 2008 financial crisis. Previous literature has investigated how the announcement of news stories has an influence on financial markets, particularly their first-moment responses to these stories (e.g. Ederington & Lee 1993; Balduzzi et al. 2001; Andersen & Bollerslev 1998). Recent work focuses on the market's second-moment responses to macro-economic news (e.g. Andersen & Bollerslev, 1998; Balduzzi et al., 2001; Huang 2018). However, no studies to date have looked at the impact of news announcements on non-parametric volatility forecasting models such as HAR-family models. Furthermore, the effect of news announcements on the volatility of stocks from different industrial sectors, and how the market's second-moment responses may vary in different financial regimes, are yet to be investigated. Therefore, in this chapter we look at the effect of the market's second-moment responses to news announcements on stock price volatility before and after the 2008 financial crisis.

Jumps are significant discontinuities in financial variables (Lee & Mykland 2007). High volatility caused by notable changes in financial asset returns within a given day is known as the jump component, and the remaining variation in returns in that day is known as the continuous component. Some of this volatility occurs between multiple assets at the same time, known as co-jumps. Co-jumps are closely related to macroeconomic news announcements (e.g. Lahaye et al, 2011; Dungey & Hvozdyk, 2012; Chatrath et al., 2014) and are caused by the co-movements of volatilities between different stocks, which is why we also investigate the impact of co-jump-related news surprise on jump components. Bollerslev et al. (2008) argue that jumps in stocks are generated by firm-specific news or market-level news. Co-jumps, in particular, are jumps generated by market-level news across different stocks. The mean cross-product (MCP) co-jump test, introduced by Bollerslev et al. (2008), tests for the presence of co-jumps. However, there is no evidence for a direct relationship between co-jumps, news and single-stock jumps. Hence, we examine how news affects the jump components for realised volatility, as well as the relationship between (i) co-jumps and jumps and (ii) co-jumps and market-level news. We investigate the extent to which co-jumps are preceded by macro-economic news announcements and whether co-jumps capture all of the volatility co-movement caused by news announcements. To further examine the impact of news announcements on volatility, we also investigate how news announcements, particularly those that co-occur with co-jumps, affect the performance of HAR models, which are commonly used to forecast realised volatility.

In this thesis, we first investigate how news affects jump components, using three different estimation methods –bi-power variation, threshold bi-power variation and corrected threshold bi-power variation (Barndorff-Nielsen & Shephard 2004; Corsi et al. 2010). Second, we focus on how co-jumps that are caused by market-level news affect realised volatility. Finally, we examine how macro-economic news and co-jumps affect forecasting in HAR models.

4.2 Data

4.2.1 Data subsets used in the analysis

The announcement of major macroeconomic news typically results in high trading volumes for stocks, as investors respond to the news by rearranging their investments. Therefore, stocks with high trading volumes are important to consider in this analysis. Macroeconomic news announcements also cause around one third of the jumps in stock volatility and contribute to a notable proportion of jump components (Evans, 2011). In order to focus on jumps caused by macroeconomic announcements, we examine returns from stocks with few intraday jumps, as this minimises the impact of jumps caused by other factors.

[Insert Tables 4.1 and 4.2 here]

The stocks included in the data set are those with either a high volume of jumps or relatively few jumps from five industrial sectors: Healthcare (HC), Information Technology (IT), Telecommunication Services (TS), Consumer Staples (CS) and Utilities from the S&P 100 index. We use two subsets of stocks from these five sectors in our analysis. The stocks with the six highest volumes before and after the financial crisis are included in the first data subset; these include AAPL, MSFT, CSCO, INTC, ORCL and GILD, and INTC, VOD, CSCO, MSFT, IBM and YHOO. INTC is the lowest jump-frequency stock in the pre-crisis period in Table 4.1, followed by VOD, CSCO, MSFT, IBM and YHOO. The lowest jump-frequency stock in the post-crisis period (Table 4.2) is VOD, followed by stocks BT, TEF, UL, AAPL and XRX. These stocks are also included in the first subset.

As the stocks with the six highest volumes and lowest jump frequencies are mainly from the IT sector (i.e. AAPL, CSCO, INTC, MSFT, ORCL, IBM, XRX, YHOO), we consider the stocks with the highest volume and lowest jump frequency from each sector separately as a second subset in order to examine how the news and co-jumps affect high-volume and low-jump-frequency stocks across different sectors. The stocks with the highest volumes in each sector before and after the 2008 financial crisis are GILD, AAPL, T, KO and EXC, and PFE, AAPL, T, KO and EXC respectively. As shown in Tables 4.1 and 4.2, the stocks with the fewest jumps in each sector before and after the crisis are AMGN, INTC, VOD, KO and AEE, and BSX, AAPL, VOD, UL and CNP respectively.

Therefore, as shown in Table 4.3, 21 stocks are considered here: GILD, PFE, AMGN and BSX from HC; AAPL, INTC, MSFT, CSCO, ORCL, IBM, YHOO and XRX from IT; T, VOD, BT and, TEF from TS; KO and UL from CS; and EXC, AEE and CNP from Utilities.

[Insert Table 4.3 here]

We use five-minute high-frequency data because studies such as Liu et al. (2015) find little evidence that five-minute realised variance outperforms other volatility measures such as 1- and 5-second realised kernels. Data from 01/01/2000 to 31/12/2007 indicate the pre-crisis period, while data from 01/07/2009 to 30/12/2016 is the post-crisis period.

4.2.2 News outlets used in the analysis

[Insert Table 4.4 here]

As seen in Table 4.4, 62 news outlets were used in the analysis. Each outlet was included in the regression as both a positive and a negative news surprise calculated according to equation (4.1), which meant that 124 independent variables were initially used in the regression in equation (4.2) before non-significant variables were removed. Many of the news outlets, such as Initial Jobless Claims (INJCJC) and Change in Nonfarm Payrolls (NFP TCH), have been used in previous research (e.g. Balduzzi et al. 2001; Huang 2018), but we also include other news outlets not previously studied that are closely related to economic performance (e.g. the University of Michigan Consumer Sentiment Index [CONSSENT]). The survey data and released values for the 63 news outlets were obtained from Bloomberg (2018) for the period from 2000 to 2016.

Table 4.4 also shows the descriptive statistics for the absolute values of positive and negative news surprise for each news outlet. The news surprise values are calculated from survey data and released values for news, which can be used to describe the positive and negative impact of each news outlet on the financial market. In the regression analysis in this chapter, the absolute values of positive and negative news surprise are used to examine the impact of news on stock volatility. Further details regarding these calculations are described in Section 4.3.

The descriptive statistics in Table 4.4 reveal that the mean absolute values for positive and negative news surprise range between 0.55 and 1.60. The majority of the maximum absolute values are smaller than 5. Thus, the descriptive statistics do not show large differences between absolute positive and negative news surprise for each news outlet.

4.3 Methodology

4.3.1 News surprise

In order to examine the relationship between news announcements and market response, we first follow the method from Balduzzi et al. (2001) to calculate the standardised news surprise by using release values and survey data for news. This offers a mathematical measure of how 'surprising' a particular news item is. Standardised news surprise can be calculated using equation (4.1):

$$S_{kt}=\frac{A_{kt}-E_{kt}}{\hat{\sigma}_k} \tag{4.1}$$

where A_{kt} is the released value for a news outlet k at a point in time t, and E_{kt} is the median of its survey forecast values. $\hat{\sigma}_k$ is the sample standard deviation of $A_{kt} - E_{kt}$ and is used to standardise the news surprises in order to make the calculated surprises for different news outlets more comparable.

In order to find out how news affects the market's second-moment response for each stock, we follow the regressions in Huang (2018) and regress the jump components estimated using each stock on the news surprises calculated from equation (4.1). The regression for jump components on individual news outlets is shown as:

$$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t \tag{4.2}$$

where J_t are the jump components for each stock and S_{kt} is the standardised news surprise for news outlet *k* at time *t*. The jump components are estimated using realised variance (RV) and corrected threshold bi-power variation (CTBV) and are defined as $\hat{J}_t = I_{\{z_t > \phi_\alpha\}} \cdot \max\left[(\mathrm{RV_t} - \mathrm{CTBV_t}), 0\right]$. Positive and negative news surprises have asymmetrical effects on financial markets, so we consider them as separate variables in the regression. Second-moment market responses are non-negative, so the absolute value of news surprise is considered in equation (4.2).

We also aim to examine the impact of same-sign (positive or negative) news surprises on jump components. The regression of jump components that co-occur with co-jumps on all news can be described as:

$$log(J_t^c + 1) = \beta_{J,p}|S_t^+| + \beta_{J.n}\ |S_t^-| + \varepsilon_t \tag{4.3}$$

Where J_t^c denotes the jump components in days with co-jumps and S_t^+ and S_t^- refers to the same-sign (positive or negative) news surprise for all news outlets on day *t*, calculated based on S_{kt} in equation (4.1). $S_t^+ = \sum_{k=1}^{k=n} |S_{kt}| 1(S_{kt} \geq 0)$ and $S_t^- = \sum_{k=1}^{k=n} |S_{kt}| 1(S_{kt} < 0)$ represent the aggregated impact of positive and negative news surprise respectively on day *t*. *n* refers to the number of news outlets. We consider the aggregated impact of news surprise on a given day *t* instead of individual news outlets in order to avoid the issue of sparse data given that the number of days with co-jumps is relatively small. The error terms in these two equations follow ε_t ~ i.i.d. N $(0,\sigma^2)$.

4.3.2 HAR models

The heterogeneous autoregressive (HAR) models used in this study are shown in equations (4.4) to (4.6). The standard HAR model (Corsi, 2009) is shown in equation (4.4). The HAR model with TBV-related jump components (HAR-TJ) in equation (4.5) is developed by Corsi et al. (2010), who also proposes the HAR model with CTBV-related jump components (HAR-CTJ), as shown in equation (4.6).

$$\mathrm{RV}_{t:t+h-1}=\beta_0 + \beta_d \mathrm{RV}_{t-1} + \beta_w \mathrm{RV}_{t-5:t-1} + \beta_m \mathrm{RV}_{t-22:t-1} + \varepsilon_t \tag{4.4}$$

$$\mathrm{RV}_{t:t+h-1}=\beta_0 + \beta_d \mathrm{RV}_{t-1} + \beta_w \mathrm{RV}_{t-5:t-1} + \beta_m \mathrm{RV}_{t-22:t-1} + \beta_j \widehat{TJ}_{t-1} + \varepsilon_t \tag{4.5}$$

$$\mathrm{RV}_{t:t+h-1}=\beta_0 + \beta_d \mathrm{RV}_{t-1} + \beta_w \mathrm{RV}_{t-5:t-1} + \beta_m \mathrm{RV}_{t-22:t-1} + \beta_j \widehat{CTJ}_{t-1} + \varepsilon_t \tag{4.6}$$

Where $\mathrm{RV}_{t_1:t_2} = \frac{1}{t_2-t_1+1}\sum_{t=t_1}^{t_2} RV_t$,with $t_1 \leq t_2$.

$\mathrm{RV}_{t-1}, \mathrm{RV}_{t-5:t-1}$ and $\mathrm{RV}_{t-22:t-1}$ are the daily, weekly and monthly lags of realised volatility, which are used to capture the long-memory dynamic dependence of RV. The jump parts in (4.5) and (4.6) are estimated based on bi-power variation and threshold bi-power variation, which are expressed as $\widehat{TJ}_t = I_{\{z_t > \phi_\alpha\}} \cdot \max\left[(\mathrm{RV}_t - \mathrm{TBV}_t), 0\right]$ and $\widehat{CTJ}_t = I_{\{z_t > \phi_\alpha\}} \cdot \max\left[(\mathrm{RV}_t - \mathrm{CTBV}_t), 0\right]$ respectively. The error terms are independent and identically distributed (i.i.d.) random variables with mean 0 and variance σ^2.

Based on the HAR-TJ and HAR-CTJ models, we propose a new HAR model that separates the jump components into two parts: (i) a part with jumps affected by significant news or captured by co-jumps; and (ii) a part which is not affected by significant news or captured by co-jumps. The co-jumps are estimated based on the MCP co-jump test (Bollerslev et al. 2008). The same separation method is used for RV in the HAR model. By separating the jumps related to the news announcements from the other jumps, the new models have an additional variable which encompasses only the jump components that are related to news, reflecting the seasonal patterns of macroeconomic news announcements.

The new models can be written as follows:

$$\mathrm{RV}_{\mathrm{t:t+h-1}} = \beta_0 + \beta_{\mathrm{d1}}\mathrm{RV}_{\mathrm{t-1}}1(S_t^A > 0) + \beta_{\mathrm{d2}}\mathrm{RV}_{\mathrm{t-1}}1(S_t^A = 0) + \beta_{\mathrm{w}}\mathrm{RV}_{\mathrm{t-5:t-1}} + \beta_{\mathrm{m}}\mathrm{RV}_{\mathrm{t-22:t-1}} + \varepsilon_{\mathrm{t}} \tag{4.7}$$

$$\mathrm{RV}_{\mathrm{t:t+h-1}} = \beta_0 + \beta_{\mathrm{d}}\mathrm{RV}_{\mathrm{t-1}} + \beta_{\mathrm{w}}\mathrm{RV}_{\mathrm{t-5:t-1}} + \beta_{\mathrm{m}}\mathrm{RV}_{\mathrm{t-22:t-1}} + \beta_{j1}\widehat{TJ}_{t-1}1(S_t^A > 0) + \beta_{j2}\widehat{TJ}_{t-1}1(S_t^A = 0) + \varepsilon_{\mathrm{t}} \tag{4.8}$$

$$\mathrm{RV}_{\mathrm{t:t+h-1}} = \beta_0 + \beta_{\mathrm{d}}\mathrm{RV}_{\mathrm{t-1}} + \beta_{\mathrm{w}}\mathrm{RV}_{\mathrm{t-5:t-1}} + \beta_{\mathrm{m}}\mathrm{RV}_{\mathrm{t-22:t-1}} + \beta_{j}\widehat{CTJ}_{t-1}1(S_t^A > 0) + \beta_{j2}\widehat{CTJ}_{t-1}1(S_t^A = 0) + \varepsilon_{\mathrm{t}} \tag{4.9}$$

Where $S_t^A = S_t^+ + S_t^+ = \sum_{k=1}^{k=n} |S_{kt}|$ refers to the aggregated impact of positive and negative news surprise on day *t* for each stock, as well as the news surprise that leads to jumps. *n* refers to the number of news outlets. The error terms are i.i.d. random variables with mean 0 and variance σ^2. Instead of considering the positive and negative news surprise separately as in equations (4.2) and (4.3), we use S_t^A , defined as the days where news surprise is present (both negative and positive) in equations (4.8) and (4.9). This is because this method can better separate the seasonal patterns of jumps that co-occur with news announcements. Therefore, the jump components $\widehat{TJ}_{t-1}1(\, S_t^A > 0)$ and $\widehat{CTJ}_{t-1}1(\, S_t^A > 0)$ are those that are related to news announcements and share their seasonal patterns.

4.4 Results

4.4.1 Descriptive statistics for stocks

4.4.1.1 Comparison of stocks with and without extreme jumps

This section discusses the descriptive statistics for the two data subsets: that is, the six stocks with either the highest volumes or the lowest number of jumps for the pre-crisis and post-crisis periods (subset one); and the stocks which have the highest trading volume and the fewest jumps in each of the five industrial sectors before and after the financial crisis (subset two). It is important to observe stocks' trading volumes and jump frequencies as they are closely related to stock volatility. The descriptive statistics for trading volume and jumps help to classify the differences between stocks, so that the effect of news announcements on stocks with different properties can be analysed in the following sections.

Previous literature suggests that macroeconomic news affects trading volumes and volatility spikes of financial assets. For example, Balduzzi et al. (2001) find that macroeconomic news increases volatility and trading volume for Treasury bonds. The spike in volatility caused by the announcements observed in the Treasury bonds are the discontinuous part of the stock returns (i.e. jumps). Therefore, it is worthwhile to look at the impact of news announcements on the trading volumes of stocks from different sectors by considering the stocks with the highest trading volumes overall (subset one) and per sector (subset two). Stocks with few jumps (i.e. those that are less sensitive to shocks) are also considered, as they allow us to find out whether the primary cause of jumps is the announcement of macroeconomic news.

The impact of news can also be compared between less sensitive stocks and sensitive stocks with high trading volumes. Studies suggest that asset returns show different levels of volatility in different financial regimes. For example, Manda (2010) finds that the average value of the VIX volatility measures from 2005 to 2009 are the highest during the 2008 financial crisis from 17-Mar-08 to 31-Mar-09, followed by the post-crisis period (1-Apr-09 to 30-Nov-09) and pre-crisis period (3-Jan-05 to 16-Mar-08). Our analysis considers a longer period of time either side of the crisis, including the burst of the dot-com bubble period from 2000 to 2002, and the European sovereign debt crisis period starting at the end of 2009, which means we can compare the average volatility level before and after the financial crisis over long horizons. By observing the average volatility level and jump frequency for different stocks, we can see how they are affected by news announcements (see discussion in Sections 4.4.2 and 4.4.3).

Panels A and B in Table 4.5 show the stocks with few jumps and high average volumes respectively before and after the financial crisis, while Panel C is comprised of stocks with both high average volumes and few jumps. From Table 4.5, we can observe that the AAPL stock from the IT sector has the largest trading volume, which is much larger than most of the other stocks before and after the financial crisis. On the other hand, BT from the TS sector has a relatively low trading volume compared to stocks from the other sectors during both periods. In addition, we find that TEF and BT from the TS sector have the fewest jumps before and after the financial crisis. Stocks with high volumes typically have many jumps per day, which are more frequent than the average value across different stocks (0.08 for the pre-crisis period and 0.12 for the post-crisis period).

It is also clear that the number of jumps per day is different in the two periods. For example, the number of jumps per day for the GILD and EXC stocks increases from 0.14 and 0.11 to 0.16 and 0.14 respectively. However, the change in the jump frequency per day for stocks with few jumps overall is mixed. For example, BT and TEF have 0.01 jumps per day, which is the lowest amount of the 21 stocks before the financial crisis, which increases to 0.05 jumps per day after the financial crisis. The number of jumps per day for the UL stock decreases from 0.11 to 0.5.

[Insert Table 4.5 here]

From Table 4.5, it is also clear that stocks with high volumes and low jump frequencies are primarily from the Information Technology sector: AAPL, MSFT, INTC and CSCO. This suggests that although there are generally fewer jumps for the stocks in the IT sector, the jumps that are present are caused by large trading volumes in response to market-level or company-level news, and thus may be larger than the jumps for other stocks. We also find that the number of jumps for all stocks with high volumes increases to more than 0.10 jumps per day from the pre-crisis period to the post-crisis period.

In addition, the table shows that the number of jumps per day for most IT sector stocks increases dramatically after the financial crisis because they are active stocks with high trading volumes. For example, the respective number of jumps per day for the stocks INTC, CSCO, MSFT, IBM and YHOO are 0.04, 0.05, 0.06, 0.07 and 0.07 before the financial crisis, increasing to 0.13, 0.14, 0.13, 0.14 and 0.15 after the crisis. This dramatic rise may be because investors had excessively optimistic expectations, resulting in larger changes in market returns. Arnott et al. (2018) argue that this pattern is a sign of another tech bubble that may burst in the future, just as in 2000, when tech giants such as Microsoft, Cisco, Intel, IBM, AOL, Oracle, Dell, Sun, Qualcomm and HP failed to meet investors' excessively optimistic expectations.

4.4.1.2 Jump components

In this section, we investigate the size of jump components estimated using BV and TBV in different regimes (the pre-crisis and post-crisis periods) by examining the descriptive statistics for jump components with and without the highest 10% of the values, and observing plots generated from the data. This helps us understand the differences in jump size for different stocks in different financial regimes and how they are related to news announcements. This section also considers which types of news (company-related news or market-related news) are the main causes of extreme jumps for each stock. We plot the jump components estimated using TBV in Figures 4.1 to 4.6. The descriptive statistics for jump components estimated using BV and TBV are calculated in Tables 4.6 to 4.9. TBV is a frequently used volatility measure in recent studies (e.g. Haugom & Ullrich, 2012; Vortelinos & Saha, 2016; Hizmeri et al., 2019), so we include the descriptive statistics of jump components estimated using TBV to compare the different volatility measures when estimating jump components. The CTBV method is similar to TBV, but it includes some modifications to reduce bias. Unsurprisingly, then, the descriptive statistics of the jump components estimated using CTBV share a similar pattern to those estimated using TBV, but with lower values and fewer significant jump days for both the pre- and post-crisis periods. More detailed evidence is shown in Section 4.4.1.4. In order to avoid repetitive comparisons based on similar values, we did not include the descriptive statistics tables for the jump components estimated using CTBV.

[Insert Figures 4.1 to 4.6 here]

Figures 4.1 to 4.6 and Tables 4.8 and 4.9 show the features of jump components estimated using TBV for stocks with high volumes and low jump frequencies in the pre-crisis and post-crisis periods. Results are shown for the full dataset for each stock (left-hand panel) and the data up to the 90th percentile, with extreme values in the top decile excluded (right-hand panel). From Figures 4.1 to 4.3, we can see that most of the stocks, especially those in the IT sector, have volatile jump components between 2000 and 2002. This was caused by the burst of the dot-com bubble, which corrected the unusual valuation of stocks caused by a decade-long bull market before 2000. Also, we find that the jump components from 2003 to 2007 in the pre-crisis period are relatively stable, and that they are not visibly larger than in the post-crisis period (as seen in Figures 4.4 to 4.6). There are some large jumps at the end of 2009 caused by the European debt crisis; however, the number and size of the jumps are lower when compared to those caused by the burst of the dot-com bubble, which was mainly centred around American companies. This suggests that shocks in the US stock market that originate from overseas economies are more limited compared to domestic equivalents.

[Insert Tables 4.6 to 4.9 here]

Tables 4.8 and 4.9 show the jump components for stocks estimated using threshold bi-power variation (TBV) for different stocks. By comparing the BV-based results in Tables 4.6 and 4.7 with those in Tables 4.8 and 4.9, we find that the jump components estimated using TBV are normally higher than those using BV. This is in line with Corsi et al.'s (2010) suggestion that BV underestimates jump components. They also argue that the bias for BV is more obvious for large jumps, which is supported by our results since the stocks with large maximum values are much larger when estimated using TBV. For example, the estimated jump components for the CNP stock are 124.86 and 23.323 in the pre-crisis and post-crisis periods using BV, while their estimated values are 377.450 and 765.330 when using TBV. The jump components produced by these two estimators show similar volatility levels for particular stocks. For example, GILD and BSX are the stocks with the highest jump components on average in the pre-crisis and post-crisis periods, regardless of the estimator used.

Table 4.6 describes the descriptive statistics for stocks' jump components in the pre-crisis period. The descriptive statistics shows that the XRX stock has a large jump component on 19/6/2003. This large jump component is primarily caused by the announcement of the Initial Jobless Claims Index in the middle of the day. The Jobless Claims index was announced to be 421,000 at 1.30pm on 19/6/2003, which was the lowest figure for four weeks. The announcement of the news led to an increase in returns from 0.475 to 1.131, contributing to the large jump component that day. This news also resulted in a dramatic change in returns for the AMGN, UL and GILD stocks on the same day at the same time, which made the majority of the contribution to those stocks' jump components that day. The returns for ANGN, UL and GILD increased from 0.101 to 0.247, from 0.027 to 0.054 and from 0.285 to 0.360 respectively.

The descriptive statistics also show that the CNP stock has one large jump component that occurred on 11/10/2002. This large jump is mainly caused by company news – namely, Standard & Poor lowered CenterPoint Energy's corporate credit rating from BBB+ to BBB on 11/10/2002 because of its large debt. This announcement resulted in a dramatic fall in its stock price from 2.198 at the end of 10/10/2002 to -2.153 at the end of 11/10/2002.

However, the announcement of negative news for CNP does not seem to cause large fluctuations in prices for other stocks. Therefore, although announcements in both macroeconomic news and company-related news produces jumps in stock returns, the impact of macroeconomic news is wider as it normally affects more than one stock. The impact of macroeconomic news is more predictable, as announcements are usually made seasonally; company-related news, however, typically comes as unpredictable one-off announcements.

The variance in jump size for the XRX stock is the largest and the second largest in Tables 4.6 and 4.8 (96.395 and 145.800 respectively). Similarly, XRX's mean is the second largest among all the stocks in both Tables 4.6 and 4.8, and many of its daily jump components are over 50 in Figure 4.1. This indicates that XRX is an extremely volatile stock with relatively large jump components and high variance compared to other stocks. However, CNP is less volatile than XRX despite its high maximum jump components (124.86 and 377.450). (These are visible in the left-hand panels of Tables 4.6 and 4.8.) As can be seen from the mean and mode of CNP's jump size, most of its jump components are much smaller than its maximum value, as the mean jump size is only 2.2. Its mean, variance, mode and maximum value for 90^{th} percentile jump components are smaller than many stocks in Table 4.6, and most of its daily jump components (see Figure 4.1) are under 5.

The mean and mode of the jump size for stock GILD are the largest in Table 4.6, and this stock's maximum value and variance are above average in the left panel, showing that it is volatile. The mean, mode, variance and the maximum value of the 90th percentile jump components for GILD (the right side of the table) are the highest, with 2.864, 0.213, 6.718 and 10.323, reflecting the results for the full data set (the left side of the table). In addition, the scatter plot for GILD (see Figure 4.2) shows that the daily jump components displayed the most volatility in 2000, with many TBV-estimated jumps over 50 occurring that year. The top 10% jump sizes for the XRX, TEF and CNP stocks would appear to be large in the pre-crisis period, as their variance fell by more than 95% when excluding the top 10% extreme values in Tables 4.6 and 4.8.

In the post-crisis period results in Table 4.7 (BV), the jump components with the highest maximum value and variance are from AEE, yet the jump components decreased dramatically from 70.683 and 18.169 to 1.084 and 0.052 before and after the crisis respectively when excluding the 10% most extreme values. Similar patterns can be found for TBV in Table 4.9 and Figure 4.4. Most of the jump components for AEE are under 2, although it has a few daily jumps close to 10 (see Figure 4.4).

In Tables 4.7 and 4.9 (BV and TBV), the BSX stock has the largest mean jump component values (1.208 and 1.799) in the left panel, and the largest maximum values (2.107 and 3.272), average values (0.833 and 1.192) and variance (0.214 and 0.523) for the jump components in the right panel. This is also visible in Figure 4.4.

It is clear that the maximum value, mean and variance for the jump components for the UL stock are the smallest in both the full and 90^{th}-percentile data using BV, which are 0.882, 0.161, 0.020 and 0.288, 0.122, 0.004 (see Table 4.7). Also, UL's jump components estimated via TBV have the lowest maximum, mean and mode values the 10% extreme values are excluded, as shown in the right-hand panel of Table 4.9. Furthermore, Figure 4.4 shows that the UL has the most stable and smallest jumps among all the stocks on average (see Figures 4.4 to 4.6). The top decile for the AEE, CNP, EXC, T and XRX stocks appear to be have volatile jump components in the post-crisis period, as the variance for these stocks decreases by more than 95% when taking out the 10% most extreme values (see Tables 4.7 and 4.9).

The tables and figures in both the current and previous sections show that the stocks from the IT sector are on average more volatile than the stocks from other sectors, especially in the pre-crisis period. In the next sub-section, we compare the volatility levels for stocks in different financial regimes in order to examine the factors that contribute to high volatility. The impact of news announcements on stocks with different volatility levels (IT sector stocks versus other sector stocks) are compared in the remainder of Section 4.4.1 and in Section 4.4.2.

4.4.1.3 Comparison of stocks in different time periods

In this section, we compare in detail the descriptive statistics for the jump components estimated using BV and TBV in different years. By comparing the jump components in different years, we can investigate how market-level and company-related news announcements affect the jump components in different market conditions (e.g. the bear market caused by the burst of the tech bubble). In Figures 4.1 to 4.3, there are more jumps for most stocks in 2000-2002 and 2007. This is because a large financial market correction started in 2000 for the unusually high stock evaluations caused by the bull market in the late 1990s. The correction started with the bankruptcy of the energy company Enron and many internet-based companies such as Webvan, Exodus Communications and Pets.com in 2001. The bear market reached its low point in 2002, with the Dow Jones Industrial Average and Nasdaq reaching 7286.27 and 1114.11 in October 2002. In addition, the terrorist attacks on 11 September 2001 resulted in a heavy downturn in the stock market as investors were unsure about the impact of the attack on the American economy.

The 2008 financial crisis developed from problems in the subprime mortgage market in 2007. The subprime mortgage market crisis began with the bankruptcy of the American real-estate investment trust New Century in April 2007, causing the Federal Reserve to supply short-term credit to banks with sub-prime mortgages. On 15 September 2008, the subprime mortgage market crisis developed into an international financial crisis as the Lehmann Brothers investment bank went bankrupt. Stock prices changed dramatically in 2007 due to the impact of the subprime mortgage crisis. Therefore, we examine the years with large jumps separately, such as the bear market from 2000 to 2002 and the year of the sub-prime mortgage crisis in 2007, as the stock returns are more volatile in these years. The jump components estimated using BV and TBV for different periods in the pre-crisis period are shown in Tables 4.10 and 4.12 respectively.

For the post-crisis period, the global financial crisis is regarded as having come to an end in June 2009, as the recession, which began in December 2007, reached its end and financial stocks began recovering from their dramatic drops during the crisis. However, the European sovereign debt crisis occurred at the end of 2009, caused by the failure to repay, or refinance, government debts for some members of the European Union. The debt crisis was only pronounced in a handful of Eurozone countries such as Greece, Cyprus, Ireland, Italy, Portugal and Spain before 2011, but it resurfaced in May 2011 and started affecting the whole Eurozone because of rising concerns about the Greek government debt crisis. European Union leaders agreed on bailout programmes to help countries with debt crises, especially Greece. The first, second and third economic adjustment programmes for Greece were implemented between May 2010 and February 2012; March 2012 and June 2015; and July 2015 and August 2018. Although some literature (e.g. Allegret et al. 2017) suggests that the impact of the European debt crisis on US stock markets was limited, some effects were certainly present given that the European Union is the second largest economic union in the world and is closely linked to the US economy (BBC, 2011). Therefore, we examine the descriptive statistics for jump components estimated using BV and TBV in Tables 4.11 and 4.13 for two time periods: (i) the recovery from the financial crisis despite concerns about the European sovereign debt crisis from July 2009 to June 2010; and (ii) the resurfacing of the crisis in the whole Eurozone and the implementation of bailout programmes in the EU from July 2010 to December 2016.

[Insert Tables 4.10 to 4.13 here]

Tables 4.10 and 4.11 show the descriptive statistics for jump components estimated using BV for different stocks for the pre-crisis and post-crisis periods. Tables 4.12 and 4.13 do the same for TBV. Tables 4.10 and 4.12, illustrated in Figures 4.1 to 4.3, show that the jumps are larger for all stocks from 2000 to 2002 than for the pre-crisis period as a whole. In the years immediately following the millennium, the stock market was in the process of correcting for the unusually high stock evaluations that characterised the dot-com bubble. The four stocks with the largest jump components in the pre-crisis period are GILD, XRX, YHOO and AAPL with 4.388, 3.771, 2.844 and 2.466 respectively, as shown in Table 4.10. They have much higher jump components between 2000 and 2002, with 7.535, 5.715, 6.058 and 4.585 respectively. This indicates that investors may have been more sensitive to news announcements during the most uncertain period in the stock market after the dot-com crash. XRX, YHOO and AAPL are from the Information Technology sector, which has relatively large jump components as a whole. The burst of the dot-com bubble had the biggest effect on IT companies (Andersen et al. 2010), which is reflected in the sector's high jump components.

As shown in Table 4.11, after the financial crisis, the stock prices of XRX and YHOO continued to have high average jump components (1.122 and 1.081), while other IT companies stabilised. These two stocks' mean jump components are only exceeded by those of BSX (healthcare). This may be because the two companies faced significant declines in profits between 2009 and 2016 compared with firms in the same sector such as AAPL (Macrotrends.net 2020). Table 4.13 shows that the equivalent results for the threshold bi-power variation method are the same.

In addition, when using the BV method (as shown in Tables 4.6, 4.7, 4.10, and 4.11), GILD and BSX are the most volatile in the pre-crisis and post-crisis periods. GILD's large average jump components in the pre-crisis period are mainly the result of a dramatic increase in its stock prices from -0.093 to 6.422, which occurred between 1.10 pm and 1.20 pm on 3rd January 2001 after the announcement of US interest rate cuts. The most volatile period for BSX during the post-crisis period occurred between 20th October to 8th November 2010. BSX's stock price increased dramatically from -2.31 to 1.76, likely as a result of the announcement of the company's strong performance during the third fiscal quarter on 19th October 2010. However, the positive impact of this announcement was not long, as BSX's returns dropped to -0.94 the following day. This may be due to the announcement from the Initial Jobless Claims index on 2first October that the US unemployment rate had increased to more than 9.5%. This continued to have a negative impact on the firm's stock prices until the Initial Jobless Claims announcement on 29th October 2010. This suggests that macroeconomic news announcements may have a notable and lasting impact on stock market volatility. These effects may be predictable due to the scheduled weekly and monthly announcements of news from various sources.

Tables 4.11 and 4.13 show that stocks' average jump components are generally higher between July 2009 to June 2010 for most companies, with the exception of GILD. Although the US National Bureau of Economic Research (NBER) announced in June 2009 that the US recession had ended and thus that the financial crisis was considered to be over, the market did not stabilise until January 2010. Many stock prices fluctuated during the recovery stage prior to January 2010, resulting in large jump components. GILD has high average jump components compared to other stocks from July 2009 to June 2010, which may be a result of the company's acquisition of several other pharmaceutical firms during this time.

4.4.1.4 Comparison of daily jump component size

While TBV is more frequently used in recent literature (e.g. Haugom & Ullrich, 2012; Vortelinos & Saha, 2016; Clinet & Potiron 2017; Hizmeri et al., 2019), its jump components show similar patterns in different financial regimes to CTBV. Corsi et al. (2010) suggest that CTBV is less biased at estimating jump components compared to TBV; therefore, the remainder of this chapter focuses on results estimated using CTBV. In this section, we compare the size of the jump components for different stocks in different financial regimes so that this information can be incorporated into the analysis of the effect of news announcements on different stocks in the following sections. Table 4.14 presents the size of the jump components estimated using corrected threshold bi-power variation (CTBV). The results show that the average daily jump components for most stocks (with the exception of AEE) are higher before the financial crisis than afterwards; this is similar to the results in Tables 4.6 to 4.13 and Figures 4.1 to 4.3. For AEE, days with high jump components, some of which approach 10, are more frequent after the crisis. This can also be seen in the TBV scatter plots for AEE (see Figures 4.1 and 4.4). The jump components in the post-crisis period in Figure 4.4 are generally larger and more frequent than those in the pre-crisis period in Figure 4.1.

[Insert Table 4.14 here]

In addition, KO is the only stock in Table 4.14 with no daily jump components greater than 5 before the financial crisis, and it has only one daily jump greater than 5 in the post-crisis period. The equivalent results for TBV, as shown in Figures 4.3 and 4.6, are similar. However, KO's maximum jump components estimated using TBV (10.904 and 8.540 for the pre-crisis and post-crisis periods respectively) in Tables 4.8 and 4.9 are slightly higher than the CTBV values (less than 5 and 10 respectively) shown in Table 4.14. This result is in line with Corsi et al.'s (2010) claim that the TBV tends to overestimate jump components. The authors developed CTBV in order to correct the biases present in TBV.

The CTBV jump components in Table 4.14 also reveal that the XRX, YHOO and BSX stocks are the most volatile stocks before the crisis, as they each have at least eight days with jump components over 5. These stocks are even more volatile after the financial crisis, as the days with large jump components become more frequent. GILD is also more volatile compared to other stocks in the post-crisis period as it has 32 days with jump components higher than 5; its pre-crisis volatility, however, is less pronounced than that of XRX, YHOO and BSX. This indicates that the patterns observed for TBV, such as IT sector stocks' behaviour in response to the burst of the dot-com bubble, are also present in estimations produced by CTBV. The main difference between the two estimators is that TBV tends to produce higher values, which are mitigated in CTBV (Corsi et al. 2010).

4.4.1.5 Section summary

In this section, we examined the impact of news on jumps and trading volume for stocks in different financial regimes. The results show that IT sector stocks tend to have higher trading volumes and volatility, likely caused by news announcements, but that they also have a relatively small number of jumps. In addition, the IT sector stocks XRX and YHOO have more extreme jumps than most stocks (Section 4.4.1.4). This provides evidence that market-level news announcements may not produce large numbers of jumps; rather, such announcements may cause dramatic changes in volatility, at least for IT sector stocks (Section 4.4.1.1). In addition, we find that jump components estimated using TBV fall after the method's bias has been corrected (i.e. when using CTBV), as shown in Section 4.4.1.4, which is in line with Corsi et al.'s (2010) findings. By examining the impact of news on jump components, we find that macroeconomic market-level news has a longer (Section 4.4.1.3), wider and more seasonal (Section 4.4.1.2) impact on stock volatility compared to company-related news, which tends to have a one-off impact on fewer stocks (Section 4.4.1.2). Throughout this section, we have studied the descriptive statistics and size of jump components in different financial regimes and individual years. We have found that the jump components from 2000 to 2002 are larger than average for the pre-crisis period, which suggests that investors respond to news more strongly when the market is uncertain.

4.4.2 Second-moment market responses to news

4.4.2.1 The effect of macroeconomic news

In this section, we first examine the impact of news on the market's second-moment responses by testing how news surprise from individual news outlets affects stocks' jump components in different financial regimes. We then test the impact of macroeconomic news as a whole (from all news outlets) on individual stocks by examining the descriptive statistics for jump components that co-occur with macroeconomic news. The impact of individual news outlets is examined based on equations (4.2) and (4.3) from Section 4.3.1, which show the regressions of the jump components on individual news outlets and all news outlets respectively. The jump components are estimated via corrected threshold bi-power variation (CTBV) of realised volatility. CTBV is used here as Corsi et al. (2010) suggest that it improves the performance of threshold bi-powder variation (TBV) in estimating jump components.

[Insert Tables 4.15 and 4.16 here]

Tables 4.15 and 4.16 show the *p*-values for the realised jump components for high-volume and low-jump-frequency stocks regressed on news surprise, measured according to individual news outlets (as shown in equation (4.2) in Section 4.3.1) before and after the financial crisis respectively. Variables that were not significant at the $p < 0.05$ level for any of the stocks were removed from the tables, leaving only news sources for which at least one stock has a significant *p*-value for positive or negative news surprise.

Tables 4.15 and 4.16 show that there are more stocks that are significantly affected by positive news surprises from Capacity Utilisation (CPTICHNG) and Initial Jobless Claims

(INJCJC) news than other types of news in the pre-crisis period. The jump components for 17 out of 21 stocks are significantly affected by these two news surprises. The jump components for the BT stock are the most affected by positive surprises from CPTICHNG and INJCJC, with coefficient values of 0.730 and 0.305 respectively for these two news outlets. The jump components for XRX, CNP, BT and BSX are affected to a greater degree by news surprises than those for other stocks. For example, the jump components for XRX are affected by many news surprises, including positive news surprises from NAPMPM and GDP PIQQ (estimated coefficients of 0.588 and 0.924 respectively) and negative news surprises from NHSPSTOT and NAPMPMI (estimated coefficients of 0.480 and 0.421 respectively). The positive surprises from the news outlets CPTICHNG, LEI CHNG and CONSSENT and the negative surprises from MAPMNMAN and PITLCHNG have a significant impact on CNP's jump components, with coefficients all larger than 0.4. In addition, positive surprise from the outlet CPTICHNG on BT and negative surprise from GDPCTOT on BSX have significant effects, with coefficients of 0.730 and 0.841 respectively.

Comparing Tables 4.15 and 4.16, it is apparent that the jump components for stocks in the post-crisis period are affected by a wider range of news than those in the pre-crisis period. For example, the jump components of the AAPL and CSCO stocks are significantly affected by three and four news surprises respectively in the pre-crisis period, yet this increases to 15 and 19 news surprises respectively in the post-crisis period. The negative surprise for the University of Michigan Consumer Sentiment Index (CONSSENT) has a significant impact on the jump components for 17 stocks in the post-crisis period, the highest coefficient being that for YHOO (0.210). UL is influenced by news surprise the least, as its jump components are only affected by positive surprises for NAPMPMI and PCE CMOM, and negative surprise for CHPMINDX, with respective estimated coefficients of 0.061, 0.050 and 0.042, all of which are smaller than 0.1.

The jump components for XRX, YHOO and BSX, however, are noteworthy as they are affected by a large range of news surprises in the post-crisis period, but also to a sizeable degree. They are significantly affected by 18, 21 and 22 news surprises respectively, and five, four and nine of their respective estimated surprise coefficients are greater than 3. Similarly, the jump components for AEE are affected by more types of news surprises (24) than any other stock, and only two of the news surprises within those 24 are smaller than 0.1. The results also show that the announcement of news does not have a different impact on high-volume stocks such as PFE and EXC compared to low-jump stocks such as BT and TEF.

In summary, the findings indicate that the jump components of high-volume and/or low-jump-frequency stocks are significantly affected by positive surprises from the Consumer Price Index (CPI CHNG) and Initial Jobless Claims (INJCJC) index before the financial crisis, and by negative surprises from the University of Michigan Consumer Sentiment Index (CONSSENT) in the post-crisis period. XRX and BSX's jump components are more likely to be affected compared to those of other stocks. In addition, jump components are sensitive to a wider variety of news announcements after the financial crisis than prior to it. Finally, we did not find significant evidence that the announcement of news has a different impact on high-volume stocks compared to low-jump-frequency stocks.

[Insert Figures 4.7 to 4.14 here]

Observation of the figures reveals that the jump components vary considerably between 2000 and 2002 in the pre-crisis period due to the burst of the dot-com bubble. It is clear that many jumps during this period occur on the same day as news announcements. There are also more co-occurrences of jumps and news in 2009-2010 than for the whole period after the financial crisis, as this is when US financial markets were recovering from the crisis and subsequent recession.

However, the number of daily jump components that may be affected by co-jump-related news (see Figures 4.7 to 4.14) is limited. As reported in previous literature, different types of stocks are affected by different news outlets. For example, Boyd et al. (2005) find that utility stocks and cyclical stocks have different responses to announcements of unemployment rates. In the present study, co-jumps occur relatively infrequently across the entire 16-year span of the data set, and they are concentrated in particular periods of time. Specifically, the largest concentrations of co-jumps occur between 2000 and 2002 in the pre-crisis period, and between June 2009 and July 2010 in the post-crisis period. This suggests that there are more news-related co-movements when the market is not very stable, such as during the burst of the dot-com bubble or the post-recession recovery. In addition, the figures show that there are some co-jumps that do not appear to be related to announcements of macroeconomic news. This indicates that there are other factors that may cause co-movements between stock prices. These factors will be discussed in the following section.

[Insert Table 4.17 here]

In the pre-crisis period, the stock AEE has the smallest mean value and variance for jumps that are related to news (as seen in Table 4.17). The proportion of its jump components that co-occur with news is the third smallest in the data set at 0.028. This reflects the overall jump results from Figure 4.1 and Table 4.8; the latter shows that 90% of the jump components estimated via TBV for AEE are lower than 1.831. Similarly, news is related to a relatively small proportion of jumps for the stock CNP at 0.024, which is the smallest proportion in the data set for the pre-crisis period. However, its news-related jumps' mean value and variance are 1.146 and 0.819, which are relatively large. Similar effects can be seen for this stock's jumps as a whole in Table 4.8, where the average jump size for CNP in the pre-crisis period is 3.616. This is relatively large compared to the majority of stocks. In summary, both AEE and CNP have small proportions of news-related jumps, but CNP appears to be affected by news announcements to a greater degree because of its larger jump components.

In addition, the PFE has the highest proportion of jumps caused by news (0.257) in the pre-crisis period, while BSX has the highest jump size mean and variance (2.941 and 23.154). Similar figures can be found for the jump components as a whole in Table 4.14 and Figures 4.7 and 4.10. For example, BSX has 18 daily jump components estimated via CTBV that are greater than 5 in Table 4.14, while PFE only has two days with jump components greater than 5. In Figures 4.7 and 4.10, most of the jumps caused by news for BSX are more volatile than for PFE, especially in 2000-2002. This suggests that the three factors are not completely correlated.

In the post-crisis period, it is clear that IBM has the lowest proportion of jumps explained by news (0.011). YHOO has the highest mean and variance for news-related jumps (1.299 and 9.969), but its proportion of news-related jumps falls close to the middle of the range (0.193). This is because YHOO has many days with large jump components, as shown in

Table 4.14, and many of them co-occur with news in the post-crisis period (see Figure 4.12). VOD has the highest proportion of jumps caused by news in the post-crisis period at 0.352, but its mean jump value and variance are fairly low at 0.251 and 0.044 (see Table 4.17). As shown in the right-hand panel of Table 4.14, VOD's daily overall jump components estimated via CTBV are mostly less than 5, and their mean size and variance when related to news are not as large as those for YHOO (as shown in Table 4.17 and Figure 4.12). This again indicates that the proportion of news-related jumps is not necessarily correlated with jump size and variance, and that different stocks show different levels of sensitivity to news announcements. That is, some stocks frequently respond to market shocks, including news announcements, while others do so rarely but on a larger scale.

4.4.2.2 The effect of co-jumps and other types of news

In this section, we first address some of the factors other than news surprise that may cause co-jumps and may influence jump components. Second, we investigate how co-jump-related news affects jump components by examining the descriptive statistics. Figure 4.15 shows how company news and overseas market news can affect the jump components estimated by CTBV from the first half of 2016. The surprising success of the 'Leave' campaign in the UK referendum on leaving the European Union (Brexit) resulted in dramatic movements in financial markets on 23rd June 2016, whose effects have persisted to the present. The price of pound sterling fell to a low point on 6th July 2016, which also led to a decrease in the price of stocks for many global companies such as AAPL, INTC and IBM. These jumps can be seen in Figure 4.15.

[Insert Figure 4.15 here]

In addition, stock prices are not only affected by overseas news, but also by domestic news related to specific companies within a particular sector. For example, Apple's (AAPL) impressive performance in Q2 2016, reported by the firm on 27/7/2016, resulted in an increase in stock prices for other IT sector companies like INTC and IBM, as seen in Figure 4.15. Although the jumps that co-occur with major overseas news and company news can cause co-jumps across different stocks, the number of co-jumps related to US macroeconomic news (that is, the type of news outlets analysed in this chapter and shown in Figures 4.7 to 4.14) tend to have predictable, seasonal patterns, as they are usually announced either weekly or monthly. In contrast, co-jumps caused by overseas and company news are normally the result of a one-off co-movement between different stocks.

In the remainder of this section, we look at how news, including market-level news and company-related news, may occur with co-jumps in different stocks, and how these co-jumps may help explain jump components. In the next section, the impact of news-related co-jumps on volatility estimation and forecasting for different stocks is considered.

[Insert Tables 4.18 and 4.19 here]

Tables 4.18 and 4.19 show the jump components that display co-movement estimated using CTBV before and after the financial crisis for the six lowest-jump-frequency and highest-volume stocks overall (data subset one) and the most extreme for each sector (data subset two). Each vertical panel within the tables (i.e. the overall top six and sector-specific stocks in terms of high volume and low jump frequency) represents a different co-jump test using a different configuration of stocks. The co-jumps are tested via the MCP co-jump test

(Bollerslev et al. 2008). AAPL, EXC, GILD, KO and T are the stocks with the highest volume in each sector during the pre-crisis period and AAPL, EXC, PFE, KO and T are the same for the post-crisis period – six stocks in total (see Table 4.19). We are interested in how jumps for stocks in one sector move simultaneously with those in other sectors, hence we estimate the co-jumps for the six stocks referred to above, which represent each sector of the market.

From Figures 4.7 to 4.14 and Tables 4.18 and 4.19, we can see that co-jumps can capture a greater proportion of changes in stock prices if they are primarily estimated based on stocks from the same sector. For example, the proportion of jumps that co-occur with co-jumps for AAPL in the pre-crisis period in Panel B of Table 4.18 is 0.106. This panel also includes the other six stocks from the IT sector. This proportion is higher than the results for AAPL in the same period in Panel A in Table 4.18, which contains fewer other stocks from the IT sector. Equivalent effects occur for CSCO and MSFT. Similarly, the results in Table 4.19 show that co-jumps can capture a greater proportion of changes in stock prices when more stocks from the same sector are included. This is despite the fact that the majority of the stocks in this subset are from other sectors (e.g. the proportion for AAPL in the pre-crisis period in Panel B is higher than that in Panel A). In other words, including more stocks from the same sector seems to improve the results, regardless of the number of stocks from other sectors that are included. This suggests that co-movements exist between stocks caused by macroeconomic news announcements whose effects are more localised to a particular sector compared to others; this is more likely to be captured when including more stocks from the same sector. The results from Section 4.4.2.1 reinforce this, as we found that negative surprise from the news outlets SPCS20SM and MTIBCHNG had a significant impact on many stocks in the IT sector in the post-crisis period, but had little effect on stocks from other sectors.

Alternatively, it is possible that company-specific news plays a bigger role in influencing stock prices of many firms in the same sector. It is understandable that news announcements from one company will have a major effect on firms that rely on that company for parts or systems, or that compete with it for market share. However, the data discussed above also show that company news is often unpredictable, and that its effects do not last for a long time in comparison to regular, periodic macroeconomic news announcements, potentially restricting the usefulness of company news for volatility forecasting. In the remaining sections of this chapter, we will limit our analysis to co-jumps that co-occur with regular macroeconomic news announcements.

4.4.2.3 The effect of news-related co-jumps

In the previous section, shocks from both regular macroeconomic news and irregular company news were considered. In this section, we focus only on regular macroeconomic news, the results for which are shown in Tables 4.20 and 4.21. These tables describe the results for the jump component regressions on news surprises which co-occur with co-jumps. Co-jumps are generated using the MCP co-jump test (Bollerslev et al. 2008). Thus far, we have seen that the number of co-jumps between stocks that occur at the same time as macroeconomic news announcements is relatively small. In order to study this type of co-jump, we use the regression in equation (4.3) to estimate the coefficients (see Section 4.3.1). In equation (4.3), the positive and negative same-sign news surprises for the significant news outlets in Tables 4.15 to 4.16 are combined into two independent variables: positive and negative news surprise. This helps avoid the problem that the number of observations of news surprise captured by co-jumps for a single news outlet are too few to be considered as a separate independent variable. The estimated regression coefficients for the pre-crisis and post-crisis data are shown in Tables 4.20 and 4.21.

[Insert Tables 4.20 and 4.21 here]

In Panel B of Table 4.20, positive news surprise contributes to the jump components of AAPL and INTC from the IT sector when they are considered as the top two stocks from this sector, with estimated coefficients of 0.167 and 0.110 in the pre-crisis period and 0.066 and 0.093 in the post-crisis period. These coefficients are higher than those generated in the overall top-six stocks calculation in both panels of Table 4.21. This suggests that the co-jumps are more likely to co-occur with positive news surprise when they are estimated together with a diverse range of stocks from different industrial sectors (Table 4.20; stocks from each sector) instead of only with stocks mainly from the IT sector (Table 4.21; overall top six stocks).

This finding is in line with Bollerslev et al. (2008), who suggest that co-jumps based on stocks from a diverse range of industrial sectors are good at capturing the impact of macro-economic news announcements. However, this appears to contradict the results from the previous section, in which the proportion of co-jump-related jump components were higher for IT stocks when considered together with a number of other stocks from that sector. We interpreted this in the previous section by comparing the possible influence of regular macroeconomic news to that of one-off company news. The results here make a stronger case for the effect of company news. In the previous section, all co-jumps were considered, but here, the data are restricted only to co-jumps that co-occur with macroeconomic news. The strong results for the cluster of IT sector stocks in the previous section are not borne out in the more limited data set here, suggesting that the findings from the earlier analysis likely cannot be attributed to the effect of regular macroeconomic news that has a greater impact on certain sectors. This raises the question of the potential impact of other factors on co-jumps such as one-off company-specific news. This kind of news may have a stronger but more localised influence on stock prices within a particular sector.

The generally larger coefficients in the left-hand column of Table 4.20 indicate that news surprises captured by co-jumps are better at explaining the jump components for the pre-crisis period than for the post-crisis period. This is the case for the vast majority of stocks, with the exception of AEE (Utilities). This complements the data from Table 4.15, which shows that there are fewer news outlets with a significant impact on stocks' jump components in the pre-crisis period compared to the post-crisis period. For example, there are only three estimated news surprise coefficients for AAPL in the pre-crisis period that are significant (see Table 4.15), but this figure is 19 for the post-crisis period (Table 4.16). However, all of these post-crisis coefficients are smaller or equal to 1.80, yet the three pre-crisis coefficients are all greater than 2.4. Table 4.20 acts as an aggregate of the co-jump-related news surprises, and it is clear that the collective effect of the co-jumps that occur with news surprises is higher before the crisis than after it. In other words, similarly to Table 4.15, the co-jump-related collective impact of news surprises is greater before the crisis than after it.

The one exception to this pattern, AEE, has a small number of very large jumps captured by co-jumps after the crisis (e.g. daily jump components of 70.683 [Table 4.7; BV] and

82.358 [Table 4.9; TBV] on 17/11/2015), which increases the significance value of the estimated coefficient dramatically. This finding is important when considering the impact of news surprise on stock market volatility during different periods of time and market conditions, as the data suggest that the relationship between co-jumps and news surprise differs between periods with different economic conditions.

4.4.3 The impact of news and co-jumps on forecasting models

In this section, we examine the impact of news and co-jumps on the modified HAR, HAR-TJ and HAR-CTJ volatility forecasting models, which were shown in equations (4.7), (4.8) and (4.9) in Section 4.3.2. We do this by comparing the results to those of the standard HAR, HAR-J and HAR-TJ models, shown in equations (4.4), (4.5) and (4.6) in Section 4.3.2, which we use as a benchmark. Table 4.22 presents the forecasting results for the standard HAR, HAR-J and HAR-TJ models. It shows that the forecasting results for almost all stocks are better in the post-crisis period than in the pre-crisis period. This may be due to the fact that the high volatility caused by the burst of the dot-com bubble in the pre-crisis period is unusual and does not continue into later years. This suggests that the data from 2000 to 2002 is less useful for forecasting, as it affects the results for the whole of the pre-crisis period.

An example of this issue can be seen in AAPL, which is one of the most active IT companies and whose value is closely related to the development of new technology compared to firms from other industrial sectors. AAPL has the worst forecasting results before the financial crisis, shown by its high mean standard error (MSE) values, probably because it is affected the most by the burst of the tech bubble. This can also be seen in its trading volume, shown in the left-hand panel of Table 4.5 in Section 4.4.1.1, which is the highest among all the stocks in the pre-crisis period. Large trading volumes often cause dramatic changes in volatility, presenting difficulties for the HAR model's ability to predict stocks' future volatility. As shown in the right-hand panel of Table 4.5, AAPL's trading volume falls after the financial crisis, resulting in a more stable volatility pattern and yielding better forecasting results. Similar forecasting patterns can be found for the other IT sector stocks YHOO and ORCL, as they have large MSEs in the pre-crisis period (8.100 and 7.388 respectively for the standard HAR-TJ model), but these fall to 1.951 and 0.346 in the post-crisis period.

[Insert Tables 4.22 and 4.23 here]

By comparing the results in Tables 4.22 and 4.23, it is clear that considering the impact of news in forecasting models can improve predictions for many stocks. More than half of the stocks have better forecasting results in the pre-crisis period from the modified HAR-TJ model, which considers news-related jump components separately, than from the standard HAR-TJ model. Nine out of 21 stocks show better forecasting results from the modified HAR-TJ model in the post-crisis period. The HAR-TJ model shows the greatest improvements after modification, but similar results can be seen for the HAR-TCJ model. Nearly half of the forecasting results in the pre-crisis period from the HAR-TCJ model improve after modification. The gains in forecasting accuracy are particularly pronounced for stocks with very large news-related jump components. For example, most stocks from the IT sector, such as XRX, YHOO, ORCL, AAPL and INTC, have relatively large jumps overall from 2000 to 2002, as shown in Figures 4.1 to 4.3 from Section 4.4.1.2. This indicates that considering news announcements in HAR-class models can improve the usefulness of data from highly volatile periods for forecasting, helping to solve the problem identified in the benchmark models above.

These results are reflected in IT stocks' large average news-related jumps (typically higher than 1) in the pre-crisis period, as shown in Table 4.17 in Section 4.4.2.1. Table 4.23 shows that separating stocks' jump components based on whether they co-occur with news surprises improves the pre-crisis forecasting results for all these stocks in at least one of the three HAR-family models. In addition, YHOO and BSX have relatively large average news-related jumps in the post-crisis period (see the right-hand panel of Table 4.17); the forecasting results for these stocks also improve in the modified version of the HAR-TJ model. Therefore, the forecasting results in Table 4.23 suggest that for many stocks, especially those with large news-related jump components, the modified versions of the HAR-family models that consider these large news-related jump components separately from other jump components yield better forecasting results.

[Insert Tables 4.24 and 4.25 here]

Comparing the results in Table 4.22 (benchmark HAR-family models) to those in Tables 4.24 and 4.25 (HAR-family models modified to account for the impact of news that co-occurs with co-jumps), we see that the modified models do not yield significant gains in forecasting, as the MSE results for most of the stocks in the three tables are similar to one another. We showed in Section 4.4.2.3 that co-jump-related news surprises have a significant effect on jump components; however, including this information in HAR models does not seem to improve forecasting. This may be because the number of co-movements that co-occur with announcements of news is relatively small. In addition, the announcement of news (or news-related co-jumps) does not seem to have a different effect on forecasting for high-volume stocks such as PFE and EXC compared to low-jump stocks such as BT and TEF.

4.5 Conclusion

In this chapter, we have examined the impact of news and co-jumps on the market's second-moment responses. Our results show that market-level news announcements have a significant impact on second-market responses for all stocks examined in this study, supporting previous literature that looks at similar issues in the S&P 500 index futures, US Treasury bond futures and long- and short-term US interest rates (e.g. Ederington & Lee 1993; Balduzzi et al. 2001; Wongswan 2006; Huang 2018).

Some stocks, such as BSX and XRX, are significantly affected by a wider variety of news announcements, while other stocks, such as VOD in the pre-crisis period and UL in the post-crisis period, are influenced by announcements from a narrower range of news outlets. We also find that positive surprises from the Consumer Price Index (CPI CHNG) and Initial Jobless Claims (INJCJC) index before the financial crisis, and negative surprise for the University of Michigan Consumer Sentiment Index (CONSSENT) in the post-crisis period, have a significant impact on most stocks' jump components. The jump components of the XRX and BSX stocks are more likely to be affected by macroeconomic news surprises compared to those of other stocks. In addition, news announcements from a wider range of outlets have a significant impact on jump components after the financial crisis compared to before it.

When considering co-jump-related news, we find that news surprises significantly affect the jump components of each stock when positive and negative news surprises for different news outlets are combined. In addition, we find that the co-jumps estimated from stocks from a diverse range of sectors are better at capturing the impact of market-level positive news surprises. However, some important negative market-level news surprises for IT sector stocks are better captured by the co-jumps estimated for these stocks only, as this sector seems to be highly sensitive to negative news from the SPCS20SM and MTIBCHNG outlets. Also, co-jump-related news has a bigger impact on stocks in the pre-crisis period than in the post-crisis period. This can be seen in the relatively large co-movements of stocks from 2000 to 2002 caused by the burst of the dot-com bubble. This period is marked by the announcement of market-level news that significantly affects the jump components for many stocks. Finally, there is little difference between stocks with large trading volumes and those with low jumps.

The benchmark forecasting results from the standard HAR models showed lower mean standard errors for the more stable post-crisis period than the relatively volatile pre-crisis period. By separating the large macroeconomic news-related jump components from 2000-2002 from those not related to macroeconomic news, and by incorporating these two variables into the (modified) HAR models, forecasting is notably improved for many stocks, especially BSX and most of those in the IT sector. This is because the jumps that co-occur with macroeconomic news in this period are large and have regular patterns, which means they can contribute significantly to forecasting models when considered separately. This can improve the usefulness of highly volatile data for forecasting, which may be a potential solution for the problem with benchmark models discussed in Section 4.4.3. However, we find that considering co-jump-related news surprises separately does not result in significant improvements for most of the forecasting results, since co-jumps are not able to capture the majority of market-level news that contributes to jump components.

The positive and negative news surprises for all news outlets together are significant in explaining the jump components; separating the jump components based on this combined news surprise value does not significantly improve the forecasting model. This is because the impact of news captured by co-jumps is limited, and so these separated jump components minimally contribute to the forecasting models.

This chapter contributes to the literature on stock volatility forecasting by incorporating information from macroeconomic news announcements into non-parametric HAR forecasting models. By running regressions of jumps on news announcements, we find that many news announcements have a significant effect on stocks' jump components. Thus, the forecasting results show that including news announcements in models can improve volatility forecasting and should be considered in future research. A useful direction for further work in this area would be to incorporate information from macroeconomic news announcements into models such as HAR-GARCH (Corsi et al. 2008) that combine parametric and non-parametric methods for volatility forecasting.

Appendix

Figures

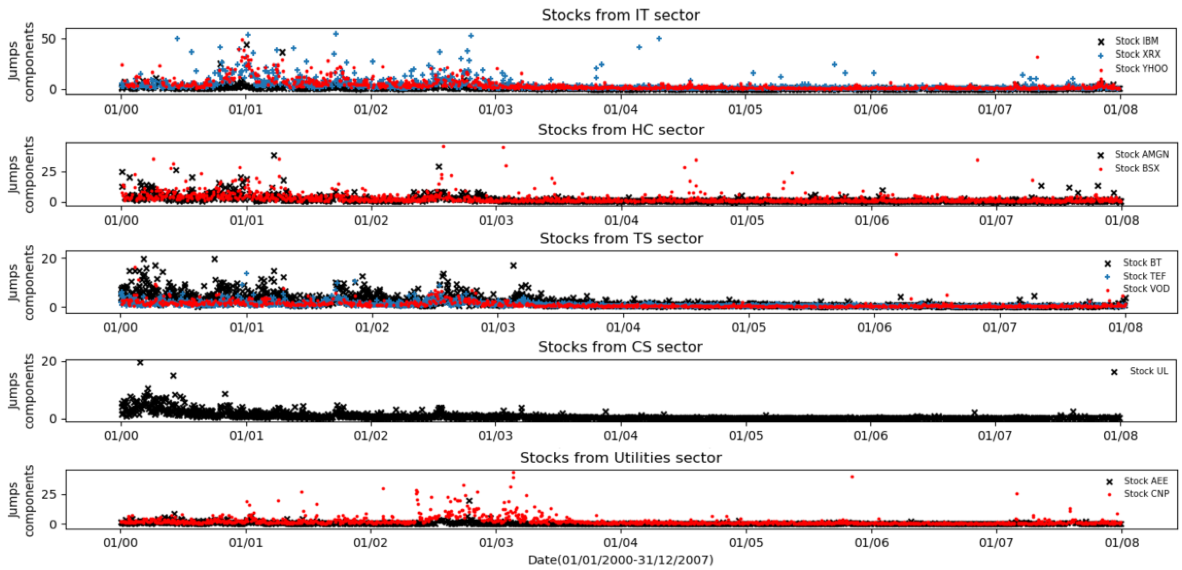

Figure 4.1: Jump components estimated using TBV for low jump frequency stocks in the pre-crisis period. Some extreme values for the XRX, YHOO, BSX, BT, VOD, TEF, UL and CNP stocks are discarded as they are too large to be displayed in the graph.

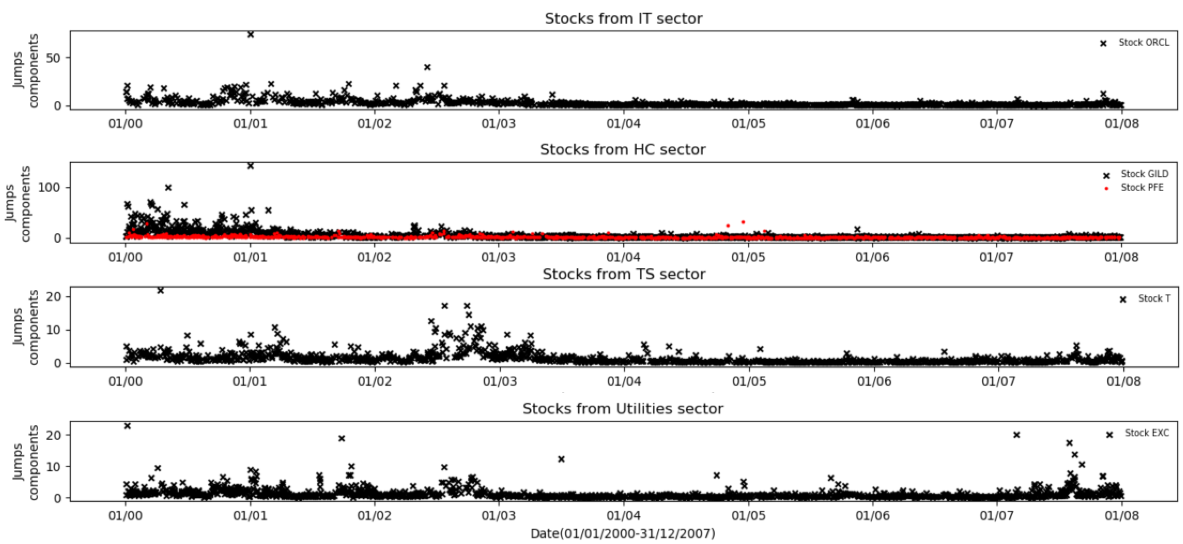

Figure 4.2: Jump components estimated using TBV for stock with high trading volumes in the pre-crisis period.

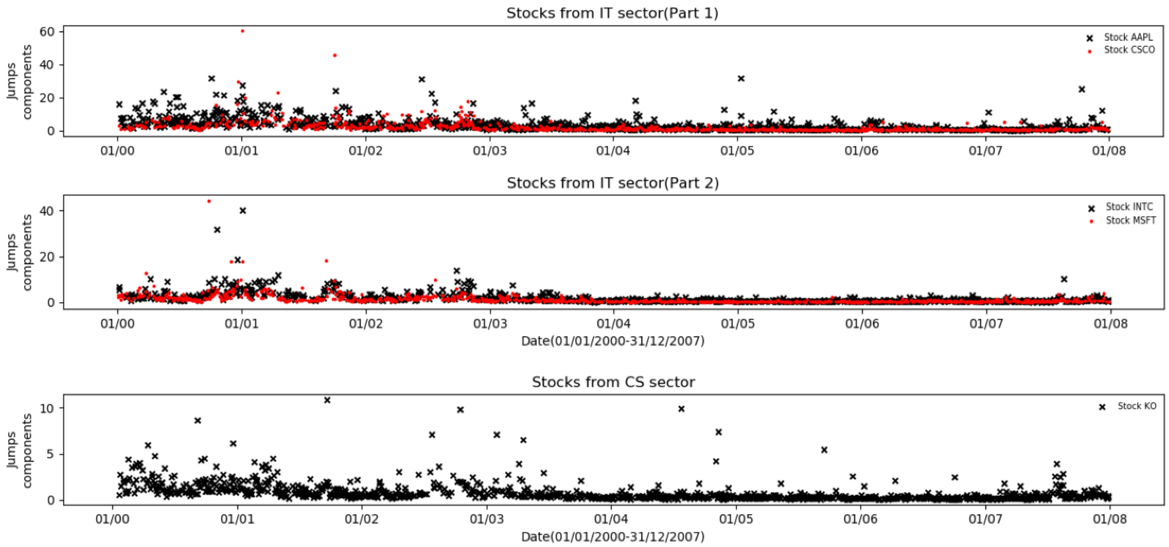

Figure 4.3: Jump components estimated using TBV for stocks with few jumps and high trading volumes in the pre-crisis period.

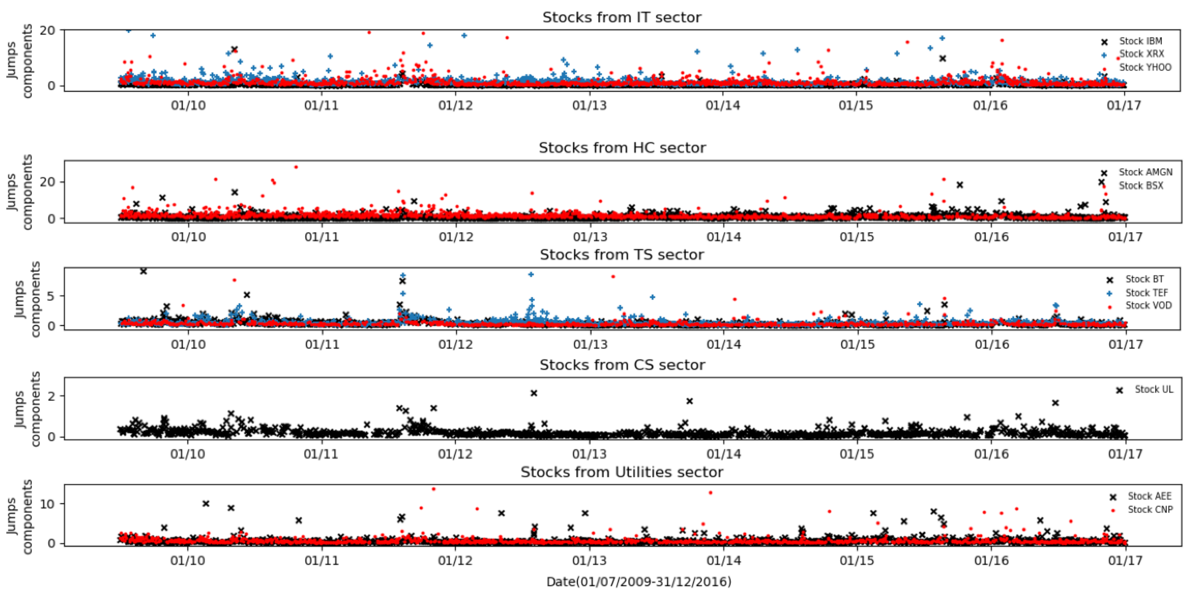

Figure 4.4: Jump components estimated using TBV for low jump frequency stocks in the post-crisis period. Some extreme values for each stock are discarded as they are too large to be displayed in the graph.

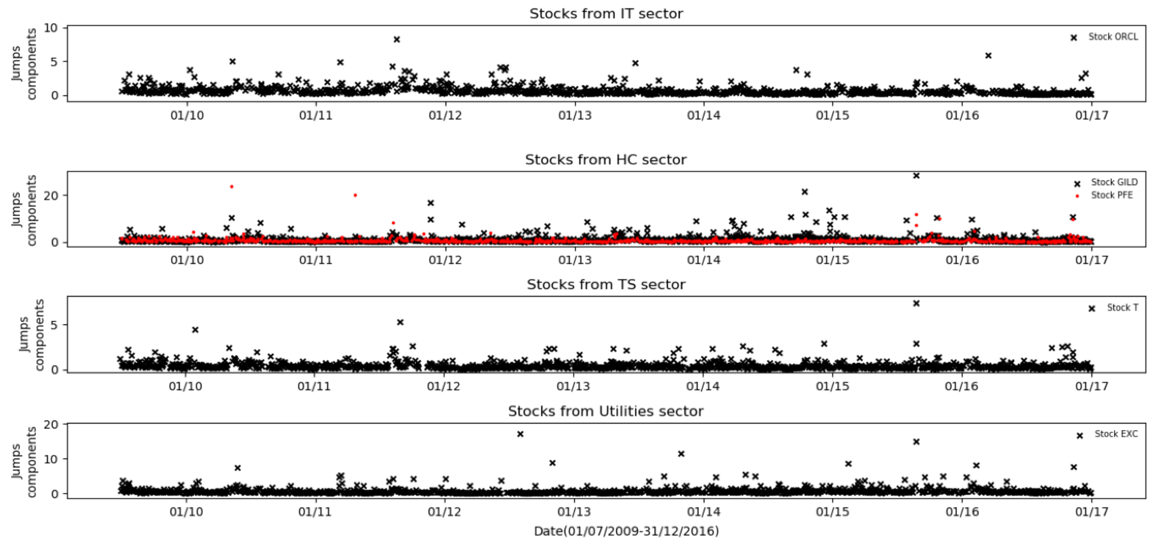

Figure 4.5: Jump components estimated using TBV for stocks with high trading volumes in the post-crisis period. Some extreme values for the ORCL, T and EXC stocks are discarded as they are too large to be displayed in the graph.

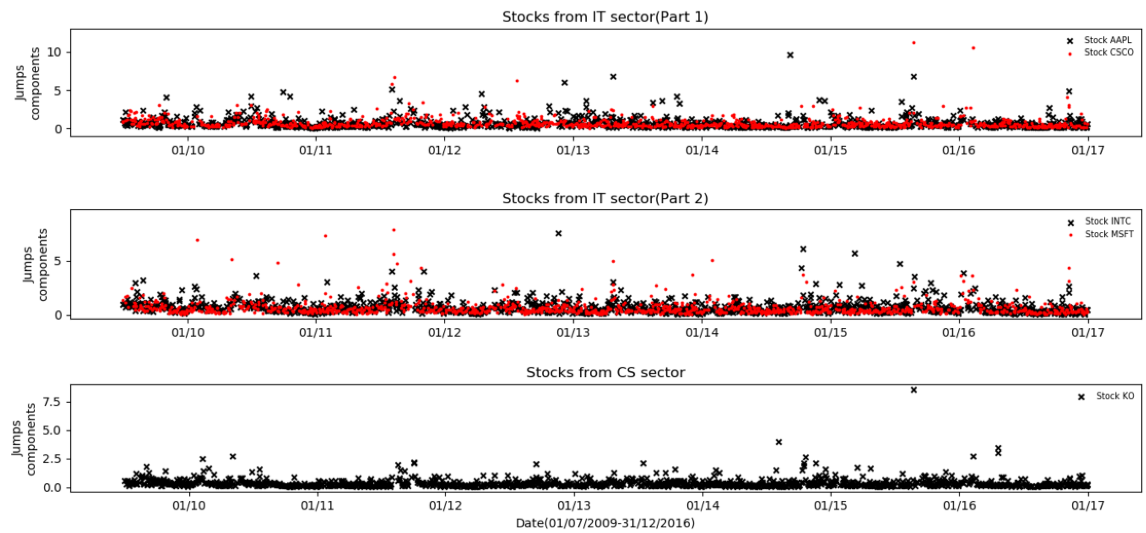

Figure 4.6: Jump components estimated using TBV for stocks with few jumps and high trading volumes in the post-crisis period. Some extreme values for the AAPL, CSCO, INTC and MSFT stocks are discarded as they are too large to be displayed in the graph.

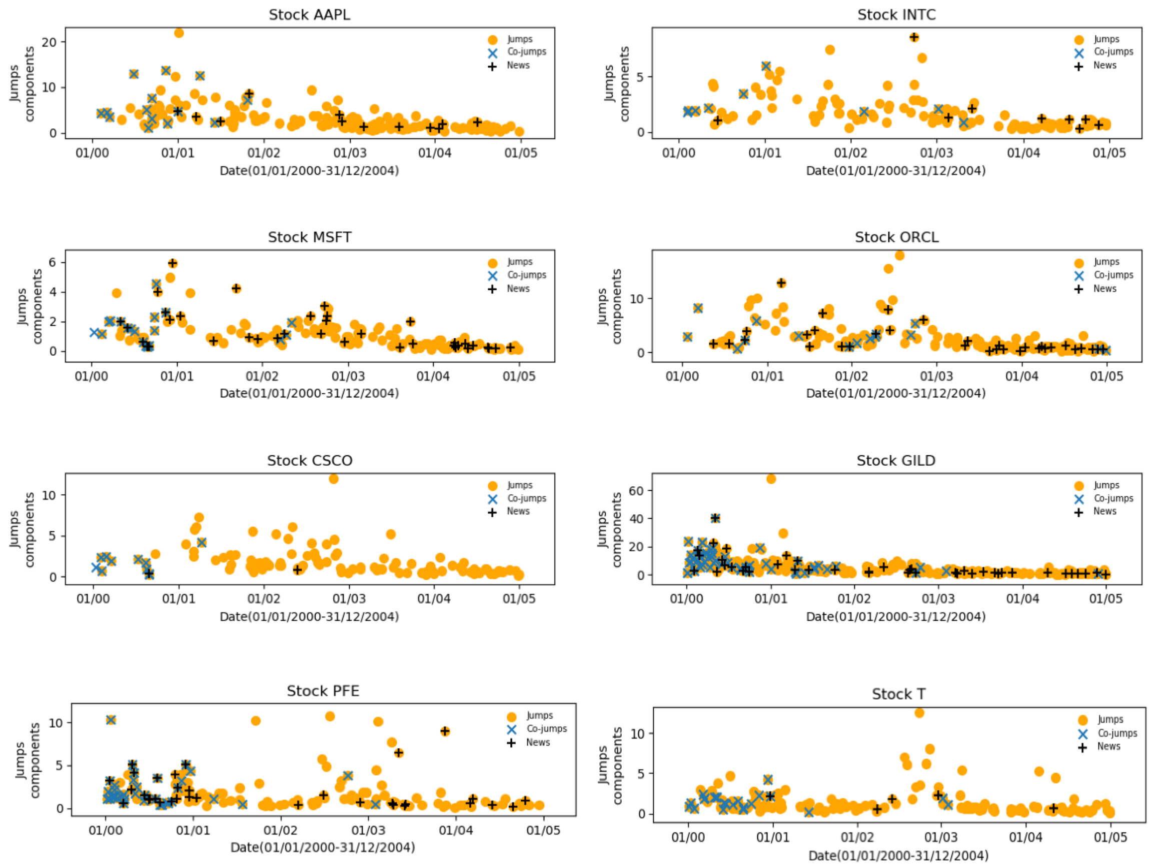

Figure 4.7: Jumps, co-jumps and news for the top six highest volume stocks overall (subset one) in the pre-crisis period. More than six stocks are shown here because there is variation in the top six stocks between the pre- and post-crisis periods. The list of news outlets is shown in Table 4.4. The co-jumps are estimated using the MCP co-jump test.

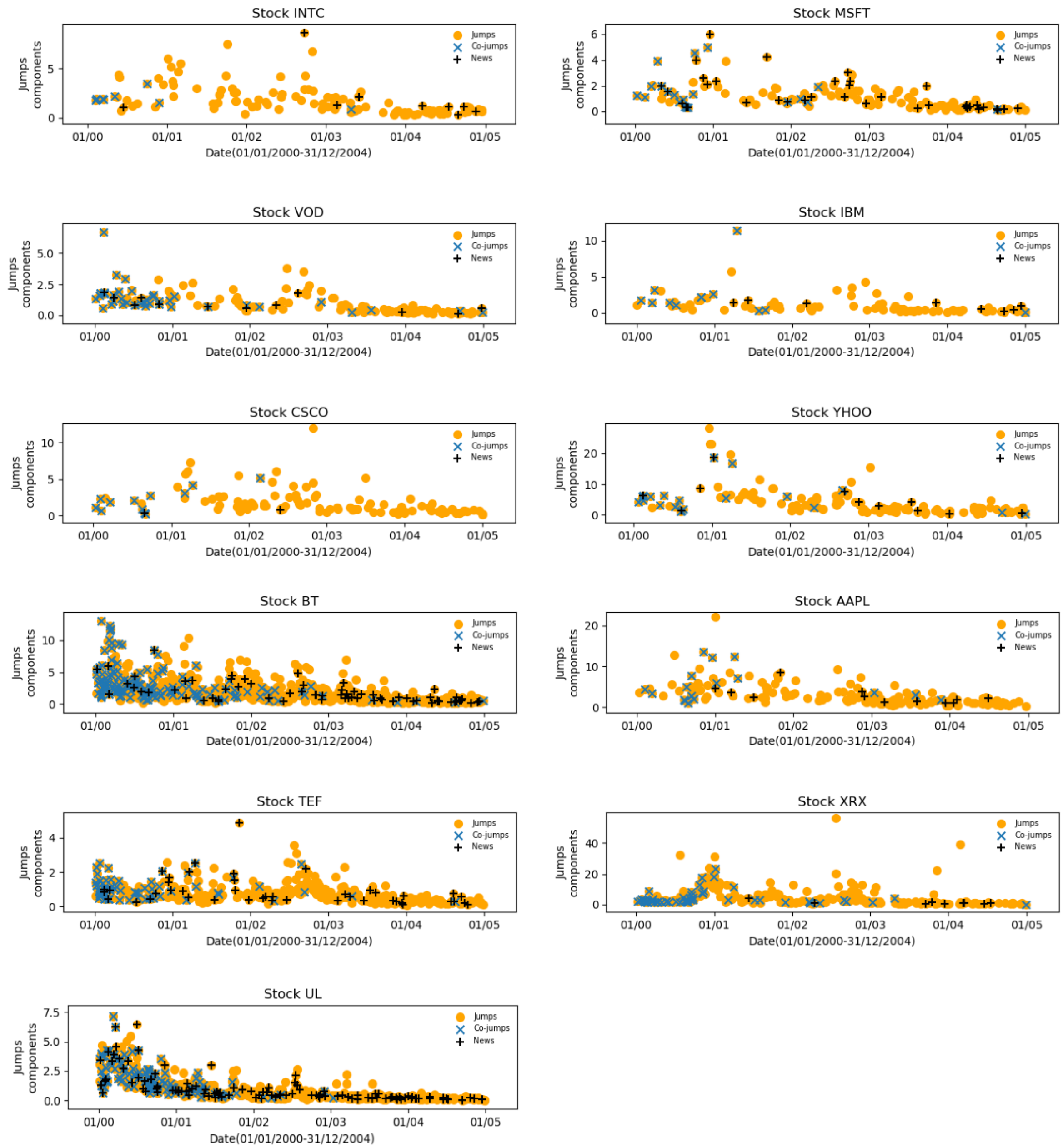

Figure 4.8: Jumps, co-jumps and news for the six stocks with the fewest jumps overall (subset one) in the pre-crisis period. More than six stocks are shown here because there is variation in the top six stocks between the pre- and post-crisis periods. The list of news outlets is shown in Table 4.4. The co-jumps are estimated using the MCP co-jump test.

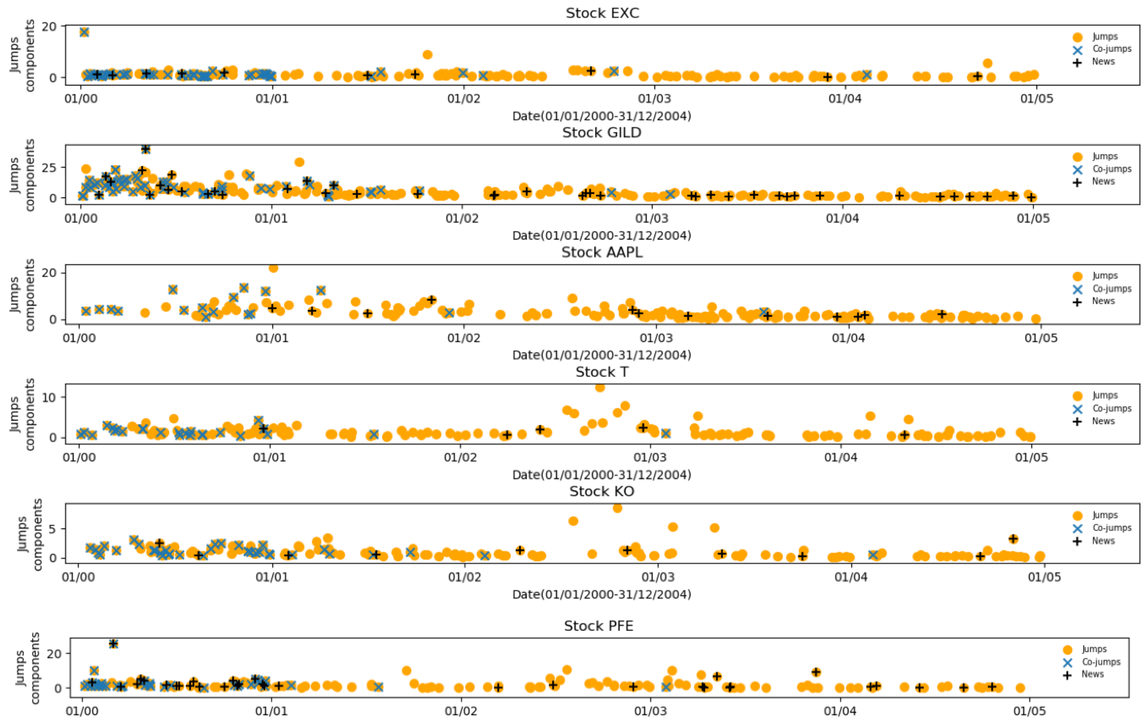

Figure 4.9: Jumps, co-jumps and news for the stocks with the highest volume in each sector (subset two) in the pre-crisis period. The list of news outlets is shown in Table 4.4. The co-jumps are estimated using the MCP co-jump test.

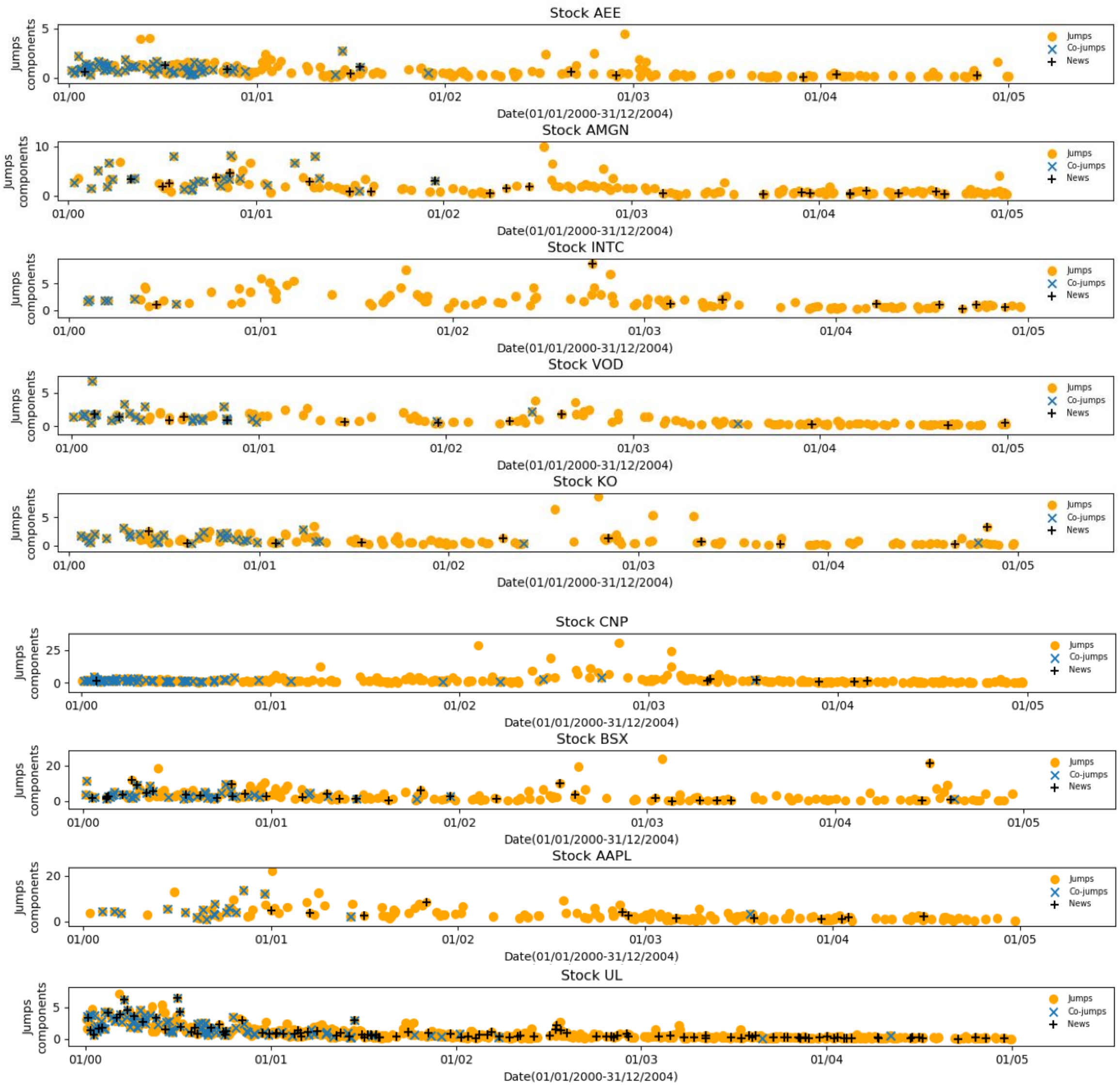

Figure 4.10: Jumps, co-jumps and news for the stocks with the fewest jumps in each sector (subset two) in the pre-crisis period. More than one stock per sector is shown here because there is variation in the top stocks between the pre- and post-crisis periods. The list of news outlets is shown in Table 4.4. The co-jumps are estimated using the MCP co-jump test.

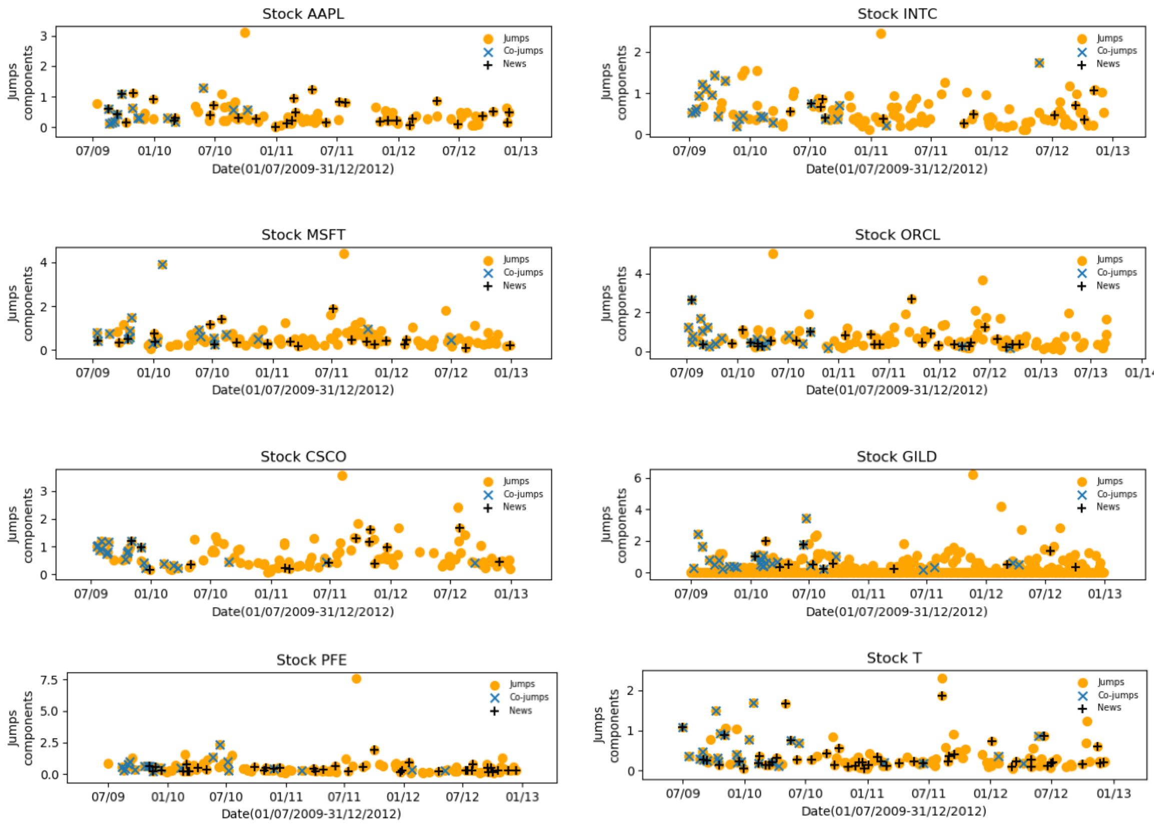

Figure 4.11: Jumps, co-jumps and news for the six stocks with the highest volumes overall (subset one) in the post-crisis period. More than six stocks are shown here because there is variation in the top six stocks between the pre- and post-crisis periods. The list of news outlets is shown in Table 4.4. The co-jumps are estimated using the MCP co-jump test.

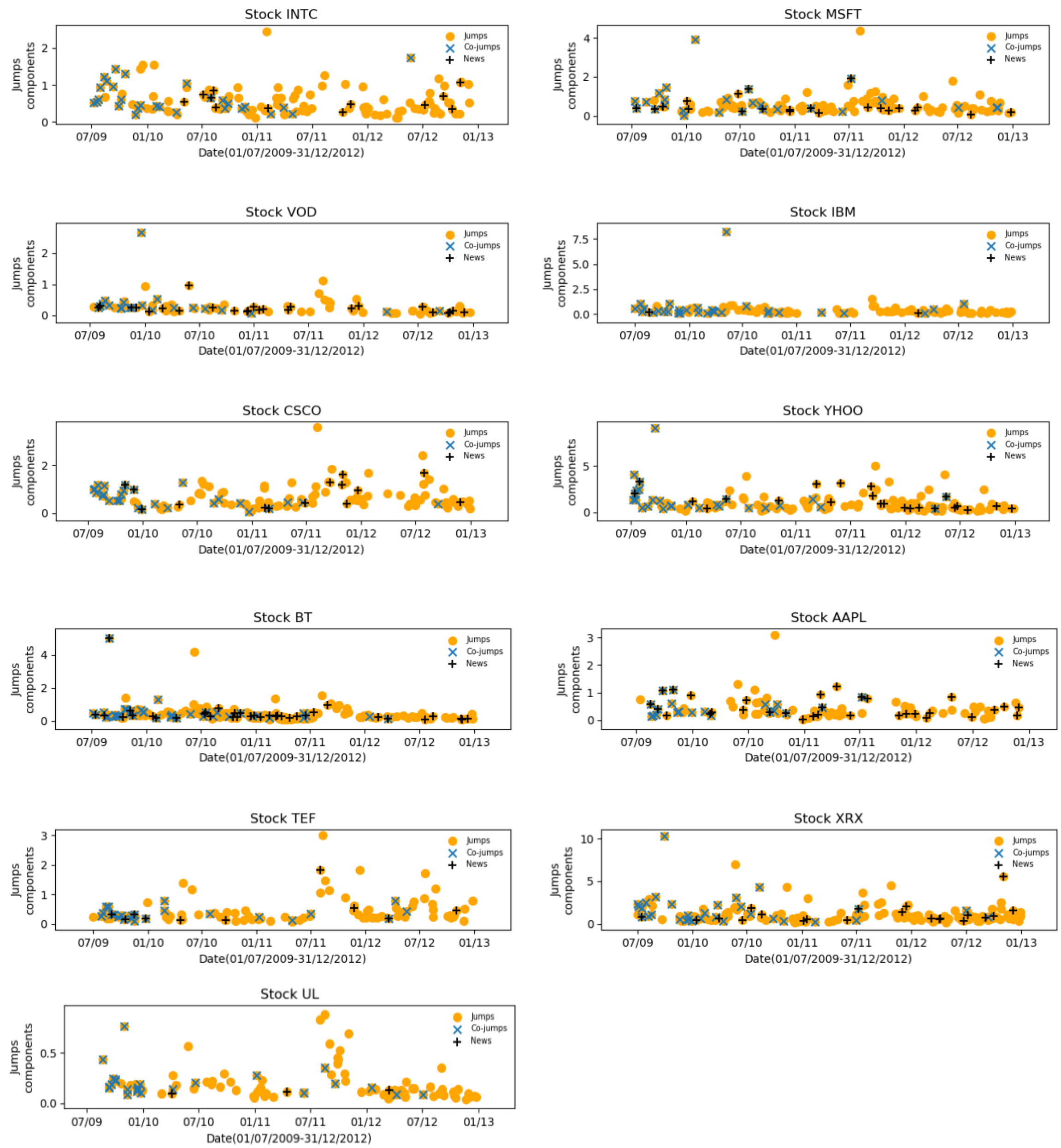

Figure 4.12: Jumps, co-jumps and news for the six stocks with the fewest jumps overall (subset one) in the post-crisis period. More than six stocks are shown here because there is variation in the top six stocks between the pre- and post-crisis periods. The list of news outlets is shown in Table 4.4. The co-jumps are estimated using the MCP co-jump test.

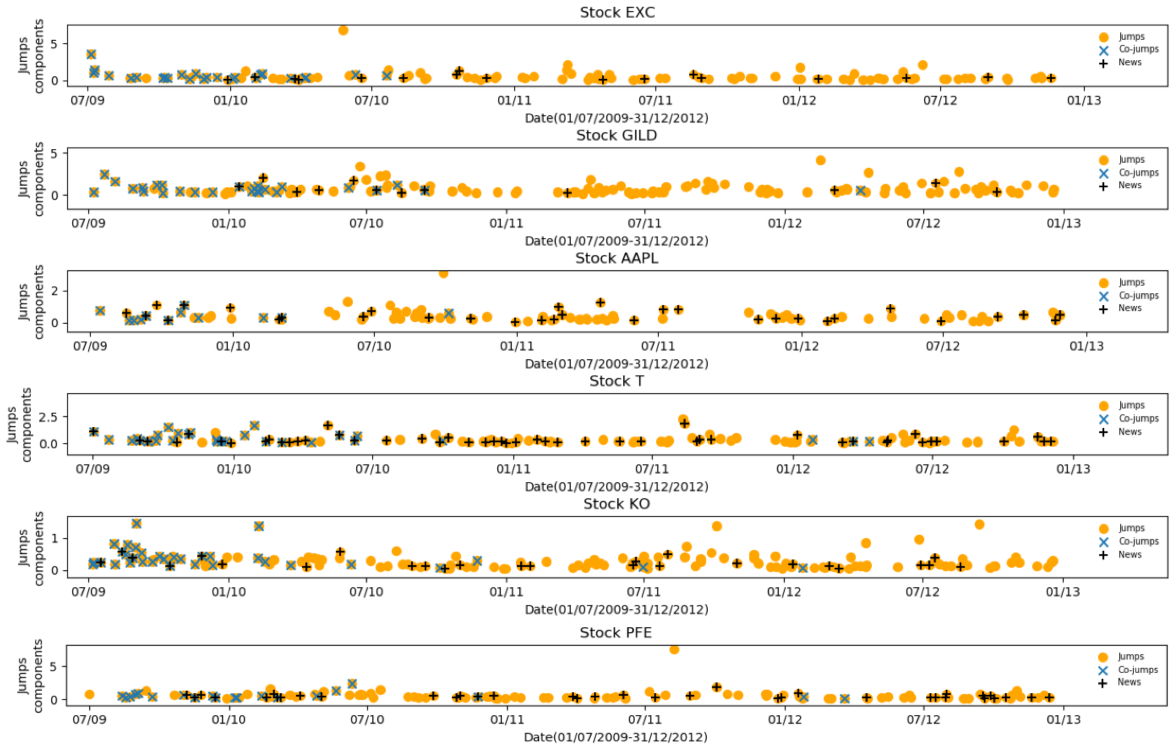

Figure 4.13: Jumps, co-jumps and news for the stocks with the highest volume in each sector (subset two) in the post-crisis period. The list of news outlets is shown in Table 4.4. The co-jumps are estimated using the MCP co-jump test.

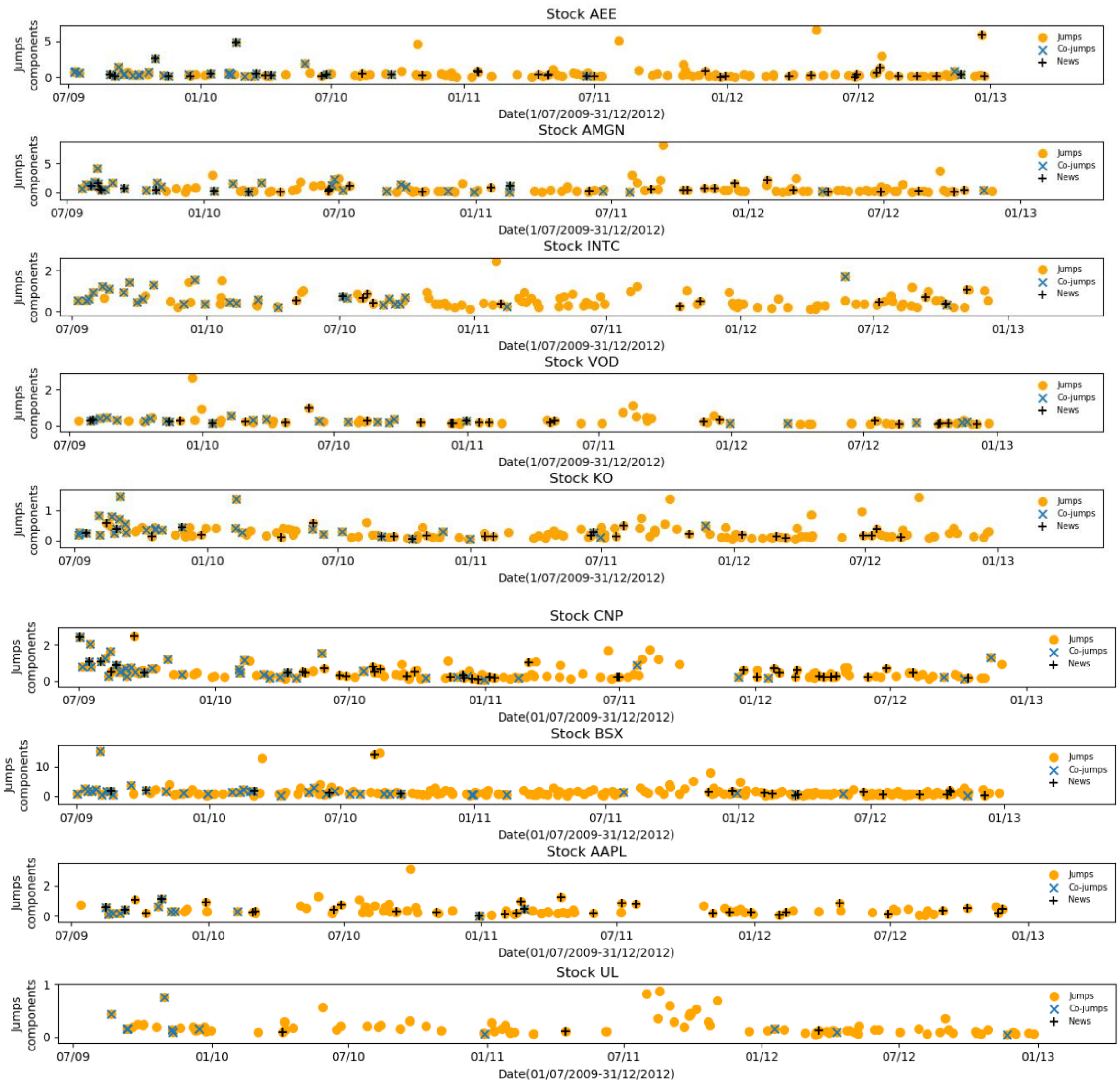

Figure 4.14: Jumps, co-jumps and news for the stocks with the fewest jumps in each sector (subset two) in the post-crisis period. More than one stock per sector is shown here because there is variation in the top stocks between the pre- and post-crisis periods. The list of news outlets is shown in Table 4.4. The co-jumps are estimated using the MCP co-jump test.

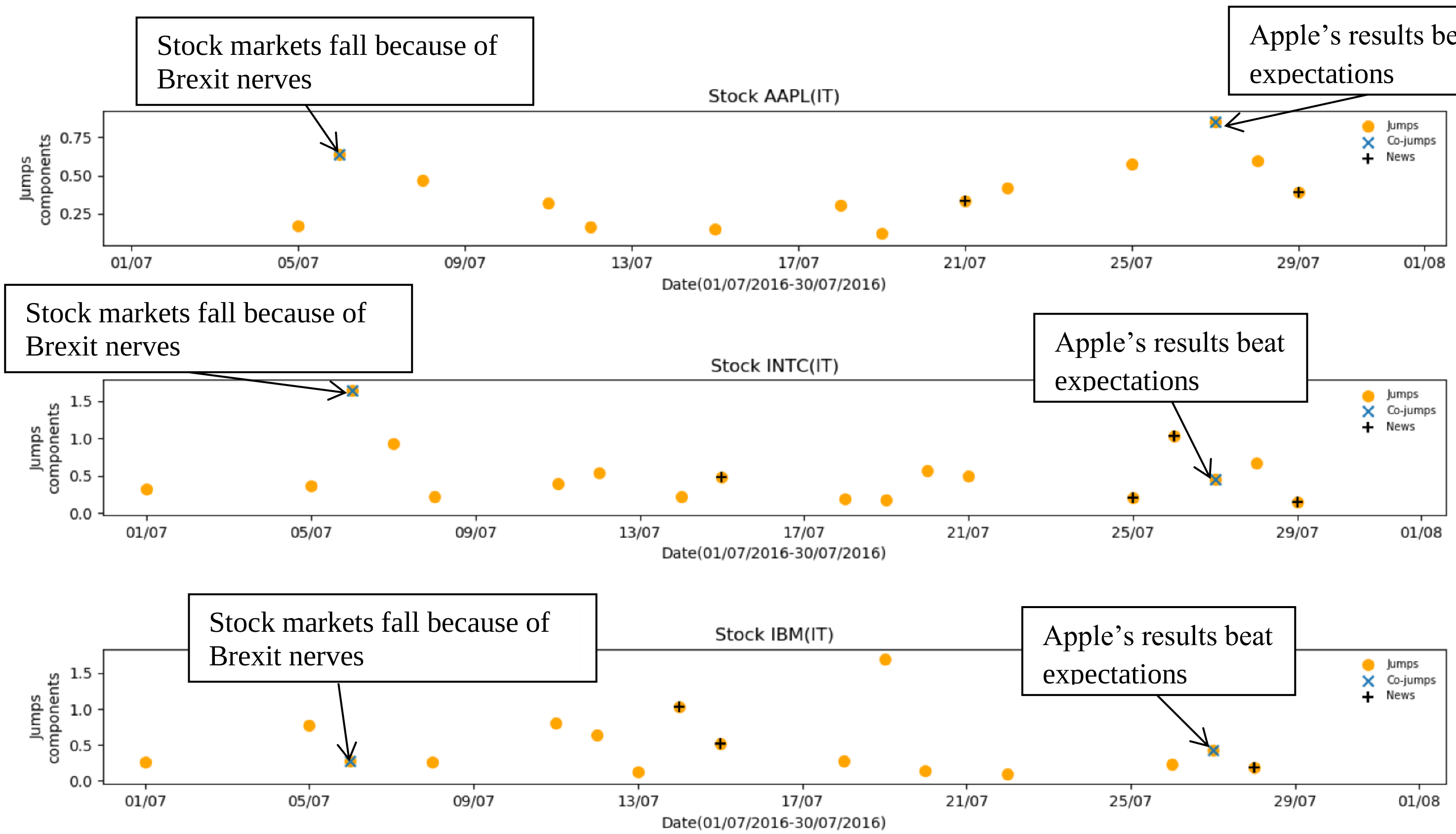

Figure 4.15: Jumps, co-jumps and news estimated using CTBV for three stocks in the IT sector (AAPL, INTC and IBM) in July 2016. Co-jumps are estimated using the MCP co-jump test based on the stocks from data subset one. The annotations show changes in the jump components that co-occur with news.

Tables

Table 4.1 Volume, realised volatility and jumps for stocks from five industrial sectors in the pre-crisis period

	Volume (in 100000)	**RV**	**Jumps**		**Volume (in 100000)**	**RV**	**Jumps**
	HC Sector				***IT Sector***		
ABT	40.467	2.761	183	**AAPL**	**<u>1325.400</u>**	8.083	170
AMGN	93.262	4.851	***168***	**CSCO**	**<u>575.390</u>**	6.158	***<u>103</u>***
BSX	60.276	5.378	256	**EBAY**	201.100	9.779	169
GILD	**<u>222.350</u>**	10.101	267	**HPQ**	104.170	5.476	184
HUM	12.153	8.914	299	**IBM**	66.218	2.455	***<u>134</u>***
JNJ	76.434	1.692	213	**INTC**	**<u>559.200</u>**	5.600	***<u>81</u>***
MDT	41.984	2.751	234	**MSFT**	**<u>642.640</u>**	3.208	***<u>118</u>***
MRK	77.056	2.505	207	**ORCL**	**<u>415.130</u>**	7.890	174
PFE	206.990	2.687	197	**XRX**	44.267	8.888	224
UNH	72.903	2.799	231	**YHOO**	213.030	11.190	***<u>145</u>***
	TS Sector				***CS Sector***		
AMT	20.347	17.360	380	**AVP**	24.239	3.074	253
BT	1.806	3.106	512	**BFB**	8.291	1.501	387
CHL	7.976	2.521	446	**COST**	41.027	4.501	240
CTL	7.832	3.086	244	**EL**	20.414	3.045	352
FTR	15.108	5.925	345	**KMB**	15.931	1.977	189
LVLT	7.775	24.788	325	**KO**	**<u>107.920</u>**	1.908	***148***
T	**102.950**	3.405	170	**PEP**	41.056	2.192	160
TEF	13.429	2.023	141	**PG**	72.613	1.811	174
VOD	15.800	2.650	***<u>94</u>***	**UL**	4.127	1.699	195
VZ	71.167	2.863	215	**WMT**	100.190	2.777	171
	Utilities Sector						
AEE	6.420	1.676	***185***	**EXC**	**24.714**	2.661	213
AEP	14.680	3.054	213	**OKE**	9.416	2.822	525
CNP	16.378	6.320	266	**PCG**	16.011	6.675	247
DUK	13.639	3.358	211	**PEG**	19.820	2.533	207
ETR	9.995	2.442	244	**SO**	20.461	2.219	194

Note. The table indicates the stocks with the highest volume (bold) and the stocks with the fewest jumps (bold and italics) in each sector. The six stocks with the highest volume (bold and underline) and fewest jumps (bold, italics and underline) overall are also shown.

Table 4.2 Volume, realised volatility and jumps for five sector stocks in post-crisis period

	Volume (in 100000)	RV	Jumps		Volume (in 100000)	RV	Jumps
	HC Sector				***IT Sector***		
ABT	66.357	1.124	322	**AAPL**	**<u>891.740</u>**	1.661	***<u>189</u>***
AMGN	42.244	1.874	330	**CSCO**	**<u>382.300</u>**	1.601	258
BSX	142.750	3.319	***218***	**EBAY**	114.680	2.468	285
GILD	128.09	2.295	310	**HPQ**	158.740	2.582	267
HUM	16.158	2.840	405	**IBM**	42.836	0.938	271
JNJ	85.417	0.795	312	**INTC**	**<u>386.820</u>**	1.719	253
MDT	52.627	1.33	333	**MSFT**	**<u>411.220</u>**	1.474	245
MRK	116.260	1.289	282	**ORCL**	206.930	1.621	240
PFE	**<u>330.860</u>**	1.397	254	**XRX**	103.730	3.057	***<u>205</u>***
UNH	57.095	2.261	377	**YHOO**	184.720	2.733	283
	TS Sector				***CS Sector***		
AMT	24.741	1.436	329	**AVP**	53.650	6.704	286
BT	2.384	0.998	***<u>90</u>***	**BFB**	8.034	1.546	332
CHL	10.468	0.693	247	**COST**	23.099	1.060	347
CTL	43.346	1.655	280	**EL**	20.817	1.676	307
FTR	113.390	4.161	307	**KMB**	18.574	0.777	367
LVLT	17.354	9.506	237	**KO**	**140.820**	0.771	308
T	**<u>227.180</u>**	0.938	274	**PEP**	51.211	0.764	355
TEF	18.631	1.158	***<u>92</u>***	**PG**	85.738	0.747	316
VOD	42.476	0.834	***<u>88</u>***	**UL**	11.635	0.582	***<u>101</u>***
VZ	139.830	0.974	293	**WMT**	89.194	0.818	323
				Utilities Sector			
AEE	15.595	1.278	304	**EXC**	**54.297**	1.495	274
AEP	28.029	1.119	328	**OKE**	15.812	2.943	230
CNP	35.506	1.922	***213***	**PCG**	24.523	1.130	336
DUK	28.292	0.969	343	**PEG**	27.316	1.358	289
ETR	12.288	1.232	340	**SO**	41.143	0.861	319

Note. The table indicates the stocks with the highest volume (bold) and the stocks with the fewest jumps (bold and italics) in each sector. The six stocks with the highest volume (bold and underline) and fewest jumps (bold, italics and underline) overall are also shown.

Table 4.3 Two subsets of stocks used in the analysis

High volume		Low jump frequency	
Pre-crisis	**Post-crisis**	**Pre-crisis**	**Post-crisis**
Subset one: chosen based on all stocks			
AAPL (IT)	INTC (IT)	INTC (IT)	VOD (TS)
MSFT (IT)	VOD (TS)	VOD (TS)	BT (TS)
CSCO (IT)	CSCO (IT)	CSCO (IT)	TEF (TS)
INTC (IT)	MSFT (IT)	MSFT (IT)	UL (CS)
ORCL (IT)	IBM (IT)	IBM (IT)	AAPL (IT)
GILD (HC)	YHOO (IT)	YHOO (IT)	XRX (IT)
Subset two: chosen based on each sector			
GILD (HC)	PFE (HC)	AMGN (HC)	BSX (HC)
AAPL (IT)	AAPL (IT)	INTC (IT)	AAPL (IT)
T (HC)	T (HC)	VOD (HC)	VOD (HC)
KO (CS)	KO (CS)	KO (CS)	UL (CS)
EXC (Utility)	EXC (Utility)	AEE (Utility)	CNP (Utility)

Note. Subset one comprises the 13 stocks with the highest volume and those with the fewest jumps in the whole data set. Subset two comprises the two stocks with the highest volume and those with the fewest jumps for each industrial sector. Stocks GILD, AAPL, INTC, VOD and UL are shown in both subsets. Therefore, there are 21 stocks in these two subsets are used in our analysis.

Table 4.4 Descriptive statistics for macroeconomic news outlets

News	Description	Positive				Negative (Absolute values)			
		Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
ADP CHNG	ADP Employment Change	0.749	0.764	0.019	3.696	0.659	0.677	0.038	3.715
NHSPATOT	Building Permits	0.889	0.755	0.017	3.353	0.672	0.482	0.017	1.981
MTIBCHNG	Business Inventories	0.944	0.578	0.464	2.784	0.910	0.632	0.464	3.711
CGNOXAI%	Manufacturers' new orders for non-defence capital goods excluding aircrafts	0.607	0.671	0.050	3.172	0.850	0.742	0.050	2.971
CPTICHNG	Capacity Utilisation	0.825	0.505	0.297	2.377	0.900	0.696	0.297	4.457
USEMNCHG	Change in Household Employment	0.765	0.670	0.071	1.904	0.845	0.558	0.011	1.818
USMMMNCH	Change in Manufacturing Payrolls	0.568	0.482	0.047	2.244	0.895	0.862	0.047	4.254
NFP TCH	Change in Non-farm Payrolls	0.653	0.555	0.013	2.423	0.850	0.722	0.013	4.098
NFP PCH	Change in Private Payrolls	0.779	0.530	0.050	2.176	0.784	0.633	0.033	2.309
CFNAI	Chicago Fed Nat Activity Index	0.706	0.472	0.035	1.688	0.929	0.729	0.035	2.602
CONCCONF	Consumer confidence index	0.803	0.601	0.039	2.403	0.812	0.613	0.020	2.735
INJCSP	Continuing Claims	0.641	0.648	0.016	5.293	0.741	0.832	0.016	9.283
GDPCPCEC	Core Personal Consumption Expenditures (PCE) (QoQ)	0.824	0.614	0.026	2.496	0.752	0.654	0.013	2.909
CPUPXCHG	Consumer Price Index (CPI) excluding Food & Energy (MoM)	0.641	0.648	0.016	5.293	0.741	0.832	0.016	9.283
CPI CHNG	Consumer Price Index (CPI) (MoM)	1.137	0.633	0.735	3.674	1.124	0.579	0.735	2.939
USCABAL	Current Account Balance	1.220	0.342	1.104	2.207	1.246	0.373	1.104	2.207
DFEDGBA	Texas' Manufacturing Activity	1.090	0.582	0.788	3.151	1.135	0.551	0.788	3.151
DGNOCHNG	Durable Goods Orders	0.916	0.604	0.048	2.359	0.703	0.597	0.032	2.613
DGNOXTCH	Durables excluding Transportation	0.685	0.525	0.040	1.818	0.898	0.698	0.013	2.713
EMPRGBCI	Empire Manufacturing	0.677	0.804	0.036	5.294	0.754	0.656	0.036	2.974
ECI SA%	Employment Cost Index	0.733	0.668	0.066	3.876	0.834	0.643	0.066	3.022
ETSLTOTL	Existing Home Sales	0.806	0.531	0.032	2.150	0.816	0.639	0.021	2.646
TMNOCHNG	Factory Orders	1.027	0.760	0.597	2.983	0.937	0.499	0.597	2.386
HPIMMOM%	FHFA House Price Index (MoM)	0.704	0.534	0.050	2.059	0.779	0.830	0.050	4.117
GDP CQOQ	GDP Annualised (QoQ)	0.885	0.852	0.160	4.163	0.610	0.599	0.160	3.203
GDP DCHG	GDP Price Deflator	0.817	0.689	0.209	3.548	0.835	0.719	0.209	3.548
GDP PIQQ	GDP Price Index	0.687	0.588	0.273	3.000	0.756	1.120	0.273	5.455
NHSPSTOT	Housing Starts	0.638	0.594	0.262	2.885	1.110	1.014	0.262	3.934
ECONUSIB	IBD/TIPP Economic Optimism	0.805	0.636	0.024	3.125	0.736	0.629	0.024	3.089
IMP1CHNG	Import Price Index (MoM)	0.698	0.600	0.038	3.377	0.876	0.687	0.038	3.300
IP CHNG	Industrial Production (MoM)	0.800	0.530	0.176	2.286	0.800	0.750	0.176	3.868
INJCJC	Initial Jobless Claims	0.811	0.502	0.270	2.975	0.901	0.725	0.270	5.409

Note. MoM, QoQ and YoY refer to month-over-month, quarter-over-quarter and year-over-year respectively

Cont. … Table 4.4 Descriptive statistics for macroeconomic news outlets

News	Description	Positive				Negative (Absolute values)			
		Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
NAPMPMI	ISM Manufacturing	0.785	0.711	0.055	4.388	0.720	0.652	0.055	4.552
MAPMINDX	ISM Milwaukee	0.823	0.675	0.052	3.869	0.746	0.628	0.052	3.137
NAPMNMAN	ISM Non-Manufacturing	0.751	0.644	0.006	2.534	0.811	0.613	0.027	1.913
NAPMNMI	ISM Non-Manufacturing Index	0.802	0.570	0.029	2.394	0.807	0.584	0.058	3.202
NAPMPRIC	ISM Prices Paid	0.738	0.580	0.051	2.104	0.804	0.728	0.103	4.054
LEI CHNG	Leading Index	0.758	0.470	0.201	1.809	0.992	0.580	0.201	2.010
CHPMINDX	MNI Chicago PMI	1.114	0.668	0.543	2.713	0.934	0.512	0.543	2.713
USHBMIDX	NAHB Housing Market Index	0.503	1.106	0.008	10.585	0.390	0.554	0.008	3.096
FRNTTOTL	Net Long-term TIC Flows	0.986	0.654	0.375	3.004	0.834	0.569	0.375	3.755
NHSLTOT	New Home Sales	0.807	0.617	0.031	2.607	0.774	0.620	0.033	3.311
SBOITOTL	NFIB Small Business Optimism	0.777	0.781	0.017	4.183	0.655	0.644	0.051	2.846
PRODNFR%	Nonfarm Productivity	0.582	0.524	0.069	2.426	0.906	0.823	0.069	3.465
PCE CMOM	Core Personal Consumption Expenditures Deflator (MoM)	0.876	0.798	0.153	4.581	0.659	0.589	0.153	2.901
GDPCTOT%	Personal Consumption	0.794	0.608	0.025	2.797	0.696	0.746	0.025	3.852
PITLCHNG	Personal Income	0.883	0.593	0.267	3.469	0.833	0.668	0.267	3.736
PCE CRCH	Personal Spending	0.846	1.121	0.347	6.255	0.632	0.537	0.347	4.170
OUTFGAF	Philadelphia Fed Business Outlook	1.002	0.552	0.605	3.628	1.016	0.672	0.605	4.838
PXFECHNG	Producer Price Index (PPI) excluding Food & Energy (MoM)	0.723	0.556	0.011	2.459	0.835	0.707	0.011	3.605
PPI CHNG	Producer Price Index (PPI) (MoM)	0.878	0.676	0.387	4.261	0.974	0.773	0.387	3.874
RSTAMOM	Retail Sales Advance (MoM)	0.880	0.725	0.208	3.537	0.779	0.596	0.208	2.497
RSTAXAG%	Retail Sales Excluding Auto and Gas	0.760	0.972	0.167	7.682	0.697	0.549	0.167	2.672
RSTAXMOM	Retail Sales Excluding Auto (MoM)	0.686	0.483	0.307	1.840	1.073	0.617	0.307	3.066
RCHSINDX	Richmond Fed Manufacturing Index	0.815	0.663	0.222	3.107	0.844	0.661	0.222	3.773
SPCS20SM	S&P CoreLogic Case-Shiller Home Price SA Index (MoM)	0.775	0.684	0.115	3.577	0.868	0.525	0.115	2.308
SPCS20Y%	S&P CoreLogic Case-Shiller Home Price NSA Index (YoY)	0.984	0.788	0.257	3.861	0.783	0.528	0.257	1.802
USTBTOT	Trade Balance	0.791	0.666	0.030	3.219	0.766	0.656	0.030	2.946
CONSSENT	The University of Michigan Consumer Sentiment Index	0.655	0.589	0.034	3.158	0.814	0.746	0.034	3.398
USURTOT	Unemployment Rate	1.012	0.504	0.679	2.715	1.113	0.606	0.679	3.393
COSTNFR%	Unit Labour Costs	0.698	0.611	0.097	3.283	0.750	0.789	0.097	4.635
MWINCHNG	Wholesale Inventories (MoM)	0.842	0.646	0.215	3.869	0.841	0.571	0.215	2.794

Note. MoM, QoQ and YoY refer to month-over-month, quarter-over-quarter and year-over-year respectively

Table 4.5 Mean volumes and jumps for stocks before and after the financial crisis

Stock (Sector)	Pre-crisis Volume (in 100000)	Pre-crisis Total Jumps	Pre-crisis Jumps per day	Post-crisis Volume (in 100000)	Post-crisis Total Jumps	Post-crisis Jumps per day
			Panel A. Few jumps			
IBM (IT)	66.22	135	0.07	42.84	271	0.14
XRX (IT)	44.27	230	0.11	103.73	205	0.11
YHOO (IT)	213.03	147	0.07	184.72	283	0.15
AMGN (HC)	93.26	173	0.09	42.24	330	0.17
BSX (HC)	60.28	264	0.13	330.86	254	0.13
BT (TS)	5.60	26	0.01	2.38	90	0.05
TEF (TS)	14.75	19	0.01	18.63	92	0.05
VOD (TS)	15.80	95	0.05	42.48	88	0.05
UL (CS)	4.13	216	0.11	11.64	101	0.05
AEE (Utilities)	6.42	194	0.10	15.60	304	0.16
CNP (Utilities)	16.38	283	0.14	35.51	213	0.11
			Panel B. High volume			
ORCL (IT)	415.13	179	0.09	206.93	240	0.13
GILD (HC)	222.35	284	0.14	128.09	310	0.16
PFE (HC)	206.99	200	0.10	330.86	254	0.13
T (TS)	102.95	172	0.09	227.18	274	0.14
EXC (Utilities)	24.71	220	0.11	54.30	274	0.14
			Panel C. High volume & few jumps			
AAPL (IT)	1325.40	173	0.09	891.74	189	0.10
CSCO (IT)	575.39	106	0.05	382.30	258	0.14
INTC (IT)	559.20	87	0.04	386.82	253	0.13
MSFT (IT)	642.64	120	0.06	411.22	245	0.13
KO (CS)	107.92	153	0.08	140.82	308	0.16
Average	224.90	166	0.08	190.04	230	0.12

Note. The table reports the mean trading volume, and number of intraday jumps (total and daily) estimated using the Lee-Mykland (2008) intraday jump test for each stock in the pre- and post-crisis periods. There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.6 Descriptive statistics for stocks' jump components estimated using BV in the pre-crisis period

Stock (Sector)	Full Max	Full Mean	Full Variance	Full Mode	Up to 90th percentile Max	Up to 90th percentile Mean	Up to 90th percentile Variance	Up to 90th percentile Mode	% change in variance
					Panel A. Few jumps				
IBM(IT)	25.398	0.866	4.719	0.057	1.692	0.457	0.145	0.057	96.9%
XRX(IT)	**170.44**	3.771	**96.395**	0.093	9.026	1.889	3.980	0.093	95.9%
YHOO(IT)	28.382	2.844	15.655	0.095	6.477	1.839	2.700	0.095	82.8%
AMGN(HC)	10.014	1.396	2.628	0.125	3.366	0.951	0.552	0.125	79.0%
BSX(HC)	37.868	2.598	13.400	0.081	5.336	1.707	1.636	0.081	87.8%
BT(TS)	12.950	1.628	3.026	0.049	3.706	1.173	0.825	0.049	72.8%
TEF(TS)	125.990	0.827	29.489	0.047	***1.270***	***0.448***	***0.094***	0.047	**99.7%**
VOD(TS)	***6.680***	0.727	***0.712***	0.058	1.776	0.512	0.198	0.058	***72.1%***
UL(CS)	15.169	0.889	1.381	0.043	2.309	0.583	0.263	0.043	81.0%
AEE(Utilities)	12.179	***0.660***	0.771	0.044	1.343	0.479	0.118	0.044	84.7%
CNP(Utilities)	124.86	2.200	72.283	0.056	3.248	0.979	0.526	0.056	99.3%
					Panel B. High volume				
ORCL(IT)	18.003	1.753	5.527	0.134	3.937	1.106	0.759	0.134	86.3%
GILD(HC)	68.554	**4.397**	38.297	**0.213**	**10.232**	**2.864**	**6.718**	**0.213**	82.5%
PFE(HC)	25.584	1.347	5.219	0.064	3.159	0.788	0.435	0.064	91.7%
T(TS)	12.566	1.035	1.846	0.069	2.289	0.691	0.268	0.069	85.5%
EXC(Utilities)	17.676	1.047	2.232	0.077	1.933	0.715	0.209	0.077	90.6%
					Panel C. High volume & few jumps				
AAPL(IT)	22.044	2.466	7.013	0.147	5.482	1.775	1.586	0.147	77.4%
CSCO(IT)	11.993	1.205	2.065	0.064	2.618	0.824	0.360	0.064	82.6%
INTC(IT)	8.636	1.211	1.774	0.144	2.714	0.849	0.393	0.144	77.8%
MSFT(IT)	25.95	0.940	3.421	0.067	1.955	0.611	0.236	0.067	93.1%
KO(CS)	8.607	0.776	1.027	***0.026***	1.859	0.519	0.178	***0.026***	82.6%

Note. The table reports the descriptive statistics for jump components estimated using BV for each stock in the pre-crisis period. The bold and italic values indicate the largest and smallest values respectively across different stocks for each column. The panel on the left shows the descriptive statistics for the jump components for the whole data set. The panel on the right shows the descriptive statistics for the jump components after eliminating the highest decile in the data set (leaving only the bottom 90%). There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.7 Descriptive statistics for stocks' jump components estimated using BV in the post-crisis period

Stock (sector)	Full data: Max	Full data: Mean	Full data: Variance	Full data: Mode	Up to 90th percentile: Max	Up to 90th percentile: Mean	Up to 90th percentile: Variance	Up to 90th percentile: Mode	% change in variance
				Panel A. Few jumps					
IBM(IT)	8.207	0.365	0.307	0.053	0.731	0.268	0.025	0.053	91.8%
XRX(IT)	24.395	1.122	4.397	**0.142**	1.930	0.703	0.153	**0.142**	96.5%
YHOO(IT)	22.556	1.081	3.568	0.104	2.067	0.703	0.188	0.104	94.7%
AMGN(HC)	13.125	0.869	1.439	0.123	1.687	0.590	0.176	0.123	87.8%
BSX(HC)	15.116	**1.208**	2.849	0.139	**2.107**	**0.833**	**0.214**	0.139	92.5%
BT(TS)	5.022	0.318	0.163	0.037	0.545	0.234	0.016	0.037	90.1%
TEF(TS)	2.992	0.330	0.128	0.053	0.704	0.235	0.018	0.053	85.7%
VOD(TS)	3.887	0.272	0.136	0.042	0.469	0.195	0.009	0.042	93.1%
UL(CS)	***0.882***	***0.161***	***0.020***	0.032	***0.288***	***0.122***	***0.004***	0.032	81.0%
AEE(Utilities)	**70.683**	0.890	**18.169**	0.061	1.084	0.364	0.052	0.061	**99.7%**
CNP(Utilities)	23.323	0.681	2.739	0.072	1.103	0.403	0.056	0.072	97.9%
				Panel B. High volume					
ORCL(IT)	5.042	0.531	0.331	0.052	1.023	0.385	0.054	0.052	83.8%
GILD(HC)	7.417	0.832	0.689	0.119	1.571	0.623	0.122	0.119	82.3%
PFE(HC)	7.610	0.475	0.349	0.047	0.812	0.350	0.034	0.047	90.3%
T(TS)	20.518	0.402	1.504	0.048	0.705	0.253	0.024	0.048	98.4%
EXC(Utilities)	21.177	0.689	2.176	0.062	1.210	0.433	0.076	0.062	96.5%
				Panel C. High volume & few jumps					
AAPL(IT)	6.170	0.465	0.309	***0.020***	0.849	0.339	0.040	***0.020***	86.9%
CSCO(IT)	3.564	0.565	0.246	0.052	1.146	0.441	0.073	0.052	***70.5%***
INTC(IT)	3.633	0.535	0.165	0.094	0.979	0.434	0.049	0.094	***70.5%***
MSFT(IT)	4.387	0.516	0.235	0.040	0.993	0.395	0.046	0.040	80.5%
KO(CS)	2.729	0.312	0.110	0.031	0.583	0.228	0.018	0.031	83.9%

Note. The table reports the descriptive statistics for jump components estimated using BV for each stock in the post-crisis period. The bold and italic values indicate the largest and smallest values respectively across different stocks for each column. The panel on the left shows the descriptive statistics for the jump components for the whole data set. The panel on the right shows the descriptive statistics for the jump components after eliminating the highest decile in the data set (leaving only the bottom 90%). There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.8 Descriptive statistics for stocks' jump components estimated using TBV in the pre-crisis period

Stock (Sector)	Full data Max	Full data Mean	Full data Variance	Full data Mode	Up to 90th percentile Max	Up to 90th percentile Mean	Up to 90th percentile Variance	Up to 90th percentile Mode	% change in variance
	Panel A. Few jumps								
IBM(IT)	44.349	1.041	5.531	0.058	2.263	0.608	0.231	0.058	95.8%
XRX(IT)	181.290	4.644	145.800	0.111	9.867	2.197	4.612	0.111	96.8%
YHOO(IT)	102.200	4.219	47.883	0.095	9.556	2.535	4.833	0.095	89.9%
AMGN(HC)	37.914	2.104	9.923	0.083	5.110	1.287	1.127	0.083	88.6%
BSX(HC)	45.495	2.900	18.201	0.094	6.083	1.838	2.047	0.094	88.8%
BT(TS)	50.577	2.196	7.694	0.059	5.199	1.504	1.690	0.059	***78.0%***
TEF(TS)	144.050	1.075	16.871	0.047	2.184	0.678	0.264	0.047	98.4%
VOD(TS)	46.503	1.058	4.072	**0.330**	2.261	0.677	0.301	**0.330**	92.6%
UL(CS)	19.461	1.000	2.195	0.042	2.523	0.603	0.296	0.042	86.5%
AEE(Utilities)	19.537	0.821	***1.052***	0.038	***1.831***	0.575	0.175	0.038	83.4%
CNP(Utilities)	**377.450**	3.616	**255.000**	0.097	5.936	1.502	1.379	0.097	**99.5%**
	Panel B. High volumes								
ORCL(IT)	73.992	2.736	17.329	0.126	6.358	1.727	2.059	0.126	88.1%
GILD(HC)	141.310	**5.986**	99.793	0.134	**14.875**	**3.425**	**11.402**	0.134	88.6%
PFE(HC)	31.179	1.412	5.053	0.088	3.105	0.877	0.465	0.088	90.8%
T(TS)	21.791	1.473	3.717	0.059	3.226	0.980	0.580	0.059	84.4%
EXC(Utilities)	22.958	1.297	3.150	0.077	2.716	0.862	0.406	0.077	87.1%
	Panel C. Few jumps & High volumes								
AAPL(IT)	31.892	3.482	16.049	0.140	7.778	2.411	3.190	0.140	80.1%
CSCO(IT)	60.258	2.090	12.493	0.079	4.745	1.305	1.089	0.079	91.3%
INTC(IT)	40.134	1.812	6.769	0.115	4.136	1.191	0.805	0.115	88.1%
MSFT(IT)	43.891	1.260	4.880	0.071	2.844	0.796	0.448	0.071	90.8%
KO(CS)	***10.904***	***0.799***	1.122	***0.031***	1.843	***0.532***	***0.172***	***0.031***	84.6%

Note. The table reports the descriptive statistics for jump components estimated using TBV for each stock in the pre-crisis period. The bold and italic values indicate the largest and smallest values respectively across different stocks for each column. The panel on the left shows the descriptive statistics for the jump components for the whole data set. The panel on the right shows the descriptive statistics for the jump components after eliminating the highest decile in the data set (leaving only the bottom 90%). There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.9 Descriptive statistics for stocks' jump components estimated using TBV in the post-crisis period

Stock (Sector)	Full data				Up to 90th percentile				
	Max	Mean	Variance	Mode	Max	Mean	Variance	Mode	% change in variance
	Panel A. Few jumps								
IBM(IT)	12.989	0.437	0.392	0.050	0.826	0.314	0.032	0.050	91.9%
XRX(IT)	38.664	1.513	7.067	**0.136**	2.647	0.987	0.331	**0.136**	95.3%
YHOO(IT)	40.466	1.461	5.873	0.091	2.807	0.931	0.336	0.091	94.3%
AMGN(HC)	40.169	0.997	3.128	0.068	1.986	0.644	0.172	0.068	94.5%
BSX(HC)	88.628	**1.799**	12.353	0.105	**3.272**	**1.192**	**0.532**	0.105	95.7%
BT(TS)	14.278	0.443	0.477	0.038	0.820	0.317	0.035	0.038	92.7%
TEF(TS)	11.146	0.514	0.532	0.056	0.986	0.353	0.041	0.056	92.3%
VOD(TS)	12.756	0.361	0.436	0.044	0.613	0.245	0.016	0.044	96.3%
UL(CS)	18.080	***0.236***	0.417	0.027	***0.419***	***0.165***	***0.008***	0.027	98.1%
AEE(Utilities)	82.359	0.752	7.448	0.058	1.150	0.446	0.065	0.058	99.1%
CNP(Utilities)	**765.330**	1.436	**530.360**	0.066	1.276	0.498	0.073	0.066	**100.0%**
	Panel B. High volumes								
ORCL(IT)	51.973	0.729	2.998	0.042	1.375	0.504	0.091	0.042	97.0%
GILD(HC)	28.444	1.222	3.337	0.073	2.257	0.799	0.238	0.073	92.9%
PFE(HC)	23.618	0.685	1.508	0.060	1.227	0.465	0.075	0.060	95.0%
T(TS)	21.420	0.478	0.818	0.038	0.862	0.324	0.035	0.038	95.7%
EXC(Utilities)	41.653	0.839	3.951	0.039	1.459	0.517	0.102	0.039	97.4%
	Panel C. Few jumps & high volumes								
AAPL(IT)	36.449	0.764	2.327	***0.023***	1.514	0.507	0.103	***0.023***	95.6%
CSCO(IT)	44.622	0.744	2.410	0.070	1.408	0.526	0.096	0.070	96.0%
INTC(IT)	48.472	0.798	3.051	0.055	1.408	0.574	0.095	0.055	96.9%
MSFT(IT)	29.383	0.693	1.288	0.041	1.206	0.488	0.070	0.041	94.5%
KO(CS)	***8.540***	0.369	***0.196***	0.030	0.737	0.269	0.026	0.030	***86.7%***

Note. The table reports the descriptive statistics for jump components estimated using TBV for each stock in the post-crisis period. The bold and italic values indicate the largest and smallest values respectively across different stocks for each column. The panel on the left shows the descriptive statistics for the jump components for the whole data set. The panel on the right shows the descriptive statistics for the jump components after eliminating the highest decile in the data set (leaving only the bottom 90%). There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.10 Descriptive statistics for jump components estimated using BV at different times during the pre-crisis period

Stock (Sector)	Full data		2000-2002		2003-2006		2007	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
	Panel A. Few jumps							
IBM (IT)	0.866	4.719	2.234	15.620	0.399	0.300	0.510	0.155
XRX (IT)	3.771	**96.395**	5.715	45.035	**2.354**	**165.080**	**1.477**	**6.899**
YHOO (IT)	2.844	15.655	**6.058**	26.786	1.249	2.266	0.970	0.566
AMGN (HC)	1.396	2.628	2.668	3.993	0.694	0.424	0.836	1.240
BSX (HC)	2.598	13.400	3.522	8.252	2.010	20.985	0.904	0.342
BT (TS)	1.628	3.026	2.572	3.780	0.686	0.412	0.477	0.405
TEF (TS)	0.827	29.489	1.537	69.544	0.325	0.069	***0.256***	0.083
VOD (TS)	0.727	***0.712***	1.449	***0.838***	0.341	0.182	0.406	0.490
UL(CS)	0.889	1.381	1.349	1.831	***0.248***	***0.047***	0.260	0.089
AEE(Utilities)	***0.660***	0.771	***0.942***	1.139	0.325	0.237	0.390	0.096
CNP(Utilities)	2.200	72.283	2.998	**90.239**	1.858	76.371	1.194	3.545
	Panel B. High volume							
ORCL(IT)	1.753	5.527	4.272	11.352	0.887	0.468	0.819	0.453
GILD(HC)	**4.388**	38.216	**7.535**	56.742	1.323	0.741	1.131	1.049
PFE(HC)	1.347	5.219	2.188	8.558	0.859	2.651	0.511	0.259
T(TS)	1.035	1.846	1.722	3.333	0.670	0.779	0.713	0.327
EXC(Utilities)	1.047	2.232	1.216	2.888	0.564	0.576	1.462	2.884
	Panel C. Few jumps & high volume							
AAPL(IT)	2.466	7.013	4.585	11.274	1.404	0.842	0.920	1.212
CSCO(IT)	1.205	2.065	2.471	3.894	0.680	0.380	0.618	0.162
INTC(IT)	1.211	1.774	2.672	2.919	0.702	0.243	0.496	***0.075***
MSFT(IT)	0.940	3.421	1.879	8.396	0.428	0.143	0.567	0.194
KO(CS)	0.776	1.027	1.149	1.222	0.505	0.837	0.483	0.320

Note. The table reports the descriptive statistics for jump components estimated using BV for each stock at different times during the pre-crisis period. The bold and italic values indicate the largest and smallest values respectively across different stocks for each column. The panel on the left shows the descriptive statistics for the jump components for the whole data set. The panel on the right shows the descriptive statistics for the jump components after eliminating the highest decile in the data set (leaving only the bottom 90%). There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.11 Descriptive statistics for jump components estimated using BV at different times during the post-crisis period

Stock (Section)	Full data Mean	Full data Variance	01/07/2009-30/06/2010 Mean	01/07/2009-30/06/2010 Variance	01/07/2010-30/12/2016 Mean	01/07/2010-30/12/2016 Variance
			Panel A. Few jumps			
IBM(IT)	0.365	0.307	0.616	1.671	0.323	0.070
XRX(IT)	1.122	4.397	1.864	**12.745**	0.982	2.739
YHOO(IT)	1.081	3.568	1.405	2.575	1.032	3.712
AMGN(HC)	0.869	1.439	0.912	0.718	0.861	1.577
BSX(HC)	**1.212**	2.847	**1.925**	5.950	**1.078**	2.167
BT(TS)	0.318	0.163	0.625	0.690	0.261	0.047
TEF(TS)	0.330	0.128	0.360	0.083	0.326	0.134
VOD(TS)	0.272	0.136	0.405	0.216	0.244	0.116
UL(CS)	***0.161***	***0.020***	***0.214***	***0.023***	***0.152***	***0.019***
AEE(Utilities)	0.890	**18.169**	0.601	0.665	0.940	**21.202**
CNP(Utilities)	0.681	2.739	0.699	0.288	0.677	3.251
			Panel B. High volumes			
ORCL(IT)	0.531	0.330	0.816	0.769	0.487	0.251
GILD(HC)	0.832	0.689	0.791	0.422	0.840	0.742
PFE(HC)	0.475	0.349	0.662	0.178	0.440	0.374
T(TS)	0.402	1.504	0.502	0.196	0.386	1.702
EXC(Utilities)	0.690	2.168	0.684	1.235	0.691	2.368
			Panel C. Few jumps & high volumes			
AAPL(IT)	0.465	0.309	0.475	0.110	0.464	0.335
CSCO(IT)	0.565	0.246	0.630	0.124	0.558	0.260
INTC(IT)	0.535	0.164	0.714	0.162	0.510	0.160
MSFT(IT)	0.516	0.235	0.629	0.416	0.498	0.207
KO(CS)	0.312	0.110	0.384	0.074	0.299	0.115

Note. The table reports the descriptive statistics for jump components estimated using BV for each stock at different times during the post-crisis period. The bold and italic values indicate the largest and smallest values respectively across different stocks for each column. The panel on the left shows the descriptive statistics for the jump components for the whole data set. The panel on the right shows the descriptive statistics for the jump components after eliminating the highest decile in the data set (leaving only the bottom 90%). There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.12 Descriptive statistics for jump components estimated using TBV at different times during the pre-crisis period

Stock (Section)	Full data		2000-2002		2003-2006		2007	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
				Panel A. Few jumps				
IBM(IT)	1.041	5.531	2.178	14.659	0.452	0.180	0.713	0.695
XRX(IT)	4.644	145.800	9.282	262.060	1.916	**56.256**	1.878	**52.269**
YHOO(IT)	4.219	47.883	9.060	80.820	1.860	15.634	1.842	8.599
AMGN(HC)	2.104	9.923	4.271	18.788	0.876	0.565	1.209	3.875
BSX(HC)	2.900	18.201	4.719	25.019	1.869	12.778	1.408	2.497
BT(TS)	2.196	7.694	3.964	11.251	0.974	1.190	0.512	0.395
TEF(TS)	1.075	16.871	1.958	39.554	0.486	0.187	0.377	***0.077***
VOD(TS)	1.058	4.072	2.186	8.560	0.509	1.022	0.472	0.274
UL(CS)	1.000	2.195	1.809	3.537	***0.341***	***0.105***	***0.314***	0.084
AEE(Utilities)	0.821	***1.052***	1.380	1.794	0.410	0.151	0.491	0.144
CNP(Utilities)	3.616	**255.000**	6.269	**599.980**	**2.189**	47.767	1.412	6.141
				Panel B. High volumes				
ORCL(IT)	2.736	17.329	6.020	37.254	1.327	1.456	1.194	2.032
GILD(HC)	**5.983**	99.729	**12.105**	180.270	1.930	2.524	1.469	1.665
PFE(HC)	1.412	5.053	2.317	6.280	0.922	4.198	0.550	0.232
T(TS)	1.473	3.717	2.510	6.613	0.881	1.129	0.895	0.692
EXC(Utilities)	1.297	3.150	1.927	3.884	0.620	0.602	**1.881**	7.085
				Panel C. Few jumps & High volumes				
AAPL(IT)	3.482	16.049	6.156	23.965	2.138	5.441	1.717	7.512
CSCO(IT)	2.090	12.493	4.544	29.520	0.975	0.635	0.924	0.725
INTC(IT)	1.812	6.769	3.910	15.462	0.987	0.651	0.882	1.074
MSFT(IT)	1.260	4.880	2.534	10.831	0.553	0.289	0.679	0.404
KO(CS)	***0.799***	1.122	***1.364***	***1.565***	0.475	0.663	0.490	0.293

Note. The table reports the descriptive statistics for jump components estimated using TBV for each stock at different times during the pre-crisis period. The bold and italic values indicate the largest and smallest values respectively across different stocks for each column. The panel on the left shows the descriptive statistics for the jump components for the whole data set. The panel on the right shows the descriptive statistics for the jump components after eliminating the highest decile in the data set (leaving only the bottom 90%). There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.13 Descriptive statistics for jump components estimated using TBV at different times during the post-crisis period

Stock (Section)	Full data Mean	Full data Variance	01/07/2009-30/06/2010 Mean	01/07/2009-30/06/2010 Variance	01/07/2010-30/12/2016 Mean	01/07/2010-30/12/2016 Variance
			Panel A. Few jumps			
IBM(IT)	0.437	0.392	0.579	1.375	0.419	0.266
XRX(IT)	1.514	7.064	2.458	16.924	1.349	5.171
YHOO(IT)	1.460	5.864	1.676	3.350	1.428	6.227
AMGN(HC)	0.998	3.126	1.121	2.797	0.979	3.178
BSX(HC)	**1.802**	12.325	2.545	14.323	**1.666**	**11.853**
BT(TS)	0.443	0.476	0.873	2.055	0.369	0.172
TEF(TS)	0.514	0.532	0.646	1.123	0.492	0.429
VOD(TS)	0.362	0.437	0.523	0.583	0.338	0.411
UL(CS)	***0.236***	0.416	0.496	2.779	***0.195***	***0.035***
AEE(Utilities)	0.753	7.442	0.953	5.362	0.720	7.778
CNP(Utilities)	1.436	**529.880**	**5.848**	**3795.000**	0.723	2.024
			Panel B. High volumes			
ORCL(IT)	0.729	2.990	1.275	19.260	0.646	0.468
GILD(HC)	1.222	3.337	1.043	1.380	1.250	3.639
PFE(HC)	0.686	1.506	1.114	4.054	0.622	1.094
T(TS)	0.478	0.818	0.644	1.170	0.452	0.761
EXC(Utilities)	0.840	3.940	1.094	10.840	0.796	2.758
			Panel C. Few jumps & High volumes			
AAPL(IT)	0.764	2.323	1.106	10.303	0.715	1.189
CSCO(IT)	0.743	2.408	1.198	15.319	0.681	0.617
INTC(IT)	0.798	3.043	1.296	17.424	0.730	1.059
MSFT(IT)	0.693	1.288	0.753	0.631	0.684	1.384
KO(CS)	0.369	***0.196***	***0.484***	***0.149***	0.351	0.201

Note. The table reports the descriptive statistics for jump components estimated using TBV for each stock at different times during the post-crisis period. The bold and italic values indicate the largest and smallest values respectively across different stocks for each column. The panel on the left shows the descriptive statistics for the jump components for the whole data set. The panel on the right shows the descriptive statistics for the jump components after eliminating the highest decile in the data set (leaving only the bottom 90%). There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.14 Size of daily jump components estimated using CTBV

	Pre-crisis period			Post-crisis period		
	J>10	10>J>5	5>J>0	J>10	10>J>5	5>J>0
IBM(IT)	1	1	124	0	1	988
XRX(IT)	**5**	**3**	**161**	**12**	**13**	**941**
YHOO(IT)	**2**	**6**	**144**	**10**	**32**	**1000**
AMGN(HC)	2	3	161	3	13	1030
BSX(HC)	**4**	**14**	**168**	**13**	**24**	**977**
BT(TS)	1	2	143	0	1	854
TEF(TS)	1	1	131	0	3	775
VOD(TS)	0	1	113	1	1	760
UL(CS)	1	1	114	0	0	732
AEE(Utilities)	2	1	157	2	10	974
CNP(Utilities)	1	1	153	4	8	941
ORCL(IT)	1	1	137	0	2	907
GILD(HC)	1	3	155	10	22	982
PFE(HC)	1	1	140	2	4	929
T(TS)	1	1	146	1	3	964
EXC(Utilities)	1	1	162	4	6	949
AAPL(IT)	1	1	127	1	5	902
CSCO(IT)	1	1	127	2	3	927
INTC(IT)	1	1	131	1	3	964
MSFT(IT)	0	2	136	1	4	936
KO(CS)	0	0	158	0	1	991

Note. The table reports the number of extreme jump components (greater than 10 and greater than 5) and relatively small jump components (smaller than 5) estimated using CTBV for each stock in both the pre- and post-crisis periods. The bold values indicate the three stocks with the most extreme values in both periods.

Table 4.15 Estimated coefficients for regression of jump components on news surprises in the pre-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	AAPL (IT)	CSCO (IT)	IBM (IT)	INTC (IT)	MSFT (IT)	ORCL (IT)	XRX (IT)	YHOO (IT)	AEE (Utilities)	CNP (Utilities)	EXC (Utilities)
CPTICHNG (+)	**0.246***	0.110	**0.173*****	**0.249*****	**0.125***	**0.240****	**0.588*****	**0.513*****	**0.144***	**0.606*****	**0.212****
USMMMNCH (-)	0.031	0.082	-0.005	0.048	-0.046	0.016	**0.197***	0.104	0.012	0.158	-0.027
NFP TCH (-)	0.029	0.004	**0.072***	0.016	**0.107****	**0.133***	0.171	0.009	**0.109****	0.092	**0.166****
INJCSP (+)	0.077	0.061	0.047	0.088	0.043	0.054	0.096	0.127	0.070	0.174	0.081
INJCSP (-)	0.026	-0.013	0.053	0.033	**0.098***	0.095	0.042	**0.276****	0.037	0.181	0.059
GDPCPCEC (+)	0.094	-0.005	**0.214***	-0.066	0.077	0.064	0.232	-0.077	-0.025	-0.063	0.101
CPUPXCHG (+)	0.021	0.029	0.069	0.024	-0.037	-0.062	-0.081	-0.072	0.046	0.136	0.056
CPUPXCHG (-)	0.084	0.043	**0.170*****	0.027	0.042	**0.210***	0.174	0.109	**0.172****	0.116	0.012
CPI CHNG (+)	-0.022	0.093	-0.023	0.012	0.108	**0.262***	0.252	**0.308***	-0.004	-0.034	0.061
DGNOCHNG (+)	0.130	-0.009	0.004	0.011	-0.007	-0.022	0.043	-0.026	-0.023	0.174	-0.036
DGNOXTCH (+)	0.115	0.024	-0.006	0.030	0.004	0.020	0.284	0.220	0.105	-0.049	0.053
ECI SA (+)	0.043	-0.028	-0.028	0.002	-0.021	-0.039	-0.252	0.025	**0.147***	0.232	**0.260****
ECI SA (-)	0.281	-0.040	-0.069	**0.312****	0.043	-0.031	0.229	0.179	-0.076	0.141	0.058
TMNOCHNG (-)	0.051	0.046	0.026	**0.129***	0.090	0.154	0.132	0.157	0.077	**0.274***	**0.325*****
GDP CQOQ (-)	0.050	-0.002	-0.010	0.038	0.068	**0.208****	0.138	0.007	**0.218*****	**0.377*****	0.121
GDP DCHG (-)	0.110	0.012	-0.001	0.013	0.081	-0.032	-0.005	0.040	0.001	0.169	0.092
GDP PIQQ (+)	0.051	0.066	0.186	-0.195	-0.100	-0.099	**0.924****	-0.076	-0.114	-0.367	-0.224
NHSPSTOT (+)	0.075	0.016	0.064	0.050	0.069	0.110	**0.227***	**0.228****	**0.101***	0.138	**0.126***
NHSPSTOT (-)	0.153	-0.004	0.056	0.037	0.051	0.051	**0.421*****	0.053	**0.149****	**0.242***	0.081

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the pre-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Cont. …Table 4.15. Estimated Coefficients for regression of jump components on news surprises in the pre-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	AAPL (IT)	CSCO (IT)	IBM (IT)	INTC (IT)	MSFT (IT)	ORCL (IT)	XRX (IT)	YHOO (IT)	AEE (Utilities)	CNP (Utilities)	EXC (Utilities)
IMP1CHNG (-)	0.023	0.024	0.044	0.077	0.067	0.094	0.176	0.052	**0.112***	0.036	**0.175****
INJCJC (+)	0.094	**0.110*****	0.033	0.049	**0.080*****	**0.149*****	**0.138***	0.094	**0.099*****	**0.191*****	**0.084***
INJCJC (-)	0.086	**0.079***	0.033	0.049	0.050	**0.104***	**0.171***	0.086	**0.082****	**0.238*****	**0.078***
NAPMPMI (+)	0.155	0.054	**0.132*****	0.050	0.016	**0.187***	0.192	0.086	0.064	0.194	**0.306*****
NAPMPMI (-)	0.100	0.094	0.051	0.055	**0.158****	**0.261****	**0.480*****	0.091	0.009	**0.372****	0.110
MAPMINDX (+)	-0.011	-0.003	-0.005	-0.004	-0.005	**0.785*****	-0.029	-0.016	0.206	-0.019	0.069
NAPMNMAN (-)	0.164	**0.288*****	0.071	**0.173***	0.107	0.182	**0.356***	0.210	0.115	**0.409****	**0.218***
NAPMPRIC (-)	0.071	0.096	0.009	0.071	0.018	-0.097	0.225	0.119	0.036	**0.328***	0.080
LEI CHNG (+)	0.224	0.029	0.029	**0.226*****	0.087	0.024	0.284	0.090	**0.209*****	**0.421****	0.148
CHPMINDX (-)	0.139	0.086	**0.104***	**0.186****	0.098	0.091	0.197	0.181	0.119	**0.275***	0.109
PRODNFR (+)	0.125	-0.042	0.033	0.145	0.060	**0.242***	0.218	-0.157	0.091	0.328	0.084
PCE CMOM (-)	-0.088	-0.031	-0.025	0.068	-0.043	-0.086	-0.077	0.089	0.094	0.020	-0.041
GDPCTOT (+)	0.042	0.004	-0.024	0.139	0.079	0.152	0.388	-0.017	-0.083	0.078	0.057
GDPCTOT (-)	**0.407***	0.027	**0.178***	0.122	0.026	-0.001	0.108	0.124	0.080	0.392	0.242
PITLCHNG (-)	**0.443***	0.089	0.134	**0.246***	**0.195***	**0.437****	0.290	0.149	0.073	**0.445***	0.131
PPI CHNG (-)	0.086	**0.136***	0.067	**0.143***	0.079	**0.223****	0.218	0.192	**0.139***	0.178	0.068
CONSSENT (+)	0.103	0.076	**0.102****	0.062	**0.128****	**0.303*****	**0.286****	**0.217****	**0.159*****	**0.497*****	**0.170****
CONSSENT (-)	0.131	0.034	0.060	0.055	0.056	0.104	**0.348*****	**0.190****	**0.090***	**0.231****	**0.117***
USURTOT (+)	0.073	-0.047	**0.158***	-0.020	**0.261****	0.154	-0.149	0.019	0.017	0.198	0.195
COSTNFR (-)	-0.019	**0.285****	0.029	-0.052	-0.035	-0.080	0.120	0.297	-0.061	-0.138	0.003

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the pre-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Cont. … Table 4.15 Estimated Coefficients for regression of jump components on news surprises in the pre-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	KO (CS)	UL (CS)	BT (TS)	T (TS)	TEF (TS)	VOD (TS)	AMGN (HC)	BSX (HC)	GILD (HC)	PFE (HC)
CPTICHNG (+)	**0.110***	**0.305*****	**0.730*****	0.083	**0.404*****	**0.286*****	0.172	**0.390****	0.068	**0.210****
USMMMNCH (-)	0.034	0.017	0.166	0.048	0.035	0.023	0.069	0.138	0.048	0.051
NFP TCH (-)	0.026	0.053	0.187	0.050	**0.100***	0.074	**0.164****	**0.284*****	0.022	0.101
INJCSP (+)	-0.004	**0.184****	**0.509*****	0.001	**0.141***	0.065	0.126	0.171	**0.087***	**0.163****
INJCSP (-)	0.007	0.033	0.100	**0.171****	-0.016	0.086	0.107	0.143	**0.102***	0.101
GDPCPCEC (+)	-0.005	**0.487****	0.290	0.162	0.204	-0.035	-0.033	-0.230	-0.015	**0.541*****
CPUPXCHG (+)	**0.112***	0.022	0.174	0.045	0.034	0.071	-0.131	-0.011	0.070	-0.006
CPUPXCHG (-)	**0.120***	**0.214***	**0.350***	-0.018	**0.202****	0.055	0.097	0.125	0.084	**0.321*****
CPI CHNG (+)	-0.079	0.189	0.229	0.112	**0.227***	0.005	**0.445*****	0.242	0.002	0.132
DGNOCHNG (+)	0.016	0.017	0.024	0.008	0.025	0.008	**0.175***	**0.532*****	-0.015	-0.044
DGNOXTCH (+)	0.029	0.070	0.314	0.042	0.131	0.076	0.024	-0.302	**0.195****	0.047
ECI SA (+)	0.017	-0.075	0.248	-0.012	0.004	0.129	0.056	**0.468****	-0.026	0.077
ECI SA (-)	-0.019	-0.026	**0.592***	0.019	0.141	0.015	0.085	0.176	-0.019	**0.401*****
TMNOCHNG (-)	0.008	**0.293*****	**0.418****	**0.165***	**0.230*****	0.025	0.069	**0.477*****	0.094	0.106
GDP CQOQ (-)	**0.115***	0.074	**0.316***	0.034	0.097	0.033	0.051	0.031	-0.004	-0.027
GDP DCHG (-)	-0.013	0.084	0.234	0.019	-0.009	-0.003	0.032	**0.411*****	**0.184*****	-0.019
GDP PIQQ (+)	-0.103	0.362	-0.228	0.209	0.089	-0.057	-0.055	-0.171	0.019	-0.187
NHSPSTOT (+)	0.058	**0.142***	**0.255***	0.057	**0.144****	0.036	**0.121***	**0.218***	0.040	0.079
NHSPSTOT (-)	0.036	**0.175***	**0.375****	0.070	0.083	0.032	**0.170***	**0.332****	0.082	0.075

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the pre-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Cont. … Table 4.15 Estimated Coefficients for regression of jump components on news surprises in the pre-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	KO (CS)	UL (CS)	BT (TS)	T (TS)	TEF (TS)	VOD (TS)	AMGN (HC)	BSX (HC)	GILD (HC)	PFE (HC)
IMP1CHNG (-)	0.031	0.086	**0.264***	0.042	0.076	**0.091***	0.117	**0.212***	0.032	0.019
INJCJC (+)	**0.063****	**0.173*****	**0.305*****	**0.065***	**0.084****	**0.064****	**0.117****	**0.190*****	**0.091*****	**0.136*****
INJCJC (-)	**0.074****	**0.082***	**0.362*****	**0.085***	**0.132*****	0.043	0.061	**0.170****	0.002	**0.127*****
NAPMPMI (+)	0.022	**0.256*****	**0.516*****	**0.120***	**0.170****	0.011	0.074	0.116	0.088	**0.150***
NAPMPMI (-)	**0.238*****	**0.250****	**0.649*****	0.063	**0.240*****	0.073	**0.238****	**0.463*****	0.091	0.013
MAPMINDX (+)	-0.009	-0.012	0.271	0.293	0.102	-0.007	-0.004	-0.025	-0.006	-0.015
NAPMNMAN (-)	0.090	**0.269****	**0.586*****	**0.255****	**0.173***	0.043	0.178	**0.291***	0.071	0.165
NAPMPRIC (-)	-0.029	-0.029	0.022	**0.162***	-0.053	0.035	0.086	0.147	0.054	0.141
LEI CHNG (+)	0.070	0.131	0.319	0.077	0.120	0.017	**0.230***	**0.364****	0.035	**0.269*****
CHPMINDX (-)	0.042	**0.180***	**0.563*****	**0.188***	0.099	0.080	0.156	**0.312***	0.047	0.130
PRODNFR (+)	**0.180****	0.121	**0.483***	0.092	0.119	-0.036	**0.259***	0.304	0.007	0.165
PCE CMOM (-)	0.077	**0.233***	0.254	-0.024	-0.063	**0.170***	0.042	0.178	0.007	**0.198***
GDPCTOT (+)	0.069	-0.063	0.099	-0.048	-0.004	0.002	0.033	0.132	0.078	**0.328****
GDPCTOT (-)	0.008	0.180	0.037	0.077	0.025	0.098	0.114	**0.841*****	-0.015	0.120
PITLCHNG (-)	-0.005	0.218	0.488	0.080	0.184	0.113	0.173	**0.470***	**0.387*****	**0.267***
PPI CHNG (-)	0.080	0.134	**0.402****	0.122	0.077	-0.005	**0.203***	**0.232***	0.018	0.064
CONSSENT (+)	**0.102****	**0.224*****	**0.555*****	**0.175*****	**0.112***	0.026	**0.295*****	0.159	0.071	**0.175*****
CONSSENT (-)	**0.105*****	**0.152****	**0.462*****	**0.126****	**0.197*****	**0.085***	0.043	**0.308*****	**0.072***	**0.188*****
USURTOT (+)	0.032	0.120	0.203	-0.049	0.020	-0.049	0.103	0.072	-0.058	0.037
COSTNFR (-)	-0.109	-0.013	-0.188	-0.105	-0.028	0.124	-0.075	-0.114	-0.016	-0.053

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the pre-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Table 4.16 Estimated Coefficients for regression of jump components on news surprises in the post-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	AAPL (IT)	CSCO (IT)	IBM (IT)	INTC (IT)	MSFT (IT)	ORCL (IT)	XRX (IT)	YHOO (IT)	AEE (Utilities)	CNP (Utilities)	EXC (Utilities)
NHSPATOT (+)	0.045	0.032	0.012	**0.192*** **	**0.102***	**0.103***	**0.193** **	0.081	**0.181*** **	0.103	0.080
NHSPATOT (-)	0.133	0.007	0.051	**0.170***	0.141	-0.062	0.060	**0.356** **	0.154	**0.377*** **	0.094
MTIBCHNG (+)	**0.133** **	**0.131***	**0.108** **	0.080	**0.194*** **	**0.125***	0.059	0.131	0.113	**0.138***	0.051
MTIBCHNG (-)	0.060	**0.145***	**0.166*** **	**0.273*** **	0.063	**0.210*** **	0.139	**0.267** **	0.073	0.078	0.096
CGNOXAI (-)	-0.052	**-0.147***	0.008	0.042	**-0.123***	-0.002	0.005	0.026	-0.042	-0.014	**-0.166***
CPTICHNG (+)	0.097	0.090	-0.023	0.040	0.035	0.071	0.270	0.062	**0.323** **	0.022	0.046
USMMMNCH (-)	0.082	-0.018	0.045	0.130	0.062	**0.205***	0.017	**0.415** **	0.127	-0.048	0.015
NFP TCH (+)	0.010	0.073	0.033	**0.207** **	0.077	**0.183** **	0.096	0.090	0.024	**0.250** **	0.088
NFP TCH (-)	0.071	0.115	0.050	**0.137***	0.081	0.054	0.175	-0.019	0.068	0.068	0.070
CFNAI (+)	0.095	0.071	0.110	0.055	**0.171***	-0.011	**0.339***	0.131	0.030	0.017	0.054
CONCCONF (-)	0.067	**0.121***	0.026	**0.117***	**0.127***	0.074	0.047	**0.219***	**0.257*** **	0.061	0.085
INJCSP (+)	0.012	0.024	0.036	0.013	0.055	0.054	0.076	0.097	**0.097** **	0.039	**0.084***
CPUPXCHG (-)	**0.112***	0.080	0.059	0.074	0.025	0.018	**0.170***	0.141	**0.159** **	**0.145***	0.041
CPI CHNG (+)	0.096	0.030	**0.145** **	0.066	0.091	**0.234*** **	0.061	0.067	0.020	**0.239*** **	0.070
DFEDGBA (+)	**0.130***	0.075	**0.168*** **	**0.141***	**0.142***	0.103	0.094	0.121	0.115	0.066	0.047
DFEDGBA (-)	**0.140*** **	0.069	**0.148*** **	**0.128** **	0.069	0.051	**0.142***	**0.195** **	**0.171*** **	0.061	**0.101***
DGNOCHNG (+)	0.054	0.071	0.013	**0.140***	**0.179*** **	-0.002	-0.015	0.048	-0.004	-0.004	-0.062
DGNOXTCH (+)	0.097	0.149	0.031	0.037	0.131	0.057	**0.337***	**0.583*** **	**0.444*** **	0.149	**0.247***
DGNOXTCH (-)	**0.142***	**0.233*** **	0.025	0.028	**0.187** **	0.009	0.088	0.066	0.115	0.140	**0.299*** **
EMPRGBCI (+)	**0.142***	0.046	0.000	**0.218** **	**0.164***	0.060	0.055	**0.343** **	0.067	**0.224** **	0.089
EMPRGBCI (-)	0.030	0.047	-0.003	0.033	0.044	**0.164*** **	**0.177***	0.120	**0.121***	0.052	0.104
TMNOCHNG (+)	0.038	0.043	0.024	0.091	0.021	0.079	0.111	**0.245***	0.113	**0.133***	**0.143***

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the post-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Cont. …Table 4.16 Estimated Coefficients for regression of jump components on news surprises in the post-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	AAPL (IT)	CSCO (IT)	IBM (IT)	INTC (IT)	MSFT (IT)	ORCL (IT)	XRX (IT)	YHOO (IT)	AEE (Utilities)	CNP (Utilities)	EXC (Utilities)
HPIMMOM (+)	**0.119***	0.096	**0.102***	0.061	0.059	**0.183****	0.112	0.170	0.070	**0.327*****	**0.187****
GDP PIQQ (+)	-0.007	0.027	0.029	0.042	0.013	0.005	0.088	**0.190***	0.065	0.002	-0.005
NHSPSTOT (-)	-0.015	-0.014	0.033	-0.015	-0.030	**0.216****	0.043	-0.079	-0.006	-0.151	0.107
IMP1CHNG (+)	**0.113***	**0.129***	**0.138****	**0.156***	0.077	**0.137***	**0.286****	0.095	**0.200****	0.072	0.082
IP CHNG (+)	0.027	0.017	**0.188***	0.077	0.143	-0.002	-0.087	0.165	-0.146	0.025	0.012
INJCJC (+)	**0.067****	**0.091****	0.031	**0.063***	**0.055***	0.037	0.053	**0.135****	-0.013	**0.143*****	0.042
INJCJC (-)	**0.095*****	**0.066***	**0.048***	0.053	0.028	0.052	**0.142****	**0.129****	**0.069***	**0.074***	**0.073***
NAPMPMI (+)	**0.096***	0.041	0.033	0.077	0.056	**0.122***	**0.187***	**0.213***	0.093	0.040	0.072
NAPMPMI (-)	**0.168****	-0.006	0.028	0.090	0.091	**0.158***	0.066	0.024	0.122	-0.021	0.016
MAPMINDX (+)	0.025	**0.319*****	0.053	0.029	0.121	0.039	0.032	0.108	0.071	0.112	0.044
MAPMINDX (-)	0.063	0.107	0.065	0.068	0.015	0.030	0.146	0.058	0.034	**0.136***	**0.172***
NAPMNMI (+)	0.046	0.045	0.071	0.031	0.060	0.101	0.066	0.104	0.087	0.096	0.124
NAPMPRIC (-)	0.041	**0.161***	0.055	0.003	0.071	-0.041	0.083	0.086	-0.013	**0.151***	0.124
LEI CHNG (+)	0.058	0.079	**0.081****	0.064	**0.077***	-0.032	0.007	**0.157***	**0.112***	0.073	0.032
CHPMINDX (-)	0.059	0.041	0.023	**0.162****	0.073	0.085	**0.243****	0.103	0.090	0.054	0.058
USHBMIDX (+)	0.033	0.043	0.056	**0.106***	0.064	0.060	0.133	0.128	**0.163****	**0.155****	**0.146***
USHBMIDX (-)	0.087	0.018	0.059	**0.173****	0.069	**0.117***	0.129	**0.192***	0.098	0.084	0.084
NHSLTOT (+)	**0.146***	0.050	0.047	0.016	0.059	**0.176***	0.002	0.085	**0.186***	**0.188***	**0.204***
SBOITOTL (+)	0.036	0.106	**0.114***	**0.166***	0.019	0.089	**0.394*****	**0.245***	**0.167***	**0.221****	**0.214****
PRODNFR (-)	0.012	0.066	-0.010	0.098	-0.026	0.051	**0.424****	0.040	-0.069	-0.115	-0.178
PCE CMOM (+)	0.097	-0.024	-0.019	0.031	-0.011	**0.131***	-0.015	0.100	0.028	-0.014	-0.037
PCE CMOM (-)	0.008	**0.121***	0.021	0.016	0.003	0.055	0.075	**0.217***	0.044	0.077	0.080

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the post-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Cont. … Table 4.16 Estimated Coefficients for regression of jump components on news surprises in the post-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	AAPL (IT)	CSCO (IT)	IBM (IT)	INTC (IT)	MSFT (IT)	ORCL (IT)	XRX (IT)	YHOO (IT)	AEE (Utilities)	CNP (Utilities)	EXC (Utilities)
GDPCTOT (+)	**0.109***	**0.129***	0.067	0.065	0.081	0.036	0.073	**0.215***	0.032	0.039	0.108
GDPCTOT (-)	0.011	0.045	0.038	0.062	0.036	0.043	0.079	0.033	**0.115***	0.025	0.052
PITLCHNG (-)	0.038	-0.037	0.022	-0.030	0.029	0.008	-0.068	-0.031	0.000	-0.039	-0.058
PCE CRCH (+)	0.014	0.066	0.030	0.113	0.042	0.073	0.149	0.118	0.071	0.055	0.112
PCE CRCH (-)	0.070	0.035	0.030	0.073	0.076	0.021	0.127	0.044	**0.118***	**0.137***	**0.278*****
OUTFGAF (+)	0.025	0.080	0.070	0.067	0.045	**0.173****	**0.238****	0.014	**0.200****	0.108	0.022
OUTFGAF ((-)	**0.093***	0.032	0.031	0.071	0.086	0.077	**0.288*****	0.096	0.056	0.070	**0.120***
PXFECHNG (+)	0.023	-0.044	0.079	**0.166***	0.026	-0.007	0.146	0.036	0.099	0.139	0.092
PXFECHNG (-)	-0.005	**0.151****	**0.138*****	0.057	**0.130***	0.102	0.023	-0.035	-0.024	0.026	0.121
PPI CHNG (-)	0.026	0.115	-0.065	0.037	-0.073	0.023	0.098	0.272	0.187	0.020	0.017
RSTAMOM (+)	-0.012	**0.356***	0.044	0.044	-0.128	0.002	0.266	-0.013	-0.013	-0.152	-0.016
RSTAMOM (-)	0.084	0.080	0.027	0.084	-0.055	**0.288***	**0.526****	0.260	0.212	0.117	0.080
RSTAXAG (+)	**0.180***	-0.072	0.012	0.157	0.149	-0.014	-0.046	0.126	0.143	0.187	0.064
RSTAXMOM (-)	-0.059	0.007	0.067	0.030	0.084	-0.064	-0.199	-0.105	-0.062	0.009	0.045
RCHSINDX (-)	0.029	0.062	0.030	**0.163****	0.045	-0.002	**0.217****	0.039	**0.159****	0.075	0.097
SPCS20SM (+)	0.081	0.087	0.068	**0.195*****	0.095	**0.117***	**0.261****	**0.225****	**0.221*****	**0.132***	0.072
SPCS20SM (-)	**0.102***	0.038	**0.097****	**0.132****	**0.119****	**0.156*****	0.126	**0.283*****	**0.211*****	**0.148****	0.109
CONSSENT (+)	**0.095***	0.071	**0.085***	**0.179*****	0.082	**0.159*****	**0.157***	**0.262*****	**0.210*****	0.083	**0.126***
CONSSENT (-)	**0.082****	**0.162*****	0.029	**0.196*****	**0.124*****	**0.077***	0.080	**0.210*****	**0.124*****	**0.092***	**0.082***
USURTOT (+)	0.090	0.116	**0.122***	-0.017	0.080	-0.016	0.104	0.086	**0.181***	0.002	0.005
COSTNFR (+)	-0.029	0.021	0.068	0.005	0.123	0.043	0.012	0.065	**0.271****	**0.261****	**0.404*****

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the post-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Cont. … Table 4.16 Estimated Coefficients for regression of jump components on news surprises in the post-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	KO (CS)	UL (CS)	BT (TS)	T (TS)	TEF (TS)	VOD (TS)	AMGN (HC)	BSX (HC)	GILD (HC)	PFE (HC)
NHSPATOT (+)	0.046	0.012	0.012	0.026	0.042	0.005	**0.188****	0.156	0.064	**0.132****
NHSPATOT (-)	**0.192*****	0.012	0.030	0.104	0.026	0.048	0.170	0.089	0.114	0.065
MTIBCHNG (+)	**0.070***	0.025	0.055	0.074	**0.074***	0.031	0.139	0.140	0.038	**0.151****
MTIBCHNG (-)	0.032	0.041	**0.090***	**0.154****	0.038	0.011	**0.198***	**0.584*****	0.101	**0.183****
CGNOXAI (-)	-0.024	0.002	0.017	0.040	-0.008	0.054	**0.177***	0.030	0.026	**0.146***
CPTICHNG (+)	0.087	0.064	0.088	0.088	0.037	0.014	**0.426****	0.209	0.095	0.099
USMMMNCH (-)	0.107	-0.008	0.079	**0.148***	0.029	0.054	-0.003	0.212	0.115	0.043
NFP TCH (+)	0.068	0.030	0.067	0.110	0.054	0.002	0.142	**0.259***	0.026	0.066
NFP TCH (-)	0.038	0.015	0.053	**0.202*****	0.010	0.017	**0.327*****	**0.301****	0.050	0.084
CFNAI (+)	0.041	0.013	-0.010	0.064	0.045	**0.323*****	0.105	0.238	0.135	-0.014
CONCCONF (-)	0.027	0.008	0.039	**0.085***	0.042	0.032	0.133	0.137	0.088	0.049
INJCSP (+)	**0.044***	0.002	0.037	0.007	0.008	-0.012	-0.021	0.086	0.024	0.025
CPUPXCHG (-)	0.066	0.006	0.042	0.026	0.018	**0.071***	0.119	**0.270****	0.067	0.069
CPI CHNG (+)	0.050	-0.002	0.021	-0.007	0.022	-0.006	-0.034	**0.221***	0.007	**0.117***
DFEDGBA (+)	0.073	0.036	**0.111****	0.068	0.087	0.017	0.147	**0.362****	**0.124***	**0.208*****
DFEDGBA (-)	0.030	0.014	**0.051***	**0.143*****	0.035	**0.071*****	**0.203*****	0.128	0.024	**0.160*****
DGNOCHNG (+)	0.028	0.012	0.015	0.050	0.012	0.041	0.092	0.036	0.010	0.005
DGNOXTCH (+)	0.064	0.017	0.074	0.126	0.064	**0.158*****	0.185	0.212	0.025	**0.176***
DGNOXTCH (-)	**0.091***	0.007	0.036	-0.036	0.008	-0.006	-0.059	-0.039	0.000	0.006
EMPRGBCI (+)	**0.132****	0.011	-0.006	0.061	0.074	0.075	0.000	**0.266***	0.057	0.088
EMPRGBCI (-)	0.025	0.017	0.038	0.033	0.041	**0.079*****	**0.167****	**0.174***	0.031	0.047
TMNOCHNG (+)	0.046	0.026	0.073	**0.110***	**0.082***	**0.102****	0.027	**0.275***	**0.151****	0.015

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the post-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Cont. … Table 4.16 Estimated Coefficients for regression of jump components on news surprises in the post-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	KO (CS)	UL (CS)	BT (TS)	T (TS)	TEF (TS)	VOD (TS)	AMGN (HC)	BSX (HC)	GILD (HC)	PFE (HC)
HPIMMOM (+)	**0.088***	0.012	0.033	**0.096***	0.000	0.001	0.099	0.199	0.054	**0.121***
GDP PIQQ (+)	**0.082***	-0.003	0.011	-0.018	0.022	-0.002	0.006	0.011	0.002	0.054
NHSPSTOT (-)	-0.018	0.029	0.056	0.008	0.017	-0.030	-0.043	0.105	-0.053	-0.028
IMP1CHNG (+)	**0.092***	0.035	0.049	0.079	0.045	**0.166*****	0.147	**0.633*****	**0.128***	0.101
IP CHNG (+)	-0.029	-0.022	0.045	-0.009	0.031	0.006	**-0.284***	-0.011	-0.039	-0.005
INJCJC (+)	0.026	0.010	**0.051****	**0.076*****	**0.044***	**0.035***	**0.100***	0.082	0.036	0.052
INJCJC (-)	0.035	0.020	**0.062****	**0.137*****	**0.052***	**0.078*****	**0.126****	**0.210*****	0.032	**0.067***
NAPMPMI (+)	**0.118*****	**0.061****	**0.090****	**0.165*****	0.022	0.031	**0.355*****	0.102	0.060	**0.207*****
NAPMPMI (-)	0.085	0.028	**0.132****	0.028	0.014	**0.110****	**0.244***	0.144	0.018	0.005
MAPMINDX (+)	0.051	0.017	0.048	**0.178*****	0.007	**0.220*****	0.157	0.154	0.021	0.064
MAPMINDX (-)	0.063	0.008	0.072	0.075	0.038	0.038	0.086	**0.309****	0.019	0.082
NAPMNMI (+)	**0.082***	0.035	**0.070***	0.075	0.031	0.013	**0.341*****	0.068	0.062	**0.175*****
NAPMPRIC (-)	0.035	-0.011	0.000	0.050	0.010	0.018	-0.041	0.104	0.008	0.092
LEI CHNG (+)	0.009	0.012	0.009	0.064	**0.078****	0.018	**0.130***	**0.178***	0.036	**0.138*****
CHPMINDX (-)	**0.086***	**0.042***	0.051	0.064	0.030	0.018	0.113	0.078	0.093	0.101
USHBMIDX (+)	**0.126*****	0.013	**0.137*****	0.031	0.022	**0.093*****	**0.150***	0.164	0.045	**0.190*****
USHBMIDX (-)	0.062	0.015	**0.089***	**0.101***	0.050	**0.059***	**0.279*****	**0.397*****	**0.115***	**0.145****
NHSLTOT (+)	0.043	0.007	0.087	0.116	-0.004	**0.115****	0.058	0.164	**0.138***	0.085
SBOITOTL (+)	**0.110***	0.034	**0.141****	**0.217*****	0.064	0.020	**0.230***	0.180	0.049	0.069
PRODNFR (-)	-0.051	-0.019	0.046	0.006	-0.007	0.018	0.024	0.046	0.017	-0.033
PCE CMOM (+)	0.062	**0.050***	0.040	-0.079	0.007	0.002	-0.067	-0.068	-0.017	0.028
PCE CMOM (-)	**0.085***	0.005	0.026	-0.062	0.006	-0.031	-0.147	-0.036	0.041	0.105

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the post-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Cont. … Table 4. Estimated Coefficients for regression of jump components on news surprises in the post-crisis period

$log(J_t + 1) = \beta_{J,k,p}|S_{kt}|\ 1\ (S_{kt} \geq 0) + \beta_{J,k,n}|S_{kt}|\ 1\ (S_{kt} < 0) + \varepsilon_t$

Negative (-) or Positive (+) news surprise	KO (CS)	UL (CS)	BT (TS)	T (TS)	TEF (TS)	VOD (TS)	AMGN (HC)	BSX (HC)	GILD (HC)	PFE (HC)
GDPCTOT (+)	0.035	0.008	0.038	0.030	0.029	0.005	**0.142***	0.178	0.045	0.078
GDPCTOT (-)	0.032	0.004	0.033	**0.131*****	0.003	0.032	**0.244*****	0.117	0.024	0.078
PITLCHNG (-)	0.002	0.024	0.044	0.081	-0.017	**0.110****	**0.195***	-0.020	0.037	-0.006
PCE CRCH (+)	0.042	0.005	0.030	0.097	0.032	0.017	**0.280****	**0.291****	0.028	**0.147***
PCE CRCH (-)	**0.118*****	0.008	0.004	**0.086***	0.042	-0.004	**0.267*****	**0.281****	0.008	**0.129***
OUTFGAF (+)	0.068	0.004	0.061	0.047	0.006	0.051	0.008	**0.365*****	0.003	0.106
OUTFGAF (-)	**0.112*****	0.009	**0.145*****	0.006	0.011	0.016	**0.151***	0.129	0.060	0.016
PXFECHNG (+)	0.045	0.003	0.025	0.026	0.017	-0.020	**0.230***	0.112	0.076	**0.132***
PXFECHNG (-)	**0.074***	-0.004	0.020	**0.145****	0.008	-0.020	**0.285*****	**0.397*****	0.028	**0.223*****
PPI CHNG (-)	-0.029	0.055	**0.133***	0.060	0.024	0.055	-0.206	0.032	0.049	-0.185
RSTAMOM (+)	0.004	-0.030	0.138	0.040	-0.067	**-0.148***	0.022	0.259	0.025	-0.088
RSTAMOM (-)	**0.207***	-0.013	**0.486*****	0.046	-0.025	0.046	0.192	0.036	-0.106	0.107
RSTAXAG (+)	0.065	0.031	0.001	0.065	0.040	**0.149****	0.043	0.309	**0.196***	0.129
RSTAXMOM (-)	-0.083	0.019	**-0.283*****	0.020	0.034	0.042	-0.001	0.095	0.118	0.031
RCHSINDX (-)	**0.125*****	0.019	0.021	0.042	0.059	**0.052***	0.036	**0.211***	0.076	0.049
SPCS20SM (+)	**0.087***	0.037	0.045	0.061	0.058	**0.112*****	**0.230****	**0.302****	0.084	0.086
SPCS20SM (-)	0.044	0.018	0.042	**0.118****	0.048	**0.071****	0.110	0.114	**0.113****	**0.151****
CONSSENT (+)	**0.092****	0.011	**0.076***	**0.161*****	0.055	**0.061****	**0.158***	**0.186***	0.053	**0.090***
CONSSENT (-)	**0.074*****	0.014	**0.079*****	**0.099*****	**0.093*****	0.024	**0.106***	**0.119***	**0.099*****	**0.078****
USURTOT (+)	0.015	0.007	0.067	-0.011	0.034	0.023	0.157	-0.031	0.053	0.138
COSTNFR (+)	0.094	0.046	0.095	0.083	0.001	0.003	0.092	0.133	-0.013	0.022

Note. The table reports the OLS coefficient estimates of the regression for jump components on negative (-) or positive (+) news surprise on individual news outlets in the post-crisis period. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. The significant coefficients are highlighted in bold. The jump components J_t is estimated using corrected threshold bi-power variation (CTBV).

Table 4.17 Proportions of news-related daily jump components, estimated using CTBV

Stock (Sector)	Pre-crisis Proportion	Pre-crisis Mean	Pre-crisis Variance	Post-crisis Proportion	Post-crisis Mean	Post-crisis Variance
	Panel A: Few jumps					
IBM(IT)	0.134	0.467	0.168	***0.011***	0.543	0.116
XRX(IT)	0.046	2.421	12.682	0.093	1.102	0.476
YHOO(IT)	0.122	1.646	2.612	0.193	**1.299**	**9.969**
AMGN(HC)	0.108	1.824	3.559	0.227	0.885	1.251
BSX(HC)	0.199	**2.941**	**23.154**	0.108	1.103	0.553
BT(TS)	0.094	1.316	2.357	0.173	0.270	0.034
TEF(TS)	0.132	0.637	0.295	0.095	0.302	0.060
VOD(TS)	0.053	0.841	0.301	**0.352**	0.251	0.044
UL(CS)	0.226	0.911	2.371	0.056	***0.144***	***0.024***
AEE(Utilities)	0.028	***0.456***	***0.068***	0.196	0.829	1.823
CNP(Utilities)	***0.024***	1.146	0.819	0.286	0.955	8.162
	Panel B: High volumes					
ORCL(IT)	0.138	1.414	2.516	0.242	0.473	0.099
GILD(HC)	0.119	0.644	1.170	0.067	0.361	0.084
PFE(HC)	**0.257**	0.843	1.446	0.227	0.525	0.172
T(TS)	0.080	1.187	1.993	0.311	0.596	4.629
EXC(Utilities)	0.072	0.746	0.387	0.106	0.555	0.343
	Panel C: High volumes & few jumps					
AAPL(IT)	0.085	1.969	3.593	0.292	0.523	0.755
CSCO(IT)	0.027	2.601	7.090	0.109	0.484	0.138
INTC(IT)	0.123	1.046	1.046	0.076	0.445	0.082
MSFT(IT)	0.198	0.807	1.041	0.172	0.457	0.146
KO(CS)	0.068	0.531	0.269	0.110	0.390	0.230

Note. The table reports the descriptive statistics for jump components that co-occur with macroeconomic news announcements, estimated using CTBV for each stock. The stocks with the highest and lowest values in each column are marked in bold and italics respectively. There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.18 Proportions of co-jump-related jump components estimated using CTBV for the six stocks with the highest volume and fewest jumps in the data set (subset one)

	Pre-crisis			Post-crisis		
	Proportion	**Mean**	**Variance**	**Proportion**	**Mean**	**Variance**
	Panel A: Six stocks with the highest volume					
AAPL(IT)	0.099	5.549	27.362	0.103	0.809	0.610
CSCO(IT)	0.093	2.648	10.194	0.094	0.870	0.329
GILD(HC)	0.096	1.284	1.456	0.101	1.005	0.743
INTC(IT)	0.102	2.406	5.059	0.098	0.795	0.293
MSFT(IT)	0.106	1.943	4.194	0.100	0.751	0.567
ORCL(IT)	0.100	4.053	19.234	0.105	0.789	0.349
PFE(HC)	0.109	2.248	10.693	0.101	0.793	0.361
T(TS)	0.101	1.888	4.717	0.109	0.497	0.240
	Panel B: Six stocks with the fewest jumps					
AAPL(IT)	0.106	5.511	25.222	0.126	0.753	0.430
BT(TS)	0.124	4.303	13.234	0.142	0.650	0.697
CSCO(IT)	0.102	2.672	10.041	0.118	0.888	0.347
IBM(IT)	0.109	1.993	15.361	0.108	0.476	0.105
INTC(IT)	0.105	2.369	4.882	0.129	0.832	0.324
MSFT(IT)	0.118	2.181	6.043	0.113	0.696	0.522
TEF(TS)	0.123	1.329	1.259	0.138	0.463	0.126
UL(CS)	0.133	2.486	5.718	0.130	0.303	0.037
VOD(TS)	0.115	1.937	5.080	0.114	0.415	0.147
XRX(IT)	0.107	5.663	250.39	0.128	2.185	9.967
YHOO(IT)	0.093	7.084	85.634	0.117	1.546	2.693

Note. The table reports the descriptive statistics for jump components that co-occur with co-jumps, estimated using CTBV for the six stocks with the highest volume and the six stocks with the fewest jumps in the data set (subset one). More than six stocks are shown in each panel of the table because there is variation in the top six stocks between the pre- and post-crisis periods.

Table 4.19 Co-jump-related jump components estimated using CTBV for the two stocks with the highest volume and fewest jumps in each sector (subset two)

	Pre-crisis			Post-crisis		
	Proportion	**Mean**	**Variance**	**Proportion**	**Mean**	**Variance**
	Panel A: Highest volume in each sector					
AAPL(IT)	0.101	5.506	27.828	0.117	1.117	11.074
EXC(Utilities)	0.122	1.605	1.700	0.127	1.056	12.204
GILD(HC)	0.093	1.565	2.976	0.117	1.010	0.638
KO(CS)	0.109	1.422	1.480	0.118	0.478	0.165
PFE(HC)	0.130	2.389	10.340	0.116	0.966	4.440
T(TS)	0.104	1.778	1.820	0.120	0.597	1.251
	Panel B: Lowest jump frequency in each sector					
AAPL(IT)	0.102	5.802	26.491	0.119	0.764	0.455
AEE(Utilities)	0.130	1.358	1.126	0.127	1.258	46.372
AMGN(HC)	0.118	4.294	17.120	0.125	0.941	0.987
BSX(HC)	0.119	4.772	21.557	0.123	2.393	17.443
CNP(Utilities)	0.115	2.762	40.231	0.118	0.881	0.529
INTC(IT)	0.104	2.431	6.154	0.123	0.853	0.497
KO(CS)	0.103	1.316	1.108	0.117	0.450	0.147
UL(CS)	0.128	2.604	6.257	0.122	0.287	0.043
VOD(TS)	0.112	1.748	4.825	0.112	0.369	0.054

Note. The table reports the descriptive statistics for jump components that co-occur with co-jumps, estimated using CTBV for the two stocks from each sector with the highest volume and the fewest jumps (subset two). More than one stock per sector is shown in each panel of the table because there is variation in the highest stocks between the pre- and post-crisis periods.

Table 4.20 Coefficients for news surprise (positive or negative) that co-occurs with co-jumps on jump components estimated using CTBV for the stocks with the highest volumes and fewest jumps for each sector (data subset two)

$log(J_t + 1) = \beta_{J,p}|S_t|\ 1\ (S_t \geq 0) + \beta_{J.n}\ |S_t|\ 1\ (S_t < 0) + \varepsilon_t$

	Pre-crisis		Post-crisis	
News Surprise	**Positive**	**Negative**	**Positive**	**Negative**
	Panel A: Highest-volume stocks in each sector			
AAPL(IT)	0.115***	0.113***	0.075***	0.026**
GILD(HC)	0.087***	0.045***	0.045***	0.040***
PFE(HC)	0.079***	0.071***	0.068***	0.040***
T(TS)	0.080***	0.071***	0.060***	0.048***
KO(CS)	0.100***	0.034**	0.050***	0.040***
EXC(Utilities)	0.111***	0.084***	0.082***	0.074***
	Panel B: Stocks with the fewest jumps in each sector			
AAPL(IT)	0.167***	0.132***	0.066***	0.046***
INTC(IT)	0.110***	0.067***	0.093***	0.059***
AMGN(HC)	0.113***	0.095***	0.092***	0.073***
BSX(HC)	0.174***	0.217***	0.136***	0.134***
VOD(TS)	0.078***	0.061***	0.028***	0.029***
KO(CS)	0.082***	0.054***	0.059***	0.048***
UL(CS)	0.125***	0.140***	0.018***	0.013**
AEE(Utilities)	0.068***	0.081***	0.090***	0.062***
CNP(Utilities)	0.293***	0.224***	0.090***	0.061***

Note. The table reports the OLS coefficient estimates of the regression for jump components on news announcements that co-occur with co-jumps for the stocks with the highest volumes and fewest jumps in each sector (subset two), for the pre- and post-crisis periods. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively.

Table 4.21 Coefficients for news surprise (positive or negative) that co-occurs with co-jumps on jump components estimated using CTBV for the six stocks with the highest volume and fewest jumps in the data set overall (subset one).

$log(J_t^c + 1) = \beta_{J,p}|S_t^+| + \beta_{J.n}\,|S_t^-| + \varepsilon_t$

	Pre-crisis		**Post-crisis**	
News Surprise	**Positive**	**Negative**	**Positive**	**Negative**
	Panel A: Top six highest-volume stocks			
AAPL(IT)	0.161***	0.104***	0.065***	0.042***
CSCO(IT)	0.050*	0.064***	0.074***	0.066***
INTC(IT)	0.071***	0.053***	0.071***	0.074***
MSFT(IT)	0.031	0.053***	0.074***	0.061***
ORCL(IT)	0.117***	0.106***	0.082***	0.048***
GILD(HC)	0.069***	0.051***	0.050***	0.027**
PFE(HC)	0.092***	0.067***	0.066***	0.042***
T(TS)	0.084***	0.088***	0.063***	0.043***
	Panel B: Top six lowest-jump-frequency stocks			
AAPL(IT)	0.133***	0.126***	0.053***	0.049***
CSCO(IT)	0.050*	0.084***	0.088***	0.080***
IBM(IT)	0.032*	0.064***	0.048***	0.052***
INTC(IT)	0.080***	0.071***	0.076***	0.069***
MSFT(IT)	0.050**	0.059***	0.069***	0.057***
XRX(IT)	0.146***	0.300***	0.148***	0.108***
YHOO(IT)	0.148***	0.118***	0.176***	0.115***
BT(TS)	0.325***	0.322***	0.038***	0.035***
TEF(TS)	0.127***	0.116***	0.035***	0.040***
VOD(TS)	0.048**	0.056***	0.022***	0.028***
UL(CS)	0.149***	0.132***	0.014***	0.014**

Note. The table reports the OLS coefficient estimates of the regression for jump components on news announcements that co-occur with co-jumps for the six stocks with the highest volumes and fewest jumps in the whole data set (subset one), for the pre- and post-crisis periods. The superscript asterisks *, ** and *** denote statistical significance at the 10%, 5% and 1% levels respectively. More than six stocks are shown in each panel of the table because there is variation in the top six stocks between the pre- and post-crisis periods.

Table 4.22 MSE results for standard HAR, HAR-TJ and HAR-CTJ models for different stocks (forecast horizon h=1)

	Pre-crisis			Post-crisis		
	HAR	**HAR-TJ**	**HAR-CTJ**	**HAR**	**HAR-TJ**	**HAR-CTJ**
	Panel A. Few jumps					
IBM(IT)	2.543	2.456	2.568	0.316	0.295	0.304
XRX(IT)	4.544	5.044	4.891	1.097	1.117	1.062
YHOO(IT)	7.536	8.100	7.735	2.140	1.951	1.928
AMGN(HC)	4.529	4.108	4.298	8.987	8.637	8.503
BSX(HC)	5.077	5.041	5.129	3.581	3.702	3.533
BT(TS)	0.598	1.961	1.552	0.227	0.245	0.238
TEF(TS)	0.576	0.815	0.541	0.136	0.136	0.136
VOD(TS)	1.111	1.141	1.104	0.060	0.061	0.067
UL(CS)	0.290	0.411	0.343	0.035	0.035	0.034
AEE(Utilities)	0.636	0.697	0.643	0.501	0.396	0.413
CNP(Utilities)	4.833	4.603	4.702	1.078	1.158	1.206
	Panel B. High volume					
ORCL(IT)	7.462	7.388	7.514	0.431	0.346	0.389
GILD(HC)	1.661	1.655	1.656	0.246	0.236	0.242
PFE(HC)	0.586	0.645	0.597	2.246	2.123	2.150
T(TS)	2.460	2.268	2.489	1.513	1.531	1.564
EXC(Utilities)	7.419	6.956	6.766	1.223	1.128	1.078
	Panel C. High volume & few jumps					
AAPL(IT)	41.446	41.846	42.823	0.992	0.957	1.024
CSCO(IT)	2.861	2.955	2.918	0.676	0.627	0.716
INTC(IT)	6.562	5.953	6.480	0.395	0.357	0.345
MSFT(IT)	1.152	1.220	1.190	0.624	0.620	0.617
KO(CS)	0.531	0.502	0.522	0.067	0.071	0.064

Note. This tables reports the mean squared error (MSE) forecasting results from standard HAR-family models with the forecast horizon set to h=1. These models are used as benchmarks for the modified HAR-family models that incorporate the impact of news in the following tables. There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.23 MSE results for HAR, HAR-TJ and HAR-CTJ models modified to take into account the impact of news on different stocks (forecast horizon h=1)

	Pre-crisis			Post-crisis		
	HAR	**HAR-TJ**	**HAR-CTJ**	**HAR**	**HAR-TJ**	**HAR-CTJ**
			Panel A. Few jumps			
IBM(IT)	**2.500**	2.457	**2.565**	0.323	0.306	0.304
XRX(IT)	4.738	**4.987**	4.987	1.122	1.120	1.086
YHOO(IT)	7.547	**7.942**	7.750	**2.120**	**1.947**	**1.926**
AMGN(HC)	4.624	4.352	**4.285**	**8.951**	**8.625**	8.564
BSX(HC)	**5.072**	**5.037**	5.144	3.624	**3.699**	3.535
BT(TS)	**0.591**	**1.902**	**1.525**	**0.218**	0.248	0.238
TEF(TS)	**0.564**	**0.809**	**0.533**	0.136	0.136	**0.135**
VOD(TS)	1.112	1.146	1.140	0.063	0.061	0.067
UL(CS)	0.293	**0.408**	**0.341**	0.037	0.035	0.034
AEE(Utilities)	0.684	**0.695**	0.645	**0.487**	0.426	0.439
CNP(Utilities)	6.381	6.168	6.059	1.740	**1.128**	1.216
			Panel B. High volumes			
ORCL(IT)	7.493	7.465	**7.502**	**0.410**	0.346	**0.385**
GILD(HC)	1.834	1.694	1.668	0.251	0.238	**0.236**
PFE(HC)	0.586	0.647	0.601	**2.382**	2.521	2.621
T(TS)	2.495	**2.248**	**2.465**	1.620	**1.525**	**1.546**
EXC(Utilities)	8.750	7.919	**6.670**	**1.133**	**1.072**	1.085
			Panel C. Few jumps & high volumes			
AAPL(IT)	41.538	**41.642**	42.946	**0.958**	**0.956**	**0.996**
CSCO(IT)	2.864	2.982	**2.894**	**0.620**	**0.603**	**0.692**
INTC(IT)	**6.537**	**5.702**	6.690	0.398	0.360	0.345
MSFT(IT)	**1.150**	**1.218**	1.206	**0.592**	**0.540**	**0.515**
KO(CS)	**0.521**	**0.496**	**0.516**	**0.066**	0.071	0.064
Improved results	**7**	**12**	**10**	**11**	**9**	**8**

Note. This table reports the mean squared error (MSE) forecasting results from HAR-family models that take into account the impact of macroeconomic news announcements (forecast horizon h=1). Bold MSE values are lower than the MSE of the corresponding benchmark models in Table 4.22. The number of improved results for each model is given at the bottom of each column. There are more stocks in Panel A than in the other panels, because the stocks with the fewest jumps vary considerably between the pre-crisis and post-crisis periods.

Table 4.24 MSE results for the HAR, HAR-TJ and HAR-CTJ models (forecast horizon h=1) modified to account for the impact of news that co-occurs with co-jumps for the stocks with the highest volume and fewest jumps in each sector (subset two).

	Pre-crisis			Post-crisis		
	HAR	**HAR-TJ**	**HAR-CTJ**	**HAR**	**HAR-TJ**	**HAR-CTJ**
	Panel A. Highest volume stock in each sector					
AAPL(IT)	**41.389**	**41.814**	42.824	0.992	0.958	1.024
GILD(HC)	1.665	**1.641**	**1.656**	0.246	0.236	0.242
PFE(HC)	0.599	0.647	0.597	2.246	2.123	2.150
T(TS)	2.695	2.874	**2.061**	1.513	1.531	1.564
KO(CS)	**0.525**	**0.499**	**0.520**	0.067	0.071	0.064
EXC(Utilities)	8.244	7.467	6.774	1.224	1.128	1.078
	Panel B. Lowest jump stock in each sector					
AAPL(IT)	42.203	41.925	42.823	0.993	0.959	1.024
INTC(IT)	6.583	5.958	6.493	0.396	0.357	0.345
AMGN(HC)	4.547	**4.107**	4.301	8.987	8.639	8.503
BSX(HC)	5.089	5.048	5.130	3.587	3.705	3.533
VOD(TS)	1.144	1.142	1.104	0.060	0.061	0.067
KO(CS)	0.534	**0.501**	0.522	0.067	0.071	0.064
UL(CS)	0.297	0.413	**0.342**	0.035	0.035	0.034
AEE(Utilities)	0.682	0.698	0.643	0.501	0.396	**0.412**
CNP(Utilities)	5.170	4.652	4.758	1.080	1.170	1.214

Note. This table reports the mean squared error (MSE) forecasting results from HAR-family models that take into account the impact of macroeconomic news announcements (forecast horizon h=1). Bold MSE values are lower than the MSE of the corresponding benchmark models in Table 4.22.

Table 4.25 MSE results for the HAR, HAR-TJ and HAR-CTJ models (forecast horizon h=1) modified to account for the impact of news that co-occurs with co-jumps for the six stocks with the highest volume and fewest jumps in the data set overall (subset one)

	Pre-crisis			Post-crisis		
	HAR	**HAR-TJ**	**HAR-CTJ**	**HAR**	**HAR-TJ**	**HAR-CTJ**
	Panel A. Top six high volume stock					
AAPL(IT)	41.715	41.852	42.824	**0.991**	**0.956**	1.024
CSCO(IT)	2.976	2.961	2.920	0.676	0.627	0.716
INTC(IT)	6.885	5.955	6.492	**0.394**	**0.356**	0.345
MSFT(IT)	1.227	1.301	1.190	0.624	0.621	0.617
ORCL(IT)	**7.404**	**7.328**	**7.407**	0.432	0.346	0.389
GILD(HC)	**1.627**	**1.600**	1.658	**0.246**	0.236	0.242
PFE(HC)	0.603	0.646	**0.596**	2.246	2.123	2.150
T(TS)	2.748	2.713	2.702	**1.513**	1.531	1.564
	Panel B. Top six low jump frequency stock					
AAPL(IT)	41.892	41.886	42.824	0.993	0.958	1.024
CSCO(IT)	2.951	2.956	2.918	0.676	0.627	0.716
IBM(IT)	2.558	2.456	2.568	0.316	0.296	0.304
INTC(IT)	6.653	5.955	6.483	0.395	0.357	0.345
MSFT(IT)	1.211	1.243	1.190	0.625	0.621	0.617
XRX(IT)	4.665	5.075	4.898	1.099	1.118	1.062
YHOO(IT)	7.580	8.145	7.743	2.140	1.951	1.928
BT(TS)	0.618	1.977	1.571	0.228	0.245	0.238
TEF(TS)	0.556	0.813	0.541	0.136	0.136	0.136
VOD(TS)	**1.160**	**1.142**	1.104	**0.060**	0.061	0.067
UL(CS)	0.290	0.411	**0.342**	0.035	0.035	0.034

Note. This table reports the mean squared error (MSE) forecasting results from HAR-family models that take into account the impact of macroeconomic news announcements (forecast horizon h=1). Bold MSE values are lower than the MSE of the corresponding benchmark models in Table 4.22. More than six stocks are shown in each panel of the table because there is variation in the top six stocks between the pre- and post-crisis periods.

Concluding Remarks

In this thesis, we have reviewed recent developments in the study of high-frequency financial volatility analysis. We have discussed the stylised facts of high-frequency returns and eleven different volatility measures, alongside different volatility patterns between calendar-time sampling returns and business-time sampling returns. In doing so, we have helped enhance the modelling and forecasting of financial volatility by incorporating trading volume, different intraday periodicity estimators and information from macroeconomic news announcements into volatility forecasting models.

In Chapter 1, we initially focused on the development of the estimation of intraday periodicity, from the parametric Flexible Fourier form method to recent non-parametric estimators such as weighted standard deviation (WSD) and Shortest Half. We also discussed the stylised facts of high-frequency data reported in a limited number of previous studies, such as the presence of jumps. Various methods of detecting intraday jumps, estimating daily jump components and incorporating jumps into volatility models were compared. In the second part of the chapter, we reviewed the development of parametric conditional variance (GARCH) models, from using low-frequency to high-frequency data, along with how they have incorporated different features of financial data (e.g. leverage effects) in recent decades. In addition, the development of non-parametric HAR-family models is outlined in this chapter, as they are important for high-frequency volatility forecasting. In the final part of the chapter, we reviewed studies that use different methods to examine the impact of news on financial volatility, from visibly observing volatility graphs to regressing the volatility components (e.g. jump components) on quantifiable variables representing news (e.g. standardised news surprise). We also discuss the impact of news announcements on co-movements between financial asset returns. Forecasting intraday volatility is challenging, as there are many features of volatility that are more visible at the intraday level. It is therefore important to accurately identify the stylised facts of volatility and efficiently incorporate them in volatility models. Since volatility forecasting will continue to be an important topic for researchers, investors and policymakers, more advanced methods and models for estimating and incorporating the stylised facts of intraday financial volatility need to be developed.

Chapter 2 discussed the stylised facts of high-frequency data. The results show that both the absolute intraday returns and their volatility measures have slow-decaying autocorrelations and that the aggregated daily returns exhibit fat tails and leverage effects. In addition, by standardising aggregated returns using eleven different volatility measures, we find that the standardised returns follow normal distributions. In addition, the various volatility measures have long-memory properties and are significantly correlated with trading volumes. We hence incorporate trading volume in a volatility forecasting model and find that the new model performs better than three other HAR-class models for data from less volatile financial regimes (the post-crisis periods, as opposed to the pre-crisis and crisis period).

Chapter 3 first investigates the impact of intraday periodicity on jump frequency, jump components and volatility forecasting, for both calendar-time and business-time sampling data. The results show that filtering for intraday periodicity results in a reduction in jump frequency for all returns and reduces the jump components for most returns regardless of the sampling scheme employed. However, the reduction in jump components for business-time sampling data is more consistent than for calendar-time data. In addition, we find that filtering for intraday periodicity yields better forecasting results from HAR models for less volatile data such as healthcare stocks and data from the post-crisis period for both sampling schemes. We also compare the intraday periodicity patterns, jump frequencies and jump components for business-time sampling and calendar-time sampling returns. The results show that intraday periodicity is more evident and that the jump frequency and jump components are higher for calendar-time sampling data compared to business-time sampling data.

Chapter 4 extends non-parametric volatility (HAR) models by considering information regarding macroeconomic news announcements. This chapter first runs a regression of the jump components on the standardised news surprises from news announcements, including some news outlets which are not considered in previous literature, and on co-jump-related news announcements. The results show that many news announcements, including those specifically related to co-jumps, have a significant impact on stocks' jump components. The second part of the study focuses on incorporating information from these news announcements in HAR-class models. We treat significant news announcements as an index to separate the jump components between those that are related to news and those that are caused by other factors. The out-of-sample forecasting results show that incorporating news announcements improves the models' forecasting performance. However, considering co-jump-related news does not seem to yield different forecasting results for HAR models. In addition, we did not find a significant effect of news announcements on the co-movements between stocks when considering the impact of co-jumps in HAR models. Since the impact of news announcements on these co-movements can be observed from their impact on covariance in multivariate models, it would therefore be worthwhile for future research to take news announcements into account in multivariate volatility models such as multivariate GARCH models.

In terms of the limitations of the thesis, we required high-frequency stock price data in order to study intraday patterns, which meant that data from a fairly limited range of sources (i.e. US stock data) could be used. The phenomena studied in this thesis may show different effects on stock prices from other countries such as European or Asian markets, or on other kinds of financial data such as foreign exchange markets and commodity markets. Future research in this area would therefore benefit from considering a wider range of data from various markets and financial domains. Another limitation is that the analysis of the effect of news announcements on stock volatility in Chapter 4 was restricted to macroeconomic news announcements from sources such as government bodies and credit agencies. Company-level news announcements, such as individual firms' quarterly earnings announcements, also have an impact on stock volatility, but were not explicitly taken into account in the analysis. We recommend that future work on this topic investigates both types of news announcements and compares their relative effects on stock prices in order to gain a greater understanding of the effects of news on stock volatility. Finally, the COVID-19 pandemic in 2020 has led to another period of uncertainty in financial markets, which would be worthy of inclusion in future research on stock market volatility.

In summary, this thesis contributes to the field by conducting a detailed examination of the impact of trading volume, intraday periodicity and macroeconomic news announcements on stock volatility by incorporating these factors into volatility forecasting models. In Chapter 2 we investigate the stylised facts of volatility measures for stocks, revealing the presence of long-memory properties and correlations between them and trading volumes. We find that incorporating the lag of trading volume can improve volatility forecasting. In Chapter 3 we find that business-time sampling data have fewer jumps and jump components and weaker intraday periodicity patterns than calendar-time sampling data. Filtering for intraday periodicity produces larger improvements in forecasting for the less volatile business-time sampling data. Chapter 4 examines the impact of macroeconomic news announcements on stocks' jump components, including news outlets which have not previously been analysed in work to date. We find that macroeconomic news has a significant impact on jump components and that incorporating news surprise in HAR-family models improves their forecasting performance. However, considering only co-jump-related news does not have a significant effect on volatility forecasting models. Overall, the thesis demonstrates the importance of trading volume, intraday periodicity and macroeconomic news in stock volatility by providing innovations to existing models, thus furthering our understanding of stock market volatility forecasting.

References

Aït-Sahalia, Y., Jacod, J. & Li, J. (2012). Testing for jumps in noisy high frequency data. *Journal of Econometrics, 168*(2), pp. 207-222.

Aït-Sahalia, Y. & Xiu, D. (2016). Increased correlation among asset classes: Are volatility or jumps to blame, or both? *Journal of Econometrics*, *194*(2), pp. 205-219.

Akgiray, V. (1989). Conditional Heteroscedasticity in time series models of stock returns: Evidence and forecasts. *The Journal of Business*, 62(1), pp. 55-80.

Alford, A.W. & Boatsman, J.R. (1995). Predicting long-term stock return volatility: Implications for accounting and valuation of equity derivatives. *The Accounting Review*, 70(4), pp. 599-618.

Allegret, J.P., Raymond, H. & Rharrabti, H. (2017). The impact of the European sovereign debt crisis on banks stocks. Some evidence of shift contagion in Europe. *Journal of Banking & Finance*, 74, pp. 24-37.

Andersen, T.G., Benzoni, L. & Lund, J. (2002). An empirical investigation of continuous-time equity return models. *The Journal of Finance,* 57(3), pp. 1239-1284.

Andersen, T.G. & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2-3), pp. 115-58.

Andersen, T.G. & Bollerslev, T. (1998). Deutsche mark–dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *The Journal of Finance*, 53(1), pp. 219-265.

Andersen, T.G., Bollerslev, T. & Diebold, F.X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, 89(4), pp. 701-720.

Andersen, T.G., Bollerslev, T., Diebold, F.X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), pp. 579-625.

Andersen, T.G., Bollerslev, T., Diebold, F.X. & Vega, C. (2007). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, 73(2), pp. 251-277.

Andersen, T.G., Bollerslev, T. & Huang, X. (2011). A reduced form framework for modelling volatility of speculative prices based on realized variation measures. *Journal of Econometrics*, 160(1), pp. 176-189.

Andersen, T.G., Bollerslev, T. & Lange, S. (1999). Forecasting financial market volatility: Sample frequency vis-a-vis forecast horizon. *Journal of Empirical Finance*, 6(5), pp. 457-477.

Andersen, T. G., Dobrev, D., & Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1), pp. 75-93.

Anderson, K., Brooks, C. & Katsaris, A. (2010). Speculative bubbles in the S&P 500: Was the tech bubble confined to the tech sector? *Journal of Empirical Finance*, 17(3), pp. 345-361.

Arnott, R., Cornell, B. & Shepherd, S. (2018). Yes. It’s a Bubble. So What? *Research Affiliates.*

Baillie, R.T., Bollerslev, T. & Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1), pp. 3-30.

Bakshi, G., Cao, C. & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, 52(5), pp. 2003-2049.

Balduzzi, P., Elton, E.J. & Green, T.C. (2001). Economic news and bond prices: Evidence from the U.S. treasury market. *Journal of Financial and Quantitative Analysis,* 36(4), pp. 523-543.

Baillie, R.T., & DeGennaro, R.P. (1990). Stock returns and volatility. *Journal of Financial and Quantitative Analysis*, 25(2), pp. 203-214.

Bandi, F.M. & Russell, J.R. (2008). Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies*, 75(2), pp. 339-369.

Barndorff-Nielsen, O.E. & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1), pp. 1-37.

Barndorff-Nielsen, O.E. & Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1), pp. 1-30.

Barndorff-Nielsen, O.E. & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 64(2), pp. 253-280.

Barndorff-Nielsen, O.E., Kinnebrock, S. & Shephard, N. (2008). Measuring downside risk-realised semivariance. *CREATES Research Paper* (2008-42).

BBC. (2011). Stock markets fall on fears over Europe's debt crisis. *BBC News*, 18 July 2011. Available at: https://www.bbc.co.uk/news/business-14183041 (Accessed: 19 March 2020).

Będowska-Sójka, B. (2015). Daily VAR forecasts with realized volatility and GARCH models. *Argumenta Oeconomica*, 1(34), pp. 157-173.

Bekierman, J. & Manner, H. (2018). Forecasting realized variance measures using time-varying coefficient models. *International Journal of Forecasting*, 34(2), pp. 276-287.

Bera, A.K. & Higgins, M.L. (1992). A test for conditional heteroskedasticity in time series models. *Journal of Time Series Analysis*, 13(6), pp. 501-519.

Bloomberg (2018). Bloomberg Professional. [Online] (Accessed: 30 November 2018).

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), pp. 307-327.

Bollerslev, T., Cai, J. & Song, F.M. (2000). Intraday periodicity, long-memory volatility, and macroeconomic announcement effects in the US Treasury bond market. *Journal of Empirical Finance*, 7(1), pp. 37-55.

Bollerslev, T., Law, T.H. & Tauchen, G. (2008). Risk, jumps, and diversification. *Journal of Econometrics*, 144(1), pp. 234-256.

Bollerslev, T., Li, J. & Xue, Y. (2018). Volume, volatility, and public news announcements. *The Review of Economic Studies*, 85(4), pp. 2005-2041.

Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of econometrics*, 73(1), 151-184.

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

Boudt, K., Croux, C. & Laurent, S. (2011a) Outlyingness weighted covariation. *Journal of Financial Econometrics*, 9(4), pp. 657-684.

Boudt, K., Croux, C. & Laurent, S. (2011b). Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Empirical Finance*, 18(2), pp. 353-367.

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

Carr, P. & Wu, L. (2003). The finite moment log stable process and option pricing. *The Journal of Finance*, 58(2), pp. 753-777.

Chan, K.F., Bowman, R.G. & Neely, C.J. (2017). Systematic cojumps, market component portfolios and scheduled macroeconomic announcements. *Journal of Empirical Finance*, 43, pp. 43-58.

Chatrath, A., Miao, H., Ramchander, S. & Villupuram, S. (2014). Currency jumps, cojumps and the role of macro news. *Journal of International Money and Finance*, 40, pp. 42-62.

Chernov, M., Gallant, A.R., Ghysels, E. & Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2), pp. 225-257.

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

Choudhry, T. (1995). Integrated-GARCH and non-stationary variances: Evidence from European stock markets during the 1920s and 1930s. *Economics Letters*, 48(1), pp. 55-59.

Christensen, K. & Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2), pp. 323-349.

Clinet, S., & Potiron, Y. (2017). *Estimation for high-frequency data under parametric market microstructure noise*. arXiv preprint. arXiv:1712.01479.

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

Corsi, F., Mittnik, S., Pigorsch, C., & Pigorsch, U. (2008) The volatility of realized volatility. *Econometric Reviews*, 27(1-3), pp. 46-78.

Corsi, F., Pirino, D. & Renò, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2), pp. 276-288.

Duffie, D., Pan, J. & Singleton, K. (2000) Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6), pp. 1343-1376.

Dumitru, A. & Urga, G. (2012) Identifying jumps in financial assets: A comparison between nonparametric jump tests. *Journal of Business & Economic Statistics,* 30(2), pp. 242-255.

Dungey, M., & Hvozdyk, L. (2012). Cojumping: Evidence from the US Treasury bond and futures markets. *Journal of Banking & Finance*, 36(5), pp. 1563-1575.

Dungey, M., McKenzie, M. & Smith, L.V. (2009). Empirical evidence on jumps in the term structure of the US Treasury market. *Journal of Empirical Finance*, 16(3), pp. 430-445.

Dong, S. & Feng, Y. (2018). Does index futures trading cause market fluctuations? *China Finance Review International*, 8(2), pp. 173-198.

Duong, D. & Swanson, N.R. (2015). Empirical evidence on the importance of aggregation, asymmetry, and jumps for volatility prediction. *Journal of Econometrics*, 187(2), pp. 606-621.

Ederington, L.H. & Lee, J.H. (1993) How markets process information: News releases and volatility. *The Journal of Finance,* 48(4), pp. 1161-1191.

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

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

Engle, R.F. & Kraft, D. (1983). Multiperiod forecast error variances of inflation estimated from ARCH models. *Applied Time Series Analysis of Economic Data*, pp. 293-302.

Engle, R.F., Lilien, D.M. & Robins, R.P. (1987) Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica*, 55(2), pp. 391-407.

Engle, R.F. & Sokalska, M.E. (2012) Forecasting intraday volatility in the US equity market: Multiplicative component GARCH. *Journal of Financial Econometrics,* 10(1), pp. 54-83.

Eraker, B., Johannes, M. & Polson, N. (2003) The impact of jumps in volatility and returns. *The Journal of Finance,* 58(3), pp. 1269-1300.

Erdemlioglu, D., Laurent, S. & Neely, C.J. (2015). Which continuous-time model is most appropriate for exchange rates? *Journal of Banking & Finance*, 61, pp. 256-268.

Evans, K.P. (2011). Intraday jumps and US macroeconomic news announcements. *Journal of Banking & Finance*, 35(10), pp. 2511-2527.

Figlewski, S. (1997) Forecasting volatility. *Financial Markets, Institutions & Instruments*, 6(1), pp. 1-88.

French, K.R., Schwert, G.W. & Stambaugh, R.F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19(1), p.3.

Gilder, D., Shackleton, M.B. & Taylor, S.J. (2014). Cojumps in stock prices: Empirical evidence. *Journal of Banking & Finance*, 40, pp. 443-459.

Glosten, L.R., Jagannathan, R. & Runkle, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5), pp. 1779-1801.

Gong, X. & Lin, B. (2018). Structural changes and out-of-sample prediction of realized range-based variance in the stock market. *Physica A: Statistical Mechanics and its Applications*, 494, pp. 27-39.

Goodhart, C. A., & O'Hara, M. (1997). High frequency data in financial markets: Issues and applications. *Journal of Empirical Finance*, 4(2-3), pp. 73-114.

Guillaume, D.M., Dacorogna, M.M., Davé, R.R., Müller, U.A., Olsen, R.B., & Pictet, O.V. (1997). From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics*, 1(2), pp. 95-129.

Hansen, P. R., & Lunde, A. (2004). *An unbiased measure of realized variance*. Available at SSRN 524602.

Haugom, E. & Ullrich, C.J. (2012). Forecasting spot price volatility using the short-term forward curve. *Energy Economics*, 34(6), pp. 1826-1833.

Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics*, 39(1), pp. 71-104.

Hizmeri, R., Izzeldin, M., Murphy, A. & Tsionas, E.G. (2019). The contribution of jump signs and activity to forecasting stock price volatility. Available at SSRN 3361972.

Huang, X., & Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4), pp. 456-499.

Huang, X. (2018) Macroeconomic news announcements, systemic risk, financial market volatility, and jumps. *Journal of Futures Markets*, 38(5), pp. 513-534.

Hull, J. & White, A. (1987) The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2), pp. 281-300.

Jain, P.C. (1988) Response of hourly stock prices and trading volume to economic news. *The Journal of Business*, 61(2), pp. 219-231.

Jiang, G.J. & Oomen, R.C. (2008). Testing for jumps when asset prices are observed with noise–a "swap variance" approach. *Journal of Econometrics*, 144(2), pp. 352-370.

Kleinnijenhuis, J., Schultz, F., Oegema, D., et al. (2013) Financial news and market panics in the age of high-frequency sentiment trading algorithms. *Journalism,* 14(2), pp. 271-291.

Koopman, S.J., Jungbacker, B. & Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 12(3), pp. 445-475.

Lahaye, J., Laurent, S. & Neely, C.J. (2011) Jumps, cojumps and macro announcements. *Journal of Applied Econometrics*, 26(6), pp. 893-921.

Lee, S.S. (2011) Jumps and information flow in financial markets. *The Review of Financial Studies*, 25(2), pp. 439-479.

Lee, S.S., & Mykland, P. A. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *The Review of Financial Studies*, 21(6), pp. 2535-2563.

Macrotrends.net (2020). *Apple Revenue and Profit*, Macrotrends.net. Available at: https://www.macrotrends.net/stocks/charts/AAPL/apple/revenue (Accessed: 19 March 2020).

Ma, F., Wei, Y., Huang, D. & Chen, Y. (2014). Which is the better forecasting model? A comparison between HAR-RV and multifractality volatility. *Physica A: Statistical Mechanics and its Applications*, *405*, pp. 171-180.

Manda, K. (2010). Stock market volatility during the 2008 financial crisis. *Glucksman Fellowship Program Student Research Reports: 2009-2010*, *87*.

Martens, M. & Van Dijk, D. (2007). Measuring volatility with the realized range. *Journal of Econometrics*, *138*(1), pp. 181-207.

Martin, R.D. & Zamar, R.H. (1993). Bias robust estimation of scale. *The Annals of Statistics*, 21(2), pp. 991-1017.

Merton, R.C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Econometrics 3*(1-2), pp. 125-144.

Miao, H., Ramchander, S. & Zumwalt, J.K. (2014). S&P 500 index-futures price jumps and macroeconomic news. *Journal of Futures Markets*, 34(10), pp. 980-1001.

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

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, 116(2), pp. 257-270.

Officer, R.R. (1973). The Variability of the Market Factor of the New York Stock Exchange. *The Journal of Business*, 46(3), pp. 434-453.

Peng, H., Chen, R., Mei, D. & Diao, X. (2018). Forecasting the realized volatility of the Chinese stock market: Do the G7 stock markets help?. *Physica A: Statistical Mechanics and its Applications*, 501, pp. 78-85.

Pu, W., Chen, Y. & Ma, F. (2016). Forecasting the realized volatility in the Chinese stock market: further evidence. *Applied Economics*, 48(33), pp. 3116-3130.

Rousseeuw, P.J. & Leroy, A.M. (1988). A robust scale estimator based on the shortest half. *Statistica Neerlandica*, 42(2), pp. 103-116.

Russell, J. & Bandi, F. (2004) Microstructure noise, realized volatility, and optimal sampling. *Review of Economic Studies*, pp. 75.

Schwert, G.W. (1989). Why does stock market volatility change over time? *The Journal of Finance*, 44(5), pp. 1115-1153.

Schwert, G.W. (1990). Stock market volatility. *Financial Analysts Journal*, 46(3), pp. 23-34.

Sentana, E. (1995). Quadratic ARCH models. *The Review of Economic Studies*, 62(4), pp. 639-661.

Shakeel, M., & Srivastava, B. (in press). Stylized facts of high-frequency financial time series data. *Global Business Review*.

Taylor, S.J. (1987). Forecasting the volatility of currency exchange rates. *International Journal of Forecasting*, *3*(1), pp. 159-170.

Taylor, S.J. (1986). *Modelling financial time series*. Singapore: World Scientific.

Taylor, S.J. & Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance*, *4*(4), pp. 317-340.

Theodosiou, M.G. & Zikes, F. (2011). *A comprehensive comparison of nonparametric tests for jumps in asset prices*. Available at SSRN 1895364.

Tiwari, A.K., Raheem, I.D, and Kang, S.H. (2019). Time-varying dynamic conditional correlation between stock and cryptocurrency markets using the copula-ADCC-EGARCH model. *Physica A: Statistical Mechanics and its Applications,* 535, 122295.

Vortelinos, D.I. & Saha, S. (2016). The impact of political risk on return, volatility and discontinuity: Evidence from the international stock and foreign exchange markets. *Finance Research Letters*, *17*, pp. 222-226.

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

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

Wiggins, J.B. (1992). Estimating the volatility of S&P 500 futures prices using the extreme-value method. *The Journal of Futures Markets*, 12(3), pp. 265-273.

Wongswan, J. (2006). Transmission of information across international equity markets. *The Review of Financial Studies,* 19(4), pp. 1157-1189.

Zakoïan, J.M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, *18*(5), pp. 931-955.

Zhao, W. & Li, H. (2010). Realized covariance matrix is good at forecasting volatility. *2010 International Conference on Logistics Systems and Intelligent Management (ICLSIM), Harbin*, pp. 1761-1764.