



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

DIPLOMARBEIT

Nichtparametrische Volatilitätsschätzer unter Market Microstructure Noise

Ausgeführt am Institut für

Wirtschaftsmathematik

Forschungsgruppe: Finanz- und Versicherungsmathematik
der Technischen Universität Wien

unter der Anleitung von

Univ.Prof. Dipl.-Math. Dr.rer.nat. Thorsten Rheinländer

durch

Rainer Hirk
Jesuitensteig 5
1230 Wien

Wien, September 2014

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Aufbau der Arbeit	3
2	Grundlegende Annahmen und Definitionen	5
2.1	Modellannahmen	5
2.2	Market Microstructure Noise	9
2.3	Sampling Schemes	11
2.4	Weitere wahrscheinlichkeitstheoretische Definitionen	11
2.5	Verschiedene Arten der Volatilität	13
3	Volatilitätschätzer	15
3.1	Realized Variance	17
3.1.1	Bias des RV unter allgemeinem Microstructure Noise	18
3.1.2	Asymptotisches Verhalten	18
3.1.3	Schätzer für den Noiseprozess	19
3.1.4	MSE	20
3.1.5	RV bei abhängigem Noiseprozess	20
3.2	Zhou Estimator	22
3.2.1	Asymptotisches Verhalten	22
3.2.2	MSE	24
3.3	Two Scaled Realized Volatility Estimator	24
3.3.1	Subsampling und Averaging	25
3.3.2	Definition TSRV	29
3.3.3	Asymptotisches Verhalten	29
3.3.4	Optimales Sampling und Biaskorrektur	30
3.3.5	Erweiterung des TSRV bei abhängigem Noiseprozess	31
3.4	Multi Scales Realized Volatility Estimator	36
3.4.1	Motivation des MSRVE Estimators	36

Inhaltsverzeichnis

3.4.2	Asymptotisches Verhalten des Noiseterms und Bestimmung der Gewichte	39
3.4.3	Diskretisierungsfehler	41
3.4.4	Asymptotisches Verhalten des MSRV	43
3.5	Realized Kernel Estimator	44
3.5.1	Annahmen an den Microstructure Noiseprozess	45
3.5.2	Definition Realized Kernel	45
3.5.3	Asymptotisches Verhalten	46
3.5.4	Anpassungen an den Microstructure Noiseprozess	50
3.6	Fourier Estimator	52
3.6.1	Definition Fourier Volatility Estimator	52
3.6.2	Bias	55
3.6.3	MSE	55
3.6.4	Abhängiger Microstructure Noiseprozess	58
3.6.5	Wahl der optimalen Frequenz	59
4	Empirische Studie	61
4.1	Die Daten	61
4.2	Analyse der einzelnen Schätzer	64
4.2.1	RV	64
4.2.2	Zhou Estimator	66
4.2.3	TSRV	68
4.2.4	MSRV	68
4.2.5	Kernel	69
4.2.6	Fourier	70
4.3	Vergleich der verschiedenen Schätzer	71
4.4	Zusammenhang zwischen der Gestalt des Limit Order Books und Volatilität	75
5	Schlussfolgerung	79
6	Programmcodes	81
	Literaturverzeichnis	91

1 Einleitung

1.1 Motivation¹

In der Finanzwirtschaft ist die *Volatilität*, besonders bei der Bewertung von Optionen im Black-Scholes-Modell, ein wesentlicher Parameter und kann allgemein als Schwankung von Zeitreihen interpretiert werden. Die ständige Veränderung der Vermögenspreise ist ein natürliches Phänomen der Finanzmärkte. Oft wird Volatilität zu Unrecht ausschließlich mit Unsicherheit verbunden. Tatsächlich erweist sich aber nicht jede Kursschwankung hinsichtlich der Stabilität der Märkte problematisch. Volatilität an sich ist vielmehr das Ausmaß, in dem die Preise der am Markt gehandelten Finanzinstrumente innerhalb einer bestimmten Zeitspanne schwanken. Sie macht sich zunehmend in immer mehr verschiedenen wirtschaftlichen Bereichen bemerkbar. Dabei nimmt oft die Frequenz und die Stärke der unerwarteten Ausschläge zu.

Preise und Renditen werden im Finanzmarkt durch ständig neue Informationen sowie die ständige Veränderung von Angebot und Nachfrage beeinflusst. Mit Hilfe der Volatilität ist es nicht möglich, eine Richtung der Ausschläge anzugeben, sie spiegelt lediglich die Intensität der Veränderung der zugrundeliegenden Faktoren sowie der dadurch entstehenden Unsicherheit der Einschätzung künftiger Entwicklungen wider. Dies führt zwangsläufig zu einer gewissen Streuung der Preise. In der Finanzwirtschaft lässt sich die Volatilität als ein Fluktuationsmaß definieren, welches die Schwankungsintensität von Wertpapierkursen, Rohstoffpreisen oder auch Zinsen innerhalb eines bestimmten Zeitraums misst. Je höher die Volatilität, umso stärker schlägt der Kurs nach oben bzw. unten aus und desto risikoreicher ist das Produkt.

Im Allgemeinen kann zwischen *historischer Volatilität* und *impliziter Volatilität* unterschieden werden. Als *historische Volatilität* wird jene Volatilität bezeichnet, die aus Zeitreihen historischer Wertänderungen berechnet wird. Im Gegensatz dazu basiert die *implizite Volatilität* nicht auf historischen Zeitreihen, sondern wird aus den Marktpreisen von Optionen ermittelt. Sie ist jene Volatilität des Basiswertes einer Option, die in ein Optionspreismo-

¹vgl. [Vol14b], [Vol14c]

1 Einleitung

dell eingesetzt den tatsächlich beobachteten Optionspreis ergibt.

Wir werden in dieser Diplomarbeit den Fokus auf die Schätzung der historischen Volatilität legen. Die Volatilität kann als Standardabweichung der Renditen betrachtet werden und wird oft für jährliche Veränderungen angegeben. Es ist leicht möglich, eine Beziehung zwischen Tagesvolatilität und Jahresvolatilität herzustellen.

$$\text{Tagesvolatilität} \approx \frac{\text{Jahresvolatilität}}{\sqrt{\text{Anzahl der Handelstage}}}$$

Ein Hauptproblem der Volatilität ist, dass sie am Markt nicht direkt beobachtbar ist. Sie kann also nur geschätzt werden und es wird daher immer eine gewisse Unsicherheit für die exakten Werte der Volatilität geben.

Aus folgenden Gründen haben Investoren Interesse an der Volatilität:

- Die Volatilität kann ausschlaggebend für die Platzierung des Produkts sein.
- Je größer die Ausschläge, desto schwieriger ist es, die Sachlage emotionslos zu betrachten.
- Falls Erträge zu einem bestimmten zukünftigen Zeitpunkt benötigt werden, erhöht eine hohe Volatilität die Wahrscheinlichkeit eines Ausfalls.
- Produkte mit hoher Volatilität haben einen großen Einfluss auf das Portfolio.
- Ein Portfolio mit Produkten hoher Volatilität benötigt eine große Diversifizierung.
- Hohe Volatilität kann durch Spekulation hohe Kursgewinne ermöglichen.
- Die Volatilität ist ein wichtiger Parameter bei Optionspreismodellen wie dem Black-Scholes-Modell.

Mittlerweile ist es möglich, die Volatilität direkt über Indizes, Optionen oder Variance Swaps zu handeln. Auf diese Weise ist es möglich, sich gegen die Volatilität abzusichern.

Die vorliegende Diplomarbeit wird sich insbesondere mit den in [GR10] vorgestellten Volatilitätsschätzern beschäftigen. Aufgrund des steigenden Interesse an der Volatilität, wird der Stellenwert solcher Schätzer immer weiter zunehmen. Wir werden die Vor- und Nachteile sowie das asymptotische Verhalten der diversen Schätzer für die Volatilität begutachten. Die richtige Handhabung der Volatilitätsschätzer ist von großer Bedeutung, da die richtige Implementierung eine große Menge an Wissen und Erfahrung voraussetzt.

1.2 Aufbau der Arbeit

Im zweiten Kapitel werden grundlegende Modellannahmen und wichtige Definitionen, die im weiteren Verlauf dieser Arbeit große Bedeutung haben, angeführt. Weiters werden der *Market Microstructure Noise*prozess sowie verschiedene *Sampling Schemes* genauer betrachtet. Zusätzlich werfen wir noch einen Blick auf die verschiedenen Arten der Volatilität.

Das dritte Kapitel stellt den Hauptteil dieser Arbeit dar. Dabei werden verschiedene Schätzer für die Volatilität, von der einfachen *Realized Variance* über den *Zhou Estimator*, den *TSRV* (Two Scaled Realized Volatility), den *MSRV* (Multi Scale Realized Volatility) bis hin zum *Kernel Estimator* und dem *Fourier Estimator* vorgestellt. Diese diversen Volatilitätsschätzer werden zusammen mit ihrem asymptotischem Verhalten, ihrer sonstigen Performance und ihren anderen Eigenschaften begutachtet.

Im vierten Kapitel wird eine empirische Studie durchgeführt. Zunächst werden die verwendeten Daten beschrieben und analysiert, damit in weiterer Folge die Volatilitätsschätzer einzeln analysiert werden können. Dabei werden auch die optimalen Samplingfrequenzen und der MSE (Mean Squared Error) der Schätzer bestimmt. Des Weiteren werden die Schätzer untereinander verglichen. Schlussendlich werden die Ergebnisse des Versuchs, einen Zusammenhang zwischen Volatilität und der Gestalt des Limit Order Books zu finden, beschrieben. Dieser Teil wurde gemeinsam mit meinem Studienkollegen David Hinterkörner durchgeführt.

Nach der Schlussfolgerung sind im sechsten Kapitel die Programmcodes für die Implementierung der Volatilitätsschätzer zu finden. Die Programme sind in MATLAB geschrieben. Es wurde die Version MATLAB 7.7.0 (R2008b) verwendet.

2 Grundlegende Annahmen und Definitionen

2.1 Modellannahmen²

Das Fundamental Theorem of Asset Pricing besagt, dass die Preise X_t in einem arbitragefreiem Markt *Semimartingalprozessen* ($X_t \in \mathcal{SM}$) auf einem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq T^*}, \mathbb{P})$ folgen. Dabei ist $T^* \leq 0$. Es ist zwar unüblich Semimartingalprozesse vor dem Zeitpunkt 0 zu starten, diese Annahme jedoch erleichtert oft die Anwendung, da einerseits oft Daten des vorherigen Tages benötigt werden und andererseits durch einfache Substitution beim Zeitpunkt 0 gestartet werden kann. Also stellt dies kein Problem dar.

Zur Erinnerung, Ergänzung und besseren Verständlichkeit werden noch einige Definitionen angeführt, die im Laufe dieser Arbeit große Bedeutung haben.

Definition 2.1.1. Ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit zugehöriger Filtrierung $\mathcal{F}_t, t \geq 0$ heißt *filtrierter Wahrscheinlichkeitsraum*.

Definition 2.1.2.³ Zu gegebenem filtrierten Wahrscheinlichkeitsraum ist ein stochastischer Prozess X ein *Semimartingal*, falls gilt:

- X ist \mathcal{F}_t adaptiert,
- X hat càdlàg (continue à droite, limite à gauche - rechtsseitig stetig und die linksseitigen Limites existieren) Pfade,
- es existiert eine Darstellung der Form

$$X = X_0 + M + A,$$

wobei X_0 fast sicher endlich und \mathcal{F}_0 -messbar, M ein lokales Martingal und A ein Prozess mit lokal endlicher Variation ist.

Definition 2.1.3.⁴ Ein *stochastischer Prozess* ist eine Familie von Zufallsvariablen ζ_t mit $t \in T$. Falls $T = [0, \infty)$ wird von einem stochastischen Prozess in stetiger Zeit gesprochen.

²vgl. [BNHLS08], S.1484

³vgl. [Sem14]

⁴vgl. [BZ05], S.139

2 Grundlegende Annahmen und Definitionen

Definition 2.1.4. ⁵ Ein stochastischer Prozess $(W_t)_{t \geq 0}$ heißt *Brown'sche Bewegung*, falls folgende Bedingungen erfüllt sind:

1. $W_0 = 0$ f.s.
2. die Pfade $t \rightarrow W_t$ sind f.s. stetig
3. W_t hat unabhängige und stationäre Inkremente
4. die Inkremente $W_t - W_s$ sind normalverteilt mit Mittelwert 0 und Varianz $t - s$ für alle $0 \leq s < t$

Die beliebteste Familie von Semimartingalen ist die Familie der *Brown'schen Semimartingalen*.

Definition 2.1.5. ⁶ Unter einem *Brown'schen Semimartingal* $(X_t \in \mathcal{BSM})$ verstehen wir einen stochastischen Prozess, der sich folgendermaßen darstellen lässt

$$X_t = \int_0^t a_s ds + \int_0^t \sigma_s dW_s,$$

wobei a_t ein vorhersehbarer lokal beschränkter Driftprozess, σ_t ein càdlàg Prozess und W_t eine Brown'sche Bewegung ist.

Definition 2.1.6. ⁷ Sei $\{X_n\}$ eine Folge von stochastischen Größen und X eine weitere stochastische Größe. Dann sprechen wir von *Konvergenz in der Wahrscheinlichkeit*, falls für alle $\epsilon > 0$ gilt

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| \geq \epsilon) = 0. \tag{2.1}$$

Wir verwenden dafür folgende Schreibweise

$$X_n \xrightarrow{P} X.$$

Unser Interesse liegt später an der quadratischen Variation des Preisprozesses, da eben diese uns einen ersten Schätzer für die Volatilität liefert.

Definition 2.1.7. ⁸ Sei $X_t(\cdot) : \Omega \rightarrow \mathbb{R}$ ein stetiger stochastischer Prozess, dann ist der *p-te Variationsprozess* von X_t , $\langle X, X \rangle_t^{(p)}$ definiert durch

$$\langle X, X \rangle_t^{(p)}(\omega) = \lim_{\Delta_{t_k} \rightarrow 0} \sum_{t_k \leq t} |X_{t_{k+1}}(\omega) - X_{t_k}(\omega)|^p, \tag{2.2}$$

⁵vgl. [BZ05], S.155f

⁶vgl. [BNHLS08], S.1484

⁷vgl. [Gur11], S.54

⁸vgl. [Oks00], S.19

2 Grundlegende Annahmen und Definitionen

wobei der Limes in Wahrscheinlichkeit zu verstehen ist und $0 = t_0 < t_1 < \dots < t_k = t$ sowie $\Delta t_k = t_{k+1} - t_k$ ist. Im Fall von $p = 1$ sprechen wir von der *totalen Variation* und im Fall $p = 2$ von der *quadratischen Variation*.

Der Preisprozess X_t folgt dabei einer stochastischen Differentialgleichung.

Definition 2.1.8. ⁹ Eine *stochastische Differentialgleichung* ist eine Gleichung der Form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \quad (2.3)$$

wobei $\mu, \sigma : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ zwei messbare Funktionen sind und W_t eine Brown'sche Bewegung ist.

Falls der Preisprozess gewisse Anforderungen erfüllt, erhalten wir einen wichtigen Bezug zur quadratischen Variation.

Annahme 2.1.9. Für $X_t \in \mathcal{BSM}$ sind die effizienten logarithmischen Preise durch folgende stochastische Differentialgleichung gegeben

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

In Integralschreibweise

$$\int_0^t dX_s = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s. \quad (2.4)$$

Für die quadratische Variation eines Prozesses $X_t \in \mathcal{BSM}$ erhalten wir

$$\langle dX, dX \rangle_t = \underbrace{\mu_t dt \sigma_t dW_t}_{=0} + \underbrace{\sigma_t dW_t \mu_t dt}_{=0} + \underbrace{\mu_t dt \mu_t dt}_{=0} + \underbrace{\sigma_t dW_t \sigma_t dW_t}_{=\sigma_t^2 \underbrace{(dW_t)^2}_{=dt} = \sigma^2 dt} = \sigma_t^2 dt.$$

Also in Integralschreibweise erhalten wir für die quadratische Variation von X_t

$$\langle X, X \rangle_t = \int_0^t \sigma_s^2 ds.$$

Unser Hauptinteresse liegt nun darin, einen stabilen Volatilitätsschätzer zu finden. Falls X_t tatsächlich dem beobachteten Preis entspricht, würde die quadratische Variation einen konsistenten Schätzer für die *integerierte Varianz*¹⁰

$$IV = \int_a^b \sigma_t^2 dt$$

liefern.

Wertänderungen können auf verschiedene Arten angegeben werden.

⁹vgl. [Gra12]

¹⁰vgl. [HL06], S.129

2 Grundlegende Annahmen und Definitionen

- absolute Veränderungen

$$r_{\text{abs}} = Y_{t_1} - Y_{t_0}$$

- relative Veränderungen

$$r_{\text{rel}} = \frac{Y_{t_1} - Y_{t_0}}{Y_{t_0}}$$

- logarithmische Veränderungen

$$r_{\text{log}} = \ln(Y_{t_1}) - \ln(Y_{t_0}) = \ln\left(\frac{\ln(Y_{t_1})}{\ln(Y_{t_0})}\right)$$

Dabei wird mit t_0 der Anfangszeitpunkt und mit t_1 der Endzeitpunkt bezeichnet. Die logarithmischen Erträge zwischen zwei Beobachtungen werden in weiterer Folge mit $Z_{t_i} = Y_{t_i} - Y_{t_{i-1}}$ bezeichnet.

Bemerkung 2.1.10. Für Finanzinstrumente, deren Preise einer Brown'schen Bewegung folgen, wächst die Volatilität mit der Quadratwurzel der Zeit.

Absolute Wertänderungen werden hauptsächlich bei Zinsen verwendet. Bei kleinen Veränderungen unterscheiden sich relative und logarithmische Volatilität kaum. Da die Preise in unserer Modellannahme durch eine stochastische Differentialgleichung beschrieben werden, werden wir hier logarithmische Veränderungen verwenden.

Nun ist es leider in der Praxis so, dass bei Hochfrequenz-Börsendaten ein sogenannter Market Microstructure Noiseprozess auftritt. Die beobachteten logarithmischen Preise Y_t setzen sich dann aus den unbeobachtbaren effizienten logarithmischen Preisen X_t und einem Noiseterm ϵ_t zusammen

$$Y_t = X_t + \epsilon_t.$$

Dabei wird angenommen, dass der Noiseprozess ϵ_t unabhängig vom Preisprozess X_t ist und für den Transaktionspreisprozess Stichprobendaten mit Zeitintervallen der Länge Δ vorliegen. Unter Ultra High Frequency ist gemeint, dass die Daten in Zeitabständen von Sekunden gemessen werden. In diesem Fall kann der Driftterm vernachlässigt werden. Wir setzen also $\mu_t = 0$ und betrachten den Volatilitätsprozess σ_t genauer. In der Praxis ist es üblich Transaktions- und Börsendaten an zufälligen Zeitpunkten zu beobachten. Der Einfachheit halber nehmen wir hier bei der asymptotischen Analyse der verschiedenen Schätzer Δ als nicht zufällig an. An den Volatilitätsprozess σ_t stellen wir keine wesentlichen Anforderungen, er kann ohne Weiteres mit der Brown'schen Bewegung W_t korreliert sein und hat nicht notwendigerweise stetige Pfade.

2.2 Market Microstructure Noise¹¹

Nun wollen wir den Microstructure Noiseprozess genauer betrachten. Der Noiseprozess ϵ_t fasst verschiedenste Störeffekte zusammen und ist hauptverantwortlich für die Inkonsistenz des Realized Variance Schätzers. Er kann im Großen und Ganzen in drei Klassen eingeteilt werden.

- Unstimmigkeiten im Tradingprozess selbst, wie Bid-Ask Bounces, unstetige Preisänderungen, oder Rundungsfehler. Bid-Ask Bounces entstehen durch folgendes Phänomen: Beobachtete Transaktionen erfolgen an den Bid- bzw. Ask-Preisen und verursachen durch die negative Autokorreliertheit im Hochfrequenzbereich Sprünge von der Bid zur Ask Seite und umgekehrt. Im Sekundenbereich tragen diese Sprünge zur Volatilität des beobachteten Preisprozess Y_t bei. Dieser Beitrag hat aber keinen Bezug zur Volatilität des wahren Preisprozess X_t . Aus diesem Grund verursachen Bid-Ask Bounces eine Erhöhung der Verzerrung von Hochfrequenzvolatilitätsschätzern.
- Informationelle Effekte, wie Unterschiede in den Tradegrößen, oder der informationelle Gehalt bei Preisänderungen, aber auch strategische Komponenten im Orderfluss oder aufsichtliche Kontrolleffekte.
- Mess- oder Aufzeichnungsfehler, wie zum Beispiel falsche Kommastellen oder falsch übernommene Preise. Aus diesem Grund ist es wichtig, die vorhandenen Daten zu säubern und somit den Noiseprozess zu vermindern um im Endeffekt stabile Voraussetzungen zu schaffen.

An den Microstructure Noiseprozess stellen wir in dieser Arbeit zumeist folgende Anforderungen.

Annahme 2.2.1. ¹² Der Noiseprozess ϵ_t erfüllt folgende Annahmen:

1. $X_t \perp \epsilon_t$, $\epsilon_s \perp \epsilon_t$ für alle $s \neq t$ und $\mathbb{E}[\epsilon_t] = 0$ für alle t
2. $\omega^2 = \mathbb{E}[\epsilon_t^2] < \infty$ für alle t
3. $\mu_4 = \mathbb{E}[\epsilon_t^4] < \infty$ für alle t

Der unabhängige Noiseprozess ϵ_t induziert dabei einen MA(1) Prozess auf den Erträgen $\eta_t = \epsilon_t - \epsilon_{t-1}$.

Diese Annahmen führen oft zu guten Ergebnissen, obwohl sie für die Praxis meistens ungeeignet sind. In der Praxis haben wir zumeist folgende Situation:

¹¹vgl. [BMV12], [ASMZ08], [HL06], S.132f

¹²vgl. [HL06], S.132f

2 Grundlegende Annahmen und Definitionen

- der Noiseprozess ϵ_t ist korreliert mit den tatsächlichen Preisen X_t
- der Noiseprozess ist zeitabhängig
- die Eigenschaften des Noiseprozesses ändern sich über die Jahre

Aus diesen Gründen schrauben wir oft die Anforderungen an den Noiseprozess zurück und betrachten die diversen Schätzer unter abhängigem Noise.

Zeitreihenabhängigkeiten in High Frequency Market Microstructure Noise¹³

Im Fall von iid Noise folgen die logarithmischen Erträge

$$Y_{\tau_i} - Y_{\tau_{i-1}} = \int_{\tau_{i-1}}^{\tau_i} \sigma_t dW_t + \epsilon_{\tau_i} - \epsilon_{\tau_{i-1}}$$

einem MA(1) Prozess, da die Inkremente $\int_{\tau_{i-1}}^{\tau_i} \sigma_t dW_t$ unkorreliert sind und $\epsilon_t \perp W_t$. Falls σ_t nicht zufällig ist, erhalten wir

$$\mathbb{E}[(Y_{\tau_j} - Y_{\tau_{j-1}})(Y_{\tau_i} - Y_{\tau_{i-1}})] = \begin{cases} \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt + 2\mathbb{E}[\epsilon^2], & \text{wenn } j = i \\ -\mathbb{E}[\epsilon^2], & \text{wenn } j = i + 1 \\ 0, & \text{wenn } j > i + 1. \end{cases}$$

Unter der iid-Annahme sind die logarithmischen Erträge negativ autokorreliert 1. Ordnung.

Ein einfaches Modell mit abhängigen Noisetermen

Um die Abhängigkeit des Noiseprozesses zu erfassen, setzen wir

$$\epsilon_{t_i} = U_{t_i} + V_{t_i}, \tag{2.5}$$

wobei U unabhängig identisch verteilt ist, V $AR(1)$ mit Koeffizient ρ ist, $\|\rho\| < 1$ und $U \perp V$. Dann erhalten wir

$$\mathbb{E}[(Y_{\tau_j} - Y_{\tau_{j-1}})(Y_{\tau_i} - Y_{\tau_{i-1}})] = \begin{cases} \int_{\tau_{i-1}}^{\tau_i} \sigma_t^2 dt + 2\mathbb{E}[U^2] + 2(1 - \rho)\mathbb{E}[V^2], & \text{wenn } j = i \\ -\mathbb{E}[U^2] - (1 - \rho)^2\mathbb{E}[V^2], & \text{wenn } j = i + 1 \\ -\rho^{j-i-1}(1 - \rho)^2\mathbb{E}[V^2], & \text{wenn } j > i + 1. \end{cases}$$

¹³vgl. [ASMZ08]

2.3 Sampling Schemes¹⁴

Es gibt verschiedene Möglichkeiten Stichproben zu entnehmen. Die üblichste Methode ist das sogenannte *Calendar Time Sampling* (CTS). Dies ist der Spezialfall, wo die Beobachtungszeiten t_i mit $i = 1, \dots, n$ äquidistante Abstände $\delta_i = \frac{(b-a)}{n}$ haben. Dabei ist a der Startzeitpunkt und b der Endzeitpunkt. Falls die Stichproben unregelmäßige Zeitabstände aufweisen, muss ein Datensatz mit äquidistanten Zeitintervallen konstruiert werden. Sind die Preise zu den Zeiten $t_0 < \dots < t_N$ gegeben, dann können die Preise zum Zeitpunkt $\tau \in [t_j, t_{j+1})$ einerseits mit der *Previous Tick Methode*

$$Y_\tau = Y_{t_j},$$

oder andererseits mit der *Linear Interpolation Methode*

$$\tilde{Y}_\tau = Y_{t_j} + \frac{\tau - t_j}{t_{j+1} - t_j} (Y_{t_{j+1}} - Y_{t_j})$$

konstruiert werden. Im Falle von Hochfrequenz-Börsendaten hat die Linear Interpolation Methode folgende ungünstige Eigenschaft:

Proposition 2.3.1.¹⁵ Sei N fest und betrachten wir die RV basierend auf der Linear Interpolation Methode, dann gilt

$$RV_{\text{all}}^{(m)} \xrightarrow{P} 0 \quad \text{für } m \rightarrow \infty.$$

Bemerkung 2.3.2. Das Resultat aus der obigen Proposition läuft darauf hinaus, dass die quadratische Variation einer Geraden Null ist. Aus diesem Grund ist die Linear Interpolation Methode für hochfrequente Daten nicht geeignet. Daher bevorzugen wir die *Previous Tick Methode*.

Falls t_i Transaktionszeiten sind, sprechen wir von *Tick Time Sampling* (TTS). Wenn hingegen die Beobachtungszeiten t_0, \dots, t_n so gewählt werden, dass $\sigma_i^2 = \frac{IV}{n}$ für alle $i = 1, \dots, n$ gilt, sprechen wir vom *Business Time Sampling* (BTS).

2.4 Weitere wahrscheinlichkeitstheoretische Definitionen

Da wir in weiterer Folge asymptotische Verteilungen betrachten, wollen wir noch weitere grundlegende wahrscheinlichkeitstheoretische Definitionen anführen.

¹⁴vgl. [HL06], S.129f

¹⁵Def. $RV_{\text{all}}^{(n)}$ siehe 3.1

2 Grundlegende Annahmen und Definitionen

Definition 2.4.1. ¹⁶ Sei $\{X_n\}$ eine Folge von stochastischen Größen und X eine weitere stochastische Größe. F_{X_n} und F_X seien die Verteilungsfunktionen von X_n und X . $C(F_X)$ bezeichne die Menge der Stetigkeitspunkte von F_X , dann sprechen wir von *Konvergenz in der Verteilung*, falls für alle $x \in C(F_X)$ gilt

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x).$$

Wir verwenden dafür folgende Schreibweise

$$X_n \xrightarrow{D} X.$$

Definition 2.4.2. ¹⁷ Sei X eine stochastische Größe mit Verteilungsfunktion $F(x; \theta)$, $\theta \in \Theta$. Ist X_1, \dots, X_n eine Stichprobe von X und T_n eine Statistik, so ist T_n *konsistenter Schätzer* von θ , falls $T_n(X_1, \dots, X_n)$ in der Wahrscheinlichkeit gegen θ konvergiert.

Die Verteilung der Statistik T wird auch Stichprobenverteilung genannt. Eine für uns besonders wichtige Größe ist der *Mean Squared Error* (MSE)

$$MSE(T; \theta) = \mathbb{E}_\theta(T - \theta)^2.$$

Der MSE hat folgende Beziehung zur Varianz:

Proposition 2.4.3. Falls die $Var(T)$ existiert, gilt für den MSE von T

$$MSE(T; \theta) = Var_\theta(T) + [\mathbb{E}_\theta(T) - \theta]^2.$$

Der Ausdruck in den eckigen Klammern ist dabei die *Verzerrung*. Daher können wir obige Proposition auch folgendermaßen anschreiben:

$$MSE = \text{Varianz} + (\text{Verzerrung})^2.$$

Proposition 2.4.4. Ein Schätzer T mit $[\mathbb{E}_\theta(T) - \theta] \rightarrow 0$ und $Var_\theta(T) \rightarrow 0$ ist konsistent für θ .

Im nächsten Kapitel werden wir häufig Landau-Symbole¹⁸ verwenden. Diese haben den Zweck das asymptotische Verhalten von Funktionen und Folgen zu beschreiben. Dabei wird zwischen großem Landau-Symbol \mathcal{O} und kleinem Landau-Symbol o unterschieden. Sei $x \in \mathbb{R}$ und der Grenzwert $a \in \{-\infty, +\infty\}$, dann gilt:

- $f \in \mathcal{O}(g)$ bedeutet

$$\limsup_{x \rightarrow a} \left| \frac{f(x)}{g(x)} \right| < \infty.$$

In Worten, f wächst nicht wesentlich schneller als g .

¹⁶vgl. [Gur11], S.55

¹⁷vgl. [Gur11], S.67

¹⁸vgl. [Lan14]

- $f \in o(g)$ bedeutet

$$\lim_{x \rightarrow a} \left| \frac{f(x)}{g(x)} \right| = 0.$$

In Worten, f wächst langsamer als g .

2.5 Verschiedene Arten der Volatilität¹⁹

Volatilität ist ein theoretisches Konstrukt und hängt im Wesentlichen auch vom Modell des zugrundeliegendem Basiswertes ab. So erhalten wir verschiedene Arten der Volatilität:

- *Bedingte Standardabweichung*

Zu gegebener Information \mathcal{F}_n ist die Varianz des Ertrags r_{n+1} gefragt:

$$\text{Var}(r_{n+1} | \mathcal{F}_n).$$

Die Quadratwurzel dieser Größe ist dann die bedingte Standardabweichung und hängt von der gegebenen Information ab. Um diese explizit zu bestimmen, müssen Annahmen an die Renditen getroffen werden. Falls die Erträge als iid normalverteilt angenommen werden, ist die Volatilität konstant und der übliche Schätzer ist

$$\hat{\sigma} = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (r_n - \bar{r})^2},$$

wobei \bar{r} das arithmetische Mittel der Renditen ist.

- *Time Series Volatility*

Zeitreihenmodelle befassen sich mit zeitabhängiger Volatilität. Die Erträge diskreter zeitabhängiger Modelle haben oft folgende Struktur:

$$r_n = \sigma_n \eta_n.$$

Die Rendite r_n ist also das Produkt aus der Volatilität σ_n und einem Erneuerungsprozess η_n mit Erwartungswert Null und Varianz eins.

- *Spot Volatility*

Die Spot Volatility ist gewissermaßen die augenblickliche Volatilität. Dazu benötigen wir ein Modell für das Underlying wie zum Beispiel eine stochastische Differentialgleichung der Form:

$$dX_t = \mu_t dt + \sigma_t dW_t. \tag{2.6}$$

¹⁹vgl. [Vol14a]

2 Grundlegende Annahmen und Definitionen

Dabei ist W_t eine Brown'sche Bewegung und X_t der logarithmische Preisprozess. Die Variable σ_t wird dann Spot Volatility genannt und wird oft bei Optionspreismodellen, wie zum Beispiel dem Black-Scholes-Modell, verwendet.

- *Integrated Volatility*

Betrachten wir einen stetigen stochastischen Prozess über ein bestimmtes Zeitintervall und teilen dieses in mehrere kleine angrenzende Intervalle. Dann liefert die Approximation für die quadratische Variation $\sum_{i=1}^n (X_i - X_{i-1})^2$ einen Schätzer für die Volatilität. Falls die Preise die SDE (2.6) erfüllen, gleicht die quadratische Variation der integrierten Volatilität $\int_0^T \sigma_t^2 dt$. Diese Größe hat eine wesentliche Bedeutung bei der Bepreisung von Optionen. Unter der Annahme, dass der Volatilitätsprozess σ_t unabhängig von der Brown'schen Bewegung W_t ist, sind die logarithmischen Erträge $r_n \sim N(0, QV_n)$.

Wir werden in dieser Arbeit ausschließlich die integrierte Volatilität behandeln. Unser Hauptziel ist es, einen stabilen Schätzer für diese zu finden.

3 Volatilitätsschätzer²⁰

Wir betrachten ausschließlich den nichtparametrischen Fall, in dem σ_t ein stochastischer Prozess ist. Das Hauptinteresse liegt darin, einen konsistenten Schätzer für die integrierte Volatilität über ein fixes Intervall T zu finden. Typischerweise werden dabei bei empirischen Anwendungen Tagesintervalle verwendet.

$$IV = \langle X, X \rangle_T = \int_0^T \sigma_t^2 dt.$$

Falls kein Noiseterm vorliegt, liefert der *Realized Variance* Estimator (RV)

$$RV_{\text{all}}^{(n)} = \sum_{i=1}^n (Y_{t_{i+1}} - Y_{t_i})^2$$

eine gute Schätzung für die quadratische Variation $\langle X, X \rangle_T$. Der Index `all` gibt an, dass alle vorhandenen Daten verwendet werden, was auch der Theorie zur Folge den bestmöglichen Schätzer liefern sollte. Die $RV_{\text{all}}^{(n)}$ konvergiert dabei gegen die quadratische Variation $\langle X, X \rangle_T$.

Im Fall der Existenz von Microstructure Noise liefert die $RV_{\text{all}}^{(n)}$ leider keinen guten Schätzer für die quadratische Variation $\langle X, X \rangle_T$. Bei der Verwendung aller Daten führt die RV zu einer Schätzung der Varianz von ϵ_t und nicht mehr wie gewünscht der quadratischen Variation. Die Verzerrung von $2n\mathbb{E}[\epsilon^2]$ ist dabei eine Größenordnung höher als erwünscht. Stellen wir die RV als Funktion des Samplingintervalls $\Delta = \frac{T}{n}$ dar, so divergiert sie, falls die Anzahl der Beobachtungen n steigt.²¹

Es ist also nicht ratsam, RV mit zu hoher Frequenz zu berechnen. In der Praxis scheint dabei ein Stichprobenintervall Δ_{sparse} zwischen 1 und 30 Minuten geeignet zu sein. Da unsere Daten nur in einem geringeren Ausmaß verunreinigt sind, erhalten wir bei unserer empirischen Studie eine optimale Samplingfrequenz von $\bar{n} = 70$ Sekunden. Wir bezeichnen diesen RV Schätzer dann mit $\Delta_{\text{sparse}} = \frac{T}{\bar{n}}$ als $RV_{\text{sparse}}^{(\bar{n})}$.

²⁰vgl. [GR10]

²¹vgl. Abbildung 4.2

3 Volatilitätschätzer

Ein großes Problem dabei ist, dass eine große Menge an Daten verworfen werden muss und dies somit im Widerspruch zu statistischen Grundregeln steht. Zum Beispiel, falls $T = 1$ und sich alle $\Delta = 1$ Sekunden Transaktionen ereignen, dann beträgt der gesamte Stichprobenumfang $n = \frac{t}{\Delta} = 23400$. Falls aber die Frequenz auf 5 Minuten reduziert wird, bleibt nur noch ein Stichprobenumfang von $n_{sparse} = 78$ übrig. Es wird also nur noch ein sehr geringer Anteil der Daten verwendet.

Der erste leicht biasangepasste Schätzer, den wir hier vorstellen, ist der *Zhou Estimator*

$$RV_{\text{Zhou}}^{(n)} = \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 + \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-1}} - Y_{t_{i-2}}) \\ + \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i+1}} - Y_{t_i}),$$

welcher eine einfache Biaskorrektur durchführt, indem er die Autokovarianzen 1. Ordnung miteinbezieht. Der Zhou Estimator ist ein inkonsistenter Schätzer und es ist dabei nicht optimal alle Daten zu verwenden.

Ziel ist es nun, möglichst alle vernünftig relevanten Daten zu verwenden und dabei trotzdem noch stabil zu bleiben. All dies vereint der sogenannte *Two Scaled Realized Volatility Estimator* (TSRV)

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} = \underbrace{RV_{\text{avg}}^{(k)}}_{\text{slow time scale}} - \frac{\bar{n}}{n} \underbrace{RV_{\text{all}}^{(n)}}_{\text{fast time scale}},$$

wobei

$$RV_{\text{avg}}^{(k)} = \frac{1}{K} \sum_{k=1}^K RV_{\text{sparse}}^{(k)}.$$

Dabei wird die quadratische Variation zweier verschiedener Frequenzen geschätzt und dann mittels geeigneter Linearkombination über die gesamte Stichprobe gemittelt. Es wird dadurch ein konsistenter, asymptotisch unverzerrter Schätzer der quadratischen Variation $\langle X, X \rangle_T$ erhalten. Die Konvergenzrate beträgt dabei $n^{-\frac{1}{6}}$. Dieser TSRV Schätzer wird in Abschnitt 3.3 vorgestellt.

Eine weitere Verbesserung erhalten wir, indem wir nun über mehrere Skalen mitteln. Dies führt zur *Multi Scale Realized Volatility* (MSRV)

$$\widehat{\langle X, X \rangle}_T^{(\text{msrv})} = \sum_{i=1}^M \alpha_i [Y, Y]_T^{(K_i)},$$

3 Volatilitätschätzer

wobei a_i Gewichte sind. M ist dabei größer als 2 und gibt die Anzahl der verwendeten Skalen an. Dieser Schätzer weist eine bessere Konvergenzrate von $n^{-\frac{1}{4}}$ vor, ist aber aufwendiger zu implementieren und der Rechenaufwand erhöht sich entscheidend.

Eine weitere Klasse von Volatilitätsschätzern sind die *Realized Kernel* Schätzer

$$K(Y_\delta) = \gamma_0(Y_\delta) + \sum_{h=1}^H k\left(\frac{h-1}{H}\right) [\gamma_h(Y_\delta) + \gamma_{-h}(Y_\delta)],$$

wobei $k(x)$ für $x \in [0, 1]$ eine nichtstochastische Gewichtsfunktion ist. Diese Klasse von Schätzern hat den großen Vorteil, dass durch die Wahl verschiedener Kernel eine große Anpassungsfreiheit besteht und somit auch die Konvergenzrate angepasst werden kann. Bei optimaler Wahl der Parameter und bei Erfüllung gewisser Anforderungen an den Kernel erhalten wir dabei einen konsistenten Schätzer mit einer Konvergenzrate von $n^{-\frac{1}{4}}$.

Mit Hilfe der Fouriermethode, welche auf einer Verbindung zwischen der Fouriertransformierten des Preisprozesses und der Fouriertransformierten des Volatilitätsprozesses aufbaut, kann der *Fourier Estimator*

$$\hat{\sigma}_{n,N}^2 = \frac{(2\pi)^2}{2N+1} \sum_{s=-N}^N \mathcal{F}(dY)_n(s) \mathcal{F}(dY)_n(-s),$$

erhalten werden. Dieser ist zwar bei geeigneter Anzahl der Fourierkoeffizienten asymptotisch unverzerrt, jedoch ein inkonsistenter Schätzer.

3.1 Realized Variance²²

Der Realized Variance Estimator (RV) ist definiert durch

$$RV_{\text{all}}^{(n)} = [Y, Y]_T^{(\text{all})} = \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2. \quad (3.1)$$

Im Falle der Abwesenheit von Microstructure Noiseprozessen ist dies ein konsistenter Schätzer für die IV für $n \rightarrow \infty$. Der große Nachteil dieses Schätzer liegt darin, dass wir, falls alle Beobachtungen zur Berechnung herangezogen werden, unter Market Microstructure Noise einen verzerrten Schätzer erhalten. Wir werden die Bezeichnung $RV_{\text{all}}^{(n)}$ bevorzugen, da bei dieser die Beobachtungsabstände besser hervorgehen.

²²vgl. [HL06], [ZMAS05]

3.1.1 Bias des RV unter allgemeinem Microstructure Noise²³

Falls die Preise wie in der Praxis üblich durch einen Market Microstructure Noiseprozess gestört werden, erhalten wir einen verzerrten Schätzer. Unter der Annahme, dass $\frac{T}{n}$ die Beobachtungsabstände zwischen den einzelnen Preisen sind, divergiert die RV für wachsendes n . Für die Verzerrung können wir folgende Aussage

$$\mathbb{E} \left[RV_{\text{sparse}}^{(n)} - IV \right] = 2n\mathbb{E} [\epsilon^2]$$

treffen.

3.1.2 Asymptotisches Verhalten²⁴

Angenommen es gilt $Y_t = X_t + \epsilon_t$, dann hat die RV basierend auf den beobachteten Preisen Y_t folgende Form:

$$RV_{\text{all}}^{(n)} = [X, X]_T^{(\text{all})} + 2[X, \epsilon]_T^{(\text{all})} + [\epsilon, \epsilon]_T^{(\text{all})}. \quad (3.2)$$

Proposition 3.1.1. Weiters nehmen wir an, dass die logarithmischen Preise die stochastische Differentialgleichung (2.3) und der Noiseprozess die Annahmen 2.2.1 Punkt 1 und 2 erfüllt, so erhalten wir für den bedingten Erwartungswert

$$\mathbb{E}(RV_{\text{all}}^{(n)} | X) = [X, X]_T^{(\text{all})} + 2n\mathbb{E}[\epsilon^2] \quad (3.3)$$

und falls Annahme 2.2.1 Punkt 3 ebenfalls gilt, erhalten wir für die Varianz

$$\text{Var}(RV_{\text{all}}^{(n)} | X) = 4n\mathbb{E}[\epsilon^4] + \mathcal{O}_p(1) \quad (3.4)$$

und für etwas stärkere Bedingungen

$$\text{Var}(RV_{\text{all}}^{(n)} | X) = 4n\mathbb{E}[\epsilon^4] + \left(8[X, X]_T^{(\text{all})}\mathbb{E}[\epsilon^2] - 2\text{Var}(\epsilon^2) \right) + \mathcal{O}_p(n^{-\frac{1}{2}}). \quad (3.5)$$

Die asymptotische Verteilung hat dann folgende Gestalt:

$$\frac{RV_{\text{all}}^{(n)} - 2n\mathbb{E}[\epsilon^2]}{\sqrt{2(\mathbb{E}[\epsilon^4])n}} \xrightarrow{D} N(0, 1), \quad \text{für } n \rightarrow \infty. \quad (3.6)$$

Bemerkung 3.1.2.²⁵ Wir sehen hier also, dass die Verzerrung im Fall von unabhängigem Microstructure Noise gegen unendlich divergiert.

²³vgl. [MS08]

²⁴vgl. [ZMAS05], S.1397

²⁵vgl. [HL06], S.134

3 Volatilitätschätzer

Ohne Microstructure Noise unter CTS ($\mathbb{E}[\epsilon^2] = 0$ und $\delta_n = \frac{b-a}{n}$) erhalten wir folgende Varianz

$$\text{Var}(RV_{\text{all}}^{(n)}) = 2 \sum_{i=1}^n \sigma_i^4 = 2 \frac{b-a}{n} \int_a^b \sigma_s^4 ds + o\left(\frac{1}{n}\right), \quad (3.7)$$

wobei $\int_a^b \sigma_s^4 ds$ die *integrierte Quartizität (IQ)* ist.

3.1.3 Schätzer für den Noiseprozess²⁶

Wir erkennen also, dass die $RV_{\text{all}}^{(n)}$ keinen verlässlichen Schätzer für die IV liefert. Die Realized Variance hat bei großem n nur noch wenig mit der wahren Volatilität zu tun. Stattdessen kann mit Hilfe des $RV_{\text{all}}^{(n)}$ Schätzers der Noiseterm $\mathbb{E}[\epsilon^2]$ geschätzt werden. Werden Hochfrequenzdaten verwendet, so liefert

$$\hat{\omega}^2 = \widehat{\mathbb{E}[\epsilon^2]} = \frac{1}{2n} RV_{\text{all}}^{(n)} \quad (3.8)$$

einen konsistenten Schätzer für die Varianz des Noiseterms. Die asymptotische Verteilung für einen festen Prozess X_t sieht dann folgendermaßen aus:

$$n^{\frac{1}{2}} (\hat{\omega}^2 - \omega^2) \xrightarrow{D} N(0, \mathbb{E}[\epsilon^4]) \quad \text{für } n \rightarrow \infty. \quad (3.9)$$

Auf dieselbe Art und Weise erhalten wir auch einen konsistenten Schätzer für die asymptotische Varianz von $\hat{\omega}^2$

$$\widehat{\mathbb{E}[\epsilon^4]} = \frac{1}{2n} \sum_i (\Delta Y_{t_i})^4 - 3 \left(\widehat{\mathbb{E}[\epsilon^2]} \right)^2. \quad (3.10)$$

Unter der Annahme eines unabhängigen Noiseprozess wissen wir, dass $\mathbb{E}([Y, Y]_T^{(n)}) = IV + 2n\mathbb{E}[\epsilon^2]$ gilt. Dabei ist zu beobachten, falls $\frac{IV}{2n}$ nicht vernachlässigbar ist, dass der Schätzer $\widehat{\mathbb{E}[\epsilon^2]}$ in der Praxis oft zur Überschätzung des Noiseprozesses ϵ_t führt. Verbesserungen des obigen Schätzers sind einerseits

$$\tilde{\omega}^2 = \frac{RV_{\text{all}}^{(n)} - RV_{\text{sparse}}^{(30\text{min})}}{2 \left(n - \frac{n}{1800} \right)} \quad (3.11)$$

und andererseits

$$\tilde{\omega}^2 = \frac{RV_{\text{all}}^{(n)} - \widehat{IV}}{2n}, \quad (3.12)$$

wobei $RV_{\text{sparse}}^{(30\text{min})}$ ein Schätzer mit Intervallfrequenz von 30 Minuten und \widehat{IV} ein unverzerrter Schätzer für die IV ist. Sowohl $\hat{\omega}^2, \tilde{\omega}^2$ als auch $\tilde{\omega}^2$ haben dieselbe asymptotische Verteilung. $\tilde{\omega}^2$ und $\tilde{\omega}^2$ sind unverzerrt, $\hat{\omega}^2$ jedoch kann in verschiedensten Fällen recht verzerrt sein.

²⁶vgl. [HL06], S.143

Bemerkung 3.1.3. Jetzt stellt sich die Frage, weshalb wir an der quadratischen Variation der unbeobachtbaren Preise X_t und nicht an der quadratischen Variation des beobachteten Prozess Y_t interessiert sind. Ein Hauptgrund dafür ist, dass ganz im Gegensatz zu dem Preisprozess des Underlyings, die Variation, verursacht durch den Noiseterm ϵ_t , eng verbunden mit jeder Transaktion ist. Aus Handelssicht stellen die ϵ_t 's Handelskosten dar, die im Gegensatz zu den Volatilitätskosten durch den zugrundeliegenden Basiswert verursacht sind. Verschiedene Marktteilnehmer zahlen unterschiedliche Handelskosten, abhängig von den Vereinbarungen der verschiedenen Institute. Des Weiteren wäre es schwierig in der Finanzwirtschaft die quadratische Variation von Y_t anzuwenden, da diese Größe sehr von der Datenfrequenz abhängt.

3.1.4 MSE²⁷

Als nächstes wollen wir den Mean Squared Error (MSE) betrachten und damit die optimale Samplingfrequenz bestimmen. Die grundlegende Idee des Subsamplings wird dann später in Abschnitt 3.3.1 beschrieben.

Proposition 3.1.4. Definieren wir $\lambda \equiv \frac{\omega^2}{IV}$ und nehmen wir $\kappa = 1$ an und seien t_0, \dots, t_n so, dass $\sigma_i^2 = \frac{IV}{n}$ (BTS), dann ist der MSE des $RV_{\text{all}}^{(n)}$ gegeben durch

$$MSE(RV_{\text{all}}^{(n)}) = IV^2 \left(4\lambda^2 n^2 + 12\lambda^2 n + 8\lambda - 4\lambda^2 + 2\frac{1}{n} \right). \quad (3.13)$$

Die optimale Samplingfrequenz des $RV^{(n)}$ ist dann die positive Lösung der Gleichung $4\lambda^2 n^3 + 6\lambda^2 n^2 - 1 = 0$. Wir erhalten

$$m_0^* \approx (2\lambda)^{-\frac{2}{3}}. \quad (3.14)$$

3.1.5 RV bei abhängigem Noiseprozess²⁸

Nun betrachten wir wieder den Standard RV Schätzer, diesmal bei abhängigem Noiseprozess. Falls die Beobachtungsfrequenz erhöht wird, erhalten wir dieselbe asymptotische Verteilung wie bei serienweise unkorrelierten Noisetermen. Wenn alle Beobachtungsdaten verwendet werden, hat das Noise sowohl auf die Verzerrung als auch die asymptotische Varianz des RV Schätzers Einfluss.

Der übliche RV Schätzer $RV_{\text{sparse}}^{(n)}$, berechnet mit der Stichprobenfrequenz $\Delta_{\text{sparse}} =$

²⁷vgl. [HL06], S.134

²⁸vgl. [ASMZ08], S.14

3 Volatilitätschätzer

$\frac{T}{n_{\text{sparse}}}$, verhält sich folgendermaßen:

$$RV_{\text{sparse}}^{(n)} \stackrel{\mathcal{D}}{\approx} \langle X, X \rangle_T + \underbrace{2n_{\text{sparse}}\mathbb{E}[\epsilon^2]}_{\text{bias due to noise}} + \underbrace{[4n_{\text{sparse}}\mathbb{E}[\epsilon^4] + \frac{2T}{n_{\text{sparse}}} \int_0^T \sigma_t^4 dt]}_{\text{total variance}} \frac{1}{n_{\text{sparse}}} N(0, 1).$$

Aus folgendem Grund ist der letzte Ausdruck genauso wie im iid Fall: Die asymptotische Varianz $\sum_i \epsilon_{t_i} \epsilon_{t_{i+J}}$ verhält sich genauso wie im unkorrelierten Fall, falls $J \rightarrow \infty$. Wenn $n_{\text{sparse}} \rightarrow \infty$ liegen die verwendeten logarithmischen Erträge Y weit genug voneinander entfernt, sodass die Abhängigkeit der ϵ_t 's keine wesentliche Rolle mehr spielt. In diesem Fall ist $\sum_i \epsilon_{t_i} (\epsilon_{t_{i+J}} - \mathbb{E}[\epsilon_{t_{i+J}} | \mathcal{F}_i])$ ein Martingal mit Varianz $\sum_i \mathbb{E} [\epsilon_{t_i}^2 (\epsilon_{t_{i+J}} - \mathbb{E}[\epsilon_{t_{i+J}} | \mathcal{F}_i])^2]$, was näherungsweise $n_{\text{sparse}} \mathbb{E}[\epsilon^2]^2$ ergibt, während das Restglied $\sum_i \epsilon_{t_i} (\epsilon_{t_{i+J}} - \mathbb{E}[\epsilon_{t_{i+J}} | \mathcal{F}_i])$ vernachlässigbar wird.

Falls alle Beobachtungen verwendet werden wie bei $RV_{\text{all}}^{(n)}$, ist die asymptotische Varianz des RV Schätzers durch die Abhängigkeit der Noiseterme beeinflusst. Für die asymptotische Varianz erhalten wir

$$\begin{aligned} \Omega_\infty &= AVAR \left[\sqrt{n} \left(\frac{[\epsilon, \epsilon]}{n} - 2\mathbb{E}\epsilon^2 \right) \right] \\ &= Var((\epsilon_1 - \epsilon_0)^2) + 2 \sum_{i=1}^{\infty} Cov((\epsilon_1 - \epsilon_0)^2, (\epsilon_{i+1} - \epsilon_i)^2). \end{aligned} \quad (3.15)$$

Falls wir nun beim RV Schätzer die höchste Samplingfrequenz $\Delta = \frac{T}{n}$ verwenden, verhält er sich wie folgt:

$$RV_{\text{all}}^{(n)} \stackrel{\mathcal{D}}{\approx} \langle X, X \rangle_T + \underbrace{2n(\mathbb{E}[\epsilon^2] + \mathbb{E}[\epsilon_{t_0} \epsilon_{t_1}])}_{\text{bias due to noise}} + \underbrace{[4n\Omega_\infty + \frac{2T}{n} \int_0^T \sigma_t^4 dt]}_{\text{total variance}} \frac{1}{n} N(0, 1),$$

wobei $\Omega_\infty = \mathbb{E}[\epsilon^4]$, im Fall von iid Noise. Falls der Noiseterm jedoch abhängig ist, erhalten wir (3.15).

Einen alternativen Ausdruck für Ω_∞ erhalten wir mit

$$\begin{aligned} Cov((\epsilon_1 - \epsilon_0)^2, (\epsilon_{i+1} - \epsilon_i)^2) &= 2Cov(\epsilon_1 - \epsilon_0, \epsilon_{i+1} - \epsilon_i)^2 \\ &+ Cum(\epsilon_1 - \epsilon_0, \epsilon_1 - \epsilon_0, \epsilon_{i+1} - \epsilon_i, \epsilon_{i+1} - \epsilon_i) \end{aligned}$$

und ähnlich für die Varianz.

3.2 Zhou Estimator²⁹

In diesem Abschnitt betrachten wir die erste Erweiterung des RV Estimators. Durch einfache Adaptionen der RV erhalten wir den Zhou Estimator, der durch

$$\begin{aligned} RV_{\text{Zhou}}^{(n)} &= \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 + \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-1}} - Y_{t_{i-2}}) \\ &\quad + \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i+1}} - Y_{t_i}) \end{aligned} \quad (3.16)$$

definiert ist.

Dieser Schätzer führt eine einfache Biaskorrektur durch, indem er die Autokovarianzen 1. Ordnung mit einbezieht. Es ist zu beachten, dass nun auch Werte außerhalb des Beobachtungsintervalls $[a, b]$ Einfluss auf den Schätzer haben. Im Speziellen können die Werte aus den Inkrementen $Y_{t_0} - Y_{t_{-1}}$ und $Y_{t_{n+1}} - Y_{t_n}$ nicht zur Verfügung stehen. Somit ist es notwendig, den Schätzer zu adaptieren:

$$\begin{aligned} \tilde{R}V_{\text{Zhou}}^{(n)} &= \sum_{i=1}^{n-1} (Y_{t_i} - Y_{t_{i-1}})^2 + \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-1}} - Y_{t_{i-2}}) \\ &\quad + \sum_{i=1}^{n-1} (Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i+1}} - Y_{t_i}). \end{aligned}$$

$\tilde{R}V_{\text{Zhou}}^{(n)}$ schätzt nun $\int_{a+\delta_1}^{b-\delta_m} \sigma_s ds = IV + \mathcal{O}(\frac{1}{m})$. Im Folgenden werden wir aber bei der Version aus (3.16) bleiben, da diese wesentlich einfacher handzuhaben ist.

3.2.1 Asymptotisches Verhalten

Das asymptotische Verhalten des Zhou Estimators kann durch folgende Proposition beschrieben werden:

Proposition 3.2.1. Unter den Annahmen, dass die logarithmischen Preise die stochastische Differentialgleichung (2.3) und der Noiseprozess die Annahme 2.2.1 Punkt 1 erfüllt, gilt

$$\mathbb{E}(RV_{\text{Zhou}}^{(n)}) = IV.$$

Falls Annahme 2.2.1 Punkt 2 ebenfalls gilt, ist die Varianz des Zhou Estimators gegeben durch

$$Var(RV_{\text{Zhou}}^{(n)}) = 8\omega^4 n + 8\omega^2 \sum_{i=1}^n \sigma_i^2 - 6\omega^4 + 6 \sum_{i=1}^m \sigma_i^4 + \mathcal{O}(\frac{1}{m^2}),$$

²⁹vgl. [Zho96], [HL06], S.134f

3 Volatilitätschätzer

und für die asymptotische Verteilung gilt

$$\frac{RV_{\text{Zhou}}^{(n)} - IV}{\sqrt{8\omega^4 n}} \xrightarrow{D} N(0, 1), \quad \text{für } n \rightarrow \infty.$$

Bemerkung 3.2.2. Ein wichtiges Resultat aus dieser Proposition ist, dass $RV_{\text{Zhou}}^{(n)}$ ein unverzerrter Schätzer der IV ist. Es ist des Weiteren zu beobachten, dass der $RV_{\text{Zhou}}^{(n)}$ im Gegensatz zum $RV_{\text{all}}^{(n)}$ eine kleinere asymptotische Varianz für $n \rightarrow \infty$ hat. Dieses Resultat ist doch ein wenig verwunderlich, da normalerweise eine Biaskorrektur die asymptotische Varianz erhöht. Ein weiterer Vorteil ist, dass in Proposition 3.2.1 die Bezugsgröße in der asymptotischen Verteilung die IV ist. Den Zhou Estimator mit der höchstmöglichen Frequenz laufen zu lassen, ist nicht zu empfehlen, da die asymptotisch Varianz mit $m!$ ansteigt. Daher können wir den IV Term streichen und erhalten $\frac{RV_{\text{Zhou}}^{(n)}}{\sqrt{8\omega^4 n}} \xrightarrow{D} N(0, 1)$, was bedeutet, dass für $n \rightarrow \infty$ der $RV_{\text{Zhou}}^{(n)}$ sich nur sehr unwahrscheinlich in der Nähe von der IV befinden wird.

Falls kein Microstructure Noise vorhanden ist ($\omega^2 = 0$), erhalten wir für die Varianz

$$\text{Var}(RV_{\text{Zhou}}^{(n)}) \approx 6 \sum_{i=1}^n \sigma_i^4,$$

was zeigt, dass die Varianz des $RV_{\text{Zhou}}^{(n)}$ rund drei mal so groß wie die des $RV^{(n)}$ ist. Bei Ausbleiben von Microstructure Noise, können wir also den erwarteten Anstieg der asymptotischen Varianz aufgrund der Biaskorrektur feststellen.

Es kann gezeigt werden, dass $\tau_i^* = \frac{c}{n}$, $i = 1, \dots, n$, das folgende Minimierungsproblem löst

$$\min_{\tau_1, \dots, \tau_n} \sum_{i=1}^n \tau_i^2, \quad \text{wobei} \quad \sum_{i=1}^n \tau_i = c.$$

Also, falls wir $\tau_i = \sigma_i^2$ und $c = IV$ setzen, sehen wir, dass $\sum_{i=1}^n \sigma_i^4$ (für festes n) unter BTS minimiert wird. Dies ist auch einer der wesentlichen Vorteile von BTS gegenüber CTS. In diesem Zusammenhang gilt unter BTS $\sum_{i=1}^n \sigma_i^4 = \frac{IV^2}{n}$, also gilt im Speziellen

$$\frac{IV^2}{n} \leq \int_a^b \sigma_s^4 ds \frac{b-a}{n}.$$

Die Varianz unter CTS ($\delta_n = \frac{b-a}{n}$) ist dann näherungsweise gegeben durch

$$\text{Var}(RV_{\text{Zhou}}^{(n)}) \approx 8\omega^4 n + 8\omega^2 \int_a^b \sigma_s^2 ds - 6\omega^4 + 6 \frac{b-a}{m} \int_a^b \sigma_s^4 ds.$$

3.2.2 MSE

Als Nächstes wollen wir den Mean Squared Error (MSE) betrachten und damit die optimale Beobachtungsfrequenz bestimmen. Dazu folgende Proposition:

Proposition 3.2.3. Definieren wir $\lambda \equiv \frac{\omega^2}{IV}$ und nehmen wir $\kappa = 1$ an und seien t_0, \dots, t_n sodass $\sigma_i^2 = \frac{IV}{n}$ (BTS), dann ist der MSE des $RV_{\text{Zhou}}^{(n)}$ gegeben durch

$$MSE(RV_{\text{Zhou}}^{(n)}) = IV^2 \left(8\lambda^2 n + 8\lambda - 6\lambda^2 + 6\frac{1}{n} - 2\frac{1}{n^2} \right).$$

Die optimale Samplingfrequenz des $RV_{\text{Zhou}}^{(n)}$ ist somit die positive Lösung der Gleichung $4\lambda^2 n^3 - 3n + 2 = 0$. Die optimale Frequenz liegt dann bei

$$m_1^* \approx \frac{\sqrt{3}}{2\lambda}.$$

In empirischen Analysen ist oft zu finden, dass $\lambda \leq 10^{-3}$ ist und dass

$$\frac{m_1^*}{m_0^*} \approx 3^{\frac{1}{2}} 2^{-\frac{1}{3}} (\lambda^{-1})^{\frac{1}{3}} \geq 10$$

ist.

m_1^* ist also um einige Male größer als m_0^* . In anderen Worten ist die optimale Beobachtungsfrequenz des $RV_{\text{Zhou}}^{(n)}$ wesentlich höher als die des $RV_{\text{sparse}}^{(n)}$.

3.3 Two Scaled Realized Volatility Estimator³⁰

Aufgrund der bereits vorgestellten Schätzer würde es nahe liegen zu vermuten, dass es von Vorteil ist, die Preise nicht allzu häufig zu beobachten. Aus statistischen Grundregeln ist jedoch klar, dass so viele Daten wie möglich verwendet werden sollten. Nun wollen wir in diesem Abschnitt einen Schätzer vorstellen, die auch bei der Verwendung aller Daten konsistent sind. Dazu verwenden wir zunächst Subsampling und Averaging, und in weiterer Folge eine Biaskorrektur. Die grundlegende Idee dahinter ist, den ursprünglichen Beobachtungszeitraum $\mathcal{G} = \{t_0, \dots, t_n\}$ in Teilnetze aufzuteilen, zum Beispiel wie folgt: $\mathcal{G}^{(1)}$ startet bei der ersten Beobachtung und nimmt Daten alle 5 Minuten auf, $\mathcal{G}^{(2)}$ startet dann bei der zweiten Beobachtung und nimmt Daten ebenfalls alle 5 Minuten auf und so weiter. Danach wird über alle Teilintervalle gemittelt. Der Vorteil liegt jetzt darin, dass der ganze Datensatz verwendet werden kann, aber die Variation des Schätzers vermindert wird. Die Vorteile von $RV_{\text{all}}^{(n)}$ und $RV_{\text{sparse}}^{(k)}$ werden also kombiniert.

³⁰vgl. [ZMAS05], S.1400ff, [ASMZ08]

3.3.1 Subsampling und Averaging

Dazu führen wir zunächst eine neue Notation ein. Wir teilen den Beobachtungszeitraum $\mathcal{G} = \{t_0, \dots, t_n\}$ in K nicht überlappende Teilnetze $\mathcal{G}^{(k)}, k = 1, \dots, K$ auf, sodass gilt

$$\mathcal{G} = \bigcup_{k=1}^K \mathcal{G}^{(k)}, \quad \text{wobei } \mathcal{G}^{(k)} \cap \mathcal{G}^{(l)} = \emptyset \text{ falls } k \neq l.$$

Die Wahl des k -ten Teilnetz $\mathcal{G}^{(k)}$ folgt meistens derart, indem mit t_{k-1} gestartet wird und dann jeder K -te Beobachtungszeitpunkt genommen wird, also

$$\mathcal{G}^{(k)} = \{t_{k-1}, t_{k-1+K}, t_{k-1+2K}, \dots, t_{k-1+n_k K}\}$$

für $k = 1, \dots, K$, wobei n_k jene Zahl ist, sodass $t_{k-1+n_k K}$ das letzte Element aus dem Teilnetz $\mathcal{G}^{(k)}$ ist, also $n_k = |\mathcal{G}^{(k)}|$. Dies ist die sogenannte *reguläre Allokation*. Die Mächtigkeit des gesamten Beobachtungsbereichs sei $n = |\mathcal{G}|$ und $RV_{\text{all}}^{(n)}$ sei die RV , bei der alle Beobachtungen verwendet werden. Falls jetzt nur die Beobachtungen des Teilnetzes $\mathcal{G}^{(k)}$ ausgewählt werden, also Y_t mit $t \in \mathcal{G}^{(k)}$, so bezeichnen wir die RV mit

$$RV_{\text{sparse}}^{(k)} = \sum_{t_i, t_{i,+} \in \mathcal{G}^{(k)}} (Y_{t_{i,+}} - Y_{t_i})^2,$$

wobei, falls $t_i \in \mathcal{G}^{(k)}$, so bezeichnet $t_{i,+}$ den nachfolgenden Zeitpunkt aus $\mathcal{G}^{(k)}$.

Eine erste leichte Verbesserung gegenüber $RV_{\text{all}}^{(n)}$, $RV_{\text{sparse}}^{(k)}$ und $RV_{\text{sparse}}^{(opt)}$ ist jetzt durch folgenden neuen Schätzer gegeben:

$$RV_{\text{avg}}^{(k)} = \frac{1}{K} \sum_{k=1}^K RV_{\text{sparse}}^{(k)}. \quad (3.17)$$

Dabei muss $\frac{n}{K} \rightarrow \infty$ für $n \rightarrow \infty$ gelten. Wir definieren

$$\bar{n} = \frac{1}{K} \sum_{k=1}^K n_k = \frac{n - K + 1}{K}.$$

Der Schätzer $RV_{\text{avg}}^{(k)}$ bleibt jedoch ein verzerrter Schätzer für die quadratische Variation $\langle X, X \rangle_T$. Die Verzerrung wächst nur noch mit dem Faktor \bar{n} statt n , also $2\bar{n}\mathbb{E}[\epsilon^2]$.

Fehler aufgrund von Microstructure Noise

Wie bereits erwähnt, werden die beobachteten Preise Y_t von einem Microstructure Noise ϵ_t gestört. Jetzt wollen wir diesen Fehlerterm $RV_{\text{avg}}^{(k)} - [X, X]_T^{(\text{avg})}$ genauer betrachten. Aus (3.3) und (3.17) wissen wir

$$\mathbb{E} \left(RV_{\text{avg}}^{(k)} | X \right) = [X, X]_T^{(\text{avg})} + 2\bar{n}\mathbb{E}[\epsilon^2]. \quad (3.18)$$

3 Volatilitätschätzer

Da die $\{\epsilon_t, t \in \mathcal{G}^{(k)}\}$ unabhängig für verschiedene k sind, gilt auch

$$\begin{aligned} \text{Var}\left(RV_{\text{avg}}^{(k)}|X\right) &= \frac{1}{K^2} \sum_{k=1}^K \text{Var}\left(RV_{\text{sparse}}^{(k)}|X\right) \\ &= 4\frac{\bar{n}}{K}\mathbb{E}[\epsilon^4] + \mathcal{O}_p\left(\frac{1}{K}\right) \end{aligned}$$

analog zu (3.4). Falls in der Varianz die Terme nächsthöherer Ordnung mit einbezogen werden, führt das zu

$$\text{Var}\left(RV_{\text{avg}}^{(k)}|X\right) = 4\frac{\bar{n}}{K}\mathbb{E}[\epsilon^4] + \frac{1}{K} \left[8[X, X]_T^{(\text{avg})} \mathbb{E}[\epsilon^2] - 2\text{Var}(\epsilon^2)\right] + o_p\left(\frac{1}{K}\right).$$

Nach [ZMAS05] Theorem A.1 haben wir dann folgende asymptotische Verteilung:

Proposition 3.3.1. Angenommen X_t erfüllt die SDE (2.3) und $Y_t = X_t + \epsilon_t$. Weiters erfüllt ϵ_t die Annahmen 2.2.1 Punkt 1 und 3. Setzen wir des Weiteren voraus, dass t_i und t_{i+1} für jedes i nicht im selben Teilnetz enthalten sind und gilt für $n \rightarrow \infty$ noch $\frac{\bar{n}}{K} \rightarrow \infty$, so gilt

$$\sqrt{\frac{k}{4\bar{n}\sqrt{\mathbb{E}[\epsilon^4]}}}\left(RV_{\text{avg}}^{(k)} - [X, X]_T^{(\text{avg})} + 2\bar{n}\mathbb{E}[\epsilon^2]\right) \xrightarrow{D} 2\sqrt{\mathbb{E}[\epsilon^4]}N(0, 1). \quad (3.19)$$

Der neue Schätzer $RV_{\text{avg}}^{(k)}$ verbessert im Vergleich zur RV das asymptotische Verhalten. Sowohl Verzerrung als auch Varianz sind kleiner.

Diskretisierungsfehler

In diesem Abschnitt untersuchen wir die Abweichung von $[X, X]_T^{(\text{avg})}$ zur IV . Dazu bezeichnen wir den Diskretisierungseffekt mit D_T , wobei

$$D_t = [X, X]_t^{(\text{avg})} - \langle X, X \rangle_t \quad (3.20)$$

$$= \frac{1}{K} \sum_{k=1}^K \left([X, X]_t^{(k)} - \langle X, X \rangle_t\right), \quad (3.21)$$

mit

$$[X, X]_t^{(k)} = \sum_{t_i, t_{i,+} \in \mathcal{G}^{(k)}, t_{i,+} \leq t} (X_{t_{i,+}} - X_{t_i})^2,$$

Jetzt wollen wir das asymptotische Verhalten von D_T genauer betrachten. Wir nehmen dabei an, dass die Preise zu regelmäßigen Abständen beobachtet werden, also $\mathcal{G}^{(l)} = \{t_{l-1}, t_{K+l-1}, \dots\}$. Zusätzlich nehmen wir an, dass

$$\max_i |\Delta t_i| = \mathcal{O}\left(\frac{1}{n}\right) \quad (3.22)$$

3 Volatilitätschätzer

und

$$\frac{K}{n} \rightarrow 0. \quad (3.23)$$

Dazu definieren wir noch eine Gewichtsfunktion

$$h_i = \frac{4}{K\Delta t} \sum_{j=1}^{(K-1)\wedge i} \left(1 - \frac{j}{K}\right)^2 \Delta t_{i-j}.$$

Falls die t_i äquidistant sind, also $\Delta t_i = \Delta t$, dann sind alle h_i , mit Ausnahme der ersten $K - 1$, gleich. Es gilt

$$h_i = \frac{4}{K} \sum_{j=1}^{(K-1)\wedge i} \left(1 - \frac{j}{K}\right)^2 = \begin{cases} 4\frac{2K^2-4K+3}{6K^2} = \frac{4}{3} + o(1), & \text{für } i \geq K - 1 \\ \frac{4i}{3K^2}(3K^2 - 3Ki + i^2) + o(1), & \text{für } i < K - 1. \end{cases}$$

Also versichern die Annahmen (3.22) und (3.23), dass

$$\sup_i h_i = \mathcal{O}(1).$$

Sei D_t der stetige Prozess aus (3.20), dann liefert die quadratische Variation $\langle D, D \rangle_T$ die beste Approximation für die Varianz von D_T .

Proposition 3.3.2. Angenommen X_t erfüllt die Annahme 2.1.9 und die Beobachtungszeiten weisen regelmäßige Abstände vor und (3.22) sowie (3.23) sind erfüllt, dann gilt folgende Approximation für die quadratische Variation von D_T

$$\langle D, D \rangle_T = \frac{TK}{n} \eta_n^2 + o_p\left(\frac{K}{n}\right),$$

wobei

$$\eta_n^2 = \sum_i h_i \sigma_{t_i}^4 \Delta t_i.$$

Im Speziellen gilt $D_T = \mathcal{O}_p\left(\frac{k}{n}^{\frac{1}{2}}\right)$. Weiters wollen wir uns nun noch das asymptotische Verhalten von D_T anschauen. Dazu folgende Proposition:

Proposition 3.3.3. Angenommen es gelten die Bedingungen aus Proposition 3.3.2 und zusätzlich, dass

$$\eta_n^2 \xrightarrow{P} \eta^2,$$

wobei η zufällig ist, dann gilt

$$\frac{D_T}{\sqrt{\frac{K}{n}}} \xrightarrow{D} \eta \sqrt{T} N(0, 1).$$

Falls die t_i äquidistant gewählt werden, gilt

$$\eta^2 = \frac{4}{3} \int_0^T \sigma_s^4 ds.$$

Gesamtfehler

Nun wollen wir die beiden Fehlerterme kombinieren. Aus Proposition 3.3.1 und 3.3.3 folgt

$$RV_{\text{avg}}^{(k)} - \langle X, X \rangle_T - 2\bar{n}\mathbb{E}[\epsilon^2] = \xi Z_{\text{total}} + o_p(1), \quad (3.24)$$

wobei Z_{total} asymptotisch standardnormalverteilt ist und unabhängig von X ist.

$$\xi^2 = \underbrace{4 \frac{\bar{n}}{K} \mathbb{E}[\epsilon^4]}_{\text{due to noise}} + \underbrace{t \frac{1}{T} \eta^2}_{\text{due to discretization}} \quad (3.25)$$

Wählen wir $K = cn^{\frac{2}{3}}$, dann werden beide Terme von ξ im Grenzwert gegenwärtig sein. Bei anderer Wahl von K wird einer der beiden Terme dominieren.

Bemerkung 3.3.4. $RV_{\text{avg}}^{(k)}$ ist aber leider noch immer ein verzerrter Schätzer für die IV. Die asymptotische Verzerrung verbessert sich zwar im Gegensatz zum $RV_{\text{all}}^{(n)}$, da sie nur noch mit dem Faktor $\bar{n} \leq n$ anwächst. Also müssen wir unseren Schätzer weiter adaptieren, um so zum TSRV zu gelangen. Davor wollen wir noch kurz eine Untersuchung der optimalen Samplingfrequenz durchführen.

Optimale Samplingfrequenz bei Subsampling

Bei der Suche nach dem optimalen \bar{n} ist es wichtig, eine gute Balance zwischen Verzerrung und Varianz zu finden. Um den MSE von $RV_{\text{avg}}^{(k)}$ zu minimieren, setzen wir $\frac{\partial MSE}{\partial \bar{n}} = 0$. Aus (3.24)-(3.25)folgt

$$\text{Verzerrung} = 2\bar{n}\mathbb{E}[\epsilon^2]$$

und

$$\xi^2 = 4 \frac{\bar{n}}{K} \mathbb{E}[\epsilon^4] + \frac{T}{\bar{n}} \eta^2.$$

Dann gilt

$$\begin{aligned} MSE &= \text{Verzerrung}^2 + \xi^2 = 4 (\mathbb{E}[\epsilon^2])^2 \bar{n}^2 + 4 \frac{\bar{n}}{K} \mathbb{E}[\epsilon^4] + \frac{T}{\bar{n}} \eta^2 \\ &= 4 (\mathbb{E}[\epsilon^2])^2 \bar{n}^2 + \frac{T}{\bar{n}} \eta^2. \end{aligned}$$

Demnach erfüllt das optimale \bar{n}^*

$$\bar{n}^* = \left(\frac{t\eta^2}{8(\mathbb{E}[\epsilon^2])^2} \right)^{\frac{1}{3}}.$$

Es können also alle n Beobachtungen verwendet werden, falls $K^* \approx \frac{n}{\bar{n}^*}$ Teilnetze verwendet werden. Das ist auch der große Vorteil des $RV_{\text{avg}}^{(k)}$. Für asymptotische Konvergenz ist allerdings notwendig, dass $\mathbb{E}[\epsilon^2] \rightarrow 0$ konvergiert.

3.3.2 Definition TSRV

Nun wollen wir mit Hilfe einer Biaskorrektur einen ersten unverzerrten Schätzer vorstellen. Wie bereits in Abschnitt 3.1.3 gezeigt wurde, kann $\mathbb{E}[\epsilon^2]$ durch $RV_{\text{all}}^{(n)}$ approximiert werden:

$$\hat{\omega}^2 = \widehat{\mathbb{E}[\epsilon^2]} = \frac{1}{2n} RV_{\text{all}}^{(n)}.$$

Dann kann ein biasangepasster TSRV Schätzer für $\langle X, X \rangle_T$ wie folgt konstruiert werden:

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} = \underbrace{RV_{\text{avg}}^{(k)}}_{\text{slow time scale}} - \frac{\bar{n}}{n} \underbrace{RV_{\text{all}}^{(n)}}_{\text{fast time scale}}. \quad (3.26)$$

3.3.3 Asymptotisches Verhalten

Um das asymptotische Verhalten von $\widehat{\langle X, X \rangle}_T^{(\text{tsrv})}$ zu betrachten, stellen wir zuerst fest, dass unter den Bedingungen von Proposition 3.3.1

$$\begin{aligned} \left(\frac{K}{\bar{n}}\right)^{\frac{1}{2}} \left(\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} - [X, X]_T^{(\text{avg})} \right) &= \left(\frac{K}{\bar{n}}\right)^{\frac{1}{2}} \left(RV_{\text{avg}}^{(k)} - [X, X]_T^{(\text{avg})} - 2\bar{n}\mathbb{E}[\epsilon^2] \right) \\ &\quad - 2(K\bar{n})^{\frac{1}{2}} \left(\widehat{\mathbb{E}[\epsilon^2]} - \mathbb{E}[\epsilon^2] \right) \xrightarrow{D} N(0, 8(\mathbb{E}[\epsilon^2])) \end{aligned}$$

gilt. Falls wir nun dieses Resultat mit denen aus 3.3.1 kombinieren, können wir ein optimales K für $n \rightarrow \infty$ bestimmen.

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} - IV = \left(\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} - [X, X]_T^{(\text{avg})} \right) + \left([X, X]_T^{(\text{avg})} - IV \right) \quad (3.27)$$

$$= \mathcal{O}_p \left(\frac{\bar{n}^{\frac{1}{2}}}{K^{\frac{1}{2}}} \right) + \mathcal{O}_p \left(\bar{n}^{-\frac{1}{2}} \right). \quad (3.28)$$

Unser optimales K bestimmen wir, indem wir die zwei Terme auf der rechten Seite minimieren. Dadurch erhalten wir die optimale Schrittweite für $RV_{\text{avg}}^{(k)}$ $K = \mathcal{O} \left(n^{\frac{2}{3}} \right)$. Die rechte Seite von (3.27) ist dann von der Ordnung $\mathcal{O}_p \left(n^{-\frac{1}{6}} \right)$. Im Speziellen erhalten wir also

$$K^* = cn^{\frac{2}{3}} \quad (3.29)$$

und damit folgende asymptotische Verteilung bei optimalen K :

Proposition 3.3.5. Angenommen X_t erfüllt die SDE (2.3) und $Y_t = X_t + \epsilon_t$ und ϵ_t erfüllt die Annahme 2.2.1 Punkt 1 und 2, dann gilt unter Annahme (3.29)

$$\begin{aligned} n^{\frac{1}{6}} \left(\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} - IV \right) &\xrightarrow{D} N(0, 8c^{-2}(\mathbb{E}[\epsilon^2])^2) + \eta\sqrt{T}N(0, c) \\ &= (8c^{-2}(\mathbb{E}[\epsilon^2])^2 + c\eta^2 T)^{\frac{1}{2}} N(0, 1). \end{aligned}$$

Die Konstante c hat dabei den Zweck, die gesamte asymptotische Varianz zu minimieren.

3.3.4 Optimales Sampling und Biaskorrektur

Um die Anzahl der Teilnetze K so gering wie möglich zu halten, minimieren wir die asymptotische Varianz und erhalten so ein optimales

$$c^* = \left(\frac{16(\mathbb{E}[\epsilon^2])^2}{T\mathbb{E}[\eta^2]} \right)^{\frac{1}{3}}.$$

Diese Konstante c^* kann berechnet werden, indem Schätzer für $\mathbb{E}[\epsilon^2]$ und η^2 verwendet werden.

Eine weitere kleine Verbesserung des TSRV kann noch durch folgende Biasanpassung erhalten werden:

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})} = \left(1 - \frac{\bar{n}}{n}\right)^{-1} \widehat{\langle X, X \rangle}_T^{(\text{tsrv})}. \quad (3.30)$$

Der Unterschied zwischen (3.26) und (3.30) ist von der Ordnung $\mathcal{O}(\frac{\bar{n}}{n}) = \mathcal{O}(K^{-1})$, also weisen beide dasselbe asymptotische Verhalten auf. $\widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})}$ hat den besonderen Anreiz, dass er unverzerrt bzgl. höherer Ordnungen ist. Für beliebige (a, b) betrachten wir nun Schätzer der Form

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})} = aRV_{\text{avg}}^{(k)} - b\frac{\bar{n}}{n}RV_{\text{all}}^{(n)}.$$

Aus (3.3) und (3.18) folgt dann

$$\begin{aligned} \mathbb{E} \left(\widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})} | X \right) &= a \left([X, X]_T^{(\text{avg})} + 2\bar{n}\mathbb{E}[\epsilon^2] \right) - b\frac{\bar{n}}{n} \left([X, X]_T^{(\text{all})} + 2n\mathbb{E}[\epsilon^2] \right) \\ &= a[X, X]_T^{(\text{avg})} - b\frac{\bar{n}}{n}[X, X]_T^{(\text{all})} + 2(a-b)\bar{n}\mathbb{E}[\epsilon^2]. \end{aligned} \quad (3.31)$$

Falls $a = b$ gewählt wird, kann der Effekt von $\mathbb{E}[\epsilon^2]$ komplett entfernt werden. Um den angepassten Schätzer aus (3.30) zu erhalten, muss $a = \frac{1}{1-\frac{\bar{n}}{n}}$ gewählt werden. Auf ähnliche Weise erhalten wir einen angepassten Schätzer für $\mathbb{E}[\epsilon^2]$

$$\widehat{\mathbb{E}[\epsilon^2]}^{(\text{adj})} = \frac{1}{2}(n - \bar{n})^{-1} \left(RV_{\text{all}}^{(n)} - RV_{\text{avg}}^{(k)} \right).$$

Dieser Schätzer erfüllt

$$\mathbb{E} \left(\widehat{\mathbb{E}[\epsilon^2]}^{(\text{adj})} | X \right) = \mathbb{E}[\epsilon^2] + \frac{1}{2}(n - \bar{n})^{-1} \left([X, X]_T^{(\text{all})} - [X, X]_T^{(\text{avg})} \right)$$

und ist daher unverzerrt höherer Ordnung. Für die asymptotische Verteilung erhalten wir

$$\begin{aligned} \widehat{\mathbb{E}[\epsilon^2]}^{(\text{adj})} - \mathbb{E}[\epsilon^2] &= \left(\widehat{\mathbb{E}[\epsilon^2]} - \mathbb{E}[\epsilon^2] \right) (1 + \mathcal{O}(K^{-1})) + \mathcal{O}_p \left(Kn^{-\frac{3}{2}} \right) \\ &= \widehat{\mathbb{E}[\epsilon^2]} - \mathbb{E}[\epsilon^2] + \mathcal{O}_p \left(n^{-\frac{1}{2}}K^{-1} \right) + \mathcal{O}_p \left(Kn^{-\frac{3}{2}} \right) \\ &= \widehat{\mathbb{E}[\epsilon^2]} - \mathbb{E}[\epsilon^2] + \mathcal{O}_p \left(n^{-\frac{5}{6}} \right). \end{aligned}$$

Es folgt, dass $n^{\frac{1}{2}} \left(\widehat{\mathbb{E}[\epsilon^2]}^{(\text{adj})} - \mathbb{E}[\epsilon^2] \right)$ und $n^{\frac{1}{2}} \left(\widehat{\mathbb{E}[\epsilon^2]} - \mathbb{E}[\epsilon^2] \right)$ die gleiche asymptotische Verteilung haben.

3.3.5 Erweiterung des TSRV bei abhängigem Noiseprozess

In der Praxis gibt es zahlreiche Situationen, bei denen die Annahme von iid Noise problematisch ist. Aus diesem Grund versuchen wir nun den TSRV Schätzer so zu verbessern, dass dieser auch bei abhängigen Noisetermen stabil bleibt. Die Idee dahinter ist die Frequenz der Fast Time Scale zu vermindern, um so den Abhängigkeitsgrad zu reduzieren.

Das Setup

Wie bisher seien Y_t die logarithmischen Transaktionspreise, die zu den Zeiten $0 = t_0, t_1, \dots, t_n = T$ beobachtet werden. Die beobachteten logarithmischen Preise Y_t setzen sich wieder aus den unbeobachtbaren logarithmischen Preisen X_t und einem Noiseterm ϵ_t zusammen

$$Y_t = X_t + \epsilon_t.$$

Annahme 3.3.6. Der Noiseprozess ϵ_t ist unabhängig von X_t und stationär. Weiters setzen wir voraus, dass es ein $\kappa > 0$ mit $\mathbb{E}[\epsilon^{4+\kappa}] < \infty$ gibt.

Dann gibt es eine Konstante $\rho < 1$, sodass für alle i

$$|\text{Cov}(\epsilon_{t_i}, \epsilon_{t_{i+l}})| \leq \rho^l \text{Var}(\epsilon)$$

gilt.

Wir betrachten wieder Teilnetze und zwar diesmal die Realized Variance basierend auf jeder j -ten Beobachtung und starten dabei bei der r -ten Beobachtung. Dadurch ergibt sich folgender Schätzer:

$$[Y, Y]_T^{(j,r)} = \sum_{0 \leq j(i-1) \leq n-r-j} (Y_{t_{j+i+r}} - Y_{t_{j(i-1)+r}})^2$$

Unter den meisten Annahmen verletzt dieser Schätzer das „sufficiency principle“, daher definieren wir die *average lag j realized volatility* als

$$\begin{aligned} [Y, Y]_T^{(j)} &= \frac{1}{j} \sum_{r=0}^{j-1} [Y, Y]_T^{(j,r)} \\ &= \frac{1}{j} \sum_{i=0}^{n-j} (Y_{t_{i+j}} - Y_{t_i})^2. \end{aligned}$$

3 Volatilitätschätzer

So können wir den TSRV durch Kombination von J und K , wobei $1 \leq J < K \leq n$, zu folgendem Schätzer verallgemeinern:

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} = \underbrace{[Y, Y]_T^{(K)}}_{\text{slow time scale}} - \frac{\bar{n}_K}{\bar{n}_J} \underbrace{[Y, Y]_T^{(J)}}_{\text{fast time scale}}, \quad (3.32)$$

mit $\bar{n}_K = \frac{(n-K+1)}{K}$ und $\bar{n}_J = \frac{(n-J+1)}{J}$.

Wir werden diesen Schätzer weiterhin mit TSRV bezeichnen. Der in [ZMAS05] vorgestellte TSRV Schätzer ist ein Spezialfall des obigen und zwar für $J = 1$ und $K \rightarrow \infty$ für $n \rightarrow \infty$. Der ursprüngliche TSRV Schätzer ist bei iid Noise stabil. Bei optimaler Auswahl von $K = \mathcal{O}(n^{\frac{2}{3}})$ erhalten wir folgenden Fehler:

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} - IV = \mathcal{O}_p(n^{-\frac{1}{6}}).$$

Bemerkung 3.3.7. Die iid Annahme an den Noiseprozess ist zu einschränkend und mit den in der Praxis herrschenden Bedingungen in der Regel nicht zu vereinbaren. Wählen wir aber $J > 1$, so kann das Problem der abhängigen ϵ_{t_i} 's sofort vernachlässigt werden, da der modifizierte TSRV Schätzer bei geeigneter Wahl von (J, K) konsistent ist.

Eine Signal-Noise Zerlegung

Lemma 3.3.8. Unter der Annahme von oben, sei $n \rightarrow \infty$ und $j = j_n$ eine Folge, dann gilt

$$\sum_{i=0}^{n-j} (X_{t_{i+j}} - X_{t_i})(\epsilon_{t_{i+j}} - \epsilon_{t_i}) = \mathcal{O}_p(j^{\frac{1}{2}}).$$

Dieses Lemma ist wichtig, da es zu folgender Darstellung führt

$$[Y, Y]_T^{(J)} = [X, X]_T^{(J)} + [\epsilon, \epsilon]_T^{(J)} + \mathcal{O}_p(J^{-\frac{1}{2}})$$

und mit Hilfe von Linearkombinationen erhalten wir

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv})} = \underbrace{[X, X]_T^{(K)} + \frac{\bar{n}_K}{\bar{n}_J} [X, X]_T^{(J)}}_{\text{signal term}} + \underbrace{[\epsilon, \epsilon]_T^{(K)} - \frac{\bar{n}_K}{\bar{n}_J} [\epsilon, \epsilon]_T^{(J)}}_{\text{noise term}} + \mathcal{O}_p(K^{-\frac{1}{2}})$$

solange

$$1 \leq J \leq K \text{ und } K = o(n).$$

Analyse des Noiseterms

Falls die ϵ_t 's unabhängig sind, kann gezeigt werden, dass $\mathbb{E}[\text{noise term}] = 0$ gilt. Dann ist die Linearkombination in (3.32) das was wir benötigen, um die Verzerrung aufgrund des Noiseprozess verschwinden zu lassen. Um einen allgemeineren Fall zu betrachten und so die näherungsweise Verteilung des Noiseterms zu erhalten, benötigen wir folgendes:

$$[\epsilon, \epsilon]_T^{(J)} = \frac{1}{J} \sum_{i=0}^n c_i^{(J)} \epsilon_{t_i}^2 - \frac{2}{J} \sum_{i=0}^{n-J} \epsilon_{t_i} \epsilon_{t_{i+J}},$$

wobei $c_i^{(J)} = 2$ für $J \leq i \leq n - J$ und $c_i^{(J)} = 1$ für andere i . Nach Konstruktion gilt

$$\sum_i c_i^{(J)} = 2J\bar{n}_J,$$

sodass für $J \leq \frac{n}{2}$

$$\begin{aligned} |\mathbb{E}[\text{noise term}]| &\leq 2\mathbb{E}[\epsilon^2] \left(\frac{1}{K}(n - K + 1)\rho^K + \frac{\bar{n}_K}{\bar{n}_J} \frac{1}{J}(n - J + 1)\rho^J \right) \\ &= \mathcal{O} \left(\frac{n}{K}(\rho^K + \rho^J) \right) \end{aligned}$$

und im Fall $Cov(\epsilon_{t_0}, \epsilon_{t_K}) = o(Cov(\epsilon_{t_0}, \epsilon_{t_J}))$,

$$\mathbb{E}[\text{noise term}] = 2\frac{n}{K}Cov(\epsilon_{t_0}, \epsilon_{t_J})(1 + o(1)).$$

Falls $J \rightarrow \infty$ und $n \rightarrow \infty$, wenn auch nur sehr langsam, ist die Verzerrung unwesentlich. Genauso wie im Fall von „m-dependent“ ϵ_t 's, wird die Verzerrung Null für endliches J . Wir erhalten folgende Proposition:

Proposition 3.3.9. Unter der Annahme $\limsup_{n \rightarrow \infty} \frac{J}{K} < 1$ und

$$\frac{K}{n^{\frac{1}{2}}}(\text{noise term} - \mathbb{E}[\text{noise term}]) \xrightarrow{D} \xi N(0, 1) \quad \text{für } n \rightarrow \infty.$$

Weiters falls $J, K \rightarrow \infty$ mit $n \rightarrow \infty$, haben wir $\xi^2 = \xi_\infty^2$, wobei

$$\xi_\infty^2 = 8Var(\epsilon)^2 + 16 \sum_{i=1}^{\infty} Cov(\epsilon_{t_0}, \epsilon_{t_i})^2.$$

Falls J nicht gegen unendlich strebt (m-dependent case), dann gilt

$$\xi^2 = \xi_\infty^2 + 4\alpha_0 + 8 \sum_{i=1}^{\infty} \alpha_i,$$

wobei

$$\alpha_i = Cov(\epsilon_{t_0}, \epsilon_{t_{i+J}})Cov(\epsilon_{t_i}, \epsilon_{t_J}) + Cum(\epsilon_{t_0}, \epsilon_{t_i}, \epsilon_{t_J}, \epsilon_{t_{i+J}}).$$

Im Fall $J \rightarrow \infty$ kann es ebenfalls besser sein ξ^2 statt ξ_∞^2 zu verwenden, wenn J nur langsam gegen unendlich strebt. Im Gegensatz dazu strebt K wesentlich schneller gegen unendlich.

Analyse des Signalterms

Als Signalterm erhalten wir folgenden:

$$[X, X]_T^{(K)} \xrightarrow{P} \langle X, X \rangle_T \quad \text{für } n \rightarrow \infty,$$

für geeignetes $K = o(n)$. Damit der Signalterm in $\widehat{\langle X, X \rangle}_T - \langle X, X \rangle_T$ konsistent ist, benötigen wir

$$\limsup_{n \rightarrow \infty} \frac{J}{K} < 1. \quad (3.33)$$

Diese Anforderung ist auch relativ einfach erfüllt ist. Daher nehmen wir normalerweise die stärkere Anforderung

$$\limsup_{n \rightarrow \infty} \frac{J}{K} = 0 \quad (3.34)$$

an. Wir erhalten folgende Proposition:

Proposition 3.3.10. Unter $1 \leq J \leq K$ und $K = o(n)$ gilt

$$\left(\frac{K}{n} \left(1 + 2 \frac{J^3}{K^3} \right) \right)^{-\frac{1}{2}} \left([X, X]_T^{(K)} - \frac{\bar{n}_K}{\bar{n}_J} [X, X]_T^{(J)} - \langle X, X \rangle_T \right) \xrightarrow{D} \eta \sqrt{T} N(0, 1),$$

wobei η^2 ist gegeben durch Theorem 3 in [ZMAS05]. Im Spezialfall von äquidistanten Beobachtungen erhalten wir

$$\eta^2 = \frac{4}{3} \int_0^T \sigma_t^4 dt.$$

Die Konvergenz in der Verteilung ist stabil und daher ist die normalverteilte Zufallsvariable unabhängig von η .

Ein kombinierter Schätzer

Wir betrachten den angepassten Schätzer

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})} = \left(1 - \frac{\bar{n}_K}{\bar{n}_J} \right)^{-1} \widehat{\langle X, X \rangle}_T^{(\text{tsrv})}.$$

Im iid Fall wurde diese Anpassung eingeführt, um bei kleiner Anzahl an Beobachtungen ein besseres Resultat erzielen zu können. Hier benötigen wir diese Anpassung für Konsistenzvoraussetzungen; (3.33) ist erfüllt, (3.34) jedoch nicht.

Wir erhalten aus Proposition 3.3.9 und 3.3.10 folgende Darstellung:

$$\begin{aligned} \widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})} &= \left(2\mathbb{E}[\epsilon^2] \frac{n}{K} \text{Cov}(\epsilon_{t_0}, \epsilon_{t_J}) + \frac{n^{\frac{1}{2}}}{K} \xi Z_{\text{noise}} + \left(\frac{K}{n} \left(1 + 2 \frac{J^3}{K^3} \right) \right)^{\frac{1}{2}} \eta \sqrt{T} Z_{\text{discrete}} \right) \\ &\quad \left(1 - \frac{\bar{n}_K}{\bar{n}_J} \right)^{-1} (1 + o_p(1)), \end{aligned}$$

3 Volatilitätschätzer

wobei Z_{noise} und Z_{discrete} asymptotisch normalverteilt und unabhängig sind. Die optimale Abstimmung zwischen den zwei Varianztermen wird bei $K = \mathcal{O}(m^{\frac{2}{3}})$ erhalten. Die schlechteste Verzerrung ist von der Ordnung $(\frac{n}{K})\rho^J = n^{\frac{1}{3}}\rho^J$. Also ist die Verzerrung relativ klein im Verhältnis zur Varianz. Daher wählen wir $n^{\frac{1}{3}}\rho^J = o(n^{-\frac{1}{6}})$. Dann kann gefahrlos $\frac{J}{K} \sim 0$ angenommen werden und es folgt:

Proposition 3.3.11. Das asymptotische Verhalten des Schätzers $\widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})}$ ist gegeben durch

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})} = \langle X, X \rangle_T + \left(2\mathbb{E}[\epsilon^2] \frac{n}{K} \text{Cov}(\epsilon_{t_0}, \epsilon_{t_J}) + \frac{n^{\frac{1}{2}}}{K} \xi Z_{\text{noise}} + \left(\frac{K}{n} \right)^{\frac{1}{2}} \eta \sqrt{T} Z_{\text{discrete}} \right) (1 + o_p(1)).$$

Das optimale K ist oben gegeben, für das optimale J wählen wir

$$\text{Cov}(\epsilon_{t_0}, \epsilon_{t_J}) = o(n^{-\frac{1}{2}}).$$

Falls ϵ_t abhängig der Ordnung m ist, wählen wir einfach $J = m + 1$. Im Fall von asymptotischer Varianz gibt es kein eindeutiges optimales J . Um die asymptotische Varianz zu minimieren, sollte J zwischen $\text{Cov}(\epsilon_{t_0}, \epsilon_{t_J}) = o(n^{-\frac{1}{2}})$ und $\limsup_{n \rightarrow \infty} \frac{J}{K} = 0$ liegen.

Eine weitere Anpassung an den TSRV

Falls K groß ist, unterschätzt $[X, X]_T^{(K)}$ die quadratische Variation $\langle X, X \rangle_T$ leicht. Dafür betrachten wir $\sigma_t^2 = \sigma^2$ und erhalten so $\langle X, X \rangle_T = \sigma^2 T$, wobei $[X, X]_T^{(K)} \approx \sigma^2 T \frac{(n-K+1)}{n}$. Also

$$\begin{aligned} [X, X]_T^{(K)} - \frac{\bar{n}_K}{\bar{n}_J} [X, X]_T^{(J)} &\approx \sigma^2 T \left(\frac{n-K+1}{n} - \frac{\bar{n}_K}{\bar{n}_J} \frac{n-J+1}{n} \right) \\ &= \sigma^2 T \frac{(K-J)\bar{n}_K}{n}. \end{aligned}$$

Eine weitere Adaption führt zum *Area Adjusted TSRV*

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv, aa})} = \frac{n}{(K-J)\bar{n}_K} \widehat{\langle X, X \rangle}_T^{(\text{tsrv})}.$$

Wegen

$$\frac{n}{(K-J)\bar{n}_K} \sim \left(1 - \frac{\bar{n}_K}{\bar{n}_J} \right)^{-1}$$

erhalten wir:

Proposition 3.3.12. Der Schätzer $\widehat{\langle X, X \rangle}_T^{(\text{tsrv, aa})}$ hat das gleiche asymptotische Verhalten wie $\widehat{\langle X, X \rangle}_T^{(\text{tsrv, adj})}$ aus Proposition 3.3.11.

Bemerkung 3.3.13. Meistens wird dieser adjustierte TSRV Schätzer empfohlen, da er einerseits dasselbe asymptotische Verhalten besitzt, andererseits aber Vorteile bei kleineren Stichproben hat, wobei hier bei der Biaskalkulation von konstanten σ ausgegangen wird.

Zusammenfassend verhalten sich beide TSRV Schätzer bei abhängigem Noiseprozess stabil und wie folgt:

$$\widehat{\langle X, X \rangle}_T^{(\text{tsrv, aa})} \stackrel{D}{\approx} \langle X, X \rangle_T + \frac{1}{n^{\frac{1}{6}}} \underbrace{\left[\underbrace{\frac{1}{c^2} \xi^2}_{\text{due to noise}} + \underbrace{c \frac{4T}{3} \int_0^T \sigma_t^4 dt}_{\text{due to discretization}} \right]}_{\text{total variance}}^{\frac{1}{2}} N(0, 1).$$

$K \sim cn^{\frac{2}{3}}$ und ξ gegeben wie in Proposition 3.3.9.

Bemerkung 3.3.14. Nun stellt sich die Frage, wenn zwei Skalen besser als eine sind, wie die Performance dann mit mehreren Skalen aussieht. Die Erweiterung von zwei auf mehrere Skalen wird Multi Scale Realized Volatility (MSRV) genannt. Dieser hat den Vorteil einer besseren Konvergenzrate, nämlich $n^{-\frac{1}{4}}$ im Gegensatz zu $n^{-\frac{1}{6}}$ beim TSRV. Sowohl TSRV als auch MSRV sind eng mit autokovarianzbasierten Schätzern verbunden, wie den Kernel Schätzern.

3.4 Multi Scales Realized Volatility Estimator³¹

3.4.1 Motivation des MSRV Estimators

Wie wir im letzten Abschnitt gesehen haben, ist der TSRV Schätzer ein konsistenter, asymptotisch stabiler Schätzer der quadratischen Variation $\langle X, X \rangle_T$, der auch beim Auftauchen von abhängigen Noisetermen funktioniert und eine Konvergenzrate von $n^{-\frac{1}{6}}$ hat. Jetzt ist es aber natürlich auch möglich, anstatt über zwei Skalen gleich über mehrere Skalen zu mitteln. Auf Kosten höherer Komplexität erhalten wir auf diese Weise den sogenannten Multi Scale Realized Volatility Estimator (MSRV). Diesen wollen wir nun herleiten. Dazu benötigen wir folgende Definitionen:

$$[X, \epsilon]^{(n, K)} = \frac{1}{K} \sum_{t_i \geq K} (X_{t_i} - X_{t_{i-K}})(\epsilon_{t_i} - \epsilon_{t_{i-K}}),$$

und

$$[\epsilon, \epsilon]^{(n, K)} = \frac{1}{K} \sum_{t_i \geq K} (\epsilon_{t_i} - \epsilon_{t_{i-K}})^2.$$

³¹vgl. [Zha06]

3 Volatilitätschätzer

Für $[Y, Y]^{(n, K)}$ erhalten wir durch Aufspalten

$$[Y, Y]^{(n, K)} = [X, X]^{(n, K)} + [\epsilon, \epsilon]^{(n, K)} + 2[X, \epsilon]^{(n, K)}.$$

Definition 3.4.1. Der *Multi Scale Realized Volatility Estimator* hat nun folgende Gestalt

$$\widehat{\langle X, X \rangle}_T^{(\text{msrv})} = \sum_{i=1}^M \alpha_i [Y, Y]_T^{(K_i)},$$

wobei a_i Gewichte sind.

Eine erste natürliche Anforderung erhalten wir durch folgende Beobachtung:

$$\mathbb{E}(\widehat{\langle X, X \rangle}_T^{(\text{msrv})} | X) = \sum_{i=1}^M \alpha_i [Y, Y]_T^{(K_i)} + 2\mathbb{E}[\epsilon^2] \sum_{i=1}^M \alpha_i \frac{n+1-K_i}{K_i}.$$

Da die $[X, X]^{(n, K_i)}$ ein asymptotisch unverzerrter Schätzer für IV ist, ist es logisch

$$\sum_{i=1}^M \alpha_i = 1 \quad \text{und} \quad \sum_{i=1}^M \alpha_i \frac{n+1-K_i}{K_i} = 0$$

zu fordern.

Eine einfache Adaption macht das Problem offensichtlicher:

$$a_1 = \alpha_1 - \left[(n+1) \left(\frac{1}{K_1} - \frac{1}{K_2} \right) \right]^{-1}, \quad a_2 = \alpha_2 - (a_1 - \alpha_1) \quad \text{und} \quad a_i = \alpha_i \quad \text{für} \quad i \geq 3.$$

Dann sind die obigen Bedingungen äquivalent zu:

Bedingung 3.4.2.

$$\sum_{i=1}^M a_i = 1,$$

und

Bedingung 3.4.3.

$$\sum_{i=1}^M \frac{a_i}{K_i} = 0.$$

Um die Eigenschaften des MSRv Schätzers und die Bedingungen an die a_i besser zu verstehen, betrachten wir zunächst folgende Proposition. Dabei sind a_i, K_i und M von n abhängig.

3 Volatilitätschätzer

Proposition 3.4.4. Angenommen K_1 und K_2 sind von der Ordnung $\mathcal{O}(1)$ für $n \rightarrow \infty$. Unter den Annahmen 2.1.9, sowie 2.2.1 Punkt 1 und 3 gilt

$$\widehat{\langle X, X \rangle}_T^{(\text{msrv})} = \sum_{i=1}^M a_i [Y, Y]^{(n, K_i)} - 2\mathbb{E}[\epsilon^2] + \mathcal{O}(n^{-\frac{1}{2}}).$$

In weiterer Folge schreiben wir für

$$[Y, Y]^{(n, K)} = [X, X]^{(n, K)} + \frac{2}{K} \sum_{i=0}^n \epsilon_{t_i}^2 + U_{n, K} + V_{n, K},$$

wobei der Term

$$U_{n, K} = -\frac{2}{K} \sum_{i=K}^n \epsilon_{t_i} \epsilon_{t_{i-K}}$$

hauptverantwortlich für den Fehler ist und

$$V_{n, K} = 2[X, \epsilon]^{(n, K)} - \frac{1}{K} \sum_{i=0}^{K-1} \epsilon_{t_i}^2 - \frac{1}{K} \sum_{i=n-K+1}^n \epsilon_{t_i}^2$$

das Restglied ist. Nun können wir den Einfluss von Bedingung (3.4.3) betrachten:

$$\begin{aligned} \widehat{\langle X, X \rangle}_T^{(\text{msrv})} &= \underbrace{\sum_{i=1}^M a_i [X, X]_T^{(n, K_i)}}_{\text{signal}} + \underbrace{\sum_{i=1}^M a_i U_{n, K_i}}_{\text{noise}} + \underbrace{2 \sum_{i=1}^M \frac{a_i}{K_i} \sum_{j=0}^n \epsilon_{t_j}^2}_{=0} + \underbrace{\sum_{i=1}^M a_i V_{n, K_i} - 2\mathbb{E}[\epsilon]^2}_{\text{end points of noise}} \\ &+ \mathcal{O}_p(n^{-\frac{1}{2}}) = \sum_{i=1}^M a_i [X, X]_T^{(n, K_i)} + \sum_{i=1}^M a_i U_{n, K_i} + R_n + \mathcal{O}_p(n^{-\frac{1}{2}}), \end{aligned} \quad (3.35)$$

wobei R_n das gesamte Restglied

$$R_n = \sum_{i=1}^M a_i V_{n, K_i} - 2\mathbb{E}[\epsilon^2]$$

ist. Dabei beseitigt Bedingung 3.4.3 den Biasterm bedingt durch $\sum_{i=1}^M \epsilon_{t_i}$ nicht nur in Erwartung, sondern auch fast sicher. Aus diesem Grund benötigen wir auch die Annahme, dass die ϵ_{t_i} 's iid sind. Dann liefert uns der MSRV einen gegenüber Microstructure Noise durchaus robusten Schätzer. Bedingung 3.4.2 sichert uns dabei, dass wir einen asymptotisch unverzerrten Schätzer für die IV erhalten.

Weiters sind U_{n, K_i} und U_{n, K_l} für $i \neq l$ unkorreliert. Falls wir Bedingung 3.4.3 voraussetzen und R_n vergleichsweise klein annehmen, dann reicht es

$$[X, X]^{(K_i)} + U_{n, K_i}, \quad \text{für } i = 1, \dots, M$$

zu beobachten. Unter der idealen Vorstellung von stetigen Beobachten macht es Bedingung 3.4.3 möglich, M verschiedene Maße für die IV zu erhalten. Dies begründet auch die Form des MSR_V Estimator. Unser Ziel ist es nun, die Gewichte a_i möglichst optimal zu bestimmen.

3.4.2 Asymptotisches Verhalten des Noiseterms und Bestimmung der Gewichte

Um die optimalen Gewichte a_i zu bestimmen, müssen wir zuerst das asymptotische Verhalten des Noiseterms analysieren. Dabei betrachten wir zunächst folgenden Term:

$$\zeta_n = \sum_{i=1}^{M_n} a_{n,i} U_{n,K_{n,i}}.$$

Dann können wir für die Varianz von ζ_n folgende Aussage treffen:

Proposition 3.4.5. Sei $\gamma_n^2 = 4 \sum_{i=1}^{M_n} (\frac{a_{n,i}}{K_{n,i}})^2$. Angenommen die ϵ_{t_i} sind iid mit Erwartungswert Null und $\mathbb{E}[\epsilon^2] < \infty$ und es gilt $M_n = o(n)$ für $n \rightarrow \infty$, dann gilt

$$\text{Var}(\zeta_n) = \gamma_n^2 n (\mathbb{E}[\epsilon^2])^2 (1 + o(1)).$$

Sind Bedingungen 3.4.2 und (3.4.3) erfüllt, dann wird γ_n^2 durch die geeignete Wahl der

$$a_{n,i} = \frac{K_{n,i}(K_{n,i} - \bar{K}_n)}{M_n \text{Var}(K_n)}$$

minimiert, wobei $\bar{K}_n = \frac{1}{M_n} \sum_{i=1}^{M_n} K_{n,i}$ und $\text{Var}(K_n) = \frac{1}{M_n} \sum_{i=1}^{M_n} K_{n,i}^2 - (\frac{1}{M_n} \sum_{i=1}^{M_n} K_{n,i})^2$.

Daraus ergibt sich der folgende minimale Wert von γ_n :

$$\gamma_n^{*2} = \frac{4}{M_n \text{Var}(K_n)}.$$

Diese Eigenschaft wird noch genauer durch folgende Proposition beschrieben:

Proposition 3.4.6. Angenommen die ϵ_{t_i} sind iid mit Erwartungswert Null und $\mathbb{E}[\epsilon^2] < \infty$ und es gilt $M_n = o(n)$ für $n \rightarrow \infty$. Des Weiteren nehmen wir an, dass $\max_{1 \leq i \leq M_n} |\frac{a_{n,i}}{i \gamma_n}| \rightarrow 0$ für $n \rightarrow \infty$. Dann gilt

$$\frac{\zeta_n}{n^{\frac{1}{2}} \gamma_n} \xrightarrow{D} N(0, \mathbb{E}[\epsilon^2]^2).$$

Darauf aufbauend entwickeln wir eine Gewichtsklasse, welche wir dann in dem restlichen Abschnitt verwenden werden. Wir nehmen dabei an, dass alle Skalen $i = 1, \dots, M$ verwendet werden, also $K_{n,i} = i$. In diesem Fall ist $\bar{K}_n = \frac{M_n+1}{2}$ und $\text{Var}(K_n) = \frac{M_n^2-1}{12}$. Die optimalen Gewichte aus Proposition 3.4.5 sind gegeben durch

$$a_{n,i} = 12 \frac{i}{M_n^2} \frac{\left(\frac{i}{M_n} - \frac{1}{2} - \frac{1}{2M_n}\right)}{\left(1 - \frac{1}{M_n^2}\right)}. \quad (3.36)$$

3 Volatilitätschätzer

Das Minimum der γ_n ist dann

$$\gamma_n^{*2} = \frac{48}{M_n(M_n^2 - 1)}$$

und somit gilt für die minimale Varianz

$$\text{Var}(\zeta_n) = \frac{48n(\mathbb{E}[\epsilon^2])^2}{M_n(M_n^2 - 1)}. \quad (3.37)$$

Die Gleichung aus (3.36) motiviert uns, Gewichte folgender Form zu betrachten

$$a_{n,i} = \frac{1}{M_n} w_{M_n} \left(\frac{i}{M_n} \right), \quad i = 1, \dots, M_n, \quad (3.38)$$

mit

$$w_M(x) = xh(x) + M^{-1}xh_1(x) + M^{-2}xh_2(x) + M^{-3}xh_3(x) + o(M^{-3}), \quad (3.39)$$

wobei h und h_1 Funktionen unabhängig von M sind. Der Grund, weshalb diese spezielle Funktion $w_M(x)$ verwendet wird, ist das $w_M(x)$ bei Null verschwinden muss und aufgrund von Bedingung 3.4.3, $\int_0^1 \frac{w_M(x)}{x} dx$ ungefähr Null sein sollte. Also müssen wir infolge von der Bedingungen 3.4.2 und 3.4.3 folgende Anforderungen an die Funktion h stellen:

Bedingung 3.4.7.

$$\int_0^1 xh(x)dx = 1,$$

und

Bedingung 3.4.8.

$$\int_0^1 h(x)dx = 0.$$

Für etwas stärkere Annahmen an h kann gezeigt werden, dass (3.35) auch etwas allgemeiner gilt:

Proposition 3.4.9. Sei $h_0 = h$ und nehmen wir an, dass für $i = 0, \dots, 2$, h_i $(3 - i)$ mal stetig differenzierbar auf $[0, 1]$ ist, und h_3 stetig auf $[0, 1]$ ist. Des Weiteren nehmen wir an, dass die Bedingungen 3.4.7 und 3.4.8 erfüllt sind und folgende Gleichungen Gültigkeit haben:

$$\begin{aligned} \int_0^1 h_1(x)dx + \frac{1}{2}(h(1) - h(0)) &= 0, \\ \int_0^1 h_2(x)dx + \frac{1}{2}(h(1) - h(0)) + \frac{1}{12}(h'(1) - h'(0)) &= 0, \\ \int_0^1 h_3(x)dx + \frac{1}{12}(h'(1) - h'(0)) &= 0. \end{aligned} \quad (3.40)$$

Seien $a_{n,i}$ gegeben durch (3.38)-(3.39) und nehmen wir noch an, dass die ϵ_{t_i} iid mit $\mathbb{E}[\epsilon^2] < \infty$ sind. Dann bleibt die Approximation (3.35) gültig, bis zu $o_p\left(\frac{n}{M_n^3}\right)$.

Bemerkung 3.4.10. Falls in (3.39) weitere Terme hinzugefügt werden, kann die Approximation in (3.35) weiter verbessert werden. Wir werden in der Folge $M_n = \mathcal{O}(n^{\frac{1}{2}})$ verwenden, weshalb wir auch den Ausdruck in (3.39) gewählt haben. Weiters sollte noch erwähnt werden, dass die Approximation in Bedingung 3.4.3 feiner als die in Bedingung 3.4.2 sein muss, da $\sum_{i=1}^M \frac{a_i}{K_i} \sum_{j=0}^n \epsilon_{t_j}^2 = n \left(\sum_{i=1}^M \frac{a_i}{K_i} \right) \mathbb{E}[\epsilon^2](1 + o_p(1))$ vernachlässigbar für asymptotische Untersuchungen werden soll.

Die spezielle Wahl von h_1, h_2 und h_3 spielt dabei keine wesentliche Rolle für das Verhalten der asymptotischen Varianz. Eine einfache Wahl von h_1 , die (3.40) erfüllt, ist gegeben durch $h_1(x) = -\frac{h'(x)}{2}$ mit Konstanten $h_2(x) = h_2$ und $h_3(x) = h_3$. In diesem Fall ist $h_2 = -\frac{h_1'(1) - h_1''(0)}{6}$ und $h_3 = \frac{h_1''(1) - h_1'''(0)}{24}$. Dadurch erhalten wir für die Gewichte

$$a_{n,i} = \frac{i}{M_n^2} h\left(\frac{i}{M_n}\right) - \frac{1}{2} \frac{i}{M_n^3} h'\left(\frac{i}{M_n}\right) + \frac{i}{M_n^3} h_2 + \frac{i}{M_n^4} h_3. \quad (3.41)$$

Für die optimalen Gewichte in (3.36) wählen wir

$$h_\zeta^* = 12 \left(x - \frac{1}{2} \right). \quad (3.42)$$

Bei dieser Wahl von h sind die $a_{n,i}$ gegeben durch (3.41) bis auf einen vernachlässigbaren multiplikativen Faktor von $((1 - M_n^{-2})^{-1})$ identisch zu denen aus (3.37).

Folgende Proposition ist nun eine Erweiterung von Proposition 3.4.6, da

$$\gamma_n^2 = 4M_n^{-3} \int_0^1 h(x)^2 dx (1 + o(1)) \quad \text{für } n \rightarrow \infty.$$

Proposition 3.4.11. Angenommen die ϵ_{t_i} sind iid mit Erwartungswert Null und $\mathbb{E}[\epsilon^2] < \infty$ und es gilt $M_n = o(n)$ für $n \rightarrow \infty$. Nehmen wir zusätzlich an, dass die $a_{n,i}$ durch (3.38) gegeben sind und die Bedingungen aus Proposition 3.4.9 erfüllt sind, dann gilt

$$\left(\frac{M_n^3}{n} \right)^{\frac{1}{2}} \zeta_n \xrightarrow{D} N(0, 4(\mathbb{E}[\epsilon^2])^2 \int_0^1 h(x)^2 dx).$$

3.4.3 Diskretisierungsfehler

Wenn wir Darstellung (3.35) betrachten, treten zwei Fehlerterme auf. Einerseits der Diskretisierungsfehler, da die Beobachtungen nur an diskreten Zeitpunkten auftreten und andererseits das Residuum R_n . Dies wollen wir nun untersuchen, um dann eine vernünftige Aussage über das gesamte asymptotische Verhalten des MSRV Schätzers treffen zu können.

Für den Diskretisierungsfehler benötigen wir folgende Resultate:

3 Volatilitätschätzer

Definition 3.4.12. Seien die Beobachtungszeiten $0 = t_0 < t_1 < \dots < t_n = T$ nicht zufällig und \mathcal{G}_n ein Teilnetz, dann ist die *Asymptotic Quadratic Variation of Time* (AQVT) definiert durch

$$H(t) = \lim_{n \rightarrow \infty} \frac{n}{T} \sum_{t_{i+1} \leq t} (t_i - t_{i-1})^2,$$

vorausgesetzt der Grenzwert existiert.

Zusätzlich nehmen wir an, dass

$$\max_{1 \leq i \leq n} |t_{i+1} - t_i| = \mathcal{O}\left(\frac{1}{n}\right). \quad (3.43)$$

Aufgrund dieser Annahme ist $H(t)$ Lipschitz-stetig. Es gilt dabei folgende Regel für die AQVT:

Proposition 3.4.13. Angenommen es gilt (3.43) und die AQVT $H(t)$ existiert. Sei $G : [0, T] \rightarrow [0, T]$ Lipschitz-stetig und setzen wir $u_i = G(t_i)$, dann existiert

$$K(u) = \lim_{n \rightarrow \infty} \frac{n}{T} \sum_{u_i \leq u} (u_i - u_{i-1})^2$$

und es gilt

$$H'(t)G'(t) = K'(G(t))$$

fast überall auf $[0, T]$.

Weiters benötigen wir noch eine weitere Proposition:

Proposition 3.4.14. Angenommen es gelten die Bedingungen aus Proposition 3.4.13, dann gilt $K(T) = T$ genau dann, wenn

$$\sum_{i=0}^n \left(u_i - u_{i-1} - \frac{T}{n}\right)^2 = o(n^{-1}).$$

Bemerkung 3.4.15. Mit Hilfe dieser zwei Propositionen ist es nun möglich unregelmäßige und fast äquidistante Stichproben zu vergleichen. Falls $H'(t)$ existiert, stetig ist und nach unten durch eine Konstante $c > 0$ beschränkt ist, dann können wir $G(t) = \int_0^t H'(s)^{-1} ds$ definieren und den adaptierten Prozess $\tilde{X}_u = X_{G(u)}$ betrachten, wobei nun die Zeitpunkte $u_i = G(t_i)$ nahezu äquidistant sind.

Wir definieren η als die Wurzel von

$$\eta^2 = \int_0^T H'(t) \sigma_t^4 dt.$$

Nun können wir das Hauptresultat für das Verhalten des Diskretisierungsfehlers von $[X, X]^{(K)}$ wie folgt formulieren:

Proposition 3.4.16. Angenommen der Preisprozess X_t erfüllt (2.3) und die Beobachtungszeiten t_i sind nicht zufällig und genügen (3.43). Weiters existiert die AQVT von $H(t)$ und ist stetig differenzierbar. Dazu nehmen wir an, dass $\min_{0 \leq t \leq T} H'(t) > 0$ ist und $M_n \rightarrow \infty$ für $n \rightarrow \infty$, mit $M_n = o(n)$. Sei $\frac{(K_1, \dots, K_L)}{M_n} \rightarrow (\kappa_1, \dots, \kappa_L)$ für $n \rightarrow \infty$. Sei Γ eine $L \times L$ Matrix mit Einträgen

$$\Gamma_{I,J} = \frac{2}{3}T \min(\kappa_I, \kappa_J) \left(3 - \frac{\min(\kappa_I, \kappa_J)}{\max(\kappa_I, \kappa_J)} \right),$$

und Z ein normalverteilter Zufallsvektor mit Kovarianzmatrix Γ , unabhängig von X , dann gilt

$$\left(\frac{n}{M_n} \right)^{\frac{1}{2}} \left([X, X]^{(K_1)} - \langle X, X \rangle, \dots, [X, X]^{(K_L)} - \langle X, X \rangle \right) \xrightarrow{D} \eta Z.$$

Es lässt sich für den Diskretisierungsfehler des MSR_V Schätzers nun folgende Aussage treffen:

Proposition 3.4.17. Angenommen die a_i erfüllen (3.38)-(3.39) und seien die Bedingungen von Proposition 3.4.9 erfüllt. Weiters sei angenommen, dass der Preisprozess X_t die SDE (2.3) erfüllt und die Beobachtungszeiten t_i nicht zufällig sind und (3.43) genügen. Darüber hinaus existiert die AQVT von $H(t)$ und ist stetig differenzierbar. Dazu nehmen wir an, dass $\min_{0 \leq t \leq T} H'(t) > 0$ ist und $M_n \rightarrow \infty$ für $n \rightarrow \infty$, mit $M_n = o(n)$ und $\frac{M_n^3}{n} \rightarrow \infty$. Setzen wir

$$\eta_h^2 = \frac{4}{3}T\eta^2 \int_0^1 dx \int_0^x h(y)h(x)y^2(3x-y)dy,$$

dann gilt

$$\left(\frac{n}{M_n} \right)^{\frac{1}{2}} \left(\sum_{i=1}^{M_n} a_{n,i} [X, X]^{(n,i)} - \langle X, X \rangle \right) \xrightarrow{D} \eta_h N(0, 1).$$

3.4.4 Asymptotisches Verhalten des MSR_V

Bei der Berechnung des MSR_V Schätzers treten zwei Hauptfehlerquellen auf. Einerseits haben wir in Proposition 3.4.11 gesehen, dass falls wir M_n Zeitskalen verwenden, die durch Microstructure Noise ϵ_t induzierte Differenz $\widehat{\langle X, X \rangle}_T^{(\text{msrv})} - IV$, auf einen Term der Ordnung $\mathcal{O}_p(n^{\frac{1}{2}} M_n^{-\frac{3}{2}})$ reduziert werden kann. Und andererseits zeigt Proposition 3.4.17, dass der Diskretisierungsfehler von der Ordnung $\mathcal{O}_p(n^{-\frac{1}{2}} M_n^{\frac{1}{2}})$ ist. Nun ist es die Aufgabe, diese zwei Fehlerterme abzustimmen. Das optimale M_n dafür ist von der Ordnung $M_n = \mathcal{O}(n^{\frac{1}{2}})$. Dies führt zu einer Varianz-Varianz Abstimmung und die Konvergenzrate des MSR_V ist dann $\widehat{\langle X, X \rangle}_T^{(\text{msrv})} - IV = \mathcal{O}_p(n^{-\frac{1}{4}})$. Dieses Resultat stellt eine Verbesserung gegenüber dem TSR_V Estimator ($\mathcal{O}_p(n^{-\frac{1}{6}})$) dar. Das Konvergenzverhalten des MSR_V kann durch folgende Proposition beschrieben werden:

Proposition 3.4.18. Angenommen die $a_{n,i}$ erfüllen (3.38)-(3.39) und die Bedingungen aus Proposition 3.4.9 sind erfüllt, sowie die Annahmen 2.1.9 und 2.2.1 Punkt 1 und 3 und wir nehmen an, dass die Beobachtungen $t_{n,i}$ nicht zufällig sind und $\max_{1 \leq i \leq n} |t_{n,i+1} - t_{n,i}| = \mathcal{O}(\frac{1}{n})$ genügen. Des Weiteren existiert die asymptotische quadratische Variation zur Zeit $h(t)$ und ist stetig differenzierbar. Darüber hinaus nehmen wir an, dass $\min_{0 \leq t \leq T} H'(t) > 0$ und dass $\frac{M_n}{n^{\frac{1}{2}}} \rightarrow c$ für $n \rightarrow \infty$. Sei dann Z eine standardnormalverteilte Zufallsvariable unabhängig von X und setzen wir

$$\begin{aligned} \nu_h^2 &= 4c^{-3}(\mathbb{E}[\epsilon^2])^2 \int_0^1 h(x)^2 dx + c \frac{4}{3} T \eta^2 \int_0^1 dx \int_0^x h(y)h(x)y^2(3x-y)dy \\ &+ 4c^{-1}Var(\epsilon^2) \int_0^1 \int_0^y xh(x)h(y)dx dy + 8c^{-1}\mathbb{E}[\epsilon^2] \int_0^1 \int_0^1 h(x)h(y) \min(x,y)dx dy \langle X, X \rangle, \end{aligned}$$

dann gilt

$$n^{\frac{1}{4}} \left(\widehat{\langle X, X \rangle}_T^{(msrv)} - IV \right) \xrightarrow{D} \nu_h Z, \quad \text{für } n \rightarrow \infty.$$

Für die optimale h -Funktion aus Gleichung (3.42), können wir nun die asymptotische Varianz des MSRV berechnen. Falls $h(x) = 12(x - \frac{1}{2})$, erhalten wir

$$\begin{aligned} \int_0^1 dx \int_0^x h(y)h(x)y^2(3x-y)dy &= \frac{39}{35}, \\ \int_0^1 \int_0^y xh(x)h(y)dx dy &= \frac{3}{5}, \\ \int_0^1 \int_0^1 h(x)h(y) \min(x,y)dx dy &= \frac{6}{5}. \end{aligned}$$

Die asymptotische Varianz ist in diesem Fall gegeben durch

$$\nu_h^2 = 48c^{-3}(\mathbb{E}[\epsilon^2])^2 + \frac{52}{35}cT\eta^2 + \frac{12}{5}c^{-1}Var(\epsilon^2) + \frac{48}{5}c^{-1}\mathbb{E}[\epsilon^2]IV.$$

3.5 Realized Kernel Estimator³²

In diesem Abschnitt beschäftigen wir uns mit dem Realized Kernel Estimator. Genauer wollen wir uns mit der Entwicklung eines robusten Schätzers, der auch bei Auftreten von Market Microstructure Noise effizient und stabil bleibt, beschäftigen. Dabei ist es möglich durch die Wahl verschiedener Kernel, den Schätzer möglichst gut an die Bedingungen anzupassen.

³²vgl. [BNHLS08]

3.5.1 Annahmen an den Microstructure Noiseprozess

Wir gehen in diesem Abschnitt zu Beginn wieder davon aus, dass sich die beobachteten Preise Y_t aus einem Prozess X_t und einem Noiseprozess ϵ_t zusammensetzen. Also $Y_t = X_t + \epsilon_t$. Für die grundlegenden Überlegungen nehmen wir dabei an, dass die Prozesse X_t und ϵ_t unabhängig voneinander sind. In weiterer Folge wird diese Forderung dann gelockert.

Zusätzlich fordern wir, dass der Prozess ϵ_t ein sogenanntes *White Noise* ist ($\epsilon_t \in \mathcal{WN}$). Es gilt also

$$\mathbb{E}(\epsilon_t) = 0, \quad \text{Var}(\epsilon_t) = \omega^2, \quad \text{Var}(\epsilon_t^2) = \lambda^2 \omega^4, \quad \text{und} \quad \epsilon_t \perp \epsilon_s$$

für alle $t, s, \lambda \in \mathbb{R}^+$. Die White Noise Annahme ist aus verschiedensten Gründen nicht zufriedenstellend. Eine Eigenschaft eines Prozesses $\epsilon_t \in \mathcal{WN}$ ist das $[\epsilon]_t = \infty$ und somit $\epsilon_t \notin \mathcal{SM}$. Aus diesem Grund gäbe es in Märkten ohne Microstructure Noise Arbitragemöglichkeiten. Die White Noise Annahme macht daher nur Sinn, falls auch wirklich Marktunstimmigkeiten vorhanden sind. Später werden wir diese Annahme fallen lassen, um einen konsistenten Kernel Estimator zu erhalten.

3.5.2 Definition Realized Kernel

Definition 3.5.1. Erfassen wir Erträge über Zeitintervalle der Länge δ , dann definieren wir für zwei Prozesse X_t und Z_t

$$\gamma_h(Z_\delta, X_\delta) = \sum_{j=1}^n (Z_{\delta j} - Z_{\delta(j-1)})(X_{\delta(j-h)} - X_{\delta(j-h-1)}),$$

mit $h = -H, \dots, -1, 0, 1, \dots, H$. Dann heißt $\gamma_h(X_\delta) = \gamma_h(X_\delta, X_\delta)$ der *Realized Autocovariance Process*.

Dabei gilt

$$\gamma_h(Y_\delta) = \gamma_h(X_\delta) + \gamma_h(X_\delta, \epsilon_\delta) + \gamma_h(\epsilon_\delta, X_\delta) + \gamma_h(\epsilon_\delta).$$

$\gamma_0(Y_\delta)$ ist dabei die Realized Variance von Y_t , die Quadratwurzel die Realized Volatility.

Wir betrachten den Zeitraum $[0, t]$ und einen stetigen logarithmischen Preisprozess Y_t und Beobachtungen im Zeitintervall $\delta > 0$, dann hat der *Realized Kernel Estimator* folgende Gestalt:

$$K(Y_\delta) = \gamma_0(Y_\delta) + \sum_{h=1}^H k\left(\frac{h-1}{H}\right) [\gamma_h(Y_\delta) + \gamma_{-h}(Y_\delta)],$$

3 Volatilitätschätzer

wobei $k(x)$ für $x \in [0, 1]$ eine nichtstochastische Gewichtsfunktion ist. Die h -te Autokovarianzfunktion ist gegeben durch

$$\gamma_h(Y_\delta) = \sum_{j=1}^n (Y_{\delta j} - Y_{\delta(j-1)})(Y_{\delta(j-h)} - Y_{\delta(j-h-1)}),$$

mit $h = -H, \dots, -1, 0, 1, \dots, H$ und $n = \lfloor \frac{t}{\delta} \rfloor$. Für kleines δ ist $Y_{\delta j} - Y_{\delta(j-1)}$ der j -te High Frequency Return, $K(Y_\delta) - \gamma_0(Y_\delta)$ ist dabei die Realized Kernel Korrektur der Realized Variance, verursacht durch Market Microstructure Noise.

Wir verwenden hier Gewichtsfunktionen $k(x)$, die auf $[0, 1]$ zweimal differenzierbar sind und definieren zur Vereinfachung:

$$k_{\bullet}^{0,0} = \int_0^1 k(x)^2 dx, \quad k_{\bullet}^{1,1} = \int_0^1 k'(x)^2 dx, \quad k_{\bullet}^{2,2} = \int_0^1 k''(x)^2 dx.$$

Von besonders spezieller Bedeutung sind Kernel, die $k'(0)^2 + k'(1)^2 = 0$ erfüllen, die sogenannte Klasse der *Smooth Kernels*.

3.5.3 Asymptotisches Verhalten

In diesem Unterabschnitt wollen wir das asymptotische Verhalten des Kernel Estimators genauer betrachten. Hierfür benötigen wir folgende Propositionen:

Proposition 3.5.2. Für $n, H \rightarrow \infty$ und $\frac{H}{n} \rightarrow 0$ gilt

$$\begin{aligned} \sqrt{\frac{n}{H}} [K(X_\delta) - IV] &\xrightarrow{D} MN \left(0, 4k_{\bullet}^{0,0} t \int_0^t \sigma_u^4 du \right), \\ \sqrt{H} [K(X_\delta, \epsilon_\delta) + K(\epsilon_\delta, X_\delta)] &\xrightarrow{D} MN \left(0, k_{\bullet}^{1,1} 8\omega^2 \int_0^t \sigma_u^2 du \right), \\ \sqrt{\frac{H^2}{n}} [(K(\epsilon_\delta))] &\xrightarrow{D} N(0, 4\omega^4 [k'(0)^2 + k'(1)^2]). \end{aligned}$$

Falls $k'(0)^2 + k'(1)^2 = 0$, ist die asymptotische Varianz von $K(\epsilon_\delta)$ wie folgt gegeben

$$4\omega^4 \left(\left(\frac{n}{H^3} \right) k_{\bullet}^{2,2} + \left(\frac{1}{Hm} \right) k_{\bullet}^{1,1} \right)$$

und

$$\sqrt{\frac{H^3}{n}} [(K(\epsilon_\delta))] \xrightarrow{D} N(0, 4\omega^4 k_{\bullet}^{2,2}) \quad \text{falls} \quad \frac{H^2}{mn} \rightarrow 0.$$

Weiters ist es sinnvoll eine Noise-to-Signal Kennzahl

$$\xi^2 = \frac{\omega^2}{\sqrt{t \int_0^t \sigma_u^4 du}}$$

3 Volatilitätschätzer

und ein Heteroskedastizitätsmaß

$$\rho = \frac{\int_0^t \sigma_u^2 du}{\sqrt{t \int_0^t \sigma_u^4 du}} = \frac{IV}{\sqrt{tIQ}}$$

einzuführen. Falls $\rho = 1$ ist, ist die Volatilität konstant und mittels Cauchy-Schwarz Ungleichung erhalten wir $\rho \leq 1$.

Für großes n und H ist die asymptotische Varianz von $K(Y_\delta) - IV$ gegeben durch

$$4t \int_0^t \sigma_u^4 du \left[\frac{H}{n} k_{\bullet}^{0,0} + 2 \frac{k_{\bullet}^{1,1}}{H} \rho \xi^2 + n \left(\frac{k'(0)^2 + k'(1)^2}{H^2} + \frac{k_{\bullet}^{2,2}}{H^3} \right) \xi^4 + \frac{k_{\bullet}^{1,1}}{Hm} \xi^4 \right].$$

Wir sehen der Fall $h'(0)^2 + k'(1)^2 = 0$ ist ein wichtiger Spezialfall, der in folgender Proposition verdeutlicht wird:

Proposition 3.5.3. Für $H = c_0 n^{\frac{2}{3}}$, haben wir

$$n^{\frac{1}{6}} [K(Y_\delta) - IV] \xrightarrow{D} MN \left(0, 4t \int_0^t \sigma_u^4 du [c_0 k_{\bullet}^{0,0} + c_0^{-2} (k'(0)^2 + k'(1)^2) \xi^4] \right). \quad (3.44)$$

Falls $h'(0)^2 + k'(1)^2 = 0$, $m \rightarrow \infty$ und $H = c_0 n^{\frac{1}{2}}$, erhalten wir

$$n^{\frac{1}{4}} [K(Y_\delta) - IV] \xrightarrow{D} MN \left(0, 4t \int_0^t \sigma_u^4 du [c_0 k_{\bullet}^{0,0} + c_0^{-1} 2k_{\bullet}^{1,1} \rho \xi^2 + c_0^{-3} k_{\bullet}^{2,2} \xi^4] \right). \quad (3.45)$$

Dieses Resultat ist besonders interessant, da die auf diese Weise erhaltene Konvergenzrate die bestmögliche ist. Die Forderung $m \rightarrow \infty$ wird aufgrund von Endeffekten benötigt. Falls m fix gewählt werden würde, erhalten wir einen zusätzlichen Term in der asymptotischen Varianz, dessen Beitrag proportional zu $\frac{\xi^2}{m}$ ist. Falls ξ^2 ziemlich klein ist, etwa 10^{-3} , kann der letzte Term auch im Fall $m = 1$ ignoriert werden.

Spezialfälle mit Konvergenzrate $n^{-\frac{1}{6}}$

Für $H = c (\xi^2 n)^{\frac{2}{3}}$ erhalten wir die asymptotische Verteilung (3.44), indem wir $c_0 = c \xi^{\frac{4}{3}}$ setzen. Für diese Kernelklasse ist

$$c^* = \left[\frac{2(k'(0)^2 + k'(1)^2)}{k_{\bullet}^{0,0}} \right]^{\frac{1}{3}}$$

jener Wert c , der die asymptotische Varianz minimiert. Die untere Grenze der asymptotischen Varianz ist dann durch

$$4c^* \omega^{\frac{4}{3}} \left(t \int_0^t \sigma_u^4 du \right)^{\frac{2}{3}} \left[k_{\bullet}^{0,0} + c^{*-3} (k'(0)^2 + k'(1)^2) \right] = 6c^* k_{\bullet}^{0,0} \omega^{\frac{4}{3}} \left(t \int_0^t \sigma_u^4 du \right)^{\frac{2}{3}}$$

3 Volatilitätschätzer

gegeben. Mit Hilfe von $c^* k_{\bullet}^{0,0}$ kann die asymptotische Effizienz des Schätzers kontrolliert werden.

	$k(x)$	$k'(0)$	$k'(1)$	$k_{\bullet}^{0,0}$	$k_{\bullet}^{1,1}$	$k_{\bullet}^{2,2}$	c^*	$ck_{\bullet}^{0,0}$
Bartlett	$1 - x$	-1	-1	$\frac{1}{3}$	1	0	2, 28	0, 76
2nd Order	$1 - 2x + x^2$	-2	0	$\frac{1}{5}$	$\frac{4}{3}$	4	3, 42	0, 68
Epanechnikov	$1 - x^2$	0	-2	$\frac{8}{15}$	$\frac{4}{3}$	4	2, 46	1, 31

Tabelle 3.1: Kernel Estimator mit Konvergenzrate $n^{-\frac{1}{6}}$

Bemerkung 3.5.4. ³³ In Tabelle 3.1 werden Bartlett Kernel ($k(x) = 1 - x$), der Epanechnikov Kernel ($k(x) = 1 - x^2$) und Second-Order Kernel ($k(x) = 1 - 2x + x^2$) miteinander verglichen. Dabei hat der Bartlett Kernel dieselbe asymptotische Verteilung wie der bereits vorgestellte TSRV. Der Second-Order Kernel ist der Effizienteste unter den drei Kernels, gefolgt vom Bartlett Kernel und dem Epanechnikov Kernel.

Spezialfälle mit Konvergenzrate $n^{-\frac{1}{4}}$

Für $H = x\xi\sqrt{n}$ und $m \rightarrow \infty$ ist die asymptotische Varianz von (3.45) proportional zu

$$4t \int_0^t \sigma_u^4 du (ck_{\bullet}^{0,0}\xi + 2c^{-1}k_{\bullet}^{1,1}\rho\xi + c^{-3}k_{\bullet}^{2,2}\xi) = \omega \left(t \int_0^t \sigma_u^4 du \right)^{\frac{3}{4}} \underbrace{4(ck_{\bullet}^{0,0} + 2c^{-1}\rho k_{\bullet}^{1,1} + c^{-3}k_{\bullet}^{2,2})}_{g(c)}.$$

Um jenes c zu bestimmen, welches die asymptotische Varianz minimiert, müssen wir also $g(c)$ minimieren. Dazu multiplizieren wir mit c^3 und setzen $x = c^2$ und erhalten so $k_{\bullet}^{0,0}x^2 - 2\rho k_{\bullet}^{1,1}x - 3k_{\bullet}^{2,2} = 0$. Durch Lösen und Umformen erhalten wir dann

$$c^* = \sqrt{\rho \frac{k_{\bullet}^{1,1}}{k_{\bullet}^{0,0}} \left[1 + \sqrt{1 + \frac{3d}{\rho}} \right]}, \quad \text{wobei } d = \frac{k_{\bullet}^{0,0} k_{\bullet}^{2,2}}{(k_{\bullet}^{1,1})^2}.$$

Mit Hilfe von c^* kann die asymptotische Varianz dann als $g \cdot \omega \left(\int_0^t \sigma_u^4 du \right)^{\frac{3}{4}}$ ausgedrückt werden, wobei

$$g = g(c^*) = \frac{16}{3} \sqrt{\rho k_{\bullet}^{0,0} k_{\bullet}^{1,1}} \left[\frac{1}{\sqrt{1 + \sqrt{1 + \frac{3d}{\rho}}}} + \sqrt{1 + \sqrt{1 + \frac{3d}{\rho}}} \right].$$

Aus den Eigenschaften des Maximum Likelihood Schätzers in [AMZ05], kann aus $\rho = 1$ auf $g \geq 8$ geschlossen werden. Es kann gezeigt werden, dass g wächst falls ρ fällt. Also im

³³vgl. Tabelle 4.7

3 Volatilitätschätzer

heteroskedastischem Fall $\rho < 1$ erwarten wir $g > 8$.

Der kubische Kernel $k(x) = 1 - 3x^2 + 2x^3$ erfüllt sowohl die Bedingung $k'(0)^2 + k'(1)^2 = 0$, als auch $k(0) = 1$ und $k(1) = 0$. Der kubische Kernel ist jener Kernel aus der Klasse der *Smooth Kernels*, der $k_{\bullet}^{2,2}$ minimiert. Im Allgemeinen gilt also $k_{\bullet}^{2,2} \geq 12$. Weiters ist zu bemerken, dass der kubische Kernel und der MSRV diesselbe asymptotische Verteilung besitzen. Minimieren von $k_{\bullet}^{2,2}$ impliziert aber nicht automatisch eine Minimierung von g . Der sogenannte flat-top Parzen Kernel

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3, & 0 \leq x \leq \frac{1}{2} \\ 2(1 - x)^3, & \frac{1}{2} \leq x \leq 1 \end{cases}$$

hat zum Beispiel eine kleinere asymptotische Varianz als kubische Kernels.

Ebenfalls interessant ist der flat-top Tukey-Hanning Kernel

$$k(x) = \sin^2 \left\{ \frac{\pi}{2} (1 - x)^p \right\}.$$

Bemerkung 3.5.5. ³⁴ Tukey-Hanning Kernel und kubischer Kernel verhalten sich fast identisch, wie in Tabelle 3.2 zu sehen ist. Der Parzen Kernel übertrifft den kubischen Kernel, kann aber mit dem Tukey-Hanning Kernel nicht mithalten.³⁵

	$k(x)$	$k_{\bullet}^{0,0}$	$k_{\bullet}^{1,1}$	$k_{\bullet}^{2,2}$	c^*	g
Kubischer Kernel	$1 - 3x^2 + 2x^3$	0,371	1,20	12,0	3,68	9,04
Parzen	$\begin{cases} 1 - 6x^2 + 6x^3, & 0 \leq x \leq \frac{1}{2} \\ 2(1 - x)^3, & \frac{1}{2} \leq x \leq 1 \end{cases}$	0,269	1,50	24,0	4,77	8,54
Tukey-Hanning ₁	$\sin^2 \left\{ \frac{\pi}{2} (1 - x) \right\}$	0,375	1,23	12,1	3,70	9,18
Tukey-Hanning ₂	$\sin^2 \left\{ \frac{\pi}{2} (1 - x)^2 \right\}$	0,219	1,71	41,7	5,74	8,29

Tabelle 3.2: Kernel Estimator mit Konvergenzrate $n^{-\frac{1}{4}}$

ω^2 Schätzer

Um die Varianz des Microstructure Noiseprozesses zu schätzen, werden in [BNHLS08] folgende zwei Schätzer vorgeschlagen. Einerseits der unverzerrte Schätzer $\tilde{\omega}^2 = \frac{-(\gamma_1(Y_\delta) + \gamma_{-1}(Y_\delta))}{2n}$ und andererseits der Schätzer $\hat{\omega}^2 = \frac{\gamma_0(Y_\delta)}{2n}$, welcher eine Verzerrung von $\frac{\int_0^t \sigma_u^2 du}{2n}$ hat. Beide Schätzer haben also ihre Vor- und Nachteile. So kann $\tilde{\omega}^2$ negativ werden und $\hat{\omega}^2$ kann

³⁴vgl. Tabelle 4.7

³⁵vgl. Tabelle 4.7

3 Volatilitätschätzer

recht verzerrt sein. Für die Varianzen gilt

$$\begin{aligned} \text{Var} \left(n^{\frac{1}{2}}(\tilde{\omega}^2 - \omega^2) \right) &= \omega^4(5 + \lambda^2) \\ \text{Var} \left(n^{\frac{1}{2}}(\hat{\omega}^2 - \omega^2) \right) &= \omega^4(1 + \lambda^2). \end{aligned}$$

Falls der Noiseprozess einer Normalverteilung unterliegt, gilt $\lambda^2 = 2$ und somit haben $\tilde{\omega}^2$ und $\hat{\omega}^2$ eine Varianz, die 3,5 beziehungsweise 1,5 mal so groß wie die des MLE ist. Eine einfache Biaskorrektur von $\hat{\omega}^2$, welche uns auch nicht negative Schätzwerte garantiert, führt uns zu unserem bevorzugtem Schätzer. Es gilt

$$\log \mathbb{E}(\hat{\omega}^2) = \log \omega^2 + \log \left(1 + \frac{IV}{(2n\omega^2)} \right).$$

Substituieren wir nun IV durch $K(Y_\delta)$ und $\tilde{\omega}^2$ durch $\frac{\gamma_0(Y_\delta)}{2n}$ so erhalten wir

$$\tilde{\omega}^2 = e^{\log \hat{\omega}^2 - \frac{K(Y_\delta)}{\gamma_0(Y_\delta)}}.$$

$\frac{K(Y_\delta)}{\gamma_0(Y_\delta)}$ ist ein Schätzer für die relative Verzerrung von $\hat{\omega}^2$, welcher für $n \rightarrow \infty$ verschwindet. Es gilt also $\tilde{\omega}^2 - \hat{\omega}^2 \xrightarrow{P} 0$.

3.5.4 Anpassungen an den Microstructure Noiseprozess

Auswirkungen von Abhängigkeiten

Bis jetzt sind wir davon ausgegangen, dass der Microstructure Noiseprozess $\epsilon_t \in \mathcal{WN}$. Jetzt wollen wir die Forderung $\epsilon_s \perp \epsilon_r$ abschwächen und erlauben Abhängigkeiten im Ausmaß von

$$\sum_{h=1}^H a_{h,H} \epsilon_{h\delta} = \mathcal{O}_p(1) \quad \text{für alle} \quad \sum_{h=1}^H a_{h,H}^2 = \mathcal{O}(1). \quad (3.46)$$

Wir haben jetzt eine Situation, in der die reihenweisen Abhängigkeiten eng mit der Stichprobenfrequenz δ zusammenhängen. Die Abhängigkeit zwischen ϵ_s und ϵ_r für $s \neq r$ verschwindet für $\delta \rightarrow 0$, während die Abhängigkeit von $\epsilon_{h\delta}$ und $\epsilon_{(h-j)\delta}$ für jedes beliebige δ erhalten bleibt. Ein einfaches Beispiel hierfür ist ein Moving Average Prozess 1. Ordnung mit zeitweiser Abhängigkeit, welche bei verändertem δ unverändert bleibt.

Proposition 3.5.6. Angenommen es sei (3.46) erfüllt. Falls $k'(0) = k'(1) = 0$, dann gilt

$$K(\epsilon_\delta) = -2H^{-2} \sum_{h=1}^H k'' \left(\frac{h}{H} \right) \sum_{i=1}^n \epsilon_{i\delta} \epsilon_{(i-h)\delta} + \mathcal{O}_p(nH^{-3}) + \mathcal{O}_p(H^{-\frac{1}{2}}).$$

Bei Auftreten von abhängigem Noise stimmt es nicht mehr, dass $H^{-\frac{1}{2}} \sum_{h=1}^H k'' \left(\frac{h}{H} \right) n^{-\frac{1}{2}} \cdot \sum_{i=1}^n \epsilon_{i\delta} \epsilon_{(i-h)\delta} \xrightarrow{L} N(0, k_{\bullet}^{2,2} \omega^4)$. Dieser Term jedoch ist von der Ordnung $\mathcal{O}_p(1)$. Falls eine „inefficient rate“ für H verwendet wird, wie zum Beispiel $H \propto n^{\frac{2}{3}}$, hat das abhängige Noise keinen Einfluss auf die asymptotische Verteilung von $K(Y_\delta)$.

3 Volatilitätschätzer

Proposition 3.5.7. Wir nehmen an, dass $k''(0) = 0$, $|k'''(0)| < \infty$ und $\epsilon_{j\delta}$ für $j = \dots, 0, 1, 2, \dots$, ein AR(1) Prozess mit „persistence parameter“ ϕ ($\phi < 1$) ist. Dann gilt

$$(nH)^{-\frac{1}{2}} \sum_{h=1}^H k''\left(\frac{h}{H}\right) \sum_{i=1}^n \epsilon_{i\delta} \epsilon_{(i-h)\delta} \xrightarrow{D} N\left(0, \omega^4 \frac{1+\phi^2}{1-\phi^2} k_{\bullet}^{2,2}\right).$$

Das bedeutet

$$K(\epsilon_\delta) = \mathcal{O}_p\left(\frac{n^{\frac{1}{2}}}{H^{\frac{3}{2}}}\right) + \mathcal{O}_p\left(\frac{n}{H^3}\right) + \mathcal{O}_p\left(\frac{1}{m^{\frac{1}{2}} H^{\frac{1}{2}}}\right).$$

Falls also $H \propto n^{\frac{1}{2}}$, dann gilt

$$K(Y_\delta) = \mathcal{O}_p(n^{-\frac{1}{4}}).$$

Falls wir annehmen, dass $X_t \perp \epsilon_t$ und $\delta \downarrow 0$, dann verändert die zeitweise Abhängigkeit des Noiseprozesses das asymptotische Verhalten von $\gamma_h(\epsilon_\delta, X_\delta)$ nicht.

Bemerkung 3.5.8. Die obige Proposition besagt, dass falls $H \propto n^{\frac{1}{2}}$, $K(\epsilon_\delta) = \mathcal{O}_p(n^{-\frac{1}{4}})$ und somit die Konvergenzrate unverändert bleibt. Die asymptotische Verteilung hingegen verändert sich.

Endogenes Microstructure Noise

Eine andere wichtige Annahme war bislang, dass $X_t \perp \epsilon_t$, also der Microstructure Noiseprozess als exogener Prozess betrachtet wurde. Jetzt untersuchen wir, wie sich der Kernel Estimator verhält, falls ϵ_t endogen ist. Dazu betrachten wir ein einfaches lineares endogenes Modell

$$\epsilon_{\delta i} = \sum_{h=0}^{\bar{H}} \beta_h (X_{\delta(i-h)} - X_{\delta(i-1-h)}) + \bar{\epsilon}_{\delta i},$$

wobei $X_t \perp \bar{\epsilon}_t$. Der Einfachheit halber nehmen wir $\bar{\epsilon}_t \in \mathcal{WN}$ an, dann gilt

$$\gamma_h(X_\delta, \epsilon_\delta) = \sum_{j=0}^{\bar{H}} \beta_j \gamma_{h+j}(X_\delta) - \sum_{j=0}^{\bar{H}} \beta_j \gamma_{h+j+1}(X_\delta) + \gamma_h(X_\delta, \bar{\epsilon}_\delta).$$

Mit Hilfe der asymptotischen Methoden für Realized Kernels unter exogenem Noiseprozess erhalten wir

$$\gamma_h(X_\delta, \epsilon_\delta) - \gamma_h(X_\delta, \bar{\epsilon}_\delta) = \begin{cases} \beta_0[X] + \mathcal{O}_p(n^{-\frac{1}{2}}), & h = 0 \\ -\beta_0[X] + \mathcal{O}_p(n^{-\frac{1}{2}}), & h = -1, \\ \mathcal{O}_p(n^{-\frac{1}{2}}) & |h| \neq 1. \end{cases}$$

Daher sind Realized Kernel robust gegenüber endogenen Microstructure Noiseprozessen.

3.6 Fourier Estimator³⁶

Mit Hilfe der Fouriermethode können wir die Volatilität als Reihenentwicklung mit Koeffizienten, welche aus den Fourierkoeffizienten der Variation des zugrundeliegenden Preisprozesses erhalten werden, darstellen. Der Fourier Estimator benötigt dabei keine Korrektur, um statistisch effizient und robust gegenüber Marktunstimmigkeiten zu sein. Falls er korrekt verwendet wird, können alle vorhandenen Daten verwendet werden, ohne einen verzerrten Schätzer zu erhalten. Es werden dabei auch Autokovarianzen höherer Ordnungen miteinbezogen. Der Schätzer ist jedoch trotzdem inkonsistent.

3.6.1 Definition Fourier Volatility Estimator

Die Fouriermethode der Volatilitätsschätzung baut auf einer Verbindung zwischen der Fouriertransformierten des Preisprozesses und der Fouriertransformierten des Volatilitätsprozesses auf. Dieses entscheidende Resultat wird in folgender Proposition zum Ausdruck gebracht. Dabei sei erwähnt, dass die ursprüngliche Zeit durch Reskalieren immer auf das Intervall $[0, 2\pi]$ reduziert werden kann.

Proposition 3.6.1. Sei X_t ein Semimartingalprozess, der Annahme 2.1.9 erfüllt. Bezeichnen wir die *Fouriertransformierte von dX* mit

$$\mathcal{F}(dX)(k) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikt} dX_t,$$

und das *Bohrkonvolutionsprodukt* zwischen den Funktionen Φ, Ψ ist definiert durch

$$(\Phi *_B \Psi)(k) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{s=-N}^N \Phi(s) \Psi(k-s). \quad (3.47)$$

Dann gilt

$$\frac{1}{2\pi} \mathcal{F}(\sigma^2)(k) = (\mathcal{F}(dX) *_B \mathcal{F}(dX))(k), \quad \text{f.s. für alle } k \in \mathbb{Z}. \quad (3.48)$$

Die Gleichung (3.48) stimmt fast sicher, was bedeutet, dass der Limes auf der rechten Seite von (3.47) in Wahrscheinlichkeit existiert.

Der Einfachheit halber nehmen wir $\mu = 0$ an, da der Driftterm nichts zu dem Volatilitätsschätzer beiträgt. Für den Fall $k = 0$ in Gleichung (3.48), können wir nun die integrierte Volatilität über das Zeitintervall $[0, 2\pi]$ wie folgt berechnen:

$$\int_0^{2\pi} \sigma_t^2 dt = (2\pi)^2 (\mathcal{F}(dX) *_B \mathcal{F}(dX))(0). \quad (3.49)$$

³⁶vgl. [MM02], [MM09], [MS08]

3 Volatilitätschätzer

In der Folge gehen wir davon aus, dass der logarithmische Preisprozess X_t Werte auf einem diskreten uneinheitlichen Netz $0 = t_{0,n} \leq t_{1,n} \leq \dots \leq t_{k_n,n} \leq 2\pi$ für jedes $n \geq 1$ annimmt, wobei wir einfachheitshalber $k_n = n$ wählen können. Die einzige Bedingung ist, dass $\rho(n) := \max_{0 \leq h \leq n} |t_{h+1,n} - t_{h,n}| \xrightarrow{n \rightarrow \infty} 0$. Aufgrund von (3.49) ist der Fourier Estimator der integrierten Volatilität für endliche Stichproben über $[0, 2\pi]$ folgendermaßen definiert:

$$\frac{(2\pi)^2}{2N+1} \sum_{s=-N}^N \mathcal{F}(dX)_n(s) \mathcal{F}(dX)_n(-s),$$

wobei

$$\mathcal{F}(dX)_n(s) = \frac{1}{2\pi} \sum_{j=1}^n e^{(-ist_j)\delta_j(X)}. \quad (3.50)$$

Wie in [MM09] gezeigt, gilt folgende Konvergenz in Wahrscheinlichkeit:

$$\lim_{n, N \rightarrow \infty} \frac{1}{2N+1} \sum_{|s| \leq N} \sum_{j=1}^n \sum_{j'=1}^n e^{is(t_j - t_{j'})} \delta_j(X) \delta_{j'}(X) = \int_0^{2\pi} \sigma_t^2 dt.$$

Bemerkung 3.6.2. Zwei wichtige Eigenschaften der Fouriermethode sind einerseits, dass der Fourier Estimator alle vorhandenen Daten durch Integration miteinbezieht und somit nicht nur die quadratischen Preisinkremente, sondern auch die Produkte der disjunkten Inkremente unseres Zeitintervalls, zum Beispiel die Autokovarianzen aller Ordnungen. Andererseits gewichtet das Konvolutionsprodukt die übergreifenden Produkte jeder gegebenen Frequenz.

Unser Hauptinteresse liegt nun darin, einen konsistenten möglichst unverzerrten Schätzer für die integrierte Volatilität des kontaminierten Preisprozesses Y_t vorzustellen. Der *Fourier Estimator* ist dann wie folgt definiert:

$$\hat{\sigma}_{n,N}^2 = \frac{(2\pi)^2}{2N+1} \sum_{s=-N}^N \mathcal{F}(dY)_n(s) \mathcal{F}(dY)_n(-s). \quad (3.51)$$

Mit Hilfe von (3.50), kann der Fourier Estimator auch anders angeschrieben werden und zwar durch

$$\hat{\sigma}_{n,N}^2 = \sum_{j=1}^n \sum_{j'=1}^n D_N(t_j - t_{j'}) \delta_j(Y) \delta_{j'}(Y),$$

wobei $D_N(t)$ der Dirichlet Kernel

$$D_N(t) = \frac{1}{2N+1} \sum_{|s| \leq N} e^{ist} = \frac{1}{2N+1} \frac{\sin[(N + \frac{1}{2})t]}{\sin \frac{t}{2}}$$

3 Volatilitätschätzer

ist.

Für numerische Anwendungen betrachten wir noch eine andere Version des Fourier Estimators. Durch Gewichten des Konvolutionsprodukt mit dem La Vallee Poussin Kernel erhalten wir folgende Form des Schätzers:

$$\hat{\Sigma}_{n,N}^2 = \frac{(2\pi)^2}{N+1} \sum_{s=-N}^N \left(1 - \frac{|s|}{N}\right) \mathcal{F}(dY)_n(s) \mathcal{F}(dY)_n(-s).$$

Die mathematischen Resultate in den nächsten zwei Unterkapiteln werden wir hier zwar nur für den Dirichlet Kernel zeigen, können aber auf gleiche Art und Weise für den obigen Fourier Estimator gezeigt werden, indem wir ihn wie folgt anschreiben:

$$\hat{\Sigma}_{n,N}^2 = \sum_{j=1}^n \sum_{j'=1}^n F_N(t_j - t_{j'}) \delta_j(Y) \delta_{j'}(Y),$$

wobei $F_N(t) = \frac{\sin^2(Nt)}{(Nt)^2}$ der Fejer Kernel ist. Diese Modifikation liefert eine Verbesserung der Performance bei hoher Stichprobenfrequenz.

Wie auch bei anderen Schätzern, sollte auch hier der Beitrag von Autokovarianzen verschiedener Ordnungen betrachtet werden, um so eine Biaskorrektur durchzuführen. Der Fourier Estimator arbeitet auf andere Weise, als die meisten anderen Volatilitätsschätzer. Um das zu illustrieren, nehmen wir einen konstanten Beobachtungsabstand τ und eine Bandweite H zur Hilfe. Dann ist die Korrektur des Realized Kernel Estimators gegeben durch

$$\sum_{h=1}^H k\left(\frac{h-1}{H}\right) \left[\sum_{j=1}^n \delta_j(Y) \delta_{j-h}(Y) + \sum_{j=1}^n \delta_j(Y) \delta_{j+h}(Y) \right],$$

wobei die Gewichtsfunktion $k(x)$ eine Funktion der Bandweite H ist. Für den Fourier Estimator ergibt sich eine Korrektur von

$$\sum_{h=1}^H D_N(\tau h) \left[\sum_{j=1}^n \delta_j(Y) \delta_{j-h}(Y) + \sum_{j=1}^n \delta_j(Y) \delta_{j+h}(Y) \right].$$

Der Dirichlet Kernel im Fourier Estimator hängt im Wesentlichen von der Anzahl der Frequenzen N und dem Abstand zwischen den einzelnen Beobachtungen ab. Es gibt also einige Anpassungsmöglichkeiten und wir werden im nächsten Unterkapitel sehen, dass eine einfache Anpassung von N den Fourier Estimator auch bei Auftreten von Microstructure Noise effizient macht.

Bemerkung 3.6.3. In den folgenden zwei Unterkapiteln versuchen wir die Performance des Fourier Estimators zu verbessern, indem wir eine Formel für die Verzerrung und den MSE herleiten und dann als Funktion der Anzahl an Frequenzen N minimieren. Diese Untersuchungen stellen eine Basis für die optimale Wahl der Anzahl der Frequenzen N und der Stichprobenintervalle dar.

3.6.2 Bias

Wie wir bereits wissen, liefert der klassische RV Estimator in der Abwesenheit von Market Microstructure Noise einen konsistenten Schätzer. Bei Auftreten von Microstructure Noise ist er jedoch verzerrt und divergiert. Deshalb betrachten wir diese Situation jetzt auch beim Fourier Estimator. Bei der Fouriermethode ist es nicht notwendig, äquidistante Beobachtungsabstände zu wählen. Der Einfachheit halber nehmen wir hier aber trotzdem äquidistante Zeitabstände von $\frac{2\pi}{n}$ an. Dabei hat eine Handelsperiode eine Länge von $[0, 2\pi]$. Für die Verzerrung des Fourier Estimators kann jetzt folgende Aussage getroffen werden:

Proposition 3.6.4. Für alle festen n, N gilt folgende Identität:

$$\mathbb{E} [\hat{\sigma}_{n,N}^2 - IV] = 2n\mathbb{E} [\epsilon^2] \left(1 - \frac{1}{2N+1} \frac{\sin([2N+1]\frac{\pi}{n})}{\sin(\frac{\pi}{n})} \right). \quad (3.52)$$

Verlangen wir $\frac{N^2}{n} \rightarrow 0$, so erhalten wir

$$\lim_{n, N \rightarrow \infty} 2n\mathbb{E} [\epsilon^2] \left(1 - \frac{1}{2N+1} \frac{\sin([2N+1]\frac{\pi}{n})}{\sin(\frac{\pi}{n})} \right) = 0.$$

Zur Erinnerung der RV Estimator hat eine Verzerrung von $2n\mathbb{E}[\epsilon^2]$. Falls also eine feste Anzahl an Beobachtungen n gewählt wird und N im Verhältnis zu n klein ist, dann erhalten wir bei dem Fourier Estimator eine kleinere Verzerrung als bei der RV . Des Weiteren geht sie sogar gegen Null, falls n und N geeignet gewählt werden.

Zusammenfassend können wir also folgende Aussage treffen:

Bemerkung 3.6.5. Der Fourier Estimator ist unter der Bedingung $\frac{N^2}{n} \rightarrow 0$ asymptotisch unverzerrt. Außerdem zeigt das Resultat aus (3.52), dass für festes n und endlicher Stichprobengröße, bei passender Wahl von N , die Verzerrung kleiner als beim RV Estimator ist.

3.6.3 MSE

In diesem Unterabschnitt leiten wir eine Formel für den Mean Squared Error (MSE) her. Dabei setzen wir voraus, dass der Volatilitätsprozess unabhängig von der Brown'schen

3 Volatilitätschätzer

Bewegung ist. Deshalb nehmen wir an, dass die „no leverage hypothesis“ hält.

Der MSE des RV Estimators ist von der Größenordnung

$$\mathbb{E} \left[(RV_{\text{all}}^{(n)} - IV)^2 \right] = 2 \frac{T}{n} (IQ + o(1)) + \Lambda_n,$$

wobei IQ die integrierte Quartizität $\int_0^T \sigma_t^4 dt$ ist und

$$\Lambda_n := n^2 \alpha + n \beta + \gamma,$$

mit

$$\begin{aligned} \alpha &= (\mathbb{E}[\eta^2])^2, & \beta &= \mathbb{E}[\eta^4] + 2\mathbb{E}[\eta^2 \eta_{-1}^2] - 3(\mathbb{E}[\eta^2])^2 \\ \gamma &= 4\mathbb{E}[\eta^2] IV - 2\mathbb{E}[\eta^2 \eta_{-1}^2] + 2([\eta^2])^2. \end{aligned} \quad (3.53)$$

Dabei verwenden wir die Notation η für $\epsilon_{t_j} - \epsilon_{t_{j-1}}$ für ein frei wählbares t_j und η_{-1} für $\epsilon_{t_{j-1}} - \epsilon_{t_{j-2}}$ für dasselbige t_j .

Mit Hilfe einfacher Berechnungen und unter Annahme 2.2.1 Punkt 1 gilt

$$\alpha = 4\mathbb{E}[\epsilon^2]^2, \quad \beta = 4\mathbb{E}[\epsilon^4], \quad \gamma = 8\mathbb{E}[\epsilon^2]^2 - \mathbb{E}[\epsilon^4].$$

Der MSE des Fourier Estimators kann jetzt wie folgt berechnet werden:

Proposition 3.6.6. Für alle festen n, N gilt

$$\mathbb{E} [(\hat{\sigma}_{n,N}^2 - IV)^2] = 2 \frac{2\pi}{n} (IQ + o(1)) + n^2 \hat{\alpha} + n \hat{\beta} + \hat{\gamma},$$

wobei

$$\begin{aligned} \hat{\alpha} &= \alpha \left[1 + D_N^2 \left(\frac{2\pi}{n} \right) - 2D_N \left(\frac{2\pi}{n} \right) \right], \\ \hat{\beta} &= \beta \left[1 + D_N^2 \left(\frac{2\pi}{n} \right) - 2D_N \left(\frac{2\pi}{n} \right) \right], \\ \hat{\gamma} &= \gamma + 4IQ \frac{2\pi}{2N+1} + 4 \left[(\mathbb{E}[\epsilon^2]^2 + \mathbb{E}[\epsilon^4]) \left(2D_N \left(\frac{2\pi}{n} \right) \right) - D_N^2 \left(\frac{2\pi}{n} \right) \right] \end{aligned}$$

mit α, β, γ aus (3.53).

Dabei hängen die Terme $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$ lediglich von dem Zeitabstand zwischen den Beobachtungen und der Anzahl der Fourierkoeffizienten ab.

Um das obige Resultat aus Proposition 3.6.6 zu verdeutlichen, bezeichnen wir mit MSE_{RV} den Mean Squared Error des RV Estimators ohne Microstructure Noise und MSE_{RV_m} den Mean Squared Error in der Gegenwart eines Noiseprozesses. Es gilt

$$MSE_{RV} = 2 \frac{2\pi}{n} (IQ + o(1))$$

3 Volatilitätschätzer

und

$$MSE_{RV_m} = MSE_{RV} + n^2\alpha + n\beta + \gamma.$$

Aus obigen Formeln ist klar ersichtlich, dass der MSE in der Abwesenheit von Microstructure Noise gegen Null konvergiert, falls $n \rightarrow \infty$. Andererseits in der Gegenwart eines Noiseprozesses divergiert der MSE für $n \rightarrow \infty$ aufgrund der n^2 und n Terme.

Auf dieselbe Art und Weise können wir das jetzt für den MSE beim Fourier Estimator betrachten und bezeichnen MSE_F und MSE_{F_m} analog. Es gilt

$$MSE_F = MSE_{RV} + c(n, N),$$

wobei $c(n, N)$ ein Term ist, der gegen Null konvergiert, falls $n, N \rightarrow \infty$. Es kann gezeigt werden, dass $c(n, N) \leq 4IQ \frac{2\pi}{2N+1}$ ist. Überdies gilt

$$MSE_{F_m} = MSE_F + n^2\tilde{\alpha}(n, N) + n\tilde{\beta}(n, N) + \tilde{\gamma}(n, N),$$

wobei

$$\begin{aligned}\tilde{\alpha}(n, N) &= 4\mathbb{E}[\epsilon^2]^2 \left[1 + D_N^2 \left(\frac{2\pi}{n} \right) - 2D_N \left(\frac{2\pi}{n} \right) \right], \\ \tilde{\beta}(n, N) &= 4\mathbb{E}[\epsilon^4] \left[1 + D_N^2 \left(\frac{2\pi}{n} \right) - 2D_N \left(\frac{2\pi}{n} \right) \right]\end{aligned}$$

und

$$\tilde{\gamma} = \gamma + 4 \left[(\mathbb{E}[\epsilon^2]^2 + \mathbb{E}[\epsilon^4]) \left(2D_N \left(\frac{2\pi}{n} \right) \right) - D_N^2 \left(\frac{2\pi}{n} \right) \right].$$

Falls $\frac{N^2}{n} \rightarrow 0$ gilt

$$\lim_{n, N \rightarrow \infty} n^2\tilde{\alpha}(n, N) + n\tilde{\beta}(n, N) = 0$$

und

$$\lim_{n, N \rightarrow \infty} \tilde{\gamma}(n, N) = 8\mathbb{E}[\epsilon^2]IV + 2\mathbb{E}[\epsilon^4] + 6\mathbb{E}[\epsilon^2]^2. \quad (3.54)$$

Bemerkung 3.6.7. Der MSE des Fourier Estimators divergiert nicht und ist auch nicht signifikant beeinflusst durch das Auftreten von Microstructure Noise. Durch geeignete Wahl von N , erhalten wir einen MSE_F bzw. MSE_{F_m} in der Größenordnung von (3.54). Der Fourier Estimator benötigt keine Biaskorrektur, um asymptotisch unverzerrt zu sein und robust gegenüber Marktunstimmigkeiten von $MA(1)$ -Typ zu sein.

3.6.4 Abhängiger Microstructure Noiseprozess

In diesem Unterabschnitt lockern wir die Voraussetzungen an unseren Microstructure Noiseprozess. Wir betrachten hier den Fall, dass ϵ_t und X_t miteinander korreliert sind.

Seien $t_0 < t_1 < \dots < t_n$ die Zeiten, in denen die Preise beobachtet werden. Wir bezeichnen mit $Z_{t_j} = X_{t_j} - X_{t_{j-1}}$ den Ertrag zwischen zwei Beobachtungen. Nehmen wir an, dass unser Noiseprozess jetzt wie folgt gegeben ist:

$$\tilde{\epsilon}_j = \alpha Z_j + \epsilon_j$$

wobei α eine Konstante ist und $\tilde{\epsilon}_j$, ϵ_j und Z_j die Kurznotationen für $\tilde{\epsilon}_{t_j}$, ϵ_{t_j} und Z_{t_j} sind. Für $\alpha = 0$ erhalten wir den Fall des unabhängigen Noiseprozesses. Sei $\eta_j := \epsilon_j - \epsilon_{j-1}$ und bezeichnen wir

$$\tilde{\eta}_j = \alpha(Z_j - Z_{j-1}) + \eta_j.$$

Dann gilt für jedes j

$$\mathbb{E}[\tilde{\eta}_j^2] = \alpha^2 \mathbb{E} \left[\int_{t_{j-1}}^{t_j} \sigma_s^2 ds \right] + \alpha^2 \mathbb{E} \left[\int_{t_{j-2}}^{t_{j-1}} \sigma_s^2 ds \right] + 2\mathbb{E}[\epsilon^2]$$

und

$$\mathbb{E}[\tilde{\eta}_j^2 Z_j] = \alpha \mathbb{E} \left[\int_{t_{j-1}}^{t_j} \sigma_s^2 ds \right].$$

Unter diesen Voraussetzungen kann die Verzerrung des RV Estimators folgendermaßen berechnet werden

$$\mathbb{E}[RV_{\text{all}}^{(n)} - IV] = 2\alpha^2 IV + 2\alpha IV + 2n\mathbb{E}[\epsilon^2].$$

Für $\alpha < 0$ wird die Verzerrung negativ. Dies ist genau dann der Fall, wenn $\tilde{\epsilon}_j$ und Z_j negativ korreliert sind. Für den Fourier Estimator erhalten wir folgendes Resultat:

Proposition 3.6.8. Für alle n, N gilt

$$\mathbb{E}[(\hat{\sigma}_{n,N}^2 - IV)^2] = (2\alpha^2 IV + 2\alpha IV + 2n\mathbb{E}[\epsilon^2]) \left(1 - D_N \left(\frac{2\pi}{n} \right) \right).$$

Bemerkung 3.6.9. Auch hier kann die Verzerrung negativ werden, falls $\alpha < 0$. Die Verzerrung für den Fourier Estimator konvergiert gegen Null, wenn $n, N \rightarrow \infty$ und $\frac{N^2}{n} \rightarrow 0$ auch im Falle eines abhängigen Noiseprozesses. Der Fourier Estimator ist also auch in diesem Fall ein asymptotisch unverzerrter Schätzer.

3.6.5 Wahl der optimalen Frequenz

Falls wir für die Frequenz $N = \frac{n}{2}$ wählen, verhält sich der Fourier Estimator wie die Realized Variance. Der Schätzer divergiert also, wenn die Intervallabstände gegen Null konvergieren. Bei geschickter Wahl der Frequenz, sodass $\frac{N^2}{n} \rightarrow 0$ gilt, kann ein unverzerrter Schätzer erhalten werden, dessen MSE gegen eine kleine Konstante konvergiert.

Nun ist es nicht so einfach diese optimale Frequenz zu bestimmen. Ab einem gewissen Wert N_{cut} zeigt sich, dass der Fourier Estimator inkonsistent wird, eine Folge der negativen Korrelation der logarithmischen Preise. Also ist es für Hochfrequenzdaten sinnvoll, die Frequenz als $N = \min(\frac{n}{2}, N_{cut})$ zu wählen. Der optimale Wert von N_{cut} kann durch Minimierung des MSE erhalten werden. Über N_{cut} können wir dann auch den Beobachtungsabstand $\delta = \frac{t*60}{2*N_{cut}}$ bestimmen.

4 Empirische Studie³⁷

Da wir nun theoretisch über die verschiedensten Schätzer Bescheid wissen, führen wir nun eine empirische Analyse durch, um die einzelnen Schätzer vergleichen zu können. Zuerst betrachten wir jeden Schätzer einzeln für sich, um dann in weiterer Folge einen Gesamtvergleich starten zu können.

4.1 Die Daten³⁸

Für die empirische Studie verwenden wir hier einen Hochfrequenzdatensatz des Istanbuler Stock Exchange (ISE) der Monate Juni und Juli 2008. Der ISE ist ein vollautomatisierter ordergetriebener Stock Exchange. Die Trades des ISE werden also elektronisch ohne Market Maker ausgeführt. Der ISE-30-Index umfasst dabei die größten an der Istanbuler Börse gehandelten Unternehmen und macht dabei 73% des gesamten Tradingvolumens aus. Im Jahr 2008 wurden im ISE-30-Index Aktien im Wert von \$248 Billionen gehandelt und die Marktkapitalisierung betrug \$11 Billionen. Das entspricht 0,7% bzw. 5% des New York Stock Exchange (NYSE). Unsere Daten erfassen dabei den gesamten Orderflow.

An der Istanbuler Börse wird von 09:30 bis 12:00 und nach einer Mittagspause von 14:00 bis 17:00 gehandelt. Der Handelstag beginnt dabei 15 Minuten früher, um mit einem Call Market Matching den Eröffnungspreis zu bestimmen. Danach wird nach dem Continuous Double Auction Konzept gehandelt. Order haben dabei eine Gültigkeit von einer Session oder einem ganzen Tag und sind für alle Broker sichtbar. Broker haben also Zugriff auf das gesamte Order Book. Der größte Stock des Index ist *T.Garanti Bankasi A.Ş.* (GARAN). Die Anzahl der Order bei dieser Aktie ist 8 mal größer als im Durchschnitt. Deshalb werden wir für die Studie unserer Schätzer diesen Stock verwenden.

³⁷vgl. [BRY08]

³⁸vgl. [VZFR13], S.6,[VZ13], S.6f

4 Empirische Studie

Datum	Trade Prices		Midquote Prices		Bid-Ask Spread	
	mean	stdev	mean	stdev	mean	stdev
2.06.08	5.5941	0.0818	5.5900	0.0809	0.0495	0.0055
3.06.08	5.6149	0.0557	5.6116	0.0463	0.0500	0.0020
4.06.08	5.5648	0.0572	5.5576	0.0559	0.0499	0.0033
5.06.08	5.6625	0.0800	5.6585	0.0841	0.0498	0.0029
6.06.08	5.6274	0.0453	5.6263	0.0454	0.0502	0.0044
9.06.08	3.3253	0.0203	3.3211	0.0190	0.0498	0.0066
10.06.08	3.3466	0.0330	3.3438	0.0329	0.0495	0.0076
11.06.08	3.3760	0.0220	3.3733	0.0198	0.0498	0.0034
12.06.08	3.3684	0.0184	3.3663	0.0163	0.0502	0.0085
13.06.08	3.3314	0.0286	3.3280	0.0300	0.0500	0.0016
16.06.08	3.2847	0.0428	3.2805	0.0440	0.0498	0.0039
17.06.08	3.2046	0.0402	3.2034	0.0414	0.0336	0.0151
19.06.08	3.1967	0.0424	3.1983	0.0429	0.0089	0.0766
20.06.08	3.2956	0.0263	3.3001	0.0278	0.0200	0.0010
23.06.08	3.2645	0.0361	3.2650	0.0355	0.0208	0.0046
24.06.08	3.2255	0.0405	3.2296	0.0403	0.0199	0.0017
25.06.08	3.2400	0.0128	3.2419	0.0121	0.0186	0.0309
26.06.08	3.1435	0.0408	3.1403	0.0413	0.0200	0.0015
27.06.08	2.9619	0.0223	2.9597	0.0208	0.0200	0.0024
30.06.08	2.8872	0.0441	2.8865	0.0423	0.0196	0.0031

Tabelle 4.1: Mittelwert und Standardabweichung der Trade Prices, Midquote Prices und des Bid-Ask Spreads für den Stock GARAN im Monat Juni 2008.

Der Noiseprozess

Da der Noiseprozess für Volatilitätsschätzer bei der Verwendung von Hochfrequenz-Börsendaten von wesentlicher Bedeutung ist, wollen wir diesen vor Beginn der empirischen Studie noch etwas genauer betrachten. Dazu verwenden wir unsere 3 verschiedenen Schätzer für den Erwartungswert des Noiseprozess und erkennen dabei, dass sich das Market Microstructure Noise in einer Größenordnung von 10^{-5} bis 10^{-6} bewegt. Dies zeigt uns, dass der Stock GARAN nicht allzu sehr durch Noise verunreinigt ist.

4 Empirische Studie

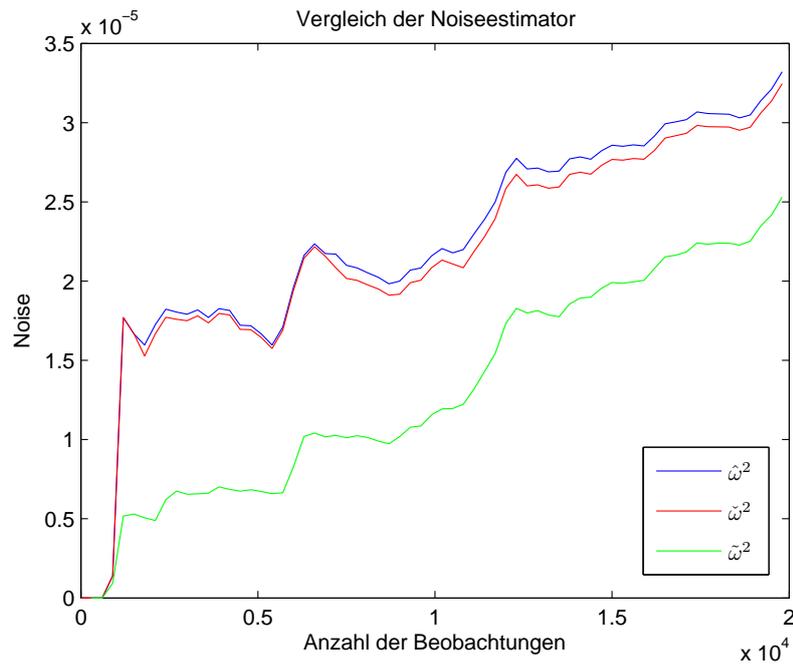


Abbildung 4.1: Die drei verschiedenen Noiseestimator im Vergleich

Datum	$\hat{\omega}^2$	$\tilde{\omega}^2$	$\tilde{\tilde{\omega}}^2$
2.06.08	$3.32e - 005$	$3.25e - 005$	$2.53e - 005$
3.06.08	$3.16e - 005$	$3.15e - 005$	$2.37e - 005$
4.06.08	$3.52e - 005$	$3.46e - 005$	$2.73e - 005$
5.06.08	$4.27e - 005$	$4.17e - 005$	$3.48e - 005$
6.06.08	$2.80e - 005$	$2.73e - 005$	$2.01e - 005$
9.06.08	$4.70e - 006$	$4.51e - 006$	$-3.21e - 006$
10.06.08	$5.25e - 006$	$4.95e - 006$	$-2.65e - 006$
11.06.08	$4.07e - 006$	$4.00e - 006$	$-3.83e - 006$
12.06.08	$3.30e - 006$	$3.25e - 006$	$-4.60e - 006$
13.06.08	$4.07e - 006$	$3.94e - 006$	$-3.83e - 006$
16.06.08	$5.27e - 006$	$5.05e - 006$	$-2.63e - 006$
17.06.08	$6.79e - 006$	$6.47e - 006$	$-1.12e - 006$
19.06.08	$4.74e - 006$	$4.61e - 006$	$-3.17e - 006$
20.06.08	$3.70e - 006$	$3.60e - 006$	$-4.21e - 006$
23.06.08	$4.21e - 006$	$3.96e - 006$	$-3.69e - 006$
24.06.08	$7.06e - 006$	$6.79e - 006$	$-8.44e - 007$
25.06.08	$3.02e - 006$	$2.95e - 006$	$-4.88e - 006$
26.06.08	$6.70e - 006$	$6.43e - 006$	$-1.21e - 006$
27.06.08	$6.72e - 006$	$6.39e - 006$	$-1.19e - 006$
30.06.08	$6.91e - 006$	$6.31e - 006$	$-9.96e - 007$

Tabelle 4.2: Die drei verschiedenen Noiseestimator im Vergleich

Einfluss von Datenaufbereitungsverfahren auf die Stabilität

Eine wesentliche Rolle bei der Performance all dieser Schätzer spielt die Datenqualität. Deshalb ist eine gute Datenaufbereitung besonders wichtig. Die rohen Daten werden aufbereitet, indem Datenfehler und Ausreißer entfernt werden. Des Weiteren können gewisse Daten gefiltert werden, um so den optimalen Datensatz zu erhalten. Es stellt sich heraus, dass die Datenvorbereitung einen wesentlichen Einfluss auf die Schätzwerte hat.

4.2 Analyse der einzelnen Schätzer

4.2.1 RV

Wir beginnen mit der Analyse der *RV*. Wie wir bereits in der Aufbereitung der Daten gesehen haben, hat der Noiseprozess eine Größenordnung von etwa 10^{-5} . Dadurch ist der Einfluss des Market Microstructure Noise auf die beobachteten Daten relativ gering. So ist eine in der Praxis meist übliche geringere optimale Samplingfrequenz von 5 Minuten zu erwarten. Diese Erwartung bestätigt sich auch. Wie in Abbildung 4.3 zu sehen ist, liegt bei unserem Stock GARAN die optimale Samplingfrequenz bei 70 Sekunden. Abbildung 4.2 und Tabelle 4.3 zeigen, dass die *RV* sehr stark von der gewählten Frequenz q abhängt und, falls alle Daten verwendet werden, die *RV* für $n \rightarrow \infty$ divergiert. Es ist daher bei der *RV* besonders auf die Samplingfrequenz zu achten. Bei optimal gewählter Samplingfrequenz $q^* = 70.4355$ beträgt der MSE 0.0011.

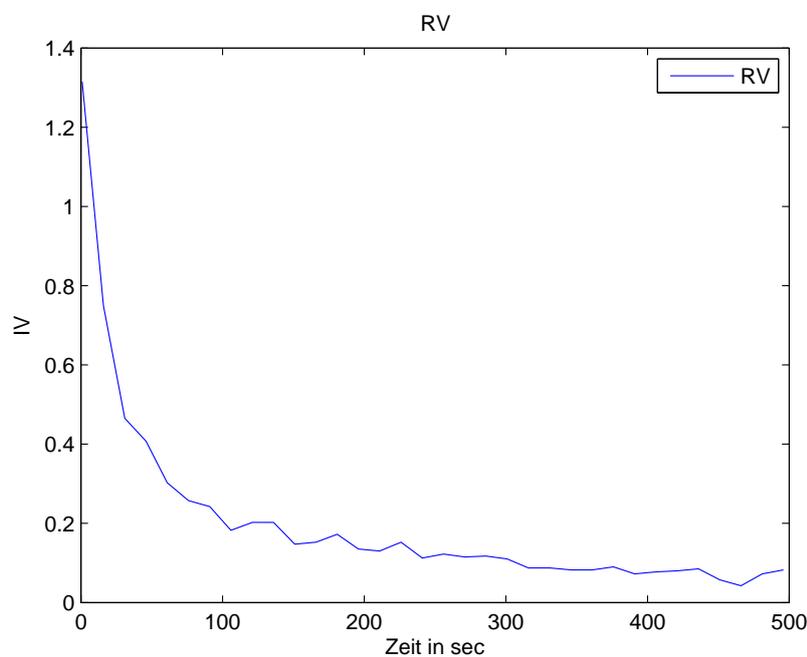


Abbildung 4.2: *RV* geplottet nach der Samplingfrequenz q

4 Empirische Studie

Datum	RV_{all}	$RV_{\text{sparse}}^{(15)}$	$RV_{\text{sparse}}^{(30)}$	$RV_{\text{sparse}}^{(60)}$	$RV_{\text{sparse}}^{(120)}$	$RV_{\text{sparse}}^{(300)}$	$RV_{\text{sparse}}^{(q^*)}$
2.06.08	1.3150	0.7750	0.5200	0.3150	0.2200	0.1150	0.2875
3.06.08	1.2500	0.6900	0.4400	0.3150	0.2000	0.1050	0.2700
4.06.08	1.3925	0.8475	0.6025	0.3625	0.1625	0.1275	0.3450
5.06.08	1.6900	0.8150	0.5050	0.3150	0.1300	0.0850	0.3400
6.06.08	1.1075	0.6125	0.3775	0.2625	0.1975	0.1025	0.2375
9.06.08	0.1860	0.1052	0.0748	0.0524	0.0372	0.0316	0.0400
10.06.08	0.2080	0.1120	0.0784	0.0520	0.0312	0.0200	0.0516
11.06.08	0.1612	0.0956	0.0700	0.0476	0.0244	0.0148	0.0408
12.06.08	0.1308	0.0788	0.0564	0.0372	0.0260	0.0108	0.0348
13.06.08	0.1612	0.0876	0.0636	0.0380	0.0236	0.0116	0.0396
16.06.08	0.2088	0.1032	0.0624	0.0376	0.0248	0.0112	0.0320
17.06.08	0.2688	0.1328	0.0800	0.0496	0.0336	0.0160	0.0460
19.06.08	0.1876	0.1060	0.0756	0.0452	0.0252	0.0164	0.0376
20.06.08	0.1464	0.0896	0.0648	0.0344	0.0240	0.0128	0.0312
23.06.08	0.1668	0.1092	0.0756	0.0540	0.0436	0.0236	0.0476
24.06.08	0.2796	0.1236	0.0780	0.0492	0.0340	0.0164	0.0448
25.06.08	0.1196	0.0780	0.0556	0.0380	0.0268	0.0116	0.0416
26.06.08	0.2652	0.1300	0.0812	0.0412	0.0236	0.0140	0.0360
27.06.08	0.2660	0.1156	0.0780	0.0484	0.0324	0.0252	0.0512
30.06.08	0.2736	0.1424	0.0840	0.0472	0.0336	0.0216	0.0448

Tabelle 4.3: Diverse RV Schätzer mit unterschiedlichen Samplingfrequenzen für den Stock GARAN im Monat Juni 2008

MSE

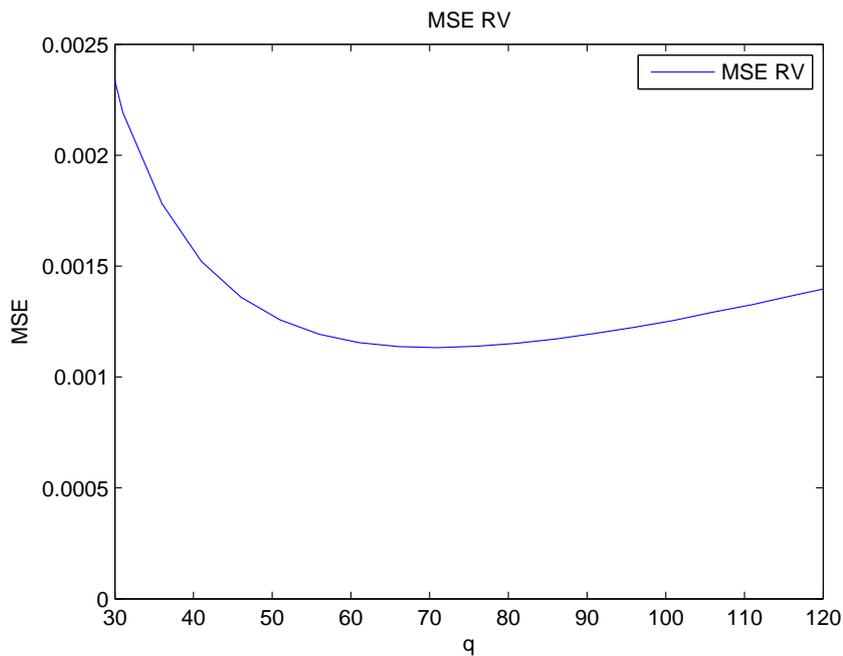


Abbildung 4.3: Der MSE der RV geplottet nach der Samplingfrequenz q

4.2.2 Zhou Estimator

Der Zhou Estimator führt eine leichte Biaskorrektur durch. Seine optimale Frequenz liegt bei $q^* = 2.4253$, wie in Abbildung 4.5 zu erkennen ist. Der MSE bei optimaler Samplingfrequenz liegt hier bei $2.2985e-004$.

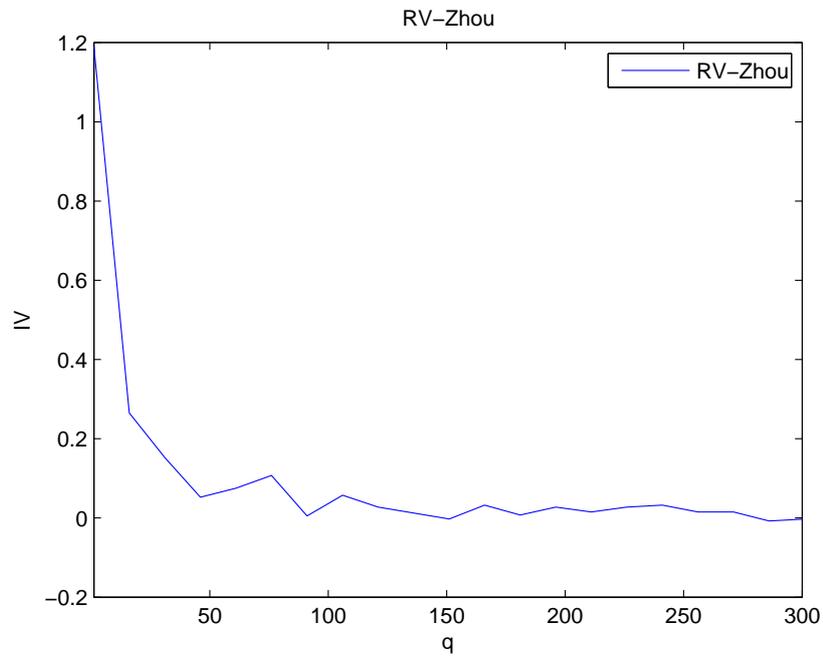


Abbildung 4.4: RV-Zhou geplottet nach der Samplingfrequenz q

4 Empirische Studie

Datum	$RV_{\text{Zhou}}^{(1)}$	$RV_{\text{Zhou}}^{(2)}$	$RV_{\text{Zhou}}^{(5)}$	$RV_{\text{Zhou}}^{(15)}$	$RV_{\text{Zhou}}^{(30)}$	$RV_{\text{Zhou}}^{(60)}$	$RV_{\text{Zhou}}^{(300)}$
2.06.08	1.1900	1.0750	0.7200	0.3000	0.1275	0.0925	0.0250
3.06.08	1.1200	0.8900	0.6400	0.2450	0.2050	0.0750	-0.0350
4.06.08	1.2375	1.0825	0.7725	0.2375	0.1050	0.0325	0.0125
5.06.08	1.4450	1.2700	0.7050	0.2300	0.1150	0.0250	0.0350
6.06.08	0.9925	0.8425	0.5900	0.2325	0.1625	0.0725	0.0125
9.06.08	0.1548	0.1400	0.0920	0.0412	0.0220	0.0184	0.0116
10.06.08	0.1784	0.1488	0.1016	0.0416	0.0220	0.0108	0.0048
11.06.08	0.1404	0.1200	0.0856	0.0400	0.0192	0.0040	0.0012
12.06.08	0.1132	0.0984	0.0724	0.0356	0.0260	0.0156	0.0044
13.06.08	0.1412	0.1156	0.0808	0.0404	0.0212	0.0124	0.0044
16.06.08	0.1696	0.1416	0.0904	0.0232	0.0208	0.0072	0.0080
17.06.08	0.2296	0.1888	0.1168	0.0340	0.0252	0.0164	0.0120
19.06.08	0.1636	0.1448	0.1028	0.0384	0.0120	0.0104	0.0024
20.06.08	0.1336	0.1088	0.0824	0.0352	0.0160	0.0104	0.0048
23.06.08	0.1516	0.1284	0.0964	0.0460	0.0268	0.0148	0.0044
24.06.08	0.2500	0.2044	0.1100	0.0284	0.0112	0.0160	0.0032
25.06.08	0.1068	0.1020	0.0764	0.0412	0.0216	0.0096	0.0036
26.06.08	0.2124	0.1700	0.1124	0.0300	0.0108	0.0076	0.0032
27.06.08	0.1956	0.1596	0.1012	0.0512	0.0216	0.0252	0.0080
30.06.08	0.2328	0.1824	0.1192	0.0352	0.0120	0.0168	0.0064

Tabelle 4.4: Vergleich RV-Zhou

MSE

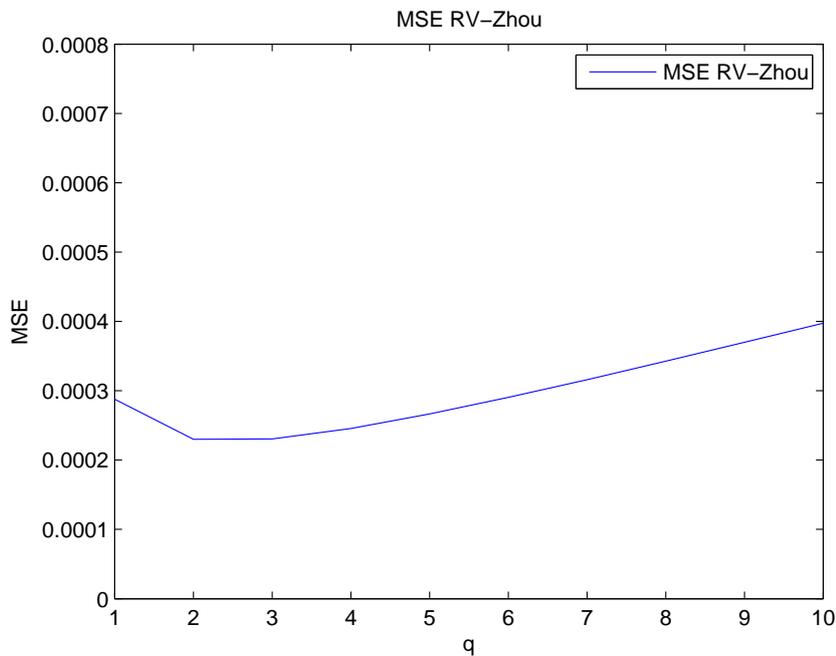


Abbildung 4.5: Der MSE der RV-Zhou geplottet nach der Samplingfrequenz q

4.2.3 TSRV

Wir betrachten bei der empirischen Analyse einerseits den normalen TSRV $\widehat{\langle X, X \rangle}_T^{(tsrv)}$ aus (3.26) und andererseits den angepassten TSRV $\widehat{\langle X, X \rangle}_T^{(tsrv, adj)}$ aus (3.30). Hier ist die Wahl der Subsamplingfrequenz K von entscheidender Bedeutung. Wir erhalten eine optimale Frequenz von 5.8519 bei einem MSE von 0.0013.

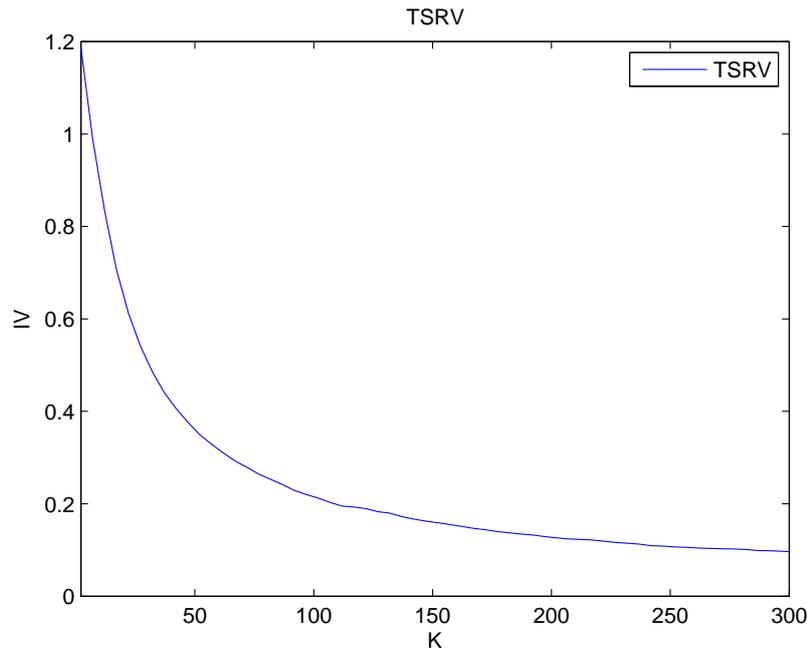


Abbildung 4.6: Vergleich des TSRV unter verschiedenen Subsamplingfrequenzen K

	6sec	30sec	1min	2min	5min	10min
$TSRV$	0.8580	0.4921	0.3135	0.1906	0.0969	0.0548
$TSRV_{adj}$	1.0295	0.5091	0.3188	0.1922	0.0973	0.0548

Tabelle 4.5: Vergleich verschiedener TSRV Schätzer

4.2.4 MSRV

Wie in Abbildung 4.7 und Tabelle 4.6 zu sehen ist, liegt die optimale Anzahl an Skalen bei 7. Der MSRV besteht durch seine gute Konvergenzrate von $n^{-\frac{1}{4}}$. Der MSE liegt dabei bei $4.4528e - 005$.

4 Empirische Studie

	Wert	M	MSE
$MSRV_5$	1.0650	5	4.7800e-005
$MSRV_6$	1.0497	6	4.4820e-005
$MSRV_7$	0.9902	7	4.4528e-005
$MSRV_8$	0.9449	8	4.5619e-005
$MSRV_9$	0.8998	9	4.7506e-005
$MSRV_{10}$	0.8575	10	4.9890e-005

Tabelle 4.6: Vergleich der verschiedenen Anzahl an Skalen beim MSR

MSE

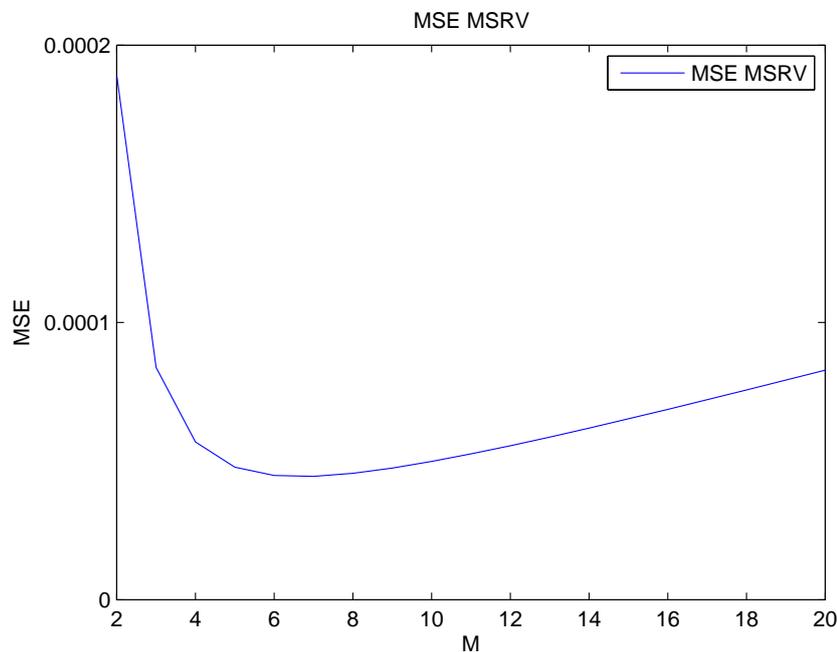


Abbildung 4.7: Der MSE der MSRV geplottet nach Anzahl der Skalen M

4.2.5 Kernel

Für den Realized Kernel Estimator betrachten wir sowohl Kernel mit Konvergenzrate von $n^{-\frac{1}{6}}$, als auch welche mit Konvergenzrate $n^{-\frac{1}{4}}$. Für den ersten Fall betrachten wir den Bartlett Kernel ($k(x) = 1 - x$), den Epanechnikov Kernel ($k(x) = 1 - x^2$) und den Second-Order Kernel ($k(x) = 1 - 2x + x^2$). Dabei hat der Bartlett Kernel dieselbe asymptotische Verteilung wie der TSRV. Der Second-Order Kernel ist der effizienteste unter den drei Kernels, gefolgt vom Bartlett Kernel und dem Epanechnikov Kernel. Für den zweiten Fall mit Konvergenzrate von $n^{-\frac{1}{4}}$ betrachten wir den Cubic Kernel ($k(x) = 1 - 3x^2 + 2x^3$), den Parzen Kernel ($k(x) = 1 - 6x^2 + 6x^3$ für $0 \leq x \leq \frac{1}{2}$ und $k(x) = 2(1-x)^3$ für $\frac{1}{2} \leq x \leq 1$)

4 Empirische Studie

und die Tukey-Hanning Kernels ($k(x) = \sin^2 \left\{ \frac{\pi}{2}(1-x)^p \right\}$) 1. und 2. Ordnung. Der Tukey-Hanning Kernel 1. Ordnung und der kubische Kernel verhalten sich fast ident. Der Parzen Kernel übertrifft den kubischen Kernel, kann aber mit dem Tukey-Hanning Kernel 2. Ordnung nicht mithalten.

Kernel	Wert	H^*	MSE
Bartlett	0.9925	5.8278	0.0029
2nd Order	0.9980	8.7418	0.0026
Epanechnikov	0.9119	6.2879	0.0050
Cubic	0.9533	7.4392	0.0010
Parzen	0.9662	9.6427	9.4596e-004
Tukey-Hanning ₁	0.9443	7.4796	0.0010
Tukey-Hanning ₂	0.8445	11.6036	9.2669e-004

Tabelle 4.7: Vergleich der verschiedenen Kernel Estimator

4.2.6 Fourier

Für den Fourier Estimator betrachten wir sowohl den Dirichlet-Kernel als auch den Fejer-Kernel. Die optimale Frequenz N_{cut} liegt bei 700. Daraus ergibt sich ein Beobachtungsabstand δ von 28 Sekunden. Der MSE liegt dabei bei $6.8218e - 004$.

Kernel	Wert	N	δ	MSE
Dirichlet	0.5322	700	28	6.8218e-004
Fejer	0.5275	700	28	6.8218e-004

Tabelle 4.8: Fourier Esitimator mit Dirichlet-Kernel und Fejer-Kernel

4 Empirische Studie

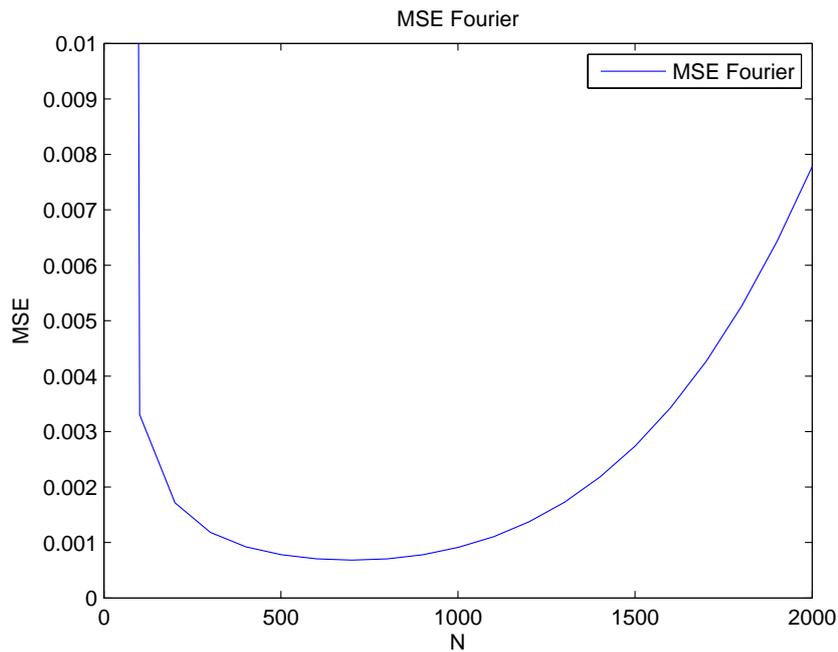


Abbildung 4.8: MSE-Fourier

4.3 Vergleich der verschiedenen Schätzer

Nachdem wir nun die Schätzer einzeln betrachtet haben, machen wir uns nun an einen Gesamtvergleich der einzelnen Schätzer.

Vergleich zwischen RV und Zhou Estimator

Wie in Abbildung 4.9 zu sehen ist, fällt der Zhou Estimator für geringe Beobachtungsfrequenzen q wesentlich schneller ab und bei größeren Samplingintervallen nähern sich die beiden Schätzer immer mehr einander an, da der Einfluss des Microstructure Noise immer geringer wird und somit die Autokovarianzen 1. Ordnung im Zhou Estimator an Bedeutung verlieren. Der Vergleich des MSE bei optimaler Samplingfrequenz geht aufgrund der Biaskorrektur klar zugunsten des Zhou Estimators aus, wie in Tabelle 4.9 zu sehen ist.

4 Empirische Studie

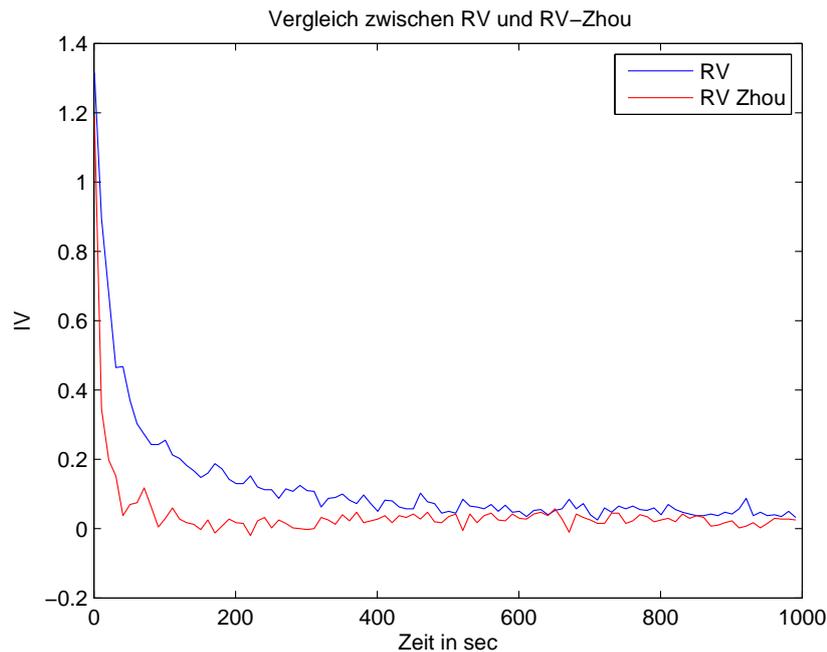


Abbildung 4.9: RV und Zhou Estimator im Vergleich

Verbesserungen des TSRV Schätzers gegenüber der RV

Im Vergleich zwischen RV mit dem TSRV Schätzer erkennen wir, dass der TSRV Schätzer die zwei Hauptprobleme der RV löst. Einerseits die Divergenz, im Fall von kleinen Abständen zwischen den Beobachtungen und andererseits die Verzerrung der RV unter Market Microstructure Noise.

Verbesserungen von MSRV gegenüber TSRV

Anstatt von zwei verschiedenen Skalen können auch mehrere verwendet werden. Beides sind konsistente Schätzer der integrierten Varianz, wobei der MSRV Schätzer die bessere Konvergenzrate ($n^{-\frac{1}{4}}$) im Gegensatz zu ($n^{-\frac{1}{6}}$) beim TSRV hat. Das Ganze geht natürlich zu Lasten des Rechenaufwands.

Gesamtvergleich

Im Gesamtvergleich liefern die konsistenten Schätzer, der MSRV Schätzer und der Kernel Estimator, die beste Performance. Einerseits empfehlen sich die beiden aufgrund ihrer hervorragenden Konvergenzraten von $n^{-\frac{1}{4}}$, andererseits durch ihre geringen MSE. Sowohl MSRV als auch Kernel Estimator mit geeignetem Kernel sind absolut zu empfehlen. Gefolgt werden die beiden vom ebenfalls konsistenten TSRV Schätzer mit Konvergenzrate von $n^{-\frac{1}{6}}$. Der Zhou Estimator hat zwar einen geringen MSE, ist aber aufgrund seiner

4 Empirische Studie

Verzerrung nicht unbedingt zu empfehlen.

In der Abwesenheit von Market Microstructure Noise liefert die RV einen effizienten Schätzer, sogar mit einer Konvergenzrate von $n^{-\frac{1}{2}}$. Da die logarithmischen Preise in der Praxis aber durch einen Noiseprozess gestört werden, ist dieser Schätzer in der Praxis verzerrt und somit nicht optimal.

Schätzer	Wert	MSE	q^*
$RV_{\text{sparse}}^{(opt)}$	0.2875	0.0011	70.4355
$RV_{\text{Zhou}}^{(opt)}$	1.0750	2.2985e-004	2.4253
$TSRV_{opt}$	1.0295	0.0013	5.8519
$MSRV_{opt}$	0.9902	4.4528e-005	7
2nd Order Kernel	0.9980	0.0026	8.7418
Tukey-Hanning ₂	0.8445	9.2669e-004	11.6036
Fourier	0.5322	6.8218e-004	28

Tabelle 4.9: Gesamtvergleich der verschiedenen Schätzer

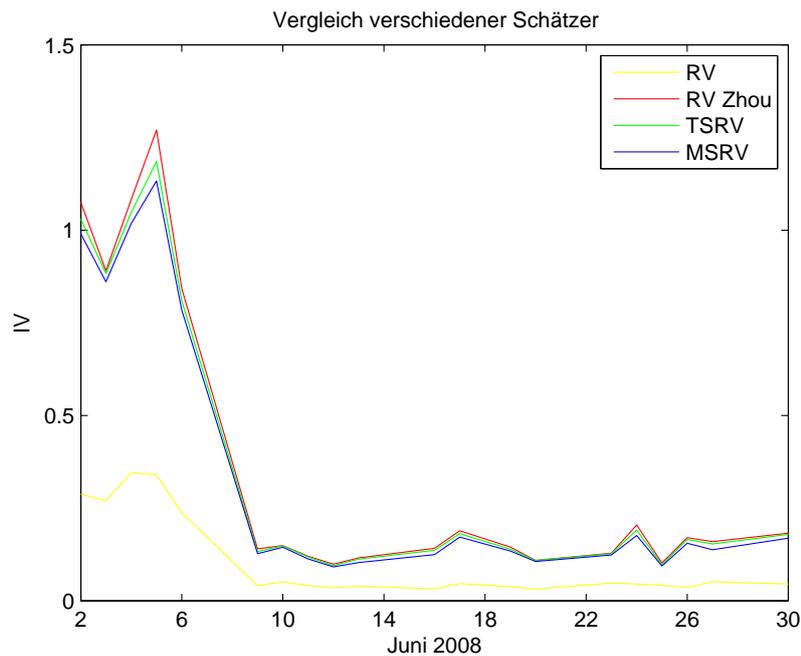


Abbildung 4.10: Vergleich verschiedener Volatilitätsschätzer über das Monat Juni 2008

Trade Prices vs. Midquote Prices

Eine weitere Möglichkeit ist die Schätzer anstatt mit Trade Prices, mit Midquote Prices zu ermitteln. Die Midquote Prices sind dabei weniger durch Market Microstructure Noise gestört, da gewisse Marktunstimmigkeiten sowie Bid-Ask Bounces wegfallen. Diesem Vorteil stehen aber eine große Anzahl an Nachteilen gegenüber. So ist die enthaltene Information in den Midquote Prices wesentlich geringer und somit ist es von Vorteil, Trade Prices unter einer guten Analyse des Noiseprozess zu verwenden. In den Tabellen 4.10 und 4.11 ist zu sehen, dass der Noiseprozess gemessen an den Midquote Prices etwa um den Faktor 10 geringer ist, als an den Trade Prices gemessen.

Datum	noise	noise _{midquote}
2.06.08	$3.32e - 005$	$2.56e - 006$
3.06.08	$3.16e - 005$	$1.09e - 006$
4.06.08	$3.52e - 005$	$5.18e - 006$
5.06.08	$4.27e - 005$	$4.67e - 006$
6.06.08	$2.80e - 005$	$4.67e - 006$
9.06.08	$4.70e - 006$	$4.82e - 007$
10.06.08	$5.25e - 006$	$6.34e - 007$
11.06.08	$4.07e - 006$	$4.65e - 007$
12.06.08	$3.30e - 006$	$2.85e - 007$
13.06.08	$4.07e - 006$	$4.29e - 007$
16.06.08	$5.27e - 006$	$3.56e - 007$
17.06.08	$6.79e - 006$	$2.76e - 006$
19.06.08	$4.74e - 006$	$4.77e - 007$
20.06.08	$3.70e - 006$	$1.26e - 006$
23.06.08	$4.21e - 006$	$5.10e - 007$
24.06.08	$7.06e - 006$	$4.19e - 007$
25.06.08	$3.02e - 006$	$4.60e - 007$
26.06.08	$6.70e - 006$	$1.32e - 006$
27.06.08	$6.72e - 006$	$1.07e - 006$
30.06.08	$6.91e - 006$	$3.89e - 007$

Tabelle 4.10: Der Noiseprozess gemessen einerseits an den Trade Prices und andererseits an den Midquote Prices

4 Empirische Studie

Datum	Trade Prices			Midquote Prices		
	$RV_{\text{sparse}}^{(70)}$	$RV_{\text{Zhou}}^{(2)}$	$TSRV^{(6)}$	$RV_{\text{sparse}}^{(42)}$	$RV_{\text{Zhou}}^{(1)}$	$TSRV^{(2)}$
2.06.08	0.2875	1.0750	1.0295	0.0575	0.1050	0.1050
3.06.08	0.2700	0.8900	0.8830	0.0256	0.0431	0.0431
4.06.08	0.3450	1.0825	1.0475	0.0388	0.2063	0.2062
5.06.08	0.3400	1.2700	1.1850	0.0625	0.1700	0.1700
6.06.08	0.2375	0.8425	0.8095	0.0613	0.0975	0.0763
9.06.08	0.0400	0.1400	0.1325	0.0117	0.0187	0.0187
10.06.08	0.0516	0.1488	0.1483	0.0147	0.0253	0.0253
11.06.08	0.0408	0.1200	0.1180	0.0064	0.0174	0.0174
12.06.08	0.0348	0.0984	0.0945	0.0057	0.0099	0.0099
13.06.08	0.0396	0.1156	0.1123	0.0074	0.0166	0.0166
16.06.08	0.0320	0.1416	0.1362	0.0085	0.0141	0.0141
17.06.08	0.0460	0.1888	0.1810	0.0138	0.1042	0.1042
19.06.08	0.0376	0.1448	0.1389	0.0091	0.0173	0.0173
20.06.08	0.0312	0.1088	0.1089	0.0148	0.0278	0.0278
23.06.08	0.0476	0.1284	0.1270	0.0100	0.0178	0.0178
24.06.08	0.0448	0.2044	0.1904	0.0136	0.0178	0.0178
25.06.08	0.0416	0.1020	0.0970	0.0068	0.0150	0.0150
26.06.08	0.0360	0.1700	0.1654	0.0483	0.0534	0.0534
27.06.08	0.0512	0.1596	0.1532	0.0228	0.0402	0.0402
30.06.08	0.0448	0.1824	0.1789	0.0116	0.0158	0.0158

Tabelle 4.11: Vergleich verschiedener Volatilitätsschätzer gemessen an den Trade- bzw. Midquote Prices

4.4 Zusammenhang zwischen der Gestalt des Limit Order Books und Volatilität³⁹

Nun wollen wir uns an die Untersuchung begeben, ob es einen Zusammenhang zwischen Volatilität und der Gestalt des Limit Order Book gibt. Es würde nahe liegen, dass eine höhere Volatilität zu erwarten ist, sofern die Limit Orders weiter weg vom Tradeprice liegen. Dabei betrachten wir den Abstand von Mittelwert, Median und Maximum der Verteilung des Limit Order Books. Diese rekonstruieren wir indem wir die Limit Orders in einem gewissen Zeitraum ansammeln.⁴⁰

³⁹in Zusammenarbeit mit David Hinterkörner

⁴⁰vgl. [VZFR13], S.6ff

4 Empirische Studie

Diese doch recht naheliegende Vermutung bestätigt sich jedoch auf den ersten Blick nicht. So erkennen wir schon auf den ersten Blick, dass im Vergleich 10 verschiedener Aktien des ISE-30-Index, kein linearer Zusammenhang besteht. Dies lässt sich auch durch Korrelationskoeffizienten und p-Wert bestätigen. Abbildung 4.11 vergleicht die Abstände mit diversen Volatilitätsschätzern und widerlegt ebenfalls unsere Vermutung. Weitere Versuche einen Zusammenhang zu Bid- oder Ask-Preis zu erkennen, haben leider auch zu keinem positivem Resultat geführt. So müssen wir daraus schließen, dass zumindest bei den verschiedenen Aktien des ISE-30-Index kein linearer Zusammenhang zwischen Volatilität und der Gestalt des Limit Order Books vorherrscht. Ein Grund dafür könnte sein, dass die Delta zu wenig Information über das Limit Order Book enthalten. Eine weitere umfangreichere Untersuchung, wie zum Beispiel mittels einer Principal Component Analyse (PCA), würde den Rahmen dieser Arbeit sprengen.

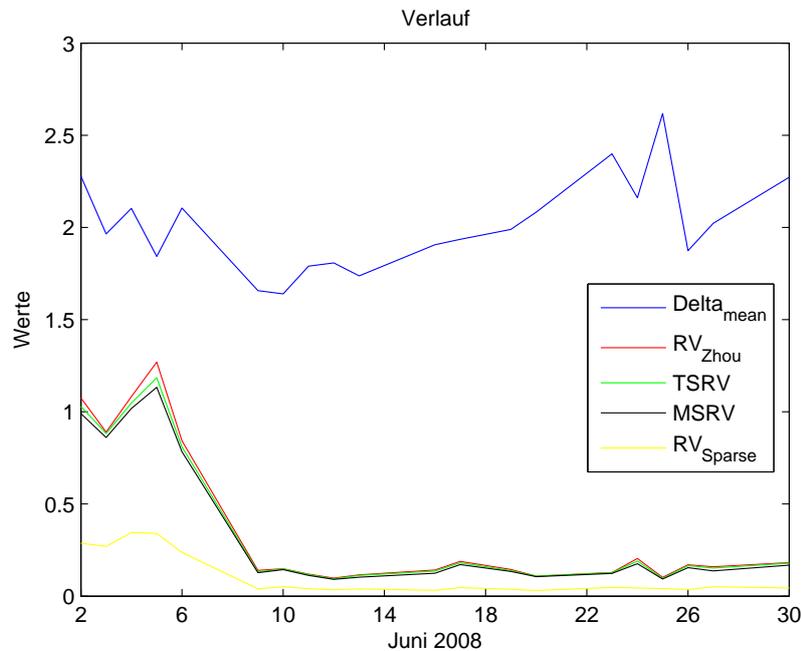


Abbildung 4.11: Vergleich zwischen dem Mittelwert der Delta und den verschiedenen Volatilitätsschätzern

4 Empirische Studie

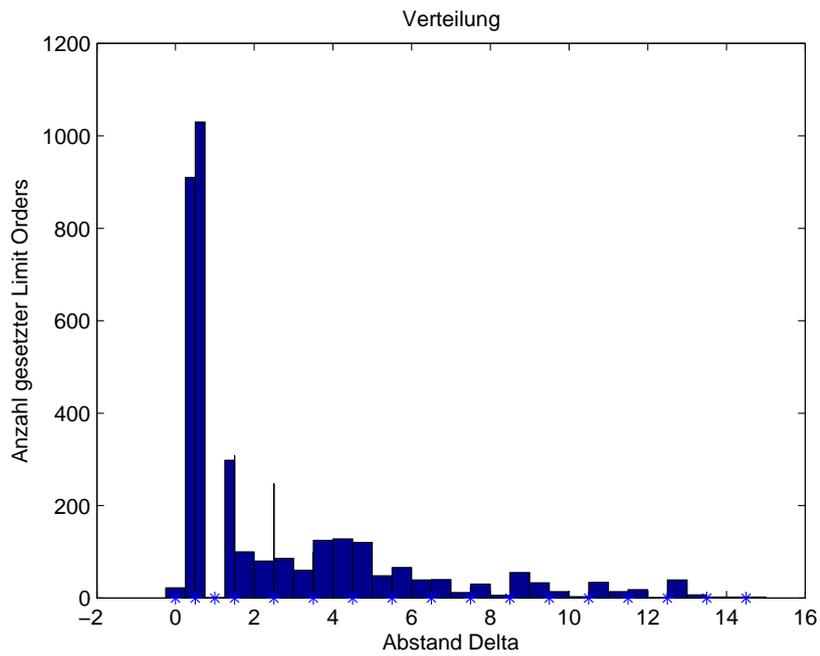


Abbildung 4.12: Verteilung der Delta des Stock GARAN

Stock	IV	Δ_{mean}
AKBNK	4.1410	2.5228
AKGRT	1.6207	3.1191
ARCLK	0.9882	4.0773
ASYAB	27.0774	3.0192
DOHOL	0.2445	3.1052
DYHOL	0.3132	3.0416
EREGL	13.3900	3.0177
GARAN	12.7298	2.0093
GSDHO	0.2528	2.2239
HALKB	4.6675	2.7768

Tabelle 4.12: Vergleich zwischen der IV für ein Monat und den kumulierten Deltas der verschiedenen Aktien, wobei die IV mit dem Zhou Estimator unter den jeweils optimalen Subsamplingfrequenzen berechnet wurde

5 Schlussfolgerung

Die Frage nach der *Volatilität* nimmt in der Finanzwirtschaft eine immer größere Bedeutung ein. Die Volatilität ist im Normalfall nicht messbar und daher müssen wir uns mit Schätzern für diese begnügen. Einige solcher Volatilitätsschätzer wurden in der vorliegenden Diplomarbeit vorgestellt.

Zuerst wird ein einfacher Schätzer für die integrierte Varianz, die so genannte *Realized Variance*, vorgestellt. Dieser Schätzer stößt jedoch sehr schnell an seine Grenzen, da er bei Auftreten von *Market Microstructure Noise* verzerrt und inkonsistent ist. Market Microstructure Noise ist ein unbeobachtbares Phänomen verursacht durch Bid-Ask Bounces und andere Marktunstimmigkeiten, welches die tatsächlichen Preise verunreinigt. Nun gibt es diverse Möglichkeiten, sich Abhilfe zu schaffen.

Wir können dieses Problem beheben, indem wir die Samplingfrequenz reduzieren, wobei jedoch der Nachteil darin besteht, dass teilweise bis zu 99% der Daten weggeworfen werden müssen. Dies ist aber nur wenig zufriedenstellend, da es das Ziel sein muss, möglichst alle vorhandenen Daten verwenden zu können. So müssen wir eine andere Möglichkeit finden, um dieses Phänomen zu beheben. Wir versuchen den Volatilitätsschätzer zu adaptieren und erhalten als erste Verbesserung den *Zhou Estimator*, einen leicht biasangepassten Schätzer für die Volatilität, der Autokovarianzen 1. Ordnung miteinbezieht. Bei der Implementierung ist es nach wie vor nicht optimal, alle Daten zu verwenden. Der Zhou Estimator ist zwar immer noch inkonsistent, die Resultate sind aber schon um einiges exakter.

Eine andere Herangehensweise ist, zwei Skalen zu verwenden um so mittels geeigneter Linearkombination zwischen Fast Time Scale und Slow Time Scale einen Schätzer zu erhalten. Dabei können nun erstmals alle Daten verwendet werden und wir erhalten auch bei geeigneter Wahl der Subsamplingfrequenz einen konsistenten Volatilitätsschätzer, den *Two Scaled Realized Volatility Estimator (TSRV)*. Zu Lasten des Rechenaufwands können wir auch über mehrere Skalen mitteln, dann sprechen wir von der *Multi Scale Realized Volatility Estimator (MSRV)*.

5 Schlussfolgerung

Eine weitere Klasse von Volatilitätsschätzern liefern die *Realized Kernel* Schätzer, welche durch eine hohe Anpassungsfreiheit bestechen. Mittels Fouriermethode erhalten wir noch einen weiteren Schätzer, den *Fourier Estimator*.

In der Praxis stellt sich nun die Frage nach dem am besten geeigneten Volatilitätsschätzer. Diese Frage ist leider nicht immer eindeutig zu beantworten. Die Auswahl des am besten geeigneten Schätzers hängt im Besonderen von der Qualität der vorhandenen Daten ab. Eine wesentliche Aufgabe und eine der Hauptschwierigkeiten bei der Implementierung von Volatilitätsschätzern ist die Wahl der optimalen Freiheitsparameter. Dabei sollten die Konvergenzrate und der Rechenaufwand auf keinen Fall außer Acht gelassen werden. Es gibt also in der Praxis meistens keinen am besten geeigneten Schätzer. Es hat zwar jeder einzelne seine Vor- und Nachteile, jedoch ist es ratsam, sich vor der Anwendung intensiv mit dem ausgewählten Schätzer zu beschäftigen. Es bedarf also einer großen Menge an Wissen, um vernünftige Schätzwerte für die Volatilität zu erhalten.

Da die Volatilität als einziger Parameter des Black-Scholes-Modells am Markt nicht beobachtbar ist, werden Volatilitätsschätzer weiter an Bedeutung gewinnen. Die Verbesserung der bereits eingesetzten Methoden sowie die Entwicklung neuer Schätzer für die Volatilität wird auch in Zukunft weiterhin ein interessantes Forschungsgebiet bleiben und somit in der Praxis eine wichtige Grundlage im Finanzbereich, insbesondere bei der Bepreisung von Optionen, bilden.

6 Programmcodes

- Programm zur Berechnung der Preise zu gegebenen Beobachtungsintervall aus dem Datensatz

```
1 function ausgabe = price_int_adpt(delta,n,j)
2 % delta...Zeitraum in Sekunden zwischen den Beobachtungen
3 % n... Sekunden bis zur letzten Beobachtung
4 % j...Startzeitpunkt
5 load dyLOB08061
6 t = [1:396040]';
7 timeprice=[t,dyLOB08061(1,8).dyLOB(:,24)];
8
9 c=1;
10 for i = j:delta:n+j-1
11     ausgabe(c) = timeprice(i,2);
12     c=c+1;
13 end
```

- Programm zur Berechnung der Midquoteprices

```
1 function ausgabe = midquote_adpt(delta,n,j)
2 % delta...Zeitraum in Sekunden zwischen den Beobachtungen
3 % n... Sekunden bis zur letzten Beobachtung
4 % j...Startzeitpunkt
5 load dyLOB08061
6 t = [1:396040]';
7 bid = dyLOB08061(1,8).dyLOB(:,15);
8 ask = dyLOB08061(1,8).dyLOB(:,16);
9 Y=[t,bid,ask];
10 c=1;
11 for i = j:delta:n+j-1
12     ausgabe(c) = (Y(i,2) + Y(i,3))/2;
13     c=c+1;
14 end
```

6 Programmcodes

- Programm zur Berechnung des Noiseestimators

```
1 function R = noise_est_adpt(n,j)
2 % n... Sekunden bis zur letzten Beobachtung
3 % j...Startzeitpunkt
4
5 R = RV_adpt(1,n,j)/(2*n);
```

- Programm zur Schätzung der IV

```
1 function R = IV_zero(n);
2 % n... Sekunden bis zur letzten Beobachtung
3
4 delta = 1;
5 j = 1;
6 M = 300;
7 q_quer = floor(n/M);
8 Y = price_int_adpt(delta,n,j);
9 R = 0;
10
11 for h=0:q_quer-1
12     m = floor((n-h)/q_quer);
13     for i=1:m-1
14         R = R + (Y((i+1)*q_quer + h) - Y(i*q_quer + h))^2;
15     end
16 end
17
18 R = n/(n-q_quer+1)*(1/q_quer)*R;
```

- Programm zur Schätzung der IQ

```
1 function R = IQ_zero(n);
2 % n... Sekunden bis zur letzten Beobachtung
3
4 delta = 1;
5 j = 1;
6 M = 300;
7 q_quer = floor(n/M);
8 Y = price_int_adpt(delta,n,j);
9 R = 0;
10
11 for h=0:q_quer-1
12     m = floor((n-h)/q_quer);
```

6 Programmcodes

```
13     for i=1:m-1
14         R = R + (Y((i+1)*q_quer + h) - Y(i*q_quer + h))^4;
15     end
16 end
17
18 R = (n^2/((n-q_quer+1)^2))*(78/(3*q_quer))*R;
```

- Programm zur Berechnung der Varianz des Microstructure Noise

```
1 function R = noise_var(n,j)
2 % n... Sekunden bis zur letzten Beobachtung
3 % j...Startzeitpunkt
4 delta = 1;
5 Y = price_int_adpt(delta,n,j);
6
7 R = 0;
8 for i=1:length(Y)-1
9     R = R + (Y(i+1)-Y(i))^4;
10 end
11
12 R = 1/(2*n)*R - 3*(noise_est_adpt(19802,1))^2;
```

- Programm zur Berechnung der RV

```
1 function R = RV_adpt(delta,n,j)
2 % delta...Zeitraum in Sekunden zwischen den Beobachtungen
3 % n... Sekunden bis zur letzten Beobachtung
4 % j...Startzeitpunkt
5 Y = price_int_adpt(delta,n,j);
6
7 R = 0;
8 for i=1:length(Y)-1
9     R = R + (Y(i+1)-Y(i))^2;
10 end
```

6 Programmcodes

- Programm zur Berechnung der optimalen Samplingfrequenz der RV

```
1 function R = n_opt(n)
2 % n... Sekunden bis zur letzten Beobachtung
3
4 x = (2*noise_est(n)/IV_zero(n))(-2/3);
5
6 R = n/x;
```

- Programm zur Berechnung des RV_{Zhou}

```
1 function R = RV_Zhou(delta,n,j)
2 % delta...Zeitraum in Sekunden zwischen den Beobachtungen
3 % n... Sekunden bis zur letzten Beobachtung
4 % j...Startzeitpunkt
5 Y = price_int_adpt(delta,n,j);
6
7 R = 0;
8 for i=1:length(Y)-3
9     R = R + (Y(i+2)-Y(i+1))2 + (Y(i+2)-(Y(i+1)))*(Y(i+1)-Y(i))
10        + (Y(i+2)-Y(i+1))*(Y(i+3)-Y(i+2));
11 end
```

- Programm zu Berechnung der optimalen Samplingfrequenz des RV_{Zhou}

```
1 function R = n_opt_Zhou(n)
2 % n... Sekunden bis zur letzten Beobachtung
3 y = sqrt(3)/(2*(noise_est(n)/IV_zero(n)));
4
5 R = n/y;
```

- Programm zur Berechnung des RV_{avg}

```
1 function R = RV_avg(n,k)
2 % n... Sekunden bis zur letzten Beobachtung
3 % k... Anzahl der Teilnetze
4 R = 0;
5 for i=1:k
6     R = R + RV_adpt(k,n-k,i);
7 end;
8 R = R/k;
```

6 Programmcodes

- Programm zur Berechnung des TSRV

```
1 function R = TSRV_adpt(n,k,j)
2 % n... Sekunden bis zur letzten Beobachtung
3 % j... Startzeitpunkt
4 % k... Frequenz Slow Time Scale
5 fast = RV_adpt(1,n,j);
6 slow = RV_avg_adpt(n,k,j);
7 n_quer=(n-k+1)/k;
8
9
10 R = slow - n_quer/n * fast;
```

- Programm zur Berechnung des angepassten TSRV

```
1 function R = TSRV_adj(n,k,j)
2 % n... Sekunden bis zur letzten Beobachtung
3 % j... Startzeitpunkt
4 % k... Frequenz Slow Time Scale
5 n_quer=(n-k+1)/k;
6
7 R = 1/(1-n_quer/n)*TSRV_adpt(n,k,j);
```

- Programm zur Berechnung der optimalen Subsamplingfrequenz des TSRV

```
1 function R = K_opt(n)
2
3 R = ((12* (noise_est(n))^2)/IQ_zero(n))^(1/3)*n^(2/3);
```

- Programm zur Berechnung des MSRV

```
1 function R = MSRV(n,M)
2 % n... Sekunden bis zur letzten Beobachtung
3 % M... Anzahl der Skalen
4 k = [1:M];
5 R = 0;
6
7 for i=1 : M
8     a = k(i);
9     R = R + ani(M,i)*RV_adpt(a,n,1);
10 end
```

6 Programmcodes

- Programm zur Berechnung der optimalen Gewichte

```
1 function R = ani(M,i)
2
3 R = 12*(i/M^2)*((i/M-1/2-1/(2*M))/(1-1/M^2));
```

- Programm zur Berechnung der Autokovarianzen für den Kernel Estimator

```
1 function R = autokov1(h,H,delta,n)
2 % h... Ordnung
3 % H... Anzahl der verwendeten Daten für Kernel Estimator
4 % delta...Zeitraum in Sekunden zwischen den Beobachtungen
5 % n... Sekunden bis zur letzten Beobachtung
6
7 t=floor(n/delta);
8 Y = price_int_adpt(1,n,1);
9 R= 0;
10
11 for i=2:delta:n-H
12     R = R + (Y(i)-Y(i-1))*(Y(i-h)-Y(i-h-1));
13 end
```

- Programm zur Berechnung des Kernel Estimator

```
1 function R = kernel_est(delta,n,H)
2 % H ... Anzahl der verwendeten Daten für Kernel Estimator (
3   entspricht in der Theorie 2H)
4 % delta...Zeitraum in Sekunden zwischen den Beobachtungen
5 % n ... Sekunden bis zur letzten Beobachtung
6 % H_0 ... adaptierte Zeitpunkt 0
7 j=1;
8 H_0=floor(H/2);
9 Y = price_int_adpt(delta,n,j);
10
11 %kernel = @(x) 1-x; %Bartlett
12 %kernel = @(x) 1-2*x+x^2; %2nd Order
13 %kernel = @(x) 1-x^2; %Epanechnikov
14 %kernel = @(x) 1-3*x^2+2*x^3; %Cubic
15 %kernel = @(x)
16 %kernel = @(x) (sin(pi/(2*(1-x))))^2; %Tukey-Hanning
17 %kernel = @(x) (sin(pi/(2*(1-x))))^2; %Tukey-Hanning2
18 %kernel = @(x) (1+x)*exp(-x); %Optimal
```

6 Programmcodes

```
19 %kernel = @(x) sin(x)/x; %Dirichlet
20 %kernel = @(x) (sin(x)/x)^2; %Fejer
21
22
23 t=n;
24 R=0;
25
26 for i=1:H
27     %if ((i-1)/H) <= 1/2      %Parzen Kernel
28         % kernel = @(x) 1-6*x^2+6*x^3;
29     %else
30         %kernel = @(x) 2*(1-x)^3;
31     %end
32     R = R + kernel((i-1)/H)*(autokov(i,H,delta,n) + autokov1(-i
33         ,H,delta,n));
34
35 R = R + autokov(0,1,delta,n);
```

- Programm zur Berechnung des optimalen H bei Konvergenzrate $n^{-\frac{1}{6}}$

```
1 function R = H_opt(n,c)
2 % n ... Sekunden bis zur letzten Beobachtung
3 % c... Konstante aus Tabelle
4 zeta2 = noise_est(n)/sqrt(IQ_zero(n))
5
6 R = c * (zeta2*n)^(2/3);
```

- Programm zur Berechnung des optimalen H bei Konvergenzrate $n^{-\frac{1}{4}}$

```
1 function R = H_opt2(n,c)
2 % n ... Sekunden bis zur letzten Beobachtung
3 % c... Konstante aus Tabelle
4 zeta2 = noise_est(n)/sqrt(IQ_zero(n));
5
6 R = c * sqrt(zeta2)*sqrt(n)
```

6 Programmcodes

- Programm zur Berechnung des Fourierestimator

```
1 function R = fourier_est(n,N,delta)
2 % n... Sekunden bis zur letzten Beobachtung
3 % N... Anzahl der Frequenzen
4 % delta...Zeitraum in Sekunden zwischen den Beobachtungen
5
6 x = floor(n/delta);
7 j=1;
8 Y = price_int_adpt(delta,n,j);
9 R = 0;
10
11 DN = @(t) (sin((N+1/2)*t))/((2*N+1)*sin(t/2));
12 %FN = @(t) (1-cos(N*t))/(2*(N+1)*N*(sin(t/2))^2);
13
14 for i=2:x
15     for j=2:x
16         if i == j
17             R = R + (Y(i)-Y(i-1))*(Y(j)-Y(j-1));
18         else
19             R = R + DN((2*pi*(i-j))/x)*(Y(i)-Y(i-1))*(Y(j)-Y(j-1));
20         end
21     end
22 end
23 end
```

- Programm zur Berechnung des MSE der RV

```
1 function R = MSE_RV(n,q)
2 % n... Sekunden bis zur letzten Beobachtung
3 % q... Frequenz
4
5 n_quer = floor(n/q);
6 lambda = noise_est(n)/IV_zero(n);
7
8 R = IV_zero(n)^2*(4*lambda^2*n_quer^2 + 12*lambda^2*n_quer + 8*
    lambda - 4*lambda^2 + 2/n_quer);
```

6 Programmcodes

- Programm zur Berechnung des MSE des RV_{Zhou}

```
1 function R = MSE_Zhou(n,q)
2 % n... Sekunden bis zur letzten Beobachtung
3 % q... Frequenz
4
5 n_quer = floor(n/q);
6 lambda = noise_est(n)/IV_zero(n);
7
8 R = IV_zero(n)^2*(8*lambda^2*n_quer+8*lambda-6*lambda^2+6/
   n_quer-2/n_quer^2);
```

- Programm zur Berechnung des MSE des TSRV

```
1 function R = MSE_TSRV(n,K)
2 % n... Sekunden bis zur letzten Beobachtung
3 % K... Frequenz Slow Time Scale
4
5 c = K/n^(2/3);
6 eta = sqrt(4/3*IQ_zero(n));
7 T = 1;
8
9 R = (8*(1/c^2)*noise_est(n)^2+ eta^2*T)/n^(1/3);
```

- Programm zur Berechnung des MSE des MSRV

```
1 function R = MSE_MSrv(n,M)
2 % n... Sekunden bis zur letzten Beobachtung
3 % M... Anzahl der Skalen
4
5 c = M/sqrt(n);
6 T = 1;
7 eta2 = 2*IQ_zero(n);
8
9 R = (48/c^3*(noise_est_adpt(19802,1))^2+52/35*c*T*eta2+12/(5*c)
   *noise_var(n,1)+48/(5*c)*noise_est_adpt(n,1)*IV_zero(n))/(n
   ^ (1/2));
```

6 Programmcodes

- Programm zur Berechnung des MSE des Kernel Estimator mit Konvergenzrate $n^{-\frac{1}{6}}$

```

1 function R = MSE_Kernel_16(n,k_00,k_strich0,k_strich1,c)
2
3 zeta2 = noise_est(n)/sqrt(IQ_zero(n));
4 T = 1;
5
6 R = (4*T*IQ_zero(n)*c*k_00+1/(c^2)*(k_strich0^2+k_strich1^2)*
      zeta2^2)/n^(1/3);

```

- Programm zur Berechnung des MSE des Kernel Estimator mit Konvergenzrate $n^{-\frac{1}{4}}$

```

1 function R = MSE_Kernel(n,k_00,k_11,k_22,c)
2
3 rho = IV_zero(n)/sqrt(IQ_zero(n));
4 zeta2 = noise_est(n)/sqrt(IQ_zero(n));
5 T = 1;
6
7 R = (4*T*IQ_zero(n)*c*k_00+1/c*2*k_11*rho*zeta2+1/c^3*k_22*
      zeta2^2)/n^(1/2);

```

- Programm zur Berechnung des MSE des Fourierestimator

```

1 function R = MSE_Fourier(n,N)
2 % n... Sekunden bis zur letzten Beobachtung
3 % N... Anzahl der Frequenzen
4
5 x = 19802;
6
7 D = @(t) (sin((N+1/2)*t))/((2*N+1)*sin(t/2));
8 alpha = 4*(noise_est_adpt(x,1))^2*(1+(D((2*pi)/n))^2-2*D((2*pi)
      /n));
9 beta = 4*noise_var(x,1)*(1+(D((2*pi)/n))^2-2*D((2*pi)/n));
10 gamma = 8*noise_est_adpt(x,1)*IV_zero(x) + 2*(noise_est_adpt(x
      ,1))^2 - 2*noise_var(x,1) + 4*(4*(noise_est_adpt(x,1))^2+4*
      noise_var(x,1))*(2*D((2*pi)/n)-(D((2*pi)/n))^2);
11
12 R = 4*pi/n*IQ_zero(x)+ (2*pi)/(2*N+1)*4*IQ_zero(x) + n^2*alpha
      + n*beta + gamma ;

```

Literaturverzeichnis

- [AMZ05] AÏT-SAHALIA, Y., MYKLAND, P. und ZHANG, L.: *How often to sample a continuous-time process in the presence of market microstructure noise*. Rev. Financ. Stud., 18,:351–416, 2005.
- [ASMZ08] AÏT-SAHALIA, Y., P. MYKLAND und L. ZHANG: *Ultra high frequency volatility estimation with dependent microstructure noise*. J. Econom., 2008.
- [BMV12] BANNOUH, K., MARTENS, M. und VAN DIJK, D.: *Forecasting Volatility with the Realized Range in the Presence of Noise and Non-Trading*. ERIM Report Series, 2012.
- [BNHLS08] BARNDORFF-NIELSEN, O.E., P.R. HANSEN, A. LUNDE und N. SHEPHARD: *Designing realised kernels to measure the ex post variation of equity prices in the presence of noise*. Econometrica, 76:1481–1536, 2008.
- [BRY08] BANDI, F.M., J.R. RUSSELL und C. YANG: *Realized volatility forecasting and option pricing*. J. Econom., 147:34–46, 2008.
- [BZ05] BRZEZNIAK, Z. und T. ZASTAWNIAK: *Basic Stochastic Processes*. Springer, London, 7. Auflage, 2005.
- [GR10] GATHERAL, J. und OOMEN R.: *Zero-intelligence realized variance estimation*. Finance Stoch., 14:249–283, 2010.
- [Gra12] GRANDITS, P.: *Vorlesungsmitschrift Stochastische Kontrolltheorie*, 2012.
- [Gur11] GURKER, W.: *Angewandte Statistik*, 2011.
- [HL06] HANSEN, P.R. und A. LUNDE: *Realized variance and market microstructure noise*. Journal of Business and Economic Statistics, 24:127–218, 2006.
- [Lan14] *Landau-Symbole*. <http://de.wikipedia.org/wiki/Landau-Symbole>, Wikipedia, August 2014.
- [MM02] MALLIAVIN, P. und M.E. MANCINO: *Fourier series method for measurement of multivariate volatilities*. Finance Stoch., 6:49–61, 2002.

Literaturverzeichnis

- [MM09] MALLIAVIN, P. und M.E. MANCINO: *A Fourier transform method for non-parametric estimation of multivariate volatility*. Ann. Stat., 37:1983–2010, 2009.
- [MS08] MANCINO, M.E. und S. SANFELICI: *Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise*. Comput. Stat. Data Anal., 52:2966–2989, 2008.
- [Oks00] OKSENDAL, B.: *Stochastic Differential Equations*. Springer, Heidelberg New York, Fifth Edition, Corrected Printing Auflage, 2000.
- [Sem14] *Semimartingal*. <http://de.wikipedia.org/wiki/Semimartingal>, Wikipedia, August 2014.
- [Vol14a] *Financial Market Volatility Definitions*. <http://staff.science.uva.nl/~marvisse/volatility.html>, University of Amsterdam, August 2014.
- [Vol14b] *Volatilität*. <http://de.wikipedia.org/wiki/Volatilitaet>, Wikipedia, August 2014.
- [Vol14c] *Volatility (finance)*. [http://en.wikipedia.org/wiki/Volatility_\(finance\)](http://en.wikipedia.org/wiki/Volatility_(finance)), Wikipedia, August 2014.
- [VZ13] VALENZUELA, M. und I. ZER: *Competition, signaling and non-walking through the book: Effects on order choice*. Journal of Banking and Finance, 37:5421–5435, 2013.
- [VZFR13] VALENZUELA, M., I. ZER, P. FRYZLEWICZ und T. RHEINLÄNDER: *Relative Liquidity and Future Volatility*. LSE Statistics, 2013.
- [Zha06] ZHANG, L.: *Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach*. Bernoulli, 12:1019–1043, 2006.
- [Zho96] ZHOU, B.: *High frequency data and volatility in foreign-exchange rates*. J. Bus. Econ. Stat., 14:45–52, 1996.
- [ZMAS05] ZHANG, L., P.A. MYKLAND und Y. AÏT-SAHALIA: *A tale of two timescales: Determining integrated volatility with noisy high frequency data*. J. Am. Stat. Assoc., 100:1394–1411, 2005.