Drift and Volatility

This thesis is being submitted in partial fulfilment of the requirements for the degree

of Doctor of Philosophy

Kefu Liao

September 2024

Summary

High-frequency stock price dynamics are conventionally modelled by three components: volatility, drift, and jumps. Volatility depicts how greatly an asset's prices swing around the mean price. Drift describes the movement of the mean price. Jumps refer to rare, significant, and sudden price changes that are too large to be explained by volatility alone. This thesis investigates the econometrics of volatility, jumps, and drift in high-frequency stock prices and the implications of these components for the stock markets.

Chapter 1 contains the introduction of this thesis. Chapter 2 reveals non-negligible drift-related finite sample biases in the estimation of good volatility, bad volatility, and signed jumps. This chapter suggests a modified estimation method that significantly reduces these biases. The empirical evidence indicates that the asymmetric impacts of good and bad volatility and the asymmetric effects of signed jumps in volatility forecasting, as found in the literature, are almost exclusively due to the influence of the measurement bias of these variables on future volatility.

Chapter 3 is subdivided into two parts. The first part focuses on the measurement of the occurrence, size, and intensity of drift bursts. The empirical results reported in the second part imply that drift bursts do not impact realized variance but explain the implied variance and variance risk premium.

Chapter 4 first demonstrates that applying a coexceedance criterion to a univariate drift test proposed by recent studies is feasible for detecting stock codrift variations. I show stock codrift variations are significantly associated with market drift bursts. I also find stock codrift variations have a significant and positive impact on market volatility. Models exploiting this effect lead to significantly better in-sample and out-of-sample market volatility forecasts. Chapter 5 is the conclusion of this thesis.

CONTENT

CHAPTER 1. INTRODUCTION 1

CHAPTER 2. GOOD AND BAD VOLATILITY ESTIMATION FOR DRIFT-DIFFUSION PROCESS AND THE IMPACT OF SIGNED JUMPS ON VOLATILITY FORECASTING 9

2.1. INTRODUCTION 10

2.2. THEORETICAL FRAMEWORK 18

2.3. REALIZED SEMIVARIANCE ESTIMATION FOR A CONSTANT DRIFT-DIFFUSION PROCESS 22

2.3.1. The bias of realized semivariances 23

2.3.2. Visualization of the bias 34

2.3.3. The bias of realized semivariances in the presence of jumps 47

2.4. REALIZED SEMIVARIANCE ESTIMATION FOR THE ORNSTEIN UHLENBECK PROCESSES 58

2.4.1. The bias of realized semivariances 59

2.4.2. The bias of realized semivariances in the presence of jumps 67

2.4.3. The bias of a noise-robust version of realized semivariances 79

2.5. SIGNED JUMP ESTIMATION 85

2.6. DATA DESCRIPTION 98

2.7. THE IMPACT OF GOOD AND BAD VOLATILITY, AND SIGNED JUMPS ON FUTURE VOLATILITY 109

2.8. CONCLUSION 124

CHAPTER 3. DRIFT BURSTS, VOLATILITY FORECASTING, AND THE VARIANCE RISK PREMIUM 127

3.1. INTRODUCTION 128

3.2. THE ECONOMETRICS OF DRIFT BURSTS 135

3.3. DRIFT BURST DETECTION RESULTS 142

3.4. MODELLING DRIFT BURST CLUSTERING EFFECT USING HAWKES PROCESSES 149

3.5. PREDICTING VOLATILITY USING DRIFT BURSTS 156

3.6. IMPACT ON THE VARIANCE RISK PREMIUM 170

3.7. CONCLUSION 179

CHAPTER 4. EMPIRICAL EVIDENCE OF STOCK CODRIFT VARIATION AND ITS IMPLICATIONS FOR MARKET VOLATILITY PREDICTION ... 181

4.1. INTRODUCTION ... 182
4.2. CODRIFT VARIATION IDENTIFICATION ... 186
4.2.1. Realized drift variation ... 186
4.2.2. Testing for drift variation ... 189
4.2.3. Testing for codrift variation ... 190
4.3. MONTE CARLO STUDY ... 192
4.3.1. Simulation set-up ... 192
4.3.2. Simulation results ... 195
4.4. DATA ... 199
4.5. THE STOCK CODRIFT VARIATION ... 200
4.5.1. Drift variations in the market and individual stocks ... 200
4.5.2. Codrift variation among individual stocks ... 201
4.5.3. Small drift variations in the market portfolio ... 203
4.5.4. Non-systematic codrift variation ... 204
4.6. THE IMPACT OF CODRIFT VARIATION ON VOLATILITY FORECASTING ... 208
4.6.1. Model set-up ... 208
4.6.2. In-sample results ... 211
4.6.3. Out-of-sample results ... 214
4.7. CONCLUSION ... 219

CHAPTER 5. CONCLUSION ... 220

REFERENCES ... 223

APPENDIX ... 239

A.1. MATHEMATICAL PROOFS FOR CHAPTER 2 ... 239
A. 2. WEIGHTED LEAST SQUARE ILLUSTRATION FOR CHAPTER 2 ... 255
A. 3. MORE EMPIRICAL RESULTS FOR CHAPTER 2 ... 264
A. 4. MORE EMPIRICAL RESULTS FOR CHAPTER 3 ... 268
A. 5. MORE EMPIRICAL RESULTS FOR CHAPTER 4 ... 277

Acknowledgements

I would like to express my deepest gratitude to Dr. Kevin Evans and Dr. Dudley Gilder, my research supervisors, for their invaluable guidance, support, and encouragement throughout this study. Their expertise and insights were instrumental in shaping this research.

I am grateful to Torben Andersen, Kim Christensen, Roberto Renò, Shuping Shi, Stephen Taylor, and Jun Yu for their constructive feedback and suggestions, which greatly improved the quality of my thesis. I would also like to thank the seminar participants at Cardiff Business School and the participants at the fifteenth annual conference of the Society for Financial Econometrics, the Lancaster Financial Econometrics Conference 2023, the Frontiers of Factor Investing Conference 2024, the Tri-University Annual Conference 2024, the Cardiff Fintech Conference 2023, and the Welsh Post-Graduate Research Conference 2024.

Special thanks to Cardiff University and the China Scholarship Council for providing the financial support necessary to conduct this research. Finally, I would like to thank my family for their unwavering support and understanding throughout my studies. Their love and patience have been my greatest source of strength.

Chapter 1. Introduction

The availability of high-frequency data has triggered the rapid development of financial econometrics over the past two decades. In the high-frequency literature, asset prices are typically modelled as an Ito semi-martingale process with three components: volatility; drift; and jumps. These components play distinct roles in governing high-frequency asset price dynamics. Volatility describes how an asset's prices fluctuate around the mean price. Drift depicts the price trend, reflecting the move of the mean price. Price jumps refer to rare, significant, and sudden price changes that are too large to be explained by the volatility.

Volatility is a crucial concept in finance, serving as the key form of market risk. It is essential for determining asset prices, making portfolio decisions, and regulating financial markets. Earlier literature measures the volatility over a trading day by the squared daily return. However, as argued by Andersen and Bollerslev (1998), this estimator may be subject to large measurement error. To improve the estimation accuracy of volatility, they suggest using an alternative estimator, realized variance, which is defined as the sum of squared intraday high-frequency price returns of that day. They find that realized variance contains much smaller measurement errors than the squared daily return in estimating daily volatility.

However, Barndorff-Nielsen and Shephard (2004) argue that realized variance may erroneously include jumps as the volatility, exaggerating the volatility level. They propose using a bipower variation, which extends the idea of realized variance by incorporating products of high-frequency absolute returns of different times, making it more robust to jumps in the price process.

The impact of drift on the estimation consistency of realized variance and bipower variation diminishes with the time interval of the returns going to zero. However, in practice, it is not possible to sample prices continuously, and financial econometricians rely on finite sample price data for computing the volatility estimators. In particular, the 5-minute sampling frequencies of asset price returns are preferred to calculate realized variance and bipower variation to mitigate the impact of microstructure noise (Andersen et al. 2007b; Bollerslev et al. 2009; Corsi 2009; Park and Linton 2012; Patton and Sheppard 2015; Bollerslev et al. 2016). Laurent and Shi (2020) observe that drift is not ignorable in 5-minute returns, especially when drift is large (e.g., the early 2000s the dot-com bubble burst and the subprime mortgage crisis periods). And they find that the non-negligibility of drift in 5-minute returns can significantly deteriorate the estimation accuracy of realized variance and bipower variation estimators. To alleviate such biases, Laurent and Shi (2020) purpose computing these two estimators on the high-frequency returns that are centred by their median and show significant improvement in the volatility estimation accuracy in the presence of nonnegligible drift.

The good and bad volatilities allow one to disentangle the downside from the upside risk, which is important to asset pricing, asset allocation, and portfolio risk management (Harlow 1991; Hyung and De Vries 2005; Ang et al. 2006; Lettau et al. 2014; Klebaner et al. 2017; Farago and Tédongap 2018). Another popular application of the good and bad volatilities is to measure signed jumps. Signed jumps, which indicate whether positive or negative jumps dominate the price variation of the day, are also important factors in asset pricing and volatility forecasting (Patton and Sheppard 2015; Mizrach et al. 2018; Bollerslev et al. 2020). In the above literature related to the decomposition of the upside and downside risk, good and bad volatilities are typically estimated by the positive and negative realized semivariances

proposed by Barndorff-Nielsen et al. (2008), with positive and negative realized semivariances calculated by the sums of positive and negative using 5-minute squared returns, respectively. However, this computation ignores the impact of nonzero drifts for a small sample. A nonzero drift should affect semivariances in two respects. First, since semivariances are computed by cumulative 5-minute squared returns, the estimation accuracy of semivariances is inevitably affected by the existence of drift these returns. Second, the presence of drift can result in uneven numbers of positive and negative returns within a finite sample, causing further biases in the measurement of semivariances.

Chapter 2 investigates the finite sample impact of nonzero drifts on the positive and negative semivariances, with the log prices assumed to follow a constant drift-diffusion process or an Ornstein-Uhlenbeck process. My finite sample theory, together with extensive simulations, uncovers that the drift can have a larger impact on the measurement performance of semivariances than on that of realized variance and bipower variation estimators. I demonstrate that the estimation inaccuracy of semivariances due to a nonzero drift leads to bias in the estimation of signed jumps. I find that the bias in the signed jump measurement is even greater than that in both positive and negative semivariances in the presence of a nonzero drift.

To reduce these drift-related biases, I modify the positive and negative semivariances, together with the signed jump estimator, by applying the same approach proposed by Laurent and Shi (2020) for alleviating the impact of drift on the estimation performance of realized variance and bipower variation. My analytical and simulation results suggest that the drift's impact on the measurement performance of modified semivariances and modified signed jump estimators is omittable relative to the respective original estimators, especially if the drift deviates far from zero.

Importantly, my results show that in the presence of a nonzero drift, the estimation accuracy of the modified semivariances and modified signed jump estimator is as good as those of the modified realized variance and bipower variation found in Laurent and Shi (2020).

Forecasting asset return volatility is key to financial economics, including risk management (Barone-Adesi et al. 1999; Christoffersen and Diebold 2000; Berkowitz and O'Brien 2002; Pritsker 2006; Andersen et al. 2007a), asset pricing (Black and Scholes 1973; Duffie et al. 2000; Cochrane 2009; Johannes et al. 2009), asset allocation (Fleming et al. 2001; Fleming and Kirby 2003; Fleming et al. 2003), and option valuation (Melino and Turnbull 1990; Bates 1996; Bakshi et al. 1997; Chernov and Ghysels 2000; Pan 2002; Eraker 2004).

The volatility persistence and the asymmetric effect of good and bad volatility on volatility forecasting have been well established in the literature (Corsi 2009; Patton and Sheppard 2015; Bollerslev et al. 2016). Chapter 2 reveals that the original and modified versions of realized variance or bipower variation exhibit similar effect of volatility persistence on volatility forecasting, suggesting that the volatility dependence effect is not influenced by the impact of drift on the estimation of volatility. The results for the original positive and negative semivariances corroborate the asymmetric impact of good and bad volatility on future volatility, with a stronger effect from the bad volatility. However, the modified signed semivariances indicate that the impacts of good and bad volatility are not significantly different from each other. This implies that good and bad volatilities may not differ in impacting volatility forecasting after their estimation are adjusted for alleviating the biases due to a nonzero drift. Additionally, I show that the asymmetric impacts of the original signed semivariances for volatility forecasting may be mostly because of the finite sample

bias in these estimators in the presence of a nonzero drift. This indicates that the asymmetric effects of good and bad volatility on volatility prediction found in existing studies may be almost exclusively due to the drift biases in the signed semivariances.

Using the original signed jump estimator, the literature finds the significance of signed jumps in stock volatility forecasting, with positive and negative jumps affecting future volatility asymmetrically (Patton and Sheppard 2015; Audrino and Hu 2016; Wang et al. 2016). Consistent with the literature, Chapter 2 also finds the asymmetric impacts of the signed jumps for volatility forecasting based on the original signed jump estimator. However, the modified signed jump estimator has almost no impact on future volatility. This leads to the conclusion that the signed jumps may be not important for volatility forecasting, which is quite different from the findings in the existing literature. Additionally, I show that the asymmetric effect demonstrated by the original signed jump estimator for volatility forecasting may be mainly due to the finite sample bias in this estimator due to a nonzero drift. This suggests that the significance of signed jumps found in existing studies may be almost exclusively attributed to the drift bias in the signed jump estimator.

The constant drift-diffusion and Ornstein-Uhlenbeck processes in Chapter 2 assume drift to be constant or linearly correlated with the prices, which does not explain the episodes of intraday price movements associated with an explosive trend. To model this price pattern, Christensen et al. (2022), Laurent et al. (2024), and Shi and Phillips (2024) argue that drift could burst in some intraday intervals of stock prices. Andersen et al. (2023) also find pockets of drift bursts for S&P 500 E-mini futures and large-cap stocks over intraday periods. The drift bursts have been utilized in several studies to model intraday price bubbles and crashes in financial markets (Jagannathan et al. 2019; Flora and Renò 2020; Christensen et al. 2022; Jagannathan

et al. 2022; Andersen et al. 2023; Bellia et al. 2023; Laurent et al. 2024; Shi and Phillips 2024).

Since drift bursts are associated with bubbles and crashes, understanding their time series properties has important implications for risk management and asset pricing. Chapter 3 discovers that both positive and negative drift bursts exhibit self-exciting behaviours. Self-exciting means that the presence of an event increases the probability of the arrival of this event in the future. This chapter also shows that negative drift bursts can excite positive drift bursts, but not vice versa. As far as I am aware, I am the first to discover these time series characteristics of drift bursts. Given the forecasting significance of drift found in Chapter 2, Chapter 3 also investigates the impacts of drift bursts on future volatility, with volatility estimated by realized variance. Since realized variance is computed on high-frequency returns, it is inevitably subject to the bias due to microstructure noise. As suggested by Andersen and Bollerslev (1998), the option-implied variance is not affected by microstructure noise in high-frequency return as it is extracted from daily options prices. Therefore, I also consider forecasting implied variance.

The empirical results of Chapter 3 show that the impacts of positive drift bursts on both realized and implied variance are not significant. Negative drift bursts have an insignificant impact on realized variance but significantly increase implied variance. To the best of my knowledge, only very recently have Laurent et al. (2024) investigated the impact of drift bursts on volatility forecasting, but they have not yet investigated the effect of the sign of drift bursts, which is studied extensively in Chapter 3. Additionally, Chapter 3 also demonstrates a significant and positive impact of negative drift bursts on variance risk premium, which gauges the investors' risk aversion via the willingness of investors to pay more for options or variance swaps to

hedge against potential increase in stock market volatility. Although consistent with the findings documented in the existing literature that large downside price moves raise risk aversion (Todorov 2010; Caporin et al. 2017), I am the first to document this for drift bursts.

Although the drift bursts for various financial assets are well documented in the literature, little research has studied the contemporaneous occurrence of drift variation (codrift variations) across these assets. Codrift variation could indicate common bubbles and crashes, which cannot be diversified away and thus have important implications for portfolio risk management and asset allocation (Longin and Solnik 2001; Kole et al. 2006; Anderson and Brooks 2014; Malceniece et al. 2019; Boninsegna and Candelon 2024).

Chapter 4 explores codrift variation among the underlying stocks of the US stock market. There is no codrift variation detection method in the previous studies. I contribute to the literature by demonstrating that applying a coexceedance criterion to a univariate drift test by Shi and Phillips (2024) is feasible to detect codrift variations among stocks. My empirical results show that stock codrift variations are partly associated with the drift bursts in the market portfolio, suggesting that some of the stock codrift variations are hard to diversify. Laurent et al. (2024) show that drift variation of the market portfolio positively predicts market volatility. Given the relationship between market portfolio drift variation and underlying stock codrift variations, stock codrift variations might also lead to higher market volatility. As expected, my results suggest that codrift variations among underlying stocks have a significant and positive impact on market volatility. Importantly, Chapter 4 shows that models exploiting the effects of stock codrift variations lead to significantly better in-

sample and out-of-sample market volatility forecasts compared to using the information from market drift bursts.

The rest of this thesis is organized as follows. Chapter 2 examines the impact of a nonzero drift on the estimation of good and bad volatility, along with the signed jumps, and the influence of these volatility and jump components on volatility forecasting. Chapter 3 investigates self-and mutual excitation behaviours of drift bursts and the power of drift bursts in predicting volatility and the variance risk premium. Chapter 4 explores methods for detecting stock codrift variations, the links between stock codrift variation with market codrift variation, and the impact of stock codrift variation on market volatility forecasting. Chapter 5 concludes.

Chapter 2. Good and bad volatility estimation for drift-diffusion process and the impact of signed jumps on volatility forecasting

Abstract

The logarithmic prices of financial assets are conventionally assumed to follow a drift-diffusion process. The finite sample theory provided in recent studies shows a significant impact of drift on volatility estimation accuracy. However, these studies have not yet investigated the influence of drift on the measurement precision of good and bad volatility, which are important for distinguishing upside from downside risks. This chapter examines the effect of drift on the estimation of good and bad volatility. My finite sample theory and extensive simulations reveal that the impact of drift on the estimation precision of good and bad volatility is even stronger than its impact on the measurement accuracy of overall volatility. I demonstrate that this unsatisfactory estimation of good and bad volatility leads to a dramatic bias in the signed jump estimation. This chapter then suggests an alternative construction of the estimators of good volatility, bad volatility, and signed jumps, showing significant improvement in the estimation accuracy in the presence of non-negligible drift.

Empirical results for the S&P 500 SPDR ETF indicate that both the original and modified volatility estimators consistently exhibit volatility persistence effects. The impact of the original good and bad volatility estimators on future volatility is asymmetric, with a stronger effect from the bad volatility. This contrasts with similar effects of modified good and bad volatility estimators on future volatility. While the original signed jump estimator affects volatility asymmetrically, with its negative sign exhibiting a stronger impact, the predictive value of the modified signed jump estimator on volatility is minor. Given the robustness of modified estimators in the presence of a nonzero drift, I conclude that good and bad volatility may closely impact future volatility, and signed jumps have little power in predicting volatility. My findings also suggest that the asymmetric impacts of the original semivariances and the original signed jump estimators on volatility forecasting, as found in the existing literature, are almost exclusively attributed to the biases in these measures due to a nonzero drift.

2.1. Introduction

It is widely believed that asset prices obey an Ito semi-martingale process. In the high-frequency financial econometrics literature, asset returns are typically modelled as an Ito semi-martingale process with two main components: drift and diffusion.[1] With locally bounded coefficients of drift and volatility as in the volatility estimation literature, diffusion dominates the returns as the time interval of returns becomes small, and the drift component is ignorable even if the drift is nonzero. According to this asymptotic theory, the estimation of volatility, based on aggregate functions of intraday returns over an estimation window, is unaffected by the presence of a nonzero drift. Typical volatility estimators include the realized variance (RV) of Andersen and Bollerslev (1998) and the bipower variation (BV) introduced by Barndorff-Nielsen and Shephard (2004).

For the volatility estimation for real financial markets, there is often only a finite sample of return observations. For example, studies tend to take 78 returns per day for the official trading session of the US stock markets based on a 5-minute grid, for balancing the trade-off between getting as close as possible to asymptotic properties and protecting against market microstructure impacts (Anderson et al. 2003; Hansen and Lunde 2006; Andersen et al. 2007b; Bollerslev et al. 2009; Corsi 2009; Corsi and Renò 2012; Liu et al. 2015; Patton and Sheppard 2015; Bollerslev et al. 2016; Bollerslev et al. 2020; Bollerslev et al. 2021; Buccheri and Corsi 2021; Bollerslev 2022). Recently, Laurent and Shi (2020) have argued that for finite samples, the drift component in returns is not ignorable relative to the volatility. They show that due to

[1] As in literature, the log prices are also occasionally driven by jumps. Jumps are rare thus log prices are mainly driven by drift and diffusion.

this reason RV and BV, which are computed on the returns, will capture significant drift components and thus can be upwardly biased in estimating the volatility.

However, Laurent and Shi (2020) and other existing studies have not yet investigated the finite sample impact of nonzero drifts on the estimation of volatility associated with upside and downside price moves (also known as good and bad volatility) in the high-frequency setting, although previous literature has addressed the negative effect of price drift on the measurement precision of good and bad volatility in a low-frequency setting e.g., monthly data (Markowitz 1959; Porter 1974; Markowitz 1991). The literature on such volatility estimation focuses on the asymptotic properties (Barndorff-Nielsen et al. 2008; Patton and Sheppard 2015; Bollerslev et al. 2020; Bollerslev 2022), as opposed to its finite sample properties. The estimation of the good and bad volatility is important as this helps investors disentangle the risk of the price moving up from the risk of the price falling. It has been long recognized that up and downside risks are not treated the same by investors. Agents who assign a larger weight to downside risk require additional compensation for holding financial assets with high sensitivity to downside market movements. In the asset pricing literature, downside volatility is also introduced as an explanatory variable to price the cross-section of equity, equity index options, commodity, sovereign bond, and currency returns (Ang et al. 2006; Lettau et al. 2014; Farago and Tédongap 2018).

This chapter studies the finite sample influence of nonzero drifts on the measurement of good and bad volatility based on high-frequency data. I consider popular good and bad volatility estimates defined by Barndorff-Nielsen et al. (2008), indicated by the positive and negative realized semivariances, respectively. These two signed semivariances have been applied to estimate good and bad volatility in a broad

range of research related to forecasting stock price returns and volatility, predicting cross-sectional stock returns, and option pricing (Barndorff-Nielsen et al. 2008; Patton and Sheppard 2015; Bollerslev et al. 2016; Feunou and Okou 2019; Bollerslev et al. 2020; Bollerslev et al. 2021; Bollerslev 2022).

For studying the finite sample performance of the semivariance in the presence of nonzero drifts, I consider two popular drift-diffusion models: a constant drift-diffusion process and a linear drift-diffusion process, with the latter also named the Ornstein-Uhlenbeck process. The two processes contain the important dynamic of nonzero drifts and are important to option pricing, value-at-risk assessments, and volatility forecasting in the literature (Barndorff-Nielsen and Shephard 2001; Aalen and Gjessing 2004; Wang and Yu 2016; Laurent and Shi 2020; Laurent et al. 2022b). My finite sample theory, together with extensive simulations, uncovers that a nonzero drift has opposite impacts on the estimation accuracy of signed semivariances: a positive drift can result in a positive bias in positive semivariance and a negative bias in negative semivariance, and a negative drift causes a negative bias in positive semivariance and a positive bias in negative semivariance. Importantly, the proportion of the bias in the semivariances can be even larger than that in RV and BV estimators found by Laurent and Shi (2020).

Evidence of stochastic skewness and kurtosis of asset return distributions has led to the development of models with jumps to better capture these dynamics. Jumps are rare and larger events than what can be explained by the classic drift-diffusion process. Under various cases of jumps, my simulation results show that a nonzero drift can still be not ignorable to the estimation performance of the semivariances, with a positive bias in positive semivariance and a negative bias in negative

semivariance again due to a positive drift, and a negative bias in positive semivariance and a positive bias in negative semivariance again because of a negative drift.

Based on the limiting results of semivariances under jump cases, Barndorff-Nielsen et al. (2008) discover that the difference between two semivariances can estimate the signed jumps.[2] Compared to good and bad volatility, signed jumps, perhaps due to their association with good and bad news (Evans 2011; Lahaye et al. 2011; Gilder et al. 2014), exhibit even broader applications in the literature. Signed jumps are utilized as an important risk factor in forecasting equity risk premium (Feunou et al. 2018), help explain credit default swap spreads (Lee and Hyun 2019), contribute to the predictability of the cross-section of expected stock returns (Mizrach et al. 2018; Bollerslev et al. 2020), and is particularly found important in volatility forecasting (Sévi 2014; Patton and Sheppard 2015; Bee et al. 2016; Lyócsa and Molnár 2016; Wang et al. 2016; Todorova 2017; Gong and Lin 2021; Özbekler et al. 2021; Slim et al. 2023; Zhu et al. 2023).

However, my analytical and simulative results show the estimation accuracy of signed jump estimator based on the difference between signed semivariances can also be impacted by a nonzero drift. The sign of this bias aligns with that of drift. Since the sign of the bias in positive and negative semivariances is opposite for a nonzero drift, the size of the bias of this signed jump estimator can assemble the magnitude of the biases in both semivariances. Therefore, the bias in the signed jump estimator is even greater than those in the signed semivariances.

To reduce the impact of nonzero drifts on the estimation of good and bad volatility and signed jumps, I suggest calculating the positive and negative realized

[2] This chapter uses the terms signed jump variation and signed jumps interchangeably.

semivariances on returns centred by the median of these returns over the volatility estimation window. This centred-return approach is proposed by Laurent and Shi (2020), and the intuition is that centring returns removes the drift, which is estimated by the daily median of the returns, from each of the intraday returns. To the best of my knowledge, I am the first to apply the centred-return approach to realized semivariances under high-frequency data. My analytical and simulation results of the constant and linear drift-diffusion models indicate that the modified semivariances and signed jump estimator leads to a dramatic improvement in the estimation precision of good and bad volatility and signed jumps, especially if the drift deviates far from zero. Importantly, my results show that the impact of a nonzero drift on the estimation performance of the modified semivariances and signed jump estimator is equal or even smaller than on the performance of the modified realized variance and bipower variation found in Laurent and Shi (2020).

For illustration purposes, I apply the modified positive semivariance, negative semivariance, and signed jump estimator, along with their original versions, to the returns of SPDR S&P 500 ETF (SPY) from 1997 to 2021, sampled at the most popular 5-minute frequency. When drift is positive (drift is estimated by the daily median of intraday returns), the positive semivariance tend to overestimate good volatility by 18% and negative semivariance underestimates bad volatility by 15% on average. In the presence of a negative drift, the positive semivariance deflates good volatility by about 15% on average and negative semivariance inflates bad volatility by 18% on average. The signed jump estimator excessively estimates signed jumps by 101% on average if the drift is positive, and commonly underestimates signed jumps by 107% if the drift is negative. These biases, especially the one in the signed jump

estimation, are much greater than those in BV as proposed by Laurent and Shi (2020), which typically exaggerates volatility by 1.2%.

Signed jumps play an important role in finance. Observers of financial markets have long noted that financial asset prices exhibit unusual upside and downside large moves, compared to what would be expected from a random walk. Recent financial research has suggested the importance of explicitly allowing for signed jumps in the prediction of price volatility option pricing, credit spread interpretation, and pricing of cross-sectional returns and options (Patton and Sheppard 2015; Feunou and Okou 2019; Bollerslev et al. 2020; Bollerslev 2022; Caporin 2023). The improvement in the estimation of signed jumps thus has important implications for these financial applications.

As mentioned, volatility, semivariances, and signed jumps have a wide financial application. More accurate estimation of volatility and signed jump estimation indicated in this chapter therefore may provide value to these applications. As an example, I investigate in this chapter if the finite sample bias of the estimation of these variables could affect the previous conclusions in the literature related to the impacts on volatility forecasting of the volatility persistence, the asymmetric effect of good and bad volatility, and the asymmetric impact of signed jumps. As noted in the introduction of this thesis, volatility forecasting is important to risk management, portfolio optimization, and pricing derivatives.

The literature concludes significantly persistent volatility dependence using the realized variance and bipower variation as the estimators of volatility (Andersen et al. 2007b; Corsi 2009). My results show that both original and modified versions of realized variance and bipower variation exhibit very similar volatility persistence

effect, which indicates that the persistent volatility dependence effect is robust when realized variance and bipower variation are modified for reducing drift biases.

Relying on the signed realized semivariance estimators, previous studies demonstrate that good and bad volatilities asymmetrically impact future volatility, with a stronger effect from bad volatility (Sévi 2014; Patton and Sheppard 2015; Audrino and Hu 2016; Wang et al. 2016; Slim et al. 2023). I obtain similar findings using the original signed semivariance. However, I find that the effects of the modified positive and negative semivariance on future volatility are not different from each other.

Using the signed jump estimator built on the semivariance difference, previous studies demonstrate that both positive and negative jumps are important to predicting future volatility, with positive jumps decreasing volatility and negative jumps increasing volatility more strongly (Sévi 2014; Patton and Sheppard 2015; Audrino and Hu 2016; Wang et al. 2016; Slim et al. 2023). As in the literature, I also observe an asymmetric impact of signed jumps according to the forecasting results related to the original signed jump estimator. However, my findings of the modified signed jump estimator tell a story completely different from the existing literature: neither positive jumps nor negative jumps can help forecast volatility.

My analysis also suggests that the asymmetric impacts of original signed semivariances for future volatility may be mostly due to their biases related to a nonzero drift. This implies that the asymmetric effects of good and bad volatility on volatility forecasting found in previous research may be almost exclusively attributed to the drift bias in the semivariance estimators. I show that the asymmetric effects of the original signed jump estimator for volatility forecasting may also be caused by the

drift-driven bias in this estimator, which implies that the importance of signed jumps found in existing studies may be almost exclusively attributed to the estimation bias of signed jumps given the impact of a nonzero drift.

The rest of this chapter is structured as follows. Section 2.2 introduces the good and bad volatility measurement using positive and negative realized semivariances. The impact of a nonzero drift on semivariances for the constant drift diffusion model and the Ornstein-Uhlenbeck process are reported in Sections 2.3 and 2.4, respectively. Section 2.5 studies the influence of a nonzero drift on the signed jump estimation. Section 2.6 reports stock market evidence of the effect of a nonzero drift on the measurement performance of the semivariances and signed jump estimator in the presence of a nonzero drift. Section 2.7 contains the results for the impact of both the original and modified versions of signed semivariances and signed jump estimator on volatility forecasting. Section 2.8 concludes. All the mathematical proofs of the lemmas, propositions, and corollaries in this chapter are reported in the Appendix A.1..

2.2. Theoretical framework

Let p_t denote a logarithmic asset price at time t. The process for p_t is conventionally expressed in stochastic differential equation form,

$$dp_t = \mu_t dt + \sigma_t \mathrm{d}W_t + J_t dN_t, 0 \leq t \leq T, \tag{2.2.1}$$

where the drift coefficient is a linear function of the log price $\mu_t = \theta p_t$ thus is time-varying, with θ constant. The diffusion coefficient σ_t denotes a stochastic volatility process, W_t is a standard Brownian motion, and J_t represents the random jump size at time t, and N_t is an independent Poisson counting process with a time-varying intensity, indicating that the arrival of jumps is also random.

The Quadratic Variation for the cumulative return process $r_t = p_t - p_0$, is then

$$\mathrm{QV}(0,\mathrm{t}) = \int_0^t \sigma_s^2 ds + \sum_{0<s\leq T} \kappa_s^2, \tag{2.2.2}$$

where QV_t indicates the quadratic variation, $\int_0^t \sigma_s^2$, which is termed the continuous variation or integrated variation, denotes the component of quadratic variation due to the continuous price process, and $\sum_{0<s\leq T} \kappa_s^2$ is the remaining jump variation, computed on the summation of squared jumps, with κ_s capturing the size of a jump. When jumps are absent, the summation vanishes, and the quadratic variation simply equals the integrated volatility of the continuous variation.

Let log prices p_{t_i} be observed at M equally spaced intervals $0 = t_0 < t_1 < t_2 \ldots < t_M < 1$ spanning one trading day t, where I normalized the daily time interval to unity for ease of notation. The distance between two consecutive observation times is then denoted by $\Delta = t_i - t_{i-1} = 1/M$. The returns computed using log prices at

equally spanned observation times may be written as $r_{t_i} = p_{t_i} - p_{t_{i-1}}$, with $i = 1,2,\dots,M$. Then, the Quadratic Variation of the log prices for day t is $QV_t = \int_{t-1}^{t} \sigma_s^2 ds + \sum_{(t-1)\leq s\leq t} \kappa_s^2$. To estimate this quadratic variation, Andersen and Bollerslev (1998) suggest a realized variance (RV) estimator, which is defined by the aggregation of all M squared intraday returns within day t,

$$RV_t = \sum_{i=1}^{M} r_{t_i}^2. \tag{2.2.3}$$

Andersen and Bollerslev (1998) show that RV converges to the volatility and jump components as the interval between observations gets smaller,

$$RV_t = \sum_{i=1}^{M} r_{t_i}^2 \xrightarrow{p} \int_{t-1}^{t} \sigma_s^2 ds + \sum_{(t-1)\leq s\leq t} \kappa_s^2, \text{ as } \Delta \to 0. \tag{2.2.4}$$

The above equation shows that in the presence of jumps, RV also captures the jump component and thus is not robust enough to estimate the continuous variation. For volatility estimation robust to jumps, Barndorff-Nielsen and Shephard (2006) proposed bipower variation (BV), defined by the summation of appropriately scaled cross-products of adjacent high-frequency absolute returns,

$$BV_t = \frac{\pi}{2} \sum_{i=2}^{M} |r_{t_{i-1}}||r_{t_i}| \tag{2.2.5}$$

As noted by Barndorff-Nielsen and Shephard (2006), BV only converges in probability to the continuous variation if the time gap between observations becomes small,

$$BV_t = \frac{\pi}{2}\sum_{i=2}^{M}|r_{t_{i-1}}||r_{t_i}| \xrightarrow{p} \int_{t-1}^{t}\sigma_s^2 ds\,, \text{ as } \Delta \to 0. \tag{2.2.6}$$

Both RV and BV estimate the volatility as a whole and do not differentiate the upside and downside price variation. On the other hand, pioneering researchers have argued that investors are primarily concerned about negative returns and downside risks (Roy 1952; Markowitz 1959). Accordingly, the fundamental mean-variance trade-off arguments that form the foundation of many widely applied asset pricing models and predictions, including the classic Capital Asset Pricing Model, should instead depend upon the downside part of the variation only (Hogan and Warren 1972; Bawa and Lindenberg 1977). Many studies in behavioural finance, supported by empirical evidence and more formal theoretical arguments rooted in prospect theory and loss aversion, also indicate that investors treat up and downside differently (Kahneman and Tversky, 1979).

Motivated by these studies and the idea that downside and upside risk are treated differently, Barndorff-Nielsen et al. (2008) first proposed decomposing the original realized variance (RV) measure into separate up and downside proportions, according to the summation of the squared positive and squared negative intraday returns, respectively. Specifically, the positive component of RV is termed realized positive semivariance (RS^+), defined by,

$$RS_t^+ = \sum_{i=1}^{M} r_{t_i}^2 I_{r_{t_i}>0}, \tag{2.2.7}$$

and the remaining negative part of RV is called realized negative semivariance (RS^-),

$$RS_t^- = \sum_{i=1}^{M} r_{t_i}^2 I_{r_{t_i}<0}. \tag{2.2.8}$$

Additionally, Barndorff-Nielsen et al. (2008) investigate the asymptotic properties of realized semivariances. They show that as the size of the intervals gets smaller, the positive realized semivariance converges to one-half of the continuous variation plus the jump variation only due to positive jumps,

$$RS_t^+ \xrightarrow{p} \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds + \sum_{(t-1)\le s\le t} \kappa_s^2 I_{\kappa_s>0}, \text{ as } \Delta \to 0. \tag{2.2.9}$$

This summation of one-half integrated variation and positive jumps is defined as the good volatility in this chapter, following the literature. Barndorff-Nielsen et al. (2008) also demonstrate that the negative realized semivariance converges to one-half of the continuous variation plus the jump variation only attributed to negative jumps,

$$RS_t^- \xrightarrow{p} \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds + \sum_{(t-1)\le s\le t} \kappa_s^2 I_{\kappa_s<0}, \text{ as } \Delta \to 0. \tag{2.2.10}$$

And this summation of one-half integrated variation and negative jumps is defined as the bad volatility as in the literature. In the absence of jumps, both good volatility and bad volatility are equal to the one-half continuous variation.

2.3. Realized Semivariance estimation for a constant drift-diffusion process

From previous Equations (2.2.4), (2.2.6), (2.2.9) and (2.2.10), it is important to note that the above volatility estimators do not contain the variation due to the drift μ_t, although the process of log prices indeed contains the drift component, as in Equation (2.2.1). This can be explained by noting that as the time gap between log prices becomes small $\Delta \to 0$, the continuous volatility or jumps always dominate the drift term in high-frequency returns even if the drift is not zero (Laurent and Shi 2020; Christensen et al. 2022; Laurent et al. 2022b). However, empirical research tends to use a finite sample of data where the interval between returns is not infinitesimally small, either because ultra-high-frequency data is not always available, or to balance the trade-off between the distortion from market microstructure noise in the higher frequency data versus aiming to reach the asymptotic property of the volatility estimators. Laurent and Shi (2020) show that drift may not necessarily be ignorable in the returns for a finite sample under a constant drift-diffusion model. They show that realized variance and bipower variation estimators, consequently, may be biased and their bias increases with the magnitude of drift. However, as far as I am aware, Laurent and Shi (2020) and existing research have not yet considered this finite sample bias in the positive and negative semivariances, although these measures, as mentioned in the introduction, are important estimators of upside and downside risk and widely applied in economic forecasting in the literature.

2.3.1. The bias of realized semivariances

For expositional purposes, I begin the investigation of the estimation bias of signed semivariances under a simple constant drift-diffusion model. This model assumes that both the drift and volatility of the log prices are constant,

$$dp_t = \mu dt + \sigma \mathrm{d}W_t. \tag{2.3.1}$$

Although the constant drift-diffusion model has simple assumptions for the volatility and drift processes, it provides ease of exposition for the beginning of my analysis of the biases of realized semivariances. For robustness purposes, the next section 2.4 will also study the biases of realized semivariances when the log prices follow a linear drift-diffusion model, which contains more general assumptions for drift and volatility as in the literature. For the constant drift-diffusion model, I begin the analysis by considering the case of no jumps and then introduce the scenarios with additive jumps in the following section 2.3.3. When jumps do not occur, the probability limit of positive and negative realized semivariances uniformly converge to the one-half integrated variance $\frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds$ as the sampling interval becomes small, as in Equations (2.2.9) and (2.2.10). Note that under the constant drift-diffusion model, the daily integrated variance equals the constant volatility, $\int_{t-1}^{t} \sigma_s^2 ds = \sigma^2$ as the volatility is assumed constant.

The log prices of the above constant drift-diffusion model, Equation (2.3.1) (computed using log prices at equally spaced observation times) may be written as,

$$p_{t_i} = p_{t_{i-1}} + \mu\Delta + \sigma\sqrt{\Delta}\varepsilon_{t_i} \text{ with } \varepsilon_{t_i} \sim N(0,1) \text{ and } \Delta = 1/\mathrm{M}. \tag{2.3.2}$$

The returns for day t are defined by $r_{t_i} = p_{t_i} - p_{t_{i-1}}$, and the number of returns is M. For example, $M = 78$ for the popular five-minute frequency sampling for the US stock market main trading session from 9:30 to 16:00 EST while $M = 390$ for the 1-minute frequency. When the log prices follow Equation (2.3.2) with a nonzero drift μ, the expected number, or frequency, of positive and negative returns may not be equal, which is illustrated in the Lemma 2.3.1 below,

Lemma 2.3.1. *For M returns r_{t_i} obtained from the drift-diffusion process, Equation (2.3.2), for $t = t_1, t_2, \dots, t_M$ of day t.*

(i) The expected frequency of positive returns and negative returns for day t are

$$M\mathbb{P}(r_{t_i} > 0) = M\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right], \tag{2.3.3}$$

And

$$M\mathbb{P}(r_{t_i} < 0) = M \times \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right), \tag{2.3.4}$$

respectively, where $\mathbb{P}(\cdot)$ denotes the probability of an argument and $\Phi(\cdot)$ is the cumulative function of the standard normal distribution.

(ii) if $\mu = 0$, the expected frequencies of positive and negative returns for day t are identical and equal to $M/2$, with $M\mathbb{P}(r_{t_i} > 0) = M\mathbb{P}(r_{t_i} < 0) = M/2$.

The statement (*i*) of *Lemma 2.3.1* indicates that for the constant drift-diffusion process and a finite number, M, of returns observed on day t, the returns on day t may be overwhelmed by positive or negative returns if the drift is much larger in

magnitude than the volatility. For example, given that Δ is fixed and $\mu/\sigma \to -\infty$, the expected frequency of the negative return within day t is $M\mathbb{P}\left(r_{t_i} < 0\right) = M\Phi\left(-\mu\sqrt{\Delta}/\sigma\right) \approx M$ while the expected frequency of positive returns is $M\left[1 - \Phi\left(-\mu\sqrt{\Delta}/\sigma\right)\right] \approx 0$. This indicates that when $\mu/\sigma \to -\infty$, almost all M returns at day t are expected to be negative, indicating an extreme persistence of returns.

Literature has addressed the possibility of $\mu/\sigma \to \pm\infty$, which implies a volitation of the no-arbitrage principle of asset pricing theory (Laurent and Shi 2020; Bollerslev 2022; Christensen et al. 2022; Laurent et al. 2022b; Andersen et al. 2023), and these studies argue that the extreme return persistence caused by $\mu/\sigma \to \pm\infty$ is a stylized fact in the stock markets. Since all intraday returns at day t are expected to be negative for $\mu/\sigma \to -\infty$, the negative semivariance is expected to be identical to realized variance, which equals the volatility plus a positive bias due to the negative μ. Then negative semivariance overestimate one-half integrated variation by sum of one-half integrated variation and this positive drift bias, suggesting the impact of a large negative drift on the estimation accuracy of negative realized semivariance can be even greater than its impact on the measurement precision of realized variance. At the same time, since there is almost no positive return due to this extremely negative drift, the positive realized semivariance, which is calculated on positive returns, is close to zero. As a result, the positive realized semivariance underestimates the one-half integrated variation by nearly 100%.

When the returns of day t are overwhelmingly positive by a large, positive drift, the positive realized semivariance can be close to realized variance. For this case of the drift, analogously, the bias ratio of positive realized semivariance is much larger than that of realized variance, and the negative realized semivariance underestimates

the one-half integrated variation by 100%. These findings under the extreme cases of the drift in the constant drift-diffusion model suggest that the proportion of the bias in semivariances can be larger than that in realized variance or close to 100% when drift is extremely large. Therefore, the existence of a nonzero drift may raise a concern in the estimation performance of the realized semivariances.

The above intuition on the impact of a nonzero drift on the estimation of realized semivariances depends on the extreme cases of drift. To highlight the generality of the result, I derive the exact bias of realized semivariances as a function of drift, μ, as shown by the following Proposition 2.3.1. The bias, defined by the difference between signed semivariances with a-half integrated variance, explores how signed semivariances deviate from estimating the one-half integrated variance based on the asymptotic theory. This measure allows us to identify the size and the sign of the bias contained in semivariances. Analogously, the bias in the realized variance and bipower variation can also be derived by comparing these two estimators with integrated variance, and these two biases have been derived by Laurent and Shi (2020).

Semivariances attempt to estimate one-half integrated variance, which is a different (smaller) amount of volatility compared to that realized variance and bipower variation aim to measure. For a fair comparison of the impact of a nonzero drift on the estimation accuracy of these three estimators, I also consider a ratio measure of the bias. For semivariances, the bias ratio is defined by the proportion of the bias relative to the a-half integrated variance. For realized variance and bipower variation, their bias ratios are defined by the proportion of their biases relative to full integrated variance. The bias ratio measure will also be key to ensure the fairness of comparing the impacts of a nonzero drift on the estimation performance of

semivariances, realized variance and bipower variation for the cases of jumps, discussed in the next section. This is because in the presence of jumps, positive semivariance, negative semivariance realized variance, and bipower variation estimate different price variations from each other, according to the asymptotic theories as in Equations (2.2.4), (2.2.6), (2.2.9), and (2.2.10).

Proposition 2.3.1. *Under the drift-diffusion process, Equation (2.3.2).*

(i) The bias in positive realized semivariance is

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= \mu^2\Delta\left[1-\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \sigma^2\left[1-\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] - \frac{1}{2}\sigma^2, \quad (2.3.5)$$

where $\varphi(x)$ is the probability density function of the standard normal distribution and $\Phi(x)$ is the respective cumulative distribution function.

(ii) The bias ratio of positive realized semivariance is

$$\mathbb{E}\left[\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]$$

$$\frac{2\mu^2\Delta}{\sigma^2}\left[1-\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 1. \quad (2.3.6)$$

(iii) The bias in negative realized semivariance is

$$\mathbb{E}\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= \mu^2 \Delta \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \sigma^2 \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \frac{1}{2}\sigma^2. \tag{2.3.7}$$

(iv) The bias ratio of negative realized semivariance is

$$\mathbb{E}\left[\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right]$$

$$= \frac{2\mu^2\Delta}{\sigma^2}\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 1. \tag{2.3.8}$$

Proposition 2.3.1 shows that the bias and the bias ratio of realized semivariances are related to values of Δ, μ, and σ. For studying the properties of the bias and bias ratio of realized semivariances, several corollaries are deduced from Proposition 2.3.1.

Corollary 2.3.1 *When drift is zero, $\mu = 0$, the bias and the bias ratio of both positive and negative realized semivariance are zero.*

Corollary 2.3.2 *For any positive drifts, $\mu > 0$, the bias ratio of the positive realized semivariance is larger than that of realized variance,*

$$\mathbb{E}\left[\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right] > \mathbb{E}\left[\left(RV_t - \int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \int_{t-1}^{t} \sigma_s^2 ds\right], \tag{2.3.9}$$

where

$$\mathbb{E}\left[\left(RV_t - \int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \int_{t-1}^{t} \sigma_s^2 ds\right] = \frac{\mu^2\Delta}{\sigma^2},$$

denotes the bias ratio of realized variance, according to Proposition 2.1 of Laurent and Shi (2020). For any negative drifts $\mu < 0$, the bias ratio of

positive realized semivariance is more negative than that of realized variance, indicated by

$$\left[\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right] < \mathbb{E}\left[\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\int_{t-1}^{t}\sigma_s^2 ds\right]. \quad (2.3.10)$$

Corollary 2.3.3 *For any negative drifts $\mu < 0$, the bias ratio of negative realized semivariance is more positive than that of realized variance, indicated by*

$$\left[\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right] > \mathbb{E}\left[\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\int_{t-1}^{t}\sigma_s^2 ds\right]. \quad (2.3.11)$$

For any positive drifts $\mu > 0$, the bias ratio of negative realized semivariance is more negative than that of realized variance,

$$\left[\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right] < \mathbb{E}\left[\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\int_{t-1}^{t}\sigma_s^2 ds\right]. \quad (2.3.12)$$

Corollary 2.3.2 and 2.3.3 help identify how impactful the bias contained in realized semivariance is, via its comparison to that of realized variance found by Laurent and Shi (2020). These two corollaries imply that whenever the signs of drift and semivariance are equal, the impact of the drift is stronger on the semivariance than on realized variance. Note that this result holds for any sampling frequencies, magnitudes of the drift and volatility levels. However, if the signs of drift and semivariance are not equal, whether the bias ratio of semivariance is larger or smaller

relative to that of realized variance is unknown, and we only know the bias proportion of realized semivariance is more negative than that of realized variance. For comparing the bias of the realized semivariance with that of the realized variance when the signs of drift and semivariance differ, the next sub-section 2.3.2 will calculate the biases of realized semivariances on some realistic and empirically important ranges of drift, volatility, and sampling intervals.

> ***Corollary 2.3.4*** *The bias in positive realized semivariance in the presence of a nonzero drift* $\mu = \mu^*$ *with* $\mu^* \neq 0$ *is identical to that in negative realized semivariance due to the same magnitude drift with the opposite sign* $\mu = -\mu^*$*. This holds for the bias ratio measure.*

Corollary 2.3.4 indicates that the bias in positive and negative realized semivariances are symmetrical across the signs of drifts, and this symmetric pattern holds for the bias ratio measure.

To reduce the effect of drift on the estimation performance of realized variance and bipower variation, Laurent and Shi (2020) suggest modifying these two estimators by replacing the return r_{t_i} with its centred version $r_{t_i} - \mu$, where μ denotes the constant drift in the returns on day t, $r_{t_1} \dots r_{t_M}$. Since the centred returns no longer contain drift in $t_1 \dots t_M$, their summations within this estimation window are robust to drift. This removes the bias in realized variance and bipower variation due to a nonzero drift. The drift μ in returns in Equation (2.3.2) is a latent variable and thus needs to be estimated. Laurent and Shi (2020) suggest measuring the drift by the median of M returns of day t, $\widehat{m}_t = \text{median}(r_{t_1} \dots r_{t_M})$, where $\widehat{m}_t$ indicates the median estimator. They show that the median is an unbiased estimator of drift. Moreover, Laurent and Shi (2020) find that the median empirically perform well in

capturing large price drift in the stock market during the dot-com bubble period in the late 1990s (Phillips et al. 2011; Phillips et al. 2015; Shi and Song 2016) and the commodity markets over the preceding decade (Phillips and Yu 2011; Gutierrez 2013; Etienne et al. 2014). Based on this median estimator, the modified realized variance and bipower variation (indicated by RV^* and BV^*, respectively) are defined by,

$$RV_t^* = \sum_{i=1}^{M} \left(r_{t_i} - \widehat{m}_t\right)^2, \tag{2.3.13}$$

$$BV_t^* = \frac{\pi}{2} \sum_{i=2}^{M} \left|r_{t_{i-1}} - \widehat{m}_t\right| \left|r_{t_i} - \widehat{m}_t\right|, \tag{2.3.14}$$

where $\widehat{m}_t = \mathrm{median}\left(r_{t_1} \ldots r_{t_M}\right)$ denotes the median of M returns for day t that are involved in the computation of the volatilities for that day. The theoretical and simulative results of Laurent and Shi (2020) show that the finite sample biases of the modified realized variance and bipower variation in estimating the volatility are much smaller than their original versions.

It is straightforward to apply the centred returns idea to construct modified realized semivariances if the interest is to alleviate the impact of nonzero drifts on semivariances. This approach removes drift and thus reduces the effect of drift on the estimation precision, which is also in the spirit of Porter (1974) who applies the de-averaged monthly returns to semivariances to reduce the bias due to the price drift in the long run.[3] The algebraic expression of these two modified realized semivariances is defined by,

[3] As in Laurent and Shi (2020), I do not assume drift to be nonzero over a long time span but do so only during the period over which the volatility estimators are computed.

$$RS_t^{+*} = \sum_{i=1}^{M} \left(r_{t_i} - \widehat{m}_t\right)^2 I_{r_{t_i} - \widehat{m}_t > 0}, \tag{2.3.15}$$

$$RS_t^{-*} = \sum_{i=1}^{M} \left(r_{t_i} - \widehat{m}_t\right)^2 I_{r_{t_i} - \widehat{m}_t < 0}. \tag{2.3.16}$$

where RS^{+*} and RS^{-*} indicate the modified positive and negative realized semivariances, respectively, and obviously, the sum of RS^{+*} and RS^{-*} equals the modified realized variance, $RV_t^* = RS_t^{+*} + RS_t^{-*}$. Note that using centred returns does not change the asymptotic property of realized measures as the median ($\widehat{m}_t$) is negligible relative to volatility when the sampling interval becomes small. This modification aims to increase their estimation accuracy for a finite sample. The impact of drift on the modified realized semivariances is expected to be smaller compared to their original version as the drift component is treated in the centred returns. I begin the exact derivation of the bias of the modified realized semivariances by introducing the following Lemma 2.3.2.

Lemma 2.3.2. *For the returns r_{t_i} obtained from the drift-diffusion process, Equation (2.3.2) for $t = t_1 \dots t_M$ on day t, the expected number of the positive centred returns $r_{t_i} - \widehat{m}_t > 0$ and that of negative centred returns $r_{t_i} - \widehat{m}_t < 0$ are uniformly equal to $M/2$ for any drift μ.*

According to Lemma 2.3.2, for M returns under the constant drift-diffusion process, Equation (2.3.2), the expected numbers of positive and negative centred returns $r_{t_i} - \widehat{m}_t$ are equal. In the presence of a large price drift across a day, the centred returns will not be overwhelmed by positive or negative returns and the modified realized semivariances RS^{+*} and RS^{-*} are calculated using $M/2$ number of

positive and negative centred returns each. Using this result, the exact bias of the RS^{+*} and RS^{-*} can be given by Proposition 2.3.2 below.

Proposition 2.3.2. *Under the drift-diffusion process, Equation (2.3.2),*

(i) The bias in both the modified positive and negative realized semivariances equals one-half of the bias in the modified realized variance,

$$\begin{aligned}\mathbb{E}\left(RS^{+*}-\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right) &= \mathbb{E}\left(RS^{-*}-\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\\ &= \frac{1}{2}\mathbb{E}\left(RV^{*}-\int_{t-1}^{t}\sigma_s^2 ds\right),\end{aligned} \tag{2.3.17}$$

where $\mathbb{E}\left(RV^{*}-\int_{t-1}^{t}\sigma_s^2 ds\right)$ *is the bias in the modified realized variance and is given by*

$$\mathrm{M}\left[\mathbb{v}(\widehat{m}_t)-2cov\left(\frac{1}{M}\sum_{1}^{M} r_{t_i},\widehat{m}_t\right)\right],$$

as proposed by Laurent and Shi (2020), with $\mathbb{v}$ *denoting the variance of the argument and* cov *indicating the covariance between the two arguments.*

(ii) The bias ratio of both the modified positive and negative realized semivariances equals that of the modified realized variance,

$$\mathbb{E}\left[\left(RS^{+*}-\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right] \tag{2.3.18}$$

$$= \mathbb{E}\left[\left(RS^{-*} - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]$$

$$= \mathbb{E}\left[\left(RV^{*} - \int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\int_{t-1}^{t}\sigma_s^2 ds\right].$$

Proposition 2.3.2 suggests that the bias in RS^{+*} and RS^{-*} are identical and this bias is only half of that of RV^{*}, which does not depend on drift μ but relies on the variance of the median and the covariance between the sample mean and the median. The bias ratio of RS^{+*} and RS^{-*} are identical and equal to that of RV^{*}, which indicates that the bias is equally impactful to the estimation precision of RS^{+*}, RS^{-*}, and RV^{*}. As suggested by Laurent and Shi (2020), the exact variance of the median can be obtained from the formula of Gupta and Nadarajah (2005) while the covariance between the median and mean can be proxied based on their asymptotic joint distribution in Ferguson (1999).

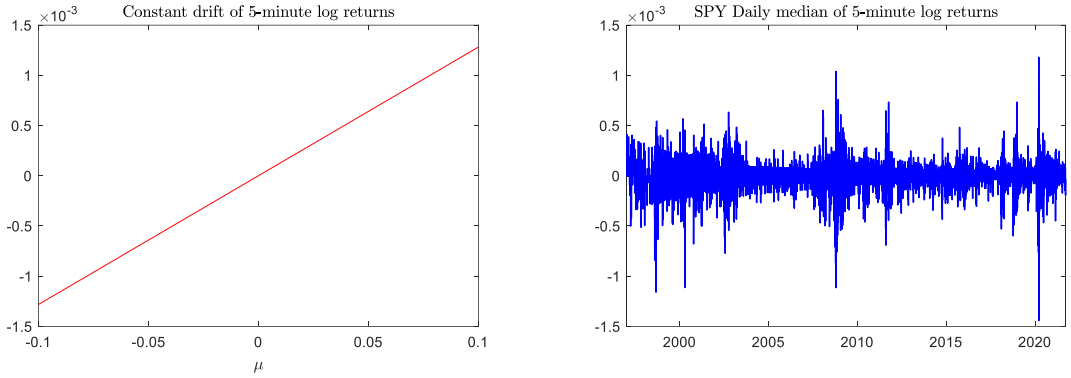
2.3.2. Visualization of the bias

To illustrate the bias in original and modified realized semivariances derived above, this section calculates these biases on some realistic parameter settings of the constant drift-diffusion model, Equation (2.3.2), for the sampling interval Δ, the volatility σ, and drift μ. The settings of these three parameters are as follows. The sampling frequency is 5 minutes, motivated by the fact that realized semivariances are generally computed on 5-minute log prices in financial studies (Patton and Sheppard

2015; Feunou and Okou 2019; Bollerslev et al. 2020; Bollerslev 2022). For this sampling grid, the sampling interval of the returns is $\Delta = 1/M = 1/78$ to be consistent with main trading session 9:30 to 16:00 EST of the SPDR S&P 500 ETF (SPY). SPY have been widely utilized to approximate the overall US stock market in the literature, and full details on SPY will be introduced in later section 2.6. The overnight returns are not considered, following an extensive body of volatility estimation literature (Andersen et al. 2007b; Corsi 2009; Duong and Swanson 2015; Patton and Sheppard 2015; Bollerslev et al. 2016; Bollerslev et al. 2021; Andersen et al. 2023; Caporin 2023). The level of the constant volatility is set to $\sigma = 0.01$, identical to around 0.159 annualized volatility, a realistic value for equity returns. The drift μ is allowed to change from -0.1 to 0.1 with increments of 0.01. This range of μ is consistent with Laurent and Shi (2020) but further truncated to better fit into my SPY data.

Figure 2.3.1 compares the constant drift of the 5-minute frequency return (left panel) with the empirical estimate of such drift of the SPY data (right panel). The sample length of SPY data is from 1997 to 2021. As depicted in the left panel, the range of drift from -0.1 to 0.1 corresponds to between $\pm 1.5 \times 10^{-3}$ at the 5-minute frequency. The right panel shows that my setting of drift in the 5-minute return aligns with the empirical range of the drift of the 5-minute returns of SPY data, with the drift estimated via the daily median of 5-minute returns.

Figure 2.3.1. Constant drift and the daily median of returns



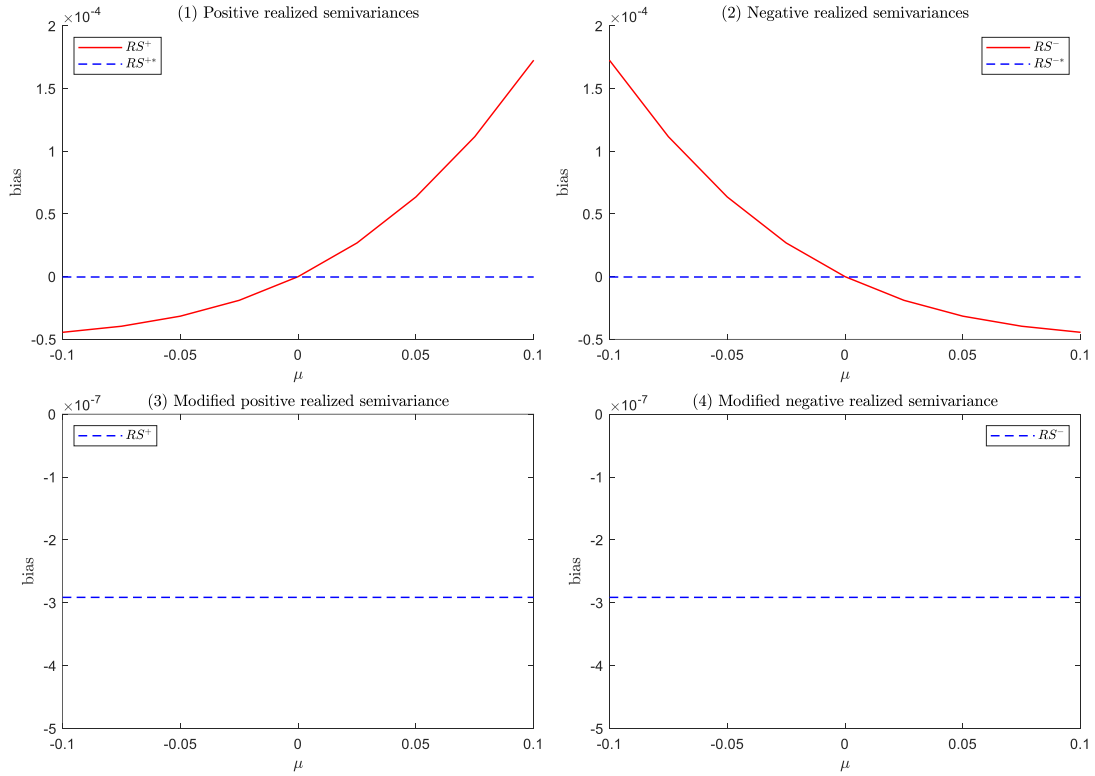
Notes: The left panel is the drift in 5-minute returns as a function of my settings of μ in Equation (2.3.2) The right panel is the empirical drift in 5-minute returns, with the drift estimated by the daily median of 5-minute returns of SPDR S&P 500 ETF (SPY).

Using these realistic settings of the parameters Δ, σ, and μ, the biases in the positive and negative semivariances are calculated by Proposition 2.3.1. With the same settings of parameters, the biases in the modified realized variance and modified realized semivariances are computed via Proposition 2.3.2, with the quantities $\mathbb{v}(\widehat{m}_t)$ and $2cov\left(\frac{1}{M}\sum_1^M r_{t_i}, \widehat{m}_t\right)$ obtained from Monte Carlo Simulations following Laurent and Shi (2020). The simulations of the 5-minute frequency log prices observed for 6.5 hours are obtained by the following procedures. I first generate the 1-second log price based on Equation (2.3.2) for replicating the 6.5h trading session of SPY from 9:30 to 16:30 EST. Then, I obtain the 5-minute log prices by skipping every 300 observations. Finally, a set containing $M = 78$ 5-minute returns is obtained by calculating the difference between the log prices for the day. The expectation $\mathbb{E}$ of the argument is realized by averaging the estimators computed on 10^4 repetitions of Monte Carlo simulation.

Figure 2.3.2 depicts the biases in positive and negative realized semivariances, RS_t^+ and RS_t^-, along with their modified forms RS_t^{+*} and RS_t^{-*}, when these

estimators are computed on 5-minute returns from the discretized constant drift-diffusion model, Equation (2.3.2). The upper panels present the biases in RS_t^+, RS_t^-, RS_t^{+*}, and RS_t^{-*}, and the biases in the two modified semivariance estimators are displayed in the bottom panel for more details. The upper panels visualize the biases in the original semivariances RS_t^+ and RS_t^-, along with their modified versions RS_t^{+*} and RS_t^{-*}.

Figure 2.3.2. The finite sample bias in positive realized semivariance, negative realized semivariance, modified positive realized semivariance, and modified negative realized semivariance under the constant drift-diffusion process



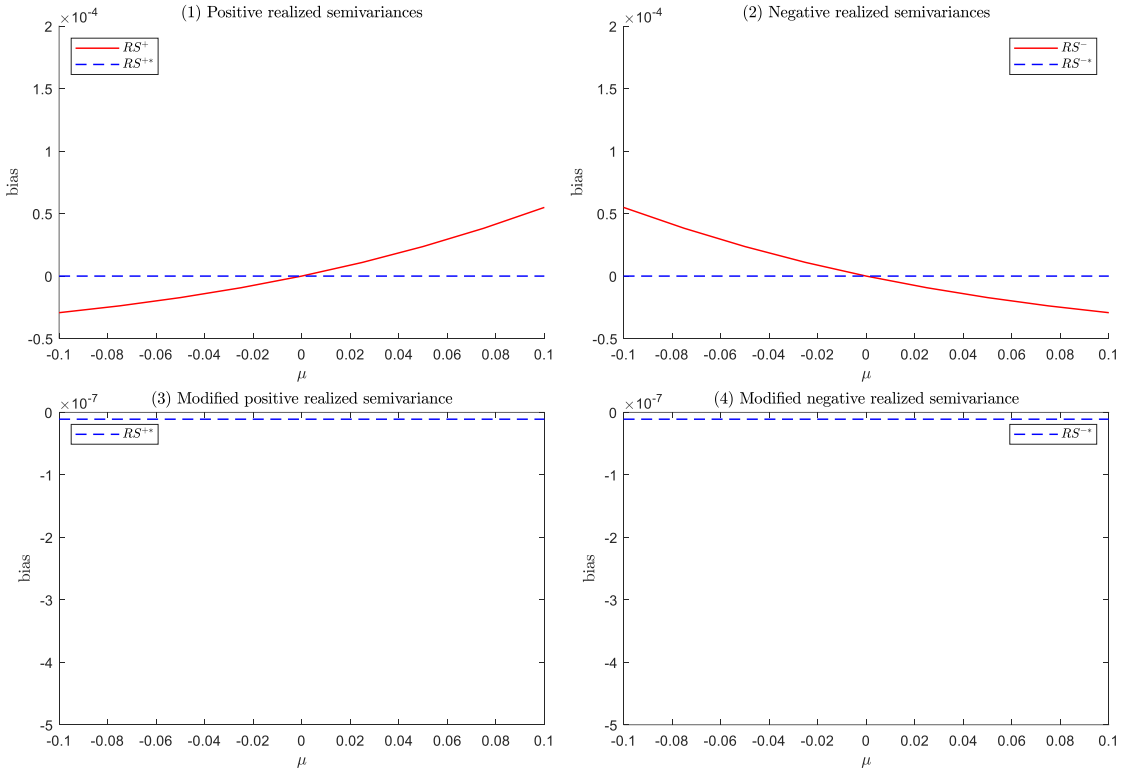
Notes: This figure depicts the finite sample bias in the positive realized semivariance (RS_t^+), negative realized semivariance (RS_t^-), modified positive realized semivariance (RS_t^{+*}), and modified negative realized semivariance (RS_t^{-*}) under the constant drift-diffusion process, Equation (2.3.2). The biases of RS_t^+ and RS_t^- are calculated by Proposition 2.3.1 with parameter settings introduced in the earlier paragraphs of section 2.3.2 while the biases in RS_t^{+*} and RS_t^{-*} are computed by Proposition 2.3.2 with these parameters using simulated log prices from the constant drift-diffusion model, Equation (2.3.2).

As in Corollary 2.3.2, the bias in the original semivariance is positive when the semivariance and drift share the same sign. When the signs of the drift and semivariance differ, the biases in both semivariances are negative. As the drift moves away from zero, the bias of the semivariances becomes larger in size, but this increase is of a sharper slope when the sign of the drift is the same with that of the semivariance. The upper two panels of Figure 2.3.2 also show that the biases in the modified semivariances $RS_t^{+\,*}$ and $RS_t^{-\,*}$, displayed as the dashed lines, are indistinguishable from zero. These biases are again presented in the lower two panels for more details. As the results show, the biases of $RS_t^{+\,*}$ and $RS_t^{-\,*}$ are negative. The order of magnitude is 10^{-7}, which is much smaller than that of the bias of the original semivariances (10^{-4}), suggesting a significantly better estimation performance of the modified semivariances in the presence of a nonzero drift.

The results shown in Figure 2.3.2 are based on the sampling frequency of 5 minutes, with $\Delta = 1/78$. As shown by Propositions 2.3.1 and 2.3.2, the biases of semivariances are also a function of Δ. How do the biases vary with Δ? To explore the relationship between Δ with the biases of the semivariances, Figure 2.3.3 calculates the biases in RS_t^+, RS_t^-, $RS_t^{+\,*}$, and $RS_t^{-\,*}$ for a smaller sampling frequency $\Delta = 1/390$, consistent with the 1-minute resolution for the main trading session 9:30 to 16:00 EST of SPY, with all other parameters in the constant drift-diffusion model unchanged. To facilitate comparison, the Y-axis of the four panels of Figure 2.3.3 is the same as that of the panels of Figure 2.3.2. As the results show, the biases in RS_t^+, RS_t^-, $RS_t^{+\,*}$, and $RS_t^{-\,*}$ for $\Delta = 1/390$ is systematically smaller in size than those for $\Delta = 1/78$ in Figure 2.3.2, respectively, suggesting that the biases in both semivariances and modified semivariances may be positively correlated with the

sampling frequency. Laurent and Shi (2020) also find a positive relationship between the sampling frequency with the biases in realized variance and bipower variation, together with their modified versions.

Figure 2.3.3. The finite sample bias in positive realized semivariance, negative realized semivariance, modified positive realized semivariance, and modified negative realized semivariance under the constant drift-diffusion process for 1-minute frequency $\Delta = 1/390$



Notes: This figure depicts the finite sample bias in the positive realized semivariance (RS_t^+), negative realized semivariance (RS_t^-), modified positive realized semivariance (RS_t^{+*}), and modified negative realized semivariance (RS_t^{-*}) under the constant drift-diffusion process, Equation (2.3.2). The biases of RS_t^+ and RS_t^- are calculated by Proposition 2.3.1 with parameter settings introduced in the earlier paragraphs of section 2.3.2, except that $\Delta = 1/390$, and the biases in RS_t^{+*} and RS_t^{-*} are computed by Proposition 2.3.2 with these parameters using simulated log prices from the constant drift-diffusion model, Equation (2.3.2).

As mentioned, the literature has investigated the impact of a nonzero drift on the estimation performance of other volatility estimators including realized variance and bipower variation, together with their modified versions (Laurent and Shi 2020). To help assess the inaccuracy of semivariances due to nonzero drifts and the effectiveness of the modified semivariances more broadly, I therefore also consider how impactful the bias in these semivariances is compared to that in those realized variance and bipower variations. As discussed previously in section 2.2, the asymptotic theories show that the semivariances converges to the different amounts of volatility from realized variance and bipower variations. For fairness purposes, all comparisons are based on the bias ratio measure, which describes the proportion of the bias in each estimator due to a nonzero drift.

The bias ratio of semivariances, realized variance, and bipower variations under the constant drift-diffusion model, Equation (2.3.2), is described as follows. The bias ratio of the positive and negative semivariances is obtained by Proposition 2.3.1. The bias ratio of the realized variance is computed by Equation (2.3.9) while the bias ratio of the realized semivariances is from Proposition 2.3.1. The bias ratio of the bipower variation is calculated via the following Equation (2.3.16), proposed by Laurent and Shi (2020),

$$\mathbb{E}\left\{\left[BV_t - \int_{t-1}^{t} \sigma_s^2 ds\right] \Big/ \int_{t-1}^{t} \sigma_s^2 ds \right\}$$

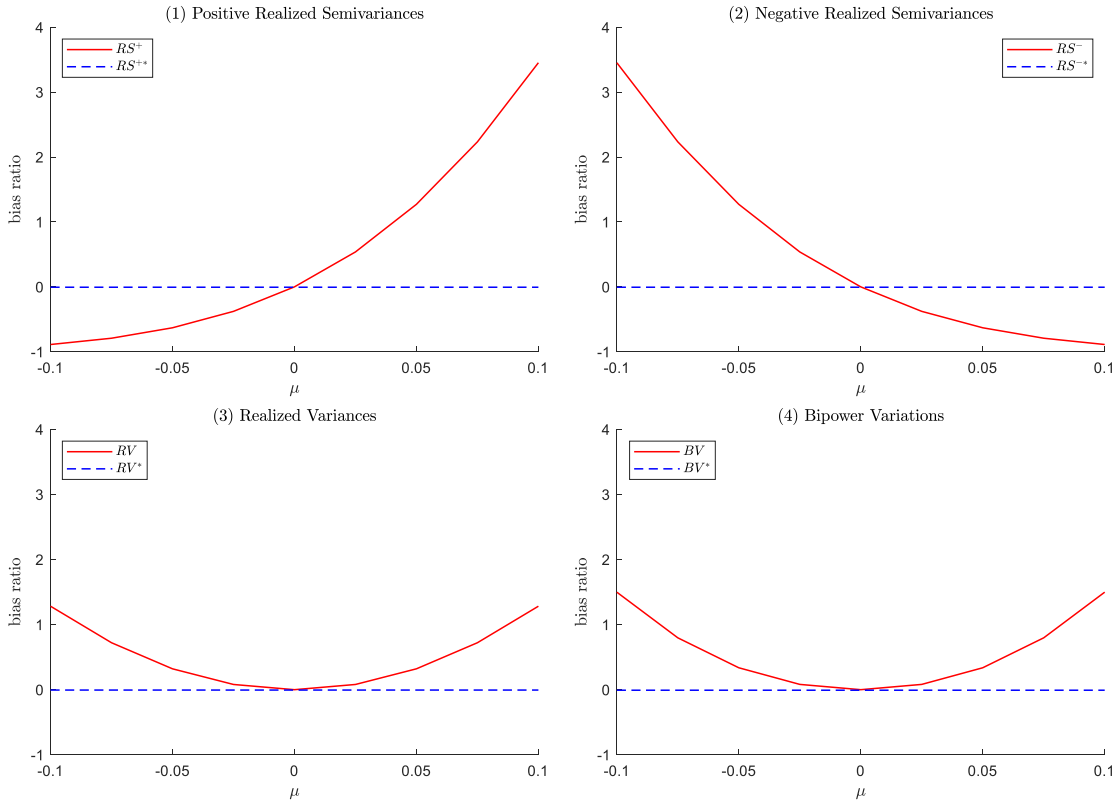
$$= \exp\left(-\frac{\mu^2 \Delta}{\sigma^2}\right) + \frac{\pi}{2}\frac{\mu^2 \Delta}{\sigma^2}\left[1 - 2\Phi\left(-\frac{\mu\sqrt{\Delta}}{\sigma}\right)\right]^2 \tag{2.3.19}$$

$$+ \frac{\mu\sqrt{2\pi}\sqrt{\Delta}}{\sigma}\exp\left(-\frac{\mu^2 \Delta}{2\sigma^2}\right)\left[1 - 2\Phi\left(-\frac{\mu\sqrt{\Delta}}{\sigma}\right)\right] - 1.$$

Since there is no exact formula for the bias ratio of the modified bipower variation (Laurent and Shi 2020), I obtain it from the above Monte Carlo simulations. The bias ratios of the modified realized variance and modified realized semivariances are calculated by Proposition 2.3.2 using simulations. For comparison purposes, all these different volatility estimators are calculated or simulated under the same settings of the parameters including the sampling interval Δ, the level of volatility σ, and the range of drift μ.

The following Figure 2.3.4 depicts the bias ratios of realized semivariances, realized variance, and bipower variation, along with their modified forms for 5-minute returns from the discretized constant drift-diffusion model, Equation (2.3.2). The upper panels visualize the bias ratios of the original semivariances RS^+ and RS^-, and the modified semivariances RS^{+*} and RS^{-*}. The lower panels show the bias ratios of RV and BV, along with their modified versions RV^* and BV^*. The bias ratio of these semivariances equals the bias relative to one-half integrated variation ($0.5\sigma^2 = 0.5 \times 0.01^2$), suggesting the bias ratio is just a scaled version of the bias. Thus, it is not surprising that the bias ratio of semivariances exhibits the same pattern as the bias in semivariances reported in the upper panels of Figure 2.3.1. The bias ratios of RV, BV, RV^* and BV^* estimators are consistent with the results by Laurent and Shi (2020): when drift deviates from zero, both RV and BV are upward biased, and the bias ratio of RV^* and BV^* are much smaller than that of RV and BV, respectively. I also find that the bias ratio of BV due to a nonzero drift appears to be slightly larger than that of the RV, which is similar to Laurent and Shi (2020).

Figure 2.3.4. The finite sample bias ratio of realized semivariances, realized variance, bipower variation, and their modified versions under the constant drift-diffusion process



Notes: This figure depicts the finite sample bias ratio of the realized measures under a constant drift-diffusion process, Equation (2.3.2). The realized measures include positive realized semivariance (RS_t^+), and negative realized semivariance (RS_t^-), realized variance (RV_t,) and bipower variation (BV_t), along with the modified estimators, indicated by RV_t^*, BV_t^*, $RS_t^{+\,*}$, $RS_t^{-\,*}$. The bias ratios of RV_t and BV_t are calculated by Equations (2.3.9) and (2.3.19) while the bias ratios of RS_t^+ and RS_t^- are computed by Proposition 2.3.1. The bias ratios of RV_t^*, $RS_t^{+\,*}$, and $RS_t^{-\,*}$ are obtained by Proposition 2.3.2, using simulating the log prices from the constant drift-diffusion model, Equation (2.3.2). The bias ratio of BV_t^* is calculated on simulated log prices from the constant drift-diffusion model, Equation (2.3.2).

As in Corollary 2.3.2, the bias ratio of the semivariance is greater than zero if the semivariance and drift μ are equal in sign. For example, when μ is positive the bias ratio of the positive semivariance RS^+ is larger than zero, and when μ is negative the bias ratio of the negative semivariance RS_t^- is larger than zero. When μ reaches its maximum ($\mu = 0.1$), the bias ratio of RS_t^+ is larger than those of RV_t and BV_t, and when μ equals its minimum the bias ratio of RS_t^- is larger than those of RV_t and BV_t.

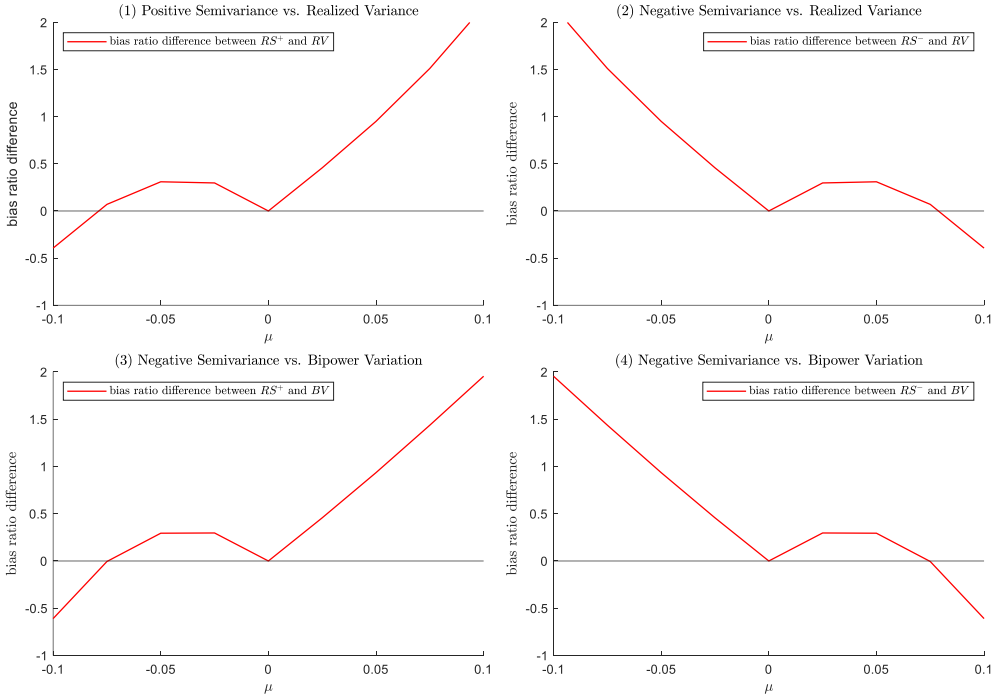
This corroborates my former intuition that when the signs of semivariance and drift are the same and the drift size is large, the semivariance is almost identical to RV_t thus overestimating the one-half integrated variation by a greater ratio than RV_t. Additionally, it can also be observed that when the semivariance and drift differ in signs, the bias of semivariance is negative. The bias ratio of RS_t^+ and RS_t^- approaches -100% when μ reaches its most negative and positive magnitudes, respectively, which supports my previous intuition that when the signs of semivariance and drift are not equal, the semivariance can be close to zero when the drift is large thus underestimating the one-half integrated variation by almost one hundred percent. However, when the bias ratio of RS_t^+ and RS_t^- is getting -100%, the bias ratio of both RV_t and BV_t exceed 100%, suggesting that when the signs of the drift and semivariance are equal and the drift size is large, the biases in realized variance and bipower variation can be more impactful than in semivariances. Comparing the positive semivariance RS_t^+ with the negative counterpart RS_t^-, the biases of these two estimators appear to be symmetric, which is consistent with Corollary 2.3.4.

From the visual observation of Figure 2.3.4, it is not straightforward to rank the impacts of drifts on different estimators for all magnitude of drift considered. To highlight the generality, I compare the bias ratio of the original semivariances (RS_t^+ and RS_t^-) with those of the original realized variance and bipower variation (RV_t and BV_t) across all drift levels considered. The bias ratio of RS_t^+ and RS_t^- can be negative and positive while the bias ratio of RV_t or BV_t are non-negative. To facilitate comparison, I consider the absolute bias ratios of RS_t^+ and RS_t^- and calculate the difference between the absolute bias ratio of that semivariance with the bias ratio of RV_t or BV_t. A positive difference indicates that the drift has a greater impact on the semivariance than RV_t or BV_t.

Figure 2.3.5 presents the difference between semivariances with RV_t or BV_t in terms of the bias ratio. Panel (1) shows the bias ratio difference between RS_t^+ and RV_t, indicated by $\left|\mathbb{E}\left[\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]\right| - \mathbb{E}\left[\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)/\int_{t-1}^{t}\sigma_s^2 ds\right]$. Panel (2) depicts the difference between RS^- and RV, indicated by $\left|\mathbb{E}\left[\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]\right| - \mathbb{E}\left[\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)/\int_{t-1}^{t}\sigma_s^2 ds\right]$. Panel (3) reports the difference in bias ratio size between the RS^+ and BV, denoted by $\left|\mathbb{E}\left[\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]\right| - \mathbb{E}\left[\left(BV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)/\int_{t-1}^{t}\sigma_s^2 ds\right]$. Panel (4) presents the difference in bias ratio size between the RS^- and BV, denoted by $\left|\mathbb{E}\left[\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]\right| - \mathbb{E}\left[\left(BV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)/\int_{t-1}^{t}\sigma_s^2 ds\right]$. When the drift $\mu > 0$, the bias ratio difference between RS^+ and RV is systematically positive, which is consistent with Corollary 2.3.2 that the bias ratio of semivariance exceeds that of RV_t if the sign of the semivariance and drift is positive. When $\mu < 0$, this bias ratio difference is positive except for some extreme negative drifts. The extreme cases of the drift are not common as can be seen in the right panel of Figure 2.3.1, indicating that for the realistic range of negative drift considered, the drift's impact on the positive semivariance RS_t^+ tends to be larger than that of RV_t. For $\mu < 0$, the bias ratio difference between RS_t^- and RV_t is systematically positive, which corroborates Corollary 2.3.3, which states that the drift has a greater impact on negative semivariance than realized variance if the drift is negative. For $\mu > 0$, this difference, for some exceptions in the existence of the extremely positive μ, is overwhelmingly larger than zero, and extremely positive μ is rare as can be observed from the right panel of Figure 2.3.1. This suggests that the bias in the negative semivariance due to

nonzero drift is often more influential than that in realized variance even when the drift is positive.

Figure 2.3.5. Comparison of bias ratios between realized semivariances and realized variance under the constant drift-diffusion model



Notes: Panel (1) of Figure 2.3.3 shows the difference between the absolute bias ratio of RS^+ and the bias ratio of RV, $\left|\mathbb{E}\left[\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]\right| - \mathbb{E}\left[\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)/\int_{t-1}^{t}\sigma_s^2 ds\right]$. Panel (2) depicts the difference between RS^- and RV for the bias ratio size, $\left|\mathbb{E}\left[\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]\right| - \mathbb{E}\left[\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)/\int_{t-1}^{t}\sigma_s^2 ds\right]$. Panel (3) reports the difference in bias ratio size between the RS^+ and BV, $\left|\mathbb{E}\left[\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]\right| - \mathbb{E}\left[\left(BV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)/\int_{t-1}^{t}\sigma_s^2 ds\right]$. Panel (4) presents the difference in bias ratio size between the RS^- and BV $\left|\mathbb{E}\left[\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right]\right| - \mathbb{E}\left[\left(BV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)/\int_{t-1}^{t}\sigma_s^2 ds\right]$.

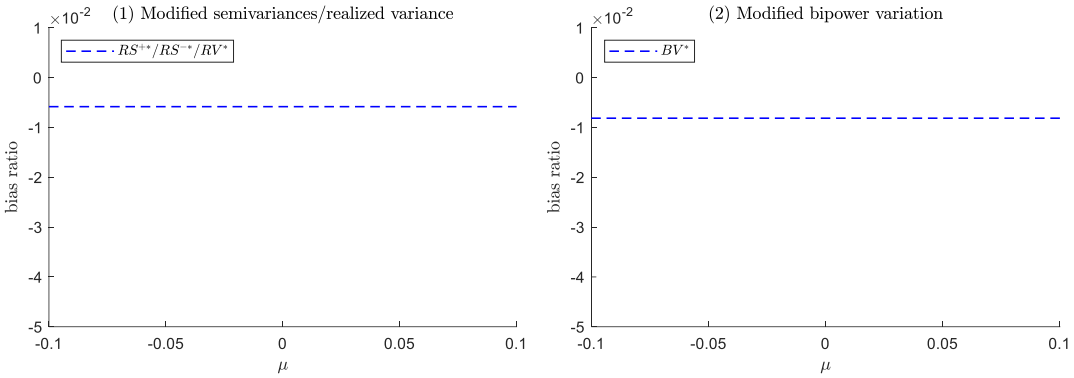
The lower two panels show the bias ratio differences between semivariances and bipower variation (BV_t). The patterns of these bias ratio discrepancies are very similar to those reported in the upper panels. Again, when drift and semivariance have the same sign (e.g. RS_t^- in the presence of $\mu < 0$), the bias ratio difference is systematically positive. If the signs of the drift and the semivariance are not equal

(e.g. RS_t^- with $\mu > 0$), the bias ratio is also positive, except for the extreme values of the drift. These results imply that a nonzero drift generally influences the estimation of semivariances more dramatically than that of bipower variation.

The bias ratios of the modified estimators reported in Figure 2.3.4 are visually indistinguishable from zero, and these measures include the modified realized variance RV_t^*, modified bipower variation BV_t^*, modified positive realized semivariance RS_t^{+*}, and modified negative realized semivariance RS_t^{-*}. To see these biases in more detail, I plot them in the separate Figure 2.3.6 with appropriate scales of the Y-axis. The bias ratios of RV_t^*, RS_t^{+*}, and RS_t^{-*} are equal according to Proposition 2.3.2 and thus are reported together in Panel (1) while the bias of BV_t^* is presented in Panel (2). Consistent with Laurent and Shi (2020), the bias ratios of RV_t^* and BV_t^* are systemically negative and the sizes of both bias ratios are below 0.01 for almost all of these drift levels. The order of magnitude of the bias ratio (1×10^{-2}) of RS_t^{+*} and RS_t^{-*} estimators are much smaller relative to the order of magnitude of the bias ratio of their original versions (1×10^{0}) reported in Figure 2.3.4, indicating that the estimation performance of the modified semivariances is much better than their original versions.

Turning to the results of the modified realized semivariances reported as the blue dashed line in the lower Panels (3) and (4), I observe that the bias ratio of the modified semivariances is much smaller than that of their original versions for all of these nonzero drifts. This indicates that the modified realized semivariance has much better estimation performance.

Figure 2.3.6. The bias ratios of modified semivariances, realized variance, and bipower variation under the constant drift-diffusion process



Notes: This figure reports the bias ratio of the modified realized variance (RV_t^*), modified positive realized semivariance ($RS_t^{+\,*}$), and modified negative realized semivariance ($RS_t^{-\,*}$) are equal and thus are reported together in Panel (1), with the bias of the modified Bipower variation (BV_t^*) separately reported in Panel (2). The bias ratio of $RS_t^{+\,*}$, $RS_t^{-\,*}$, and RV_t^* is calculated by Proposition 2.3.2 while the bias of BV_t^* is obtained by the simulations, with the parameters setting of the calculations and simulations introduced in the earlier paragraphs of section 2.3.2.

2.3.3. The bias of realized semivariances in the presence of jumps

Evidence of stochastic skewness and kurtosis of asset return distributions has led to the development of models with jumps to better capture these dynamics. Jumps are larger events than what can be explained by the classic drift-diffusion process, and jumps are found to be rare but a stylized fact of asset prices in financial markets (Huang and Tauchen 2005; Andersen et al. 2007b; Tauchen and Zhou 2011; Christensen et al. 2014; Bajgrowicz et al. 2016; Kolokolov and Renò 2023). The limiting results of Barndorff-Nielsen et al. (2008) show that the positive and negative semivariance estimators also capture the sum of the squared positive and negative jumps, respectively, in addition to one-half of the integrated variation. In the presence of jumps, is the impact of nonzero drift still a problem for the estimation performance of semivariances? Laurent and Shi (2020) show that the median of intraday returns

could be biased due to the presence of jumps although they show that this bias is small. This suggests that the estimation accuracy of the modified semivariances, which depends on the median estimator, could be influenced. For this case, do the modified semivariances still exhibit more satisfactory estimation performance than the original semivariances for a nonzero drift?

To explore these research questions, I consider the constant drift-diffusion model, Equation (2.3.1) with k additive jumps,

$$p_{t_i} = p_{t_{i-1}} + \mu\Delta + \sigma\sqrt{\Delta}\varepsilon_{t_i} + \sum_{j=1}^{k} \phi_{t_i}^{j} I_{t_i}^{j}, \tag{2.3.20}$$

where $I_{t_i}^{j}$ is a dummy variable that randomly assigns the occurrence of the j^{th} jump with corresponding jump size $\phi_{t_i}^{j}$, and k indicates the number of jumps. Incorporating additive jumps into the drift-diffusion model is first proposed by Merton (1976) and has been widely applied in the recent literature (Huang and Tauchen 2005; Barndorff-Nielsen and Shephard 2006; Corsi et al. 2010; Andersen et al. 2012; Gilder et al. 2014; Laurent and Shi 2020). I consider five scenarios of jumps for Equation (2.3.14), including (1) one positive jump with $\phi_{t_i}^{1} = 0.6\sigma$; (2) one negative jump with $\phi_{t_i}^{1} = -0.6\sigma$; (3) two jumps with $\phi_{t_i}^{j} = 0.6\sigma$ for $j = 1,2$; (4) two jumps with $\phi_{t_i}^{j} = -0.6\sigma$ for $j = 1,2$; (5) two jumps with $\phi_{t_i}^{1} = 0.6\sigma$ and $\phi_{t_i}^{2} = -0.6\sigma$. For these five cases, the size of all jumps is fixed at 0.6 of the constant volatility as in Laurent and Shi (2020). This magnitude of jumps indicates that jumps contribute approximately 26% to the Quadratic Variation on the day of a jump arrival.

To check the robustness of my results to the changes in the jump size, I also allow the jump magnitude to vary, with the results discussed in the latter parts of this

section, and this has not yet been done by Laurent and Shi (2020). Since semivariances capture jumps, the component (one-half integrated variation plus squared positive or negative jumps) that semivariances attempt to estimate varies over these different jump scenarios. For the fairness of comparing the estimation accuracy of semivariances across these different jump cases, I consider the bias ratio measure, which describes the proportion of the bias in semivariances. The bias ratio also allows us to quickly compare the impacts of drift on the estimation performance between original and modified semivariances, which is the primary interest of this chapter. Given various cases of additive jumps discussed, in this section the bias ratios of original and modified realized semivariances (RS_t^+, RS_t^-, RS_t^{+*} and RS_t^{-*}) are calculated by Monte Carlo simulations of Equation (2.3.14), with the same parameters as section 2.3.2.

Table 2.3.1 reports the estimation biases of the positive realized semivariance RS_t^+ and its modified version RS_t^{+*} for the five scenarios of jumps, along with the results for a no jump case attached to the top of the table for comparison. The biases of semivariances in the absence of jumps are copied from those reported in Figure 2.3.3. Within each jump scenario, the drift μ is allowed to change from -0.1 to 0.1. The top panel reports the bias ratios of RS_t^+ in the absence of jumps, which have been depicted in Figure 2.3.3. But additional findings could be obtained from the number of presentations here. The bias ratio of RS_t^+ is larger in degree in the presence of positive drift than that in the presence of equal-sized negative drift. A positive drift and an equal-sized negative drift result in identical drift bias in each 5-minute return, thus shaping the same drift bias in RV_t.

Table 2.3.1. The estimation biases of the positive realized semivariances for different sizes of drifts and across various cases of jumps.

Notes: This table compares the bias ratio between positive realized semivariance RS_t^+ with its modified version RS_t^{+*} for a range of drifts and under five scenarios of jumps. The bias ratios of RS_t^+ and RS_t^{+*} are defined by,

$$\left\{\left[RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds - \sum_{(t-1)\le s\le t}\kappa_s^2 I_{\kappa_s>0}\right] \Big/ \left[\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds + \sum_{(t-1)\le s\le t}\kappa_s^2 I_{\kappa_s>0}\right]\right\}$$

$$\left\{\left[RS_t^{+*} - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds - \sum_{(t-1)\le s\le t}\kappa_s^2 I_{\kappa_s>0}\right] \Big/ \left[\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds + \sum_{(t-1)\le s\le t}\kappa_s^2 I_{\kappa_s>0}\right]\right\},$$

respectively. The biases are calculated on the log prices simulated from the constant drift-diffusion model with additive jumps, Equation (2.3.20).

Drift μ	-0.1	-0.075	-0.05	-0.025	0	0.025	0.05	0.075	0.1
No jump									
RS_t^+	-0.888	-0.791	-0.630	-0.378	0.000	0.538	1.271	2.234	3.452
RS_t^{+*}	-0.006	-0.006	-0.006	-0.006	0.000	-0.006	-0.006	-0.006	-0.006
(1) One jump with $\phi_{t_i}^1 = 0.6\sigma$									
RS_t^+	-0.663	-0.571	-0.440	-0.254	0.006	0.360	0.829	1.432	2.184
RS_t^{+*}	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013
(2) One jump with $\phi_{t_i}^1 = -0.6\sigma$									
RS_t^+	-0.890	-0.795	-0.636	-0.387	-0.015	0.515	1.239	2.188	3.390
RS_t^{+*}	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
(3) Two jumps with $\phi_{t_i}^j = 0.6\sigma$ for $j = 1,2$									
RS_t^+	-0.563	-0.473	-0.354	-0.195	0.017	0.296	0.657	1.112	1.673
RS_t^{+*}	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012
(4) Two jumps with $\phi_{t_i}^j = -0.6\sigma$ for $j = 1,2$									
RS_t^+	-0.892	-0.797	-0.641	-0.395	-0.027	0.496	1.211	2.147	3.334
RS_t^{+*}	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021
(5) Two jumps with $\phi_{t_i}^1 = 0.6\sigma$ and $\phi_{t_i}^2 = -0.6\sigma$									
RS_t^+	-0.667	-0.576	-0.447	-0.264	-0.007	0.343	0.806	1.401	2.143
RS_t^{+*}	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009

But why is it the drift with the same size and with opposite signs causes asymmetric bias in RS_t^+? The intuition behind this asymmetry is mainly due to the finite sample restriction. For the 5-minute frequency finite sample, the overall number of returns of one trading day is $M = 78$. For both positive and negative drift, RV_t is always calculated on all intraday returns. But a positive linear drift makes us pick up more than one-half of the overall number of returns $M/2 = 39$ for calculating RS_t^+

while a negative linear drift lets us use less $M/2$. Even if the positive and negative drift have the same size, the positive drift causes a larger bias in RS_t^+ as it includes more returns for calculating RS_t^+. It is the unbalanced use of the number of intraday returns (due to a nonzero drift) that causes the asymmetric bias in RS_t^+.

Scenario (1) reports the bias ratios of RS_t^+ and RS_t^{+*} in the presence of a positive jump. As Panel (1) shows, RS_t^+ is negatively biased when $\mu < 0$ and is positively biased when $\mu > 0$. Comparing RS_t^+ with RS_t^{+*}, the bias ratio of the latter is systematically much smaller in size for a nonzero drift. It should also be noted that the bias of RS_t^{+*} is -0.013 for all of these drifts, which is a remarkably small ratio and thus suggests the impact of drift on the estimation accuracy of RS_t^{+*} is very minor.

Comparing the scenario (1) of jumps with that of no jump, in the presence of a nonzero drift, the bias ratio of RS_t^+ with no jump is systematically larger in size than that of RS_t^+ with one positive jump, scenario (1). This may be because the presence of a positive jump increases the component that RS_t^+ attempts to measure, thus reducing the proportion of the bias. When drift deviates from zero, the bias ratio of RS_t^{+*} with one positive jump is systematically more negative than that of RS_t^{+*} without jumps, which may be because the median is no longer an unbiased estimator of drift but upward biased by the positive jump (Laurent and Shi 2020). Using an upwardly biased drift estimator indicates that we remove an amount that is larger than the actual drift size from each intraday return when drift is positive, and we reduce an amount that is smaller than the actual drift size when drift is negative. This implies that whenever drift is positive or negative, there will remain a negative drift for every intraday return. This negative drift in the returns might cause additional negative bias in RS_t^{+*} (recall from Figure 2.3.2 that the bias in RS_t^{+*} is already negative when there is no jump for both positive and negative drift). However, this additional negative

drift bias is very small, comparing RS_t^{+*} without jumps and with one positive jump, the bias ratio only becomes slightly more negative, with 0.7% difference.

Scenario (2) assesses the biases of the positive semivariances RS_t^+ in the presence of a negative jump. With one negative jump, both RS_t^+ and RS_t^{+*} attempt to estimate one-half integrated variation only. RS_t^+ again, contains a downward bias if drift is negative, and an upward bias if drift is positive. Compared to RS_t^+, the bias of RS_t^{+*} again has a much lower size across all of these nonzero drifts. Comparing scenario (2) with the scenario of no jumps, the bias ratios of RS_t^+ are very similar, which may be explained by noting that RS_t^+ consistently attempts to estimate one-half integrated variation for both cases. A difference is that the sign of the bias ratio of RS_t^{+*} in scenario (2) tends to be positive while that in the top case is negative.

The reason why RS_t^{+*} in scenario (2) overestimates the one-half integrated variation when this negative jump occurs may be because the median underestimates the drift by negative jumps (Laurent and Shi 2020). Applying a downward biased drift estimator implies that we remove an amount that is smaller than the actual size of drift from each return when the drift is positive, and we over-remove the negative drift in each return. Consequently, there will always be a positive drift left, thus resulting in an expected positive bias in RS_t^{+*}. However, this positive drift bias is minor as the bias ratio of RS_t^{+*} in scenario (2) is only 0.7%, indicating a very high level of the estimation accuracy of RS_t^{+*}. Panels (3), (4) and (5) compare the biases between RS_t^+ and RS_t^{+*} for two positive small jumps, two negative small jumps, and two small jumps with opposite signs, respectively. Similar to the results in above (2) and (3), the bias ratio of RS_t^+ for any cases related to positive jump(s) becomes smaller in size compared to that without jumps, and the bias ratio of RS_t^+ in the presence of two negative jumps are comparable to that without jumps. Still, the bias ratio of RS_t^{+*} is

much smaller than that of RS_t^+ when drift is not zero. Also, the bias of RS_t^{+*} has only an ignorable impact on its estimation accuracy for all drifts considered, although the median overestimates drift in the presence of positive jumps and underestimates drift when negative jumps occur.

Table 2.3.2 evaluates the bias ratios of negative realized semivariances RS_t^- and RS_t^{-*}, which attempts to estimate the one-half integrated variation plus the sum of squared negative jumps. Panels (1) to (5) of this table contain the same five scenarios of jumps as those in Table 2.3.1 above.

Table 2.3.2. The estimation bias ratios of negative realized semivariances for different sizes of drifts across various cases of jumps.

Notes: This table compares the estimation bias ratios of negative realized semivariance RS_t^- with its modified version RS_t^{-*} for the same scenarios of drifts and jumps as in Table 2.3.1. The biases are calculated on the log prices simulated from the constant drift-diffusion model with additive jumps, Equation (2.3.20).

Drift μ	-0.1	-0.075	-0.05	-0.025	0	0.025	0.05	0.075	0.1
(1) One jump with $\phi_{t_i}^1 = 0.6\sigma$									
RS_t^-	3.398	2.195	1.244	0.519	-0.012	-0.385	-0.635	-0.794	-0.890
RS_t^{-*}	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
(2) One jump with $\phi_{t_i}^1 = -0.6\sigma$									
RS_t^-	2.189	1.436	0.832	0.363	0.008	-0.253	-0.439	-0.570	-0.662
RS_t^{-*}	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014
(3) Two jumps with $\phi_{t_i}^j = 0.6\sigma$ for $j = 1,2$									
RS_t^-	3.342	2.153	1.215	0.500	-0.025	-0.393	-0.640	-0.797	-0.891
RS_t^{-*}	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020
(4) Two jumps with $\phi_{t_i}^j = -0.6\sigma$ for $j = 1,2$									
RS_t^-	1.679	1.117	0.661	0.299	0.020	-0.193	-0.352	-0.471	-0.561
RS_t^{-*}	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011
(5) Two jumps with $\phi_{t_i}^1 = 0.6\sigma$ and $\phi_{t_i}^2 = -0.6\sigma$									
RS_t^-	2.149	1.406	0.810	0.346	-0.004	-0.261	-0.445	-0.575	-0.666
RS_t^{-*}	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008

Panel (1) of Table 2.3.2 compares the negative semivariance RS_t^- with its modified version RS^{-*} for their estimation biases in the case of a positive jump. As

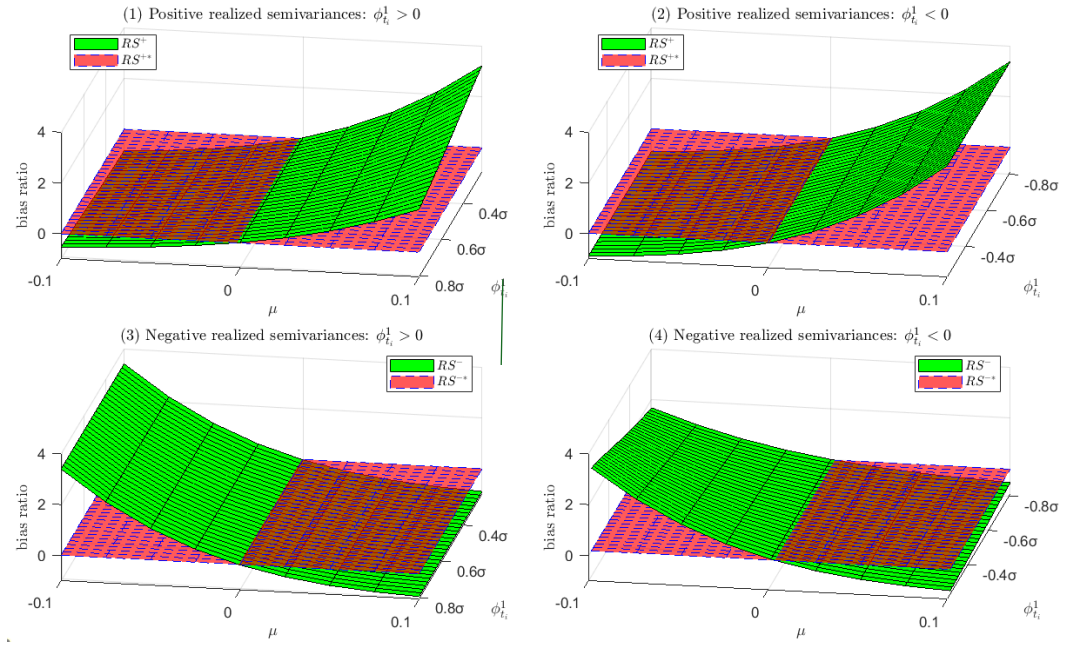
the panel shows, RS_t^- is substantially biased for these nonzero drifts in my simulations, with a positive bias ratio when $\mu < 0$ and a negative bias ratio when $\mu > 0$. In contrast, the modified estimator RS_t^{-*} is systematically much more accurate and does not vary with these different drift values. The bias of RS_t^{-*} is positive as the median is slightly upwardly biased by the positive jump. This may be interpreted as follows. Since the median is upward distorted by the positive jump, the centring modification thus overreduces the actual drift from each intraday return. This leaves a negative drift and causes a positive bias of RS_t^{-*} (recall from Figure 2.3.2 that the bias of the negative semivariance is positive when the drift is negative). However, this bias is small (0.7%) such that it has little influence on the estimation performance of RS_t^{-*}. For the remaining four jump cases reported in Panels (2)-(5), the bias ratio of RS_t^{-*} is always much lower compared to RS_t^- in the presence of the nonzero drifts in my simulations. While the median is no longer an unbiased estimator of the drift due to the presence of jumps, the size of the bias of RS_t^{-*} is consistently very small thus the bias of the median due to nonzero jumps only has a little effect on the estimation precision of RS_t^{-*}.

The results in Tables 2.3.1 and 2.3.2 for comparing the bias ratios of the original realized semivariance estimators with their modified versions rely on one certain size of jumps, which equals 0.6 of volatility. To assess the robustness of the results to a broader range of jump sizes, I calculate the bias ratios of realized semivariances on taking one positive or negative jump in Equation (2.3.20), with its size $\left|\phi_{t_i}^1\right|$ changing from $\sqrt{1/19}\sigma \approx 0.2294\sigma$ to $\sqrt{2/3}\sigma \approx 0.8165\sigma$ with an increment of 0.02σ. The minimum and maximum values in this range correspond to jumps contributing approximately 5% (very small jumps) and 40% (very large jumps) to the Quadratic Variation on the day of a jump arrival. Allowing one positive or negative jump for one

day is motivated by the findings from recent studies that jumps are not common (Christensen et al. 2014; Bajgrowicz et al. 2016; Li et al. 2022; Kolokolov and Renò 2023). For each jump size, I allow drift μ to vary from -0.1 to 0.1. The remaining parameters in Equation (2.3.20) are those used for computing the results in Tables 2.3.1 and 2.3.2.

Figure 2.3.7 reports the bias ratio of the realized semivariances and their modified versions for various combinations of drift μ and jump $\phi_{t_i}^1$.

Figure 2.3.7. The bias ratio of realized semivariances for various combinations of drift and jump



Notes: Panels (1) and (2) present the bias ratios of positive semivariances (RS_t^+ and RS_t^{+*}) for various combinations of the drift μ and a jump with its size and sign indicated by $\phi_{t_i}^1$. Panels (3) and (4) present the bias ratios of negative semivariances (RS_t^- and RS_t^{-*}) for various combinations of the drift μ and a jump with its size equal to $\phi_{t_i}^1$. The biases are calculated on the log prices simulated by the constant drift-diffusion model, Equation (2.3.20).

The upper panels compare the bias ratio of the positive semivariance RS_t^+ with that of its modified version RS_t^{+*}, with Panel (1) for the case of a positive jump and

Panel (2) for the case of a negative jump. Panels (3) and (4) report the comparative result between the negative semivariance RS_t^- and its modified form RS_t^{-*} for a positive jump and a negative jump, respectively. The upper two panels show that across all of these $\phi_{t_i}^1 > 0$ and $\phi_{t_i}^1 < 0$, the RS_t^+ is upward biased for $\mu > 0$ and is downward for $\mu < 0$. In these two panels, throughout the different sizes of the jump $\phi_{t_i}^1$, the bias ratio of RS_t^+ increases in size nonlinearly as μ deviates from zero and is consistently substantial, especially when μ is large and positive. The bias ratio of RS_t^+ increase with μ in a deeper slope when $\mu > 0$ than when $\mu < 0$.

It can be observed from Panel (1) that the bias proportion of RS_t^+ for these nonzero μ systematically decreases in magnitude as $\phi_{t_i}^1$ increases. This can be explained by noting that RS_t^+ estimates the one-half integrated variation plus the sum of all squared positive jumps, which increases with the size of the positive jumps. The proportion of bias in RS^+ is thus smaller. Meanwhile in Panel (2), the bias of RS_t^+ is not sensitive to the changes of the negative $\phi_{t_i}^1$ (e.g., the line of $\mu = 0.1$ is parallel to the $\phi_{t_i}^1$ axis). This is because RS_t^+ asymptotically converges to the half volatility plus the sum of squared positive jumps, which is not related to negative jumps. Comparing RS_t^+ with the modified estimator RS_t^{+*} in Panels (1) and (2), the bias of RS_t^{+*} due to the existence of a nonzero μ is much smaller for all of these sizes of the positive jump.

Panels (3) and (4) of Figure 2.3.6 compare the negative realized semivariance RS_t^- with its modified form RS_t^{-*} for the same ranges of a drift and a jump (μ and $\phi_{t_i}^1$) as those reported in Panels (1) and (2), respectively. As can be seen from Panels (3) and (4), the sign of the bias of RS_t^- is opposite to that of μ for all of these $\phi_{t_i}^1$, and the size of the bias of RS_t^- is large, especially if μ is far below zero. The lower two

panels show that the bias ratio of RS_t^- does not change with $\phi_{t_i}^1 > 0$ but declines in size as $\phi_{t_i}^1$ moves from -0.2294σ to -0.8165σ. This can be explained by noting that the component of RS_t^- attempt to estimate is not affected by positive jumps but grows with the magnitude of negative jumps. Comparing the RS_t^- with RS_t^{-*} in Panels (3) and (4), the bias of the latter estimator for nonzero μ is systematically lower for all these $\phi_{t_i}^1$. Overall, based on the findings of Panels (1)-(4) of Figure 2.3.6, it can be concluded that the estimation superiority of my modified semivariances over the original semivariances in the presence of nonzero drifts and jumps is robust for a broad range of the jump sizes, from 5% of Quadratic Variation (very small) to 40% of Quadratic Variation (very large).

2.4. Realized Semivariance estimation for the Ornstein Uhlenbeck processes

The constant drift-diffusion model introduced in section 2.3 is a simplified model of returns for facilitating exposition. This model is a special case of the Ornstein Uhlenbeck process and follows the same model used by Laurent and Shi (2020). To highlight the generality of results of the bias in realized semivariances, this section investigates the bias of the semivariances for the log prices that follow the Ornstein Uhlenbeck process, which contains more realistic assumptions on the drift and volatility processes. I start the analysis in section 2.4.1 without considering a jump component in the Ornstein Uhlenbeck model. Then I introduce the estimation analyses including jump cases in the subsequent section 2.4.2.

The linear drift-diffusion process is defined by,

$$dp_t = \mu_t dt + \sigma_t \mathrm{d}W_t, 0 \leq t \leq 1, \tag{2.4.1}$$

where the drift coefficient is a linear function of the log prices $\mu_t = \theta p_t$ with θ constant, thus the drift is time-varying. The diffusion coefficient σ_t follows the GARCH(1,1) process of Nelson (Nelson 1991), and W_t is a standard Brownian motion. Since $\mu_t = \theta p_t$, the drift is not zero on day t if both θ and the initial value of the log price of day t denoted by p_{t_0} deviate from zero. Since drift is a linear function of the log prices, the Ornstein-Uhlenbeck process can be seen as a linear drift-diffusion model. The Ornstein-Uhlenbeck process has also been applied extensively for volatility estimation and option pricing in the literature (Barndorff-Nielsen and Shephard 2001; Wang and Yu 2016; Laurent and Shi 2020; Laurent et al. 2022).

2.4.1. The bias of realized semivariances

Laurent and Shi (2020) show that under the Ornstein Uhlenbeck model, the non-zero linear drift in returns has a similar effect as the constant drift-diffusion model on the estimation of both realized variance (RV_t) and bipower variation (BV_t) estimators: both RV_t and BV_t are upward biased with if θ and p_{t_0} deviates from zero for the finite sample data. To mitigate these biases, Laurent and Shi (2020) again suggest using the modified realized variance (RV_t^*) and modified bipower variation (BV_t^*) defined by Equations (2.3.8) and (2.3.9), which are computed on the intraday returns centred by their daily median, $\widehat{m}_t = \text{median}(r_{t_1} \dots r_{t_M})$. They find that for the Ornstein-Uhlenbeck process, using the de-median technique reduces the average drift over M returns of day t, and this average drift is defined by,

$$\bar{\mu}_t = \frac{1}{M} p_{t_0}(e^{\theta} - 1), \tag{2.4.2}$$

where $\bar{\mu}_t$ indicates that the average linear drift over day t, and how $\bar{\mu}_t$ changes with θ and p_{t_0}, which will be discussed later. The theoretical and simulation results of Laurent and Shi (2020) show that RV_t^* and BV_t^* are much more immune to the changes in θ and p_{t_0} than their original versions.

Does a nonzero linear drift have a similar effect to the nonzero constant drift on the estimation of realized semivariances? If the bias in realized semivariances for the Ornstein-Uhlenbeck process is large, will the modified semivariances have a reduced magnitude of bias? This section investigates the impact of a nonzero linear drift on the measurement precision of positive and negative semivariances and their modified versions, denoted by RS_t^+, RS_t^-, RS_t^{+*}, and RS_t^{-*}. The exact derivation of the finite

sample biases of these semivariances under the linear drift-diffusion process, however, is rather complicated and left for future research. This section computes these biases via a Monte Carlo simulation study. There is a precedent for this approach in the literature. The Monte Carlo approach without technical derivation follows Laurent and Shi (2020) to gauge the bias of bipower variations under the Ornstein-Uhlenbeck process since the derivation of such bias is also not available due to its complexity.

Additionally, I also replicate the bias of RV_t, RV_t^*, BV_t, and BV_t^* in Laurent and Shi (2020). Since the derivation of the biases of BV_t and BV_t^* are unknown, I also calculate these biases via simulations. Although the biases of RV_t and RV_t^* are derived analytically in Laurent and Shi (2020), I also compute these biases by simulations for fair comparison. Furthermore, the simulation results in the chapter are not intended to be comprehensive, but rather to reflect a realistic application of the linear drift-diffusion models.

As proposed by Laurent and Shi (2020), the simulation of the Ornstein-Uhlenbeck process can be conducted by a discrete version of Equation (2.4.1),

$$p_{t_i} = \exp(\theta\Delta_M)p_{t_{i-1}} + \eta_{t_i}, \tag{2.4.3}$$

where p_{t_i} is the log price and η_{t_i} is the diffusive volatility process, which is obtained by the Euler discretization of the continuous $\mathrm{GARCH}(1,1)$ process of Nelson (1990),

$$\eta_{t_i} = \sigma_{t_i}\sqrt{\Delta_M}\epsilon_{t_i}, \tag{2.4.4}$$

where

$$\sigma_{t_i}^2 = \kappa\left(\omega - \sigma_{t_{i-1}}^2\right)\Delta_M + \sqrt{2\lambda\kappa}\sigma_{t_{i-1}}^2\sqrt{\Delta_M}v_{t_i},$$

with $\kappa > 0$, $\omega > 0$, and $0 < \lambda < 1$, and ϵ and v are two independent standard normal random variables. As shown by Andersen and Bollerslev (1998), the unconditional variance of this discretized GARCH(1,1) model is $\mathbb{E}(\sigma_t^2) = \omega$. Then the expected integrated variance or continuous variation from $t-1$ to t is

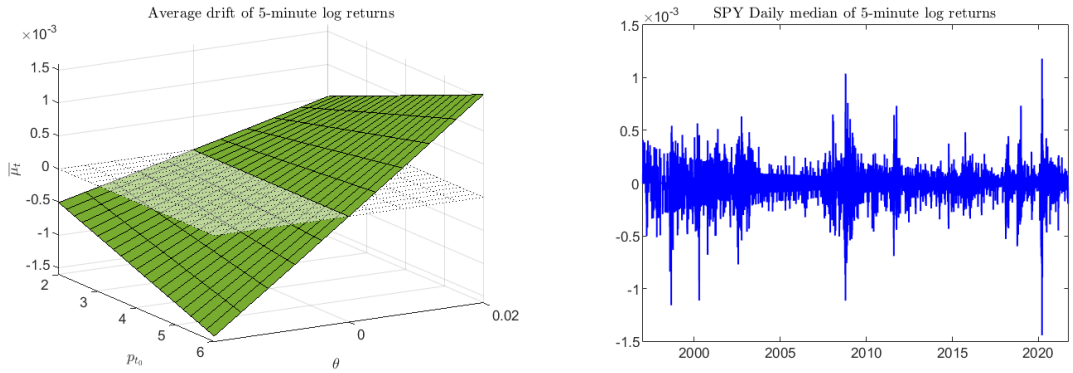
$$\mathbb{E}\left(\int_{t-1}^{t} \sigma_s^2 \, ds\right) = \int_{t-1}^{t} \mathbb{E}(\sigma_s^2) \, ds = \omega. \tag{2.4.5}$$

This expected integrated variance will be used for evaluating the estimation accuracy of all realized measures in the simulation analysis, including RS_t^+, RS_t^-, RV_t, BV_t, and their modified versions, $RS_t^{+\,*}$, $RS_t^{-\,*}$, RV_t^*, and BV_t^*.

The parameter settings in Equation (2.4.4) are the same as in Andersen and Bollerslev (1998) and Laurent and Shi (2020): $\kappa = 0.035$, $\lambda = 0.296$, $\omega = 10^{-4}$, the initial volatility for the GARCH(1,1) process is set as the unconditional volatility of this process ($\sqrt{\omega}$), and W and B are independent standard Brownian motions. I allow θ to vary from -0.02 to 0.02 and the initial log price p_{t_0} to change from 2 to 6. For checking if these ranges of p_{t_0} and θ are realistic, the left panel of Figure 2.4.1 depicts the average drift calculated on these p_{t_0} and θ based on Equation (2.4.2) under 5-minute sampling frequency ($M = 78$) while the right panel reports the empirical average drift (estimated by the daily median) of the 5-minute returns of the real SPY data. Comparing these two panels, the average drifts decided by the combinations of p_{t_0} and θ defined in my simulation are consistent with the empirical average drift observed from the SPY data. This indicates that my simulation settings of p_{t_0} and θ are practical and realistic for the US stock market. Additionally, similar to the simulation of the constant drift-diffusion model in section 2.3, I first simulate 23,401 (6.5 hours duration) 1-second log prices from Equation (2.4.3), then obtain the

5-minute log prices by skipping every 300 observations of the 1-second log prices. The simulations are repeated 10^4 times and the expectation ($\mathbb{E}$) is computed by averaging across these repetitions.

Figure 2.4.1. Comparing average drift in the simulation with the daily median of the return of SPY

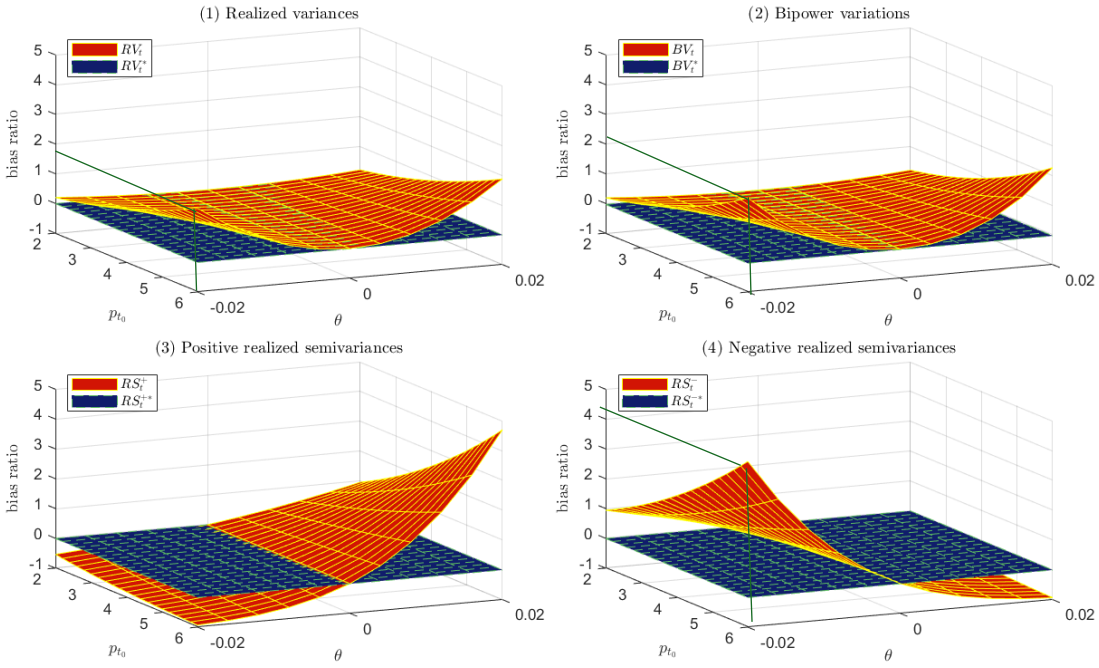


Notes: The left panel shows the range of the average drift $\bar{\mu}_t$ of 5-minute returns over one day under my simulation settings of θ and p_{t_0}. This average drift $\bar{\mu}_t$ is calculated by Equation (2.4.2). The right panel shows the daily median of the 5-minute returns on the SPDR S&P 500 ETF (SPY) from 1997 to 2021, which estimates a realistic range of the average drift of the 5-minute returns over a day.

Figure 2.4.2 compares the finite sample bias ratio of realized variance, bipower variation, positive semivariance, and negative semivariance RV_t, BV_t, RS_t^+, and RS_t^-, along with their modified versions, RV_t^*, BV_t^*, $RS_t^{+\,*}$, and $RS_t^{-\,*}$, under various combinations of parameter θ and initial log price p_{t_0}. Panel (1) depicts the comparative result for RV_t and RV_t^*, Panel (2) compares BV_t with BV_t^*, and Panels (3) and (4) contain the bias ratios of and RS_t^+, RS_t^-, $RS_t^{+\,*}$ and $RS_t^{-\,*}$. The figures for the bias of the realized variance and bipower variation measures exhibit a consistent pattern with Laurent and Shi (2020): the biases of the original estimators RV_t and BV_t increase with $|\theta|$ and p_{t_0}, and both biases are symmetric around $\theta = 0$. The biases of

RV_t and BV_t can be influential, especially if θ and p_{t_0} deviates far from zero. For example, BV_t overestimates the volatility by 200% for $\theta = 0.02$ and $p_{t_0} = 6$. Compared to RV_t and BV_t, the biases of the modified volatility estimators RV_t^* and BV_t^* are not distinguishable from zero for all combinations of nonzero θ and p_{t_0}. The much better estimation performance of RV_t^* and BV_t^* relative to RV_t and BV_t is also consistent with the findings of Laurent and Shi (2020).

Figure 2.4.2. The finite sample bias ratio of realized semivariances, realized variance, bipower variation, and their modified versions under the Ornstein-Uhlenbeck process



Notes: Each panel of this figure compares the bias ratios of a realized variation measure with its modified version for various combinations of the initial log prices p_{t_0} and parameter θ. The realized measures include realized variance (RV_t), Bipower variation (BV_t), positive realized semivariance (RS_t^+), and negative realized semivariance (RS_t^-) while their modified counterparts are indicated by RV_t^*, BV_t^*, RS_t^{+*}, RS_t^{-*}. The biases of all of these eight realized measures in this figure are calculated on the log prices simulated from the linear drift-diffusion model, Equation (2.4.3), with the parameter settings introduced in section 2.4.1.

Panel (3) contains the bias ratios of positive realized semivariance RS_t^+ and its respective modified measure RS_t^{+*}. As the panel shows, RS_t^+ is upward biased for

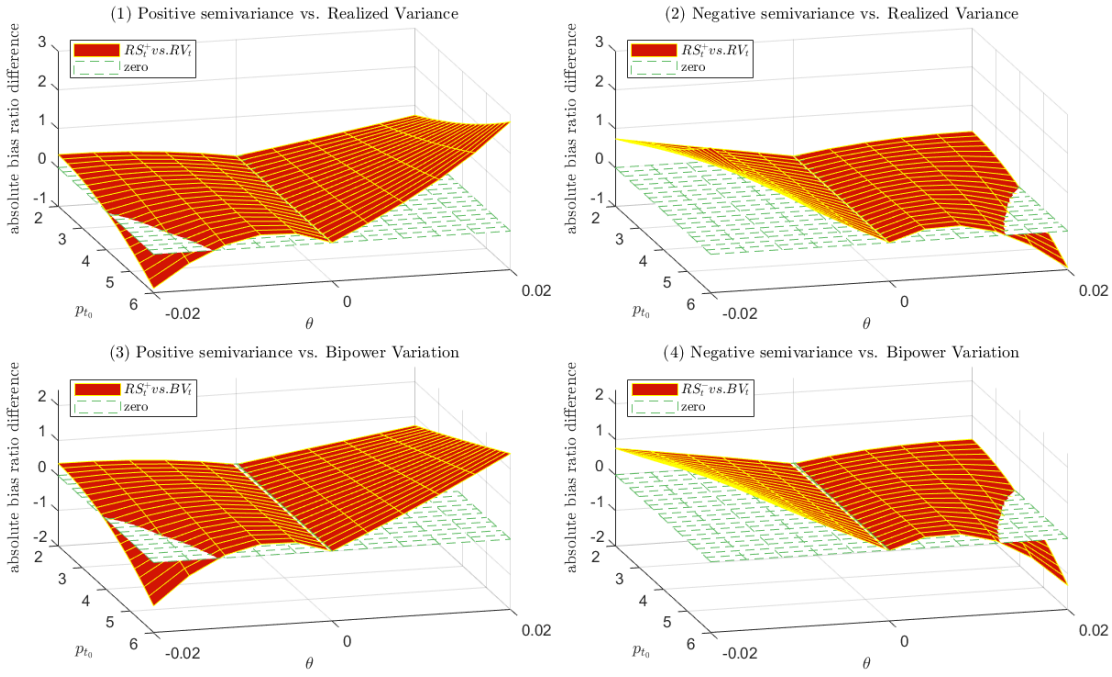
$\theta > 0$ and is downward biased for $\theta < 0$, across p_{t_0}. Recall that the magnitude of the (average) linear drift over day t is positively related to both $|\theta|$ and p_{t_0} and the sign of the linear drift is consistent with the sign of θ. Therefore, RS^+ is upward biased for a positive linear drift and is downward biased for a negative linear drift, suggesting that the effect of a nonzero linear drift on the estimation of RS_t^+ is similar to that of a nonzero constant drift. Given the consistency inherent in the price dynamic between the constant and linear drift models (Laurent et al. 2022b), it is hardly surprising that the constant and linear drift have a similar impact on semivariances from an estimation perspective. The size of the bias ratio of RS_t^+ can be very large. For example, when $\theta = 0.02$ and $p_{t_0} = 6$, the bias ratio is about 4.5, indicating that RS_t^+ overestimates the one-half integrated variation by around 450%. In contrast to the imprecision of RS_t^+ in the presence of a nonzero drift, the bias ratio of RS_t^{+*} appears much smaller for all of these p_{t_0} and nonzero θ. This implies that RS_t^{+*} performs much better in estimating the one-half integrated variation.

Panel 4 exhibits the biases of negative realized semivariance RS_t^-. The sign of the bias ratio of RS_t^- and that of θ are always opposite for all of p_{t_0}, suggesting that a positive linear drift leads to negative bias in RS_t^- and a negative linear drift results in positive bias in RS^-. This is consistent with the effect of a nonzero constant drift on the estimation bias of RS_t^-. Similar to RS_t^+, the bias ratio of RS_t^- can also be very influential, particularly when θ reaches its most negative value. While the bias ratio of RS_t^- is large, the bias ratio of RS_t^{-*} is much lower for all of these nonzero θ and p_{t_0}.

For a moderate linear drift (moderate values of θ and p_{t_0}), it is not straightforward to compare the degree of the bias ratio of semivariances with that of realized variance and bipower variation from Figure 2.4.2. To illustrate further the

bias ratio of realized semivariances RS_t^+ and RS_t^-, Figure 2.4.3 compares the bias ratios of each of RS_t^+ and RS_t^- with that of realized variance RV_t and bipower variation BV_t under various combinations of parameter θ and initial log price p_{t_0}. Panel (1) compares the bias ratio of RS_t^+ with that of RV_t based on the difference between the bias ratio size of RS_t^+ with that of RV_t, while Panel (2) compares the bias ratio size between RS_t^- and RV_t. Panels (3) and (4) evaluate the size of the bias ratios of RS_t^+ and RS_t^- relative to BV_t, based on the bias ratio difference.

Figure 2.4.3. Comparison of the magnitude of the bias ratio between the realized semivariances and realized variance and the bipower variation under Ornstein-Uhlenbeck process

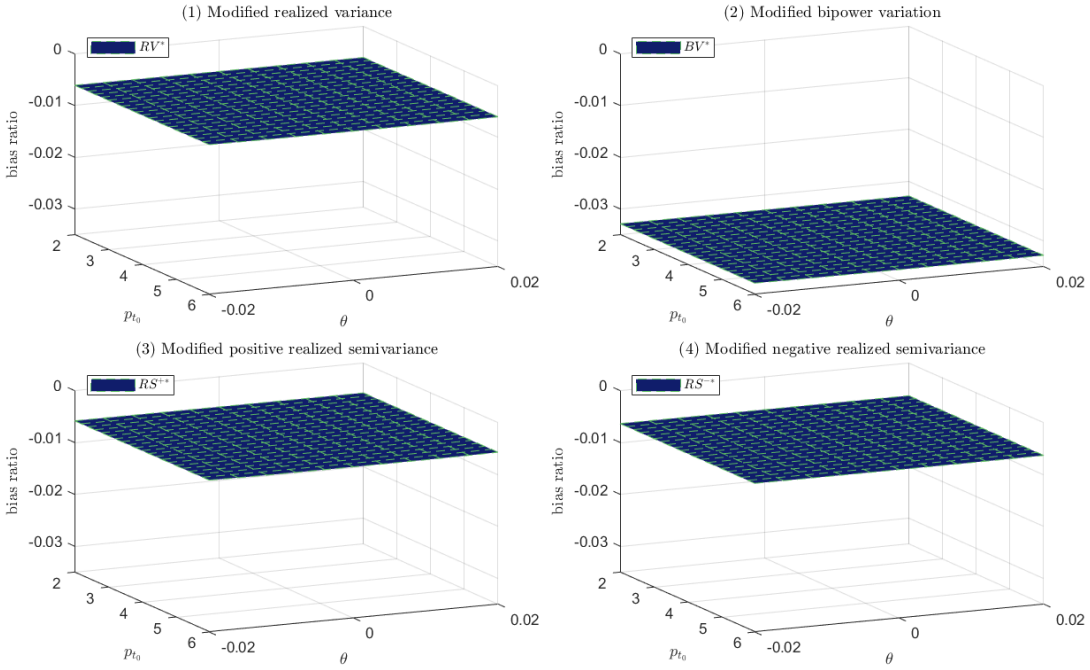


Notes: Panel (1) reports the difference between the bias ratio magnitude of the positive realized semivariance (RS_t^+) with that of realized variance for various combinations of parameter θ and initial log price P_{t_0}. Panel (2) reports the difference between the bias ratio magnitude of the negative realized semivariance (RS_t^-) with that of realized variance (RV_t) Panel (3) and Panel (4) contain the difference between the bias ratio size of RS_t^+ and RS_t^- with that of bipower variation (BV_t), respectively. The magnitude of the bias ratios of RV_t, BV_t, RS_t^+, and RS_t^- are calculated on the simulated log prices from the linear drift-diffusion model, Equation (2.4.3), with the settings of the parameters introduced in section 2.4.1.

I observe that the discrepancy of the bias ratio between RS_t^+ and RV_t and that between RS_t^- and RV_t are positive for these p_{t_0} and nonzero θ, with only exceptions due to some extreme values of $|\theta|$, which are rare as observed from Figure 2.4.1. This implies that the biases of RS_t^+ and RS_t^- are generally larger in magnitude than that of RV_t, which does not alter my conclusions from the constant drift-diffusion model. Again, this consistency may be due to the similarity between the constant and linear drift-diffusion model. The results exhibit very similar patterns as Panels (1) and (2): the two differences in the discrepancy of the bias ratio between RS_t^+ and RV_t and that between RS_t^- and RV_t are larger than zero for these different combinations of nonzero θ and p_{t_0}, notwithstanding some very large $|\theta|$ and p_{t_0}. This indicates that the bias ratios of RS_t^+ and RS_t^- also tend to be larger in size than that of BV_t.

The bias ratios of RV_t^*, BV_t^*, $RS_t^{+\,*}$ and $RS_t^{-\,*}$ are indisguisable from zero reported in Figure 2.4.2. To see these bias ratios in more detail, Figure 2.4.4 depicts the bias ratios of RV_t^*, BV_t^*, $RS_t^{+\,*}$ and $RS_t^{-\,*}$ in Panels (1), (2), (3), and (4), respectively. The biases of RV_t^*, $RS_t^{+\,*}$, and $RS_t^{-\,*}$ are all around -0.6%, which is much smaller in magnitude than that of BV_t^* at -3.3%, indicating that the impact of a nonzero drift considered on the estimation of modified positive and modified negative semivariance is as small as that on the estimation of realized variance and bipower variation. Comparing Figure 2.4.4 with Figure 2.4.2, the order of the magnitude of the bias ratios (10^{-3}) of $RS_t^{+\,*}$ and $RS_t^{-\,*}$ is much smaller than that (10^{-0}) of their original counterparts RS_t^+ and RS_t^-, indicating a far more satisfactory estimation of the one-half integrated variation when the drift is not zero.

Figure 2.4.4. Bias ratio of modified realized semivariances, modified realized variance, and modified bipower variation under the Ornstein-Uhlenbeck process



Notes: This figure reports the bias ratios of modified realized variance (RV_t^*), modified bipower variation (BV_t^*), the modified positive semivariance (RS_t^{+*}), and modified negative semivariance (RS_t^{-*}), for various combinations of the initial log prices p_{t_0} and parameter θ. The bias ratios of all of these four modified estimators in this figure are calculated on the log prices simulated from the discrete form of the Ornstein-Uhlenbeck process, Equation (2.4.3), with the parameter settings introduced in section 2.4.1.

2.4.2. The bias of realized semivariances in the presence of jumps

As discussed in section 2.3.3, jumps are rare but an important stylized fact of asset prices in financial markets, therefore this section studies the impact of drift on the estimation of semivariances by augmenting the Ornstein-Uhlenbeck process to allow additive jumps. Recall that the positive and negative semivariance estimators

converge to the sum of the squared positive and negative jumps, respectively, together with the one-half integrated variation. This indicates that the component that the semivariances attempt to estimate when jumps occur is larger than that when jumps are absent. Therefore, the proportion of the drift-driven bias in the semivariances may become smaller on the arrival of jumps. [4] In other words, large jumps may imply less impact of a nonzero drift on the estimation of the semivariances. Additionally, the occurrence of jumps may influence the estimation bias of modified semivariances since the daily median of returns which estimate drift could be no longer unbiased due to jumps. How influential is the bias of the median to the measurement performance of modified semivariances? Do the modified semivariances still exhibit much less estimation bias than their original counterparts?

The Ornstein-Uhlenbeck process with k additive jumps is defined by,

$$p_{t_i} = \exp(\theta\Delta_M)p_{t_{i-1}} + \eta_{t_i} + \sum_{j=1}^{k} \phi_{t_i}^{j} I_{t_i}^{j}, \tag{2.4.6}$$

with the jump parameters $I_{t_i}^{j}$ and $\phi_{t_i}^{j}$ defined as in section 2.3.3. Semivariances may estimate different sizes of components across various cases of jumps. To facilitate comparison, I again consider the bias ratio measure to explore the estimation performance of semivariances. The bias ratios of the original and modified semivariances, indicated by RS_t^+, RS_t^-, RS_t^{+*}, and RS_t^{-*}, are calculated on the 10^4 repeated simulations of the 5-minute log prices from Equation (2.4.6), with the parameter settings and simulation procedures the same as those introduced in section 2.4.1.

[4] I assume drift is independent of jumps thus the drift bias itself is hardly influenced by jumps, but whether there is a relationship between jumps and drift is interesting and left for future research.

Table 2.4.1 compares the estimation biases of positive semivariance (RS_t^+), with its modified version (RS_t^{+*}) for the same five scenarios of jumps, together with the cas scenario of no jump, as those in Table 2.3.1. The initial log price for the results in Table 2.4.1 is set as $p_{t_0} = 6$. The results are qualitatively very consistent for the initial log prices equal to $p_{t_0} = 2$ or 4.

Table 2.4.1. The estimation bias of positive realized semivariance for different sizes of parameter θ and across various cases of jumps.
Notes: This table compares the estimation biases of positive Realized Semivariance, RS^+, with its modified version, RS^{+*}, for a range of drifts and under five scenarios of jumps. The biases are calculated on the log prices simulated from the linear drift-diffusion model with additive jumps, Equation (2.4.6).

θ	-0.2	-0.16	-0.12	-0.08	-0.04	0	0.04	0.08	0.12	0.16	0.2
No jump											
RS^+	-0.932	-0.873	-0.773	-0.613	-0.366	0.000	0.510	1.206	2.117	3.273	4.696
RS^{+*}	-0.006	-0.006	-0.006	-0.006	-0.006	0.000	-0.006	-0.006	-0.006	-0.006	-0.006
(1) One jump with $\phi_{t_i}^1 = 0.6\sigma$											
RS^+	-0.713	-0.647	-0.555	-0.425	-0.244	0.006	0.344	0.790	1.362	2.077	2.949
RS^{+*}	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013
(2) One jump with $\phi_{t_i}^1 = -0.6\sigma$											
RS^+	-0.933	-0.874	-0.776	-0.617	-0.374	-0.015	0.491	1.177	2.076	3.215	4.620
RS^{+*}	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
(3) Two jumps with $\phi_{t_i}^j = 0.6\sigma$ for $j = 1,2$											
RS^+	-0.616	-0.547	-0.458	-0.342	-0.187	0.016	0.283	0.627	1.060	1.594	2.239
RS^{+*}	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012
(4) Two jumps with $\phi_{t_i}^j = -0.6\sigma$ for $j = 1,2$											
RS^+	-0.933	-0.876	-0.779	-0.622	-0.381	-0.028	0.471	1.149	2.035	3.159	4.544
RS^{+*}	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022
(5) Two jumps with $\phi_{t_i}^1 = 0.6\sigma$ and $\phi_{t_i}^2 = -0.6\sigma$											
RS^+	-0.716	-0.651	-0.560	-0.432	-0.254	0.000	0.327	0.768	1.332	2.037	2.897
RS^{+*}	-0.007	-0.007	-0.007	-0.007	-0.007	0.000	-0.007	-0.007	-0.007	-0.007	-0.007

The top panel in Table 2.4.1 reports the bias ratio of RS_t^+ and RS_t^{+*} when no jump occurs, which replicates the results from Panel (3) of Figure 2.4.2. As in Figure 2.4.2, RS^+ suffers an upward estimation bias if $\theta > 0$ and a downward bias if $\theta < 0$. This ratio deviates from zero as $|\theta|$ becomes larger and tends to be larger for a positive θ than a negative, suggesting that the proportion of the bias in RS^+ is larger in the

presence of positive linear drift than in the presence of an equally-sized negative drift, similar to the conclusions made in the constant drift-diffusion model. In contrast, the bias ratio of RS_t^{+*} is much smaller in magnitude than that of RS_t^+ for a nonzero drift.

Panel (1) reports the bias ratios of RS_t^+ and RS_t^{+*} in the presence of a positive jump, with the size of this jump equal to 60% of the spot volatility. The sign of the bias ratio of RS_t^+ aligns with that of θ, with a larger proportion of bias when $\theta > 0$, which is qualitatively consistent with the properties of the estimation bias in RS_t^+ in the absence of jumps reported on the top. Note that the bias ratio of RS_t^+ for case (1) is positive when $\theta = 0$. This positive but small bias is due to the finite sample effect of that positive jump: the presence of positive jumps increases the expected frequency of positive returns thus causing an upward bias in RS_t^+. However, this bias appears to be very minor (only 0.006). Comparing the bias ratio of RS_t^+ in the absence of jumps on the top panel with that of RS_t^+ in case (1), the latter is generally smaller in size. This may be explained as follows. The finite sample effect of this positive jump is very minor thus the bias in RS_t^+ due to a nonzero drift ($\theta \neq 0$) is dominated by the effect of the finite sample drift. This suggests that the bias in RS_t^+ in the presence of this positive jump are very close to that in RS_t^+ without jumps (I confirm that this is also evidenced by my simulations). However the estimation target of RS_t^+ include positive jump variation, leading to a contraction of the proportion of that drift bias in RS_t^+ (recall that the bias ratio of RS^+ is defined by the bias in RS_t^+ relative to the one-half integrated variation and the sum of squared positive jumps).

For Panel (1), the modified measure RS_t^{+*} has a much lower magnitude of the bias ratio compared to RS_t^+, for all of these nonzero θ, indicating the estimation performance of RS_t^{+*} in the presence of a nonzero linear drift is much better than RS_t^+. The bias ratio of RS_t^{+*} for Panel (1) tends to be more negative than that of RS_t^{+*}

in the top panel. The main reason may be that the median is upward biased due to the occurrence of a positive jump. The positively biased median over-reduces the drift when drift is positive and under-removes the drift when drift is negative, leaving a negative drift in each intraday return, which adds further negative bias in RS_t^{+*}. However, the bias of the median estimator does not largely change the estimation superiority of RS_t^{+*} to RS_t^{+}. The bias ratio of RS_t^{+*} is still very small (1.3%) across nonzero θ, much smaller than that of RS_t^{+}.

Panel (2) reports the bias ratios of RS_t^{+} and RS_t^{+*} for a range of θ, under the scenario of one negative jump. The bias ratios of RS_t^{+} are qualitatively very similar but slightly more negative than those of RS_t^{+} with no jump. The high degree of similarity is because RS_t^{+} always attempt to estimate the one-half integrated variation either in the presence of negative jumps or in the absence of jumps. The small difference is due to the finite sample effect from the presence of the negative jump: the occurrence of negative jumps decreases the expected frequency of positive returns thus causing a downward bias in RS_t^{+}. Additionally, the sign of the bias ratio of RS_t^{+*} for case (2) is positive, which is opposite to that of RS_t^{+*} in the top panel. This is attributed to the negatively biased median estimator of drift due to the existence of negative jumps. The downside biased drift measure under-removes the drift when drift is positive and turns drift to positive when drift is negative, resulting in remaining positive drift, and thus causing a positive bias to RS_t^{+*}. However, the bias of the median due to this positive jump only has a little influence on the workings of RS_t^{+*}. Compared to RS_t^{+}, the bias ratio of RS_t^{+*} is still much lower in size for all of these nonzero θ.

Panel (3) contains the drift-driven bias ratios of RS_t^{+} and RS_t^{+*} for two positive jumps. The bias ratio of RS_t^{+} is further reduced compared to that of RS_t^{+} with one

positive jump, Panel (1), as RS_t^+ of the Panel (3) estimate an even larger component due to more jumps. For Panel (3), RS_t^{+*} still performs very well in estimation practice although the median is upwardly biased by two positive jumps. This shows that the median is resilient or little affected by outliers or jumps, consistent with the simulation by Laurent and Shi (2020) and the fact that it is widely recognized that the median is commonly applied to measure the central tendency (average) when there are outliers.

Panel (4) presents the bias ratios of RS_t^+ and RS_t^{+*} in the presence of nonzero linear drift for two negative jumps. The bias ratio of RS_t^+ is slightly reduced compared to that of RS_t^+ with one positive jump, Panel (1), due to the finite sample effect of positive jumps. Relative to RS_t^+, the bias of RS_t^{+*} is much smaller for all of these nonzero linear drifts. Panel (5) which compares the bias ratios between RS_t^+ and RS_t^{+*} also evidences the much more accurate estimation of RS_t^{+*} than RS_t^+, although the median is distorted by these two negative jumps. The bias ratio of RS_t^{+*} between this case and the case of null jump is very similar, which may be explained by the inclusion of two symmetric jumps.

Table 2.4.2 reports the bias ratio of negative semivariance (RS_t^-), with the respective modified estimator (RS_t^{-*}) for the same five cases of jumps as in Table 2.4.1. The initial log price is set as $p_{t_0} = 6$ and the results are very similar for $p_{t_0} = 2$ or 4. The top panel replicate the results from Panel (4) of Figure 2.4.2 for the bias ratio of RS_t^{-*} and RS_t^- without jump occurrence. As in Figure 2.4.2, the bias ratio of RS_t^- increases nonlinearly with $|\theta|$, and the sign of this bias ratio is opposite to that of θ. This ratio results also demonstrate that the bias ratio of RS_t^- is larger in the presence of positive θ than in the presence of equally sized negative θ, which is due

to the symmetry of the bias ratios of RS_t^+ and RS_t^- around zero θ. In contrast, the bias ratio of RS_t^{-*} is much smaller in magnitude than that of RS_t^- for a nonzero drift.

Table 2.4.2. The estimation bias of negative realized semivariance for different sizes of parameter θ and across various cases of jumps.
Notes: This table compares the estimation biases of negative Realized Semivariance, RS_t^-, with its modified version, RS_t^{-*}, for a range of drifts and under five scenarios of jumps. The biases are calculated on the log prices simulated from the linear drift-diffusion model with additive jumps, Equation (2.4.6).

θ	-0.2	-0.16	-0.12	-0.08	-0.04	0	0.04	0.08	0.12	0.16	0.2
No jump											
RS_t^-	4.555	3.201	2.088	1.200	0.513	0.001	-0.364	-0.613	-0.776	-0.876	-0.935
RS_t^{-*}	-0.006	-0.006	-0.006	-0.006	-0.006	0.000	-0.006	-0.006	-0.006	-0.006	-0.006
(1) One jump with $\phi_{t_i}^1 = 0.6\sigma$											
RS_t^-	4.487	3.150	2.050	1.172	0.494	-0.012	-0.372	-0.619	-0.779	-0.878	-0.936
RS_t^{-*}	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
(2) One jump with $\phi_{t_i}^1 = -0.6\sigma$											
RS_t^-	2.858	2.030	1.342	0.785	0.345	0.008	-0.243	-0.426	-0.557	-0.650	-0.717
RS_t^{-*}	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014	-0.014
(3) Two jumps with $\phi_{t_i}^j = 0.6\sigma$ for $j = 1,2$											
RS_t^-	4.421	3.100	2.013	1.146	0.475	-0.024	-0.380	-0.623	-0.782	-0.880	-0.937
RS_t^{-*}	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020
(4) Two jumps with $\phi_{t_i}^j = -0.6\sigma$ for $j = 1,2$											
RS_t^-	2.169	1.557	1.043	0.622	0.284	0.018	-0.187	-0.342	-0.460	-0.551	-0.621
RS_t^{-*}	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013	-0.013
(5) Two jumps with $\phi_{t_i}^1 = 0.6\sigma$ and $\phi_{t_i}^2 = -0.6\sigma$											
RS_t^-	4.493	3.154	2.053	1.173	0.494	0.000	-0.372	-0.619	-0.779	-0.878	-0.936
RS_t^{-*}	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007

Panel (1) shows the bias ratios of RS_t^- and RS_t^{-*} for one positive jump, with the size of this jump identical to 60% of the spot volatility. The bias ratio of RS_t^- in this case, exhibits qualitatively consistent properties as that of RS_t^- in the absence of jumps: RS_t^- suffers an upward estimation bias if $\theta < 0$ and a downward bias if $\theta > 0$, with a greater ratio of bias if $\theta < 0$. A difference is that the bias ratio of RS_t^- with this positive jump is systematically more negative than that of RS_t^- without jumps, which is due to the finite sample effect of this positive jump: the presence of this positive jump results in less often negative returns, contributing a downward bias in RS_t^-. Of course, the finite sample impact of this positive jump on the estimation of RS_t^- is

small thus the bias in RS_t^- is mostly driven by the effect of the finite sample drift, which explains why the bias ratio of RS_t^- in the presence of the positive jump in Panel (1) are qualitatively very similar to that of RS_t^- in the absence of jumps.

In contrast, the bias ratio of the modified measure RS_t^{-*} of Panel (1) is much lower in size compared to RS_t^- in the presence of nonzero θ considered, indicating the estimation performance of RS_t^{-*} in the presence of nonzero linear drifts may be much better than RS_t^-. The bias ratio of RS_t^{-*} for Panel (1) is generally more positive than that of RS_t^{-*} in the absence of jumps in the top panel. An intuition to explain this result may be that the positively biased median due to the occurrence of a positive jump over removes the drift when drift is positive and underreduces the drift when drift is negative, leaving a negative drift and thus causing extra positive bias in RS_t^{-*}. However, the bias of the median estimator due to the positive jump considered has little influence on the estimation robustness of RS_t^{-*} to RS_t^- in the presence of the nonzero linear drift. The bias ratio of RS_t^{-*} is still much smaller in size than that of RS_t^- for $\theta \neq 0$.

Panel (2) contains the bias ratios of RS_t^- and RS_t^{-*} for different θ and one negative jump. The bias ratios of RS_t^- for $\theta \neq 0$ are qualitatively very similar in properties but smaller in size than those of RS_t^+ with no jump. This difference is because the estimation target RS_t^- include this negative jump, thus diluting the proportion of the bias. The sign of the bias ratio of RS_t^{-*} for Panel (2) is more negative than that of RS_t^{-*} in the top panel. This is because the downwardly biased drift measure (due to the existence of negative jumps) under-removes the drift when drift is positive and turns drift to positive when drift is negative, leaving a positive drift, causing additional negative bias to RS_t^{-*}. However, this extra negative bias

appears to have little impact on the effectiveness of RS_t^{-*}: the bias ratio of RS_t^{-*} is still much lower in size relative to RS_t^- for all of these nonzero θ.

Panel (3) reports the bias ratios of RS_t^- and RS_t^{-*} for various θ and two positive jumps. The bias ratio of RS_t^- is more negative relative to that of RS_t^- with one positive jump, Panel (1), because including more positive jumps further increases the chance of positive return in the finite sample. For case (3), RS_t^{-*} still performs much better than RS_t^- in estimation practice when θ is not zero even though the median is upwardly biased by two positive jumps. Case (4) presents the bias ratios of RS_t^- and RS_t^{-*} across different θ for two negative jumps. The bias ratio of RS_t^- is reduced compared to that of RS_t^- with one positive jump, Panel (1), due to more negative jumps are included in the component that RS_t^- attempts to estimate. Relative to RS_t^-, the bias of RS_t^{-*} is much smaller for all of these nonzero linear drifts, regardless of the negative bias of the median by the two negative jumps. Panel (5) compares the bias ratios between RS_t^- and RS^{-*} in the presence of two symmetric jumps. The results again evidence the much more accurate estimation of RS_t^{-*} than RS_t^-. The bias ratio of RS_t^{-*} between this case and the case of no jump is very similar, which may be interpreted by the inclusion of two symmetric jumps.

Overall, the estimation performance of original and modified semivariances under different jump scenarios is shaped by the finite sample impact of jumps, the proportion of the jump size relative to drift size, the finite sample effect of linear drift, and the effect of the median when biased by the jumps. Notwithstanding the complexity of these factors that influence the result, the influence of a nonzero linear drift on the estimation accuracy of the modified positive realized semivariance can be always much smaller than that of the original positive realized semivariance across

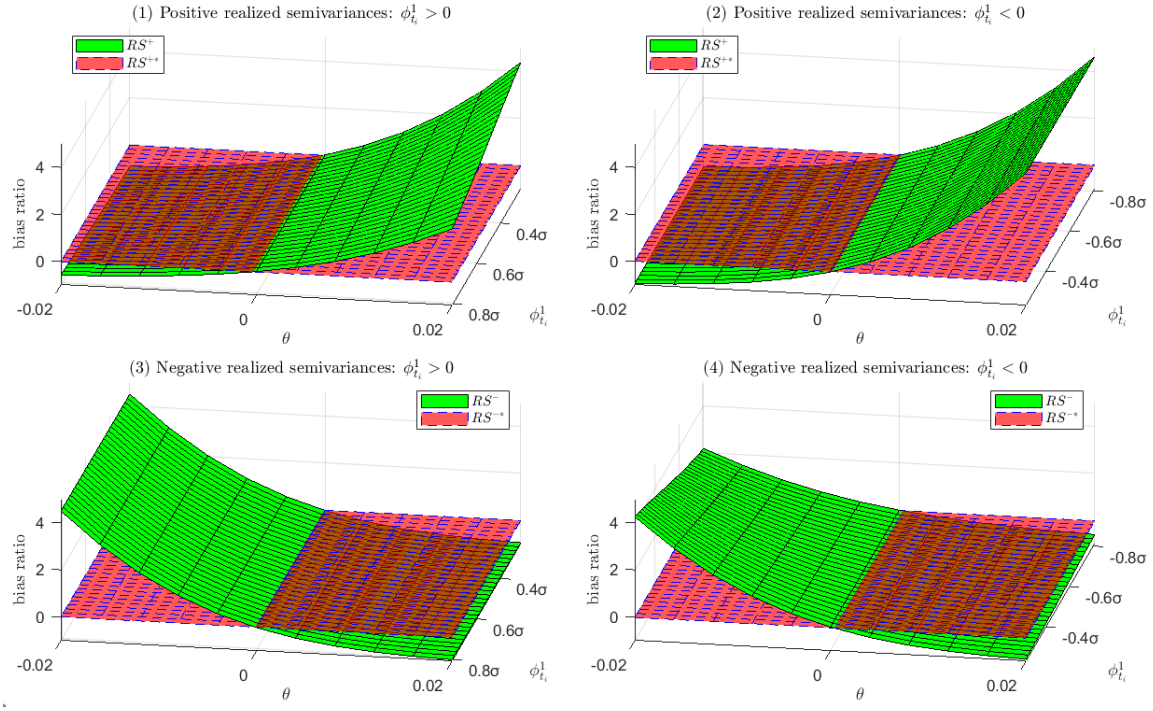
different scenarios of jumps considered, especially when the linear drift is large (within the realistic range of θ).

The simulation results discussed so far rely on a fixed jump size, which equals 60% of spot volatility. To highlight the generality of the results, I also check the bias ratio of semivariances to a broader range of jump sizes, I allow one positive jump in Equation (2.4.6) to vary from $\phi_{t_i}^1 = \sqrt{1/19}\sigma \approx 0.2294\sigma$ to $\phi_{t_i}^1 = \sqrt{2/3}\sigma \approx 0.8165\sigma$ in 0.02 increments, corresponding to around 5% (a very small jump) to 40% proportion (a very large jump) of jumps relative to the Quadratic Variation on the day of a jump arrival. I also consider the case of one negative jump ($\phi_{t_i}^1 < 0$), and the size of this negative jump ranges from $\phi_{t_i}^1 = -\sqrt{2/3}\,\sigma \approx -0.8165\sigma$ to $\phi_{t_i}^1 = -\sqrt{1/19}\sigma \approx -0.2294\sigma$. These ranges of the positive and negative jump are the same as those defined in section 2.3.3. As in Tables 2.4.1 and 2.4.2, I allow the parameter θ to change from -0.02 to 0.02 for each jump size. The remaining parameters in Equation (2.4.6) are those used for Tables 2.4.1 and 2.4.2.

Figure 2.4.5 reports the bias ratios of semivariances and the respective modified semivariance for various combinations of the drift parameter θ and the jump parameter $\phi_{t_i}^1$. Panel (1) displays the bias ratio of RS_t^+ and RS_t^{+*} for $\sqrt{1/19}\sigma < \phi_{t_i}^1 < \sqrt{2/3}\sigma$, Panel (2) presents the bias ratio of RS_t^+ and RS_t^{+*} for $-\sqrt{2/3}\sigma < \phi_{t_i}^1 < -\sqrt{1/19}\,\sigma$, and the lower panels depict the bias ratio of RS_t^- and RS_t^{-*} for $\sqrt{1/19}\sigma < \phi_{t_i}^1 < \sqrt{2/3}\sigma$ and $-\sqrt{2/3}\sigma < \phi_{t_i}^1 < -\sqrt{1/19}\,\sigma$, respectively. The upper panels show that RS_t^+ is upward biased for $\theta > 0$ and is downward for $\theta < 0$ for all of these $|\phi_{t_i}^1| > 0$. In these two panels, the proportion of the bias in RS_t^+ can be very large (up to 400%) especially when θ is largely positive, regardless of the different

values of $\phi_{t_i}^1$. For both panels, the level of the bias of RS_t^+ becomes larger as $|\theta|$ increases, for all of these values of $\phi_{t_i}^1$. It can be seen from Panel (1) that the bias ratio of RS_t^+ for all these nonzero θ is lower in degree as $\phi_{t_i}^1$ becomes more positive. This phenomenon can be explained as follows. The finite sample effect of one positive jump on RS_t^+ is very minor and not sensitive to different sizes of this jump. Therefore, the bias in RS_t^+ with one positive jump is almost driven by the nonzero drift and thus is consistent for different sizes of this jump. But in the presence of one positive jump, RS_t^+ capture the size of this positive jump, thus there is a contraction of the proportion of the drift bias in RS_t^+ if the size of the positive jump becomes larger.

Figure 2.4.5. Biases of semivariances for various combinations of the parameter θ and a jump with its size equal to $\phi_{t_i}^1$



Notes: Panels (1) and (2) present the bias ratio of the positive realized semivariance and its modified version (RS_t^+) and their modified forms (RS_t^{+*}) for various combinations of the parameter θ and a jump with its size equal to $\phi_{t_i}^1$. Panels (3) and (4) present the bias ratio of negative semivariance (RS_t^-) and its modified form (RS_t^{-*}) for various combinations of the parameter θ and a jump with its size equal to $\phi_{t_i}^1$. The bias ratios are calculated on the log prices simulated by the linear drift-diffusion model with additive jumps, Equation (2.4.6).

Panel (2) shows that the bias of RS_t^+ almost does not change with the value of the negative $\phi_{t_i}^1$. For instance, the line corresponding to $\theta = 0.01$ is virtually parallel to the $\phi_{t_i}^1$ axis. I might give a reason for this. The finite sample impact of one negative jump on RS_t^+ is also very small and similar across these jump sizes. This suggests that bias in RS_t^+ with this negative jump is dominated by the finite sample bias due to a nonzero drift and thus is consistent across different jump sizes $\phi_{t_i}^1$. For one negative jump, RS_t^+ is defined only to estimate the one-half integrated variation, thus different sizes of the positive jump almost do not cause changes in the proportion of the drift bias. Comparing RS_t^+ with the modified estimator RS_t^{+*} in Panels (1) and (2), the bias of RS^{+*} due to the existence of nonzero θ is much smaller for all of these $\phi_{t_i}^1$. This suggests that for a wider range of jump sizes the finite sample estimation performance of RS_t^{+*} is consistently much better than the original RS_t^+ estimator.

Turning to Panel (3) of Figure 2.4.5, I find that the bias ratio of RS_t^- almost does not change with the value of $\phi_{t_i}^1 > 0$. This may be interpreted by the following reasons. Due to the small finite sample impact of one positive jump on RS_t^-, the bias in the RS_t^- estimator in the presence of this positive jump is mainly due to the finite sample bias of drift therefore consistent across the different sizes of this jump. The estimation target of RS_t^- does not contain the squared negative jump, and this implies that the ratio of the drift bias in RS_t^- will also be similar across different sizes of one negative jump. As Panel (3) shows, the bias of RS_t^{-*} due to the existence of nonzero θ is much smaller for all of these $\phi_{t_i}^1$. This suggests that for a wider range of jump sizes the finite sample estimation performance of RS_t^{-*} is consistently much better than the original RS^+ estimator.

From Panel (4), I observe that the bias ratio of RS_t^- for all these nonzero θ is lower in degree as $\phi_{t_i}^1$ becomes more positive. This is because the finite sample effect of one negative jump on RS_t^- is minor and similar across different magnitudes of this jump, thus the bias in RS_t^- with one negative jump is almost due to a nonzero drift thus similar over different sizes of this jump. For the scenario of one negative jump, the estimation target of RS_t^- include the squared negative jump, thus explaining why there is a decrease in the proportion of the drift-driven bias ratio of RS_t^- if the magnitude of the negative jump is greater.

2.4.3. The bias of a noise-robust version of realized semivariances

Recent studies argue that the 5-minute frequency sampling is too sparse and thus associated with measurement error, which results in the estimation inefficiency of volatility (Jacod et al. 2009; Podolskij and Vetter 2009; Aït-Sahalia et al. 2012; Lee and Mykland 2012; Hautsch and Podolskij 2013; Christensen et al. 2014). As the estimation inefficiency tends to decrease in proportion to the observed increments of the process as the sampling frequency decreases, these studies suggest using the ultrahigh-frequency sampling scheme (e.g., tick-by-tick) as a better alternative for estimating volatility. However, for the ultrahigh-frequency sampling, prices might be substantially contaminated by noise, induced by microstructure effects that arise from market imperfections such as bid-ask spreads and price discreteness (e.g., Niederhoffer and Osborne 1966; Roll 1984; Black 1986). These noise-contaminated prices invalidate the asymptotic properties of volatility estimators (see Bandi and Russell, 2008 and Hansen and Lunde, 2006 among others).

Assume that the log prices p_{t_i} are contaminated by noise, such that observed noisy prices $p^{\circ}_{t_i}$ are expressed as

$$p^{\circ}_{t_i} = p_{t_i} + u_{t_i}, \text{with } 0 < t_i < T, \tag{4.1}$$

where u_{t_i} is a white noise process with mean zero and variance q^2 as in the literature (Barndorff-Nielsen et al. 2008; Christensen et al. 2014; Laurent and Shi 2020). Barndorff-Nielsen et al. (2008) show that the noise will dominate the positive and negative semivariance as the sampling interval becomes small.

To make inferences about good and bad volatility using noisy ultrahigh-frequency prices, I follow the suggestion by Barndorff-Nielsen et al. (2008) to use the pre-averaging approach of Jacod et al. (2009). This approach is also applied by Jacod et al. (2009), Podolskij and Vetter (2009), Christensen et al. (2014), and Bajgrowicz et al. (2016) to reduce the impact of the microstructure noise on the estimation performance of realized variance, bipower variation, and semivariances. Intuitively, this approach locally smooths the observed price series $p^{\circ}_{t_i}$ so that the microstructure component u_{t_i} (almost) disappears under averaging. Returns on this pre-averaged price series can then be used to construct noise consistent measures of the jump-variation components. To implement pre-averaging, returns are calculated on the log prices that are pre-averaged in a local neighbourhood of H observations,

$$r^{\circ}_{t_i} = \frac{1}{H}\left(\sum_{j=H/2}^{H-1} p^{\circ}_{t_{i+j}} - \sum_{j=0}^{H/2-1} p^{\circ}_{t_{i+j}}\right), \tag{2.4.7}$$

where $H = \lceil \tau\sqrt{M} \rceil$ with the parameter $\tau = 2$ following Christensen et al. (2014).

Based on the pre-averaged return $r_{t_i}^{\diamond}$, the noise-modified RV_t and BV_t (indicated by NRV_t and NBV_t, respectively) are calculated as follows,

$$NRV_t = \frac{M}{M-H+2}\frac{1}{H\psi_H}\sum_{i=1}^{M-H+2}\left|r_{t_i}^{\diamond}\right|^2 - \frac{\widehat{\omega}_t^2}{\theta^2\psi_H}, \tag{2.4.8}$$

$$NBV_t = \frac{M}{M-2H+2}\frac{1}{H\psi_H}\frac{\pi}{2}\sum_{i=1}^{M-2H+2}\left|r_{t_i}^{\diamond}\right|\left|r_{t_{i+H}}^{\diamond}\right| - \frac{\widehat{\omega}_t^2}{\theta^2\psi_H}, \tag{2.4.9}$$

where $\psi_H = (1+2H^{-2})/12$ and $\widehat{\omega}_t^2/\theta^2\psi_H$ is a bias correction, which compensates for the residual microstructure noise that remains after pre-averaging, $\widehat{\omega}_t^2$ denotes the estimator for the noise variance given by,

$$\widehat{\omega}_t^2 = \frac{1}{2(\mathrm{M}-1)}\sum_{i=2}^{M}\left|r_{t_i}^{\diamond}\right|\left|r_{t_{i-1}}^{\diamond}\right|. \tag{2.4.10}$$

Recall that RV_t can be decomposed into RS_t^+ and RS_t^-, respectively, based on decomposing the sign of returns. Analogously, based on splitting the sign of the pre-averaged returns, NRV_t may also be appropriately decomposed into the noise-modified RV_t^+ and RV_t^- (indicated by NRV_t^+ and NRV_t^-),

$$NRV_t^+ = \frac{M}{M-H+2}\frac{1}{H\psi_H}\sum_{i=1}^{M-H+2}\left|r_{t_i}^{\diamond}\right|^2 I(r_{t_i}^{\diamond}>0) - \frac{\widehat{\omega}_t^2}{2\theta^2\psi_H}, \tag{2.4.11}$$

$$NRV_t^- = \frac{M}{M-H+2}\frac{1}{H\psi_H}\sum_{i=1}^{M-H+2}\left|r_{t_i}^{\diamond}\right|^2 I(r_{t_i}^{\diamond}<0) - \frac{\widehat{\omega}_t^2}{2\theta^2\psi_H}, \tag{2.4.13}$$

where for simplicity I assume equally half bias-correction ($\widehat{\omega}_t^2/2\theta^2\psi_H$) for NRV_t^+ and NBV_t^- such that the bias correction drops out when making the difference between NRV_t^+ and NRV_t^-.

Laurent and Shi (2020) find that the bias of NRV_t and NBV_t may not be small due to the non-ignorable impact of a nonzero drift on the pre-average returns. To reduce such drift bias, they modify the NRV_t and NBV_t estimators (indicated by NRV_t^* and NBV_t^*) by computing NRV_t and NBV_t on centred pre-averaged returns,

$$NRV_t^* = \frac{M}{M-H+2}\frac{1}{H\psi_H}\sum_{i=1}^{M-H+2}\left|r_{t_i}^{\diamond} - \widehat{m}_t^{\diamond}\right|^2 - \frac{\widehat{\omega}_t^2}{\theta^2\psi_H}, \tag{2.4.14}$$

$$NBV_t^* = \frac{M}{M-2H+2}\frac{1}{H\psi_H}\frac{\pi}{2}\sum_{i=1}^{M-2H+2}\left|r_{t_i}^{\diamond} - m_t^{\diamond}\right|\left|r_{t_{i+H}}^{\diamond} - \widehat{m}_t^{\diamond}\right| - \frac{\widehat{\omega}_t^2}{\theta^2\psi_H}, \tag{2.4.15}$$

where $\widehat{m}_t^{\diamond} = \text{median}\left(r_{t_1}^{\diamond} \dots r_{t_M}^{\diamond}\right)$ denotes the median of $M-H+2$ pre-averaged returns involved in the computation of the volatilities for day t. Since the impact of a nonzero drift on the pre-average returns is not small, a nonzero drift might also violate the estimation performance of NRV_t^+ and NRV_t^-. Therefore, I also compute $NRV_t^{+\,*}$ and $NRV_t^{-\,*}$ on centred pre-averaged returns,

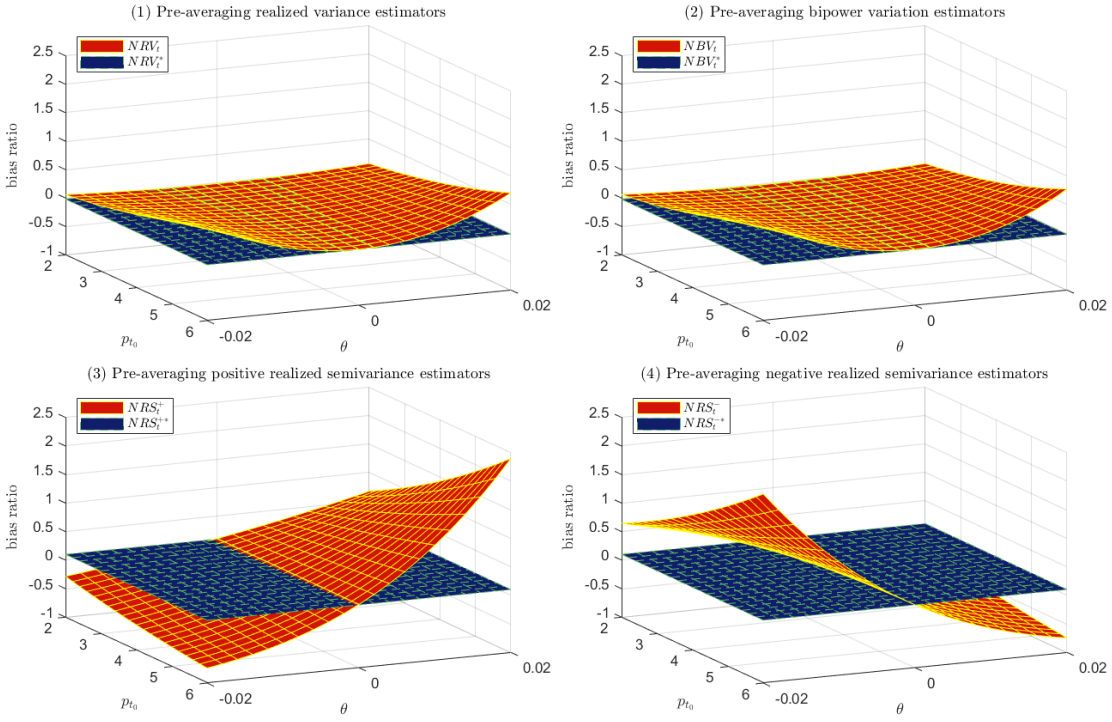
$$NRV_t^{*,+} = \frac{M}{M-H+2}\frac{1}{H\psi_H}\sum_{i=1}^{M-H+2}\left|r_{t_i}^{\diamond} - \widehat{m}_t^{\diamond}\right|^2 I(r_{t_i}^{\diamond} - \widehat{m}_t^{\diamond} > 0) - \frac{\widehat{\omega}_t^2}{2\theta^2\psi_H}, \tag{2.4.16}$$

$$NRV_t^{*,-} = \frac{M}{M-H+2}\frac{1}{H\psi_H}\sum_{i=1}^{M-H+2}\left|r_{t_i}^{\circ} - \widehat{m}_t^{\circ}\right|^2 I(r_{t_i}^{\circ} - \widehat{m}_t^{\circ} < 0) - \frac{\widehat{\omega}_t^2}{2\theta^2\psi_H}, \quad (2.4.17)$$

To evaluate the estimation accuracy of the above pre-averaging estimators and modified pre-averaging estimators I compute the bias of noise-modified volatility estimators and respective modified versions, via Monte Carlo simulations with 10^4 replications of the Ornstein-Uhlenbeck process. I consider no jumps for simplicity. The sampling frequency is at 1 second and the volatility of the noise is set $q = 0.1\%$, as in Laurent and Shi (2020).

The simulated bias ratios of pre-averaged realized variance, bipower variation, positive realized semivariance, and negative realized semivariance, along with their modified versions are reported in Figure 2.4.5. As $|\theta|$ and p_{t_0} deviate from zero, the original estimator, NRV_t and NBV_t, are upwardly biased. In contrast, the bias ratios of the modified estimator, NRV_t^* and NBV_t^*, are not disguisable from zero. The gap between original and modified estimators increases with a larger magnitude of $|\theta|$ and p_{t_0}. This finding is consistent with the pre-averaged realized variance and bipower variation presented in Figures 6 & 7 in Laurent and Shi (2020), implying that modified versions of the pre-averaged realized variance and pre-averaged bipower variation are much more accurate in the presence of a nonzero drift than their original versions. For the NRS_t^+ and NRS_t^- estimators, I observe patterns similar to Figure 3.1 in this chapter. While the estimation accuracy of the original semivariance deteriorates asymmetrically as θ and p_{t_0} deviate from zero, the modified estimator is much more accurate. These results suggest that a nonzero drift can also cause biases in the pre-averaging versions of the semivariances due to the non-ignorable impact of a nonzero drift on the pre-averaged returns. Such biases are much smaller in modified semivariances.

Figure 2.4.5. The finite sample bias ratio of pre-averaging realized semivariances, realized variance, bipower variation, and their modified versions under the Ornstein-Uhlenbeck process



Notes: Each panel of this figure compares the bias ratio of a pre-averaging realized variation measure with its modified version for various combinations of the initial log prices p_{t_0} and parameter θ. The realized measures include pre-averaged versions of the realized variance (NRV_t), Bipower variation (NBV_t), positive realized semivariance (NRS_t^+), and negative realized semivariance (NRS_t^-) while their modified counterparts are indicated by NRV_t^*, NBV_t^*, NRS_t^{+*}, NRS_t^{-*}. The bias ratios of all eight realized measures in this figure are calculated on the log prices simulated from the Ornstein-Uhlenbeck model, Equation (2.4.3), with the parameter settings introduced in section 2.4.1.

2.5. Signed jump estimation

With increasingly finer sampled intraday returns, the discrepancy between the realized positive and negative semivariance measures formally converges to the variation due to positive minus negative price discontinuities, and this variation is also termed the signed jumps. Intuitively, since the variation associated with Brownian price increments, is symmetric and thus the same for the positive and negative semivariance estimators, their difference only formally manifests variation stemming from jumps. Signed jumps help us know whether positive or negative jumps dominate large price moves of the day and thus might reveal information on the leverage effect and investor behaviours. Motivated by this, signed jumps are popular in recent financial forecasting studies, including option pricing (Feunou and Okou 2019), cross-sectional stock return prediction (Bollerslev et al. 2020), and volatility forecasting (Patton and Sheppard 2015; Bollerslev 2022).

The signed jumps in these studies are typically measured via an estimator defined by the gap between positive and negative realized semivariances, motivated by the asymptotic results of Barndorff-Nielsen et al. (2008) that the discrepancy between positive and negative realized semivariances converges to the signed jumps as the sampling interval becomes small,

$$J_t^{\Delta} = RS_t^{+} - RS_t^{-} \xrightarrow{p} \sum_{1<s\leq t} \kappa_s^2 I_{\kappa_s>0} - \sum_{1<s\leq t} \kappa_s^2 I_{\kappa_s<0}, \text{as } \Delta \to 0. \tag{2.5.1}$$

where $J_t^{\Delta} = RS_t^{+} - RS_t^{-}$ denotes the signed jump estimator defined by the difference between positive and negative realized semivariances, $\sum_{1<s\leq t} \kappa_s^2 I_{\kappa_s>0}$ is the sum of all positive jumps, and $\sum_{1<s\leq t} \kappa_s^2 I_{\kappa_s<0}$ is the sum of all negative jumps. According to my

previous results, the modified positive and negative semivariances, denoted by RS_t^{+*} and RS_t^{-*}, can perform better in estimation than RS_t^+ and RS_t^-, respectively. Motivated by the possible benefits of this superiority of RS_t^{+*} and RS_t^{-*}, I define an alternative signed jump estimator based on the gap between RS_t^{+*} and RS_t^{-*},

$$J_t^{\Delta *} = RS_t^{+*} - RS_t^{-*}, \tag{2.5.2}$$

where $J_t^{\Delta *}$ indicates the modified signed jump estimator. For notational simplicity, I write $RS_t^+ - RS_t^-$ as J_t^{Δ} and $RS_t^{+*} - RS_t^{-*}$ as $J_t^{\Delta *}$ for the rest of this chapter.

For a drift-diffusion process with jumps, Equation (2.2.1), the estimation bias in the signed jump estimator J_t^{Δ} is defined by $J_t^{\Delta} - \left(\sum_{1<s\leq t} \kappa_s^2 I_{\kappa_s>0} - \sum_{1<s\leq t} \kappa_s^2 I_{\kappa_s<0}\right)$, and the bias in the corresponding modified estimator $J_t^{\Delta *}$ is defined by $J_t^{\Delta *} - \left(\sum_{1<s\leq t} \kappa_s^2 I_{\kappa_s>0} - \sum_{1<s\leq t} \kappa_s^2 I_{\kappa_s<0}\right)$. My previous conclusions suggest that a nonzero drift can cause a substantial bias in both RS_t^+ and RS_t^- for the finite sample. Do the biases in RS_t^+ and RS_t^- also cause the estimation imprecision of J_t^{Δ}? If the bias in J_t^{Δ} is impactful to its estimation, does $J_t^{\Delta *}$ do a better job? I begin the analysis by exploring the relationship between the bias in J_t^{Δ}, with those in RS_t^+, and RS_t^-.

Lemma 2.5.1. *(1). According to the definition of J_t^{Δ}, along with the definitions of the bias in J_t^{Δ}, RS_t^+, and RS_t^- under a drift-diffusion process with jumps, Equation (2.2.1), the bias in J_t^{Δ} equals the difference between the bias in RS_t^+ and that in RS_t^-,*

$$J_t^{\Delta} - \left(\sum_{1<s\leq t} [\kappa_s I(\kappa_s > 0)]^2 - \sum_{1<s\leq t} [\kappa_s I(\kappa_s < 0)]^2\right) \tag{2.5.3}$$

$$= \left[RS_t^+ - \left(\frac{1}{2} \int_{t-1}^{t} \sigma_s^2 ds + \sum_{1<s\leq t} [\kappa_s I(\kappa_s > 0)]^2 \right) \right]$$

$$- \left[RS_t^- - \left(\frac{1}{2} \int_{t-1}^{t} \sigma_s^2 ds + \sum_{1<s\leq t} [\kappa_s I(\kappa_s < 0)]^2 \right) \right].$$

(2). According to the definition of $J_t^{\Delta *}$, along with the definitions of the bias in $J_t^{\Delta *}$, $RS_t^{+\,*}$, and $RS_t^{-\,*}$, the bias in $J_t^{\Delta *}$ equals the difference between the bias in $RS_t^{+\,*}$ and that in $RS_t^{-\,*}$,

$$J_t^{\Delta\,*} - \left(\sum_{1<s\leq t} [\kappa_s I(\kappa_s > 0)]^2 - \sum_{1<s\leq t} [\kappa_s I(\kappa_s < 0)]^2 \right)$$

$$= \left[RS_t^{+\,*} - \left(\frac{1}{2} \int_{t-1}^{t} \sigma_s^2 ds + \sum_{1<s\leq t} [\kappa_s I(\kappa_s > 0)]^2 \right) \right] \tag{2.5.4}$$

$$- \left[RS_t^{-\,*} - \left(\frac{1}{2} \int_{t-1}^{t} \sigma_s^2 ds + \sum_{1<s\leq t} [\kappa_s I(\kappa_s < 0)]^2 \right) \right].$$

Lemma 2.5.1 is important as it identifies the links between the bias in the signed jump estimator and those in semivariances, suggesting that the bias in the signed jump estimator is exclusively due to (the difference in) the biases in semivariances. Using Lemma 2.5.1, I first derive the biases in J_t^{Δ} and $J_t^{\Delta *}$ under the constant drift-diffusion model, Equation (2.3.2).

Proposition 2.5.1. *Under the drift-diffusion process, Equation (2.3.2).*

The expected bias in the semivariance-based signed jump estimator, J_t^{Δ} is

$$\mathbb{E}\left(J_t^{\Delta} - 0\right) \tag{2.5.5}$$

$$= \mu^2\Delta + \sigma^2 - 2\mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$$

$$- 2\sigma^2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right).$$

For any $\mu > 0$ and $\sigma > 0$, $\mathbb{E}\left(J_t^{\Delta} - 0\right) > 0$ and the bias ratio of J_t^{Δ} is $\left(J_t^{\Delta} - 0\right)/0 = +\infty$. For any $\mu < 0$ and $\sigma > 0$, $\mathbb{E}\left(J_t^{\Delta} - 0\right) < 0$ and the bias ratio of J_t^{Δ} is $\left(J_t^{\Delta} - 0\right)/0 = -\infty$.

Corollary 2.5.1 *The bias in J_t^{Δ} when $\mu = \mu^*$ with $\mu^* \neq 0$ is the same in magnitude as that in J_t^{Δ} when $\mu = -\mu^*$.*

Proposition 2.5.2. *Under the drift-diffusion process, Equation (2.3.2), for any nonzero μ, the expected bias in the semivariance-based modified signed jump estimator $J_t^{\Delta\,*}$ is*

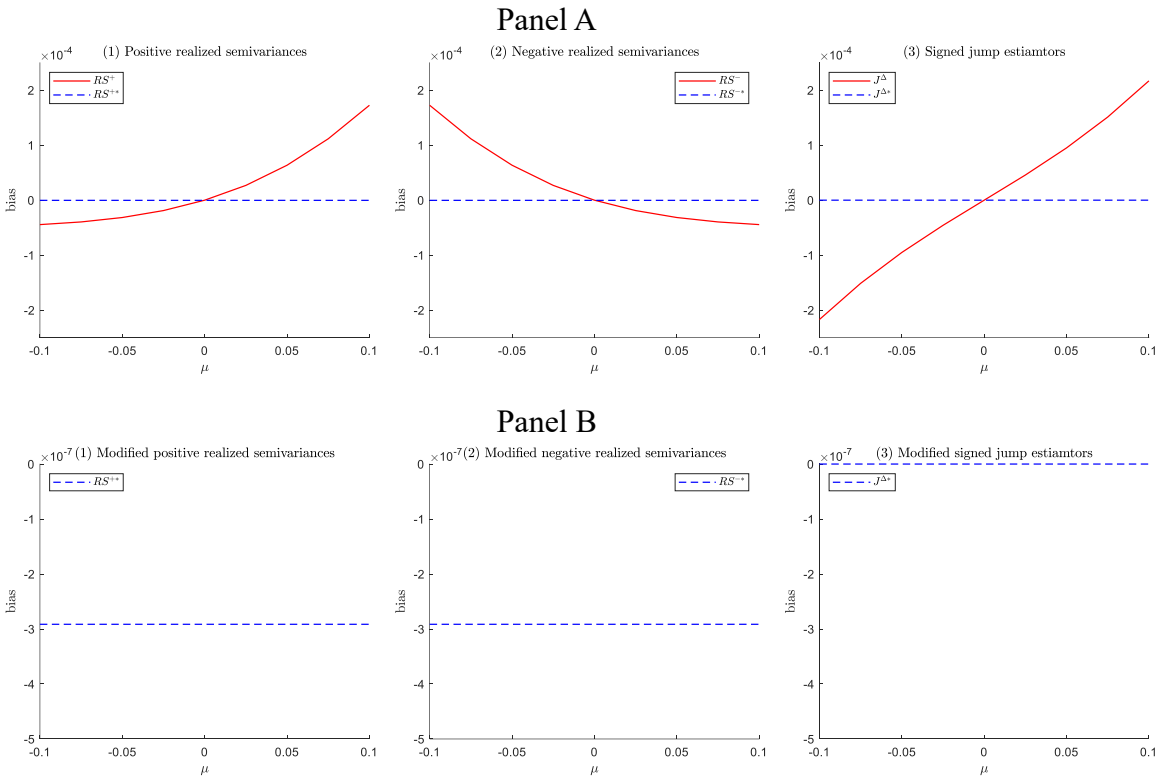
$$\mathbb{E}\left(J_t^{\Delta\,*} - 0\right) = 0. \tag{2.5.6}$$

Proportion 2.5.1 suggests that for a constant drift-diffusion process, the bias in J_t^{Δ} is not zero in the presence of a nonzero drift, and the sign of the bias in J_t^{Δ} aligns with that of the drift. This bias can be much more impactful than that in RS_t^+ or RS_t^- as when jumps are absent the estimation target of J_t^{Δ} is zero while the estimation target of both RS_t^+ or RS_t^- is the one-half integrated variation. Corollary 2.5.1 shows that the magnitude of the bias in J_t^{Δ} is symmetric around zero drift. Proportion 2.5.2 implies that the drift has no impact on the estimation $J_t^{\Delta\,*}$, with the bias expected to be zero. This indicates that $J_t^{\Delta\,*}$ outperforms J_t^{Δ} in terms of the signed jump estimation performance under the constant drift-diffusion model (in the absence of jumps).

From Lemma 2.5.1 and Proportion 2.5.1, under the constant drift-diffusion model, we derive that the bias ratio of J_t^{Δ} is much greater than that of RS_t^+ and RS_t^- but we do not yet know how the bias in J_t^{Δ} compares to those in RS_t^+ and RS_t^-. For the latter comparison, I consider visualizing the biases in J_t^{Δ}, RS_t^+ and RS_t^- using the same realistic parameters (volatility σ, drift μ, and sampling frequency Δ) of the constant drift-diffusion model as in section 2.3.2. To illustrate further Proposition 2.5.2, I also use these parameters to calculate the biases in the corresponding modified estimators, including $J_t^{\Delta\,*}$, $RS_t^{+\,*}$ and $RS_t^{-\,*}$.

I begin the analysis by reporting the bias in the estimators, with the bias ratio results reported later. Figure 2.5.1 depicts the biases in semivariances and signed jump estimators, denoted by RS_t^+, RS_t^-, and J_t^{Δ}, along with the biases in modified estimators, indicated by $J_t^{\Delta\,*}$, $RS_t^{+\,*}$, and $RS_t^{-\,*}$, which are attached in the lower panel for comparison. From the upper right panel, I observe that the sign of the bias in J_t^{Δ} is consistent with the sign of μ and the size of the bias is symmetric around $\mu = 0$, which corroborates Proposition 2.5.1 and Corollary 2.5.1, respectively. The bias in J_t^{Δ} for any nonzero μ considered appears to be greater in magnitude than those in RS_t^+ and RS_t^-. This may be explained by noting that the bias in RS_t^+ is always the opposite sign as that in RS_t^- for the nonzero μ considered. The bias in J_t^{Δ}, which is the difference between the biases in RS_t^+ and RS_t^- (Lemma 2.5.1), thus captures all the biases in RS_t^+ and RS_t^-, with no offset for these nonzero μ, resulting in the size of the bias in J_t^{Δ} is larger than that in RS_t^+ and RS_t^-.

Figure 2.5.1. The finite sample bias in positive realized semivariance, negative realized semivariance, and signed jump estimator, along with their modified versions under the constant drift-diffusion process



Notes: This figure depicts the finite sample bias in the positive realized semivariance (RS_t^+), negative realized semivariance (RS_t^-), modified positive realized semivariance (RS_t^{+*}), and modified negative realized semivariance (RS_t^{-*}) under the constant drift-diffusion process, Equation (2.3.2). The biases of RS_t^+ and RS_t^- are calculated by Proposition 2.3.1 with parameter settings introduced in the earlier paragraphs of section 2.3.2 while the biases in RS_t^{+*} and RS_t^{-*} are computed by Proposition 2.3.2 with the same parameters using simulated log prices from the constant drift-diffusion model, Equation (2.3.2).

The lower right panel reveals that the bias in the modified signed jump estimator ($J_t^{\Delta*}$) is zero, which corroborates Proposition 2.5.2. As shown by the first two plots of the lower panel, the bias in the modified semivariance RS_t^{+*} and RS_t^{-*} are equal in size and sign at about -2.8×10^4 (Proposition 2.3.2). Recall that the bias in $J_t^{\Delta*}$ equals the gap between the biases in RS_t^{+*} and RS_t^{-*} (Lemma 2.5.1). As the biases are fully offset, the bias in $J_t^{\Delta*}$ is zero, which is obviously smaller in size than that in

RS_t^{+*} and RS_t^{-*}. In summary, for a constant drift-diffusion model with the parameters realistic for the stock market, I find that the bias in the semivariance-based signed jump estimator because of a nonzero drift can be larger in magnitude than that in positive and negative semivariances, and the bias in the modified signed jump estimator due to a nonzero drift can be smaller in magnitude than that in modified positive and negative semivariances.

As in my previous analysis of the semivariance estimation, I study the biases performance of the signed jump estimator and its modified form (J_t^{Δ} and $J_t^{\Delta *}$) under constant or linear drift-diffusion model with additive jumps. Following my former methods, the bias ratios of J_t^{Δ} and $J_t^{\Delta *}$ are calculated on the simulated 5-minute log prices from these two drift models, with the sampling frequency, volatility, drift, and jump parameters same as those used in sections 2.3.3 and 2.4.2. To assess the influence of bias in J_t^{Δ} and $J_t^{\Delta *}$ on its estimation accuracy, I consider the bias ratio measure, defined by the proportion of the bias relative to the actual signed jumps.

Table 2.5.1 reports the calculated bias ratios of J_t^{Δ} and $J_t^{\Delta *}$ for the five scenarios of jumps that are used for investigating the bias of semivariances as in Tables 2.3.1 and 2.3.2. For each jump scenario, the bias ratios are computed by varying the drift μ from -0.1 to 0.1, with increments of 0.025. This range of μ is consistent with that considered in section 2.3.2 for studying the bias ratio of semivariances. Panel (1) of Table 2.5.1 reports the bias ratio of the original signed jump estimator J_t^{Δ} and its modified version $J_t^{\Delta *}$ for nonzero drift μ and the case of one positive jump with its size $\phi_{t_i}^1$ equal to 0.6 of spot volatility σ_{t_i} . In the presence of one positive jump, both J_t^{Δ} and $J_t^{\Delta *}$ attempt to estimate the squared positive jump, $0.36\sigma_{t_i}^2$. The result shows that the original estimator J_t^{Δ} is substantially biased, with the magnitude of the bias

ratio larger than 100% for all of these nonzero drifts μ. Moreover, the bias ratio of J_t^{Δ} is much greater in magnitude than the biases of the original semivariances RS_t^+ and RS_t^- reported in the same jump case (1) of Tables 2.3.1 and 2.3.2 for the same nonzero drifts. $J_t^{\Delta *}$ underestimates the signed jumps by 4% for all levels of drift, due to the difference between the negative bias in $RS_t^{+ *}$ and the positive bias in $RS_t^{- *}$ as reported in Tables 2.3.1 and 2.3.2 is not zero. Compared to J_t^{Δ}, the bias ratio of the modified estimator $J_t^{\Delta *}$ is much lower for all of these μ, which indicates that $J_t^{\Delta *}$ achieves a much higher level of estimation accuracy of the signed jumps than J_t^{Δ}.

Table 2.5.1. The bias ratio of signed jump estimators across different scenarios of jumps for the constant drift-diffusion model

Notes: This table presents the bias ratios of the signed jump estimator (J_t^{Δ}) and its modified version ($J_t^{\Delta *}$) across a range of drift μ and for different cases of jumps. The bias of J_t^{Δ} and $J_t^{\Delta *}$ is computed on the simulated 5-minute frequency log price from the constant drift-diffusion model with additive jumps, Equation 2.3.14. Since the signed jumps equal zero for Panel (5), the bias reported in this panel is modified as the ratio relative to the overall size of two jumps $2 \times (0.6\sigma)^2$.

μ	-0.1	-0.075	-0.05	-0.025	0	0.025	0.05	0.075	0.1
(1) One jump with $\phi_{t_i}^1 = 0.6\sigma_{t_i}$									
J_t^{Δ}	-6.303	-4.411	-2.779	-1.328	0.031	1.396	2.863	4.523	6.452
$J_t^{\Delta *}$	-0.041	-0.041	-0.041	-0.041	-0.041	-0.041	-0.041	-0.041	-0.041
(2) One jump with $\phi_{t_i}^1 = -0.6\sigma_{t_i}$									
J_t^{Δ}	-6.465	-4.534	-2.872	-1.404	-0.040	1.319	2.769	4.400	6.290
$J_t^{\Delta *}$	0.042	0.042	0.042	0.042	0.042	0.042	0.042	0.042	0.042
(3) Two jumps with $\phi_{t_i}^j = 0.6\sigma_{t_i}$ for $j = 1,2$									
J_t^{Δ}	-3.274	-2.297	-1.444	-0.678	0.045	0.774	1.557	2.437	3.454
$J_t^{\Delta *}$	-0.035	-0.035	-0.035	-0.035	-0.035	-0.035	-0.035	-0.035	-0.035
(4) Two jumps with $\phi_{t_i}^j = -0.6\sigma_{t_i}$ for $j = 1,2$									
J_t^{Δ}	-3.274	-2.297	-1.444	-0.678	0.045	0.774	1.557	2.437	3.454
$J_t^{\Delta *}$	-0.035	-0.035	-0.035	-0.035	-0.035	-0.035	-0.035	-0.035	-0.035
(5) Two jumps with $\phi_{t_i}^1 = 0.6\sigma_{t_i}$ and $\phi_{t_i}^2 = -0.6\sigma_{t_i}$									
J_t^{Δ}	-3.364	-2.368	-1.502	-0.729	-0.003	0.721	1.494	2.359	3.355
$J_t^{\Delta *}$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001

Case (2) compares the biases of the original signed jump estimator J_t^{Δ} and the respective modified estimator $J_t^{\Delta *}$ over different levels of drifts μ for the presence of

one negative jump with its size $\phi_{t_i}^1 = -0.6\sigma_{t_i}$. For this jump case, the goal of both J_t^{Δ} and $J_t^{\Delta *}$ is to measure the negative sign of the squared negative jump, $-0.36\sigma_{t_i}^2$. The bias ratio of J_t^{Δ} here is systematically greater than 100% in magnitude across all of these nonzero μ, indicating that a nonzero μ again causes unsatisfactory performance of J_t^{Δ} in measuring the negative jump variation. Relative to J_t^{Δ}, the bias of $J_t^{\Delta *}$ is much smaller in size. The sign of the bias of $J_t^{\Delta *}$ is negative since the median is slightly downward biased by the negative jump.

Panels (3) and (4) report the biases of J_t^{Δ} and $J_t^{\Delta *}$ in the presence of a nonzero drift μ for estimating the signed jump variation of two positive and two negative jumps, respectively. For these nonzero μ, the bias of J_t^{Δ} in Panels (3) and (4) is approximately one-half of its bias in Panels (1) and (2). This is because the size of the signed jump variation doubles. But the (absolute) proportion of the bias ratio of J_t^{Δ} in Panels (3) and (4) are still largely greater than zero, especially if μ deviates far from zero. In these two panels, the bias of $J_t^{\Delta *}$ is much smaller than that of J_t^{Δ} across all of these nonzero μ. Panel (5) exhibits the biases of J_t^{Δ} and $J_t^{\Delta *}$ as a function of μ for the scenario of two jumps with both sizes equal to 0.6 of volatility but with opposite signs. Since the signed jump variation equals zero for this jump case, I adjust the biases in this panel as the ratio relative to the overall size of two jumps $2 \times (0.6\sigma)^2$. This adjustment also facilitates comparing the bias reported in Panel (5) with that in Panels (3) and (4), where the bias is also based on the ratio relative to the sum of two squared jumps $2 \times (0.6\sigma)^2$. The bias of J_t^{Δ} in Panel (5) is close to those in Panel (3) and (4) for these nonzero μ: the bias of J_t^{Δ} due to nonzero drift is considerable and the size of the bias increases with $|\mu|$. Comparing of J_t^{Δ} and $J_t^{\Delta *}$ in Panel (5), the latter systematically exhibits a much smaller magnitude of bias across these nonzero drifts.

The comparative results between J_t^{Δ} and $J_t^{\Delta *}$ reported in Table 2.5.1 are based on the simple constant drift-diffusion models. To be consistent with the analysis of the semivariances, I also calculate the biases of J_t^{Δ} and $J_t^{\Delta *}$ via simulating the 5-minute log prices from the more comprehensive linear drift-diffusion model with additive jumps, Equation (2.4.6), with identical parameter settings as in section 2.4.1. Table 2.5.2 reports the bias ratios of J_t^{Δ} and $J_t^{\Delta *}$ for the same five scenarios of jumps as in Tables 2.5.1. For each jump setting, the biases are computed for the θ varying from -0.02 to 0.02 with the initial log price at $p_{t_0} = 6$. To avoid repetition, results for alternative initial log prices $p_{t_0} = 2$ or 4 are not reported here but are qualitatively very consistent with those reported here.

Table 2.5.2. The bias ratio of signed jump estimators across different scenarios of jumps for the linear drift-diffusion model.

Notes: This table presents the biases of the signed jump estimator (J_t^{Δ}) and its modified version ($J_t^{\Delta *}$) across a range of θ with the initial log price at $p_{t_0} = 6$ and for different cases of jumps. The bias of J_t^{Δ} and $J_t^{\Delta *}$ is computed on the simulated 5-minute frequency log price from the constant drift-diffusion model with additive jumps, Equation (2.4.6). Since the signed jumps equal zero for case (5), the bias reported in this panel is modified as the ratio relative to the overall size of two jumps $2 \times (0.6\sigma)^2$

θ	-0.2	-0.16	-0.12	-0.08	-0.04	0	0.04	0.08	0.12	0.16	0.2
(1) One jump with $\phi_{t_i}^1 = 0.6\sigma_{t_i}$											
J_t^{Δ}	-7.930	-5.916	-4.170	-2.643	-1.269	0.030	1.339	2.746	4.334	6.179	8.341
$J_t^{\Delta *}$	-0.040	-0.040	-0.040	-0.040	-0.040	-0.040	-0.040	-0.040	-0.040	-0.040	-0.040
(2) One jump with $\phi_{t_i}^1 = -0.6\sigma_{t_i}$											
J_t^{Δ}	-8.119	-6.062	-4.283	-2.732	-1.343	-0.041	1.261	2.651	4.211	6.015	8.125
$J_t^{\Delta *}$	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044
(3) Two jumps with $\phi_{t_i}^j = 0.6\sigma_{t_i}$ for $j = 1,2$											
J_t^{Δ}	-4.109	-3.075	-2.171	-1.373	-0.647	0.044	0.743	1.494	2.337	3.310	4.441
$J_t^{\Delta *}$	-0.034	-0.034	-0.034	-0.034	-0.034	-0.034	-0.034	-0.034	-0.034	-0.034	-0.034
(4) Two jumps with $\phi_{t_i}^j = -0.6\sigma_{t_i}$ for $j = 1,2$											
J_t^{Δ}	-4.319	-3.243	-2.307	-1.484	-0.745	-0.050	0.643	1.376	2.190	3.123	4.203
$J_t^{\Delta *}$	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037
(5) Two jumps with $\phi_{t_i}^1 = 0.6\sigma_{t_i}$ and $\phi_{t_i}^2 = -0.6\sigma_{t_i}$											
J_t^{Δ}	-4.215	-3.160	-2.239	-1.429	-0.696	-0.003	0.693	1.434	2.263	3.216	4.322
	<0.00	<0.00	<0.00	<0.00	<0.00	<0.00	<0.00	<0.00	<0.00	<0.00	<0.00
$J_t^{\Delta *}$	1	1	1	1	1	1	1	1	1	1	1

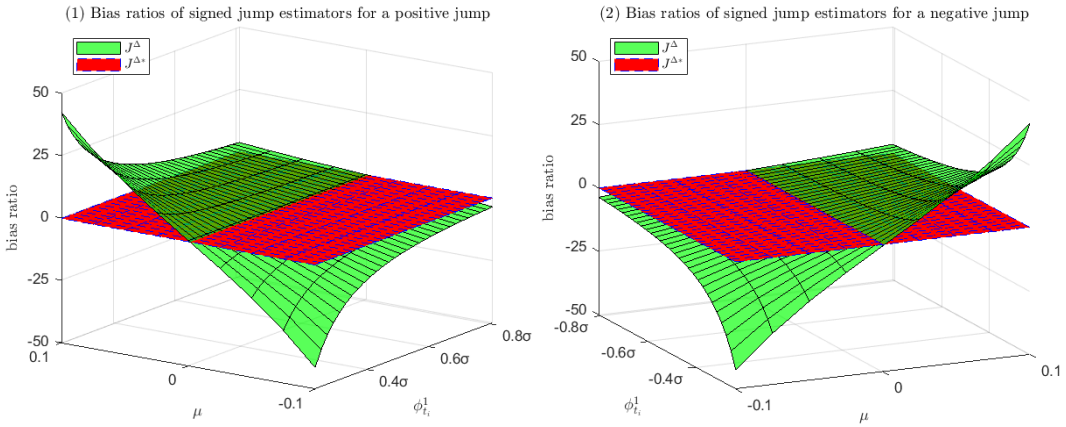
For all nonzero θ and all of these five cases of jumps in Table 2.5.1, the original estimator J_t^{Δ} is still largely biased while the bias ratio of the modified estimator, $J_t^{\Delta *}$ is again much smaller in size. Although $J_t^{\Delta *}$ either underestimate or overestimate the signed jumps due to the existence of jumps, the proportion of the bias in $J_t^{\Delta *}$ is consistently very small (less than 5%). Overall, the estimation superiority of the modified signed jump estimator over the original counterpart in the presence of nonzero drifts is consistent when the log prices are simulated from the more sophisticated linear drift-diffusion model.

The results of the bias of signed jump estimators J_t^{Δ} and $J_t^{\Delta *}$ in the above Tables 2.5.1 and 2.5.2 are investigated by assuming one size of jumps, which is 0.6 of volatility. For checking the robustness of my results in Tables 2.5.1 to different jump sizes, I allow one positive or negative jump in the constant drift-diffusion model, Equation (2.3.14) to vary from 0.2294σ to 0.8165σ or -0.8165σ to -0.2294σ as in section 2.3.3, with the remaining parameters of the model also consistent with those considered in section 2.3.3.

Panel (1) of Figure 2.5.1 depicts the bias of J_t^{Δ} and $J_t^{\Delta *}$ for different cases of the drift μ and a positive jump with size $\phi_{t_i}^1 > 0$, while Panel (2) presents the bias for various combinations of the drift and a negative jump with size $\phi_{t_i}^1 < 0$. As in Table 2.5.1, for the $\phi_{t_i}^1$ considered, the original estimator J_t^{Δ} contains an upward bias for $\mu > 0$ and a downward bias for $\mu < 0$. This is because for these $\phi_{t_i}^1$, there is always a positive bias in RS_t^+ and negative bias in RS_t^- for $\mu > 0$, and a negative bias in RS_t^+ and positive bias in RS_t^- for $\mu < 0$, as reported in Figure 2.3.6. The size of the bias of J_t^{Δ} increases with $|\mu|$ and decreases with $\left|\phi_{t_i}^1\right|$, as larger jumps further dilute the proportion of the bias due to a nonzero μ. In contrast to the unsatisfactory estimation

accuracy of J_t^{Δ}, the modified estimator $J_t^{\Delta *}$ exhibits much lower sizes of the bias for all these combinations of nonzero μ and $\phi_{t_i}^1$.

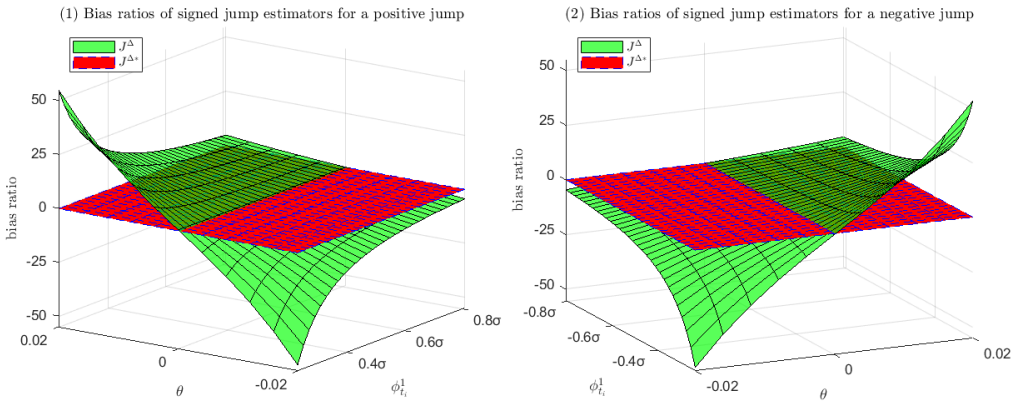
Figure 2.5.1. Bias ratios of signed jump estimators as a function of the sizes of the drift and jump under the constant drift-diffusion model



Notes: Panel (1) presents the bias ratios of the signed jump estimator (J_t^{Δ}) and its modified form ($J_t^{\Delta *}$) as a function of the sizes of the drift μ and a positive jump $\phi_{t_i}^1 > 0$. Panel (2) reports the biases of these two estimators as a function of the sizes of μ and a negative jump $\phi_{t_i}^1 < 0$. The bias ratio of J_t^{Δ} is defined by the bias in J_t^{Δ} relative to the size of the signed jumps $\left|\phi_{t_i}^1 \sigma\right|$, and the bias ratio of $J_t^{\Delta *}$ is defined by the bias in $J_t^{\Delta *}$ relative to $\left|\phi_{t_i}^1 \sigma\right|$. The biases in J_t^{Δ} and $J_t^{\Delta *}$ are calculated on the log prices simulated by the constant drift-diffusion model, Equation (2.3.14).

Figure 2.5.2 depicts the bias ratio of J_t^{Δ} and $J_t^{\Delta *}$ in the presence of one positive or negative jump in the linear drift-diffusion model, Equation (2.4.6), to change within the same range as Figure 2.5.1, with the remaining parameters of the model consistent with those in section 2.4.1. Panel (1) depicts the bias ratio of the original signed jump estimator J_t^{Δ} and its modified version $J_t^{\Delta *}$ with different sizes of θ and a positive jump $\phi_{t_i}^1 > 0$, while Panel (2) presents the bias results for the sizes of θ and a negative jump $\phi_{t_i}^1 < 0$.

Figure 2.5.2. Bias ratios of signed jump estimators as a function of the sizes of the drift and jump under the linear drift-diffusion model



Notes: Panel (1) presents the bias ratios of the signed jump estimator (J_t^{Δ}) and its modified form ($J_t^{\Delta *}$) as a function of the sizes of the parameter θ and a positive jump $\phi_{t_i}^1 > 0$. Panel (2) reports the biases of these two estimators as a function of the sizes of θ and a negative jump $\phi_{t_i}^1 < 0$. The biases are defined by the difference between the estimator and the signed jumps relative to the size of the signed jumps, and the biases are calculated on the log prices simulated by the linear drift-diffusion model, Equation (2.4.6).

The results of the bias ratios reveal a similar pattern as those reported in Table 2.5.2. For all of these $\phi_{t_i}^1$, J_t^{Δ} always overestimates when $\theta > 0$ and underestimates when $\theta < 0$. The size of the bias ratio of J_t^{Δ} increasing as $|\theta|$ becomes larger and $\left|\phi_{t_i}^1\right|$ gets smaller. Across these sizes of signed jump $\phi_{t_i}^1$ and θ, the signed jump estimator J_t^{Δ} often misestimate the actual signed jumps by over 100% and up to about $30 \times 100\%$, indicating that this bias is much more impactful than that in positive and negative semivariances reported in Figure 2.3.6. Comparing J_t^{Δ} with $J_t^{\Delta *}$, the bias ratio of $J_t^{\Delta *}$ is systematically much smaller for all of these different combinations of $\phi_{t_i}^1$ and nonzero θ, regardless of the sign of $\phi_{t_i}^1$.

In summary, theoretical and simulation results suggest that the bias ratio of the signed jump estimator is much higher than that of the volatility estimators. Can these findings of the biases also be evidenced from the real data in the stock markets? The next section will explore this research question using high-frequency SPY data.

2.6. Data Description

I present the results for the SPDR S&P 500 Growth ETF (SPY) to investigate empirical evidence for the proportion of the bias in the signed jump estimator due to a nonzero drift, compared to those of the realized semivariances, bipower variation, and realized variance. This exchange-traded fund has been constructed to be a broad representation of the overall stock market and has been applied by previous literature as the market portfolio proxies (e.g., Barigozzi et al. 2014; Patton and Sheppard 2015; Fan et al. 2016; Gao et al. 2018). Therefore, this chapter uses SPY as the proxy for the S&P 500 index. I begin this section with a brief discussion of the data sources, followed by descriptive statistics of the resulting measures for volatility, signed jumps and drift bias.

I obtain tick-by-tick SPY prices from Tick Data Inc. The sample period is from January 2, 1997, to September 21, 2021, with a total of $n = 6222$ days. The SPY tick data is cleaned according to the standard rules used in the literature (Barndorff-Nielsen et al. 2009; Christensen et al. 2014; Patton and Sheppard 2015; Bollerslev et al. 2016; Jiang and Zhu 2017; Christensen et al. 2023):

1. Transactions outside 9:30:00 to 16:00:00 were removed.
2. Transactions with a 0 price or volume were removed.
3. Only retain the transaction prices from the most active exchange of each day (the transaction prices from other exchanges were dropped).
4. Only transaction prices from regular trades were retained (I removed prices with irregular trades that are highlighted by Tickdata.com). The classification details of regular and irregular trades can be found on the official website of Tick Data Inc., https://www.tickdata.com/.

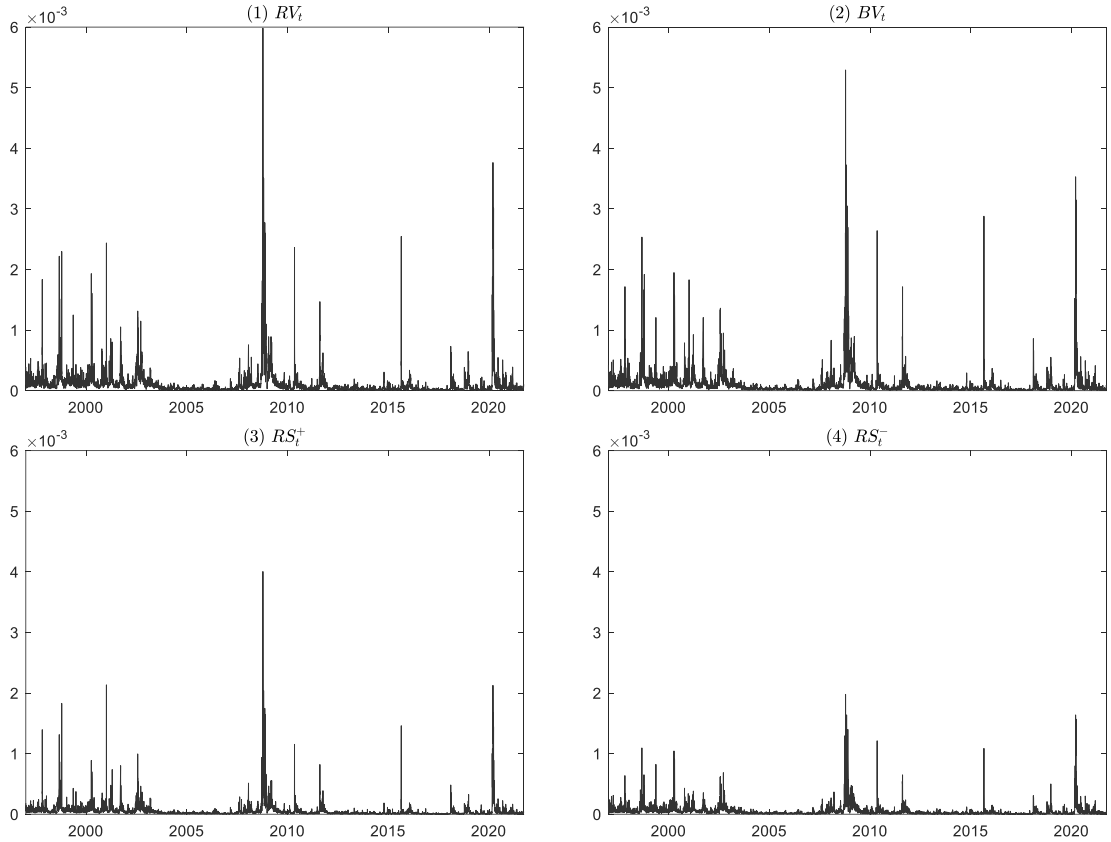
5. If multiple transactions have the same timestamp, use the median price.
6. Delete transaction prices related to corrected trades. The details of corrected trades can be found on the official website of Tick Data Inc.

Following the common practise in the realized variance literature (Andersen et al. 2007b; Corsi 2009; Patton and Sheppard 2015; Bollerslev et al. 2016; Bollerslev et al. 2020; Laurent and Shi 2020; Bollerslev 2022; Caporin 2023; Christensen et al. 2023), this chapter focuses on price volatility during the intraday session of a trading day, thus overnight returns are excluded. After cleaning, the tick-by-tick transaction prices of SPY are then sampled at the 5-minute frequency. The daily median of the 5-minute returns is applied to estimate the daily (average) price drift of these returns. The 5-minute returns are also used for calculating the estimators that are defined in section 2.3, including realized variance (RV_t), bipower variation (BV_t), positive realized semivariance (RS_t^+), negative realized semivariance (RS_t^-), and the signed jump estimator J_t^{Δ}, along with their modified versions, denoted by RV_t^*, BV_t^*, ${RS_t^+}^*$, ${RS_t^-}^*$, and ${J_t^{\Delta}}^*$. Recall that the signed jump estimator is a function of the realized semivariances, defined as $J_t^{\Delta} = RS_t^+ - RS_t^-$ and ${J_t^{\Delta}}^* = {RS_t^+}^* - {RS_t^-}^*$.

Figure 2.6.1 depicts the daily time series of RV_t, BV_t, RS_t^+, and RS_t^-, reported in Panels (1) to (4), respectively. As can be seen from this figure, both RV_t and BV_t are larger for early 2000, the financial crisis and pandemic recessions, reflecting that the stock market is more volatile during those periods. RV_t is visually very similar as BV_t, which suggests that the size of jumps may be generally very small relative to RV_t. RS_t^+ and RS_t^- are smaller than RV_t as they are decompositions of RV_t. RS_t^+ is generally very similar to RS_t^-, except that RS_t^+ has some much larger extreme values. This is surprising as an extremely negative price move is expected to be much sharper than a positive one. From my inspection of the price dynamic associated with

extremely large RS_t^+, I find that there is generally an explosive upward price trend of the day, indicating a large positive drift may exist in the price dynamic.

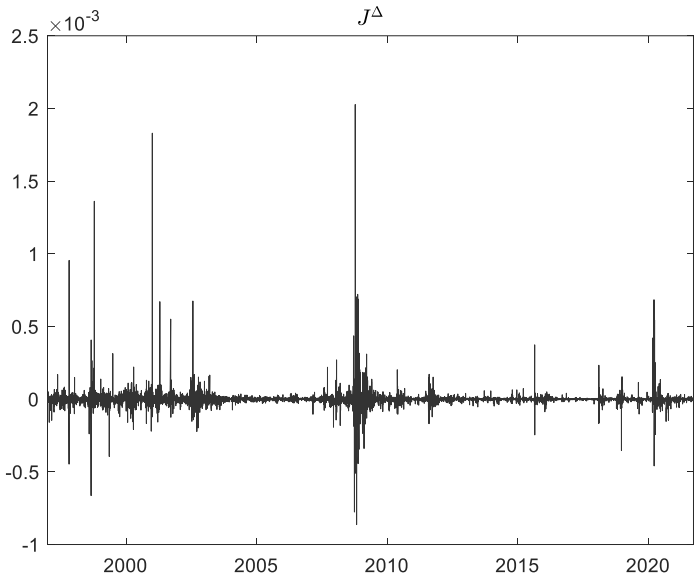
Figure 2.6.1. Daily realized variance, bipower variation, positive semivariance and negative semivariance



Notes: Panel (1), (2), (3), (4) depicts the daily realized variance (RV_t), bipower variation (BV_t), positive semivariance (RS_t^+) and negative semivariance (RS_t^-). The sample is SPDR S&P 500 ETF from 1997 to 2021.

Figure 2.6.2 depicts the daily time series of the original signed jump estimator, indicated by J_t^{Δ}, from January 2, 1997, to September 21, 2021. As the results show, J_t^{Δ} can be positive and negative over the sample length. Similar to RV_t and BV_t, J_t^{Δ} is also much greater in size for economic depressions such as the subprime mortgage bubbles from 2008 to 2009 and the pandemic recession around 2020.

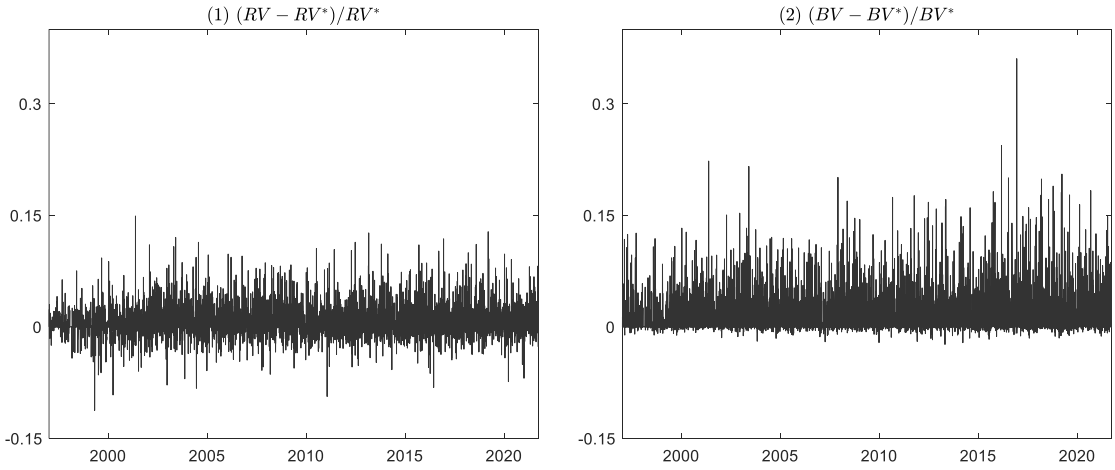
Figure 2.6.2. Daily signed jump estimator



Notes: This figure depicts the daily signed jump estimator (J_t^{Δ}) with $J_t^{\Delta} = RS_t^{+} - RS_t^{-}$. The sample is SPDR S&P 500 ETF from 1997 to 2021.

Figure 2.6.3 compares realized variance (RV_t) and bipower variation (BV_t) with their modified versions, denoted by RV_t^* and BV_t^* for SPY. The first panel depicts the discrepancy between RV_t and RV_t^*, measured as $(RV_t - RV_t^*)/RV_t^*$, while the second panel reports $(BV_t - BV_t^*)/BV_t^*$. These two discrepancies may proxy the proportion of the bias in RV_t and BV_t, as the biases in RV_t^* and BV_t^* are ignorable relative to RV_t and BV_t. The latter discrepancy is also investigated by Laurent and Shi (2020) but $(RV_t - RV_t^*)/RV_t^*$ is somewhat not reported and discussed in the literature. Additionally, these two discrepancies deviate from zero when the daily median (drift) of returns is nonzero. I find that the median is not zero for 4544 out of 6222 days, or 73% of the sample days, suggesting that a nonzero drift is prevalent across the SPY sample.

Figure 2.6.3. The discrepancy between realized variance and modified realized variance, and the discrepancy between bipower variation and modified bipower variation

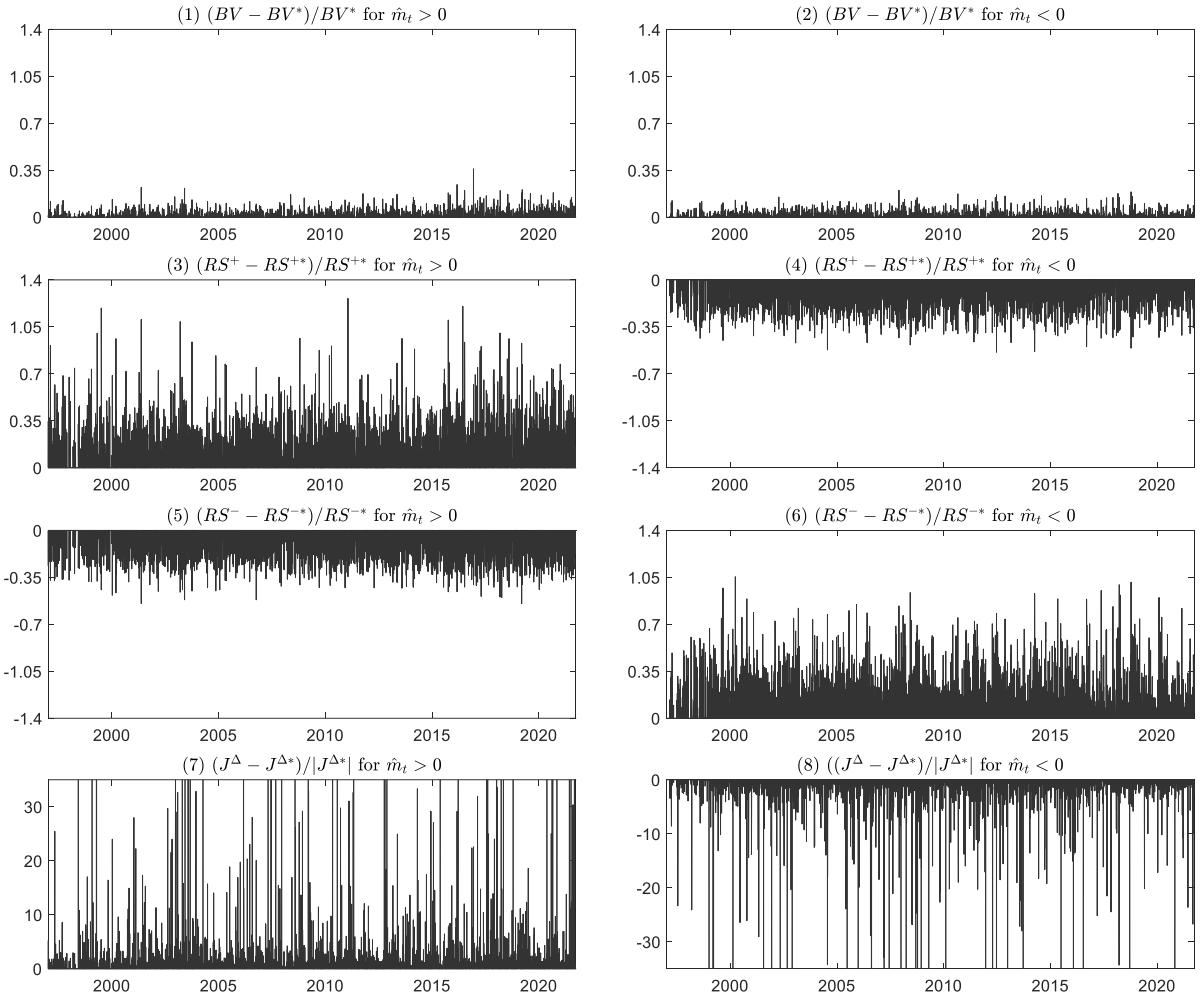


Notes: The first panel depicts the discrepancy between realized variance and modified realized variance, $(RV_t - RV_t^*)/RV_t^*$, and the second panel of, with RV_t and RV_t^* calculated by Equations (2.2.4) and (2.3.10). The second panel shows $(BV_t - BV_t^*)/BV_t^*$, where BV_t^* is computed by Equations (2.6.2) and (2.6.3). The data is SPDR S&P 500 ETF from 1997 to 2021.

From Panel (1), I observe that RV_t tends to be larger than RV_t^* with their gap more likely to be positive over the sample period. The positive discrepancy corroborates the finite sample theory of Laurent and Shi (2020) that RV can be subject to greater positive bias due to a nonzero drift than RV_t^*. The negative discrepancy might be due to the measurement error of RV_t and RV_t^*. From the second panel, I observe the gap between BV_t and BV_t^* is substantially positive, consistent with Laurent and Shi (2020). This implies that the conventional BV very often overestimates the integrated variance. Compared to the results across two panels, the discrepancy between BV and BV^* is generally more positive than that between RV_t and RV_t^*, suggesting that the impact of drift on BV_t may be stronger than that on RV_t, consistent with the findings by Laurent and Shi (2020).

Figure 2.6.4 compares the bipower variation, realized semivariances and the signed jump estimator (BV_t, RS_t^+, RS_t^-, and J_t^Δ) with their modified versions (BV_t^*, RS_t^{+*}, RS_t^{-*}, and $J_t^{\Delta *}$) in the presence of nonzero drifts.

Figure 2.6.4. The discrepancy between original and modified estimators when the daily median of log intraday returns is positive or negative



Notes: Panel (1) depicts the discrepancy between bipower variation (BV_t) with modified bipower variation (BV_t^*) by $(BV_t - BV_t^*)/BV_t^*$ when the daily median of log intraday returns is positive ($\hat{m}_t > 0$). Panel (2) contains this discrepancy for $\hat{m}_t < 0$. Panel (3) presents the gap between positive realized semivariance (RS_t^+) with modified positive realized semivariance (RS_t^{+*}) by $(RS_t^+ - RS_t^{+*})/RS_t^{+*}$ for $\hat{m}_t > 0$. Panel (4) shows this gap for $\hat{m}_t < 0$. Panel (5) reports the discrepancy between negative realized semivariance (RS_t^-) with modified negative realized semivariance (RS_t^{-*}) by $(RS_t^- - RS_t^{-*})/RS_t^{-*}$ for $\hat{m}_t > 0$. Panel (6) depicts this gap for $\hat{m}_t < 0$. Panel (7) reports the discrepancy between the signed jump estimator (J_t^Δ) with the modified signed jump estimator ($J_t^{\Delta *}$) by $\left(J_t^\Delta - J_t^{\Delta *}\right)/\left|J_t^{\Delta *}\right|$ for $\hat{m}_t > 0$. Panels (8) reports this discrepancy for $\hat{m}_t < 0$. The data sample is the 5-minute log prices of the SPDR S&P 500 ETF (SPY) from 1997 to 2021.

The comparisons are based on the discrepancy between the estimator with the modified estimator, relative to the modified estimator. Panels marked with odd numbers report the discrepancy results when the daily median of the log intraday returns ($\widehat{m}_t$) is positive, $\widehat{m}_t > 0$ while the panels with even numbers contain the results for $\widehat{m}_t < 0$. The scale of the horizontal axis is fixed across all of these panels in the figure for ease of comparison. Panels (1) and (2) compare BV_t with BV_t^* via $(BV_t - BV_t^*)/BV_t^*$ for $\widehat{m}_t > 0$ and $\widehat{m}_t < 0$, respectively. This discrepancy is positive and tends to be close in magnitude for $\widehat{m}_t > 0$ and $\widehat{m}_t < 0$, which is consistent with the theoretical and simulation findings by Laurent and Shi (2020) that the bias in BV_t is symmetrically distributed around zero drift.

Panels (3) and (4) compare RS_t^+ with $RS_t^{+\,*}$ via $(RS_t^+ - RS_t^{+\,*})/RS_t^{+\,*}$ for $\widehat{m}_t > 0$ and $\widehat{m}_t < 0$, respectively. As in my finite sample theories and simulations, the results from these two panels indicate that RS_t^+ overestimates if $\widehat{m}_t > 0$ and underestimates if $\widehat{m}_t < 0$. The bias for $\widehat{m}_t > 0$ appears to be greater in magnitude than for $\widehat{m}_t < 0$. which is consistent with the asymmetry of the estimation bias of RS_t^+ around positive and negative drifts found my previous theoretical and simulation results. Comparing Panels (1) and (2) with Panels (3) and (4), the discrepancies in the latter panels appear to be larger, implying that both positive and negative drift may be more impactful to the estimation of RS_t^+ than BV_t. This result corroborates my theoretical findings in Figures 2.3.5 and 2.4.3.

Panels (5) and (6) present the difference between RS_t^- and $RS_t^{-\,*}$ (estimated as $(RS_t^- - RS_t^{-\,*})/RS_t^{-\,*}$) in the presence of the positive and negative $\widehat{m}_t$, respectively. The results suggest that RS_t^- systematically underestimates if $\widehat{m}_t > 0$ and overestimates more severely if $\widehat{m}_t < 0$, which is consistent with my theoretical and simulative findings. Compared to Panels (3) and (4), the discrepancy between RS_t^-

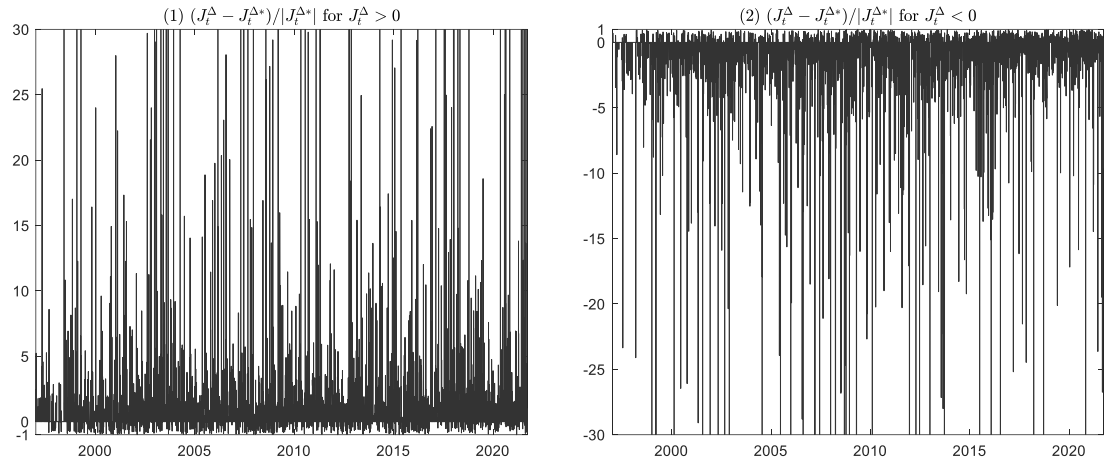
and RS_t^{-*} seem to be close to that between RS_t^+ and RS_t^{+*} with opposite sign of $\widehat{m}_t$. This empirical symmetric pattern of the positive and negative semivariances corroborates my Corollary 2.3.4 and simulation results. Comparing Panels (1) and (2) with Panels (5) and (6), the difference in the latter panels tends to be greater in size, suggesting that both positive and negative drift may be more influential to the estimation of RS_t^- relative to BV_t, thus supporting my theoretical findings in Figures 2.3.5 and 2.4.3.

Panels (7) and (8) report the bias of the signed jump estimator J_t^Δ. These two panels compare J_t^Δ with $J_t^{\Delta *}$ according to the ratio $\left(J_t^\Delta - J_t^{\Delta *}\right)/\left|J_t^{\Delta *}\right|$, for $\widehat{m}_t > 0$ and $\widehat{m}_t < 0$. The results indicate that the estimation bias of J_t^Δ seems to be symmetric around positive and negative $\widehat{m}_t$, with the sign of the bias aligning with that of $\widehat{m}_t$. This is consistent with my theoretical and simulation findings in section 2.5. By comparing the bottom panels with the other six panels, the results suggest that the proportion of the bias of J_t^Δ appears to be much larger than that of RS_t^+ and RS_t^-. This also evidences my findings in section 2.5 that the bias in J_t^Δ can account for a much larger proportion than that in RS_t^+ and RS_t^- estimators.

Since the signed jump estimator J_t^Δ has positive and negative components, I also present the discrepancy $\left(J_t^\Delta - J_t^{\Delta *}\right)/\left|J_t^{\Delta *}\right|$ conditional on $J_t^\Delta > 0$ or $J_t^\Delta < 0$. This allows us to see the estimation bias of the positive and negative components of J^Δ due to a nonzero drift. Figure 2.6.5 reports the discrepancy $\left(J_t^\Delta - J_t^{\Delta *}\right)/\left|J_t^{\Delta *}\right|$ conditional on the sign of J_t^Δ. Panel (1) reports $\left(J_t^\Delta - J_t^{\Delta *}\right)/\left|J_t^{\Delta *}\right|$ for $J_t^\Delta > 0$, while Panel (2) contains this discrepancy for $J_t^\Delta < 0$. When $J_t^\Delta > 0$, the discrepancy can be positive or negative. Interestingly, the positive discrepancy appears to be much greater in size than its negative, with the size of the negative discrepancy no greater than 100%. This

suggests that J_t^{Δ} overestimate signed jumps much more dramatically than underestimate when $J_t^{\Delta} > 0$. When $J_t^{\Delta} < 0$, the result switches, J_t^{Δ} underestimate the signed jumps much more severely.

Figure 2.6.5. The discrepancy between original and modified signed jump estimators when the original estimator is positive or negative



Notes: Panel (1) reports the discrepancy between original and modified signed jump estimators, measured as $\left(J_t^{\Delta} - J_t^{\Delta *}\right)/\left|J_t^{\Delta *}\right|$, when the original estimator is positive ($J_t^{\Delta} > 0$). Panel (2) shows the same discrepancy for $J_t^{\Delta} > 0$.

The exceedingly large positive and negative bias for $J_t^{\Delta} > 0$ and $J_t^{\Delta} < 0$ may be explained by noting the bias in J_t^{Δ} due to a nonzero drift can be much larger than the actual signed jumps thus dominate J_t^{Δ}. In this case, a positive drift can shape a positive J_t^{Δ} no matter the sign of actual signed jumps, and a negative drift can cause a negative J_t^{Δ} regardless of the sign of signed jumps. When the impact of a drift does not overwhelm the signed jumps, J_t^{Δ} could be positive or negative and its signs cannot be determined by the sign of the drift. This explains why the negative and positive discrepancy for $J_t^{\Delta} > 0$ and $J_t^{\Delta} < 0$ are much smaller. Additionally, by comparing the original and modified signed jump estimators, I find that for $J_t^{\Delta} > 0$ and $J_t^{\Delta} < 0$, the rates of misclassifying the sign of signed jumps are 24.5% and 22.1%, respectively,

and for overall J_t^{Δ} this rate equals 23.3%. This suggests that the impact of a nonzero drift on J_t^{Δ} causes that J_t^{Δ} frequently misidentifies the sign of the signed jump.

The discrepancies of volatility and signed jump estimators reported in Figure 2.6.4 contain some large spikes, therefore resulting in difficulties in making a general conclusion of the comparison between those discrepancies. For generality purposes, Table 2.6.1 produces quantile statistics for those discrepancies reported in Figure 2.4.2, with the quantile set as 5%, 25%, 50%, 75%, and 95%.

Table 2.6.1. Quantile statistics for the discrepancy between original and modified estimators when the daily median of log intraday return is positive or negative

Notes: This table reports the quantile statistics for the discrepancy between original and modified estimators between bipower variation (BV_t) with modified bipower variation (BV_t^*), $(BV_t - BV_t^*)/BV_t^*$, the discrepancy between positive realized semivariance (RS_t^+) and modified positive realized semivariance (RS_t^{+*}), $(RS_t^+ - RS_t^{+*})/RS_t^{+*}$, the discrepancy between negative realized semivariance (RS_t^-) and modified negative realized semivariance (RS_t^{-*}), $(RS_t^- - RS_t^{-*})/RS_t^{-*}$, and the discrepancy between the signed jump estimator (J_t^{Δ}) and the modified signed jump estimator ($J_t^{\Delta*}$), $\left(J_t^{\Delta} - J_t^{\Delta*}\right)/J_t^{\Delta*}$. $\widehat{m}_t$ denotes the daily median of returns.

	5%	25%	50%	75%	95%
$(BV_t - BV_t)/BV_t^*$ for $\widehat{m}_t > 0$	-0.005	0.002	0.012	0.033	0.086
for $\widehat{m}_t < 0$	-0.004	0.002	0.012	0.032	0.081
$(RS_t^+ - RS_t^{+*})/RS_t^{+*}$ for $\widehat{m}_t > 0$	0.039	0.103	0.184	0.295	0.524
for $\widehat{m}_t < 0$	-0.344	-0.229	-0.154	-0.093	-0.034
$(RS_t^- - RS_t^{-*})/RS_t^{-*}$ for $\widehat{m}_t > 0$	-0.328	-0.224	-0.154	-0.094	-0.039
for $\widehat{m}_t < 0$	0.032	0.102	0.179	0.296	0.572
$\left(J_t^{\Delta} - J_t^{\Delta*}\right)/\left\|J_t^{\Delta*}\right\|$ for $\widehat{m}_t > 0$	0.165	0.491	1.012	2.229	11.09
for $\widehat{m}_t < 0$	-12.70	-2.388	-1.072	-0.531	-0.161

The discrepancy $(BV_t - BV_t)/BV_t^*$ is systematically very similar for $\widehat{m}_t > 0$ and $\widehat{m}_t < 0$, indicating that the phenomenon of the bias in BV_t is symmetric around positive and negative drift generally exists across the quantiles considered. Both $(RS_t^+ - RS_t^{+*})/RS_t^{+*}$ and $(RS_t^- - RS_t^{-*})/RS_t^{-*}$ are asymmetric around $\widehat{m}_t = 0$: the magnitude of the discrepancy when the sign of $\widehat{m}_t$ equals that of the semivariance, with few exceptions, is overwhelmingly greater than that when the sign of $\widehat{m}_t$ is opposite to that of the semivariance. $(RS_t^+ - RS_t^{+*})/RS_t^{+*}$ is qualitatively very

similar to $(RS_t^- - RS_t^{-*})/RS_t^{-*}$ when their signs are equal. But there are still some differences, especially for the extreme quantiles: $(RS_t^- - RS_t^{-*})/RS_t^{-*}$ with $\widehat{m}_t < 0$ is larger than $(RS_t^+ - RS_t^{+*})/RS_t^{+*}$ with $\widehat{m}_t > 0$ for the quantile 95%, and $(RS_t^+ - RS_t^{+*})/RS_t^{+*}$ with $\widehat{m}_t < 0$ is greater in size than $(RS_t^+ - RS_t^{+*})/RS_t^{+*}$ with $\widehat{m}_t > 0$ for the quantile 5%. This may be explained by noting that the downside drift is often larger than the upside drift for the stock market, especially during economic recessions.

At least one-half of the discrepancy $\left(J_t^{\Delta} - J_t^{\Delta *}\right)/\left|J_t^{\Delta *}\right|$ is larger than 100% for $\widehat{m}_t > 0$ and $\widehat{m}_t < 0$, corroborates my simulations findings that J_t^{Δ} may often underestimate or overestimate the actual signed jumps by 100%. This discrepancy is overwhelmingly greater in size than those related to volatility estimators above across all of these quantiles, and in particular for 5% and 95% quantiles. Also, $\left(J_t^{\Delta} - J_t^{\Delta *}\right)/\left|J_t^{\Delta *}\right|$ exhibits a qualitatively symmetric pattern around zero $\widehat{m}_t$ but is slightly larger in magnitude for $\widehat{m}_t < 0$, especially for quantiles 5%. This may be because of the larger size of the downside drift compared to the upside drift for the stock market over the crisis periods.

2.7. The impact of good volatility and bad volatility, and signed jumps on future volatility

Forecasting price volatility allows us to predict how prices may vary in the future, which is important for investors and the financial industry involved with option pricing (Black and Scholes 1973), derivative pricing (Duffie et al. 2000), asset allocation (Merton 1969), and risk management (Christoffersen and Diebold 2000). A large volume of empirical studies has shown the predictive importance of long-memory dependence in financial market volatility, which is typically characterized by the fact that daily realized variance can have autocorrelations which are significant for many lags. This stylized fact motivates the estimation of long-memory type $ARFIMA$ models for realized volatilities in Areal and Taylor (2002), Andersen et al. (2003), Thomakos and Wang (2003), Pong et al. (2004), Koopman et al. (2005), and Deo et al. (2006).

Corsi (2009) proposed a Heterogeneous Autoregressive (HAR) class of volatility models with one-day, one-week, and one-month lagged volatility averages. He finds that this mixing of relatively short-term lagged volatility components is capable of reproducing the remarkably slow volatility autocorrelation decay that is almost indistinguishable from that of a hyperbolic (long-memory) pattern over the most empirically relevant forecast horizon. Moreover, the HAR model outperforms the $ARFIMA$ model in terms of out-of-sample volatility forecast (Corsi 2009). Due to its superiority in modelling the long memory effect of volatility, the HAR model has been widely applied in key research on realized variance forecasting for the last two decades (Andersen et al. 2007b; Corsi et al. 2010; Corsi and Renò 2012; Duong and Swanson 2015; Patton and Sheppard 2015; Bollerslev et al. 2016; Buccheri and Corsi 2021; Bollerslev 2022; Laurent et al. 2022b; Andersen et al. 2023; Caporin 2023). To

provide more evidence to the literature, I apply the HAR model for forecasting volatility for my SPY data. Specifically, the HAR model presented by Corsi (2009) is defined as follows:

$$V_{t+1} = \beta_0 + \beta_d V_t + \beta_w \left(\frac{1}{5} \sum_{i=0}^{4} V_{t-i} \right) + \beta_m \left(\frac{1}{22} \sum_{i=0}^{21} V_{t-i} \right) + \epsilon_{t+1}, \qquad (2.7.1)$$

where V_{t+1} is the daily volatility measure being forecasted, typically RV_{t+1}, and V_t denotes current volatility, which is the 1-day lagged value with respect to V_{t+1}. The remaining terms represent the average volatility over the past 5 days, $\frac{1}{5}\sum_{i=0}^{4} V_{t-i}$, and the past 22 days, $\frac{1}{22}\sum_{i=0}^{21} V_{t-i}$, and ϵ_{t+1} is the disturbance. The predictors in this HAR model have some overlapping lags. Specifically, the past 5-day average volatility $\frac{1}{5}\sum_{i=0}^{4} V_{t-i}$ includes the past 1-day volatility V_t, and the past 22-day average volatility $\frac{1}{22}\sum_{i=0}^{21} V_{t-i}$ includes the average from the shorter 5-day lagged volatility $\sum_{i=0}^{4} V_{t-i}$. To reduce the overlap between these predictors, this HAR model used throughout this chapter adopts the non-overlapping reparameterization of the HAR model suggested by Patton and Sheppard (2015),

$$V_{t+1} = \beta_0 + \beta_d V_t + \beta_w \left(\frac{1}{4} \sum_{i=1}^{4} V_{t-i} \right) + \beta_m \left(\frac{1}{17} \sum_{i=5}^{21} V_{t-i} \right) + \epsilon_{t+1}, \qquad (2.7.2)$$

where $\frac{1}{4}\sum_{i=1}^{4} V_{t-i}$ excludes the past 1-day volatility V_t and $\frac{1}{17}\sum_{i=5}^{21} V_{t-i}$ removes the average volatility over lags 1 to 4. In the above Equation (2.7.2), both the forecast target, V_{t+1}, and predictors, measured by the lagged volatility, may be proxied by realized variance such that the HAR model becomes

$$RV_{t+1} = \beta_0 + \beta_d\, RV_t + \beta_w\, \overline{RV}_{w,t} + \beta_m\, \overline{RV}_{m,t} + \epsilon_{t+1}, \qquad (2.7.3)$$

where $\overline{RV}_{w,t} = \frac{1}{4}\sum_{i=1}^{4} RV_{t-i}$ and $\overline{RV}_{m,t} = \frac{1}{17}\sum_{i=5}^{21} RV_{t-i}$. Of course, the volatility estimator, which measures the forecast target V_{t+1} of Equation (2.7.7) need not necessarily be the same as the estimator used for measuring the lagged volatilities on the right-hand side.

The good and bad volatility indicates the risk of prices going up and down. Motivated by the intuition that disentangling downside risk from upside risk may add more information beyond seeing these two risks as a whole, previous studies investigate whether dividing volatility into good and bad volatility improves volatility forecasting. The literature generally finds that this decomposition enhances volatility forecasts (Sévi 2014; Patton and Sheppard 2015; Todorova 2017; Gong and Lin 2021; Özbekler et al. 2021). To provide the literature with more evidence, I decompose RV_t in Eqaution 2.7.3 into RS_t^+, and RS_t^-, following the same RS specification by Patton and Sheppard (2015),

$$RV_{t+1} = \beta_0 + \beta_{\mathrm{d}}^{+}\, RS_t^+ + \beta_{\mathrm{d}}^{-}\, RS_t^- + \beta_w\, \overline{RV}_{w,t} + \beta_m\, \overline{RV}_{m,t} + \epsilon_{t+1}. \qquad (2.7.4)$$

Note that if this decomposition of volatility would add no information, we would see that the coefficients of RV_t in Equation (2.7.3), and those of RS_t^+ and RS_t^- are equal, as indicated by $\beta_d = \beta_d^- = \beta_d^+$.

Jumps are large moves in the asset prices, with the upside and downside jumps often associated with good and bad news (Evans 2011; Lahaye et al. 2011; Gilder et al. 2014). Signed jumps, which are the difference between positive and negative jumps, thus may indicate whether the day is dominated by positive or negative news.

Motivated by the intuition that good (bad) news reduces (increases) future volatility, a large volume of research applies signed jumps to forecast volatility and most studies find they are significant (Sévi 2014; Patton and Sheppard 2015; Bee et al. 2016; Lyócsa and Molnár 2016; Wang et al. 2016; Todorova 2017; Gong and Lin 2021; Özbekler et al. 2021; Caporin 2023; Slim et al. 2023; Zhu et al. 2023). To provide more evidence of the impact of signed jumps on volatility forecasting, I follow Patton and Sheppard (2015) and define a specification by including the signed jump estimator and bipower variation,

$$RV_{t+1} = \beta_0 + \beta_{J^\Delta} J_t^\Delta + \beta_C \, BV_t + \beta_w \, \overline{RV}_{w,t} + \beta_m \, \overline{RV}_{m,t} + \epsilon_{t+1}. \tag{2.7.5}$$

I term this regression the J^Δ model in this chapter. If signed jumps are important, β_{J^Δ} should be significantly different from zero.

To determine whether the impact of positive jump variation on volatility forecasting differs from that of negative jump variation, and thus whether the impact of jumps is driven more by the positive or negative jump variation, I follow Patton and Sheppard (2015) and extend the J^Δ model by replacing the independent signed jump variable J_t^Δ with its positive and negative components, indicated by J_t^Δ $(J_t^\Delta > 0)$ and J_t^Δ $(J_t^\Delta < 0)$, respectively. This leads to the following $J^\pm$ model,

$$\begin{aligned} RV_{t+1} = \beta_0 + \beta_{J^+} J_t^\Delta (J_t^\Delta > 0) + \beta_{J^-} J_t^\Delta \, (J_t^\Delta < 0) + \beta_C \, BV_t + \beta_w \, \overline{RV}_{w,t} \\ + \beta_m \, \overline{RV}_{m,t} + \epsilon_{t+1}. \end{aligned} \tag{2.7.6}$$

If the two signed jump components have identical forecasting power, then we would expect to find $\beta_{J^+} = \beta_{J^-} = \beta_{J^\Delta}$.

To help evaluate the forecasting significance of signed jumps, I consider a benchmark model (named by the BV model) that excludes the signed jump estimators,

$$RV_{t+1} = \beta_0 + \beta_C\, BV_t + \beta_w\, \overline{RV}_{w,t} + \beta_m\, \overline{RV}_{m,t} + \epsilon_{t+1}. \qquad (2.7.7)$$

If signed jumps add predictive power, we would expect that this BV model is outperformed by the J^{Δ} and $J^{\pm}$ models explaining the variation of future volatility.

The above models can be estimated by the Ordinary Least Squares method (OLS) if the errors, ϵ_t, are independent, normally distributed, and have fixed volatility over sample days. However, the error term appears to exhibit a pattern of heteroscedasticity, changing across the sample period in accordance with the level of the volatility (Patton and Sheppard 2015), therefore, estimation by OLS has the disadvantage that the resulting estimates focus primarily on fitting periods of high volatility and place little weight on low volatility periods. Consequently, the OLS coefficient estimator is no longer efficient. This heteroscedasticity bias is also evidenced when applying OLS to the HAR model for my SPY sample, with the details of the results stored in the Appendix A.2. for presentation purposes.

To overcome this, I follow Patton and Sheppard (2015) and use a Weighted Least Squares (WLS) method to estimate the models in this chapter. The WLS method attempts to provide a more efficient alternative to OLS by putting different weights on errors. Specifically, the WLS method puts relatively less weight on errors which are likely to have a large variance and more weight on errors which are likely to have a small variance (note that the variances mentioned here indicate the variance of the errors, which is not the variance of the returns). As for the weights, I use the inverse of the fitted value of the HAR model estimated by the OLS method as in Patton and

Sheppard (2015). This idea is motivated by the positive relationship between the variance of residuals and the level of the fitted values of the HAR model estimated by the OLS method. The statistical inference on the coefficient estimates is based on the Newey–West Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors proposed by Newey and West (1987a). The full technical details describing the calculations of the WLS model estimation are provided in the Appendix A.2..

All estimators were calculated daily using returns sampled in business time instead of the more common calendar time. This means that instead of using prices spaced evenly in calendar time, I use prices spaced evenly in "event" time. Specifically, as opposed to applying SPY transaction prices every 5 minutes, I sample SPY prices every ten transactions. Consequently, I sample more frequently during high activity periods and less frequently during quieter times. Under certain conditions, business-time sampling can yield realized measures with better statistical properties (Oomen, 2005; Patton and Sheppard, 2015), and this method is now widely used in the literature (Bollerslev and Todorov, 2011). I sample prices 79 times a day, averaging an interval of 5 minutes. The first and last prices of the day are used as the initial and final observations, with the remaining 77 observations evenly spaced between them. The choice of an approximate 5-minute sampling window is standard and aims to avoid microstructure noise like bid-ask bounce. Since price observations are available more frequently than the approximate 5-minute sampling period, there are multiple possible grids of approximate 5-minute prices, depending on the initial sample observation.

I employ ten distinct grids of 5-minute prices to generate ten different estimators. These estimators are correlated but not identical. I then average these estimators to derive our final one. This subsampling method, introduced by Zhang et al. (2005), is

expected to enhance precision compared to using a single estimator. Additionally, I find that the evidence of drift biases in realized variance, bipower variation, semivariances and signed jump estimators is qualitatively very consistent with that reported in Figures 2.6.4 and 2.6.5. This is expected as the sampling interval of the returns in each subsample is approximately 5 minutes.

Table 2.7.1 reports the in-sample estimation results of the HAR model introduced by Corsi (2009), Equation (2.7.3), BV model, Equation (2.7.7), and the RS, J^{Δ}, and $J^{\pm}$ models proposed by Patton and Sheppard (2015), Equations (2.7.4) to (2.7.6), for daily realized variance forecast. Table 2.7.1 focuses on the daily volatility forecast. The forecast results associated with longer horizons are qualitatively very similar to the daily forecast results, with those results reported in the appendix A.3.. The first column reports the estimation results for the HAR model that exploits the impacts of lagged daily, weekly, and monthly volatility on volatility forecasting. The coefficients on daily, weekly, and monthly volatilities, indicated by β_d, β_w, and β_m, are all positive and significant at the 5% level, with $\beta_d + \beta_w + \beta_m$ approximately 1. This result reveals substantial volatility persistence as found in the literature (Andersen et al. 2007b; Corsi 2009; Corsi et al. 2010; Corsi and Renò 2012; Patton and Sheppard 2015; Bollerslev et al. 2016; Buccheri and Corsi 2021; Caporin 2023).

The second column contains the estimation results of the RS model, which divides the most recent volatility component into signed semivariances. The coefficients on weekly and monthly volatility lags, denoted by β_w and β_m, are significant at the 5% level. The very small, negative, and insignificant coefficient on positive semivariance, β_d^+ contrasts with the much larger, positive, and highly significant coefficient on negative semivariance, β_d^-. This estimation result suggests

that future volatility is more strongly related to recent bad volatility than to recent good volatility, which is consistent with the finding in Patton and Sheppard (2015).

Table 2.7.1. In-sample estimation results of different volatility models for daily realized variance forecast

Notes: The table provides in-sample parameter estimates and measures of fit for the volatility models with volatility and signed jump estimators for daily realized variance forecasts. The brackets report the Heteroskedasticity and Autocorrelation Consistent (HAC) t-statistics. HAR is a specification which uses the realized variance, Equation (2.7.3), RS, Equation (2.7.4), is based on semivariances, BV is a model that depends on bipower variation, Equation (2.7.7), J^{Δ} and $J^{\pm}$, are the models where the signed jump estimators are applied, Equations (2.7.5) and (2.7.6). The intercept of the model is not reported. The sample is SPDR S&P 500 ETF prices from January 1997 to September 2021.

	HAR	RS	BV	J^{Δ}	$J^{\pm}$
β_J^{Δ}				-0.498	
				(-6.67)	
β_J^{+}					-0.182
					(-2.40)
β_J^{-}					-1.480
					(-4.87)
β_d^{+}		-0.219			
		(-0.92)			
β_d^{-}		1.366			
		(15.80)			
β_C			0.586	0.573	0.479
			(14.85)	(16.10)	(9.50)
β_d	0.587				
	(14.23)				
β_w	0.303	0.310	0.289	0.287	0.295
	(6.48)	(7.75)	(6.14)	(6.97)	(7.49)
β_m	0.070	0.079	0.076	0.088	0.085
	(2.72)	(3.37)	(2.98)	(3.71)	(3.80)
R^2	0.502	0.568	0.507	0.551	0.586

Similar to Patton and Sheppard (2015), the semivariance model explains 13.2% more of the variation in future volatility than the model that includes only realized variance. The effect of lagged realized variance implied by this specification is $(\beta_d^{+} + \beta_d^{-})/2$, and we see that it is similar in magnitude to the coefficient found in the reference specification, where I include only lagged realized variance, which indicates that models that use only realized variance are essentially averaging the vastly

different effects of positive and negative returns. Consistent with Patton and Sheppard (2015), my testing results, which are not reported here, show that the null hypothesis that the coefficients of RV_t, RS_t^+, and RS_t^- are equal ($\beta_d = \beta_d^- = \beta_d^+$) is rejected, providing strong evidence that decomposing volatility into the good and bad volatilities significantly enhances the explanatory power of this model.

The third column presents the results for the BV model, which forecasts volatility using recent bipower variation, BV_t. The results are similar to the findings in the literature: the coefficient estimate on BV_t (β_w) is positive and highly significant, again indicating the strongly persistent volatility dependence. The fourth column in Table 2.7.1 contains the results for the J^{Δ} model that evaluates if there is incremental information including the signed jump variable. As found in Patton and Sheppard (2015), the coefficient on J_t^{Δ}, denoted by β_J^{Δ}, is negative and significant at 1% level. This leads to the conclusion that negative jumps lead to higher future volatility, while positive jumps lead to lower future volatility. Comparing the BV model with J^{Δ}, the latter model has an 8.7% larger adjusted R^2, indicating that including signed jumps explains more future volatility variation, which is consistent with the conclusion of Patton and Sheppard (2015).

The final column demonstrates the estimated coefficients of the $J^{\pm}$ model. This model uses the information from the decomposed components of J_t^{Δ}. The coefficient on bipower variation (β_C) is positive and significant at the 5% level, again indicating a strong volatility persistence. The coefficient on negative jumps (β_J^-) is significant at the 5% level and negative while that on the positive jumps (β_J^+) is significant and negative, with β_J^- larger than β_J^+ in magnitude. This leads to the conclusion that the positive and negative jumps have different impacts on volatility forecasting, with the

negative (positive) jumps leading to higher (lower) future volatility. Compared to the BV model which excludes any jump components, the $J^{\pm}$ model has a 15.6% larger adjusted R^2, indicating that incorporating the signed decompositions of jumps explains more volatility variation, similar to the findings by Patton and Sheppard (2015).

As in Patton and Sheppard (2015), the effects of volatility dependence and signed jumps on future volatility may help reconcile the forecasting impacts of good and bad volatility. The impact on future volatility of one-half integrated variation and positive jumps differs in signs, therefore the effect of the sum of these two components, which is the good volatility, may average out. The effects on future volatility of both one-half integrated variation and negative jumps are positive, causing the positive impact of bad volatility on future volatility.

The conclusions so far on the effects of the volatility persistence, asymmetric effect of good and bad volatility, and the effect of signed jumps made so far are entirely based on the realized variance, bipower variation, semivariances, and semivariance-based signed jump estimators. These estimators, however, as shown in my previous theoretical and simulation findings, may be subject to biases due to a nonzero drift. My earlier results also show that the modified versions of these estimators are much more robust to the impact of a nonzero drift. To explore the impact of drift on volatility persistence, the asymmetric effect of good and bad volatility, and the significance of signed jumps on volatility forecasting, I adjust all of these volatility models by substituting all estimators, including predictors, and forecast targets with their modified versions.

Table 2.7.2 shows the estimation results for the models associated with the modified volatility and signed jump estimators, termed HAR^*, RS^*, BV^*, $J^{\Delta *}$, and $J^{\pm *}$ models.

Table 2.7.2. In-sample estimation results of the modified volatility models for daily realized variance forecast

Notes: The table provides in-sample parameter estimates and measures of fit for the volatility models with modified volatility and signed jump estimators for monthly forecasts. The brackets report the Heteroskedasticity and Autocorrelation Consistent (HAC) t-statistics. HAR^* is a specification which uses the modified realized variance for Equation (2.7.3), RS^* modifies Equation (2.7.4) based on the modified semivariances, BV^* is an alternative specification of Equation (2.7.7) which depends on modified bipower variation, $J^{\Delta *}$ and $J^{\pm *}$ are the modified versions of Equations (2.7.5) and (2.7.6), where the modified signed jump estimators are applied. The intercept of the model is not reported. The sample is SPDR S&P 500 ETF prices from January 1997 to September 2021.

	HAR^*	RS^*	BV^*	$J^{\Delta *}$	$J^{\pm *}$
$\beta_J^{\Delta *}$				-0.089	
				(-0.73)	
$\beta_J^{+ *}$					-0.037
					(-0.20)
$\beta_J^{- *}$					-0.064
					(-0.25)
$\beta_d^{+ *}$		0.402			
		(2.50)			
$\beta_d^{- *}$		0.787			
		(3.98)			
β_C^*			0.580	0.588	0.583
			(14.21)	(13.43)	(10.95)
β_d^*	0.578				
	(13.74)				
β_w^*	0.313	0.308	0.300	0.297	0.297
	(6.45)	(5.89)	(6.07)	(5.69)	(5.52)
β_m^*	0.070	0.066	0.075	0.073	0.076
	(2.65)	(2.66)	(2.88)	(2.91)	(3.02)
R^2	0.496	0.500	0.499	0.500	0.500

The HAR^* model modifies the HAR model by including the daily, weekly, and monthly lags of modified realized variance. RS^* is the model that decomposes the daily lag of the HAR^* model into the modified positive and negative semivariances. BV^* is the regression that uses the modified bipower variation. $J^{\Delta *}$ is the model that

explores the impact of modified signed jumps on volatility forecasting, and $J^{\pm *}$ denotes the model that divides the modified signed jump estimator into positive and negative components.

As the table shows, all estimated coefficients and their t-statistics in the HAR^* and BV^* models are rather similar to those of HAR and BV reported in Table 2.7.1. This suggests that the volatility persistence effect is not influenced by the estimation inaccuracy of realized variance and bipower variation because of a nonzero drift. Of course, given very small discrepancies between the original and modified volatility estimators (below 1.2% on average), it is hardly surprising that these two versions of estimators perform fairly similarly from a volatility forecasting perspective.

From the estimation results of the RS^* model, I observe that the coefficient on modified positive semivariance, β_d^{+*} is positive and significant at the 5% level. Since the estimation of RS_t^{+*} is much more accurate than RS_t^+ in the presence of a nonzero drift, we can conclude that good volatility positively predicts volatility. The coefficient on the modified negative semivariance, β_d^{-*} is positive and significant at the 5% level. Given that the measurement of RS_t^{-*} is much more satisfactory than RS_t^- when drift deviates from zero, it then can be concluded that bad volatility also leads to higher future volatility.

From my testing results (not reported), the null hypothesis that the coefficients of RV_t^*, RS_t^{+*}, and RS_t^{-*} are equal ($\beta_d^* = \beta_d^{-*} = \beta_d^{+*}$) is not rejected a the 5% level. This result does not support the asymmetric effect of good and bad volatility. The modified semivariance model explains only 0.8% more of the variation in future volatility than the model that includes only modified realized variance, which indicates that the decomposition of volatility into good and bad components does not

add much information on volatility forecasting. This is quite different from the results of Patton and Sheppard (2015), who demonstrate that separating volatility into good and bad volatilities significantly improves volatility forecast.

The penultimate column of Table 2.7.2 reports the estimation results of the $J^{\Delta *}$ model that includes the modified signed jump estimator. The coefficient on the modified signed jumps estimator ($\beta_J^{\Delta *}$) is not significant at any reasonable level. Since $J_t^{\Delta *}$ is much more accurate in the presence of a nonzero drift, I conclude that signed jumps are not important to predict volatility. This result is very different from the previous studies, which conclude the significance of signed jumps in volatility forecasting based on the original signed jumps estimator (Patton and Sheppard 2015; Audrino and Hu 2016; Wang et al. 2016). By alleviating the drift-driven bias from this estimator, I find no significant impact for the signed jumps on future volatility.

The last column contains the estimation results of the $J^{\pm *}$ model that divides the modified signed jump estimator into its positive and negative components. The coefficients on both signed components ($\beta_J^{+ *}$ and $\beta_J^{- *}$) are not significant, leading to the conclusion that neither positive nor negative jumps impact volatility, when the modified estimator is much less sensitive to a nonzero drift relative to its original counterpart. This finding is very different from Patton and Sheppard (2015), who find both positive and negative jumps are important, with positive and negative jumps decreasing and increasing volatility, respectively. By reducing the impact of drift on the measurement precision of signed jumps, I find that both signs of signed jumps have limited impacts on volatility forecasting.

This lacking evidence on signed jumps may reconcile with my previous forecasting findings of good and bad volatility: the impacts of good and bad volatility

on future volatility are not statistically different from each other. Since both positive and negative jumps have insignificant impacts on future volatility, it is very unlikely that they perform very differently for volatility forecasting when they are added with a common one-half integrated variation component.

The insignificant impact of signed jumps found in this chapter is quite counter-intuitive. The positive and negative jumps are expected to decrease and increase volatility given that positive and negative jumps may be often associated with good and bad macroeconomic news, respectively (Evans 2011; Gilder et al. 2014; Caporin et al. 2017). For the weak evidence of signed jumps, I might give an interpretation, following Caporin (2023)'s arguments. Since the S&P 500 ETF is highly liquid, the information conveyed by negative jumps provides an increase in the intraday volatility just after the jump events and by the end of the day. Consequently, these negative jumps produce an increase in the continuous volatility component within the same day as this jump. Then, the persistence induced by the recent negative jumps on the future realized variance is partly absorbed by the recent volatility component, thus leading to a contraction of the role of the negative jumps. Analogously, positive jumps cause a decrease in the continuous volatility component within the same day of these jumps. Since volatility is persistent, lower recent volatility leads to lower future volatility, thus reducing the significance of the recent positive jumps.

But, why does the original signed jump estimator have a negative impact on future volatility? The interpretation starts from the "leverage effect" theory by Black (1976). As the price of a stock drops, the value of the equity decreases while the debt remains constant. This increases the leverage ratio, which in turn raises the risk, and thus the volatility of the stock. When stock prices increase, the leverage ratio becomes smaller, leading to lower volatility. Therefore, we have a negative return-volatility

relationship, which has also been well-established in the finance literature. According to the definition of drift, a nonzero drift indicates the move of the mean of price. This suggests that drift can drive the daily return. Therefore, drift should also be negatively correlated with future volatility. Indeed, this relationship is empirically evidenced by my results (not reported) that the correlation between the daily median with the one-step-ahead realized variance is -0.225 and significant at the 5% level. As shown in my theoretical and simulation findings, the drift bias in the original signed jump estimator is positively correlated with the drift. Therefore, this drift bias may also be negatively related to volatility. From my empirical results, drift biases often dominate the original signed jump estimator, thus explaining the negative impact of the original signed jump estimator on volatility forecasting. This indicates that the negative effect of signed jumps found in the literature may be almost exclusively due to the drift-related bias of the signed jump estimator.

Why do positive and negative components of the original signed jump estimator lead to lower and higher future volatility, respectively? Given the leverage effect of the drift bias in the original signed jump estimator, the positive drift bias decreases volatility while the negative bias increases volatility. As indicated by the results in Figure 2.6.5, the positive and negative parts of the original signed jump estimator are often dominated by the positive and negative drift bias of this estimator, respectively. This suggests that the volatility forecasting effects of the positive and negative components of the original signed jump estimator may be artificially influenced by the leverage effect of its positive and negative drift bias on volatility.

2.8. Conclusion

Although the drift component is asymptotically ignorable in the log price dynamic, Laurent and Shi (2020) find that a nonzero drift can generate an influential estimation bias of realized variance and bipower variation for a realistic sample of high-frequency log prices. Good and bad volatility have important implications for investors to identify the upside and downside risks, but Laurent and Shi (2020) somehow have not yet investigated the finite sample estimation bias of good and bad volatility. This chapter investigates the effect of drift on the finite sample estimation of good and bad volatility. For studying the measurement of good and bad volatility in a high-frequency finite sample setting, I consider the most popular estimators in the literature: positive and negative realized semivariance. My finite sample theory, together with extensive simulations, show that the semivariance overestimates good or bad volatility if the signs of the drift and the semivariance are equal, and the semivariance underestimates the good or bad volatility if the signs of the drift and the semivariance differ. Importantly, I find that the drift component causes an even more severe estimation bias in realized semivariances than in realized variance and bipower variation as found by Laurent and Shi (2020). Moreover, these properties of the estimation bias of realized semivariances do not alter for a version of semivariances that is robust to contamination of the microstructure noise.

An important application of the semivariances is to estimate signed jumps. The literature commonly estimates the signed jumps by the difference between the positive and negative semivariances. My theoretical and simulation analyses also show that this signed jump estimator overestimates signed jumps when drift is positive and underestimates when drift is negative. The results also show that this signed jump estimator aggregates the drift biases in positive and negative, thus

exhibiting even more unsatisfactory estimation performance in finite samples than the semivariances.

To alleviate the biases of the semivariances and signed jump estimators due to a nonzero drift, I modify these estimators by applying the returns centred by the daily median, with this centred-return approach proposed by Laurent and Shi (2020). My finite sample theory for the alternative semivariances along with simulations uncover significant improvements in the estimation performance of good and bad volatility and the signed jumps. The modified semivariance and signed jump estimators, together with their original versions, are applied to the 5-minute returns on the SPDR S&P 500 ETF for the period from 1997 to 2021. I observe a positive discrepancy between the original and modified semivariance when the signs of the drift and the semivariance are equal, and a negative discrepancy when the signs of the drift and the semivariance differ, which is consistent with my theory and simulations. Furthermore, the discrepancy between the original and modified signed jump estimator is positive if the drift is positive, and the discrepancy between the original and modified signed jump estimator is negative if the drift is negative.

To provide more evidence on the effect of volatility persistence, good and bad volatility, and signed jumps on volatility forecasting, I applied the realized variance, bipower variation, semivariances, and signed jump estimators to predict future realized variance for the SPY sample. My results demonstrate that recent realized variance and bipower variation lags significantly increase future realized variance, which supports the well-known volatility dependence property. The results also reveal that the positive semivariance almost has no impact on volatility while the negative semivariance leads to significantly higher volatility, suggesting the asymmetric impact of good and bad volatility on volatility forecasting, similar to the findings of previous

studies. Finally, I show that positive jumps decrease volatility and negative jumps increase volatility, with the impact of negative jumps stronger, which is also consistent with the literature.

As discussed in the earlier part of this chapter, the realized variance, bipower variation, semivariance, and signed jump estimators may be subject to the bias in the presence of a nonzero drift while their modified versions are much less sensitive to a nonzero drift. To investigate the impact of drift on the forecasting results, I also apply the modified estimators to volatility forecasting. The modified realized variance and bipower variation exhibit very similar volatility persistency effects as their original counterparts, indicating that the biases in these volatility estimators are not impactful from a volatility forecasting perspective.

However, the results of the modified positive and negative semivariances do not suggest the asymmetric impact of good and bad volatility on volatility forecasting. This finding is further reconciled by the limited value of signed jumps and their positive and negative components in terms of predicting volatility, found in later analysis using the modified signed jump estimator. One possible explanation for the weak evidence of signed jumps is that the role of recent positive or negative jumps in forecasting volatility is absorbed by the recent volatility. Additionally, I also show that the asymmetric impact of the original signed semivariances and the significance of the original signed jump estimator on volatility forecasting may be artificially caused by the leverage effect on future volatility of the drift-driven bias in these estimators. suggesting that the asymmetric effect of good and bad volatility and signed jumps found in the existing literature is almost exclusively due to the impact of the drift biases in the estimators of these variables.

Chapter 3. Drift bursts, volatility forecasting, and the variance risk premium

Abstract

The conventional semi-martingale process assumes that the price trend is indicated by the drift component. The constant drift and Ornstein–Uhlenbeck models defined in Chapter 2 assume that drifts are evenly distributed intraday and thus are unable to explain the undeniable stylized fact of occasional short-lived explosive price trends, also known as drift bursts, such as gradual jumps and flash crashes in the stock markets. Chapter 3 studies these intraday drift bursts. I discover a large contribution of approximately 20% from drift burst variation to daily price variation. Applying the mutually-exciting Hawkes process, I find that both positive and negative bursts exhibit self-exciting behaviours and that negative drift bursts impact positive drift bursts, but not vice versa. I observe that a lack of evidence of drift bursts on realized variance contrast with their importance in affecting implied variance, which leads to the role of drift bursts in explaining the variance risk premium. Although consistent with the findings documented in the previous literature that large downside price moves raise investors' risk aversion, I am the first to document this for drift bursts.

3.1. Introduction

The conventional semi-martingale process assumes that the price trend is determined by the drift component. Recently, there has been growing consensus that intraday explosive drifts (trends) such as gradual jumps and flash crashes or spikes are stylized facts in financial markets. The gradual jumps denote that the price experienced a rapid descent or ascent, comprised of many small downward or upward moves, but no apparent price discontinuities, known as jumps. A flash crash is a downside and explosive trend in asset prices occurring within a few minutes followed by a quick recovery, which is also termed the "V" shape price pattern by Flora and Renò (2020). A flash spike is an inverse "V" shape price pattern with an upside and explosive trend followed by a quick price drop. One of the most infamous flash crashes occurred on May 6, 2010, in the US stock market, where the S&P 500 index collapsed for 5 minutes from 2:45 PM, resulting in a 4.5% loss, then the price persistently increased and regained most of the 450-point drop for the next 5 minutes.

The very existence of flash crashes or spikes and gradual jumps casts doubt on the orderliness of the financial market architecture. This has an adverse impact on investor confidence and raises concerns of regulators, who view the orderly functioning of financial markets as their first and foremost objective (CFTC and SEC 2010, 2011; SEC 2015). The constant drift and Ornstein–Uhlenbeck models defined in Chapter 2 are not sufficient to uncover these two classes of explosive drift or drift burst patterns as both models assume the price drifts are smoothly distributed in each interval over a whole trading session. To identify the existence of such price patterns, I follow the drift burst hypothesis proposed by Christensen et al. (2022). The drift burst hypothesis assumes that drift can be exceedingly large for a short intraday

interval, thus capturing the explosive trend behaviour of flash crashes and gradual jumps.

I start the analysis by exploring three empirical research questions to better understand the drift bursts. How often do drift bursts occur? What are the price patterns that are related to drift bursts? Do drift bursts identify gradual jumps and flash crashes? My sample data is the S&P 500 E-mini future transaction prices from June 2, 2003, to December 30, 2020. To explore the frequency of the drift bursts, I use the test proposed by Christensen et al. (2022) to identify the occurrence of drift bursts. This drift burst statistic, to the best of my knowledge, is the only test that detects the drift bursts in the literature. This statistic has been shown by Christensen et al. (2022) to have good size and power properties in their simulation study. Moreover, the drift burst test statistic is robust to volatility spikes and broad forms of stochastic jumps defined in the literature (Christensen et al. 2022), thus, picking up neither random jumps nor large volatility episodes, but genuine explosive trends.

My testing results of the E-mini futures data show that drift bursts are rare events, which occur about 1.5 times per month on average, which is consistent with Christensen et al. (2022). My results also show that drift bursts can be decomposed into positive and negative components, with the frequency of the negative burst occurrence approximately double that of the positive burst occurrence. I find almost all of these bursts are associated with gradual jumps: 99% of positive bursts and 98% of negative bursts are with positive and negative gradual jumps, respectively, and only 1% of positive bursts and 2% of negative bursts are with flash crashes or flash spikes, according to the classification criteria defined in the literature (Nanex, 2010). My finding that only a small proportion of drift bursts are related to flash crashes and spikes does not indicate the unsatisfactory performance of drift burst tests in

identifying these two price patterns, but may be due to the extreme scarcity of flash crashes and spikes relative to gradual jumps as found in the literature (Christensen et al. 2022; Jagannathan et al. 2022; Andersen et al. 2023).

Recent studies show that gradual jumps are often caused by economic news (Andersen et al. 2023; Bouamara et al. 2023). Since drift bursts are mostly related to gradual jumps, the main cause of drift bursts could be economic news. From my observations of the intraday price patterns that are identified as drift bursts, I observe that positive and negative bursts are associated with extremely large upside and downside price changes, respectively. For gauging the contribution of intraday drift bursts relative to the daily price variation, I propose a daily variation measure of drift bursts derived from the drift burst test statistic. Based on this measure, I find that 20% of daily price variation is attributed to drift bursts, on average, in the presence of drift bursts. My results also demonstrate that drift bursts differ in magnitude over the sample, and their level is much higher during the 2008 financial crisis and 2020 pandemic recessions.

According to my visual inspections of the distribution of the drift burst occurrences over my sample length, drift bursts exhibit clustering characteristics, especially during economic declines. I consider modelling the bursts' clustering behaviours using the popular mutual-excitation Hawkes processes (HPs) proposed by Hawkes (1971). Because drift bursts may be associated with large price moves, modelling their dynamics has important implications for risk management, asset pricing, and understanding of the behaviour of markets. The HPs exploit the clustering effects of positive or negative bursts by allowing the occurrence probability, also termed the intensity, of positive or negative bursts to be positively affected by the past arrivals of both positive and negative bursts. In other words, the

HPs let the past positive and negative arrivals mutually "excite" the future positive or negative arrivals. Based on the estimation results of these HPs, I discover that both positive and negative bursts strongly self-excite. I also find that positive and negative bursts mutually excite each other in an asymmetric fashion: positive bursts are excited by the negative bursts, but not vice versa.

Previous studies tend to apply mutual-excitation HPs to model the clustering effect of jumps but they have not yet investigated such modelling for the drift bursts (Bowsher 2007; Large 2007; Bauwens and Hautsch 2009; Aït-Sahalia et al. 2015; Clements and Liao 2017; Ma et al. 2019). My analysis applying HPs to drift bursts helps model the clustering behaviours of gradual jumps. As far as I am aware, my findings of the prominent self-excitation effect of both positive and negative bursts, along with the asymmetric mutual-excitation behaviour between positive and negative bursts are novel in the literature.

Since drift bursts could be related to large price movements, they might exhibit the leverage effect on future volatility, with negative bursts increasing volatility and positive bursts lowering volatility. Anticipating the magnitude of volatility is important to asset pricing (Heston and Nandi 2000; Ghysels et al. 2005; Christensen and Nielsen 2007; Christoffersen et al. 2008; Goyal and Saretto 2009; Corsi et al. 2013) and risk management (Christoffersen and Diebold 2000; Clements et al. 2008; Maheu and McCurdy 2011). Making precise volatility forecasts is of the utmost importance for many practitioners and regulators of the financial markets.

I explore the effects on volatility forecasting of all three measures of positive and negative bursts investigated in this chapter including occurrence, variation and intensity. Surprisingly, the results show that both positive and negative bursts generally have little effects on future volatility across daily, weekly and monthly

forecast horizons and for these three burst measures. This weak evidence of bursts on volatility forecasting is virtually unchanged for the most popular ex-post volatility estimators including realized variance (RV) in Andersen and Bollerslev (1998), Bi-power variation (BV) in Barndorff-Nielsen and Shephard (2004), and a drift burst-robust Rice variance ($RiceV$) measure advocated by Laurent et al. (2022c). The results that the drift burst have an insignificant impact on future volatility are puzzling since positive and negative drift bursts, which are associated with positive and negative large returns, are expected to exhibit a negative and positive effect on volatility due to the leverage effect.

As far as I am aware, my study on whether drift bursts help in predicting realized variance is related to Laurent et al. (2024). They find the significance of drift bursts in realized variance forecasting, which contrasts with my weak evidence of drift bursts in explaining future realized variance. However, they do not test for drift bursts, thus their results may be influenced by exposing to large false positives of drift burst testing or the probability that we incorrectly indicate the presence of drift bursts. This chapter focuses exclusively on the statistically significant drift bursts and finds weak evidence of drift bursts in predicting realized variance, which are quite different from their conclusions.

Because realized variance is calculated on high-frequency returns, it is inevitably subject to the bias due to microstructure noise. As suggested by Andersen and Bollerslev (1998), the option-implied variance is not affected by microstructure noise in high-frequency return as it is derived from daily options prices. Therefore, I also consider forecasting implied variance. Implied variance is a measure of the stock market's expectation of volatility based on S&P 500 index options. Thus, implied variance is different from any of the above Ex-post RV, BV, and $RiceV$ volatility

measures. Implied variance is an important indicator that measures the fear of the investors and is an essential factor for option pricing. This chapter studies whether drift bursts have an impact on implied variance estimate: the (squared) Chicago Board Options Exchange's Volatility Index, indicated by VIX^2. VIX is often referred to as the fear gauge as if something concerns the market, traders and investors tend to start buying OTM options, pushing up the option prices and thus VIX. My results show that for all three types of burst variables, negative bursts lead to a significantly higher VIX at the close of that day while the effect of positive bursts is limited. The significance of negative bursts is consistent with the intuition that a large price drop increases the concerns of investors thus increasing the VIX.

The variance risk premium (VRP) is defined by the difference between the implied variance and realized variance. VRP is found to be positive, on average, in the literature, suggesting that investors are willing to pay extra money to hedge against the increase in future volatility. VRP has been extensively studied in financial studies and has been found to have important implications for volatility prediction and asset return forecasting (Bollerslev et al. 2009; Carr and Wu 2009; Arisoy 2010; Bollerslev et al. 2014; Prokopczuk and Simen 2014; Feunou et al. 2018; Li and Zinna 2018; Kilic and Shaliastovich 2019; Pyun 2019). I investigate the impacts of all three drift burst measures on VRP, motivated by the intuition that large negative returns may increase investors' risk aversion, causing a higher VRP. The results show that all variables of negative bursts significantly increase VRP and the respective positive burst variables insignificantly affect VRP, which corroborates that larger negative returns increase risk aversion. Additionally, I also study the impacts of these burst variables on VRP when the threshold of the drift burst test statistic increases. I find that positive drift burst variation significantly decreases VRP at higher thresholds,

which contrasts with the insignificant impact of this variable on VRP at lower thresholds. The significance of negative burst variation under stricter thresholds may be because extremely large negative returns can strongly reduce the risk aversion of investors.

My results of the significance of drift bursts in explaining VRP are close to the studies which find that VRP is dominated by large price moves in the stock markets (Todorov 2010; Caporin et al. 2017). The literature also documents other factors that may influence VRP, including broker dealers' funding liquidity (Adrian and Shin 2010), monetary policy (Bekaert et al. 2013), certain macro variables (Bollerslev et al. 2011), and business conditions (Corradi et al. 2013). A possible extension of this chapter could be to control these variables and see how drift bursts perform in predicting VRP.

The remainder of this chapter is organized as follows. Section 3.2 introduces the theoretical background of drift bursts, and the drift burst test. Section 3.3 presents the descriptive analysis of drift bursts detected from my S&P 500 E-mini future price sample. Section 3.4 reports the results of modelling the clustering features of drift bursts via the mutual-excitation HPs. Section 3.5 presents the results of the impacts of drift bursts on realized and implied variance forecasting. Section 3.6 contains the analysis of the effects of drift bursts on the variance risk premium. Section 3.7 concludes.

3.2. The econometrics of drift bursts

To study the role of drift bursts in anticipating volatility for the stock markets, this section introduces the theoretical background of stock prices. Following the conventions of the literature, I assume the asset logarithmic prices p_t follows the Itô semi-martingale process below,

$$dp_t = \mu_t d_t + \sigma_t dW_t + J_t, t \in [0, T], \tag{3.2.1}$$

where T is the fixed time duration (e.g., $T = 1$ for one day), μ_t is the drift, σ_t is a strictly positive cádlág process, W_t is a standard Brownian motion, and J_t are Poisson jumps process. I assume that the log price p_{t_i} is recorded at times $0 = t_0 < t_1 < \cdots < t_n = T$ for day t, with n indicating the number of price observations, and that the interval between neighbouring time points is always equal.

As mentioned in the introduction, flash crashes and gradual jumps in the stock markets reflect short-lived directional log price movements. Christensen et al. (2022) show that neither stochastic volatility, $\sigma_t dW_t$, nor jumps, J_t, can explain short-lived locally explosive trends in the price paths of financial assets. An intuitive explanation is that stochastic volatility, $\sigma_t dW_t$, or the jump process, J_t, due to their random natures, merely lead to wider price dispersion and thus are unable to reconcile the persistent direction of price evolution over episodes of locally explosive trends. To postulate the pattern of short-lived locally explosive price trends, Christensen et al. (2022) assume drift explodes locally, which they term drift bursts. More specifically, drift, μ_t, is exceedingly large within a short intraday interval before a time point τ_{db}. To model the explosion of drift, they assume μ_t follows an exponential function with time (t) as the input:

$$\mu_t = a(\tau_{db} - t)^{-\alpha}, t \in [\tau_{db} - t_\vartheta, \tau_{db}], \quad (3.2.2)$$

where a is a constant, $1/2 < \alpha < 1$, and $[\tau_{db} - t_\vartheta, \tau_{db}]$ denotes a short intraday interval. If $a > 0$, the exponential function, Equation (3.2.2) allows drift to explode positively over the interval $[\tau_{db} - t_\vartheta, \tau_{db}]$ while if $a < 0$, drift negatively explodes for this interval. Motivated by the fact that the strongly directional price move is also very volatile, Christensen et al. (2022) assume volatility exponentially explodes at the same time,

$$\sigma_t = b|\tau_{db} - t|^{-\beta}, t \in [\tau_{db} - t_\vartheta, \tau_{db}], \quad (3.2.3)$$

with $b > 0$, $0 < \beta < 1/2$. Christensen et al. (2022) suggest that the exponential function of spot volatility, as described in Equation (3.2.3), could have a smaller power than that of spot drift such that $\alpha - \beta > 1/2$. In this scenario, volatility bursts less violently than the drift. The motivation behind this restriction is that the price movement of crashes or gradual jumps, although associated with a large level of volatility, is still highly directional, that is, dominated by the drift of the price.

The dramatic importance of this setup is that the drift standardized by volatility (μ_t/σ_t) tends to infinity when approaching the time point τ_{db}, allowing the construction of a statistical test for the presence of a drift burst. To exploit this convergence property of the drift-volatility ratio for drift burst detection in the log price process, Equation (3.2.1), Christensen et al. (2022) first propose the estimates of spot drift and volatility at time point t_i based on the log prices p_{t_i} recorded at times $0 = t_0 < t_1 < \cdots < t_n = T$ for day t. Specifically, they define the measure of spot drift at t_i by,

$$\hat{\mu}_{t_i}^n = \frac{1}{h_n}\sum_{j=1}^{n} K\left(\frac{t_{j-1}-t_i}{h_n}\right) r_{t_{j-1},k_n}, \tag{3.2.4}$$

where $\hat{\mu}_{t_i}^n$ is the spot drift estimate with h_n indicating the bandwidth of this estimate, $K(\cdot)$ is a backwards-looking exponential kernel, $K(x) = \mathrm{e}^{-|x|}$, for $x \leq 0$. And Christensen et al. (2022) define the spot volatility as,

$$\hat{\sigma}_{t_i}^n = \left[\frac{1}{h_n'}\sum_{j=1}^{n} K\left(\frac{t_{j-1}-t_i}{h_n'}\right)\left(r_{t_{j-1},k_n}\right)^2\right]^{1/2}, \tag{3.2.5}$$

where $\hat{\sigma}_{t_i}^n$ denotes the spot volatility estimate and h_n' is the bandwidth for this volatility measure. Note that the notation of the bandwidth of volatility estimate, h_n', is different from that of the drift, indicated by h_n as these two bandwidths are not necessarily equal. Christensen et al. (2022) show that a test statistic calculated on the ratio of the spot drift and volatility estimates can detect whether there is a drift burst in the price process between the interval $[t_{i-h_n}, t_i]$, where h_n denotes the bandwidth of spot drift estimate as opposed to that of spot volatility measure,

$$Z_{t_i}^n = \sqrt{h_n}\frac{\hat{\mu}_{t_i}^n}{\hat{\sigma}_{t_i}^n}. \tag{3.2.6}$$

Christensen et al. (2022) show that in the absence of drift bursts near t_i, the test statistic $Z_{t_i}^n$ will converge in probability to a standard normal distribution,

$$Z_{t_i}^n \xrightarrow{d} N(0,1). \tag{3.2.7}$$

In the presence of positive (negative) drift bursts near t_i, the test statistic will converge almost surely to a positive (negative) infinite value,

$$Z_{t_i}^n \xrightarrow{a.s.} \pm\infty. \tag{3.2.8}$$

Therefore, a positive intraday drift burst is identified within the interval $\left[t_{i-h_n}, t_i\right]$ if $Z_{t_i}^n$ exceeds a positive threshold value, $Z_{t_i}^n > \Phi$, where Φ is a threshold value that is obtained from the normal inverse cumulative distribution function and should ensure the false positives or the probability of detecting a drift burst using the test statistic when there is actually no drift burst, is extremely close to zero. A negative intraday drift burst is detected this interval if $Z_{t_i}^n < -\Phi$. This chapter follows Christensen et al. (2022) for the parameter settings of bandwidths h_n and h_n' and the threshold Φ: $h_n =$ 5-minute, $h_n' = 25$-minute, threshold value $\Phi \geq 4$.This threshold $\Phi = 4$ is the minimum threshold found by Christensen et al. (2022) that can ensure the false positives of the burst detection are extremely close to zero.

These settings of bandwidths imply that $Z_{t_i}^n$ detects 5-minute positive intraday drift bursts that last for $[t_{i-5\text{min}}, t_i]$ if $Z_{t_i}^n > \Phi$ while $Z_{t_i}^n$ detects a 5-minute length of negative intraday drift bursts for this interval if $Z_{t_i}^n < -\Phi$. This suggests that I attempt to detect the drift bursts that last for 5 minutes but does not necessarily imply that my data is 5-minute frequency.

Additionally, following the simulation study of Christensen et al. (2022), the test statistic $Z_{t_i}^n$ has satisfactory statistical performance in detecting the 5-minute length of the intraday drift bursts using 1-second frequency log price data. Therefore, I use 1-second frequency log prices for computing the test statistic $Z_{t_i}^n$. Following Christensen et al. (2022), I record the calculated $Z_{t_i}^n$ every five-seconds for a trading day and allow

at most only one burst to be detected over a rolling 5-minute window that moves along by 5 seconds, where the test statistic $Z_{t_i}^n$ attains a local extremum and exceeds the threshold.

For ultrahigh-frequency data (e.g., 1-second), the true, efficient log price p_t, Equation (3.2.1) is not available from the stock markets because transaction and quotation data are contaminated by multiple layers of ‘‘noise’’ or ‘‘friction’’ (Black 1986; Stoll 2000). Only the noisy log price is observable from the data,

$$p_{t_i}^{\diamond} = p_{t_i} + u_{t_i}, \text{with } 0 < t_i < T, \tag{3.2.9}$$

where $p_{t_i}^{\diamond}$ is the noisy price, u_{t_i} is a q-dependent noise and $q \geq 0$ is a positive constant. The spot drift and volatility estimates calculated on the observed 1-second log price $p_{t_i}^{\diamond}$ using Equations (3.2.4) and (3.2.5) are no longer accurate due to the existence of noise process u_{t_i} in the log price $p_{t_i}^{\diamond}$.

To mitigate the adverse impact of microstructure noise on the estimation precision of spot drift and volatility, I follow Christensen et al. (2022) and modify the spot drift and volatility estimators by applying the preaveraging technique. Intuitively, the preaveraging technique locally smooths the observed price series $p_{t_i}^{\diamond}$ so that the microstructure noise component u_{t_i} (almost) disappears under averaging. The modified spot drift measure is:

$$\hat{\mu}_{t_i}^{n*} = \frac{1}{h_n} \sum_{j=1}^{n-k_n+2} K\left(\frac{t_{j-1} - t_i}{h_n}\right) r_{t_{j-1},k_n}^{*}, \tag{3.2.10}$$

where $\hat{\mu}_{t_i}^{n*}$ is the modified spot drift and r_{t_i,k_n}^{*} denotes the preaveraged return, $r_{t_i,k_n}^{*} = \frac{1}{k_n}\left(\sum_{j=k_n/2}^{k_n-1} p_{(t_i+t_j)/N}^{\diamond} - \sum_{j=0}^{k_n/2-1} p_{(t_i+t_j)/N}^{\diamond}\right)$ with the $k_n = 3$ indicating the preaveraging window size. The modified spot volatility estimator is defined by:

$$\hat{\sigma}_{t_i}^{n*} = \left[\frac{1}{h_n'}\sum_{j=1}^{n-k_n+2}\left(K\left(\frac{t_{j-1}-t_i}{h_n'}\right)r_{t_{j-1},k_n}^{*}\right)^2 + 2\sum_{L=1}^{L_n}\omega\left(\frac{L}{L_n}\right)\sum_{j=1}^{n-k_n-L+2}K\left(\frac{t_{j-1}-t_i}{h_n'}\right)K\left(\frac{t_{j+L-1}-t_i}{h_n'}\right)r_{t_{j-1},k_n}^{*}r_{t_{j-1+L},k_n}^{*}\right]^{1/2}, \tag{3.2.11}$$

where $\hat{\sigma}_{t_i}^{n*}$ is a heteroscedasticity and autocorrelation consistent (HAC)-type estimate of spot volatility, the second term of $\hat{\sigma}_{t_i}^{n*}$ is required to account for any noise dependence and the serial correlation induced by the preaveraging method. The lags for the HAC $L_n = 2(k_n - 1) + 10$ following Christensen et al. (2022), and $\omega(\cdot)$ is a Parzen kernel,

$$\omega(x) = \begin{cases} 1 - 6x^2 + 6|x|^3, \text{for } 0 \le |x| < 1/2, \\ 2(1-|x|)^3, \text{for } 1/2 < |x| \le 1, \\ 0, \text{otherwise.} \end{cases} \tag{3.2.12}$$

As found by Christensen et al. (2022) and Bellia et al. (2023), the drift burst test statistic is robust to compound Poisson jumps, infinite activity small jumps, autocorrelated market microstructure noise and pre-announced jumps. The pre-announced jumps typically refer to significant price movements in a stock or other financial instrument that are expected to occur due to pre-announced events. These events can include earnings reports, economic data releases, or other significant news that is scheduled to be released at a specific time. Further, the test is robust to volatility spikes (Christensen et al. 2022). Thus, the drift burst statistics measure is

picking up neither jumps nor large volatility episodes, but genuine directional price trends.

Based on the detection of intraday drift bursts via the test statistics $Z_{t_i}^n$, I calculate the number of intraday positive drift bursts (NDB_t^+) on day t:

$$NDB_t^+ = \sum_{i=1}^{n-k_n+2} DB_{t_i}^+, \tag{3.2.13}$$

where $DB_{t_i}^+ = I\left(Z_{t_i}^n > \Phi\right)$ with $I\left(Z_{t_i}^n > \Phi\right)$ denoting the dummy which equals one if the statistic exceeds the positive threshold at the time point t_i. The number of intraday negative drift bursts on day t:

$$NDB_t^- = \sum_{i=1}^{n-k_n+2} DB_{t_i}^-, \tag{3.2.14}$$

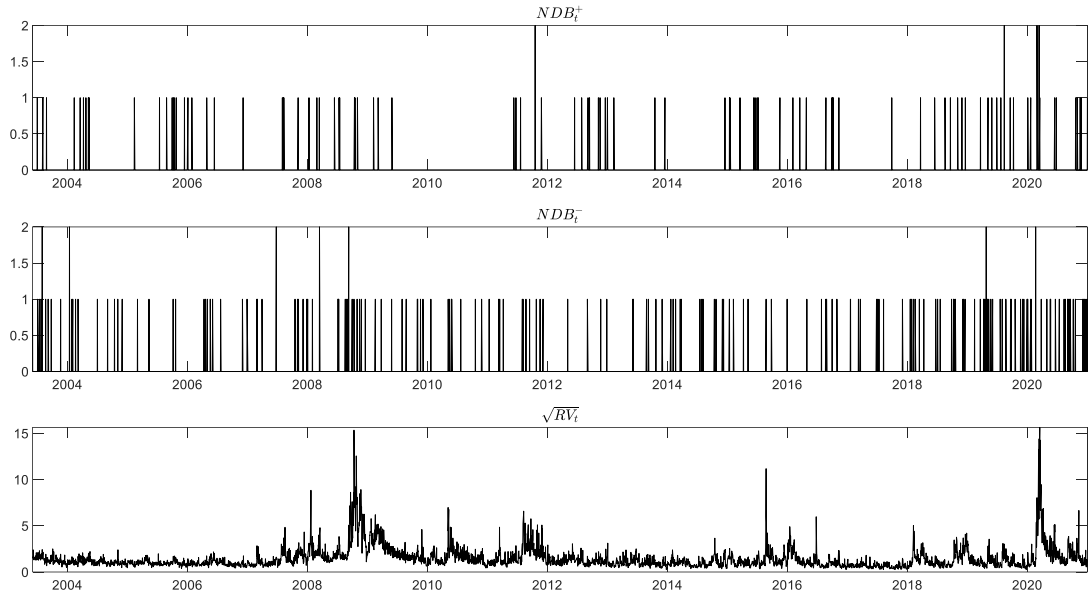
where $DB_{t_i}^- = I\left(Z_{t_i}^n < -\Phi\right)$ with $I\left(Z_{t_i}^n < -\Phi\right)$ indicating the dummy which equals one if the statistic is lower than the negative threshold at the time t_i.

3.3. Drift burst detection results

The data is based on the transaction prices of the S&P500 E-mini futures (ticker, ES), collected at the 1-second frequency, and is provided by Tick Data Inc. Applying the drift burst statistics to 1-second resolution data is consistent with the simulation study of Christensen et al. (2022), who show that the drift burst statistics have a satisfactory performance in detecting drift bursts in 1-second frequency data. The sample period length is $T_N = 4420$ trading days over 17 years, from June 2, 2003, to December 30, 2020. To ensure that the sample consists of regular trading days, I remove all non-business days and days on which the exchange closed early. This chapter focuses on the active trading session of the US stock market, thus all trades before 9:30 EST or after 16:00 EST are discarded.

The test detects a total of 313 bursts out of the SPY sample, with 207 negative bursts and 106 positive bursts. This corresponds to only 1.51 drift bursts per month on average, which indicates that drift bursts are rare events. To illustrate the distribution of the occurrence of drift bursts across my sample period, the first two panels of Figure 3.1 depict the daily number of intraday positive and negative bursts (NDB_t^+ and NDB_t^-), respectively. The bottom panel of this figure contains the daily realized variance RV_t, which is defined by the sum of 5-minute squared returns from 9:30 to 16:00 EST. I report the square root version of RV_t for presentation purposes. From the top panel, I observe the clustering effects of positive bursts. For example, positive bursts are absent from 2010 to 2011 but occur 6 times for the sample period from 2011 to 2012. I also observe some clusters of negative bursts in the middle panel. There are some days with multiple negative bursts around the 2008-2009 subprime mortgage crisis and 2020 pandemic periods, over which the level of realized variance is much higher than usual.

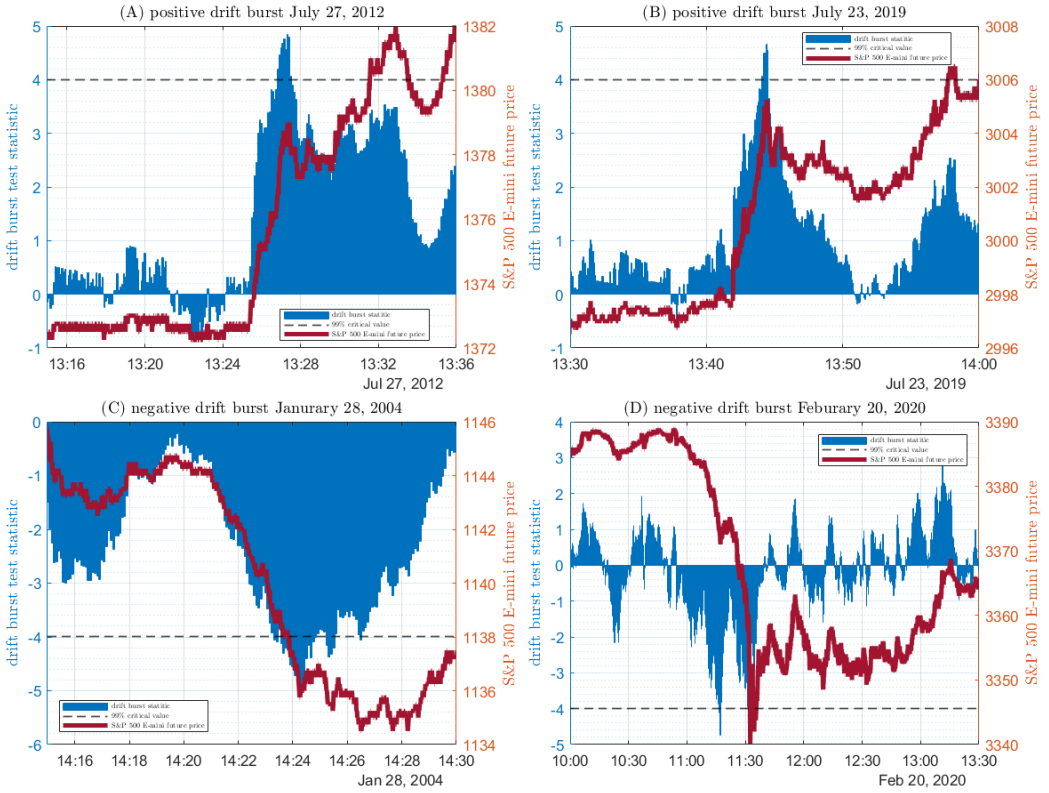
Figure 3.3.1. Daily time series of the number of intraday drift bursts and realized variance



Notes: This figure reports the time series of the daily number of positive intraday drift bursts (NDB_t^+), Equation (3.2.13), the daily number of negative intraday drift bursts (NDB_t^-), Equation (3.2.14), and daily realized variance (RV_t), defined by the sum of 5-minute squared returns from 9:30 to 16:00 EST. The realized variance is in the square root form for presentation purposes. The sample data is S&P 500 E-mini prices from June 2, 2003, to December 30, 2020.

Having investigated the daily dynamic of burst occurrences, it is also noteworthy to have a look at the details of the intraday price dynamic associated with bursts. This analysis allows us to observe how bursts influence price movements. Figure 3.3.2. depicts four randomly selected examples of intraday drift bursts. Panels (A) and (B) contain two examples of the prices related to positive bursts while the lower Panels (C) and (D) report two examples of the prices associated with negative bursts. In each panel, the left Y-axis is the drift burst statistics, the right Y-axis is the S&P 500 E-mini future prices, and the X-axis indicates the intraday times. The red thick line indicates the S&P 500 E-mini future prices while the blue bars indicate the drift burst statistics. The dashed horizontal line denotes the threshold of 4 for testing drift bursts.

Figure 3.3.2. Four example days of intraday price dynamics that experience drift bursts.



Notes: This figure reports four randomly selected examples of intraday drift bursts. Panels (A) and (B) contain two examples of the prices related to the positive bursts while the lower Panels (C) and (D) report two examples of the prices associated with negative bursts. In each panel, the left Y-axis is the drift burst statistics, the right Y-axis is the S&P 500 E-mini future prices, and the X-axis is the intraday times. The red thick line denotes the S&P 500 E-mini future prices while the blue bars indicate the drift burst statistics. The dashed horizontal line denotes the threshold of 4 for testing drift bursts.

Panel (A) illustrates that on July 27, 2012, within the intraday interval between 13:24 and 13:28, there is a large upward trend of intraday prices that are detected to be significantly associated with a positive intraday burst as the test statistic exceeds the threshold. The stock price experienced a rapid ascent from a low point of 1373 to a high of 1379, comprised of many small upward moves, but no apparent jumps. This extreme, yet smooth, transition to a new price level is labelled a ‘‘gradual jump’’ by Barndorff-Nielsen et al. (2009). On July 27, 2012, the pledge by the European Central Bank President to "do whatever it takes" to defend the euro sparked off a sharp rally in the US markets. Investors had been concerned about the new European debt, as

Spain's borrowing costs surpassed the critical 7%. Mario Draghi's comments thus boosted sentiment and the fear-gauge CBOE Volatility Index (VIX) slumped 9.4%. Panel (B) shows that price gradually jumps from 13:40 to 13:45 on July 23, 2020, where a significant positive drift burst is detected. The price increase came after reports that U.S. officials would travel to China for face-to-face trade discussions, boosting equities further following stronger-than-expected earnings results earlier in the day.

Turning to the lower panel (C), a negative burst is detected within the time interval from about 14:20 to 14:25 on January 28, 2004, over which there is a sharp downside price trend and price drops substantially by 8 points. This downside gradual jumps occurred after the Federal Reserve changed the language in its outlook for interest rates, sparking worry among investors about when the central bank might start raising rates. Panel (D) shows that there are two negative bursts detected from 11:00 to 11:30 on February 20, 2020, with the drift burst statistic exceeding the critical value twice. The price over the period from 11:00 to 11:30 dropped about 50 points, which is much more dramatic than the negative price move reported in Panel (C). On 20 February 2020, stock markets across the world suddenly plunged after growing concern due to the COVID-19 pandemic.

The above four representative examples imply that drift bursts often indicate the gradual jump price pattern defined by Barndorff-Nielsen et al. (2009), which is consistent with Christensen et al. (2022) and Andersen et al. (2023). Christensen et al. (2022) and Bellia et al. (2023) show that drift bursts may also detect flash crashes, which are "V" shape price moves with a downside and explosive trend in prices occurring within a few minutes followed by a price recovery. However, I find that only about 1% of those negative bursts detected are associated with flash crashes,

following Nanex (2010), Braun et al. (2018), Félez-Viñas (2017) and Laly and Petitjean (2020), who define a flash crash by the phenomenon that the stock has to reverse (recover) to at least 90% of the initial price within a time gap of maximum 5-minute. Using the same approach, I observe that around 2% of positive bursts are flash spikes. These results imply that both positive and negative drift bursts are only very rarely related to flash crashes or spikes, but are dominated by gradual jumps. This may be interpreted by noting that the flash crashes themselves are extremely uncommon events in the financial markets. Andersen et al. (2023) and Jagannathan et al. (2022) also find that the majority of intraday drift bursts are related to gradual jumps with flash crashes very rare. Christensen et al. (2022) only link two flash crash events to drift bursts in a 7-year sample of S&P 500 E-mini future prices.

The economic reasons that determine the gradual jumps is a very new research question in the literature. Recent studies tend to link the causality of gradual jumps with economic news. Andersen et al. (2023) show that gradual jumps reflect the investors' concern regarding the lack of transparency of the monetary policies. Bouamara et al. (2023) argue that gradual jumps in the equity index can be caused by the heterogeneous reaction of the stock constituents to major economic news, such as pre-scheduled announcements, natural disasters, or geopolitical conflicts. They find that the stock constituents do not respond to the news at the same time, with less-liquid individual stocks typically lagging their reaction.

The different degrees of the gradual jumps found in Figure 3.3.1 imply that the drift burst might have different magnitudes. For example, in the negative burst reported in the left-bottom panel, the price associated with a negative burst decreases by 8 points while the negative burst in the right-bottom panel causes a much larger price drop, with the price decreasing over 50 points. Motivated by the possible

difference in the sizes of bursts, I constructed a variation measure to gauge the contribution of drift bursts to total price variation of that day.

To obtain this measure, first let $R_{t_i}^{5\text{min}}$ denote the return over the 5-minute interval $[t_{i-5\text{min}}, t_i]$, in which a 5-minute drift burst is detected. I assume that $R_{t_i}^{5\text{min}}$ is dominated by that burst such that the size of that burst is proxied by the magnitude of $R_{t_i}^{5\text{min}}$. This assumption is based on the observation that the price returns are almost exclusively due to the strong upward and downward drift burst. Then, the variation of bursts for day t may be defined by the sum of squared $R_{t_i}^{5\text{min}}$:

$$VDB_t^+ = \sum_{i=1}^{n-k_n+2} \left(R_{t_i}^{5\text{min}}\right)^2 \cdot DB_{t_i}^+, \tag{3.3.1}$$

and

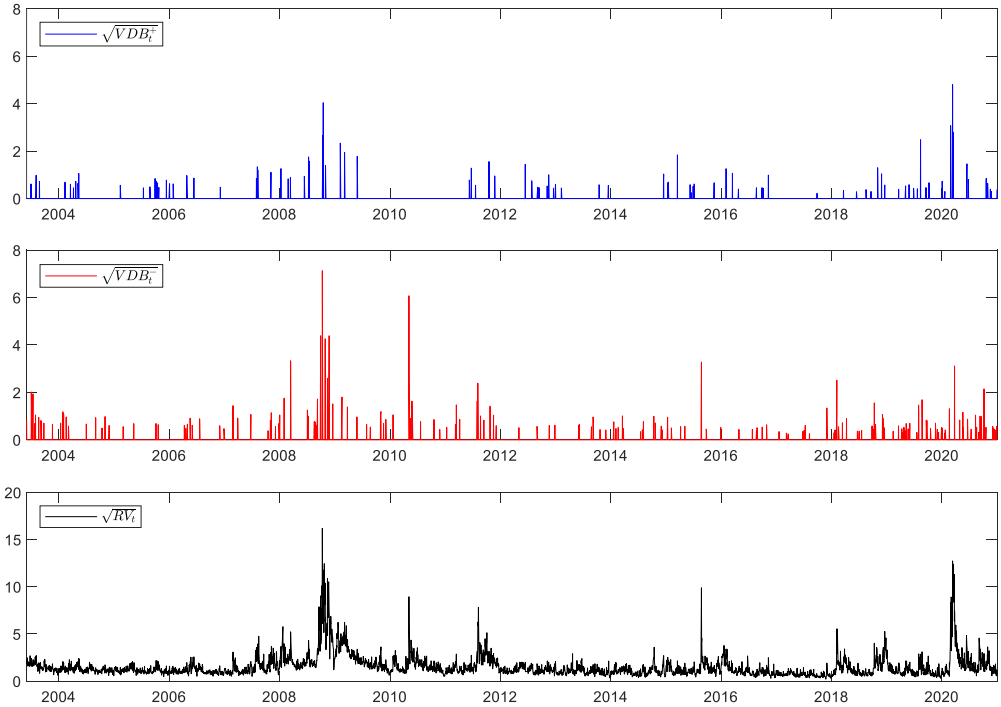
$$VDB_t^- = \sum_{i=1}^{n-k_n+2} \left(R_{t_i}^{5\text{min}}\right)^2 \cdot DB_{t_i}^-, \tag{3.3.2}$$

where VDB_t^+ is the variation of positive bursts on day t, $DB_{t_i}^+$ is a dummy which equals one if a positive burst is detected at a time point t_i, n denotes the overall number of price observations of day t, k_n is the previously defined preaveraging window, VDB_t^- is the variation of negative bursts on that day, and $DB_{t_i}^-$ is the dummy that equals one when a negative burst is identified at t_i. Laurent et al. (2022c) recently proposed a realized drift measure which attempts to estimate the variation of drift bursts. However, their measure does not distinguish the large daily constant drift, or a linear drift from the intraday drift bursts, but captures them as a whole.

Based on the calculated VDB_t^+ and VDB_t^-, I find that the average ratio of VDB_t^+ relative to RV_t is 18.77% while that of VDB_t^- to RV_t is 19.66%, suggesting that the

relative contribution of drift bursts to the daily price variation is about 20% on average, which suggests that drift bursts are associated with large price movements. Figure 3.3.3 presents the daily time series of drift burst variations and the realized variance. The top, middle, and bottom panels contain the positive and negative burst variations VDB_t^+ and VDB_t^-, and realized variance (RV_t), respectively. All measures are in square root form for presentation purposes. From the upper two panels, I observe that both VDB_t^+ and VDB_t^- are not constant but vary over the sample period. These two burst variations appear to be larger in the late 2008, early 2009, and 2020, during which the realized variance RV_t, reported in the bottom panel, is also higher than normal.

Figure 3.3.3. Daily time series of drift burst variations and realized variance



Notes: The upper two panels of this figure contain the daily variation of positive drift burst (VDB_t^+) and the daily variation of negative drift bursts (VDB_t^-), respectively. The burst variation is defined by the sum of the squared 5-minute returns up to the intraday time points when bursts are detected. The bottom panel depicts the daily realized variance (RV_t), defined by the sum of 5-minute squared returns from 9:30 to 16:00 EST. The sample data is S&P 500 E-mini prices from June 2, 2003, to December 30, 2020. All variables are in square root form.

3.4. Modelling drift burst clustering effect using Hawkes processes

Given the visual observation of the burst clusters in section 3.3, the burst arrivals may be a self-exciting process, which means that the intensity, or the probability, of the occurrence of positive or negative bursts is positively affected by their own past arrivals. The intensity of positive or negative bursts might be impacted by the past arrivals of their counterparts as well, resulting in mutually exciting behaviour. The visual inspection of burst clusters requires rigorous statistical modelling. To model bursts' mutual-excitation characteristic, I consider the Hawkes (1971) processes (HPs). HPs exploit the mutual-excitation effects by allowing the intensity of an event to be determined by the past arrivals of this event and or other events. Previous studies tend to apply HPs to model the mutual-excitation effect of jumps (Bowsher 2007; Large 2007; Bauwens and Hautsch 2009; Aït-Sahalia et al. 2015; Clements and Liao 2017; Ma et al. 2019). As far as I am aware, my analysis of fitting HPs to the clustering behaviours of bursts is novel. Drift bursts are associated with large price movements, and modelling their dynamics thus has important implications for risk management, asset pricing, and understanding of the behaviour of markets.

To begin, let the arrivals of intraday positive and negative drift bursts be denoted by two sequences of increasing event times: $\left\{t_{q^+}^+\right\}_{q^+\in 1,\ldots,n^+}$, $0 < t_1^+ < t_2^+ < t_3^+ < \cdots < t_{n^+}^+ < T_N$; and $\left\{t_{q^-}^-\right\}_{q^-\in 1,\ldots,n^-}$, $0 < t_1^- < t_2^- < t_3^- < \cdots < t_{n^-}^- < T_N$, where $\left\{t_{q^+}^+\right\}_{q^+\in 1,\ldots,n^+}$ indicates the time of positive burst arrivals, $\left\{t_{q^-}^-\right\}_{q^-\in 1,\ldots,n^-}$ is the time per negative burst arrivals, n^+ and n^- denote the overall numbers of positive and negative bursts over the sample length $(0, T_N)$, respectively. Let $N^+(t) = \sum_{q^+\geq 1} \mathbf{1}_{t_{q^+}^+ \leq t}$ and $N^-(t) = \sum_{q^-\geq 1} \mathbf{1}_{t_{q^-}^- \leq t}$ be a counting process for recording the

number of positive bursts and negative bursts up to t, with corresponding intensities $P(dN^+(t) = 1) = IDB_t^+ dt$ and $P(dN^-(t) = 1) = IDB_t^- dt$.

Then, the mutual-excitation HPs for the intensities of positive and negative bursts at time t are defined by:

$$IDB_t^+ = C^{+,+} + \int_0^t \delta^{+,+}(t-s)dN^+(s) + \int_0^t \delta^{+,-}(t-s)dN^-(s), \quad (3.4.1)$$

and

$$IDB_t^- = C^{-,-} + \int_0^t \delta^{-,+}(t-s)dN^+(s) + \int_0^t \delta^{-,-}(t-s)dN^-(s). \quad (3.4.2)$$

IDB_t^+ is the intensity of positive bursts excited by both the arrivals of itself $dN^+(s)$ and its negative counterpart $dN^-(s)$, IDB_t^- denotes the intensity of negative burst excited by the arrivals of positive bursts $dN^+(s)$ and its own occurrence $dN^-(s)$, $C^{+,+}$ and $C^{-,-}$ denote the intensity of an underlying Poisson process and is a non-negative constant, $\delta^{+,+}(t-s)$, $\delta^{+,-}(t-s)$, $\delta^{-,+}(t-s)$, and $\delta^{-,-}(t-s)$ are non-negative decaying functions. The intuition is that the first intensity specification, Equation (3.4.1), allows the positive and negative drift burst before t, indicated by $dN^+(s)$ and $dN^-(s)$, $0 < s < t$, to increase the intensity of positive bursts at time t, IDB_t^+, by $\delta^{+,+}(t-s) + \delta^{+,-}(t-s)$. The second intensity specification, Equation (3.4.2), is allowing all positive and negative drift bursts prior to t, indicated by $dN^+(s)$ and $dN^-(s)$, $0 < s < t$, to increase the intensity of negative bursts at time t, IDB_t^-, by $\delta^{-,+}(t-s) + \delta^{-,-}(t-s)$.

Following Aït-Sahalia et al. (2015) and Bormetti et al. (2015), I use the exponentially decaying functions to parametrize those decaying functions. Then the discretized versions of Equations (3.4.1) and (3.4.2) may be written as,

$$IDB_t^+ = C^{+,+} + \sum_{t_{q^+}^+ < t} \mathrm{A}^{+,+} e^{-B^{+,+}(t - t_{q^+}^+)} + \sum_{t_{q^-}^- < t} \mathrm{A}^{+,-} e^{-B^{+,+}(t - t_{q^-}^-)}, \tag{3.4.3}$$

and

$$IDB_t^- = C^{-,-} + \sum_{t_{q^+}^+ < t} \mathrm{A}^{-,+} e^{-B^{-,-}(t - t_{q^+}^+)} + \sum_{t_{q^-}^- < t} \mathrm{A}^{-,-} e^{-B^{-,-}(t - t_{q^-}^-)}, \tag{3.4.4}$$

where $C^{+,+}, C^{-,-}$ are coefficients related to the underlying Poisson intensities, $A^{+,+}, A^{-,-}$ are self-excitation coefficients, $A^{+,-}, A^{-,+}$ are mutual-excitation coefficients, and $B^{+,+}, B^{-,-}$ are decay rates. The coefficients $A^{+,+}, A^{-,-}$, $A^{+,-}$, and $A^{-,+}$ are important to test the self- and mutual- excitation behaviours of positive and negative drift bursts. $A^{+,+}$ summarise to what extent all arrivals (exponentially weighted) of positive bursts before day t impact the probability of the occurrence of positive bursts on that day. $A^{-,-}$ describes the same effect for the negative bursts. $A^{+,-}$ denotes to what extent all arrivals (exponentially weighted) of negative bursts before day t impact the probability of the occurrence of positive bursts on day t while $A^{-,+}$ indicates the impact of all arrivals of positive bursts before day t on the chance of the presence of negative bursts on that day. Following Ogata (1981), the parameters in Equations (3.4.3) and (3.4.4) are estimated by maximizing the negative log-likelihood function below,

$$\begin{aligned} &\log L_{T_N}(C^{+,+}, C^{-,-}, A^{+,+}, A^{+,-}, A^{-,+}, A^{-,-}, B^{+,+}, B^{-,-}) \\ &\quad = \log L_{T_N}^+(C^{+,+}, A^{+,+}, A^{+,-}, B^{+,+}) \\ &\quad + \log L_{T_N}^-(C^{-,-}, A^{-,+}, A^{-,-}, B^{-,-}), \end{aligned} \tag{3.4.5}$$

with the two log function components $\log L_{T_N}^+(C^{+,+}, A^{+,+}, A^{+,-}, B^{+,+})$ and $log L_{T_N}^-(C^{-,-}, A^{-,+}, A^{-,-}, B^{-,-})$ defined by,

$$\log L^{+}_{T_N}(C^{+,+}, A^{+,+}, A^{+,-}, B^{+,+})$$

$$= -C^{+,+}T_N - \frac{A^{+,+}}{B^{+,+}} \sum_{q^+=1}^{n^+} \left[1 - e^{-B^{+,+}\left(T_N - t^{+}_{q^+}\right)}\right]$$

$$- \frac{A^{+,-}}{B^{+,+}} \sum_{q^-=1}^{n^-} \left[1 - e^{-B^{+,+}\left(T_N - t^{-}_{q^-}\right)}\right]$$

$$- \sum_{q^+=1}^{n^+} \log\left[C^{+,+} + A^{+,+} \cdot R^{+,+}_{q^+} + A^{+,-} \cdot R^{+,-}_{q^+}\right],$$

and

$$\log L^{-}_{T_N}(C^{-,-}, A^{-,+}, A^{-,-}, B^{-,-})$$

$$= -C^{-,-}T_N - \frac{A^{-,+}}{B^{-,-}} \sum_{q^+=1}^{n^+} \left[1 - e^{-B^{-,-}\left(T_N - t^{+}_{q^+}\right)}\right]$$

$$- \frac{A^{-,-}}{B^{-,-}} \sum_{q^-=1}^{n^-} \left[1 - e^{-B^{-,-}\left(T_N - t^{-}_{q^-}\right)}\right]$$

$$- \sum_{q^-=1}^{n^-} \log\left[C^{-,-} + A^{-,+} \cdot R^{-,+}_{q^-} + A^{-,-} \cdot R^{-,-}_{q^-}\right]$$

where $R^{+,+}_{q^+}$, $R^{+,-}_{q^+}$, $R^{-,+}_{q^-}$, and $R^{-,-}_{q^-}$ are given recursively by

$$R^{+,+}_{1} = R^{+,-}_{1} = R^{+,-}_{1} = R^{-,-}_{1} = 0,$$

$$R^{+,+}_{q^+} = e^{-B^{+,+}\left(t^{+}_{q^+} - t^{+}_{q^+-1}\right)} \cdot \left(1 + R^{+,+}_{q^+-1}\right),$$

$$R^{+,-}_{q^+} = e^{-B^{+,+}\left(t^{+}_{q^+} - t^{+}_{q^+-1}\right)} \cdot R^{+,-}_{q^+-1} + \sum_{\left\{q^-:\, t^{+}_{q^+-1} \le t^{-}_{q^-} < t^{+}_{q^+}\right\}} e^{-B^{+,+}\left(t^{+}_{q^+} - t^{-}_{q^-}\right)},$$

$$R^{-,+}_{q^-} = e^{-B^{-,-}\left(t^{-}_{q^-} - t^{-}_{q^--1}\right)} \cdot R^{-,+}_{q^--1} + \sum_{\left\{q^+:\, t^{-}_{q^--1} \le t^{+}_{q^+} < t^{-}_{q^-}\right\}} e^{-B^{-,-}\left(t^{-}_{q^-} - t^{+}_{q^+}\right)}$$

and

$$R^{-,-}_{q^-} = e^{-B^{-,-}\left(t^{-}_{q^-} - t^{-}_{q^--1}\right)} \cdot \left(1 + R^{-,-}_{q^--1}\right).$$

The statistical inference on the parameter estimates in Equations (3.4.3) and (3.4.4) is based on the variances of parameters that are estimated by the inverse of the diagonal of the Hessian matrix of the log-likelihood function. Table 3.4.1 presents the estimated parameters, along with the loglikelihood, Equation (3.4.5).

Table 3.4.1. The maximum likelihood estimation results of the parameters in the intensity function of the mutual-excitation Hawkes processes.

Notes: This table reports the maximum likelihood estimation results for the coefficients in the intensity function of the mutual-excitation Hawkes (1971) process, Equations (3.4.3) and (3.4.4). Log-like denotes the loglikelihood of the coefficient estimation, Equation (3.4.5). The coefficients $C^{+,+}, C^{-,-}$ are the underlying Poisson intensities, $A^{+,+}, A^{-,-}$ are self-excitation coefficients, $A^{+,-}, A^{-,+}$ are mutual-excitation coefficients, and $B^{+,+}, B^{-,-}$ are decay rates. All the coefficients are scaled by 10^3 for presentation purposes. The parentheses below the coefficients show the t-statistics.

$C^{+,+}$	$B^{+,+}$	$A^{+,+}$	$A^{+,-}$	$C^{-,-}$	$B^{-,-}$	$A^{-,+}$	$A^{-,-}$	Log-like
7.76	13.87	7.39	1.24	7.82	9.42	1.07	7.31	1340.52
(3.66)	(7.61)	(5.32)	(1.98)	(3.04)	(11.36)	(0.95)	(10.83)	

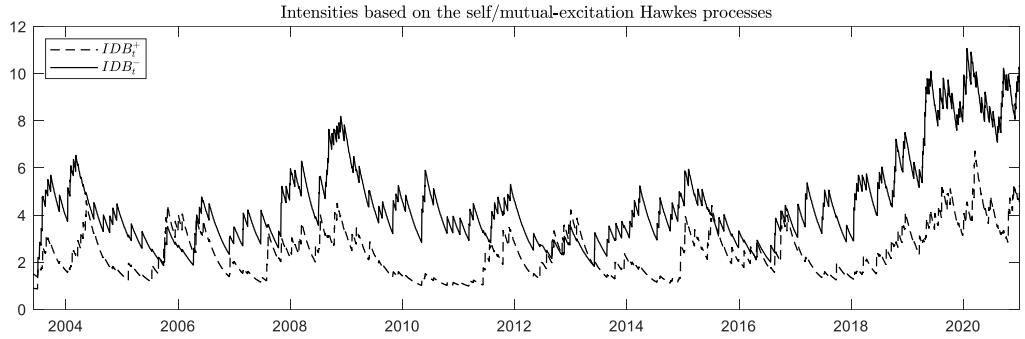
As can be seen from this table, the decay rate $B^{+,+}, B^{-,-}$ and initial intensity $C^{+,+}, C^{-,-}$ are all positive and significant at the 5% level. The decay rate $B^{+,+}$ and $B^{-,-}$ are larger than the corresponding mutual-excitation factors $A^{+,+}, A^{+,-}$ and $A^{-,+}, A^{-,-}$, which meets the requirement of the stationarity of the intensity processes of the mutually-excitation HPs (Hawkes 1971; Bormetti et al. 2015). The coefficient estimates of the self-excitation factor $A^{+,+}$ and $A^{-,-}$ are positive and significant at the 5% significance level. This result suggests that both positive and negative bursts significantly self-excite.

The mutual-excitation parameter $A^{+,-}$ is significant at the 5% level, and this indicates that the probability of a positive burst at day t is significantly impacted by the arrivals of negative bursts before that day. Another mutual-excitation factor, denoted by $A^{-,+}$, is not significant at the 10% level, implying that negative bursts are

not significantly excited by previous positive burst arrivals. The phenomenon that positive and negative bursts affect each other in an asymmetric fashion may be explained by noting that there could be some positive policies to restore the confidence of the stock market after a large price downturn. Governments and regulatory bodies can implement policies to improve the market, such as lowering interest rates, providing financial assistance to struggling sectors, or enacting measures to prevent market manipulation. These good policies may lead to a subsequent surge in the stock market prices.

Based on estimated parameters and the arrival times of positive and negative bursts in advance to day t ($t_{q^+}^{+} < t$ and $t_{q^-}^{-} < t$), the mutual-excitation HP-based intensity of positive and negative bursts for day t, denoted by IDB_t^+ and IDB_t^-, for all trading days of the sample can be constructed recursively from Equations (3.4.3) and (3.4.4). Figure 3.4.1 visualizes the mutual-excitation intensities of positive and negative bursts, IDB_t^+ and IDB_t^-, denoted by the dashed and solid lines. As the figure shows, IDB_t^+ and IDB_t^- fluctuate over the sample period and are at their higher levels during the 2008-2009 crisis and the 2020 pandemic periods. Both IDB_t^+ and IDB_t^- have an upward trend from 2010 to 2021, indicating that from 2010 there are increased probabilities of burst arrivals. This is consistent with Christensen et al. (2022), who find that the occurrence rate of drift bursts appears to have increased over time. The intensity of drift bursts has become relatively higher since 2018. The causes of this phenomenon are interesting and left for future research. Comparing IDB_t^+ with IDB_t^-, the latter is overwhelmingly, with a few exceptions, greater in size across the sample period. This may be interpreted by noting that the unconditional probability of the occurrence of negative bursts is much greater than that of positive bursts in my sample under the threshold considered.

Figure 3.4.1. Time series of drift burst intensities



Notes: This figure depicts daily intensities of positive and negative drift bursts based on the mutual-excitation Hawkes (1971) processes, indicated by IDB_t^+ and IDB_t^-, respectively (the dashed and solid lines). IDB_t^+ is defined by the probability of the occurrence of a positive burst at the close of day t conditional on the arrivals of both positive and negative bursts. IDB_t^- is defined by the probability of the occurrence of a negative burst at the close of day t conditional on the arrivals of both positive and negative bursts prior to day t. The sample period is from June 2, 2003, to December 30, 2020. All intensities are in percentages.

3.5. Predicting volatility using drift bursts

Chapter 2 shows that daily measures of constant drift and linear drift can forecast volatility. Laurent et al. (2024) also demonstrate that unsigned drift bursts can positively predict volatility. Intuitively, drift bursts are associated with large price movements relative to other intraday price variation, and thus could have an impact on future volatility. To rich empirical evidence, this section explores the effect of drift bursts on volatility forecasting. Anticipating the appropriate level of volatility has a significant impact on pricing financial assets such as stocks and options (Heston and Nandi 2000; Ghysels et al. 2005; Christensen and Nielsen 2007; Christoffersen et al. 2008; Goyal and Saretto 2009; Corsi et al. 2013). Volatility forecasting also plays a critical role in financial risk management (Christoffersen and Diebold 2000; Clements et al. 2008; Maheu and McCurdy 2011). Making precise volatility forecasts is of the utmost importance for many practitioners and regulators of the financial markets.

For the volatility proxy, I use the realized variance, which is defined by the sum of the squared returns over a period (e.g., one trading day). Realized variance is one of the most popular estimators in measuring the variance of asset returns in the volatility forecasting, asset pricing, and risk management literatures (Andersen et al. 2000; Koopman et al. 2005; Andersen et al. 2007b; Bollerslev et al. 2009; Corsi 2009; Feunou and Okou 2019; Bollerslev et al. 2020). Consistent with Chapter 2, the volatility forecasting target is h-day ahead cumulative realized variance, indicated by $RV_{t,t+h}$, where,

$$RV_{t,t+h} = RV_{t+1} + RV_{t+2} + \cdots + RV_{t+h}, \tag{3.5.1}$$

and RV_t is defined as the sum of the squared 5-minute returns of S&P 500 E-mini futures recorded on that day from 9:30 EST to 16:00 EST on day t. This frequency is consistent with that used in Chapter 2 for constructing realized variance. Many papers in the literature on realized variance prediction use "standard" 5-minute data as it balances the trade-off between the desire to use very finely sampled data to minimize the estimation error on the one hand, and not be overwhelmed by "noise" in the high-frequency prices on the other (Andersen et al. 2007b; Corsi 2009; Duong and Swanson 2015; Liu et al. 2015; Patton and Sheppard 2015; Bollerslev et al. 2016; Buccheri and Corsi 2021; Clements and Preve 2021; Christensen et al. 2023; Laurent et al. 2024). Also, there is very limited evidence on the predictive superiority of a particular frequency versus another (Liu et al. 2015).

I investigate the impacts on the h-day ahead cumulative realized variance ($RV_{t,t+h}$) forecasting of all drift burst variables defined previously in this chapter, including burst occurrences NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14), burst variation, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2), and mutual-excitation intensity of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4). To explore the impacts of these different burst variables on $RV_{t,t+h}$, $RV_{t,t+h}$is regressed on each type of the variable of positive and negative bursts.

All these regression models, in which drift burst measures are used, all control for lagged daily, weekly, and monthly cumulative realized variance, denoted by RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$, respectively, where,

$$RV_{t-h',t} = RV_{t-h'+1} + RV_{t-h'+2} + \cdots + RV_t, \tag{3.5.2}$$

with RV_t denoting the same daily realized variance on day t. These RV lags are all annualized: RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ are multiplied by 252, 252/5, and 252/22, respectively. [5] These realized variance lags are the same predictors of the heterogeneous autoregressive (HAR) model proposed by Corsi (2009), which is widely applied as the benchmark model in the volatility forecasting literature (Andersen et al. 2007b; Corsi et al. 2010; Busch et al. 2011; Corsi and Renò 2012; Duong and Swanson 2015; Bollerslev et al. 2016; Andersen et al. 2023). Therefore, my regression models which harness the information of drift burst variables are various extensions of the HAR model.

According to the previous literature, various extensions of the HAR model have been proposed, for example, including jumps (Andersen et al. 2007b; Corsi et al. 2010; Duong and Swanson 2015), leverage effects (Bollerslev et al. 2006; Corsi and Renò 2012; Patton and Sheppard 2015), or allowing for non-linearities in the volatility process (Andersen et al. 2007b; McAleer and Medeiros 2008; Fengler et al. 2015; Clements and Preve 2021). However, several recent studies have argued the significance of drift in understanding and modelling asset prices: Christensen et al. (2022) emphasize that drift forms an integral part of price dynamics across currencies, fixed-income investments, equities, and commodities; Laurent and Shi (2020) show that neglecting drift can result in significant finite sample bias in realized variance and severe size distortion in jump testing; and Laurent et al. (2024) demonstrate the importance of drift bursts in volatility forecasting. My extensions of the HAR model using various drift burst variables thus provide more evidence on the importance of drift bursts in the literature.

[5] For the RV lags in the right-hand side of the HAR model, I again apply their non-overlapping versions, which is consistent with Chapter 2.

The details of these regression models are as follows. The first regression model is the standard HAR model, which contains the lagged daily, weekly, and monthly cumulative realized variance, RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ as the explanatory variables and does not depend on any drift burst measures and used as a benchmark. The second model extends the HAR model by exploiting the effects of the occurrences of positive and negative drift bursts, NDB_t^+ and NDB_t^-. The third regression augments the HAR by including the information on the variation of drift bursts, VDB_t^+ and VDB_t^-. The fourth model attempt to improve the HAR using mutual-excitation intensity of bursts, IDB_t^+ and IDB_t^-. [6]

The level of realized variance changed substantially across our sample period. For example, Figure 3.3.3 exhibits that the level of realized variance is much larger from 2008 to 2009 than during the period from 2004 to 2005. Financial econometricians find that residuals of the HAR model estimated using Ordinary Least Square (OLS) are correlated with the level of the realized variance (Corsi et al. 2008; Patton and Sheppard 2015). In other words, the resulting OLS coefficient estimates focus primarily on fitting periods of high realized variance and place little weight on more tranquil periods. This may increase the variance of the coefficient estimate, resulting in efficiency of the coefficient estimation of the HAR model.

To alleviate this issue, I modify all the regression models by using the logarithmic realized variance as the forecast target. The logarithmic realized variance varies much less dramatically over my sample period than the realized variance, therefore

[6] The intensity of positive and negative bursts can also be calculated exclusively using self-excitation behaviours, without considering the mutual-excitation effects across positive and negative bursts. However, ignoring the mutual-excitation impacts may cause omitted-variable bias, leading to incorrect parameter estimates. Consequently, biased parameters may result in unsatisfactory estimations of the intensity of drift bursts.

producing less heteroskedastic residuals (Andersen et al. 2003; Thomakos and Wang 2003; Clements and Preve 2021). Andersen et al. (2007b) and Caporin et al. (2017) also use the logarithmic converted HAR-style models for modeling and forecasting volatility. All variation predictors of the models including the lagged daily, weekly, and monthly average realized variance, together with the drift burst variations are also converted into their logarithmic form for consistency. An alternative method to improve the inefficiency due to heteroscedasticity is the weighted least square (WLS) estimation method introduced in Chapter 2. I observe that the conclusions related to the effects of drift burst variables on the future realized variance is not altered when the models are estimated by WLS. The results for WLS are attached in Appendix A.4.

Table 3.5.1 reports the daily forecast results ($h = 1$) including the coefficient estimates and goodness of fit for the four models that depend on the drift burst information, along with the benchmark model that is only based on the realized variance lags. The parentheses contain the Heteroskedasticity and Autocorrelation Consistent (HAC) robust t-statistics according to Newey and West (1987b) with the number of the lags equal to $2 \times (h + 1)$. The HAC adjustment corrects the impact of heteroskedasticity in the residuals on the accuracy of the standard errors of the model coefficient estimates. The first column of this table contains the variables that are used as predictors in the regressions. The second column shows the HAR model suggested by Corsi (2009). As this column shows, daily, weekly, and monthly lagged averages of realized variance, denoted by RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$, significantly increase the one-day ahead realized variance $RV_{t,t+1}$ at the 0.1% level of significance. This is consistent with the findings by Corsi (2009), indicating that the time series volatility is persistent. The third column of Table 3.5.1 shows the estimation results for the model that explores the impact of positive and negative drift burst occurrences,

indicated by NDB_t^+ and NDB_t^-, on the next day realized variance, $RV_{t,t+1}$. As shown by the results, NDB_t^+ significantly decreases $RV_{t,t+1}$. The significance of NDB_t^+ could be explained by noting that positive return reduces the future volatility. However, I find that NDB_t^- does not significantly affect $RV_{t,t+1}$ at the 5% level, which is puzzling as the presence of large negative price movements are expected to increase future volatility according to the negative return-volatility relationship.

Table 3.5.1. The impact of drift bursts on daily realized variance forecasts

This table reports the results of coefficient estimation and goodness of fit (adjusted R^2) for the four models that investigate the impacts of different drift burst measures on daily realized variance $RV_{t,t+1}$ forecasts. The first column contains the variables used in these models. All regressions contain RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ as the independent variables, where RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ indicate the lagged daily, weekly, and monthly average realized variance, Equation (3.5.1). The second column is the model that depends on RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$. The third column contains the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14). The fourth column is for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2). The fifth column reports the results of the specifications that are based on the mutual-exciting intensity of bursts, IDB_t^+ and IDB_t^-, Equations 3.4.3 and 3.4.4. The sample is 1-second S&P 500 E-mini future transaction prices from June 2003 to December 2020.

Constant	-0.012	-0.007	-0.010	-0.006
	(-1.12)	(-0.63)	(-0.92)	(-0.21)
RV_t	0.563	0.571	0.572	0.562
	(30.51)	(29.77)	(29.99)	(30.68)
$RV_{t-5,t}$	0.273	0.270	0.271	0.273
	(12.61)	(12.38)	(12.51)	(12.60)
$RV_{t-22,t}$	0.109	0.105	0.106	0.109
	(6.22)	(5.98)	(6.04)	(6.03)
NDB_t^+		-0.177		
		(-3.28)		
NDB_t^-		-0.023		
		(-0.52)		
VDB_t^+			-0.156	
			(-2.03)	
VDB_t^-			-0.062	
			(-1.16)	
IDB_t^+				-0.020
				(-1.84)
IDB_t^-				0.009
				(1.33)
R^2	0.748	0.749	0.749	0.748

The fourth column contains the estimation results of the model that anticipates $RV_{t,t+1}$ using the variation information contained in positive and negative drift bursts, VDB_t^+ and VDB_t^-. The results show that drift burst variation reveals a significant and negative effect of VDB_t^+ and an insignificant impact of VDB_t^- (at the 5% level), which is consistent with the effects of the occurrence measures in column 3. Turning to the fifth column pertaining to the results of the model using the HPs' intensity of positive and negative drift bursts, IDB_t^+ and IDB_t^-, I observe that both IDB_t^- and IDB_t^+ do not significantly affect $RV_{t,t+1}$ at the 5% level. The significance of NDB_t^+ and VDB_t^+ contrasts with the insignificance of IDB_t^+, even though NDB_t^+, VDB_t^+, and IDB_t^+ may be expected to have similar impacts because they both depend on the occurrence information of positive drift bursts. The impact on future volatility of the measures related to the negative drift bursts, including NDB_t^-, VDB_t^-, and IDB_t^-, are always not significant.

Table 3.5.2 presents the coefficient estimation and goodness of fit of the same four models reported in Table 3.5.1 for the 5-day ahead cumulative realized variance, or weekly realized variance forecast, $RV_{t,t+5}$. I again observe the significant effect of the volatility persistence, with the coefficients of the lags of daily, weekly, and monthly average realized variance, RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$, significant at the 0.1% level. The results also show that none of the variables of negative drift bursts (NDB_t^-, VDB_t^-, IDB_t^-, and IDB_t^-) have a significant impact on $RV_{t,t+5}$, which is similar to Table 3.5.1. In Table 3.5.2, all variables related to the upside drift bursts including NDB_t^+, VDB_t^+, IDB_t^+, and IDB_t^+ negatively predict $RV_{t,t+5}$, but only the impact of NDB_t^+ is statistically significant, which is qualitatively consistent with the effects of these variables on the daily realized variance forecast.

Table 3.5.2. The impact of drift bursts on weekly realized variance forecasts

Notes: This table reports the regression estimation results for investigating the impact of different drift burst measures on weekly realized variance, $RV_{t,t+5}$, forecasts. The detailed information of these models and the sample are as in the notes of Table 3.5.1.

Constant	0.098	0.098	0.099	0.043
	(5.05)	(5.08)	(5.10)	(0.79)
RV_t	0.436	0.438	0.440	0.434
	(20.73)	(20.17)	(20.45)	(21.49)
$RV_{t-5,t}$	0.292	0.293	0.291	0.288
	(9.58)	(9.61)	(9.61)	(9.56)
$RV_{t-22,t}$	0.161	0.159	0.159	0.154
	(5.07)	(5.03)	(5.05)	(4.62)
NDB_t^+		-0.109		
		(-2.12)		
NDB_t^-		0.036		
		(0.87)		
VDB_t^+			-0.064	
			(-0.78)	
VDB_t^-			-0.027	
			(-0.53)	
IDB_t^+				-0.023
				(-1.18)
IDB_t^-				0.026
				(1.59)
R^2	0.741	0.742	0.741	0.743

Table 3.5.3 contains the results of the monthly realized variance forecast, $RV_{t,t+22}$ for the same four models reported in Tables 3.5.1 and 3.5.2. The second column demonstrating the HAR model estimation result again reveals the importance of volatility persistence, with the effects of all lagged realized volatilities, indicated by RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$, significant at the 5% level. Across the third, fourth, and fifth columns, the drift burst variables systematically fail to predict the monthly realized variance, even for the occurrence of positive bursts.

Table 3.5.3. The impact of drift bursts on monthly realized variance forecasts

Notes: This table reports the regression estimation results for investigating the impact of different drift burst measures on monthly realized variance, $RV_{t,t+22}$, forecasts. The detailed information of these models and the sample are as in the notes of Table 3.5.1.

Constant	0.251	0.248	0.249	0.084
	(5.27)	(5.26)	(5.24)	(0.62)
RV_t	0.296	0.290	0.289	0.290
	(11.29)	(11.39)	(11.23)	(12.54)
$RV_{t-5,t}$	0.270	0.274	0.272	0.260
	(7.94)	(7.90)	(8.01)	(8.33)
$RV_{t-22,t}$	0.190	0.190	0.191	0.170
	(3.40)	(3.42)	(3.45)	(2.74)
NDB_t^+		-0.061		
		(-1.09)		
NDB_t^-		0.087		
		(1.63)		
VDB_t^+			0.032	
			(0.40)	
VDB_t^-			0.070	
			(1.41)	
IDB_t^+				-0.038
				(-0.92)
IDB_t^-				0.063
				(1.26)
R^2	0.589	0.589	0.589	0.598

The above models that exploit the impacts of drift bursts on realized variance (RV) forecasting are controlled for the daily, weekly, and monthly lags of RV. Recent studies argue that RV overestimates the return variance as it may also capture the drift burst variation (Laurent et al. 2022c; Andersen et al. 2023). Therefore, the right-hand side of these drift burst regressions may be subject to the double counting issue of the explanatory variable information. To address this problem, I modify these regression models by replacing all RV lags with the lags of the recent $RiceV$ variance estimator proposed by Laurent et al. (2022c). $RiceV$, simply defined as half of the sum of squared first order difference of returns, is robust to the drift burst variation thus including this variance measure is devoid of any overlapping of the drift burst

information in the explanatory variables in the regression models. The first-order difference of returns denotes the difference between itself and the previous observation. The intuition is that the first-order difference detrends return data thus reducing the impact of drift bursts, which are associated with extreme trends, on volatility estimation.

The results (attached in Appendix A.4.) show that the effects of drift bursts on the realized variance forecast based on the modified models are again generally not significant, which is consistent with those from the original models. Additionally, I also further modify the regression models by using $RiceV$ for the forecast target. This ensures that we forecast a more accurately estimated volatility. The results again reveal weak effects of drift bursts on $RiceV$ forecasts, similar to the insignificant impacts of bursts on RV.

The sample correlation between RV and $RiceV$ equals 99.33%. Given the high degree of similarity in these two volatility estimators inherent in the S&P 500 returns, it is hardly surprising that the effects of drift burst on volatility forecasting are not sensitive to the different choices of volatility estimators in the models. The weak impact of drift bursts on volatility forecasting is quite counter-intuitive since large price moves are expected to exhibit a leverage effect on volatility, with positive return reducing volatility and negative return increasing volatility. The interpretation of this puzzling result is left for future research.

Realized variance is calculated using high-frequency returns and is therefore inevitably subject to bias due to microstructure noise. Although various methods have been applied in the literature to address this bias, these methods require certain assumptions regarding the microstructure noise structure (Jacod et al. 2009; Hansen and Lunde, 2006; Podolskij and Vetter 2009; Aït-Sahalia et al. 2012; Lee and

Mykland 2012; Hautsch and Podolskij 2013; Christensen et al. 2014). As suggested by Andersen and Bollerslev (1998), the option-implied variance, derived from the options prices, is not affected by the microstructure noise present in high-frequency stock price returns thus does not depend on specific hypotheses about the microstructure noise structure. Due to the superior measurement performance of implied variance in the presence of microstructure noise, I also consider forecasting implied variance.

The implied variance measure of the S&P 500 index is defined in the literature by the squared CBOE implied variance index VIX^2, which is the variance of the S&P 500 index derived from the S&P 500 options using the Black–Scholes formula. Realized variance measures the historical volatility while implied variance estimates the volatility expected by investors in the future. Implied variance has important applications in financial markets. This volatility measure indicates the investors' fear of the risk of the S&P 500 index. Traders use implied variance to identify potential opportunities. For example, high implied variance might indicate a good time to sell options, while low implied variance could be a buying opportunity. Implied variance is a key component in options pricing models, such as the Black-Scholes model. It helps determine the premium of an option, influencing trading strategies and hedging. This chapter studies the impact of drift bursts on implied variance, measured by VIX^2. Since the end-of-day VIX_t^2 captures investors' expectation of the realized variance of the S&P 500 index over the next month, my analysis of the effect of bursts on VIX_t^2 is consistent with the study of the impact of drift bursts on the monthly realized variance forecast. Of course, one can also synthesize daily and weekly realized variance using the CBOE 1-Day Volatility Index (VIX1D Index) and or implied variance from end-

of-day option prices. This chapter focuses on forecasting the popular one-month implied variance VIX^2, which is consistent with monthly realized variance.

To explore the effects of different burst measures on VIX_t^2, I consider four regression models. All models contain the S&P 500 index option implied variance at the close of day t, VIX_t^2 as the response variable (response variable means the dependent variable or forecast target in the model), including the daily lag of implied variance, denoted by VIX_{t-1}^2. The inclusion of this lag tries to take into account the autocorrelation in the daily time series of VIX_t^2. These four models differ in the drift burst variables that are included. The first model excludes all drift burst information as regarded as the benchmark. The second model uses the drift burst occurrence, and the third model is based on the drift burst variation. The fourth model depends on the mutual-excitation intensities of drift bursts. Consistent with the realized variance model, these implied variance models are estimated in the logarithmic form. The results for these models estimated by the alternative WLS method demonstrate qualitatively similar conclusions and are reported in Appendix A.4.

Table 3.5.4 reports the parameter estimates and the coefficient of determination for the four regression models that exploit the effects of different drift burst information variables on the implied variance at the close of day t, indicated by VIX_t^2. The second column of Table 3.5.4 contains the estimation results for the model that solely depends on the daily lag of the implied variance, indicated by VIX_{t-1}^2. This is the model that studies if the daily implied variance is persistent. As the results show, the coefficient on VIX_{t-1}^2 is significant at the 5% level, suggesting that the daily S&P 500 option implied variance is strongly persistent. The third column shows the results for the model that exploits the information of the positive and negative drift burst

occurrence, indicated by NDB_t^+ and NDB_t^-, respectively. It can be seen from this column that NDB_t^- significantly increases VIX_t^2 while the impact of NDB_t^+ on implied variance is not significant at any reasonable level. The significance of NDB_t^- may be interpreted by noting that large price drops increase investors' fears of downside risk, thus, investors tend to hedge this risk via trading in S&P 500 index options, increasing the VIX index.

Table 3.5.4. The impact of drift bursts on implied variance

This table reports the regression estimation results for investigating the impact of different drift bursts measure on the implied variance, which is defined by the squared daily close value of the Chicago Board Options Exchange (CBOE) S&P 500 implied variance index VIX_t^2. All regressions contain VIX_t^2 as the dependent variable and VIX_{t-1}^2 as the independent variable, where VIX_{t-1}^2 denotes the lagged daily implied variance. The first column contains the variables used in the regression models. The second column reports the coefficient estimation results and measure of fit (R^2) for the model that depends on VIX_{t-1}^2. The third column shows the results for the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14). The fourth column contains the results for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2). The fifth column reports the results of the specifications that are based on the mutual-excitation intensity of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4) The sample is the S&P 500 E-mini future transaction prices and the daily close of the CBOE VIX index from June 2003 to December 2020.

Constant	0.014	0.010	0.014	0.004
	(3.55)	(2.55)	(2.95)	(0.62)
VIX_{t-1}^2	0.907	0.902	0.893	0.905
	(52.13)	(53.28)	(54.38)	(52.30)
NDB_t^+		-0.018		
		(-0.73)		
NDB_t^-		0.110		
		(7.71)		
VDB_t^+			0.028	
			(0.55)	
VDB_t^-			0.152	
			(2.46)	
IDB_t^+				-0.004
				(-1.38)
IDB_t^-				0.006
				(2.84)
R^2	0.963	0.964	0.966	0.964

The fourth column reports the estimation results for the models that exploit the effect of the positive and negative drift burst variations (VDB_t^+ and VDB_t^-) on VIX_t^2. The impact of VDB_t^+ is not significant, which contrasts with the positive effect of VDB_t^- and significant at the 5% level. This result is consistent with that of the occurrence measure, again corroborating that negative drift bursts, which are associated with large price drops, lead to a significantly higher degree of fear among market participants while positive bursts have limited effects on the market fears. The final column exhibits the impact of mutually-exciting intensity of positive bursts and negative bursts, denoted IDB_t^+ and IDB_t^-, on the implied variance VIX_t^2. I again observe an asymmetric effect, with IDB_t^+ not significantly affecting VIX_t^2 and IDB_t^- significantly increasing VIX_t^2 (at the 1% level). This suggests that days with a higher probability of negative drift burst occurrence leads to higher implied variance at the close of the day, whereas days with a chance of positive drift bursts almost have no influence on implied variance.

3.6. Impact on the variance risk premium

The variance risk premium is a phenomenon in the variance swap markets. The variance swap is an over-the-counter financial derivative. The seller of the two parties involved in a variance swap transaction will pay an amount based on the realized variance of the underlying asset. The buyer who wants to hedge the risk will pay a fixed amount, called the strike, specified at the start of the contract. At the end of the contract, the net payoff to the counterparties will depend on the difference between the actual, realised variance over the life of the swap contract and the strike variance.

Based on the historical trading data, the variance swap strike is observed to be greater than the realised variance, on average. In other words, the buyer of variance commonly ends up with a loss on the trade while the seller profits. This is also observed in volatility swap markets, where a forward contract on the future standard deviation of a given underlying asset. One reason for this phenomenon may be that swap buyers tend to dislike risk or be risk averse such that they are willing to pay extra money to hedge against the risk of the underlying assets in the future as they lock the risk in the future. This indicates that the loss of the swap buyer is the premium they pay for their risk aversion, and this premium is conventionally termed the variance risk premium (VRP).

VRP has been investigated extensively in financial studies and has been found to have important implications for volatility prediction and asset return forecasting (Bollerslev et al. 2009; Carr and Wu 2009; Arisoy 2010; Bollerslev et al. 2014; Prokopczuk and Simen 2014; Feunou et al. 2018; Li and Zinna 2018; Kilic and Shaliastovich 2019; Pyun 2019). In financial research, the VRP of an asset is generally proxied by the discrepancy between the implied and realized variance. The option implied variance, calculated from option prices, is a proxy for the swap strike

but is more accessible than swap strikes as options are publicly traded on exchanges. For the S&P 500 index-based variance swap contract with a one-month maturity, the variance risk premium is proxied by:

$$VRP_t = VIX_t^2 - E\left(RV_{t,t+22}\right), \tag{3.6.1}$$

where VRP_t denotes the proxied variance risk premium for the swap contract that starts from the opening of day $t+1$ and ceases at the end of day $t+22$, with 22 indicating the number of business days of one month. VIX_t^2 is the squared Chicago Board Options Exchange (CBOE) S&P 500 option implied variance index at the close of day t, which proxies the swap strike initially agreed by participants, according to the no-arbitrage theories by Carr and Madan (1998), Demeterfi et al. (1999), and Britten-Jones and Neuberger (2000), and $E\left(RV_{t,t+22}\right)$ is the expected next-month realized variance of the underlying asset (over the life of the swap contract). Following Caporin et al. (2017), $E\left(RV_{t,t+22}\right)$ is proxied by the exponentials of the fitted value of the model that uses logarithmic next-month realized variance $RV_{t,t+22}$ as the dependent variable, along with a constant, the logarithmic realized variance of the previous month, $RV_{t-22,t}$, and the logarithmic realized variance of the previous three month, $RV_{t-66,t}$, as the independent variables.

The daily realized variance RV_t for constructing these monthly realized volatilities, defined previously as the sum of squared 5-minute returns from 9:30 to 16:00 EST, is multiplied by the ratio between the sum of squared close-to-close S&P 500 daily returns and the average of RV_t in the sample. This follows Caporin et al. (2017) which is meant to take into account the contribution of overnight returns to the total variance. This scaling makes this section comparable to and consistent with the previous ones. Both $E\left(RV_{t,t+22}\right)$ and VIX are converted to the annualized percentage.

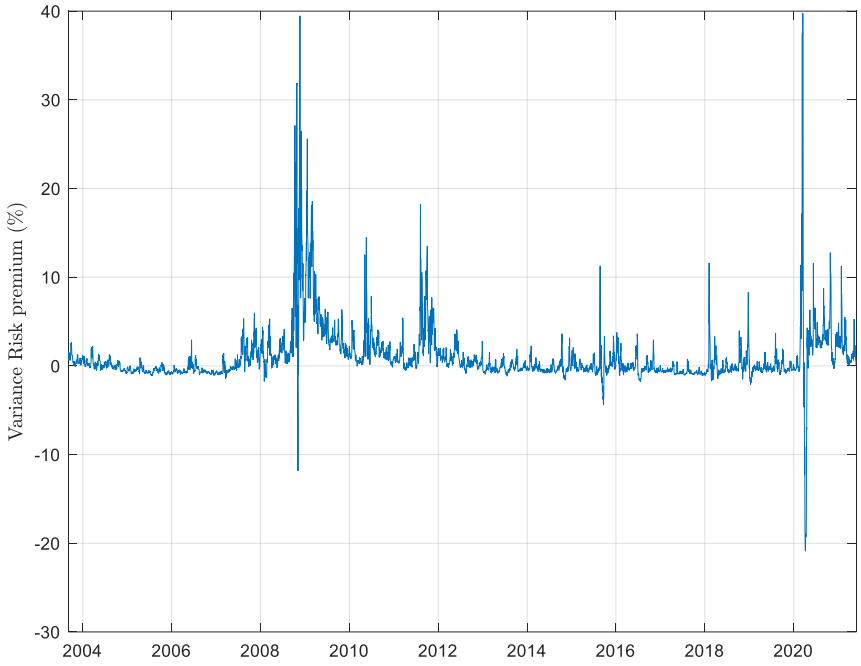
My findings in section 3.5 show that negative drift bursts have almost no impact on next-month realized variance while the negative bursts strongly increase the end-of-day implied variance, implying that negative bursts might widen the gap between the implied and realized variance thus increasing the variance risk premium. Additionally, negative bursts may be associated with large negative returns, therefore increasing the investors' willingness to hedge the downside risk. This intuition also suggests that the negative drift bursts might increase VRP.

Financial studies related to VRP centre on showing that VRP contributes importantly to future returns and volatility of the underlying asset (Bollerslev et al. 2009; Bollerslev et al. 2014; Prokopczuk and Simen 2014; Feunou et al. 2018; Kilic and Shaliastovich 2019). Studies also investigate the factors that may influence the VRP. Empirical studies by Carr and Wu (2009) show that the market portfolio, the size factor, and the book-to-market factor, introduced by Fama and French (1993), cannot explain VRP. Consistent with the results from theoretical models that non-Gaussian shocks to fundamentals have a substantial impact on risk premium (Barro 2006; Gabaix 2012; Drechsler 2013), Todorov (2010) makes use of a parametric model and shows that price jumps are linked to the variation in the variance risk premium. Caporin et al. (2017) also find that positive contemporaneous jumps (cojumps) among individual stocks increase VRP, and negative cojumps decrease VRP. To the best of my knowledge, the investigation of the impact of drift bursts on the variance risk premium is novel in the literature.

The time series of the variance risk premium (VRP) in my sample is shown in Figure 3.6.1. As expected, VRP is often positive and becomes much higher for the subprime mortgage crisis and pandemic recessions, similar to the findings in the related literature (Bollerslev et al. 2014; Caporin et al. 2017; Feunou et al. 2018; Kilic

and Shaliastovich 2019) and consistent with the intuition that extreme downside price moves dramatically raises the demand of options, leading to an increase in the implied variance thus levelling up the VRP. There are somewhat extreme negative VRP observations also during the period of the subprime mortgage crisis and pandemic recessions, which are also observed by these previous studies, but there lacks an explanation of the reason for their presence.

Figure 3.6.1. Variance risk premium.



Notes: The figure reports the daily variance risk premium, computed as in Equation (3.6.4), using the Chicago Board Options Exchange's CBOE Volatility Index VIX and expected monthly realized variance from the S&P 500 E-mini futures. The expected monthly realized variance is calculated by the exponentials of the fitted value of the model that uses logarithmic next-month realized variance as the dependent variable, along with a constant, the logarithmic realized variance of the previous month, and the logarithmic realized variance of the previous three months as the independent variables. The sample length is from June 2003 to December 2020.

To study the impact of drift bursts on the variance risk premium (VRP), I consider four regressions, with the dependent variable of these regressions uniformly set as VRP_t of day t. All models include the one-day lag of this premium, indicated by VRP_{t-1}, as an explanatory variable, and differ in the drift burst variables are used. The first model is the baseline model which only contains VRP_{t-1}. This model uses the persistence effect of VRP and is also applied in Caporin et al. (2017). The second model augments the first model by including recent positive and negative drift burst occurrences NDB_t^+ and NDB_t^-. The third is the regression that is based on the return variation caused by positive and negative drift bursts, indicated by VDB_t^+ and VDB_t^-. The fourth model explores the effect of mutual-excitation intensity of positive and negative drift bursts, IDB_t^+ and IDB_t^-.

Since the dependent variable VRP is constructed using monthly realized variance, the residuals of the regression may be autocorrelated, which results in underestimation of the variance of coefficient estimates and exaggerates the t-statistics of the coefficient, leading to an unreliable test of significance. To reduce this bias, the statistical inference of the coefficient is based on the Heteroskedasticity and Autocorrelation Consistent (HAC) Standard Errors with the number of the lags equal to $2 \times (21 + 1)$, where 21 is the size of overlapping window of two neighbouring realized variance. The detail of HAC calculations can be found in the appendix of Chapter 2.

Table 3.6.1 reports the results for these four models, including the coefficient estimation, along with the goodness of fit. The dependent variable of these models is the variance risk premium VRP_t. The first column contains the independent variables

that are used in the regression models. Each of the subsequent four columns reports the estimated coefficients and the measure of fit R^2 for one of the four models.

Table 3.6.1. The impact of drift bursts on the variance risk premium

This table reports the estimation results for four regressions, aiming to investigate the impact of different drift burst measures on variance risk premium. The dependent variable for all models is uniformly set as VRP_t, which is the variance risk premium on day t, and all these models include an independent variable VRP_{t-1}, which is the one-day lag of the variance risk premium. The first column contains the names of the explanatory variables used in the regression. The second column contains results for the model that depend on the one-day lag of variance risk premium. The third column reports the results for the model that uses the occurrence of positive and negative bursts, NDB_t^+ and NDB_t^-. The fourth column is for the specification that includes the variation of positive and negative drift bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2), and the fifth column contains results for the model that is based on the mutual-excitation intensity of positive and negative bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4), respectively. The sample data is S&P 500 E-mini future prices and daily close of the CBOE VIX index from June 2003 to December 2020.

Constant	0.102	0.075	0.089	-0.126
	(2.98)	(2.03)	(2.44)	(-1.28)
VRP_{t-1}	0.851	0.848	0.826	0.840
	(30.45)	(29.58)	(22.85)	(28.44)
NDB_t^+		0.415		
		(0.52)		
NDB_t^-		0.419		
		(3.28)		
VDB_t^+			0.610	
			(1.22)	
VDB_t^-			0.121	
			(3.02)	
IDB_t^+				-0.042
				(-1.46)
IDB_t^-				0.075
				(3.19)
R^2	0.723	0.725	0.737	0.725

The second column of the table contains the estimation results for the model that only relies on the one-day lagged variance risk premium, VRP_{t-1}. This is a variable measured by monthly implied and realized variance but with a one-day lag. The coefficient on VRP_{t-1} is positive and significant at the 1% level and the measure of fit, R^2, of the model is 0.78, which is very high, implying that the daily VRP is highly

persistent and this effect explains a substantial proportion of the variation of the VRP time series. This is consistent with Caporin et al. (2017), who find that VRP is positively and significantly correlated with its daily lag and the regression model exploiting this effect enjoys a large goodness of fit. The third column presents the model that uses the information on positive and negative drift burst occurrences, denoted by NDB_t^+ and NDB_t^-. As the results show, NDB_t^+ reduces VRP but this effect is not statistically significant. NDB_t^- has a significant and positive impact, at the 5% level, on VRP. The fourth column presents the estimation results for the regression model that is based on positive and negative drift burst variation, VDB_t^+ and VDB_t^-. I find that the impact of VDB_t^+ on VRP is not statistically distinguishable from zero, and VDB_t^- has a significant and positive impact on VRP.

The fourth column of Table 3.6.1 contains the estimation results of the model that includes the mutually-exciting intensity of positive and negative bursts, IDB_t^+ and IDB_t^-. IDB_t^+ does not play an important role in explaining the premium, consistent with the effect of NDB_t^+ reported in the second column of the table. I observe a strong positive effect of IDB_t^-, in line with the significance of NDB_t^- and VDB_t^-. Given the high degree of similarity in the information on intensity and occurrence (both are based on the arrival information of drift bursts), it is not surprising that these two measures perform fairly similarly in explaining VRP. The strong positive effects of all these four negative burst measures on VRP are consistent with the intuition that the presence of large negative return increase the investors' risk aversion as investors dislike downside risk.

Additionally, my conclusion of the significance of drift bursts in explaining VRP is not altered for alternative transformations of VRP proposed by Carr and Wu (2009),

including the volatility risk premium defined by $sqrt\_VRP = \sqrt{VIX_t^2} - \sqrt{E\left(RV_{t,t+22}\right)}$, and the log variance risk premium $log\_VRP = \log\left[VIX_t^2 / E\left(RV_{t,t+22}\right)\right]$. The results are attached in Appendix A.4.

The above results on the impact of various drift burst measures on VRP are based on the minimum threshold of detecting drift bursts suggested by Christensen et al. (2022). To reduce the false positives, I consider higher thresholds for detecting drift bursts including 4.5 and 5, which are also considered in Christensen et al. (2022). A higher threshold also allows one to detect those drift bursts that may be larger in size and more impactful. However, these drift bursts are extremely rare. For example, the test statistic with a threshold equal to 5 only detects 5 positive bursts and 12 negative bursts out of 4544 days. This is too rare to allow appropriate estimation of the mutual-excitation intensity. Thus, the analysis of the effects of the Hawkes (1971) mutual-excitation intensity, with drift bursts tested at thresholds 4.5 and 5, is excluded.

Table 3.6.2 reports the results of the coefficient estimation and measure of fit for models that harness the information of the occurrence measure of positive and negative bursts (NDB_t^+ and NDB_t^-) and exploit the variation of positive and negative bursts (VDB_t^+ and VDB_t^-), for explaining the variance risk premium (VRP). Columns two and three contain the estimation results of these two models for the drift burst test statistics with a threshold equal to 4.5. The coefficient estimates on these occurrences and return variations of drift bursts including NDB_t^+, NDB_t^-, VDB_t^+ and VDB_t^-, together with their statistical significance are qualitatively very consistent with those variables in Table 3.6.1. Specifically, the coefficients on positive burst measures (NDB_t^+ and VDB_t^+) are not significant, while these on negative drift burst variables (NDB_t^- and VDB_t^-) significantly increase VRP at the 5% level. These consistent

results imply that my findings of the effects of drift burst occurrence and variation on VRP are robust when bursts are identified using the threshold 4.5.

Table 3.6.2. The impacts of drift bursts detected by thresholds of 4.5 and 5 on the variance risk premium

Notes: This table replicates the results of Table 3.6.1 which investigates the impact of different drift burst measures on the variance risk premium, with the threshold of testing bursts of 4.5 and 5. All variables and models are the same as in Table 3.6.1, except for mutually exciting intensities that are not included.

threshold	4.5		5	
Constant	0.090	0.097	0.098	0.101
	(2.65)	(2.79)	(2.87)	(2.95)
VRP_{t-1}	0.848	0.835	0.850	0.850
	(29.32)	(25.47)	(30.46)	(30.46)
NDB_t^+	0.880		-0.429	
	(0.52)		(-1.29)	
NDB_t^-	0.642		1.616	
	(2.47)		(1.97)	
VDB_t^+		0.690		-0.315
		(1.68)		(-2.73)
VDB_t^-		0.103		0.132
		(3.02)		(2.03)
R^2	0.725	0.734	0.724	0.724

The fourth and fifth columns of Table 3.6.2 present the impacts of drift burst occurrences and variations on VRP with the threshold for testing drift bursts raised to 5. Again, NDB_t^- and VDB_t^- lead to significantly higher VRP at the 5% level. The results also reveal that VDB_t^+ has a significant and negative effect on VRP at the 5% level. This significance of VDB_t^+ contrasts with the weak evidence of VDB_t^+ reported in the left panel of Table 3.6.2 and Table 3.6.1, suggesting that the positive burst variation detected by increasing the threshold is important in predicting VRP. This may be because the positive bursts detected by a high threshold might be related to very large positive price movements that strongly decrease the investors' fears.

3.7. Conclusion

The constant and linear drift-diffusion models studied in Chapter 2 are unable to explain the stylized fact of short-lived and explosive drifts including gradual jumps and flash crashes or spikes in the intraday stock price patterns. To identify drift bursts, I apply the test statistic proposed by Christensen et al. (2022). Consistent with Christensen et al. (2022), my results show that the occurrence of drift bursts is not frequent. Drift bursts can be decomposed into positive and negative components, with the positive bursts capturing the positive gradual jumps and flash spikes and the negative counterparts identifying negative gradual jumps and flash crashes. I also find that the majority of drift bursts are associated with gradual jumps, due to the scarcity of flash crashes and flash spikes in the stock market. I find that the presence of drift bursts contributes approximately 20% of daily price variation on average, according to the return variation measure attributed to drift bursts. This novel burst variation measure also shows that drift bursts differ in magnitude over the sample and their levels are much greater in the subprime mortgage crisis and pandemic recessions.

Motivated by my observation of the clustering behaviour of drift bursts, I apply the mutual-excitation Hawkes (1971) process to model the arrival intensity of drift bursts. The model estimation results reveal significant self-exciting characteristics of both positive and negative drift bursts, and that positive and negative bursts mutually excite each other asymmetrically, with the impact from the negative bursts to positive drift bursts being much stronger.

As drift bursts are found often associated with large price returns relative to other intraday returns, they may impact volatility. I investigate if drift bursts forecast volatility. All three measures of drift bursts studied in this chapter, including occurrence, return variation, and intensity are applied to forecast volatility.

Surprisingly, the results show that both positive and negative drift bursts tend to have a very little impact on future realized variance (RV). However, I find that drift bursts have an impact on the volatility expected by investors measured by VIX. Negative drift bursts lead to a significantly higher VIX at the close of that day for all three drift burst variables. This may be because negative returns may raise the fear of investors. But, it is still puzzling that drift bursts affect VIX, but not RV.

I also investigate the impact of drift bursts on the variance risk premium (VRP), which is used in the literature to gauge the degree of investors' risk aversion. My empirical results show that across all three measures of drift bursts, negative bursts significantly increase VRP while positive drift bursts do not significantly affect VRP. The positive and significant effect of negative drift bursts is explained by noting that the large negative returns increase risk aversion. I also observe that the effect of positive drift burst variation becomes significant and negative when drift bursts are tested using a stricter threshold, which is in line with the intuition that a very large positive return can lower risk aversion.

Chapter 4. Empirical evidence of stock codrift variation and its implications for market volatility prediction

Abstract

This chapter examines contemporaneous drift variation (codrift variation) among individual stocks. I show, through a Monte Carlo study, that applying a coexceedance criterion to a univariate drift test by Shi and Phillips (2024) is feasible to detect codrift variation. However, one should not expect to detect all stocks that are associated with a codrift variation. My empirical results provide evidence of a relationship between market drift variation and underlying stock codrift variation. Importantly, I show that stock codrift variation helps forecast market-level volatility: Codrift variation among stocks have positively predict market volatility. Finally, I show that models exploiting the effects of codrift variation lead to significantly better out-of-sample market volatility forecasting performance.

4.1. Introduction

Recently, non-parametric tests employing high-frequency data have been developed to detect whether, in addition to a volatility diffusion, the variation of asset prices is also driven by drift. Tests which explicitly identify drifts have been developed by Flora and Renò (2020), Andersen et al. (2023), Laurent et al. (2022), and Christensen et al. (2022). The seminal work in this area is Laurent et al. (2022) and Shi and Phillips (2024) who separate realised measures of volatility into a component driven by continuous price changes and another component driven by drift variation.

The application of these non-parametric tests to various markets has supported the presence of price drift variation. For example, drift variation in equity indices and individual stocks have been documented in Jagannathan et al. (2019), Bellia et al. (2020), Flora and Renò (2020), Jagannathan et al. (2022), Christensen et al. (2022), and Laurent et al. (2022), among others. Moreover, evidence for the presence of drift variation in foreign exchange and Treasury bond markets is documented in Christensen et al. (2022) and Flora and Renò (2020).

Jumps are sudden large price moves, which are larger than what can be interpreted by the drift-diffusion model alone. Contemporaneous jump (cojump) variation among underlying stocks may indicate systematic large price move, which challenges the functionality of portfolio for diversification. Motivated by the importance of cojumps, literature investigates the contemporaneous jump variation among underlying stocks of a market portfolio or (Gilder et al. (2014).

However, little research has studied the simultaneous drift variation, termed contemporaneous drift variation (codrift variation), across different stocks. The importance of codrift variation is not ignorable. Codrift variation may capture the common bubbles and crashes, as drift variation are found to play a role in measuring

price bubbles and crashes (Christensen et al. 2022; Jagannathan et al. 2022; Shi and Phillips 2024). Since common bubbles and crashes cannot be diversified away, they have important implications for portfolio selection and asset allocation. Boninsegna and Candelon (2024) find that when common bubbles explode, while some techniques exist to diversify risk in individuals' portfolios, stock bubbles can erode all investors' profits. Anderson and Brooks (2014) demonstrate that much of the common variation in stock returns that can be attributable to market risk is due to common bubbles as opposed to being driven by fundamentals. Malceniece et al. (2019) show that a collective crash can increase the cost of capital for companies. Kole et al. (2006) analysis shows that in the presence of systematic crashes, the diversification possibilities erode rapidly, causing investors who face short sales constraints to withdraw from equity markets. Longin and Solnik (2001) show that equity market correlations increase after a common crash in the market, dampening the diversification potential of portfolio managers.

In this chapter, I explore codrift variation among individual stocks. Specifically, I investigate three research questions associated with codrift variation. The first equation is straightforward: how to detect codrift variation among stocks? As far as I am aware, there is no such method of detecting codrift variation in the literature. My first contribution is to demonstrate, through a Monte Carlo simulation study, that it is possible to detect codrift variation among stocks by applying the coexceedance rule to the univariate drift variation test by Shi and Phillips (2024). However, I show that using this codrift variation testing approach, one should not expect to detect all common drift variations comprising a codrift variation. Additionally, I also find that this codrift variation testing method may uncover some moderate and small systematic drift in the market that are not able to find via applying the single drift variation test to the market portfolio index.

My second research question is: is there an association between codrift variation among the underlying stocks and drift variation in the market portfolio? I refer to codrift variation which involves the market portfolio as systematic codrift variation since they represent non-diversifiable codrift variation. Intuitively, a systematic codrift variation results from market-level factors initiating a codrift variation amongst the underlying stocks, which is ultimately reflected as a drift variation in the market portfolio.

My second contribution is to document that there is an association between drift variation in the market portfolio and codrift variation among the underlying stocks. The number of stocks detected to be involved in a systematic codrift variation identified by our coexceedance-based detection methods is often moderate relative to the number of stocks included in my sample. However, given the evidence from my Monte Carlo study, I argue that the moderate numbers are consistent with the occurrence of systematic codrift variation, and it is very unlikely that they indicate independent or idiosyncratic drifts. Importantly, the results suggest there is a relationship between large drift variation in the market portfolio and large codrift variation in the underlying stocks. In other words, market-level factors are able to generate simultaneous drift variation in stock prices.

I also document non-systematic codrift variation which involve relatively large numbers of stocks. I am unable to explain these as being industry-specific codrift variation or related to drift variation in (proxies for) the Fama and French (FF) SMB and HML factors (Fama and French 1992,1993). I conclude that these non-systematic codrift variation are misclassified and result from a failure to detect some moderate and small drift variation in the market portfolio. Therefore, systematic codrift variation are likely to be more common than is suggested by the number of drift variation detected in the market portfolio.

Forecasting price volatility has important implications for the financial industry for asset pricing (Black and Scholes 1973), derivative pricing (Duffie et al. 2000), asset allocation (Merton 1969), and risk management (Christoffersen and Diebold 2000). Laurent et al. (2024) show that market drift variation positively predicts market volatility. Given the clear link between drift variation in the market portfolio and codrift variation among the underlying stocks, codrift variation should also predict market volatility. Therefore, my third research question is: do stock codrift variation forecast the volatility of the market portfolio?

My third contribution is to uncover the forecasting significance of codrift variation. Our empirical evidence shows that codrift variation in stocks positively predict the volatility of the market portfolio. Additionally, I also find the stock codrift variation outperform the market drift variation in forecasting market volatility for both in-sample and out-of-sample performance. The reason why codrift variation are more important may be that they can reflect some small and moderate drift variation that are unable to detect from the market portfolio index.

The rest of the chapter is organized as follows. Section 4.2 introduces the theories and methods to testing stock codrift variation. Section 4.3 presents the simulation study of evaluating the efficacy of the codrift variation test. Section 4.4 contains the data description of the market portfolio and individual stocks. Section 4.5 investigate the links between stock codrift variation and market drift variation. Section 4.6 discusses the implications of codrift variation for volatility forecasting. Section 4.7 concludes.

4.2. Codrift variation identification

4.2.1. Realized drift variation

This chapter assumes that the log prices p_t follow an Ito semi-martingale process,

$$dp_t = \mu_t dt + \sigma_s dW_t + J_t dN_t, 0 \leq t \leq T, \tag{4.2.1}$$

where μ_t is the drift coefficient, already introduced in Chapter 3, σ_s is a càdlàg and strictly positive volatility coefficient, W_t is a standard Brownian motion, J_t represents the random jump size at time t, and N_t is an independent Poisson counting process with a time-varying intensity.

As in Chapters 2 and 3, I assume that the log prices p_{t_i} are observed at M equally spaced intervals spanning the period $(0, \mathrm{T})$, $0 = t_0 < t_1 < t_2 \ldots < T$. The distance between two consecutive observation times is then denoted by $\Delta = t_i - t_{i-1} = 1/M$. The returns computed using log prices at equally spanned observation times may be written as $r_{t_i} = p_{t_i} - p_{t_{i-1}}$, with $i = 1,2, \ldots, M$.

The quadratic variation for the cumulative return process $r_t = p_t - p_0$, is given by,

$$\mathrm{QV}(0, \mathrm{T}) = \int_0^T \sigma_s^2 ds + \sum_{0<s\leq T} J_s^2, \tag{4.2.2}$$

where $\int_0^T \sigma_s^2 ds$ is the integrated volatility, and $\sum_{0<s\leq t} J_s^2$ is the sum of squared jumps that occurred between time 0 and time T. Andersen and Bollerslev (1998) define the realized variance or variation over the period $(0, \mathrm{T})$, by the summation of the M high-frequency intradaily squared returns,

$$RV_t = \sum_{i=1}^{M} r_{t_i}^2. \tag{4.2.3}$$

As shown by Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2002), RV_t converges to the quadratic variation as the interval between observations gets smaller,

$$RV_t = \sum_{i=1}^{M} r_{t_i}^2 \to \int_0^T \sigma_s^2 ds + \sum_{0<s\leq T} J_s^2 \quad \text{as } \Delta \to 0. \tag{4.2.4}$$

Note that the quadratic variation, Equation (4.2.2), does not contain the integrated drift variation, indicated by $\int_0^T (\mu_s)^2 ds$. This is because the integrated drift variation will be ignorable relative to the integrated volatility as the sampling interval becomes small (Laurent and Shi 2020; Christensen et al. 2022; Laurent et al. 2024; Shi and Phillips 2024). The negligibility of integrated drift variation with a diminishing sampling interval, results in a challenge of its measurement. However, nonzero drifts ($\mu_t \neq 0$) do exist in the log prices as price trends and are important to financial applications. Christensen et al. (2022) demonstrate that drift variation form an integral part of price dynamics across currencies, fixed-income investments, equities, and commodities; Laurent and Shi (2020) argue that ignoring drift variation can lead to significant finite sample bias in realized variance and severe size distortion in jump testing. Laurent et al. (2024) find the usefulness of drift variation in forecasting volatility.

Shi and Phillips (2024) and Laurent et al. (2024) allow for nonparametric identification of the integrated drift variation. Specifically, they define realized drift variation by,

$$RDV_t = M \sum_{i=2}^{M} r_{t_i} r_{t_{i-1}}, \tag{4.2.5}$$

where RDV_t denotes the realized drift variation of day t, and they uncover that for $\Delta \to 0$,

$$RDV_t \to \int_0^T (\mu_s)^2 ds. \tag{4.2.6}$$

The above asymptotic property indicates that the realized drift variation can capture the integrated drift variation. As discussed in Laurent et al. (2024), realized drift variation is robust to jumps. This is because jumps indeed will be asymptotically eliminated in the realized drift via the same mechanism which eliminates the bias in bipower variation, that is the product of two consecutive returns.

The realized drift is the central insight on which the theoretical and empirical results of this chapter build. Of course, nothing prevents the estimates of the squared drift variation defined by Equation (4.2.6) from becoming negative in a given finite sample. Therefore, following the suggestion of Shi and Phillips (2024), I truncate the actual empirical measurements at zero,

$$RDV_t^+ = \max(RDV_t, 0). \tag{4.2.7}$$

where RDV_t^+ denotes the positively truncated realized drift variation.

4.2.2. Testing for drift variation

The time series properties and forecasting value of the drift variation have been extensively studied by Laurent et al. (2024). The empirical results discussed by Laurent et al. (2024) rely on the above simple nonparametric drift variation estimates defined by positively truncated realized drift. As mentioned, the theoretical justification for this measurement is based on the notion of increasingly lower frequency, $\Delta \to 0$. Of course, any practical implementation with a fixed sampling interval, or $\Delta > 0$, is invariably subject to measurement error. The nonnegativity truncation in Equation (4.2.7) alleviates part of this finite sample issue by eliminating theoretically infeasible negative estimates for the squared drift variation. However, the resulting drift variation series in Laurent et al. (2024) arguably also displays an unreasonably large number of nonzero small positive values. It may be desirable to treat these small drift variations as measurement errors, linking only large values of RDV_t^+ with the drift variation. This section provides a theoretical framework for doing so.

The distributional results developed by Laurent et al. (2024) and extended in Shi and Phillips (2024), imply that, in the absence of drift variation, the following statistic is,

$$Z_t = \frac{RDV_t}{\sqrt{RiceQ_t * M}} \to N(0,1), \tag{4.2.8}$$

for $\Delta \to 0$, where $RiceQ_t$ that appears in the denominator is a realized quarticity measure proposed by Laurent et al. (2024) with $RiceQ_t = \frac{M}{6}\sum_{i=2}^{M} \left(r_{t_i} - r_{t_{i-1}}\right)^2 \left(r_{t_{i-1}} - r_{t_{i-2}}\right)^2$. The "significant" drift variation may therefore be identified by comparing the test statistics (Z_t) to the standard normal distribution.

As in Shi and Phillips (2024), I detect the "significant" drift variation by the realization of Z_t in excess of a critical value,

$$RDV_t^* = \max(RDV_t, 0)\, I_{Z_t > \Phi_{1-\alpha/2}^{-1}}, \tag{4.2.9}$$

where RDV_t^* denotes the significant realized drift variation estimator, I_x denotes the indicator function and Φ_x^{-1} denotes the inverse cumulative function of the normal distribution, α indicates the two-tailed significance level. Note that for $\Phi_{1-\alpha/2}^{-1} > 0$, the definitions in Equations (4.2.9) automatically guarantee that RDV_t^* is positive. Of course, the nonnegativity truncation imposed in Equation (4.2.7) underlying the empirical drift variation measurements employed in Laurent et al. (2024) corresponds directly to $\alpha = 1$.

4.2.3. Testing for codrift variation

Stock codrift variation indicate the tendency for drifts to arrive simultaneously among different stocks. As already discussed in the Introduction, codrift may help identify the large price comovement in asset returns, which is important to risk diversification of portfolios. I take one approach to codrift testing for individual stocks. Specifically, I use the following co-exceedance rule:

$$\sum_{j=1}^{N} I_{RDV_t^{(j)} > 0} \begin{cases} \geq 2 \text{ Codrift variation} \\ \leq 1 \text{ No Codrift variation} \end{cases},$$

where $I_{RDV_t^{(j)} > 0}$ is an indicator function taking the value one when a realized drift variation is larger than zero for asset j on day t. I use the coexceedance rule in

conjunction with the test of Shi and Phillips (2024) introduced above. Note that this codrift variation test is not able to identify intraday codrift variation. The intraday version of the codrift variation test could be constructed by applying the co-exceedance rule to the test statistic proposed by Christensen et al. (2022). However, I do not build this intraday codrift test in this chapter since the minimum available sampling frequency of my individual stock data (1 minute) is too sparse for this test. Of course, intraday codrift variation could a more accurate in terms of timing and are left for future research.

4.3. Monte Carlo study

This section aims to study the efficacy of the co-exceedance rule for detecting co-drift variation among individual assets. I consider the number of 41 assets, to be consistent with my later empirical analysis for 41 Dow Jones Industry (historical) constituents. I conduct a multivariate Monte Carlo simulation study to examine the functionality of the codrift variation test. The simulation results presented below are not intended to be comprehensive, but to reflect my empirical application of the codrift variation test.

4.3.1. Simulation set-up

I generate log prices for 41 assets from a modified version of the multivariate model used in Barndorff-Nielsen et al. (2011). My modifications insert a common intraday volatility pattern f_t as in Gilder et al. (2014) and use a drift component as in Shi and Phillips (2024) and Laurent et al. (2024). When $f_t = 1$ and the common drift component, indicated by $\phi(x)M_t$, are constant, for all t, our model has the same notation and equations as Barndorff-Nielsen et al. (2011). My simulated model is:

$$dY_t^{(j)} = \phi(x)M_t dt + f_t\left(dV_t^{(j)} + dF_t^{(j)}\right) + J_t, \tag{4.3.1}$$

$$dV_t^{(j)} = \rho^{(j)}\sigma^{(j)}dB_t^{(j)},$$

$$dF_t^{(j)} = \sqrt{1-(\rho^{(j)})^2}\sigma_t^{(j)}dW_t,$$

where $j = 1,2,3, \ldots, 41$, $dB_t^{(j)}$ and dW_t are independent Brownian motions, and $\phi(x)M_t$ denotes the systematic drift component for all 41 stocks. The instantaneous volatility of the continuous components is $\sigma_t^{(j)} = \exp\left(\beta_0^{(j)} + \beta_1^{(j)}\varrho_t^{(j)}\right)$ with $d\varrho_t^{(j)} = \alpha^{(j)}\varrho_t^{(j)}dt + dB_t^{(j)}$. The multiplicative term f_t introduces a U-shape intraday volatility pattern. I follow Andersen et al. (2012) and model f_t according to Hasbrouck (1999) where,

$$f_t = C + Aexp(-at) + Bexp\big(-b(1-t)\big), t \in [0,1], \tag{4.3.2}$$

with $A = 0.75$, $B = 0.25$, $C = 0.8893$, $a = 10$, and $b = 10$. As stated by Andersen et al. (2012), these parameters mean that volatility at $t = 0$ is three times the volatility at $t = 0.5$ and volatility at $t = 1$ is 1.5 times the volatility at $t = 0.5$. The innovation correlation (statistical leverage) for the continuous component of $Y_t^{(j)}$ and $\varrho_t^{(j)}$ is given by $\rho_t^{(j)}$ and there is a perfect correlation between the continuous components of the innovations in $Y_t^{(l)}$and $Y_t^{(m)}$. The correlation between the continuous components of $Y_t^{(l)}$and $Y_t^{(m)}$ is $\sqrt{1-(\rho^{(l)})^2}\sqrt{1-(\rho^{(m)})^2}$. J_t indicates price jumps. Following Andersen et al. (2023), J_t is a compound Poisson process with the intensity $p_x = 1/5$ and the size $\delta_t \sim N\big(0, \sigma_J^2\big)$, where $\sigma_J = 0.9\%$.

I use the following parameter values $\left(\beta_0^{(j)}, \beta_1^{(j)}, \alpha^{(j)}, \rho^{(j)}\right) = (-5/16, 1/8, -1/40, -0.83)$, for all i. Except for the common drift component $\phi(x)M_t dt$, these are identical to those used in Barndorff-Nielsen et al. (2011) and Gilder et al. (2014). For this drift component, the function, $\phi(x)$ takes a value of either -1 or 1 with equal probability, and common drift variation coefficient, M_t, is specificized through a drift burst model, as in Christensen et al. (2022), Laurent et al.

(2024), and Shi and Phillips (2024), and is also the same drift burst model introduced in Chapter 3,

$$M_t = \mu^e (1-t)^{-\alpha}. \tag{4.3.3}$$

Following Laurent et al. (2024), I fix $\alpha = 0.9$ and only allow this drift process to be controlled through the μ^e parameter only. A zero-drift model can be obtained by setting $\mu^e = 0$. Consistent with Laurent et al. (2024), volatility bursts are not considered in this simulation, and I use the same volatility trajectory for all models (and all replications) for ease of comparison.

Following Laurent et al. (2024), I allow the drift parameter μ^e to vary between 0 and 0.0071 so that the daily price change (absolute) is from 0 to 5%. The values of μ^e are not very informative except that the larger they are the stronger the drift. For ease of comparison and to offer more economically meaningful insights, I report the average daily returns caused by μ^e rather than the values of μ^e themselves, for my later analysis related to the size and power evaluation of the drift and codrift variation tests.

To simulate prices from the above differential Equation (4.3.1), I use its Euler discretisation. I simulate prices every second for 10^4 days, where I assume that there are 390 min in a day to match our empirical data. Hence, for each trading day, I simulate prices at times $t_i = i/n$, for $i = 1, \dots, n$ and $n = 23400$. In addition, I introduce microstructure noise by adding an i. i. d. random variable to the (log) price so that I sample,

$$P_{t_i}^{(j)} = Y_{t_i}^{(j)} + \varepsilon_{t_i}^{(j)}, \tag{4.2.4}$$

where $P_{t,i}^{(j)}$ is the log price of asset i sampled on day t at intraday time j, $\varepsilon_{t_i}^{(i)} \sim NID\left(0, \left(\omega_{t_i}^{(j)}\right)^2\right)$ and $\left(\omega_{t_i}^{(j)}\right)^2 = 0.001 \sum_{k=1}^{n} (\sigma_t^{(j)}(t_k/n)^2$. A noise-to-signal ratio of 0.001 is consistent with the estimates in Hansen and Lunde (2006). Following Laurent et al. (2024) and Shi and Phillips (2024), the simulated stock data is finally resampled at the 5-minute frequency for evaluating the efficacy of the codrift variation test.

4.3.2. Simulation results

The efficacy of the codrift variation test based on the co-exceedance rule is evaluated by its size and power performance. The size of the test denotes the maximum probability of incorrectly rejecting the null hypothesis of no codrift when it is true. The power of the test is the probability that the test correctly rejects the null hypothesis when the alternative hypothesis that a codrift exists is true. A lower size or a higher power indicates more satisfactory testing performance. Literature has applied the single test to detect the drift variation in the market portfolio. For the comparison, I also include the size and power analysis for applying a single drift variation test to the log prices of the market portfolio, which are proxied by aggregating the simulated log prices of those 41 stocks.

Table 4.2.1 reports the size and power of the single test of the market drift variation and the stock codrift variation test calculated on the simulated data. The column “Truncated drift” contains the power and size assessments related to the positively truncated realized drift of the market prices without pretesting. The columns named “Drift test” and “Codrift test” present the evaluation results of the

market drift variation test and stock codrift variation test, respectively, with the nominal significance level of these two tests uniformly set at 5%. The first column contains the different daily returns caused by the respective drift variations. A daily return of 0% indicates there is zero drift variation. Daily returns of 0.1% to 5% are caused by increased drift variation coefficients from $\mu^e = 0.00016$ (small drift variations) to 0.0031 (large drift variations). The notation "med.no" in the second column denotes the median of the number of stocks detected to be involved in the codrift variation.

Table 4.3.1. Simulated size and power of the drift variation and codrift variation tests

Notes: This table reports the simulated size and power of the market drift variation and stock codrift variation tests. The first column contains the absolute daily returns from 0% to 5% caused by an increased level of a drift variation from zero. The column "Truncated drift" contains the results related to the market drift variation, estimated by the positively truncated realized drift without pretesting. The columns "Drift test" and "Codrift test" present the results of the market drift variation test and the stock codrift variation test, respectively, with the nominal significance level of these two tests uniformly set at 5%. "med.no" denotes the median of the number of stocks detected to be involved in the codrift variation.

		Truncated drift	Drift test	Codrift test
Daily return 0%	size	0.508	0.052	0.625
	med.no			2
Absolute daily return 0.1%	power	0.511	0.072	0.728
	med.no			3
Absolute daily return 0.5%	power	0.631	0.161	0.863
	med.no			5
Absolute daily return 1%	power	0.844	0.264	0.986
	med.no			10
Absolute daily return 2%	power	0.900	0.886	1.000
	med.no			23
Absolute daily return 5%	power	0.986	0.914	1.000
	med.no			38

The first row reports the size of the positively truncated realized drift, the drift variation test with the 5% nominal level, and the codrift variation test with the 5% level when the drift variation is absent. I observe the drift variation test on the market is very close to the nominal size of 5%. This is consistent with Shi and Phillips (2024), who find that there is almost no size distortion of the drift variation test. The size of the codrift variation test, however, is larger than that of both the positive truncated drift variation and the single drift variation test at 5% normal level, indicating that the codrift variation test may be noisier than the market drift variation test for detecting a common drift variation.

The remaining rows contain the power performance of the positively truncated market drift variation, the market drift test, and the codrift variation test, as the daily average return increases from 0.1% to 5% due to a larger drift variation. The powers of both the positively truncated market drift variation and the market drift variation test are close to one when the return becomes large. This suggests the good power of these two methods for identifying large drift variation. However, when the drift variation is small and moderate, which leads to 0.1% and 0.5% returns, the powers of the market drift variation test appear not satisfactory, 0.072 and 0.16. This indicates the difficulty of the market drift variation test in finding small and moderate systematic drift variations in the market.

The power of the stock codrift test also increases with the level of return. The median number of stocks associated with codrift variation increases when return is larger in magnitude but is always smaller than the overall number of stocks (41). This indicates the difficulty in identifying all common drifts comprising a codrift when using the coexceedance-based detection methods, which may be due to the impact of idiosyncratic characteristics of individual stocks. For example, drift variations in the

stock with a larger volatility may be more difficult to detect. Compared to the market drift variation test, the codrift variation test exhibits much better power performance, especially for small and moderate drift variations (0.1% and 0.5% return levels). Moreover, the power of the codrift variation test is even greater than that of the positively truncated market drift variation, which does not depend on the statistical test. These results suggest that the stock codrift variation test can be more powerful than either testing or truncating drift variation from the overall market portfolio index alone, in terms of identifying small and moderate common drift variations in the market.

4.4. Data

I obtain tick-by-tick transaction prices of the SPDR S&P 500 Growth ETF (SPY) from Tick Data Inc. SPY is used as a proxy for the portfolio of the US stock market. As already introduced in Chapter 2, the sample ranges from January 2, 1997, to September 21, 2021, with 6222 trading days. The SPY Tick data is cleaned as in Chapter 2.

I retrieve data of individual stocks from Pi-trading Inc. The sample contains 1-minute transaction prices from December 2002 to July 2019. For the choice of the stocks, I consider all stocks that were ever a constituent of the Dow Jones Industry Average Index between December 2002 to July 2019, for $T = 3220$ days. Of the distinct constituents of the Dow Jones Industry Average over this period, I retain for our analysis the 41 that were continuously available for at least four years (Dow Inc. is excluded due to insufficient observations). These large-cap stocks represent a significant portion of the U.S. equity market. To be aligned with the sample period of individual stocks, the observations on SPY outside December 2002 to July 2019 are discarded.

I also obtain data of some other ETFs from Pi-trading Inc.: iShares Russell 2000 Index Fund (IWM), iShares S&P 500 Value Index Fund (IVE), and iShares S&P 500 Growth Index Fund (IVW). The sample length of these assets is uniformly set as December 2002 to July 2019, same as the sample period for the 41 individual stocks and SPY. Following Chapter 2 and 3, the prices of all of these stock and ETF are uniformly sampled at the 5-minute frequency from 9:30 to 16:00 (EST). Typically, the five-minute frequency is to alleviate the distortion from market microstructure noise (Ait-Sahalia et al. 2005; Bandi and Russell 2006; Hansen and Lunde 2006; Andersen et al. 2007b; Bandi and Russell 2008).

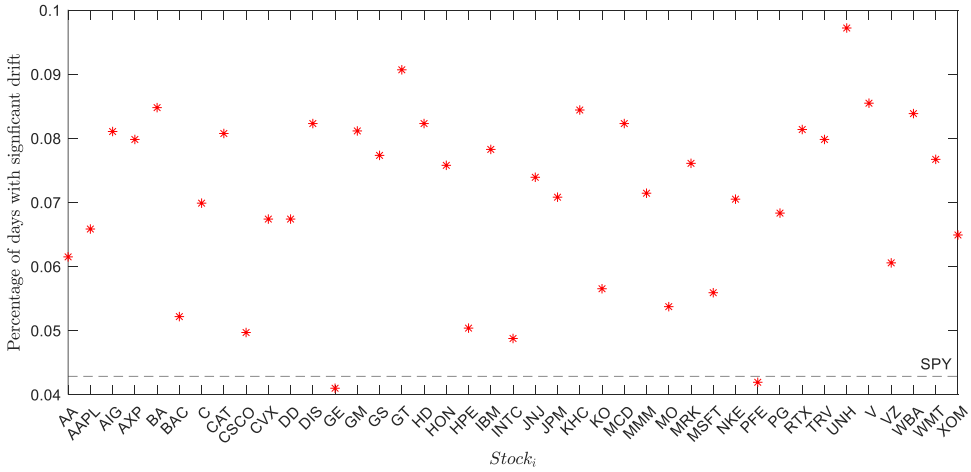
4.5. The stock codrift variation

This section aims to explore my second research equation of this chapter: are the codrift variation among the underlying stocks associated with the drift variation in the market portfolio? To clarify terminology, codrift variation involving the market portfolio proxy are said to be systematic codrift variation, whilst codrift variation amongst the individual stocks that exclude the market proxy are said to be non-systematic codrift variation.

4.5.1. Drift variations in the market and individual stocks

Before examining the presence of codrift variation based on the co-exceedance approach, I summarize results from applying the single drift test to the individual securities and SPY. The following Figure 4.5.1 depicts the proportion of sample days with significant drift variations detected in stocks and SPY. The dashed line denotes the percentage of drift variation in SPY while the red stars are those for stocks.

Figure 4.5.1. Proportion of sample days with significant drift variation detected



Notes: This figure depicts the proportion of sample days with significant drift variations in stocks and SPY. The dashed line denotes the percentage of drift variation in SPY while the red stars are that of stocks.

As the results show, drift variations are rare in market portfolios, with 4.2% of sample length. The significant drift variations appear to be more in individual stocks than in the market portfolio, which may be due to the idiosyncratic drift variation component in these stocks.

4.5.2. Codrift variation among individual stocks

Table 4.5.1 summarizes the codrift variation detected by the coexceedance-based codrift variation detection method for the 41 stocks. In Table 4.5.1, I report the codrift variation days, the median numbers of stocks found to be involved in the codrift variation, and the mean proportions of drift variations detected in the individual stocks which are involved in a codrift variation. The significance level for SPY drift variation testing is always fixed at 5%. Panel A summarises the results for all codrift variation, whilst Panels B and C summarise the results for systematic and non-systematic codrift variation, respectively. As the table shows, across all significance levels, a majority of codrift variation are non-systematic, with a relatively low proportion (about 5 %) of individual stock drifts being involved in systematic codrift variation. However, the median number of stocks involved in systematic codrift variation is significantly higher than the median number involved in non-systematic codrift variation for most detection methods. This provides answers to research question two: yes, there appears to be an association between drift variations in the market portfolio and codrift variation among the underlying stocks.

Table 4.5.1. Summary of codrift variation detected by coexceedance codrift detection method.

Notes: Panel A contains results for all codrift variation. Panel B contains results for systematic codrift variation. Panel C contains results for non-systematic codrift variation. The significance levels for testing drift in the market portfolio proxied by the SPDR S&P 500 ETF is fixed at 5%. Two sample Wilcoxon rank sum tests were conducted to test for a difference in the median number of stocks involved in systematic and non-systematic codrift variation. A * next to the median number of stocks involved in a systematic/non-systematic codrift indicates that the median is significantly larger at the 1% level.

Sig. stock	Days	Min no. stks	Med. no. stks	Max no. stks	Codrifts (%)
Panel A: All Codrift variation					
5.00%	2141	2	3	18	76.1
1.00%	1365	2	3	11	58.9
0.10%	687	2	2	10	40.8
0.05%	571	2	2	8	37.7
0.01%	359	2	2	6	30.3
Panel B: Systematic Codrift variation					
5.00%	131	2	6*	18	4.7
1.00%	116	2	4*	10	5.0
0.10%	85	2	3*	10	5.1
0.05%	75	2	3*	8	5.0
0.01%	54	2	3*	6	4.6
Panel C: Non-systematic Codrift variation					
5.00%	2010	2	3	14	71.5
1.00%	1249	2	3	11	53.9
0.10%	602	2	2	8	35.8
0.05%	496	2	2	8	32.7
0.01%	305	2	2	6	25.7

My Monte Carlo simulations in Section 4.3 demonstrates that it is difficult to detect all common drift variations comprising a codrift variation when using the coexceedance-based detection methods. This was true whether the common drift variations were large or small. Hence, although a moderate proportion of stocks were detected to be involved in the systematic codrift variation, they could represent true systematic codrift variation and more stocks (if not all) are involved in the systematic codrift variation.

4.5.3. Small drift variations in the market portfolio

One reason for detecting non-systematic codrift variation that involve a large number of stocks might be because they occur concomitantly with small, undetected, drifts in the market proxy. To investigate the likelihood of a type II error in the detection of a small drift variation in the market proxy when a large number of stocks are found to be involved in a non-systematic codrift variation, I calculate the p-values of the single drift variation test applied to the SPY and examine whether non-systematic codrift variation are associated with low p-values. Small p-values, greater than the significance level of 2.5% (one-tail), would be indicative of type II errors.

Table 4.5.2 summarises the number of non-systematic codrift variation and the maximum, median and minimum number of stocks detected to be involved in the non-systematic codrift variation when the p-values for the SPY are 5%, 25%, 37.5% and 50% (one-tail). I also report the number of non-systematic codrift variation associated with these p-values in the SPY which involve more than 5 stocks, where I consider 5 or more stocks to be a large number. The p-value below the median row indicates the two sample Wilcoxon rank test for the difference in the median number of the associated stocks and non-associated stocks. The significance level for detecting codrift variation is fixed at 5% two-tail or 2.5% one-tail, but I confirm that the results are qualitatively similar for more conservative significance levels.

Table 4.5.2. Non-systematic codrift variation associated with p-values in the SPY which are less than 5%, 25%, 37.5%, and 50%.
Notes: The significance level for detecting codrift is fixed at 5%.

SPY p-value (one tail)	<5%	<25%	<37.5%	<50%
No.	61	475	743	1018
No. (>5)	26	138	176	199
Max.	14	14	14	14
Med.	5	4	4	4
p-value	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
Min.	2	2	2	2

As the table shows, there is some evidence of non-systematic codrift variation involving large numbers of stocks being associated with type II errors in the detection of drift variation in the market portfolio proxy. By comparing Table 4.5.2 with Panel C of Table 4.5.1, it can be seen that many of the non-systematic codrift variation detected to involve more than 5 stocks are associated with p-values in the SPY for these less strict p-values. And for all of these scenarios, non-systematic codrift variation involving the maximum number of stocks are associated with a p-value for the SPY. Moreover, the median number of stocks associated with larger p-values of SPY is significantly greater than that of non-associated stocks. These findings indicate the link between large stock codrift variation with small market drift variation. This result may be interpreted by noting that, as found in my simulation results, the codrift variation test can be much more powerful than the drift variation test in detecting systematic drift variations.

4.5.4. Non-systematic codrift variation

After taking into consideration the non-systematic codrift variation that are likely to have been misclassified, there remain some that involve relatively large numbers of stocks. Table 4.5.3 explores two further possible explanations. The first is that they represent industry codrift variation. If all the stocks involved in a non-systematic codrift are from the same industry, then I refer to the non-systematic codrift variation as industry codrift variation. The second is that the codrift variation could be a result of sensitivity to other risk factors. In particular, I explore whether these codrift variation occur concomitantly with drifts in proxies for the FF SMB and HML factors (Fama and French 1992,1993).

Table 4.5.3. Summary of industry codrift variation (Panel A) and codrift variation with the Fama and French SMB and HML factors (Panel B).

Notes: In Panels A and B, the number of industry and FF codrift variation are reported along with the maximum, median and minimum number of stocks detected to be involved in the industry and FF codrift variation. In Panel A column 1, the number of stocks belonging to each industry is also reported. In Panel B, only FF codrift variation that occur in isolation of codrift variation with the market portfolio are reported, i.e., the association between drift variations in the SMB and HML factors and non-systematic codrift variation in the stocks is reported. In addition, I also report FF codrift variation in which both the SMB and HML factors drift (SMB and HML).

	Stks	No. codrift variation	Median no. stks	pval	Min no. stks	Max no. stks
Panel A: Industry codrift variation						
Industry						
Finance, Insurance and Real Estate	9	431	2	-	2	6
Manufacturing	24	1117	2	-	2	8
Other	2	14	2	-	2	2
Retail Trade	4	129	2	-	2	3
Services	2	11	2	-	2	2
Transportation, Communications, Electric, Gas and Sanitary service	3	27	2	-	2	2
Panel B: Industry codrift variation						
FF factor						
SMB	-	168	4	(0.0071)	2	9
HML	-	49	4	(0.0109)	2	8
SMB and HML	-	4	3	(0.6173)	2	5

To examine industry codrift variation, the stocks were sorted into 6 industry portfolios by SIC code. I used the industry allocations available from the website of the U.S. Securities and Exchange Commission. In Panel A of Table 4.5.3, I report the number of stocks belonging to each of the 6 industries, the number of industry codrift variation detected and the maximum, median and minimum number of stocks detected to be involved in each type of industry codrift variation. From Panel A, it can be seen that my sample of stocks is concentrated in the manufacturing industry. Thus, some of our non-systematic codrift variation which involve relatively large numbers of stocks might be explained by codrift variation occurring exclusively amongst the stocks within this industry. However, I do not find any evidence for this. The maximum number of stocks involved in the Manufacturing industry codrift variation tends to be low (<8 stocks), with many of the codrift variation involving just 2 stocks. The membership of stocks to the remaining industries is relatively low so codrift variation in these industries cannot explain non-systematic codrift variation involving large numbers of stocks.

In order to construct FF factor proxies, I follow Bannouh et al. (2012) in using intraday returns for ETFs that proxy small capitalisation, value and growth portfolios. Specifically, the SMB mimicking portfolio is formed by taking returns from the SPY and subtracting returns from the iShares Russell 2000 Index Fund (IWM). The HML mimicking portfolio is formed by taking returns on the iShares S&P 500 Value Index Fund (IVE) and subtracting returns from the iShares S&P 500 Growth Index Fund (IVW). Panel B of Table 4.5.3 reports the number of codrift variation between the SMB and HML factors and the stocks. To ensure that I isolate the influence from the FF factors, the FF codrift variation reported exclude times where the SMB and HML factors codrift with the SPY. However, I do report codrift variation in which both the

SMB and HML factors drift variations. I report the maximum, median and minimum number of stocks detected to be involved in the FF codrift variation. And I also report the p-value for comparing the median number of stocks related to these factor-based codrift variation with that of the remaining codrift variation.

From Panel B of Table 4.5.3, I do not find that codrift variation with the FF factor proxies explain non-systematic codrift variation involving large numbers of stocks. In general, the number of stocks involved in the FF codrift variation is low. Moreover, the difference between the median number of stocks associated with these factor-based codrift variation with that of the remaining codrift variation is not statistically significant, as indicated by the large p-values. These findings are unable to explain these as being a consequence of industry codrift variation or codrift variation associated with the FF, SMB, and HML risk factors. Hence, more systematic codrift variation might occur than is suggested by the drifts detected in the proxy for the market portfolio.

4.6. The impact of codrift variation on volatility forecasting

Volatility forecasting is important to the financial markets. Investors can use volatility data to align their portfolios with expected returns. Volatility is also an important input for investment, option pricing, and financial market regulation. Laurent et al. (2024) show that drift variations in the market positively predict the market's volatility. Given the strong link between the codrift variation among stocks and the drift variation in the market proxy, codrift variation among stocks might also lead to a higher market volatility. Motivated by this reason, this section tests the hypothesis that codrift variation positively forecast market volatility.

4.6.1. Model set-up

As in Chapters 2 and 3, I employ the Heterogeneous Autoregressive (HAR) type model suggested by Corsi (2009) for volatility forecasting. The original HAR model is the HAR-RV model, which exploits the effect of volatility persistence via the lags of daily, weekly, and monthly realized variance. The HAR-RV model has been widely applied to forecasting volatility due to its superior out-of-sample performance compared to autoregressive moving average models. This chapter uses the HAR-RV model as a benchmark model for exploring the impact of market drift variations on future market volatility.

Model 1- HAR-RV model:

$$RV_{t,t+h} = \beta_0 + \beta_d\, RV_t + \beta_w\, RV_{t-5,t} + \beta_m\, RV_{t-22,t} + \varepsilon_t, \qquad (4.6.1)$$

where

$$RV_{t-h,t} = \frac{1}{\mathrm{h}}\sum_{i=1}^{h} RV_{t-i}\,, h = 1,2,\dots,$$

and $RV_{t,t+h}$ is the h-day ahead average cumulative volatility,

$$RV_{t,t+h} = \frac{1}{\mathrm{h}}\sum_{i=1}^{h} RV_{t+i}.$$

Laurent et al. (2024) show that including the realized drift variation in the HAR-RV model significantly improves the in-sample and out-of-sample volatility forecast. They propose the following RDV$^+$ model,

Model 2- RDV$^+$ model: the RDV$^+$ model is formulated by including daily drift variation component to the HAR-RV model,

$$RV_{t,t+h} = \beta_0 + \beta_{rd}^{+} RDV_t^{+} + \beta_d RV_t + \beta_w RV_{t-5,t} + \beta_m RV_{t-22,t} + \varepsilon_t, \qquad (4.6.2)$$

where RDV_t^+ denotes the positive truncated realized drift variation, Equation (4.2.5). If the drift variation does not affect future volatility, we would expect to find $\beta_{rd}^{+} = 0$. Laurent et al. (2024) also consider weekly and monthly lags of realized drift variation. In this chapter, I focus on the daily lag only, to better evaluate the impact of a single drift variation on future volatility. The regression results that include the weekly and monthly components are attached in Appendix A.5.

As mentioned, the positively truncated realized drift may be subject to bias due to measurement errors, and a pretested realized drift variation may be more robust to this bias. Motivated by this, I modify the RDV$^+$ model by substituting the RDV_t^+ component with its statistically significant version, RDV_t^*, Equation (4.2.7). The

significance level for testing the realized drift variation is set as 5%, following Laurent et al. (2022a).

Model 3- RDV* model: the RDV* model is formulated by including recent pretested daily realized drift variation to the HAR-RV model,

$$RV_{t,t+h} = \beta_0 + \beta^*_{rdv}\, RDV^*_t + \beta_d\, RV_t + \beta_w\, RV_{t-5,t} + \beta_m\, RV_{t-22,t} + \varepsilon_t. \quad (4.6.3)$$

If drift variation does not predict future volatility, it is expected to see $\beta^*_{rdv} = 0$.

As discussed in the previous section, the codrift variation among stocks is closely related to the drift variation in the market portfolio. Therefore, the stock codrift variation might also lead to an increase in market volatility. To teat this hypothesis, I formulate the following coRDV model by augmenting the RDV* model by including the stock codrift variation.

Model 4- coRDV model: this model exploits the information from stock codrift variation,

$$RV_{t,t+h} = \beta_0 + \beta^*_{rdv}\, RDV_t + \beta^c_{rdv}\, coRDV_t + \beta_d\, RV_t + \beta_w\, RV_{t-5,t} + \beta_m\, RV_{t-22,t} + \varepsilon_t, \quad (4.6.4)$$

where $coRDV_t$ is defined by the mean of contemporaneous realized drift variation among the 41 individual stocks, which is to estimate the stock codrift variation. If there is no impact of the stock codrift variation on the market portfolio's volatility, we would expect to observe $\beta^c_{rdv} = 0$.

4.6.2. In-sample results

Table 4.6.1 reports the results of in-sample estimation and goodness of fit, for the HAR-RV, RDV$^+$, RDV*, and coRDV models, defined in the previous section. The column head contains the coefficient names in these models. Each row reports the estimation results for one model. The tables include four panels for the daily, weekly, monthly, and quarterly realized variance forecasts, respectively. The first row of each panel contains the estimation results for the HAR-RV model, which depends on the daily, weekly, and monthly volatility lags. I observe that the coefficients of these three volatility lags, indicated by β_d, β_w, and β_m, are positive and significant at the 5% level for all of these forecast horizons, implying statistically significant volatility persistence. This finding is consistent with Corsi (2009). The second row of these four panels reports the estimation results for the RDV$^+$ model exploiting the impact of the positively truncated realized drift estimator on volatility forecasting. I find that the coefficient of this version of the realized drift (β_{rdv}^+) is overwhelmingly positive for these forecast horizons. β_{rdv}^+ is significant at the 5% level for the shorter two horizons (h = 1 and h = 5) and the longest horizon (h = 66). This indicates that the recent drift variation can significantly increase future volatility. Compared to the HAR-RV model, the goodness of fit, indicated by R^2, of the RDV$^+$ model is 2.1% higher, indicating that including market drift variation also helps explain more of the changes in future volatility than the model that contains only realized variance. The usefulness of drift variation for stock market volatility forecasting is consistent with the findings of Laurent et al. (2024).

Table 4.6.1. In-sample estimation results for drift and codrift variation models for realized variance forecasting

Notes: This table reports the in-sample coefficient estimation and goodness of fit for the different models for daily (h=1), weekly (h=5), monthly (h=22), and quarterly realized variance forecast (h=66). The first row of each panel contains the results for the Heterogeneous Autoregressive with realized variance (HAR-RV) model, Equation (4.6.1). The second row of each panel reports the results of the RDV^{+} model which augments the HAR-RV by including the positively truncated realized drift variation, Equation (4.6.2). The third row of these panels presents the results of the RDV^{*}, which is the specification including the realized drift variation pretested by the 5% critical value, Equation (4.6.3). The final row in each panel reports the results of the coRDV model which adds the contemporaneous realized drift variation to the RDV^{*} model, Equation (4.6.4). The brackets are the Heteroskedasticity- and autocorrelation-consistent (HAC) robust t-statistics by Newey and West (1987). The estimation for the intercept is not reported.

	β_{rdv}^{+}	β_{rdv}^{*}	β_{rdv}^{c}	β_d	β_w	β_m	R^2
h=1				0.528	0.320	0.113	0.529
				(6.14)	(4.34)	(2.07)	
	1.211			0.486	0.328	0.107	0.540
	(1.96)			(5.55)	(4.55)	(1.94)	
		0.059		0.528	0.320	0.113	0.529
		(0.07)		(6.10)	(4.35)	(2.07)	
		-0.060	0.088	0.497	0.306	0.119	0.547
		(-0.07)	(2.18)	(6.25)	(4.39)	(2.21)	
h=5				0.372	0.354	0.189	0.632
				(6.86)	(4.21)	(2.70)	
	1.314			0.327	0.367	0.178	0.640
	(2.30)			(6.52)	(4.28)	(2.56)	
		-0.264		0.370	0.357	0.190	0.633
		(-0.52)		(6.70)	(4.24)	(2.75)	
		-0.378	0.059	0.355	0.350	0.187	0.642
		(-0.74)	(2.89)	(6.70)	(4.26)	(2.72)	
h=22				0.210	0.353	0.236	0.561
				(5.32)	(2.23)	(2.66)	
	0.395			0.199	0.355	0.232	0.563
	(1.27)			(4.46)	(2.27)	(2.65)	
		-0.250		0.211	0.352	0.238	0.562
		(-0.47)		(5.26)	(2.22)	(2.73)	
		-0.383	0.070	0.206	0.333	0.235	0.573
		(-0.73)	(2.94)	(4.55)	(2.32)	(2.74)	
h=66				0.133	0.221	0.247	0.305
				(4.02)	(2.51)	(4.27)	
	0.436			0.122	0.224	0.242	0.306
	(1.93)			(3.62)	(2.58)	(4.21)	
		-0.176		0.134	0.220	0.250	0.305
		(-0.47)		(3.97)	(2.49)	(4.38)	
		-0.329	0.096	0.124	0.204	0.239	0.319
		(-0.86)	(1.86)	(3.42)	(2.65)	(4.35)	

The third row of each panel presents the RDV^{*} model which uses the realized drifts pretested at the 5% significance level. As the results show, the coefficient of the

pretested realized drift variation, denoted by β^*_{rdv}, is not significant at any reasonable levels over the forecast horizons considered, and in all cases, the gain in R^2 from including this variable is very small, 0.1%. These results suggest that the significant drift variation provides limited information on forecasting volatility, which reconciles the findings in Chapter 3, where I find that the volatility forecasting value of the drift variation is weak when the drift variation is pretested using the method in Christensen et al. (2022). My results from both drift variation testing approaches in Shi and Phillips (2024) and Christensen et al. (2022) consistently indicate that significant drift variation does not help anticipate volatility. The limited impact of the significant drift variation on volatility forecasting is also found by Laurent et al. (2022a), who demonstrate that the significant realized drift variation is less useful than positively truncated realized drift variation for anticipating volatility. The reason why significant drift variation fails to forecast volatility may be explained by noting that, as found in my simulation results, although pretesting drift variation alleviates the nonzero values due to the measurement error, it may also neglect a large number of moderate and small systematic drift variations, which could contain important information on future volatility.

The last row of each panel reports the coRDV model, which is based on the recent stock codrift variation ($coRDV_t$). Across all of these horizons, $coRDV_t$ significantly increases future volatility at the 10% level for daily and quarterly forecasts, and at the 1% level for weekly and monthly forecasts, which contrasts with consistently weak evidence of the market's significant realized drift across all three horizons. The R^2 of the coRDV model always ranks highest for the forecast horizons considered, even greater than that of the RDV^+ model, which includes some small and moderate market drift variations by not testing for market drift variation. These results suggest that

codrift variation among underlying stocks may be more important than market drift variations for the market volatility forecasting. This superiority of codrift variation may be interpreted by noting that, as revealed in my simulation analysis, the codrift variation can better capture small and moderate systematic drift variations in the market, compared to the drift variations of the market portfolio. In other words, codrift variation can uncover some small and moderate systematic drift variations that are not detectable by looking at the market portfolio as a whole. These small and moderate systematic drift variations could improve market volatility forecasting.

4.6.3. Out-of-sample results

My in-sample analysis evaluates model performance by comparing the fitted values to the actual realized variance, using all available data up to and including T. However, this procedure could draw an overly optimistic picture of the model forecasting ability, because the least squared error method tends to take pains to avoid large prediction errors and is thus susceptible to overfitting.

An out-of-sample analysis may more realistically uncover the model forecasting ability. We can estimate the model based on data up to and including today, construct a forecast of tomorrow's realized variance, wait until tomorrow, record the forecast error defined as the difference between the realized variance forecast with the actual realized variance, re-estimate the model, make a new forecast of the next day and so forth. At the end of this exercise, we would have a sample of forecast errors, which would be truly out-of-sample and would provide a more realistic picture of the model's performance.

However, evaluating real-time forecasts is very time-consuming. Researchers commonly resort to "pseudo", or "simulated", out-of-sample analysis, which means mimicking the procedure described in the last paragraph, using some historical date earlier than the end of the sample as a starting point. The resulting forecasting errors are then used to get an estimate of the model's out-of-sample forecasting ability. In this chapter, I rely on this pseudo-out-of-sample forecasting method, following the implementation of this approach in the literature (Corsi 2009; Duong and Swanson 2015; Patton and Sheppard 2015; Bollerslev et al. 2016; Andersen et al. 2023). Specifically, I consider the h-step-ahead cumulative realized variance forecasts for the SPY starting on December 15, 2006, through to the end of the sample. The forecasts are based on re-estimating the parameters of the different models each day with a fixed length Rolling Window (RW) comprised of the previous 1000 days, as well as an Increasing Window (IW) or expanding window based on all of the available observations. The sample sizes for the increasing window for the SPY thus range from 1000 to 3201 days.

The following two loss functions are applied for measuring the forecast error of the different models:

(a) mean square error (MSE),

$$\mathrm{MSE}\left(RV_{t,t+h}, F_{t,t+h}\right) = \left(RV_{t,t+h} - F_{t,t+h}\right)^2,$$

(b) gaussian Quasi-likelihood (QLIKE) loss function:

$$\mathrm{QLIKE}\left(RV_{t,t+h}, F_{t,t+h}\right) = \frac{RV_{t,t+h}}{F_{t,t+h}} - \ln\frac{RV_{t,t+h}}{F_{t,t+h}} - 1,$$

where $F_{t,t+h}$ denotes the h-day-ahead forecast. The MSE and QLIKE are both unbiased loss functions (Patton 2011) and are widely applied by key research (Bollerslev et al. 2016; Andersen et al. 2023). A model with a loss function lower than another model on average indicates that this model generates smaller forecast errors thus more accurate forecasts.

Table 4.6.2 reports the ratio of the average MSE and QLIKE for the different models, relative to the losses of the HAR-RV model, for daily, weekly (h=5), monthly (h=22), and seasonal realized variance forecast (h=66).

Table 4.6.2. Out-of-sample realized variance forecast losses.

Notes: This table reports the average MSE and QLIKE for the different models relative to the losses of the HAR-RV model, for daily (h=1), weekly (h=5), monthly (h=22), and quarterly realized variance forecast (h=66). The HAR-RV model denotes the Heterogeneous Autoregressive with realized variance (HAR-RV) model, Equation (4.6.1). The RDV^{+} model denotes the model that augments the HAR-RV by including the positively truncated realized drift, Equation (4.6.2). The RDV^{*} is the specification which includes the HAR-RV model the realized drift pretested by the 5% critical value, Equation (4.6.3). The coRDV is the model which adds the realized drift to the HAR-RV model, Equation (4.6.4). The lowest loss in each row is indicated in **bold**.

			HAR-RV	RD^{+}	RD^{*}	coRD
h=1	RW	MSE	1.000	**0.971**	1.002	0.987
		QLIKE	1.000	**0.955**	0.998	0.985
	IW	MSE	1.000	**0.961**	1.000	0.986
		QLIKE	1.000	0.966	1.000	**0.941**
h=5	RW	MSE	1.000	0.980	1.005	**0.974**
		QLIKE	1.000	0.987	1.002	**0.974**
	IW	MSE	1.000	0.974	1.009	**0.973**
		QLIKE	1.000	0.981	1.002	**0.942**
h=22	RW	MSE	1.000	0.990	1.002	**0.965**
		QLIKE	1.000	0.990	1.000	**0.981**
	IW	MSE	1.000	0.992	1.004	**0.954**
		QLIKE	1.000	0.996	1.000	**0.953**
h=66	RW	MSE	1.000	0.996	1.003	**0.960**
		QLIKE	1.000	0.996	1.002	**0.987**
	IW	MSE	1.000	0.997	1.002	**0.950**
		QLIKE	1.000	0.999	1.000	**0.973**

The RDV^{+} model outperforms the HAR-RV model for all of these different scenarios while the RDV^{*} model tends to fail to improve on the standard HAR. This again reflects the importance of including moderate and small realized drift for volatility forecasting. Aside from some mixed evidence for the daily forecast, the coRDV model systematically exhibits the lowest average loss, which supports the superiority of including codrift variation in terms of volatility forecasting.

Of course, my samples of the forecast losses may be naturally subject to sampling variability. Sampling variability refers to the fact that the average loss will vary from one sample to the next. Therefore, it is important to use statistical tests to account for these variances when concluding whether the results of the average loss comparisons found in the sample data can be inferred as true for the full population. Therefore, I consider testing the difference in the loss function between two models via the Diebold–Mariano-West (DMW) statistic developed by Diebold and Mariano (1995) and West (1996), with adjustment to the Newey and West (1987) Heteroskedasticity and Autocorrelation Corrected (HAC) standard errors with lags set as $2(h-1)$.

Table 4.6.3 contains forecasting results tested by the DMW statistic. Each column contains the value of the DMW test statistic from comparing one pair of forecasting models. The left-most column compares the standard HAR-RV with the model that includes the recent contemporary realized drift ($coRDV_t$). The DMW test statistic is positive across all forecast horizons, and I also see that this coRDV model, with few exceptions, significantly outperforms HAR-RV. This indicates the superior out-of-sample performance of the codrift variation information over the lags of realized variance. The middle column compares models that differ in the daily drift variation information. The first model uses the drift variation of the market proxy, estimated by

the positively truncated realized drift (RDV_t^+), while the second model uses stock codrift variation ($coRDV_t$). I observe that $coRDV_t$ performs on par with RDV_t^+ at daily forecast but outperforms RDV_t^+ at other horizons. This again supports the better forecasting power of codrift variation among stocks over the drift variations of the market proxy, which does not alter the conclusions from my in-sample analysis.

Table 4.6.3. Out-of-sample realized variance forecast losses.

Notes: This table reports different pair comparisons of the models for the daily (h=1), weekly (h=5), monthly (h=22), and quarterly (h=66) realized variance forecast performance. The HAR-RV model denotes the Heterogeneous Autoregressive with realized variance (HAR-RV) model, Equation (4.6.1). The RDV$^+$ model denotes the model that augments the HAR-RV by including the positively truncated realized drift, Equation (4.6.2). The RDV* is the specification which includes the HAR-RV model the realized drift variation pretested by the 5% critical value, Equation (4.6.3). The coRDV is the model which adds the codrift variation to the RDV* model, Equation (4.6.4).

			HAR-RV vs. coRDV	RDV$^+$ vs. coRDV	RDV* vs. coRDV
h=1	RW	MSE	1.767	-0.821	1.639
		QLIKE	1.686	-0.107	4.071
	IW	MSE	1.286	-0.302	2.552
		QLIKE	4.725	1.596	4.374
h=5	RW	MSE	1.676	0.466	1.269
		QLIKE	3.539	0.946	1.459
	IW	MSE	1.715	0.063	1.767
		QLIKE	4.375	2.528	3.863
h=22	RW	MSE	2.949	1.896	1.987
		QLIKE	2.436	0.726	1.973
	IW	MSE	4.933	3.643	2.098
		QLIKE	6.286	5.866	2.897
h=66	RW	MSE	2.247	2.037	2.772
		QLIKE	2.929	1.533	2.157
	IW	MSE	5.704	5.105	2.648
		QLIKE	6.081	5.892	2.502

The final column of the table compares a forecasting model that includes recent pretested market drift variation information with the coRDV model that contains codrift variation. The information in stock codrift variation significantly improves out-of-sample forecast performance over the significant market drift variation, corroborating that stock codrift variation is superior to market drift variation for volatility forecasting.

4.7. Conclusion

Although price drift variation have been investigated in various markets, very few studies have studied codrift variation or the tendency of drift variation to occur simultaneously among different markets or assets. However, codrift variation may estimate the common bubbles and crashes, which are hard to diversify and thus have important implications for portfolio selection and asset allocation. This chapter explores codrift variation among the underlying stocks of S&P 500 index. I show, through a Monte Carlo study, that applying a co-exceedance criterion to a univariate drift variation test by Shi and Phillips (2024) is feasible to detect codrift variation among stocks, although one should not expect to detect all stocks that are involved with codrift variation. Empirical evidence shows a association between drift variation in the market portfolio proxy and codrift variation in the underlying stocks.

Volatility forecasting is important to risk management, asset pricing, and pricing derivatives. Laurent et al. (2024) find that drift variations positively forecast market volatility. Given the strong relationship between market drift variation and stock codrift variation, I also investigate the impact of codrift variation on market volatility. The results show that codrift variation lead to significantly higher market volatility. Importantly, stock codrift variation lead to better in-sample and out-of-sample market volatility forecasts than market drift variations. An explanation is that codrift variation among stocks may capture some small and moderate systematic drift variations that are not detectable in the market portfolio index.

Chapter 5. Conclusion

The logarithmic prices of financial assets are conventionally assumed to follow a drift-diffusion process with jumps. Since drift is asymptotically dominated by diffusion and jumps in high-frequency logarithmic prices, drifts have received little attention in the high-frequency literature. Chapter 2 shows that ignoring drifts may result in finite sample biases in the estimation of positive and negative realized semivariances and the semivariance-based signed jump estimator proposed by Barndorff-Nielsen et al. (2008), along with the microstructure noise robust versions of these estimators. I modified these realized estimators and jump test statistics using the same approach as Laurent and Shi (2020), and my finite sample theory for the modified volatility estimators and simulations shows significant improvement in the estimation accuracy of good and bad volatility, together with signed jumps.

The volatility forecasting results using both original and modified realized variance and bipower variation lead to volatility dependence effects. However, positive and negative realized semivariances exhibit an asymmetric effect on volatility forecasting while both modified positive and negative semivariances positively affect future volatility similarly. While the original signed jump estimator forecasts volatility asymmetrically, the modified estimator exhibits almost no forecasting significance. Based on the forecasting results of the modified estimators, I conclude that good and bad volatility may have similar positive impacts on future volatility and that signed jumps do not predict volatility, which are quite different from the findings in the existing literature. I show that the asymmetric impacts on future volatility of good and bad volatility and signed jumps found in the literature are almost exclusively due to the drift-related bias in the signed semivariance and signed jump estimators.

Chapter 2 studies the daily constant and linear drift, but does not consider intraday drift bursts, which are short-lived and explosive drift during a trading day. Recent literature argues that it is costly to neglect drift bursts in the price dynamic because drift bursts can model price fast bubbles and crashes. To help understand the dynamics of drift bursts, Chapter 3 investigates their time series behaviours. The results using the Hawkes process uncover that both positive and negative drift bursts exhibit self-exciting characteristics, and the negative drift bursts can excite positive drift bursts, but not vice versa. Motivated by the forecasting importance of drift found in Chapter 2, Chapter 3 also investigates the impacts of drift bursts on future volatility. The results show that positive drift bursts are not important for predicting both realized and implied variance, and negative drift bursts have a weak impact on realized variance, but significantly increase the implied variance. I show that negative drift bursts also leads to an increase in the variance risk premium.

Although the drift variation for one single asset is extensively studied in the literature, little research has investigated the codrift variation among various assets. Chapter 4 analyses the codrift variation across the most important underlying stocks of the US stock market. I show that it is feasible to detect codrift variation among stocks by applying a coexceedance criterion to a univariate drift test introduced by Shi and Phillips (2024). The results show that stock codrift variations detected using this method are often related to the market drift variation and thus are hard to diversify. Given the close relationship between market drift variation and underlying stock codrift variation, along with the significance of market drift variation in market volatility forecasting, I investigate the impact of codrift variation on market volatility. My results suggest that codrift variation among underlying stocks leads to significantly higher market volatility, and the models that use codrift variation

information significantly outperform those exploiting the impacts of market drift variation for out-of-sample market volatility forecast.

There are some potential extensions of this thesis. First, it will be interesting to extend the data of all chapters to broader markets such as bond, crude oil and crypto markets. Second, signed jumps have also been found useful for option pricing, forecasting credit spreads, bond risk premium, and cross-sectional returns in the literature. Later research may study whether the significance of signed jumps found in these studies is artificially shaped by the bias in the estimators due to a nonzero drift. Third, it is worthwhile to investigate if drift bursts predict equity risk premium. The downside drift bursts in Chapter 3 may be associated with large price drops. Risk-averse equity owners will demand a high equity premium to compensate for the extreme losses. Therefore, the downside drift bursts might play a role in explaining the equity risk premium. Fourth, multivariate modelling for drift bursts across different markets can be an interest to future research. For example, an extension of Chapter 3 could be developing multivariate Hawkes models to measure the drift burst contagion across global markets. Specifically, this model could study, if a drift burst in one region of the world, increases the intensity of drift both in the same region (self-excitation) as well as in other regions (cross-excitation).

References

Aalen, O. O. and Gjessing, H. K. 2004. Survival models based on the Ornstein-Uhlenbeck process. *Lifetime data analysis* 10, pp. 407-423.

Adrian, T. and Shin, H. S. 2010. Liquidity and leverage. *Journal of Financial Intermediation* 19(3), pp. 418-437.

Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R. J. 2015. Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics* 117(3), pp. 585-606.

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

Ait-Sahalia, Y., Mykland, P. A. and Zhang, L. 2005. How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies* 18(2), pp. 351-416.

Andersen, T. G. and Bollerslev, T. 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39(4), pp. 885-905. doi: 10.2307/2527343

Andersen, T. G., Bollerslev, T., Christoffersen, P. and Diebold, F. X. 2007a. Practical volatility and correlation modeling for financial market risk management. In: *The Risks of Financial Institutions*. University of Chicago Press, pp. 513-548.

Andersen, T. G., Bollerslev, T. and Diebold, F. X. 2007b. 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. and Labys, P. 2000. Great realizations. *Risk* 13, pp. 105-108.

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

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

Andersen, T. G., Li, Y., Todorov, V. and Zhou, B. 2023. Volatility measurement with pockets of extreme return persistence. *Journal of Econometrics* 237(2), pp. 1-27.

Anderson, K. and Brooks, C. 2014. Speculative bubbles and the cross-sectional variation in stock returns. *International Review of Financial Analysis* 35, pp. 20-31.

Anderson, T. G., Bollerslev, T., Diebold, F. X. and Vega, C. 2003. Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review* 93(1), pp. 38-62.

Ang, A., Chen, J. and Xing, Y. 2006. Downside risk. *The Review of Financial Studies* 19(4), pp. 1191-1239.

Areal, N. M. and Taylor, S. J. 2002. The realized variance of FTSE-100 futures prices. *Journal of Futures Markets* 22(7), pp. 627-648.

Arisoy, Y. E. 2010. Volatility risk and the value premium: Evidence from the French stock market. *Journal of Banking & Finance* 34(5), pp. 975-983.

Audrino, F. and Hu, Y. 2016. Volatility forecasting: Downside risk, jumps and leverage effect. *Econometrics* 4(1), p. 8.

Bajgrowicz, P., Scaillet, O. and Treccani, A. 2016. Jumps in high-frequency data: Spurious detections, dynamics, and news. *Management Science* 62(8), pp. 2198-2217.

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

Bandi, F. M. and Russell, J. R. 2006. Separating microstructure noise from volatility. *Journal of Financial Economics* 79(3), pp. 655-692.

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

Bannouh, K., Martens, M., Oomen, R. C. and van Dijk, D. J. 2012. Realized mixed-frequency factor models for vast dimensional covariance estimation. *ERIM Report Series Reference No. ERS-2012-017-F&A*.

Barigozzi, M., Brownlees, C., Gallo, G. M. and Veredas, D. 2014. Disentangling systematic and idiosyncratic dynamics in panels of volatility measures. *Journal of Econometrics* 182(2), pp. 364-384.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. 2011. Multivariate realised kernels: consistent positive semi-definite estimators of the

covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics* 162(2), pp. 149-169.

Barndorff-Nielsen, O. E., Kinnebrock, S. and Shephard, N. 2008. Measuring downside risk-realised semivariance. *CREATES Research Paper* 42(42). Available at: https://ssrn.com/abstract=1262194

Barndorff-Nielsen, O. E. and Shephard, N. 2002. Econometric analysis of realized variance 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. and 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. and 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., Hansen, P. R., Lunde, A. and Shephard, N. 2009. Realized kernels in practice: trades and quotes. *The Econometrics Journal* 12(3), pp. C1–C32.

Barndorff-Nielsen, O. E. and Shephard, N. 2001. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63(2), pp. 167-241.

Barone-Adesi, G., Giannopoulos, K. and Vosper, L. 1999. VaR without correlations for portfolios of derivative securities. *Journal of Futures Markets* 19(5), pp. 583-602.

Barro, R. J. 2006. Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics* 121(3), pp. 823-866.

Bates, D. S. 1996. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies* 9(1), pp. 69-107.

Bauwens, L. and Hautsch, N. 2009. Modelling financial high frequency data using point processes. In: *Handbook of Financial Time Series*. Springer, pp. 953-979.

Bawa, V. S. and Lindenberg, E. B. 1977. Capital market equilibrium in a mean-lower partial moment framework. *Journal of Financial Economics* 5(2), pp. 189-200.

Bee, M., Dupuis, D. J. and Trapin, L. 2016. Realizing the extremes: Estimation of tail-risk measures from a high-frequency perspective. *Journal of Empirical Finance* 36, pp. 86-99.

Bekaert, G., Hoerova, M. and Duca, M. L. 2013. Risk, uncertainty and monetary policy. *Journal of Monetary Economics* 60(7), pp. 771-788.

Bellia, M., Christensen, K., Kolokolov, A., Pelizzon, L. and Renò, R. 2020. High-frequency trading during flash crashes: Walk of fame or hall of shame? SAFE Working Paper No. 270.

Bellia, M., Christensen, K., Kolokolov, A., Pelizzon, L. and Renò, R. 2023. Do designated market makers provide liquidity during a flash crash? *SAFE Working Paper* No. 270.

Berkowitz, J. and O'Brien, J. 2002. How accurate are value-at-risk models at commercial banks? *The Journal of Finance* 57(3), pp. 1093-1111.

Black, F. 1976. Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business & Economic Statistics Section, 1976*.

Black, F. 1986. Noise. *The Journal of Finance* 41(3), pp. 528-543.

Black, F. and Scholes, M. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81(3), pp. 637-654.

Bollerslev, T. 2022. Realized semi (co) variation: Signs that all volatilities are not created equal. *Journal of Financial Econometrics* 20(2), pp. 219-252.

Bollerslev, T., Gibson, M. and Zhou, H. 2011. Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Journal of Econometrics* 160(1), pp. 235-245.

Bollerslev, T., Li, S. Z. and Zhao, B. 2020. Good volatility, bad volatility, and the cross section of stock returns. *Journal of Financial and Quantitative Analysis* 55(3), pp. 751-781.

Bollerslev, T., Litvinova, J. and Tauchen, G. 2006. Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics* 4(3), pp. 353-384.

Bollerslev, T., Marrone, J., Xu, L. and Zhou, H. 2014. Stock return predictability and variance risk premia: Statistical inference and international evidence. *Journal of Financial and Quantitative Analysis* 49(3), pp. 633-661.

Bollerslev, T., Medeiros, M. C., Patton, A. J. and Quaedvlieg, R. 2021. From zero to hero: Realized partial (co) variances. *Journal of Econometrics* 231(2), pp. 348-360.

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

Bollerslev, T., Tauchen, G. and Zhou, H. 2009. Expected stock returns and variance risk premia. *The Review of Financial Studies* 22(11), pp. 4463-4492.

Bollerslev, T. and Todorov, V., 2011. Tails, fears, and risk premia. *The Journal of Finance*, 66(6), pp.2165-2211.

Boninsegna, G. and Candelon, B. 2024. *Analysis of Stock Bubbles: Detection, Impact on Market Stability, and Portfolios' Reactions.* University of Louvain.

Bormetti, G., Calcagnile, L. M., Treccani, M., Corsi, F., Marmi, S. and Lillo, F. 2015. Modelling systemic price cojumps with Hawkes factor models. *Quantitative Finance* 15(7), pp. 1137-1156.

Bouamara, N., Boudt, K., Laurent, S. and Neely, C. J. 2023. Sluggish news reactions: A combinatorial approach for synchronizing stock jumps. *arXiv preprint arXiv:2309.15705*.

Bowsher, C. G. 2007. Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics* 141(2), pp. 876-912.

Braun, T., Fiegen, J. A., Wagner, D. C., Krause, S. M. and Guhr, T. 2018. Impact and recovery process of mini flash crashes: An empirical study. *Plos one* 13(5), p. e0196920.

Britten-Jones, M. and Neuberger, A. 2000. Option prices, implied price processes, and stochastic volatility. *The Journal of Finance* 55(2), pp. 839-866.

Buccheri, G. and Corsi, F. 2021. HARK the SHARK: Realized variance modeling with measurement errors and nonlinear dependencies. *Journal of Financial Econometrics* 19(4), pp. 614-649.

Busch, T., Christensen, B. J. and Nielsen, M. Ø. 2011. The role of implied variance in forecasting future realized variance and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics* 160(1), pp. 48-57.

Caporin, M. 2023. The role of jumps in realized variance modeling and forecasting. *Journal of Financial Econometrics* 21(4), pp. 1143-1168.

Caporin, M., Kolokolov, A. and Renò, R. 2017. Systemic co-jumps. *Journal of Financial Economics* 126(3), pp. 563-591.

Carr, P. and Madan, D. 1998. Towards a theory of volatility trading. *Volatility: New Estimation Techniques for Pricing Derivatives* 29, pp. 417-427.

Carr, P. and Wu, L. 2009. Variance risk premiums. *The Review of Financial Studies* 22(3), pp. 1311-1341.

CFTC and SEC. 2010. *Findings regarding the market events of may 6, 2010. report of the staffs of the CFTC and SEC to the joint advisory committee on emerging regulatory issues*. Available at: https://www.sec.gov/news/studies/2010/marketevents-report.pdf

CFTC and SEC. 2011. *Recommendations regarding regulatory responses to the market events of may 6, 2010. report of the Joint CFTC-SEC Advisory Committee on Emerging Regulatory Issues*. Available at: https://www.sec.gov/spotlight/sec-cftcjointcommittee/021811-report.pdf

Chernov, M. and Ghysels, E. 2000. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics* 56(3), pp. 407-458.

Christensen, B. J. and Nielsen, M. Ø. 2007. The effect of long memory in volatility on stock market fluctuations. *The Review of Economics and Statistics* 89(4), pp. 684-700.

Christensen, K., Oomen, R. and Renò, R. 2022. The drift burst hypothesis. *Journal of Econometrics* (227), pp. 461-497.

Christensen, K., Oomen, R. C. and Podolskij, M. 2014. Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics* 114(3), pp. 576-599.

Christensen, K., Siggaard, M. and Veliyev, B. 2023. A machine learning approach to volatility forecasting. *Journal of Financial Econometrics* 21(5), pp. 1680-1727.

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

Christoffersen, P. F. and Diebold, F. X. 2000. How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics* 82(1), pp. 12-22.

Clements, A. and Liao, Y. 2017. Forecasting the variance of stock index returns using jumps and cojumps. *International Journal of Forecasting* 33(3), pp. 729-742.

Clements, A. and Preve, D. P. 2021. A practical guide to harnessing the har volatility model. *Journal of Banking & Finance* 133, p. 106285.

Clements, M. P., Galvão, A. B. and Kim, J. H. 2008. Quantile forecasts of daily exchange rate returns from forecasts of realized variance. *Journal of Empirical Finance* 15(4), pp. 729-750.

Cochrane, J. 2009. *Asset pricing: Revised edition*. Princeton university press.

Corradi, V., Distaso, W. and Mele, A. 2013. Macroeconomic determinants of stock volatility and volatility premiums. *Journal of Monetary Economics* 60(2), pp. 203-220.

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

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

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

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

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

Demeterfi, K., Derman, E. and Kamal, M. 1999. A guide to volatility and variance swaps. *Journal of Derivatives* 6(4), pp. 9-32.

Deo, R., Hurvich, C. and Lu, Y. 2006. Forecasting realized variance using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment. *Journal of Econometrics* 131(1-2), pp. 29-58.

Diebold, F. X. and Mariano, R. S. 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* 13(3).

Drechsler, I. 2013. Uncertainty, time-varying fear, and asset prices. *The Journal of Finance* 68(5), pp. 1843-1889.

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

Duong, D. and 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.

Eraker, B. 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *The Journal of Finance* 59(3), pp. 1367-1403.

Etienne, X. L., Irwin, S. H. and Garcia, P. 2014. Bubbles in food commodity markets: Four decades of evidence. *Journal of International Money and Finance* 42, pp. 129-155.

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

Fama, E. F. and French, K. R. 1992. The cross-section of expected stock returns. *The Journal of Finance* 47(2), pp. 427-465.

Fama, E. F. and French, K. R. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), pp. 3-56.

Fan, J., Imerman, M. B. and Dai, W. 2016. What does the volatility risk premium say about liquidity provision and demand for hedging tail risk? *Journal of Business & Economic Statistics* 34(4), pp. 519-535.

Farago, A. and Tédongap, R. 2018. Downside risks and the cross-section of asset returns. *Journal of Financial Economics* 129(1), pp. 69-86.

Félez-Viñas, E. 2017. *Market fragmentation, mini flash crashes and liquidity*.

Fengler, M. R., Mammen, E. and Vogt, M. 2015. Specification and structural break tests for additive models with applications to realized variance data. *Journal of Econometrics* 188(1), pp. 196-218.

Ferguson, T. S. 1999. Asymptotic joint distribution of sample mean and a sample quantile. *unpublished: http://www. math. ucla. edu/~ tom/papers/unpublished/meanmed. pdf* 5.

Feunou, B., Jahan-Parvar, M. R. and Okou, C. 2018. Downside variance risk premium. *Journal of Financial Econometrics* 16(3), pp. 341-383.

Feunou, B. and Okou, C. 2019. Good volatility, bad volatility, and option pricing. *Journal of Financial and Quantitative Analysis* 54(2), pp. 695-727.

Fleming, J. and Kirby, C. 2003. A closer look at the relation between GARCH and stochastic autoregressive volatility. *Journal of Financial Econometrics* 1(3), pp. 365-419.

Fleming, J., Kirby, C. and Ostdiek, B. 2001. The economic value of volatility timing. *The Journal of Finance* 56(1), pp. 329-352.

Fleming, J., Kirby, C. and Ostdiek, B. 2003. The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics* 67(3), pp. 473-509.

Flora, M. and Renò, R. 2020. V-shapes. *Available at SSRN 3554122*.

Gabaix, X. 2012. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *The Quarterly Journal of Economics* 127(2), pp. 645-700.

Gao, L., Han, Y., Li, S. Z. and Zhou, G. 2018. Market intraday momentum. *Journal of Financial Economics* 129(2), pp. 394-414.

Ghysels, E., Santa-Clara, P. and Valkanov, R. 2005. There is a risk-return trade-off after all. *Journal of Financial Economics* 76(3), pp. 509-548.

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

Gong, X. and Lin, B. 2021. Effects of structural changes on the prediction of downside volatility in futures markets. *Journal of Futures Markets* 41(7), pp. 1124-1153.

Goyal, A. and Saretto, A. 2009. Cross-section of option returns and volatility. *Journal of Financial Economics* 94(2), pp. 310-326.

Gupta, A. K. and Nadarajah, S. 2005. On the moments of the beta normal distribution. *Communications in Statistics-Theory and Methods* 33(1), pp. 1-13.

Gutierrez, L. 2013. Speculative bubbles in agricultural commodity markets. *European Review of Agricultural Economics* 40(2), pp. 217-238.

Hansen, P. R. and Lunde, A. 2006. Realized variance and market microstructure noise. *Journal of Business & Economic Statistics* 24(2), pp. 127-161.

Harlow, W. V. 1991. Asset allocation in a downside-risk framework. *Financial Analysts Journal* 47(5), pp. 28-40.

Hasbrouck, J. 1999. The dynamics of discrete bid and ask quotes. *The Journal of Finance* 54(6), pp. 2109-2142.

Hautsch, N. and Podolskij, M. 2013. Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: theory, implementation, and empirical evidence. *Journal of Business & Economic Statistics* 31(2), pp. 165-183.

Hawkes, A. G. 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58(1), pp. 83-90.

Heston, S. L. and Nandi, S. 2000. A closed-form GARCH option valuation model. *The Review of Financial Studies* 13(3), pp. 585-625.

Hogan, W. W. and Warren, J. M. 1972. Computation of the efficient boundary in the ES portfolio selection model. *Journal of Financial and Quantitative Analysis* 7(4), pp. 1881-1896.

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

Hyung, N. and De Vries, C. G. 2005. Portfolio diversification effects of downside risk. *Journal of Financial Econometrics* 3(1), pp. 107-125.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M. and Vetter, M. 2009. Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic processes and their applications* 119(7), pp. 2249-2276.

Jagannathan, R., Pelizzon, L., Schaumburg, E., Getmansky Sherman, M. and Yuferova, D. 2019. Liquidity provision: Normal times vs Crashes. *Available at SSRN 3079593*.

Jagannathan, R., Pellizon, L., Schaumburg, E., Sherman, M. G. and Yuferova, D. 2022. Recovery from fast crashes: Role of mutual funds. *Journal of Financial Markets*, p. 100646.

Jiang, G. J. and Zhu, K. X. 2017. Information shocks and short-term market underreaction. *Journal of Financial Economics* 124(1), pp. 43-64.

Johannes, M. S., Polson, N. G. and Stroud, J. R. 2009. Optimal filtering of jump diffusions: Extracting latent states from asset prices. *The Review of Financial Studies* 22(7), pp. 2759-2799.

Kilic, M. and Shaliastovich, I. 2019. Good and bad variance premia and expected returns. *Management Science* 65(6), pp. 2522-2544.

Klebaner, F., Landsman, Z., Makov, U. and Yao, J. 2017. Optimal portfolios with downside risk. *Quantitative Finance* 17(3), pp. 315-325.

Kole, E., Koedijk, K. and Verbeek, M. 2006. Portfolio implications of systemic crises. *Journal of Banking & Finance* 30(8), pp. 2347-2369.

Kolokolov, A. and Renò, R. 2023. Jumps or staleness? *Journal of Business & Economic Statistics* 00(0), pp. 1-27.

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

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

Laly, F. and Petitjean, M. 2020. Mini flash crashes: Review, taxonomy and policy responses. *Bulletin of Economic Research* 72(3), pp. 251-271.

Large, J. 2007. Measuring the resiliency of an electronic limit order book. *Journal of Financial Markets* 10(1), pp. 1-25.

Laurent, S., Renò, R. and Shi, S. 2022a. Realized drift. *Available at SSRN 4084647*.

Laurent, S., Renò, R. and Shi, S. 2022b. Realized drift. SSRN.

Laurent, S., Renò, R. and Shi, S. 2022c. *Realized Drift*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4084647

Laurent, S., Renò, R. and Shi, S. 2024. Realized drift. *Journal of Econometrics*, p. 105813.

Laurent, S. and Shi, S. 2020. Volatility estimation and jump detection for drift–diffusion processes. *Journal of Econometrics* 217(2), pp. 259-290.

Lee, H. H. and Hyun, J.-S. 2019. The asymmetric effect of equity volatility on credit default swap spreads. *Journal of Banking & Finance* 98, pp. 125-136.

Lee, S. S. and Mykland, P. A. 2012. Jumps in equilibrium prices and market microstructure noise. *Journal of Econometrics* 168(2), pp. 396-406.

Lettau, M., Maggiori, M. and Weber, M. 2014. Conditional risk premia in currency markets and other asset classes. *Journal of Financial Economics* 114(2), pp. 197-225.

Li, J. and Zinna, G. 2018. The variance risk premium: Components, term structures, and stock return predictability. *Journal of Business & Economic Statistics* 36(3), pp. 411-425.

Li, Y., Nolte, I., Nolte, S. and Yu, S. 2022. Testing for Jumps in a Discretely Observed Price Process with Endogenous Sampling Times. *Sandra and Yu, Shifan, Testing for Jumps in a Discretely Observed Price Process with Endogenous Sampling Times (October 6, 2022)*.

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

Longin, F. and Solnik, B. 2001. Extreme correlation of international equity markets. *The Journal of Finance* 56(2), pp. 649-676.

Lyócsa, Š. and Molnár, P. 2016. Volatility forecasting of strategically linked commodity ETFs: gold-silver. *Quantitative Finance* 16(12), pp. 1809-1822.

Ma, F., Liao, Y., Zhang, Y. and Cao, Y. 2019. Harnessing jump component for crude oil volatility forecasting in the presence of extreme shocks. *Journal of Empirical Finance* 52, pp. 40-55.

Maheu, J. M. and McCurdy, T. H. 2011. Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics* 160(1), pp. 69-76.

Malceniece, L., Malcenieks, K. and Putniņš, T. J. 2019. High frequency trading and comovement in financial markets. *Journal of Financial Economics* 134(2), pp. 381-399.

Markowitz, H. M. 1959. *Portfolio Selection: Efficient Diversification of Investments*. London: Yale University Press.

Markowitz, H. M. 1991. Foundations of portfolio theory. *The Journal of Finance* 46(2), pp. 469-477.

McAleer, M. and Medeiros, M. C. 2008. A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics* 147(1), pp. 104-119.

Melino, A. and Turnbull, S. M. 1990. Pricing foreign currency options with stochastic volatility. *Journal of Econometrics* 45(1-2), pp. 239-265.

Merton, R. C. 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics* 51(3), pp. 247-257.

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

Mizrach, B., Swanson, N. R. and Yu, B. 2018. Disentangling the Effects of News, Small Jumps, and Large Jumps on Stock Return Predictability.
Available at:
https://econweb.rutgers.edu/nswanson/papers/Signed%20jump%20factors%2008-01-2018.pdf

Nanex. 2010. *Flash Equity Failures in 2006, 2007, 2008, 2009, 2010, and 2011*.
Available at:
http://www.nanex.net/FlashCrashEquities/FlashCrashAnalysis_Equities.html

Nelson, D. B. 1990. ARCH models as diffusion approximations. *Journal of Econometrics* 45(1-2), pp. 7-38.

Nelson, D. B. 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pp. 347-370.

Newey, W. and West, K. 1987a. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3), pp. 703-708.

Newey, W. K. and West, K. D. 1987b. Hypothesis testing with efficient method of moments estimation. *International Economic Review*, pp. 777-787.

Niederhoffer, V. and Osborne, M. F. M. 1966. Market making and reversal on the stock exchange. *Journal of the American Statistical Association* 61(316), pp. 897-916.

Ogata, Y. 1981. On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory* 27(1), pp. 23-31.

Oomen, R.C., 2005. Properties of bias-corrected realized variance under alternative sampling schemes. *Journal of Financial Econometrics*, 3(4), pp.555-577.

Özbekler, A. G., Kontonikas, A. and Triantafyllou, A. 2021. Volatility forecasting in European government bond markets. *International Journal of Forecasting* 37(4), pp. 1691-1709.

Pan, J. 2002. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics* 63(1), pp. 3-50.

Park, S. and Linton, O. 2012. Realized variance: Theory and applications. In: *Handbook of Volatility Models and Their Applications*. Wiley, pp. 317-345.

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

Patton, A. J. and Sheppard, K. 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics* 97(3), pp. 683-697.

Phillips, P. C., Shi, S. and Yu, J. 2015. Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review* 56(4), pp. 1043-1078.

Phillips, P. C., Wu, Y. and Yu, J. 2011. Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review* 52(1), pp. 201-226.

Phillips, P. C. and Yu, J. 2011. Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics* 2(3), pp. 455-491.

Podolskij, M. and Vetter, M. 2009. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli* 15(3), pp. 634-658.

Pong, S., Shackleton, M. B., Taylor, S. J. and Xu, X. 2004. Forecasting currency volatility: A comparison of implied volatilities and AR (FI) MA models. *Journal of Banking & Finance* 28(10), pp. 2541-2563.

Porter, R. B. 1974. Semivariance and stochastic dominance: A comparison. *The American Economic Review*, pp. 200-204.

Pritsker, M. 2006. The hidden dangers of historical simulation. *Journal of Banking & Finance* 30(2), pp. 561-582.

Prokopczuk, M. and Simen, C. W. 2014. The importance of the volatility risk premium for volatility forecasting. *Journal of Banking & Finance* 40, pp. 303-320.

Pyun, S. 2019. Variance risk in aggregate stock returns and time-varying return predictability. *Journal of Financial Economics* 132(1), pp. 150-174.

Roll, R. 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance* 39(4), pp. 1127-1139.

Roy, A. D. 1952. Safety first and the holding of assets. *Econometrica: Journal of the Econometric Society*, pp. 431-449.

SEC. 2015. *Equity market volatility on august 24, 2015*. Available at: https://www.sec.gov/marketstructure/research/equity_market_volatility.pdf.

Sévi, B. 2014. Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research* 235(3), pp. 643-659.

Shi, S. and Phillips, P. C. 2024. Uncovering Mild Drift in Asset Prices with Intraday High-Frequency Data. *Available at SSRN 4786774*.

Shi, S. and Song, Y. 2016. Identifying speculative bubbles using an infinite hidden Markov model. *Journal of Financial Econometrics* 14(1), pp. 159-184.

Slim, S., Tabche, I., Koubaa, Y., Osman, M. and Karathanasopoulos, A. 2023. Forecasting realized variance of Bitcoin: The informative role of price duration. *Journal of Forecasting* 42(7), pp. 1909-1929.

Stoll, H. R. 2000. Presidential address: friction. *The Journal of Finance* 55(4), pp. 1479-1514.

Tauchen, G. and Zhou, H. 2011. Realized jumps on financial markets and predicting credit spreads. *Journal of Econometrics* 160(1), pp. 102-118.

Thomakos, D. D. and Wang, T. 2003. Realized variance in the futures markets. *Journal of Empirical Finance* 10(3), pp. 321-353.

Todorov, V. 2010. Variance risk-premium dynamics: The role of jumps. *The Review of Financial Studies* 23(1), pp. 345-383.

Todorova, N. 2017. The asymmetric volatility in the gold market revisited. *Economics Letters* 150, pp. 138-141.

Wang, X. and Yu, J. 2016. Double asymptotics for explosive continuous time models. *Journal of Econometrics* 193(1), pp. 35-53.

Wang, Y., Ma, F., Wei, Y. and Wu, C. 2016. Forecasting realized variance in a changing world: A dynamic model averaging approach. *Journal of Banking & Finance* 64, pp. 136-149.

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

Zhu, H., Bai, L., He, L. and Liu, Z. 2023. Forecasting realized variance with machine learning: Panel data perspective. *Journal of Empirical Finance* 73, pp. 251-271.

Appendix

A.1. Mathematical proofs for Chapter 2

Proof of Lemma 2.2.1.

Proof. (i) Under the constant drift-diffusion model, Equation (2.3.2), $r_{t_i} = \mu\Delta + \sigma\sqrt{\Delta}\varepsilon_{t_i}$ with $\varepsilon_{t_i} \sim N(0,1)$, the distribution of the return is normal, $r_{t_i} \sim N(\mu\Delta, \sigma^2\Delta)$. The probability that r_{t_i} is positive equals,

$$\mathbb{P}(r_{t_i} > 0) = 1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$$

where $\mathbb{P}(\cdot)$ denotes the probability of an argument and $\Phi(x)$ is the cumulative distribution function of the standard normal distribution.

The expected frequency of $r_{t_i} > 0$ over an observation window $t = 1 \ldots M$ equals the M multiplying the probability that a return is positive, denoted by,

$$M\mathbb{P}(r_{t_i} > 0) = 1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$$

$$= M\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right],$$

The expected frequency of $r_{t_i} < 0$ for the observation window $t = 1 \ldots M$ can be obtained by multiplying M with the chance of a negative return,

$$M\mathbb{P}(r_{t_i} < 0) = M\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right).$$

(ii) If drift is zero ($\mu = 0$), then $\Phi(-\mu\sqrt{\Delta}/\sigma) = 1/2$. Thus, the expected frequency of $r_{t_i} > 0$ and that of $r_{t_i} < 0$,

$$M\mathbb{P}(r_{t_i} > 0) = M\mathbb{P}(r_{t_i} < 0) = \frac{M}{2} \quad \blacksquare$$

Proof of Proposition 2.2.1.

Proof. (i) Give that the return r_{t_i} follows the content drift-diffusion model, Equation (2.3.2), $r_{t_i} = \mu\Delta + \sigma\sqrt{\Delta}\varepsilon_{t_i}$ with $\varepsilon_{t_i} \sim N(0,1)$ and $\Delta = 1/M$, we have $r_{t_i} \sim N(\mu\Delta, \sigma^2\Delta)$. Let $r_{t_i}|r_{t_i} > 0$ denotes the return conditional on its positive, then $r_{t_i}|r_{t_i} > 0$ follows a truncated normal distribution, with

$$\mathbb{E}(r_{t_i}|r_{t_i} > 0) = \mu\Delta + \sigma\sqrt{\Delta} \cdot \frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}$$

$$\mathbb{V}(r_{t_i}|r_{t_i} > 0) = \left(\sigma\sqrt{\Delta}\right)^2 \cdot \left[1 + \frac{\frac{-\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)} - \left(\frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}\right)^2\right]$$

where $\mathbb{E}(\cdot)$ and $\mathbb{V}(\cdot)$ denote the mean and variance of the argument, $\varphi(x)$ is the probability density function of the standard normal distribution

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right),$$

and $\Phi(x)$ is the respective cumulative distribution function.

Since the volatility σ_s equals a constant σ for $r_{t_i} \sim N(\mu\Delta, \sigma^2\Delta)$, we have the expected integrated volatility equals,

$$\mathbb{E}\left(\int_{t-1}^{t} \sigma_s^2 ds\right) = \int_{t-1}^{t} \mathbb{E}(\sigma_s^2) ds = M\sigma^2\Delta = \sigma^2.$$

The expected bias in the positive realized semivariance (RS_t^+) is

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= \mathbb{E}\left(\sum_{i=1}^{M} r_{t_i}^2 I_{r_{t_i}>0}\right) - \frac{1}{2}\sigma^2$$

Due to the linearity of expectation, we have

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= M \times \mathbb{E}\left(r_{t_i}^2 I_{r_{t_i}>0}\right) - \frac{1}{2}\sigma^2$$

$$= M \times \mathbb{P}(r_{t_i} > 0)\mathbb{E}\left(r_{t_i}^2 | r_{t_i} > 0\right) - \frac{1}{2}\sigma^2$$

$$= M \times \mathbb{P}\left(r_{t_i} > 0\right)\left[\mathbb{E}\left[r_{t_i} | r_{t_i} > 0\right]^2 + \mathbb{V}\left[r_{t_i} | r_{t_i} > 0\right]\right] - \frac{1}{2}\sigma^2$$

$\mathbb{P}\left(r_{t_i} > 0\right) = 1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$, according to *Lemma 2.3.1*

$$= M \times \left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right]\left\{\left[\mu\Delta + \sigma\sqrt{\Delta} \times \frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}\right]^2\right.$$

$$\left. + \left(\sigma\sqrt{\Delta}\right)^2 \times \left\{1 + \frac{\frac{-\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)} - \left[\frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}\right]^2\right\}\right\} - \frac{1}{2}\sigma^2$$

$$= M \times \left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right]\left\{\mu^2\Delta^2 + 2\mu\Delta\sigma\sqrt{\Delta} \cdot \frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)} + \left(\sigma\sqrt{\Delta}\right)^2\right.$$

$$\left. + \left(\sigma\sqrt{\Delta}\right)^2 \frac{\frac{-\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}\right\} - \frac{1}{2}\sigma^2$$

$$= \mu^2\Delta\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \sigma^2\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] - \frac{1}{2}\sigma^2$$

(ii) The bias ratio of the positive realized semivariance (RS_t^+) equals,

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right) / \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$= \left\{\mu^2\Delta\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \sigma^2\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] - \frac{1}{2}\sigma^2\right\} / \frac{1}{2}\sigma^2$$

$$= \frac{2\mu^2\Delta}{\sigma^2}\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] - 1$$

$$= \frac{2\mu^2\Delta}{\sigma^2}\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 1.$$

(iii) Let $r_{t_i}|r_{t_i} < 0$ denotes the return conditional on its negative. Given that $r_{t_i} \sim N(\mu\Delta, \sigma^2\Delta)$, then $r_{t_i}|r_{t_i} < 0$ follows a truncated normal distribution with the mean and variance defined by,

$$\mathbb{E}\left(r_{t_i}|r_{t_i} < 0\right) = \mu\Delta - \sigma\sqrt{\Delta} \cdot \frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}$$

$$\mathbb{V}\left(r_{t_i}|r_{t_i}<0\right)=\left(\sigma\sqrt{\Delta}\right)^2\cdot\left[1+\frac{\frac{\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}-\left(\frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}\right)^2\right].$$

The expected bias in the negative realized semivariance (RS_t^-) is

$$\mathbb{E}\left(RS_t^- -\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$=\mathbb{E}\left(\sum_{i=1}^{M}r_{t_i}^2 I_{r_{t_i}<0}\right)-\frac{1}{2}\sigma^2$$

Due to the linearity of expectation, we have

$$\mathbb{E}\left(RS_t^- -\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$=M\mathbb{E}\left(r_{t_i}^2 I_{r_{t_i}<0}\right)-\frac{1}{2}\sigma^2$$

$$=M\mathbb{P}(r_{t_i}<0)\mathbb{E}\left(r_{t_i}^2|r_{t_i}<0\right)-\frac{1}{2}\sigma^2$$

$$=M\mathbb{P}\left(r_{t_i}<0\right)\left[\mathbb{E}\left[r_{t_i}|r_{t_i}<0\right]^2+\mathbb{V}\left[r_{t_i}|r_{t_i}<0\right]\right]-\frac{1}{2}\sigma^2$$

$\mathbb{P}\left(r_{t_i}<0\right)=\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$, according to *Lemma 2.3.1*

$$=M\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\left\{\left[\mu\Delta-\sigma\sqrt{\Delta}\frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}\right]^2+\left(\sigma\sqrt{\Delta}\right)^2\left\{1+\frac{\frac{\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}-\left[\frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}\right]^2\right\}\right\}-\frac{1}{2}\sigma^2$$

$$= M\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\left[\mu^2\Delta^2 - 2\mu\Delta\sigma\sqrt{\Delta}\frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)} + \left(\sigma\sqrt{\Delta}\right)^2 + \frac{\mu\left(\sigma\sqrt{\Delta}\right)^2\sqrt{\Delta}}{\sigma} \times \frac{\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}{\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)}\right]$$

$$-\frac{1}{2}\sigma^2$$

$$= \left[\mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\sigma^2 + \frac{\mu\sigma^2\sqrt{\Delta}}{\sigma} \times \varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] - \frac{1}{2}\sigma^2$$

$$= \mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\sigma^2 - \frac{1}{2}\sigma^2$$

(vi) The bias ratio of the negative realized semivariance (RS_t^-) equals,

$$\mathbb{E}\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right) / \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$= \left\{\mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\sigma^2 - \frac{1}{2}\sigma^2\right\} / \frac{1}{2}\sigma^2$$

$$= \frac{2\mu^2\Delta}{\sigma^2}\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 1 \quad \blacksquare$$

Proof of Corollary 2.2.1.

Proof. According to Proposition 2.1.1, when drift is zero $\mu = 0$, the biases of both positive and negative realized semivariance are zero.

For $\mu = 0$, the bias in RS_t^+ equals,

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= 0^2\Delta\left[1-\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right]+0\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)+\sigma^2\left[1-\Phi\left(\frac{-0\sqrt{\Delta}}{\sigma}\right)\right]-\frac{1}{2}\sigma^2$$

$$= \sigma^2\left[1-\frac{1}{2}\right]-\frac{1}{2}\sigma^2 = 0,$$

the bias ratio of RS_t^+ equals,

$$\mathbb{E}\left(RS_t^+-\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$= 0/\frac{1}{2}\sigma^2 = 0,$$

the bias in RS_t^- equals,

$$\mathbb{E}\left(RS_t^--\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= 0^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)-0\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)+\Phi\left(\frac{-0\sqrt{\Delta}}{\sigma}\right)\sigma^2-\frac{1}{2}\sigma^2$$

$$= \frac{1}{2}\sigma^2-\frac{1}{2}\sigma^2 = 0,$$

and the bias ratio of RS_t^- equals,

$$\mathbb{E}\left(RS_t^--\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$= 0/\frac{1}{2}\sigma^2 = 0 \ \blacksquare$$

Proof of Corollary 2.2.2.

Proof. According to Proposition 2.1 of Laurent and Shi (2020), the bias ratio of realized variance equals

$$\mathbb{E}\left(RV_t - \int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \int_{t-1}^{t} \sigma_s^2 ds = \frac{\mu^2 \Delta}{\sigma^2}.$$

For any $\mu > 0$, the bias ratio of RS_t^+ is,

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds$$

$$= \frac{2\mu^2 \Delta}{\sigma^2}\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 1$$

Because $\varphi(x) > 0$ (all output values the probability density function of the standard normal distribution may produce are positive), we have

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds$$

$$> \frac{2\mu^2 \Delta}{\sigma^2}\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] - 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 1$$

Given that $-\Phi(x) > -1/2$ for $x < 0$, we have

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds$$

$$> \frac{2\mu^2 \Delta}{\sigma^2}\left[1 - \frac{1}{2}\right] - 2 \times \frac{1}{2} + 1$$

$$= \frac{\mu^2 \Delta}{\sigma^2} = \mathbb{E}\left(RV_t - \int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \int_{t-1}^{t} \sigma_s^2 ds$$

For any $\mu < 0$, the bias ratio of RS_t^+ is,

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right) \Big/ \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds$$

$$= \frac{2\mu^2\Delta}{\sigma^2}\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 1$$

$$< \frac{2\mu^2\Delta}{\sigma^2}\left[1 - \frac{1}{2}\right] + \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2 \times \frac{1}{2} + 1$$

$$< \frac{2\mu^2\Delta}{\sigma^2}\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$$

Given that $-\Phi(x) < -1/2$ and $\varphi(x) > 0$ for $x > 0$, we have

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$< \frac{2\mu^2\Delta}{\sigma^2}\left[1 - \frac{1}{2}\right] + \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$$

$$< \frac{2\mu^2\Delta}{\sigma^2}\left[1 - \frac{1}{2}\right] + 0$$

$$= \frac{\mu^2\Delta}{\sigma^2} = \mathbb{E}\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\int_{t-1}^{t}\sigma_s^2 ds\,. \blacksquare$$

Proof of Corollary 2.2.3.

For any $\mu > 0$, the bias ratio of RS_t^- is,

$$\mathbb{E}\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$= \frac{2\mu^2\Delta}{\sigma^2}\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 1$$

Given that $\Phi(x) < 1/2$ and $\varphi(x) > 0$ for $x < 0$,

$$\mathbb{E}\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$< \frac{2\mu^2\Delta}{\sigma^2}\times\frac{1}{2} + 0 + 2\times\frac{1}{2} - 1$$

$$= \frac{\mu^2\Delta}{\sigma^2} = \mathbb{E}\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\int_{t-1}^{t}\sigma_s^2 ds.$$

For any $\mu < 0$, the bias ratio of RS_t^- is,

$$\mathbb{E}\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$= \frac{2\mu^2\Delta}{\sigma^2}\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \frac{2\mu\sqrt{\Delta}}{\sigma}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 1$$

Given that $\Phi(x) > 1/2$ and $\varphi(x) > 0$ for $x > 0$,

$$\mathbb{E}\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$> \frac{2\mu^2\Delta}{\sigma^2}\times\frac{1}{2} + 0 + 2\times\frac{1}{2} - 1$$

$$= \frac{\mu^2\Delta}{\sigma^2} = \mathbb{E}\left(RV_t - \int_{t-1}^{t}\sigma_s^2 ds\right)\Big/\int_{t-1}^{t}\sigma_s^2 ds\,.\ \blacksquare$$

Proof of Corollary 2.2.4.

Proof. According to Proposition 2.3.1, under the drift-diffusion process, Equation (2.3.2) with $\mu = \mu^*$, the bias in RS_t^+ is,

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= (\mu^*)^2\Delta\left[1-\Phi\left(\frac{-\mu^*\sqrt{\Delta}}{\sigma}\right)\right] + \mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu^*\sqrt{\Delta}}{\sigma}\right) + \sigma^2\left[1-\Phi\left(\frac{-\mu^*\sqrt{\Delta}}{\sigma}\right)\right] - \frac{1}{2}\sigma^2$$

Since $1-\Phi(-x) = \Phi(x)$, then

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= (\mu^*)^2\Delta\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + \mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu^*\sqrt{\Delta}}{\sigma}\right) + \sigma^2\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) - \frac{1}{2}\sigma^2$$

Given that $\varphi(-x) = \varphi(x)$, we have

$$\mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= (\mu^*)^2\Delta\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + \mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + \sigma^2\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) - \frac{1}{2}\sigma^2$$

Based on Proposition 2.3.1, for the drift-diffusion process, Equation (2.3.2), with $\mu = -\mu^*$, the bias in RS_t^- is,

$$E\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)$$

$$= (\mu^*)^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + \Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right)\sigma^2 - \frac{1}{2}\sigma^2,$$

which is the bias in RS_t^+ with $\mu = \mu^*$. This symmetric property holds for the bias ratio, and the proofs are trivial ■

Proof of Lemma 2.3.2.

With an even number of intraday returns M, the median is taken halfway between the middle two values, exactly half of M returns are negative and half positive. ■

Proof of Proposition 2.3.2.

By exchangeability of return r_{t_i} for $i = 1,2,\dots,M$, the centred returns $r_{t_i} - \widehat{m}_t$ for $i = 1,2,\dots,M$ are identically distributed to each other. Comparing the joint distribution of $\left(r_{t_1} - \widehat{m}_t, r_{t_2} - \widehat{m}_t, \dots, r_{t_M} - \widehat{m}_t\right)$ with that of $\left(-r_{t_1} + \widehat{m}_t, -r_{t_2} + \widehat{m}_t, \dots, -r_{t_M} + \widehat{m}_t\right)$ and using the symmetry of normal distributions, the centred returns $r_{t_i} - \widehat{m}_t$ are each symmetrically distributed about zero.

The bias of the modified positive realized semivariance $\text{bias}(RS^+)$ is

$$\text{bias}(RS^+) = E\left\{\sum_{i=1}^{M}\left[\left(r_{t_i} - \widehat{m}_t\right)I\left(r_{t_i} - \widehat{m}_t > 0\right)\right]^2 - \frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right\}\Big/\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds$$

$$= \left\{E\left\{\sum_{i=1}^{M}\left[\left(r_{t_i} - \widehat{m}_t\right)I\left(r_{t_i} - \widehat{m}_t > 0\right)\right]^2\right\} - \frac{M}{2}\sigma^2\Delta\right\}\Big/\frac{M}{2}\sigma^2\Delta.$$

From Lemma 2.3.2, the expected number of positive returns equals $M/2$, and $r_{t_i} - \widehat{m}_t$ are identically distributed to each other. Therefore, the bias of RS^+,

$$\text{bias}(RS^+) = \frac{M}{2}E\left\{\left[\left(r_{t_i} - \widehat{m}_t\right)I\left(r_{t_i} - \widehat{m}_t > 0\right)\right]^2\right\}\Big/\frac{M}{2}\sigma^2\Delta$$

$$- 1 = E\left\{\left[\left(r_{t_i} - \widehat{m}_t\right)I\left(r_{t_i} - \widehat{m}_t > 0\right)\right]^2\right\}/\sigma^2\Delta - 1.$$

Since $r_{t_i} - \widehat{m}_t$ are each symmetrically distributed about 0, we have

$$\text{bias}(RS^+) = E\left\{\left[\left(r_{t_i} - \widehat{m}_t\right) I\left(r_{t_i} - \widehat{m}_t > 0\right)\right]^2\right\}/\sigma^2\Delta - 1$$

$$= E\left\{\left[\left(r_{t_i} - \widehat{m}_t\right) I\left(r_{t_i} - \widehat{m}_t < 0\right)\right]^2\right\}/\sigma^2\Delta - 1 = E\left\{\left(r_{t_i} - \widehat{m}_t\right)^2\right\}/\sigma^2\Delta - 1.$$

Hence, $\text{bias}(RS^+) = \text{bias}(RS^-) = \text{bias}(RV)$. ■

Proof of Proposition 2.5.1.

Proof.

According to Lemma 2.5.1, in the absence of jumps, the bias in the signed jump estimator J_t^{Δ} equals,

$$J_t^{\Delta} - 0$$

$$= \mathbb{E}\left(RS_t^+ - \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right) - \mathbb{E}\left(RS_t^- - \frac{1}{2}\int_{t-1}^{t} \sigma_s^2 ds\right).$$

By substituting the results of the biases in RS_t^+ and RS_t^- under the constant drift-diffusion model (Proposition 2.3.1), we have

$$J_t^{\Delta} - 0$$

$$= \mu^2\Delta\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] + \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \sigma^2\left[1 - \Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)\right] - \frac{1}{2}\sigma^2$$

$$- \left[\mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \sigma^2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \frac{1}{2}\sigma^2\right]$$

$$= \mu^2\Delta - \mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \sigma^2 - \sigma^2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \frac{1}{2}\sigma^2$$

$$- \mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - \sigma^2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + \frac{1}{2}\sigma^2$$

$$= \mu^2\Delta + \sigma^2 - 2\mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right).$$

If $\mu > 0$,

$J_t^{\Delta} - 0$

$$= \mu^2\Delta + \sigma^2 - 2\mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$$

Given that $-\Phi(-x) > -1/2$ for $x > 0$, we have

$J_t^{\Delta} - 0$

$$> \mu^2\Delta + \sigma^2 - 2\mu^2\Delta \times \frac{1}{2} + 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2 \times \frac{1}{2}$$

$$= 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) > 0.$$

If $\mu < 0$,

$J_t^{\Delta} - 0$

$$= \mu^2\Delta + \sigma^2 - 2\mu^2\Delta\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) + 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2\Phi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right)$$

Given that $-\Phi(-x) < -1/2$ for $x < 0$, we have

$$J_t^{\Delta} - 0$$

$$< \mu^2\Delta + \sigma^2 - 2\mu^2\Delta \times \frac{1}{2} + 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2 \times \frac{1}{2}$$

$$= 2\mu\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu\sqrt{\Delta}}{\sigma}\right) < 0.$$

Based on the above results of the nonzero bias in J_t^{Δ} in the presence of a nonzero μ. We have for $\mu > 0$ and any $\sigma > 0$, the bias ratio of J_t^{Δ} equals $\left(J_t^{\Delta} - 0\right)/0 = +\infty$, and when $\mu < 0$ and any $\sigma > 0$, the bias ratio of J_t^{Δ} equals $\left(J_t^{\Delta} - 0\right)/0 = -\infty$. ■

Proof of Corollary 2.5.1.

Proof.

If $\mu = \mu^*$

$$J_t^{\Delta} - 0$$

$$= (\mu^*)^2\Delta + \sigma^2 - 2(\mu^*)^2\Delta\Phi\left(\frac{-\mu^*\sqrt{\Delta}}{\sigma}\right) + 2\mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{-\mu^*\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2\Phi\left(\frac{-\mu^*\sqrt{\Delta}}{\sigma}\right)$$

$$= (\mu^*)^2\Delta + \sigma^2 - 2(\mu^*)^2\Delta\left[1 - \Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right)\right] + 2\mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2\left[1 - \Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right)\right]$$

$$= (\mu^*)^2\Delta + \sigma^2 - 2(\mu^*)^2\Delta + 2(\mu^*)^2\Delta\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + 2\mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2$$

$$+ 2\sigma^2\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right)$$

$$= -(\mu^*)^2\Delta - \sigma^2 + 2(\mu^*)^2\Delta\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + 2\mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + 2\sigma^2\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right)$$

If $\mu = -\mu^*$

$$J_t^{\Delta} - 0$$

$$= (-\mu^*)^2\Delta + \sigma^2 - 2(-\mu^*)^2\Delta\Phi\left(\frac{-(-\mu^*)\sqrt{\Delta}}{\sigma}\right) + 2(-\mu^*)\sigma\sqrt{\Delta}\varphi\left(\frac{-(-\mu^*)\sqrt{\Delta}}{\sigma}\right)$$
$$- 2\sigma^2\Phi\left(\frac{-(-\mu^*)\sqrt{\Delta}}{\sigma}\right)$$

$$= (\mu^*)^2\Delta + \sigma^2 - 2(\mu^*)^2\Delta\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) - 2\mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) - 2\sigma^2\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right)$$

$$= -\left[-(\mu^*)^2\Delta - \sigma^2 + 2(\mu^*)^2\Delta\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + 2\mu^*\sigma\sqrt{\Delta}\varphi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right) + 2\sigma^2\Phi\left(\frac{\mu^*\sqrt{\Delta}}{\sigma}\right)\right],$$

which equals the negative sign of the bias in J_t^{Δ} when $\mu = \mu^*$. ■

Proof of Proposition 2.3.2.

Proof.

The bias of the modified positive realized semivariance $\text{bias}(RS^+)$ is

$$J_t^{\Delta\,*} - 0$$

$$= \left[RS_t^{+\,*} - \left(\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\right] - \left[RS_t^{-\,*} - \left(\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\right]$$

From Proposition 2.3.2, the bias in $RS^{+\,*}$ equals the bias in $RS^{-\,*}$, therefore,

$$J_t^{\Delta\,*} - 0$$

$$= \left[RS_t^{+\,*} - \left(\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\right] - \left[RS_t^{+\,*} - \left(\frac{1}{2}\int_{t-1}^{t}\sigma_s^2 ds\right)\right] = 0 \ \blacksquare$$

A. 2. Weighted Least Square Illustration for Chapter 2

This section explains the motivation and calculation of using the Weighted Least Square (WLS) estimation method in this chapter. For the illustration, I consider the estimation of the J^{Δ} model as an example. The traditional method for estimating the J^{Δ} model relies on the Ordinary Least Square (OLS) method (e.g., Andersen et al. 2007b; Corsi 2009). To implement this method, for example, for the J^{Δ} model, first substitute the realized measures from the sample into this model,

$$RV_{23} = \beta_0 + \beta_{j^{\Delta}} J_{22}^{\Delta} + \beta_d \, RV_{22} + \beta_w \, \overline{RV}_{w,22} + \beta_m \, \overline{RV}_{m,22} + \epsilon_{23},$$

$$RV_{24} = \beta_0 + \beta_{j^{\Delta}} J_{23}^{\Delta} + \beta_d \, RV_{23} + \beta_w \, \overline{RV}_{w,23} + \beta_m \, \overline{RV}_{m,23} + \epsilon_{24},$$

$$\vdots$$

$$RV_n = \beta_0 + \beta_{j^{\Delta}} J_{n-1}^{\Delta} + \beta_d \, RV_{n-1} + \beta_w \, \overline{RV}_{w,n-1} + \beta_m \, \overline{RV}_{m,n-1} + \epsilon_n.$$

Note that the dependent variable starts from RV_{23} as the calculation of the respective independent variable $\overline{RV}_{m,t}$ requires a minimum of 21 observations. The matrix presentation of the above equations is,

$$\begin{bmatrix} RV_{23} \\ RV_{24} \\ \vdots \\ \vdots \\ RV_n \end{bmatrix} = \begin{bmatrix} 1 & J_{22}^{\Delta} & RV_{22} & \overline{RV}_{w,22} & \overline{RV}_{m,22} \\ 1 & J_{23}^{\Delta} & RV_{23} & \overline{RV}_{w,23} & \overline{RV}_{m,23} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & J_{n-1}^{\Delta} & RV_{n-1} & \overline{RV}_{w,n-1} & \overline{RV}_{m,n-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_{\mathrm{d}} \\ \beta_{\mathrm{w}} \\ \beta_{\mathrm{m}} \end{bmatrix} + \begin{bmatrix} \epsilon_{23} \\ \epsilon_{24} \\ \vdots \\ \vdots \\ \epsilon_n \end{bmatrix}. \qquad \text{(A. 2.1)}$$

This can be rewritten more simply as:

$$y = X\beta + \epsilon, \qquad \text{(A. 2.2)}$$

where $y = \begin{bmatrix} RV_{23} \\ RV_{24} \\ \vdots \\ \vdots \\ RV_n \end{bmatrix}$, $X = \begin{bmatrix} 1 & J_{22}^{\Delta} & RV_{22} & \overline{RV}_{w,22} & \overline{RV}_{m,22} \\ 1 & J_{23}^{\Delta} & RV_{23} & \overline{RV}_{w,23} & \overline{RV}_{m,23} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & J_{n-1}^{\Delta} & RV_{n-1} & \overline{RV}_{w,n-1} & \overline{RV}_{m,n-1} \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_{\mathrm{d}} \\ \beta_{\mathrm{w}} \\ \beta_{\mathrm{m}} \end{bmatrix}$, and $\epsilon = \begin{bmatrix} \epsilon_{23} \\ \epsilon_{24} \\ \vdots \\ \vdots \\ \epsilon_n \end{bmatrix}$. The OLS estimator of β ($\hat{\beta}_{OLS}$) is the solution to minimize the sum of squared residuals defined by,

$$(y - X\beta)'(y - X\beta),$$

and the solution of this minimization problem is

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'y. \tag{A. 2.3}$$

The residuals based on the above OLS solution of coefficients (e_{OLS}) are defined by,

$$e_{OLS} = y - X\hat{\beta}_{OLS}, \tag{A. 2.4}$$

The fitted value (or predicted value) of y is defined by,

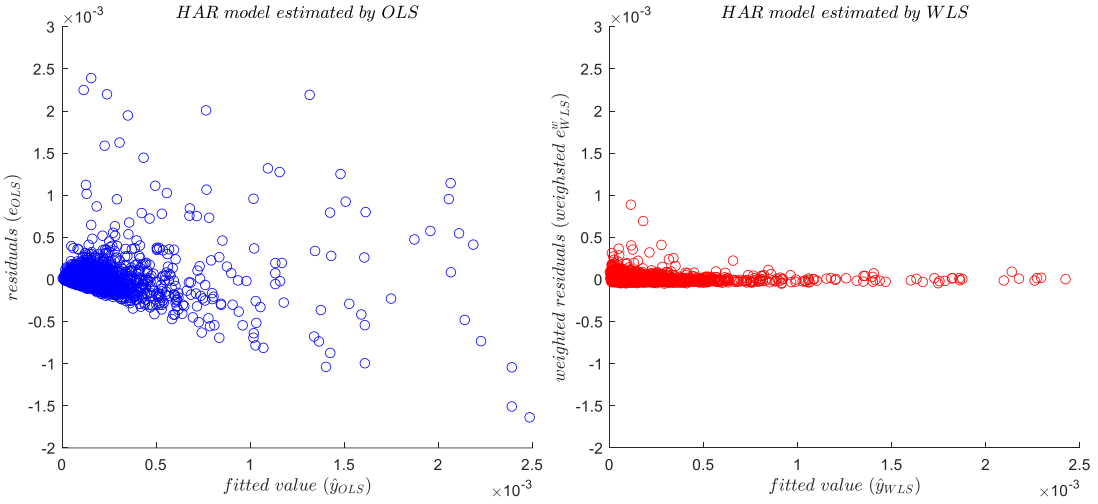
$$\hat{y}_{OLS} = \begin{bmatrix} \widehat{RV}_{23}^{OLS} \\ \widehat{RV}_{24}^{OLS} \\ \vdots \\ \vdots \\ \widehat{RV}_{n}^{OLS} \end{bmatrix} = X\hat{\beta}_{OLS}. \tag{A. 2.5}$$

One important (Gauss-Markov) assumption of Ordinary Least Square (OLS) is that the conditional variance of the disturbances should be fixed across the sample period (or homoskedasticity). For example, the OLS for the J^{Δ} model assumes the conditional variance of the disturbances of the J^{Δ} model to be fixed across the sample period, that is, $\mathrm{E}(\epsilon\epsilon'|X) = \sigma^2 I$, where σ^2 is the level of conditional variance and is fixed, and I indicates the identity matrix of size $n - 22$. However, as documented in the literature (e.g., Corsi et al. 2008; Patton and Sheppard 2015), the assumption of

homoskedasticity may be not realistic as the variance of the disturbances of the HAR model may vary substantially across the sample period (or heteroskedasticity) and is correlated its lagged value (or autocorrelation). For example, the disturbances of the HAR model around the 2008 financial crisis are generally greater than other tranquil periods (due to volatility changes much more dramatically during this period).

To check if heteroskedasticity is present, the left panel of Figure A.2.1 reports the scatter plot of residuals (e_{OLS}) and fitted value ($\hat{y}_{OLS}$) for the J^{Δ} model estimated by the OLS method for the SPY sample of this chapter.[7] As the panel shows, the spread of the residuals is increasing as the fitted values changes. This implies that the level of the conditional volatility of the disturbances is likely to change dramatically across the sample period and increases with the level of the fitted values.

Figure A.2.1. The scatter plots for residuals and fitted value for the J^{Δ} model



Notes: The left panel reports the scatter plot of residuals (e_{OLS}) and fitted value ($\hat{y}_{OLS}$) for the J^{Δ} model estimated by the Ordinary Least Square (OLS) method. The residuals (e_{OLS}) and fitted value ($\hat{y}_{OLS}$) are obtained by Equation (A.2.4) and Equation (A.2.5), respectively. The right panel reports the scatter plot of weighted residuals (e^{w}_{WLS}) and fitted value ($\hat{y}_{WLS}$) for the J^{Δ} model estimated by the Weighted Least Square (WLS) method. The weighted residuals (e^{w}_{WLS}) and fitted value ($\hat{y}_{WLS}$) are obtained by Equation (A.2.9) and Equation (A.2.10), respectively. The data used for computing these residuals and fitted values is SPDR S&P 500 ETF from 1997 to 2021.

[7] Here, I only report the results for the J^{Δ} model in Figure A.1. But I confirm that the results for the remaining models lead to similar patterns.

Since the level of the conditional volatility of the disturbances possibly changes across the sample period and is consistent with the level of the fitted values, estimation by OLS has the disadvantage that the resulting estimates focus primarily on fitting periods of high RV and place little weight on low RV periods. To overcome this, Weighted Least Squares (WLS) may be a solution. WLS attempts to provide a more efficient alternative to OLS by putting less weight on disturbances which are likely to have a large variance and more weight on disturbances which are likely to have a small variance. If each weight w is inversely proportional to the conditional variance of the corresponding disturbances, then the WLS estimator is more efficient than the OLS estimator.

Motivated by the positive relationship between the volatility of the residuals and the level of OLS fitted values, Patton and Sheppard (2015) suggest using the inverse of the OLS fitted values as the weights. Specifically, for the J^A model, first obtain the OLS fitted value $\hat{y}_{OLS}$ from Equation (A.5). Then calculate the weights by the inverse value of each element in $\hat{y}_{OLS}$. The matrix form of the weights is defined by the matrix with the weights on the diagonal and zeroes everywhere else,

$$W = \begin{bmatrix} 1/\widehat{RV}_{23}^{OLS} & 0 & \cdots & 0 \\ 0 & 1/\widehat{RV}_{24}^{OLS} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\widehat{RV}_{n}^{OLS} \end{bmatrix}. \quad \text{(A. 2.6)}$$

Finally, based on the weights W, the WLS estimator of β ($\hat{\beta}_{WLS}$) is the solution to the minimization problem of the weighted sum of squared residuals,

$$(y - X\beta)'W(y - X\beta),$$

and the solution of this minimization problem is

$$\hat{\beta}_{WLS} = (X'WX)^{-1}X'Wy. \tag{A. 2.7}$$

The residuals (e^{w}_{WLS}) based on the above WLS estimated coefficients $\hat{\beta}_{WLS}$ are defined by,

$$e_{WLS} = \alpha W\left(y - X\hat{\beta}_{WLS}\right), \tag{A. 2.8}$$

and the weighted residuals (e^{w}_{WLS}) are defined by,

$$e^{w}_{WLS} = \alpha W\left(y - X\hat{\beta}_{WLS}\right), \tag{A. 2.9}$$

where the component $\alpha = 1/\left[1/(n-22)\sum_{j=1}^{n-22} w_j\right]$ (with w_j indicating the jth diagonal elements of the W matrix) is a scale which ensures that the average weight equals one ($1/(n-22)\sum_{j=1}^{n-22}\{\alpha w_j\} = 1$), and this facilitates comparing the WLS-weighted residuals e^{w}_{WLS} with the above OLS (unweighted) residuals e_{OLS}. The fitted value (or predicted value) of y is defined by,

$$\hat{y}_{WLS} = \begin{bmatrix} \widehat{RV}_{23}^{WLS} \\ \widehat{RV}_{24}^{WLS} \\ \vdots \\ \vdots \\ \widehat{RV}_{n}^{WLS} \end{bmatrix} = X\hat{\beta}_{WLS}. \tag{A. 2.10}$$

To check if heteroskedasticity is reduced for weighted WLS residuals e^{w}_{WLS}, the right panel of Figure A.1 reports the scatter plot of weighted WLS residuals (e^{w}_{WLS}) and fitted value ($\hat{y}_{WLS}$) for the HAR model estimated by the WLS method. [8] Comparing the left panel with the right panel, the deviation of the weighted WLS residuals (e^{w}_{WLS}) is much smaller than that of the OLS residuals (e_{OLS}) as the fitted values changes. This implies that the heteroskedasticity of the weighted disturbances is reduced across the sample period. Motivated by the superiority of such WLS

[8] Here, I only report the results for the J^{A} model in Figure A.1. But I confirm that the results for the remaining models in this chapter lead to similar patterns.

estimation method in terms of treating the heteroskedasticity of disturbances, I use apply this WLS estimation method for estimating all models throughout this chapter.

The population of coefficients β is unknown and the estimate of these coefficients (e.g., $\hat{\beta}_{WLS}$) is used for testing certain hypotheses about the population of coefficients β.

To explore the impact of the predictors on the future realized variance is weak or strong, this chapter tests the null hypothesis that the population of coefficient β is zero and the alternative hypothesis that the population of coefficient β is not zero. To test these hypotheses, the distribution (type of probability distribution, mean, variance) of $\hat{\beta}_{WLS}$ must be known. To obtain the type of probability distribution of $\hat{\beta}_{WLS}$, first, obtain a formulation of $\hat{\beta}_{WLS}$ by Equation (A.2.7) and Equation (A.2.2),

$$\hat{\beta}_{WLS} = \beta + (X'WX)^{-1}X'W\epsilon. \tag{A. 2.11}$$

From Equation (A.2.11), it can then be seen that the type of probability distribution of $\hat{\beta}_{WLS}$ is consistent with that of disturbances ϵ. As often, the type of probability distribution of disturbances ϵ is assumed to be (multivariate) normal. Therefore, the type of probability distribution of the WLS estimated coefficients $\hat{\beta}_{WLS}$ is also a multivariate normal distribution. From Equation (A.2.11), it can also be easily calculated that the mean of the distribution of $\hat{\beta}_{WLS}$ ($E(\hat{\beta}_{WLS})$) equals β,

$$E(\hat{\beta}_{WLS}) = \beta \tag{A. 2.12}$$

(Note that $E((X'WX)^{-1}X'W\epsilon) = (X'WX)^{-1}X'WE(\epsilon) = 0$ as one of the Gauss-Markov assumptions states that the disturbances average out to 0 for any values of X). From Equation (A.2.11), the variance-covariance of the coefficients $\hat{\beta}_{WLS}$ is

$E\left(\left(\hat{\beta}_{WLS}-\beta\right)\left(\hat{\beta}_{WLS}-\beta\right)'\right)$ and can be obtained by substituting $\hat{\beta}_{WLS}$ with $\beta+(X'WX)^{-1}X'W\epsilon$,

$$E\left(\left(\hat{\beta}_{WLS}-\beta\right)\left(\hat{\beta}_{WLS}-\beta\right)'\right)=(X'WX)^{-1}X'WE(\epsilon\epsilon')W'X(X'WX)^{-1}. \quad \text{(A. 2.13)}$$

The proportion $X'WE(\epsilon\epsilon')W'X$ in the middle of the right-hand side of the above Equation (A.2.13) need to be estimated (Equation (A.2.13) is also known as the 'sandwich' estimator with two $(X'WX)^{-1}$ as upper and lower breads and $X'WE(\epsilon\epsilon')W'X$ as meat in the middle). To estimate the 'meat' proportion $X'WE(\epsilon\epsilon')W'X$, I use the heteroskedasticity and autocorrelation consistent (HAC) estimator by Newey and West (1987a). The HAC estimator is robust to estimation bias due to the heteroskedasticity and autocorrelation of the variance of disturbances. Specifically, the HAC estimator for $X'WE(\epsilon\epsilon')W'X$ is given by,

$$\begin{aligned}X'WE(\epsilon\epsilon')W' = &\frac{n}{n-k}\sum_{j=1}^{n-22}\left(w_j e_{WLS,j}\right)^2 x_j' x_j \\ &+\frac{n}{n-k}\sum_{l=1}^{L}\left(1-\frac{l}{L+1}\right)\sum_{j=l+1}^{n-22} w_j w_{j-l} e_{WLS,j} e_{WLS,j-l}\left(x_j' x_{j-l}\right. \\ &\left.+x_{j-l}' x_j\right)\end{aligned} \quad \text{(A. 2.14)}$$

where x_j is the jth row of the X matrix, $e_{WLS,j}$ is the jth row of the e_{WLS} vector, w_j is the jth diagonal elements of the W matrix, and k is the number of predictors (including the constant) in the model (e.g., $k=4$ for the J^{Δ} model), and L is the number of lags for compensating the estimation bias risen by the autocorrelation. Following Corsi and Renò (2012), I set the number of lags L equal to $2(h+1)$, where h is the lead length (or the forecast horizon) of the dependent variable of the model (e.g., for the J^{Δ} model, $L=4$ as $h=1$). By substituting Equation (A.2.14) to

Equation (A.2.13), the HAC variance-covariance matrix for $\hat{\beta}_{WLS}$ can be obtained. I use notation $\widehat{VCV}_{HAC}$ to indicate this HAC variance-covariance matrix.

In summary, the results of the distribution of information on $\hat{\beta}_{WLS}$ show that $\hat{\beta}_{WLS}$ is distributed multivariate normal with mean equal to β and variance-covariance matrix equal to $\widehat{VCV}^{\text{HAC}}$. Then, the test statistics T_{stat} for testing the hypotheses about the coefficients β are calculated by $T_{stat} = \hat{\beta}_{WLS} \oslash \text{diag}\left(\widehat{VCV}_{HAC}\right)$, where "$\oslash$" indicates the Hadamard division (Hadamard division is a binary operation that takes in two matrices of the same dimensions and returns a matrix of the divided corresponding elements), "diag" denotes the Matrix-to-vector diag operator (e.g., $\text{diag}\left(\widehat{VCV}^{\text{HAC}}\right)$ returns a vector of the diagonal entries of $\widehat{VCV}^{\text{HAC}}$). For the HAR model, the population of the coefficients β and T_{stat} can be presented by,

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_{\text{d}} \\ \beta_{\text{w}} \\ \beta_{\text{m}} \end{bmatrix} \text{ and } T_{stat} = \begin{bmatrix} T_{stat,0} \\ T_{stat,d} \\ T_{stat,w} \\ T_{stat,m} \end{bmatrix}.$$

Based on this formulation of β and T_{stat}, the hypotheses for each of the coefficients in the matrix β can be tested based on the corresponding test statistic in the matrix T_{stat}. For example, one can test the null & alternative hypotheses of β_{d} by the value of $T_{stat,d}$. If $T_{stat,d}$ is larger than 1.96, the false positive (or the probability of mistaken rejection of a null hypothesis) is smaller than 0.05 (two-tail). The probability of a false positive is based on the p-value that corresponds to $t = 1.96$ with the degree of freedom $\text{df} = n - 22 - k = 6222 - 22 - 4 = 6196$ for the student t-distribution is approximately 0.05 (As often, to account for small sample bias, standard normal distribution is often replaced with the student t-distribution). As the risk of false positive is small, one can reject the null hypothesis that the population

of β_{d} is zero thus accepting the alternative hypothesis that the population of β_{d} is different from zero.

The goodness of fit of a statistical model describes how well it fits a set of observations. To measure the goodness of fit, this chapter uses the coefficient of determination, denoted R^2. R^2 is defined by,

$$R^2 = 1 - \frac{\sum_{j=1}^{n-22}\left(y_j - \hat{y}_{WLS,j}\right)^2}{\sum_{j=1}^{n-22}\left(y_j - \frac{1}{n-22}\sum_{j=1}^{n-22} y_j\right)^2} \qquad \text{(A. 2.15)}$$

where y_j is the jth row of the y matrix and $\hat{y}_{WLS,j}$ is the jth row of the $\hat{y}_{WLS}$ matrix.

A. 3. More empirical results for Chapter 2

Table A.3.1. In-sample estimation results of the volatility models for weekly forecast

Notes: The table provides in-sample parameter estimates and measures of fit for the volatility models with volatility and signed jump estimators for weekly forecast. The brackets report the Heteroskedasticity and Autocorrelation Consistent (HAC) t-statistics. HAR^* is a specification which uses the realized variance, Equation (2.7.3), RS^*, Equation (2.7.4), is based on modified semivariances, BV^* is model that depends on modified bipower variation, Equation (2.7.7), $J^{\Delta *}$ and $J^{\pm *}$, are the models where the modified signed jump estimators are applied, Equations (2.7.5) and (2.7.6). The intercept is not reported.

	HAR	RS	BV	J^{Δ}	$J^{\pm}$
β_J^{Δ}				-0.294	
				(-4.96)	
β_J^{+}					-0.301
					(-5.52)
β_J^{-}					-0.444
					(-3.58)
β_d^{+}		0.009			
		(0.19)			
β_d^{-}		0.873			
		(11.58)			
β_C			0.440	0.446	0.434
			(12.78)	(13.15)	(10.36)
β_d	0.425				
	(13.38)				
β_w	0.374	0.373	0.365	0.366	0.369
	(8.57)	(8.73)	(9.03)	(9.11)	(9.32)
β_m	0.113	0.109	0.117	0.119	0.115
	(2.61)	(2.57)	(2.70)	(2.77)	(2.69)
R^2	0.598	0.618	0.603	0.618	0.618

Table A.3.2. In-sample estimation results of the volatility models for monthly forecast

Notes: The table provides in-sample parameter estimates and measures of fit for the volatility models with volatility and signed jump estimators for monthly forecast. The brackets report the Heteroskedasticity and Autocorrelation Consistent (HAC) t-statistics. HAR^* is a specification which uses the realized variance, Equation (2.7.3), RS^*, Equation (2.7.4), is based on modified semivariances, BV^* is model that depends on modified bipower variation, Equation (2.7.7), $J^{\Delta *}$ and $J^{\pm *}$, are the models where the modified signed jump estimators are applied, Equations (2.7.5) and (2.7.6). The intercept is not reported.

	HAR	RS	BV	J^{Δ}	$J^{\pm}$
β_J^{Δ}				-0.234	
				(-6.53)	
β_J^{+}					-0.172
					(-2.55)
β_J^{-}					-0.319
					(-2.84)
β_d^{+}		0.024			
		(0.53)			
β_d^{-}		0.606			
		(7.89)			
β_C			0.303	0.312	0.297
			(6.59)	(6.77)	(6.24)
β_d	0.210				
	(3.81)				
β_w	0.366	0.331	0.332	0.334	0.335
	(4.74)	(5.90)	(5.69)	(5.80)	(5.71)
β_m	0.158	0.128	0.136	0.133	0.133
	(2.27)	(1.85)	(1.97)	(1.90)	(1.90)
R^2	0.447	0.447	0.429	0.439	0.440

Table A.3.3. In-sample estimation results of the modified volatility models for weekly forecast

Notes: The table provides in-sample parameter estimates and measures of fit for the volatility models with modified volatility and signed jump estimators for weekly forecast. The brackets report the Heteroskedasticity and Autocorrelation Consistent (HAC) t-statistics. HAR^* is a specification which uses the modified realized variance for Equation (2.7.3), RS^* modifies Equation (2.7.4) based on modified semivariances, BV^* is alternative specification of Equation (2.7.7) that depends on modified bipower variation, $J^{\Delta *}$ and $J^{\pm *}$ are the modified versions of Equations (2.7.5) and (2.7.6), where the modified signed jump estimators are applied. The intercept is not reported.

	HAR^*	RS^*	BV^*	$J^{\Delta *}$	$J^{\pm *}$
$\beta_J^{\Delta *}$				-0.210	
				(-3.13)	
$\beta_J^{+ *}$					-0.015
					(-0.16)
$\beta_J^{- *}$					-0.586
					(-1.69)
$\beta_d^{+ *}$		0.051			
		(0.64)			
$\beta_d^{- *}$		0.397			
		(2.98)			
β_c^*			0.201	0.367	0.200
			(2.47)	(7.59)	(2.30)
β_d^*	0.203				
	(2.49)				
β_w^*	0.483	0.475	0.478	0.408	0.471
	(6.76)	(6.55)	(6.66)	(5.46)	(6.75)
β_m^*	0.161	0.156	0.164	0.105	0.157
	(3.20)	(3.13)	(3.28)	(1.67)	(3.25)
R^2	0.590	0.592	0.592	0.336	0.594

Table A.3.4. In-sample estimation results of the modified volatility models for monthly forecast

Notes: The table provides in-sample parameter estimates and measures of fit for the volatility models with modified volatility and signed jump estimators for monthly forecast. The brackets report the Heteroskedasticity and Autocorrelation Consistent (HAC) t-statistics. HAR^* is a specification which uses the modified realized variance for Equation (2.7.3), RS^* modifies Equation (2.7.4) based on modified semivariances, BV^* is alternative specification of Equation (2.7.7) that depends on modified bipower variation, J^{4*} and $J^{\pm *}$ are the modified versions of Equations (2.7.5) and (2.7.6), where the modified signed jump estimators are applied. The intercept is not reported.

	HAR^*	RS^*	BV^*	$J^{\Delta *}$	$J^{\pm *}$
$\beta_J^{\Delta *}$				-0.063	
				(-0.88)	
$\beta_J^{+ *}$					-0.079
					(-1.06)
$\beta_J^{- *}$					-0.418
					(-1.30)
$\beta_d^{+ *}$		-0.016			
		(-0.20)			
$\beta_d^{- *}$		0.358			
		(3.29)			
β_c^*			0.146	0.168	0.153
			(2.24)	(3.28)	(2.18)
β_d^*	0.148				
	(2.29)				
β_w^*	0.402	0.392	0.400	0.287	0.393
	(4.40)	(4.23)	(4.27)	(2.75)	(4.24)
β_m^*	0.177	0.172	0.178	0.128	0.173
	(2.47)	(2.42)	(2.52)	(1.81)	(2.46)
R^2	0.442	0.443	0.443	0.104	0.442

A. 4. More empirical results for Chapter 3

Table A.4.1. The impact of drift bursts on daily realized variance forecasts with model estimated by the Weighted Least Square method

This table reports the results of coefficient estimation and goodness of fit (adjusted R^2) for the four models that investigate the impacts of different drift burst measures on daily realized variance $RV_{t,t+1}$ forecasts. The row head contains the variables used in these models. All regressions contain RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ as the independent variables, where RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ indicate the lagged daily, weekly, and monthly average realized variance, Equation (3.5.1). The first column reports the results of the model that depends on RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$. The second column shows the results for the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14). The third column contains the results for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2). The fourth column reports the results of the specifications that are based on mutual-excitation of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4). The sample is S&P 500 E-mini future transaction prices from June 2003 to December 2020.

Constant	0.163	0.177	0.161	0.111
	(3.37)	(3.62)	(3.68)	(1.45)
RV_t	0.579	0.577	0.641	0.577
	(13.02)	(12.97)	(14.37)	(13.00)
$RV_{t-5,t}$	0.341	0.342	0.302	0.341
	(5.58)	(5.60)	(5.62)	(5.60)
$RV_{t-22,t}$	0.036	0.035	0.032	0.034
	(1.23)	(1.22)	(1.18)	(1.17)
NDB_t^+		-0.316		
		(-2.48)		
NDB_t^-		-0.013		
		(-0.11)		
VDB_t^+			-1.058	
			(-3.93)	
VDB_t^-			-0.748	
			(-1.93)	
IDB_t^+				-0.017
				(-0.67)
IDB_t^-				0.025
				(1.44)
R^2	0.574	0.574	0.596	0.574

Table A.4.2. The impact of drift bursts on weekly realized variance forecasts with model estimated by the Weighted Least Square method

This table reports the results of coefficient estimation and goodness of fit (adjusted R^2) for the four models that investigate the impacts of different drift burst measures on weekly realized variance $RV_{t,t+5}$ forecasts. The row head contains the variables used in these models. All regressions contain RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ as the independent variables, where RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ indicate the lagged daily, weekly, and monthly average realized variance, Equation (3.5.1). The first column reports the results of the model that depends on RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$. The second column shows the results for the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14). The third column contains the results for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2). The fourth column reports the results of the specifications that are based on mutual-excitation of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4). The sample is S&P 500 E-mini future transaction prices from June 2003 to December 2020.

Constant	0.442	0.429	0.355	-0.083
	(5.02)	(4.90)	(3.98)	(-0.21)
RV_t	0.256	0.256	0.570	0.486
	(3.18)	(3.12)	(11.51)	(12.27)
$RV_{t-5,t}$	0.469	0.467	0.311	0.353
	(4.84)	(4.79)	(4.89)	(5.24)
$RV_{t-22,t}$	0.099	0.101	0.049	0.053
	(1.69)	(1.72)	(0.96)	(0.98)
NDB_t^+		-0.246		
		(-1.62)		
NDB_t^-		0.448		
		(1.67)		
VDB_t^+			-0.819	
			(-2.87)	
VDB_t^-			-0.869	
			(-1.54)	
IDB_t^+				0.027
				(0.22)
IDB_t^-				0.092
				(1.38)
R^2	0.636	0.635	0.653	0.634

Table A.4.3. The impact of drift bursts on monthly realized variance forecasts with model estimated by the Weighted Least Square method

This table reports the results of coefficient estimation and goodness of fit (adjusted R^2) for the four models that investigate the impacts of different drift burst measures on monthly realized variance $RV_{t,t+22}$ forecasts. The row head contains the variables used in these models. All regressions contain RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ as the independent variables, where RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$ indicate the lagged daily, weekly, and monthly average realized variance, Equation (3.5.1). The first column reports the results of the model that depends on RV_t, $RV_{t-5,t}$, and $RV_{t-22,t}$. The second column shows the results for the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14). The third column contains the results for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2). The fourth column reports the results of the specifications that are based on the mutual-excitation intensity of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4). The sample is S&P 500 E-mini future transaction prices from June 2003 to December 2020.

Constant	1.057	1.033	0.973	-0.912
	(4.03)	(4.08)	(3.67)	(-0.56)
RV_t	0.187	0.181	0.394	0.176
	(2.64)	(2.60)	(5.59)	(3.04)
$RV_{t-5,t}$	0.395	0.401	0.294	0.372
	(3.02)	(3.05)	(3.23)	(3.24)
$RV_{t-22,t}$	0.103	0.103	0.068	0.088
	(1.20)	(1.20)	(0.84)	(0.96)
NDB_t^+		-0.301		
		(-1.35)		
NDB_t^-		0.766		
		(1.63)		
VDB_t^+			-0.619	
			(-1.44)	
VDB_t^-			-0.507	
			(-2.96)	
IDB_t^+				0.090
				(0.28)
IDB_t^-				0.433
				(1.40)
R^2	0.449	0.449	0.439	0.467

Table A.4.4. The impact of drift bursts on option implied variance with model estimated by the Weighted Least Square method

This table reports the regression estimation results for investigating the impact of different drift bursts measure on the implied variance, which is defined by the squared daily close value of the Chicago Board Options Exchange (CBOE) S&P 500 implied variance index VIX_t^2. All regressions contain VIX_t^2 as the dependent variable and VIX_{t-1}^2 as the independent variable, where VIX_{t-1}^2 denotes the lagged daily implied variance. The head of each row contains the variables used in the regression models. The first column reports the coefficient estimation results and measure of fit (R^2) for the model that depends on VIX_{t-1}^2. The second column shows the results for the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14.) The third column contains the results for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2). The fourth column reports the results of the specifications that are based on the mutual excitation intensity of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4) The sample is the S&P 500 E-mini future transaction prices and the daily close of the CBOE VIX index from June 2003 to December 2020.

Constant	0.053	0.040	0.060	-0.022
	(2.81)	(2.17)	(3.22)	(-0.39)
VIX_{t-1}^2	0.876	0.871	0.837	0.874
	(29.87)	(29.94)	(25.77)	(29.93)
NDB_t^+		-0.026		
		(-0.24)		
NDB_t^-		0.375		
		(5.09)		
VDB_t^+			0.439	
			(1.19)	
VDB_t^-			0.317	
			(3.46)	
IDB_t^+				-0.008
				(-0.85)
IDB_t^-				0.018
				(2.72)
R^2	0.931	0.932	0.936	0.931

Table A.4.5. The impact of drift bursts on daily realized variance forecasts with a drift burst robust volatility estimator

This table reports the results of coefficient estimation and goodness of fit (adjusted R^2) for the four models that investigate the impacts of different drift burst measures on daily realized variance $RiceV_{t,t+1}$ forecasts. The row head contains the variables used in these models. All regressions contain $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$ as the independent variables, where $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$ indicate the lagged daily, weekly, and monthly average drift burst robust volatility estimators, Equation (3.5.1). The first column reports the results of the model that depends on $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$. The second column shows the results for the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14). The third column contains the results for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2). The fourth column reports the results of the specifications that are based on the mutual excitation intensities of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4). The sample is S&P 500 E-mini future transaction prices from June 2003 to December 2020.

Constant	0.461	0.463	0.463	0.461
	(47.00)	(45.51)	(46.79)	(7.76)
$RiceV_t$	0.545	0.548	0.549	0.544
	(29.02)	(28.27)	(28.67)	(29.19)
$RiceV_{t-5,t}$	0.282	0.282	0.282	0.281
	(12.85)	(12.78)	(12.85)	(12.79)
$RiceV_{t-22,t}$	0.122	0.120	0.120	0.121
	(6.69)	(6.55)	(6.60)	(6.48)
NDB_t^+		-0.134		
		(-2.47)		
NDB_t^-		0.019		
		(0.43)		
VDB_t^+			-0.109	
			(-1.49)	
VDB_t^-			-0.015	
			(-0.27)	
IDB_t^+				-0.025
				(-1.22)
IDB_t^-				0.013
				(1.34)
R^2	0.741	0.741	0.741	0.741

Table A.4.6. The impact of drift bursts on weekly realized variance forecasts with a drift burst robust volatility estimator

This table reports the results of coefficient estimation and goodness of fit (adjusted R^2) for the four models that investigate the impacts of different drift burst measures on weekly realized variance $RiceV_{t,t+5}$ forecasts. The row head contains the variables used in these models. All regressions contain $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$ as the independent variables, where $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$ indicate the lagged daily, weekly, and monthly average drift burst robust volatility estimators, Equation (3.5.1). The first column reports the results of the model that depends on $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$. The second column shows the results for the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14). The third column contains the results for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2). The fourth column reports the results of the specifications that are based on the mutual-excitation intensities of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4). The sample is S&P 500 E-mini future transaction prices from June 2003 to December 2020.

Constant	0.543	0.541	0.543	0.435
	(30.88)	(30.48)	(30.77)	(3.74)
$RiceV_t$	0.423	0.420	0.423	0.420
	(20.57)	(20.02)	(20.32)	(21.39)
$RiceV_{t-5,t}$	0.300	0.303	0.300	0.295
	(9.83)	(9.93)	(9.89)	(9.77)
$RiceV_{t-22,t}$	0.171	0.171	0.171	0.165
	(5.30)	(5.30)	(5.32)	(4.90)
NDB_t^+		-0.076		
		(-1.47)		
NDB_t^-		0.069		
		(1.69)		
VDB_t^+			-0.028	
			(-0.35)	
VDB_t^-			0.009	
			(0.18)	
IDB_t^+				-0.015
				(-0.33)
IDB_t^-				0.095
				(1.64)
R^2	0.737	0.737	0.737	0.738

Table A.4.7. The impact of drift bursts on monthly realized variance forecasts with a drift burst robust volatility estimator

This table reports the results of coefficient estimation and goodness of fit (adjusted R^2) for the four models that investigate the impacts of different drift burst measures on monthly realized variance $RiceV_{t,t+22}$ forecasts. The row head contains the variables used in these models. All regressions contain $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$ as the independent variables, where $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$ indicate the lagged daily, weekly, and monthly average drift burst robust volatility estimators, Equation (3.5.1). The first column reports the results of the model that depends on $RiceV_t$, $RiceV_{t-5,t}$, and $RiceV_{t-22,t}$. The second column shows the results for the models that exploit the occurrence of drift bursts, NDB_t^+ and NDB_t^-, Equations (3.2.13) and (3.2.14). The third column contains the results for the regression that uses the variation of bursts, VDB_t^+ and VDB_t^-, Equations (3.3.)1 and (3.3.2). The fourth column reports the results of the specifications that are based on the mutual-excitation intensities of bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4). The sample is S&P 500 E-mini future transaction prices from June 2003 to December 2020.

Constant	0.629	0.625	0.626	0.338
	(15.37)	(15.45)	(15.29)	(1.16)
$RiceV_t$	0.288	0.281	0.280	0.283
	(11.18)	(11.26)	(11.09)	(12.49)
$RiceV_{t-5,t}$	0.274	0.280	0.277	0.263
	(7.90)	(7.88)	(7.98)	(8.47)
$RiceV_{t-22,t}$	0.198	0.199	0.199	0.180
	(3.44)	(3.48)	(3.50)	(2.88)
NDB_t^+		-0.039		
		(-0.70)		
NDB_t^-		0.110		
		(1.06)		
VDB_t^+			0.055	
			(0.70)	
VDB_t^-			0.093	
			(1.29)	
IDB_t^+				-0.030
				(-0.40)
IDB_t^-				0.081
				(1.55)
R^2	0.588	0.588	0.588	0.596

Table A.4.8. The impact of drift bursts on the volatility risk premium

This table reports the estimation results for four regressions, aiming to investigate the impact of different drift burst measures on volatility risk premium. The dependent variable for all models is uniformly set as $sqrt\_VRP_t$, which is the volatility risk premium of day t, and all these models include an independent variable $sqrt\_VRP_{t-1}$, which is the one-day lag of the volatility risk premium. The head of each row contains the names of the explanatory variables used in the regression. The first column contains results for the model that depend on the one-day lag of variance risk premium. The second column reports the results for the model that uses the occurrence of positive and negative bursts, NDB_t^+ and NDB_t^-. The third column is for the specification that includes the variation of positive and negative bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2), and the fourth column contains results for the model that is based on the mutual-excitation intensities of positive and negative bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4), respectively. The sample data is S&P 500 E-mini future prices and daily close of the CBOE VIX index from June 2003 to December 2020.

Constant	0.039	0.036	0.042	0.004
	(8.24)	(7.82)	(6.89)	(0.17)
$sqrt\_VRP_{t-1}$	0.899	0.897	0.884	0.888
	(101.32)	(94.86)	(59.07)	(87.25)
NDB_t^+		-0.009		
		(-0.15)		
NDB_t^-		0.086		
		(5.01)		
VDB_t^+			0.041	
			(1.07)	
VDB_t^-			0.016	
			(3.14)	
IDB_t^+				-0.007
				(-1.13)
IDB_t^-				0.012
				(3.29)
R^2	0.808	0.810	0.812	0.809

Table A.4.9. The impact of drift bursts on the logarithm variance risk premium

This table reports the estimation results for four regressions, aiming to investigate the impact of different drift burst measures on variance risk premium. The dependent variable for all models is uniformly set as log_VRP_t, which is the logarithm variance risk premium of day, and all these models include an independent variable log_VRP_{t-1}, which is the one-day lag of the logarithm variance risk premium. The head of each row contains the names of the explanatory variables used in the regression. The first column contains results for the model that depend on the one-day lag of variance risk premium. The second column reports the results for the model that uses the occurrence of positive and negative bursts, NDB_t^+ and NDB_t^-. The third column is for the specification that includes the variation of positive and negative bursts, VDB_t^+ and VDB_t^-, Equations (3.3.1) and (3.3.2), and the fourth column contains results for the model that is based on the mutual-excitation intensity of positive and negative bursts, IDB_t^+ and IDB_t^-, Equations (3.4.3) and (3.4.4), respectively. The sample data is S&P 500 E-mini future prices and daily close of the CBOE VIX index from June 2003 to December 2020.

Constant	0.042	0.040	0.042	0.053
	(9.84)	(9.55)	(9.76)	(3.84)
log_VRP_{t-1}	0.901	0.900	0.900	0.896
	(119.30)	(120.15)	(114.29)	(110.95)
NDB_t^+		-0.055		
		(-1.59)		
NDB_t^-		0.075		
		(6.32)		
VDB_t^+			0.002	
			(0.14)	
VDB_t^-			0.005	
			(3.88)	
IDB_t^+				-0.011
				(-1.80)
IDB_t^-				0.006
				(3.76)
R^2	0.812	0.815	0.812	0.812

A. 5. More empirical results for Chapter 4

Table A.5.1. In-sample estimation results of models with daily, weekly, and monthly market drift bursts lags

Notes: This table presents the results of the model that includes daily, weekly, and monthly lags of positively truncated drift bursts, indicated by RD_t^+.

$$RV_{t,t+h} = \beta_0 + \beta_{rd,d}^+ RD_t^+ + \beta_{rd,w}^+ RD_{t-5,t}^+ + \beta_{rd,m}^+ RD_{t-22,t}^+ + \beta_d RV_t + \beta_w RV_{t-5,t} + \beta_m RV_{t-22,t} + \varepsilon_t.$$

The brackets are the Heteroskedasticity- and autocorrelation-consistent (HAC) robust t-statistics by Newey and West (1987). The estimation for the intercept is not reported.

	$\beta_{rd,d}^+$	$\beta_{rd,w}^+$	$\beta_{rd,m}^+$	β_d	β_w	β_m	R^2
h=1	0.982	1.399	-0.716	0.483	0.298	0.124	0.540
	(1.33)	(1.00)	(-0.52)	(5.59)	(3.80)	(2.14)	
h=5	1.315	0.360	-1.773	0.331	0.349	0.235	0.640
	(2.38)	(0.26)	(-0.71)	(6.75)	(4.24)	(2.42)	
h=22	0.604	-0.343	-3.074	0.196	0.334	0.358	0.568
	(2.50)	(-0.33)	(-0.68)	(4.48)	(2.12)	(2.35)	
h=66	0.320	-1.247	8.828	0.127	0.253	-0.028	0.306
	(2.52)	(-0.89)	(0.86)	(3.28)	(2.09)	(-0.09)	

Table A.5.2. In-sample estimation results of models with daily, weekly, and monthly market drift bursts lags

Notes: This table presents the results of the model that includes daily, weekly, and monthly lags of pretested drift bursts, indicated by RD_t.

$$RV_{t,t+h} = \beta_0 + \beta_{rd,d} RD_t + \beta_{rd,w} RD_{t-5,t} + \beta_{rd,m} RD_{t-22,t} + \beta_d RV_t + \beta_w RV_{t-5,t} + \beta_m RV_{t-22,t} + \varepsilon_t.$$

The brackets are the Heteroskedasticity- and autocorrelation-consistent (HAC) robust t-statistics by Newey and West (1987). The estimation for the intercept is not reported.

	$\beta_{rd,d}$	$\beta_{rd,w}$	$\beta_{rd,m}$	β_d	β_w	β_m	R^2
h=1	0.324	-1.387	1.458	0.525	0.328	0.105	0.530
	(0.36)	(-1.36)	(0.57)	(6.13)	(4.47)	(2.31)	
h=5	-0.109	-1.035	2.012	0.365	0.370	0.169	0.636
	(-0.27)	(-0.75)	(0.57)	(6.77)	(4.56)	(2.93)	
h=22	-0.154	-0.438	-0.671	0.211	0.345	0.254	0.562
	(-0.74)	(-0.26)	(-0.15)	(5.29)	(2.21)	(3.24)	
h=66	-0.145	-1.181	4.699	0.131	0.221	0.218	0.301
	(-0.75)	(-0.71)	(0.57)	(4.14)	(2.26)	(3.04)	

Table A.5.3. In-sample estimation results of models with daily, weekly, and monthly codrift variations lags

Notes: This table presents the results of the model that includes daily, weekly, and monthly lags of pretested drift bursts and codrift variations, indicated by RD_t and $coRD_t$.

$$RV_{t,t+h} = \beta_0 + \beta_{rd,d}\, RD_t + \beta_{rd,w}\, RD_{t-5,t} + \beta_{rd,m}\, RD_{t-22,t} + \beta^c_{rd,d}\, coRD_t + \beta^c_{rd,w}\, coRD_{t-5,t} + \beta^c_{rd,m}\, coRD_{t-22,t} + \beta_d\, RV_t + \beta_w\, RV_{t-5,t} + \beta_m\, RV_{t-22,t} + \varepsilon_t,$$

The brackets are the Heteroskedasticity- and autocorrelation-consistent (HAC) robust t-statistics by Newey and West (1987). The estimation for the intercept is not reported.

	$\beta_{rd,d}$	$\beta_{rd,w}$	$\beta_{rd,m}$	$\beta^c_{rd,d}$	$\beta^c_{rd,w}$	$\beta^c_{rd,m}$	β_d	β_w	β_m	R^2
h=1	0.149	-1.441	1.758	0.065	0.073	0.042	0.495	0.289	0.091	0.549
	(0.18)	(-1.42)	(0.61)	(1.17)	(0.68)	(0.54)	(6.33)	(4.09)	(1.87)	
h=5	-0.138	-1.004	1.845	0.035	0.008	0.249	0.341	0.334	0.104	0.656
	(-0.37)	(-0.46)	(0.44)	(1.41)	(0.11)	(1.67)	(7.21)	(4.42)	(1.52)	
h=22	-0.299	-0.580	-1.488	0.024	0.119	0.294	0.191	0.288	0.154	0.600
	(-1.29)	(-0.41)	(-0.30)	(1.77)	(1.33)	(1.52)	(4.66)	(2.06)	(1.40)	
h=66	-0.293	-1.737	-2.355	0.016	-0.007	0.935	0.108	0.183	-0.037	0.364
	(-1.12)	(-0.94)	(-0.54)	(2.40)	(-0.10)	(1.32)	(3.75)	(2.15)	(-0.17)	