Project Proposals MatØk 9th Semester: Fall 2022

Orimar Sauri Toke C. Zinn

The following 4 project proposals from the supervisors are given as suggestions in a random (non-prioritized) order. It is allowed that the same project is written by more than one group. Note, however, that you can put forward your suggestions for discussion.

1 High-Frequency analysis of Stochastic Volatility and its applications to risk management

1.1 Motivation

Most of the risk measures used in the financial sector are typically described as functions of the volatility, e.g. the so-called integrated volatility, and the beta risk factor of a given portfolio. Moreover, the full-understanding of the volatility is crucial for pricing and hedging financial derivatives. The standard model for the behavior of the log-prices is given by the continuous Itô's semimartingale

$$X_t := \log S_t = \log S_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s, \quad t \geq 0, \tag{1}$$

in which μ and σ are random processes. Here the diffusion coefficient (σ) plays the role of the (spot) volatility in the price.

1.2 Description of the project

The main goal of this project consists on the *non-parametric estimation* of functionals of the so-called *integrated volatility*

$$C_t := \langle X \rangle_t = \int_0^t c_s ds, \quad c_t := \sigma_t \sigma_t', \quad t \geq 0,$$

and its *key role in risk management*.

The estimation procedure will be performed under a *high-frequency set-up*: We aim to estimate C_t based on equidistant observations of the process X on $[0, T]$, say

$$X_0, X_{\Delta_n}, X_{2\Delta_n}, \ldots, X_{(n-1)\Delta_n}, X_T.$$

Here $T > 0$ is a fixed finite trading horizon, and $\Delta_n > 0$ is a sequence such that $\Delta_n \downarrow 0$ as $n \to \infty$. To do this we concentrate on the asymptotic properties of some of the fundamental estimators known in the literature:

1. **Realized Covariation**:

$$V_t^{n,1} := \sum_{i=1}^{[t/\Delta_n]} \Delta_i^n X \Delta_i^n X', \;\; t \geq 0,$$

where $\Delta_n \downarrow 0$ as $n \to \infty$, X as in (1) and $\Delta_i^n X := X_{i\Delta_n} - X_{(i-1)\Delta_n}$.

2. **Spot Volatility:** For large sample size, the statistic $V_t^{n,1}$ is approximately C_t. However, in some applications (see below), it is required to estimate not only C_t but also c_t. This is can be done by "numerically differentiating" $V_t^{n,1}$: Let $k_n \in \mathbb{N}$ be such that $k_n \to \infty$ and $k_n\Delta_n \to 0$ and put

$$\hat{c}_t := \frac{V_{t+k_n\Delta_n} - V_t}{k_n\Delta_n} = \frac{1}{k_n\Delta_n}\sum_{i=1}^{k_n} \Delta_{i+j}^n X \Delta_{i+j}^n X', \;\; (j-1)\Delta_n < t \leq j\Delta_n. \qquad (2)$$

3. **The Pre-average method**. High-Frequency financial data tends to be contaminated by microstructure noise. By pre-averaging assets returns one can get rid of the noise:

$$V_t^{n,2} := \sum_{i=1}^{[t/\Delta_n]-k_n+1} \bar{X}_i^n \bar{X}_i^{n\prime}, \;\; t \geq 0,$$

where for a function $g : [0,1] \to \mathbb{R}$ and a sequence $k_n \uparrow \infty$, we have let

$$\bar{X}_i^n = \sum_{j=1}^{k_n-1} g\left(\frac{j}{n}\right) \Delta_{i+j-1}^n X.$$

4. **The Hayashi-Yoshida estimator**. Liquidity has a big impact in the performance of the preceding estimators: If one of the assets is not trade as often as the others, the number of observations used to compute e.g. the realized covariance will drop drastically. The pre-averaged Hayashi-Yoshida estimator alleviates this problem by considering

$$V_t^{n,3}(k,l) := \sum_{i,j} \bar{X}_i^{n,k} \bar{X}_i^{n,l} \mathbf{1}_{(0,t]}(t_{i+k_n}^k \vee t_{j+k_n}^l)\mathbf{1}_{I_i^k \cap I_j^l \neq \emptyset}, \;\; t \geq 0,$$

where t_i^k denotes the ith observation time of the kth asset and

$$I_i^k := (t_i^k, t_{i+k_n}^k], \;\; i = 1, \ldots, n_k, k = 1, \ldots, d.$$

A second goal of the project includes the use of some of the previous estimators for risk management purposes in *one of the following topics*:

Portfolio Allocation with risk measures

If X is as in (1) represents the log-price of d risky assets, then under suitable conditions, the daily log-returns $r_T := X_T - X_{T-1}$ satisfy for $T = 1, \ldots, N$

$$\mu_{T|T-1} := \mathbb{E}(r_T \mid \mathcal{F}_{T-1}) = \int_{T-1}^{T} b_s ds,$$

and

$$\Sigma_{T|T-1} := \mathrm{Var}(r_T \mid \mathcal{F}_{T-1}) = \int_{T-1}^{T} c_s ds,$$

where $(\mathcal{F}_t)_{t\geq 0}$ is a filtration that contains the market's available information. Thus, if one wishes to create a portfolio at time $T-1$ using:

1. *A conditional mean-variance criterion* one must solves the optimization problem

$$\min_{w\in\mathbb{R}^d} w'\Sigma_{T|T-1}w \text{ subject to } \sum w_i = 1,\ w'\mu_{T|T-1} = \mu,$$

   where $w \in \mathbb{R}^d$ represents the weights of a given portfolio.

2. The *conditional value at risk* (also known as expected short-fall) associated to the portfolio with returns $w'r_T$ can be computed as

$$\mathrm{CVaR}_\alpha(w'r_T \mid \mathcal{F}_{T-1}) = w'\mu_{T|T-1} - w'\Sigma_{T|T-1}wc_\alpha,$$

   where c_α is a constant depending on α. In this situation one must solves the optimization problem

$$\min_{w\in\mathbb{R}^d} \mathrm{CVaR}_\alpha(w'r_T \mid \mathcal{F}_{T-1}) \text{ subject to } \sum w_i = 1,\ w'\mu_{T|T-1} = \mu.$$

This makes clear that the estimation and forecasting of C_t is crucial in this framework.

Hedging with implied volatility and high-frequency data

The arbitrage-free price, say Π_t, of a European derivative with pay-off $\phi(S_T)$ at maturity $T > 0$ can be written using the Black-Scholes option formula and the *cumulated implied volatility* as

$$\Pi_t = C(S_t, r(T-t), \Xi_t),$$

where Ξ_t is the derivative's implied volatility and

$$C(S,R,v) = e^{-R}\mathbb{E}\left[\phi(S\exp(R - v/2 + \sqrt{v}Z))\right],\ \ Z \sim N(0,1).$$

If (1) holds, or equivalently

$$\frac{dS_t}{S_t} = \tilde{\mu}_t dt + \sigma_t dB_t,$$

Itô's formula implies that the process $(\Xi_t)_{t\geq 0}$ is also a continuous Itô's semimartingale. [Zhang, 2012] showed that if the implied volatility satisfies that

$$\Xi_t = \Xi_0 + \int_0^t \rho_s dS_s + \int_0^t H_s ds,$$

then Π_t can be *perfectly replicated* by

$$\Delta_t = \partial_s C(S_t, r(T-t), \Xi_t) + \rho_t \partial_v C(S_t, r(T-t), \Xi_t),$$

i.e. almost surely

$$\Pi_t = \Pi_0 + \int_0^t \Delta_s dS_s.$$

In practice, an agent that wishes to *cover against a short position on the derivative* needs to estimate and forecast ρ_t. This can be done by the relation

$$\rho_t = \frac{\langle \Xi, S\rangle_t'}{(S_t\sigma_t)^2}.$$

Hence, ρ_t can be estimated using the spot-volatility estimator in (2).

Beta risk factor

A popular way to asses the exposure to systematic risk (or volatility) is via the so-called *Beta risk factor*. The *CAPM* (in its continuous-time version) or more generally any continuous-time *Factor Model,* asserts that a well-diversified portfolio with log-price Y admits the representation

$$dY_t = \beta dX_t + dF_t,$$

where the factor X represents a set of market-portfolios while F is the so-called *unsystematic risk* and satisfies that

$$\langle X, F \rangle = 0.$$

Since

$$\beta = C_t^{-1} \langle X, Y \rangle_t \, ,$$

the parameter β measures the proportion of volatility (risk) of Y relative to the risk in the market's portfolio X. Investors typically seek for portfolios for which β is small. One can efficiently estimate β using any of the estimators described above and thus asses the risk associated to Y against market's risk.

1.3 Data and references

The main references for the theoretical part are:

- [Aït-Sahalia and Jacod, 2014] and [Jacod and Protter, 2011].

They can be complemented with

- [Boudt et al., 2021], [Barndorff-Nielsen and Shephard, 2004], [Christensen et al., 2010],

For the applications

- [Bauwens et al., 2012], [Wang and Cheng, 2022], [Nadarajah et al., 2014], [Zhang, 2012], [Mykland and Zhang, 2008], [Boudt et al., 2017].

High-Frequency data is available at https://www.histdata.com/.

2 Pricing under Rough Volatility

2.1 Motivation

The Black and Scholes option's implied volatility of a European call option with time to expiration $\tau = T - t$ and log strike $k = \log(K/S_t)$ is defined as the unique (random) number $\sigma_{BS}(k, \tau) > 0$ such that

$$\Pi_t = C_{BS}(S_t, r\tau, \sqrt{\tau}\sigma_{BS}(k, \tau)),$$

where Π_t denotes the market price of the option, and C_{BS} is the option's price under the Black and Scholes model, i.e.

$$C_{BS}(S, R, v) = e^{-R}\mathbb{E}\left[(S\exp(R - v/2 + \sqrt{v}Z) - K)^+\right], \quad Z \sim N(0, 1).$$

It is well known that for a single expiration, the function $k \mapsto \sigma_{BS}(k, \tau)$ generates the

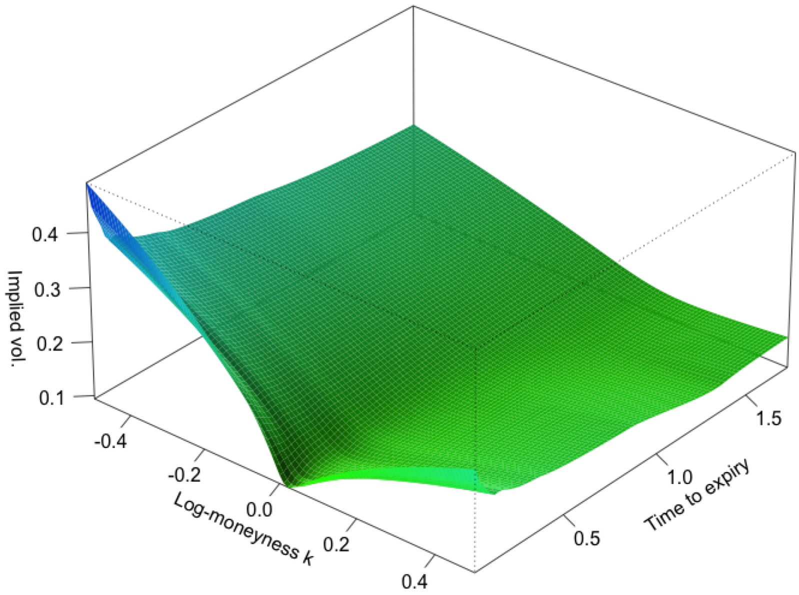

Figure 1: Volatility Surface. Source [Bayer et al., 2016].

so-called volatility smile (see Figure 1). Despite that this stylized fact is consistent with the efficient price model

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dB_t, \tag{3}$$

where σ is a stochastic process, little can be said from the volatility surface about the nature of the volatility process σ_t. Thus, in order to get a better understanding of the latter, deeper properties of the smiles need to be studied.

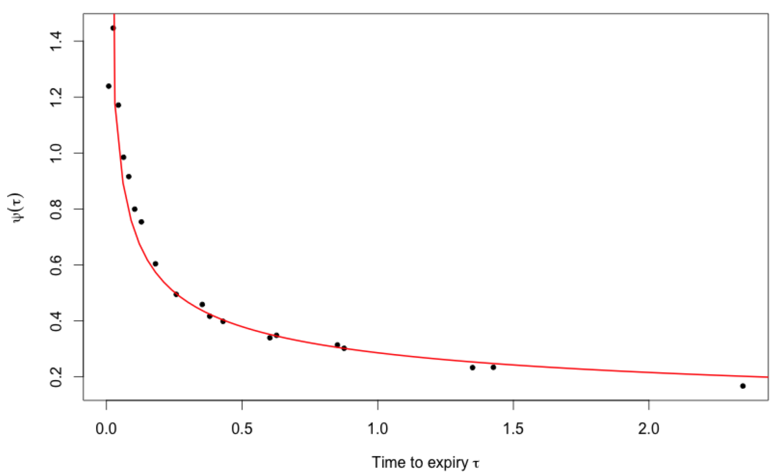

Figure 2: Non-parametric estimates of the S&P at-the-money (ATM) volatility skew. Source [Bayer et al., 2016].

One of the most unique characteristics of volatility smiles is the so-called *at-the-money*

(ATM) volatility skew: For some $0 < H < 1/2$, the function

$$\psi(\tau) = |\partial_k \sigma_{BS}(0, \tau)| .$$

satisfies the power law

$$\psi(\tau) \propto \tau^{H-1/2}.$$

Figure 2 shows an estimate of this function as well a fitted curve with $H \approx 0.1$. Unfortunately most of the classical stochastic volatility models where volatility is modelled as a diffusion (like the Heston model and local volatility models) cannot reproduce this property. On the other, several studies ([Gatheral et al., 2018], [Bennedsen et al., 2021] [Bolko et al., 2022]) have found evidence that volatility is well described by processes driven by a fractional Brownian motion and in turn, this class of process can perfectly reproduce the volatility skew. This new class of stochastic volatility models are now known as rough volatility (rVol from now on) models.

2.2 Description of the project

One of the simplest rVol model is based on the so-called fractional Ornstein-Uhlenbeck (fOU) process. Specifically, the risky asset follows the dynamics in (3) in which the volatility is given by $\sigma_t := e^{Y_t}$, where Y_t satisfies the fractional SDE

$$dY_t = -\lambda(Y_t - \theta)dt + \nu dW_t^H, \tag{4}$$

in which $\theta \in \mathbb{R}$, $\lambda, \nu > 0$ and W^H is a fractional Brownian motion of index $H \in (0, 1/2)$. Thus, the main goal of this project is *the study and implementation of option pricing techniques* in this framework as well as its calibration. Note that W^H is neither a semimartingale nor Markovian, so classical simulations techniques cannot be directly applied. As part of your application and model validation students may:

1. Under the rVol model (3)-(4), calibrate the volatility smile and skew of options written on the S&P500 with special emphasis on the estimation of the parameter H.

2. Estimation and implementation of delta strategies under rough-volatility models: Recall that the delta of a call option is given

$$\Delta_t := \partial_S C_{BS}(S_t, r\tau, \sigma_{BS}(k, \tau)) + \partial_v C_{BS} \partial_S \sigma_{BS}(k, \tau).$$

2.3 Data and references

The main references for this project are:

- [Bayer et al., 2016] , [Gatheral, 2006], [Nourdin, 2013], [Bennedsen et al., 2017]

They can be complemented with

- [Gatheral et al., 2018], [Barndorff-Nielsen et al., 2018], [Hull, 2009].

Options prices can be retrieved using the R package 'quantmod'.

3 Point Processes and Applications

It is a relatively important feature, that the supply and demand curves within the power markets are piecewise constant - specifically, they are *càglàd*. That is, they are left-continuous with right limits. To preserve this property, it is advantageous to model the bids $B = (p_i, q_i)_{i=1}^n$ as the realization of a point process and subsequently "interpolate" them. Here, the name "point process" may be misleading, as the points do not follow a process over time, albeit the structure could potentially be modeled using a time series.

In the book *Random Measures, Theory and Applications* by O. Kallenbergs, a point process is defined as a random measure over a suitable set. In our setting, we may consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a suitable random variable $n : \Omega \to \mathbb{N}_0$ as well as random variables $(p_i, q_i)_{i=1}^n$, such that $(p_i, q_i) : \Omega \to \mathbb{R} \times (0, \infty)$. Notice, that we endow $\mathbb{N}_0$ with its power set as the σ-algebra, while $\mathbb{R} \times (0, \infty)$ is endowed with the typical Borel σ-algebra. Then, we can define our point process, or random measure, $\xi : \Omega \times \mathcal{B}(\mathbb{R} \times (0, \infty)) \to \mathbb{N}_0$ by

$$\xi(\omega, B) = \sum_{i=1}^{n(\omega)} \delta_{(p_i(\omega), q_i(\omega))}(B), \quad \omega \in \Omega, B \in \mathcal{B}(\mathbb{R} \times (0, \infty)).$$

Note, that $\xi(\omega, B)$ counts the number of points (p_i, q_i) that are realized in $B \in \mathcal{B}(\mathbb{R} \times (0, \infty))$ for a given $\omega \in \Omega$.

Since the points $(p_i, q_i)_{i=1}^n$ determine the corresponding supply or demand function, we can analyze the functions by modeling the points. In addition, we can understand the risks arising from the problem's discrete nature.

Another great advantage of this approach is that we can interact with the supply and demand functions; we may be interested in placing our bid (p, q), in addition to the random points $B = (p_i, q_i)_{i=1}^n$. As B is not a prior known, we do not yet know how we will impact the supply or demand curves, but we can simulate and estimate our impact.

The project can be summarised as follows:

1 What models can be used to model the point processes of the bids to the supply and demand curves?

2 How do these depend on fundamental variables like solar and wind production?

3 As the market clears each day, is there any temporal dependence between the point processes?

The suggested literature is

- [Daley et al., 2003]
- [Kallenberg et al., 2017]
- [Cox and Isham, 1980]

In addition, it is of course worth surveying the papers of AAU's researchers Jesper Møller, Rasmus Waagepetersen, Jakob Gulddahl Rasmussen, etc. - in particular those on spatial or marked point processes.

4 Validation and Corrections of Forecasts

The most fundamental data in modelling power prices is weather data. Centrica has access to a lot of weather data and has bought several forecasts from forecast providors. However, these forecasts have proven to be erroneous at times.

To be precise, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and assume that $W : \Omega \times \mathbb{N} \to [0, M]$ is a time-series that describes the production from wind generation units; the value of M corresponds to the total installed capacity and could, in principle, change over time. Now, let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a suitable filtration for which $W(n) \equiv \omega \mapsto W(\omega, n)$ is $\mathcal{F}_n$ measurable for all $n \in \mathbb{N}$.

In addition, we have access to a forecast $\hat{W} : \Omega \times \mathbb{N} \to [0, M]$, where $\hat{W}(n)$ is $\mathcal{F}_{n-1}$ measurable. These forecasts are point forecasts and are not always correct; variations occur as they are only forecasts, but the errors look far too systematic. In particular, the data could suggest that the forecast providers filter away "outliers" to improve, say, their accuracy.

To rigorously test whether the forecasts are good, we would like them to satisfy the *martingale hypothesis*; suppose we are forecasting $W(n)$ subject to the filtration $\mathcal{F}_{n-1}$. Then, if the forecast is good, we should have

$$\hat{W}(n) = \mathbb{E}[W(n) \mid \mathcal{F}_{n-1}].$$

Following the paper, *Testing for the Martingale Hypothesis* by J. Park and Y. Whang, one can test the martingale hypothesis. Alternatively, one can look at the series

$$\varepsilon(n) = W(n) - \hat{W}(n), \quad n \in \mathbb{N}.$$

By extension, one should attempt to correct the forecasts, should they be erroneous. One approach is to utilize copulas; in the paper *"Modelling time-varying exchange rate dependence using the conditional copula."* the author A. Patton shows the existence of a *conditional copulas*, that is generalized in *Time-dependent copulas* by J. Fermanian and M. Wegkamp. The advantage of utilizing copulas is, that we can numerically find the (conditional) distribution of $\varepsilon(n)$ as defined above.

Therefore, the project can be outlined as follows:

1 Investigate whether the forecast $\hat{W}(n)$ of $W(n)$ is "good".

2 How can we correct the forecasts? Does correcting it make sense, or is it better to make new forecasts?

The suggested literature is:

- [Park and Whang, 2005]
- [Patton, 2001]
- [Fermanian and Wegkamp, 2012]

References

[Aït-Sahalia and Jacod, 2014] Aït-Sahalia, Y. and Jacod, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

[Barndorff-Nielsen et al., 2018] Barndorff-Nielsen, O. E., Benth, F. E., and Veraart, A. E. (2018). *Ambit Stochastics*. Probability Theory and Stochastic Modelling. Springer International Publishing.

[Barndorff-Nielsen and Shephard, 2004] Barndorff-Nielsen, O. E. and Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3):885–925.

[Bauwens et al., 2012] Bauwens, L., Hafner, C., and Laurent, S. (2012). *Handbook of Volatility Models and Their Applications*. Wiley Handbooks in Financial E. Wiley.

[Bayer et al., 2016] Bayer, C., Friz, P., and Gatheral, J. (2016). Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904.

[Bennedsen et al., 2017] Bennedsen, M., Lunde, A., and Pakkanen, M. S. (2017). Hybrid scheme for brownian semistationary processes. *Finance and Stochastics*, 21(4):931–965.

[Bennedsen et al., 2021] Bennedsen, M., Lunde, A., and Pakkanen, M. S. (2021). Decoupling the short- and long-term behavior of stochastic volatility. *Journal of Financial Econometrics*.

[Bolko et al., 2022] Bolko, A. E., Christensen, K., Pakkanen, M., and Veliyev, B. (2022). A gmm approach to estimate the roughness of stochastic volatility. *Available at arXiv*.

[Boudt et al., 2021] Boudt, K., Dragun, K., Sauri, O., and Vanduffel, S. (2021). Beta-adjusted covariance estimation. *Available at SSRN*.

[Boudt et al., 2017] Boudt, K., Laurent, S., Lunde, A., Quaedvlieg, R., and Sauri, O. (2017). Positive semidefinite integrated covariance estimation, factorizations and asynchronicity. *Journal of Econometrics*, 196(2):347 – 367.

[Christensen et al., 2010] Christensen, K., Kinnebrock, S., and Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics*, 159(1):116 – 133.

[Cox and Isham, 1980] Cox, D. R. and Isham, V. (1980). *Point processes*, volume 12. CRC Press.

[Daley et al., 2003] Daley, D. J., Vere-Jones, D., et al. (2003). *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer.

[Fermanian and Wegkamp, 2012] Fermanian, J.-D. and Wegkamp, M. H. (2012). Time-dependent copulas. *Journal of Multivariate Analysis*, 110:19–29.

[Gatheral, 2006] Gatheral, J. (2006). *The Volatility Surface: A Practitioner's Guide*. Wiley finance series. John Wiley & Sons.

[Gatheral et al., 2018] Gatheral, J., Jaisson, T., and Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance*, 18(6):933–949.

[Hull, 2009] Hull, J. (2009). *Options, Futures and Other Derivatives.* Eastern economy edition. Pearson/Prentice Hall.

[Jacod and Protter, 2011] Jacod, J. and Protter, P. (2011). *Discretization of Processes.* Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg.

[Kallenberg et al., 2017] Kallenberg, O. et al. (2017). *Random measures, theory and applications*, volume 1. Springer.

[Mykland and Zhang, 2008] Mykland, P. A. and Zhang, L. (2008). Inference for volatility-type objects and implications for hedging. *Statistics and Its Interface*, 1(2):255–278.

[Nadarajah et al., 2014] Nadarajah, S., Zhang, B., and Chan, S. (2014). Estimation methods for expected shortfall. *Quantitative Finance*, 14(2):271–291.

[Nourdin, 2013] Nourdin, I. (2013). *Selected Aspects of Fractional Brownian Motion.* Bocconi & Springer Series. Springer Milan.

[Park and Whang, 2005] Park, J. Y. and Whang, Y.-J. (2005). A test of the martingale hypothesis. *Studies in Nonlinear Dynamics & Econometrics*, 9(2).

[Patton, 2001] Patton, A. J. (2001). Modelling time-varying exchange rate dependence using the conditional copula.

[Wang and Cheng, 2022] Wang, M. and Cheng, Y. (2022). Forecasting value at risk and expected shortfall using high-frequency data of domestic and international stock markets. *Journal of Forecasting.*

[Zhang, 2012] Zhang, L. (2012). Implied and realized volatility: empirical model selection. *Annals of Finance*, 8:259–275.