Estimation of Volatility Using Open, Close, High, and Low Prices: Application to Stochastic Volatility Models and High-Frequency Data.

Georgi Dinolov

September, 2014

Abstract

Early empirical work strongly suggests that the volatility of financial asset prices varies over time. Understanding time-dependent volatility is an extremely important problem in econometrics and finance, especially in terms of options pricing, portfolio management, and risk evaluation. To this end, many models have been developed to estimate time-varying volatility processes from market price data. Most statistical models have used the opening and closing prices over considered sampling frequencies for estimation and forecasting. With the wider availability of higher frequency data, effort has been placed on developing models and algorithms that utilize more of this abundant information to improve inference. We focus on stochastic volatility models which, in addition to opening and closing prices, use minimum and maximum prices to summarize intraperiod data. We derive the likelihood for open, close, high, and low (OCHL) prices for a single asset. We use this likelihood to obtain maximum likelihood estimates for volatility from simulated high-frequency data and compare their performance to other commonly used estimators for high-frequency data. Further, we derive the joint likelihood for two assets with OCHL data, which is expected to provide better estimates for correlation in addition to volatility. In this way, we lay the groundwork for the implementation of a bivariate stochastic volatility algorithm. Finally, we outline a method of incorporating market mictrostructure noise, a phenomenon endogenous to high-frequency data, in the univariate OCHL stochastic volatility setting.

Acknowledgements

I want to thank my advisors Professor Abel Rodriguez and Professor Hongyun Wang for helping me write my advancement. I also want to thank the rest of my advancement committee members, Professor Raquel Prado and Professor Daniel Friedman, for taking the time to read my work.

Contents

Chapter 1

Introduction

1.1 Motivation

In financial econometrics, the geometric Brownian Motion (GBM) is the most popular method for describing the evolution of asset prices (Hull and White, 1987). A random process S_t follows geometric Brownian motion if it is the solution to the stochastic differential equation:

$$dS_t = \mu S_t dt + S_t \sigma dW_t, \tag{1.1}$$

where W_t is a standard Wiener process and μ and σ are the instantaneous drift and volatility of the process, respectively. Under the GBM, the increments of the log price ($Y_t \equiv \log S_t$) over a period Δ, which constitute the log returns of the asset, are independent, stationary, and identically distributed

$$r_{t+\Delta} \equiv Y_{t+\Delta} - Y_t = \log(S_{t+\Delta}/S_t) \sim N\left(\Delta\mu, \Delta\sigma^2\right).$$

The use of GBM to model the evolution of asset prices facilitates estimation and pricing of financial derivatives. However, early work by Mandelbrot (1967) and Fama (1965) showed that volatility changes over time. For example, Figure (1.1) shows the daily squared log returns $(r_{t+\Delta})^2$ for the S&P 500 index and Netflix, Inc., along with simulated daily squared log returns generated with constant volatility. The non-constant volatility of the assets over time is evident by volatility clustering for the asset prices: the distinct periods of low and high daily squared log returns which, under constant intraday volatility and zero drift, have expected values equal the daily volatility. Empirical evidence also shows that negative returns tend to lead to higher volatility than positive returns, which is a phenomenon known as leverage. This chapter reviews models used to capture some of these stylized features of the data.

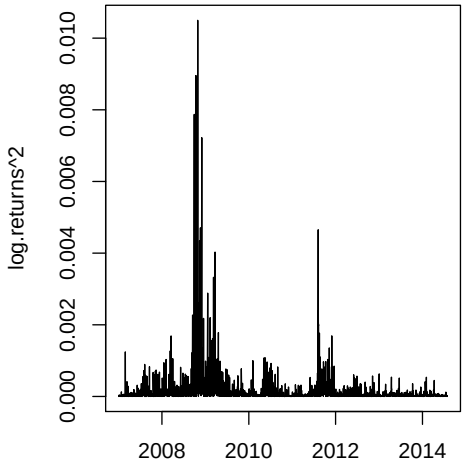
1.2 Review of Models for Time-varying Volatility

The literature on volatility as a non-constant process ranges from the early autoregressive heteroskedastic (ARCH) model by Engle (1982) and the generalized version by Bollerslev (1986), to stochastic volatility models (Hull and White, 1987). Most statistical models have used opening and closing prices over the frequency considered for estimation and forecasting.

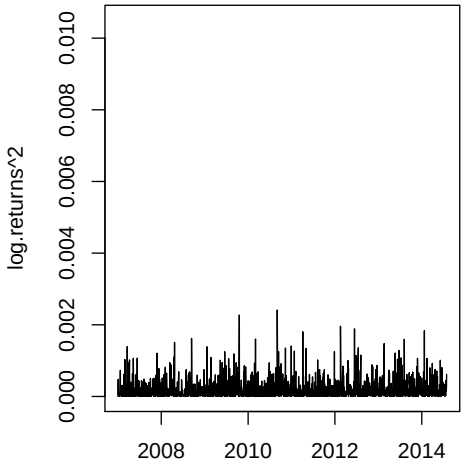
1.2.1 ARCH/GARCH Models

The most important development in modeling volatility changes has been the *autoregressive conditional heteroskedasticity* or ARCH model, introduced by Engle (1982). In results from models of inflation, Engle found that large and small forecast errors occurred in clusters, suggesting a form of heteroskedasticity in which the variance depends on the size of previous fluctuations.

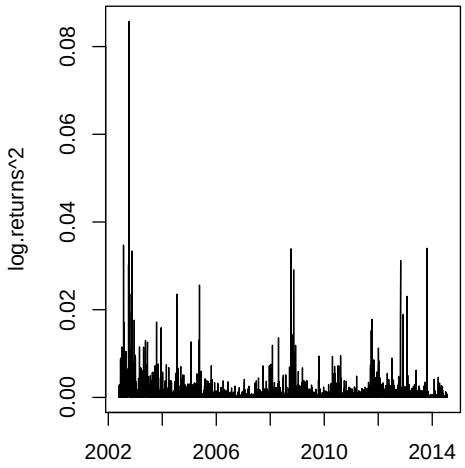
We will introduce ARCH models with the particular ARCH(p) model. ARCH models use a conditional structure to model time-varying volatility. Assuming that the log returns r_t follow the linear



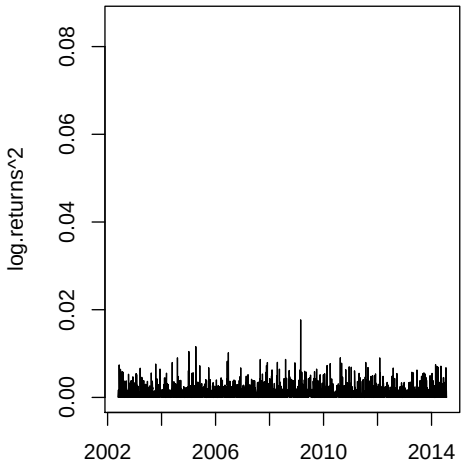
(a) Daily squared log returns for the S&P 500 index.



(b) Simulated daily squared log returns, obtained by sampling a normal distribution with the mean and standard deviation of the S&P 500 daily log returns.



(c) Daily squared log returns for Netflix, Inc.



(d) Simulated daily squared log returns, obtained by sampling a normal distribution with the mean and standard deviation of the Netflix, Inc. daily log returns.

Figure 1.1: Evident in the above figures is the clustering of periods with low and high daily changes in price for the market data, demonstrating that asset volatility fluctuates over time. The constant volatility simulated data is characteristically different from the market data.

relationship

$$r_t = \mu + \varepsilon_t, \qquad t = 1, \ldots, T,$$
$$\varepsilon_t \sim N(0, \sigma_t^2)$$

where μ is a mean level that may be dependent on regression and explanatory variables, the ARCH(p) model characterizes the distribution of the stochastic error by imposing an autocorrelated, time-dependent structure on the variance of the noise term

$$\begin{aligned} \sigma_t^2 &= \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2, \\ \varepsilon_t &= \sigma_t Z_t, \end{aligned}$$

where ω and $\{\alpha_i\}, i = 1, \ldots, p$ are nonnegative constants, and $Z_t \sim N(0,1)$. The conditional variance function is formulated to mimic the phenomenon of large volatility following large shocks to the dependent variable manifested as large deviation of r_t from the mean μ. In other words, high value for ε_t^2 increases σ_{t+1}^2, which in turn increases the expectation of ε_{t+1}^2, and so on.

Bollerslev (1986) extended the ARCH(p) model into the GARCH(p,q) model, the *generalized* ARCH, by including past variances into the conditional variance formula:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \sigma_{t-i}^2 + \sum_{i=1}^{p} \beta_i \varepsilon_{t-i}^2.$$

This additional feature captures longer-term effects outlasting the impacts of shocks to the dependent variables. One issue is that, if α_i and β_i are left unrestricted, the variance can potentially be negative.

ARCH and GARCH models have been widely successful (Bollerslev and Zhang, 2003), but they fail to capture some important features of financial and economic data. The most important feature not captured is the leverage effect, which has been addressed by Nelson's *exponential* GARCH (Nelson, 1991) model:

$$\log \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \log \sigma_{t-i}^2 + \sum_{i=1}^{q} \beta_i g(Z_{t-i}),$$

where

$$g(Z_t) = \theta Z_t + \gamma[|Z_t| - E|Z_t|],$$

with $Z_t \sim N(0,1)$. Assuming that $\gamma > 0$ and $\theta = 0$, the innovations in $\log \sigma_{t+1}^2$ are positive (negative) when the magnitude of Z_t is larger (smaller) than its expected value. If $\gamma = 0$ and $\theta < 0$, the innovation in the variance is negative (positive) when Z_t is positive (negative). In this manner the EGARCH model captures the asymmetric leverage effect of negative returns leading to greater volatility, as well as the previously included effect of above (below)-average (in terms of magnitude) residuals leading to greater (smaller) volatility.

ARCH-type models are appealing because they allow for direct parametric modeling of volatility, and consequentially varieties of models building upon the ARCH, GARCH, and EGARCH models presented so far have been developed. The immediate availability of a conditional likelihood function for the observables allows for standard and efficient MLE methods to be used for fitting (Baillie and Myers, 1991), and by an extension, Bayesian analysis is applicable as well (Bauwens and Lubrano, 1998; Nakatsuma, 2000; Vrontos et al., 2000).

1.2.2 Stochastic Volatility

Whereas GARCH uses a single stochastic process to drive both the asset and volatility process, stochastic volatility (SV) models are defined by a set of two stochastic differential equations that define separate (but possibly dependent) evolutions for the asset and volatility:

$$dS_t = \mu(S_t, \sigma_t)dt + \nu(S_t, \sigma_t)dW_{t,1}, \tag{1.2}$$
$$d\sigma_t = \alpha(S_t, \sigma_t)dt + \beta(S_t, \sigma_t)dW_{t,2},. \tag{1.3}$$

One of the first and most well-known continuous-time SV models was introduced by Hull and White (1987). In it, the log price follows Brownian motion (so that the price follows GBM) with changing volatility, and the volatility follows an Ornstein-Uhlenbeck (OU) process:

$$\begin{aligned} dY_t &= \mu dt + \sigma_t dW_{t,1} \\ d\log(\sigma_t) &= \alpha(\phi - \log(\sigma_t))dt + \omega dW_{t,2} \end{aligned}$$

Discretizing the continuous-time model above produces the state-space model

$$\begin{aligned} Y_t &= \mu + Y_{t-1} + \varepsilon_t^1 & \varepsilon_t^1 &\sim N(0, \sigma_t^2), \\ \log(\sigma_t) &= \alpha + \phi\left[\log(\sigma_{t-1}) - \alpha\right] + \varepsilon_t^2 & \varepsilon_t^2 &\sim N(0, \tau^2), \end{aligned}$$

which is the model proposed by Clark (1973), one of the first papers in which serial correlation is allowed in the volatility process.

From the late 1990s, stochastic volatility models have taken center stage in econometric analysis of volatility forecasting (Shephard, 2005). Stochastic volatility models can capture time clustering (Hull and White, 1987), as well as leverage effects (Yu, 2005) by introducing correlation between $W_{t,1}$ and $W_{t,2}$.

Empirical evidence suggests that, across different asset types and timescales, the dependence in the volatility structure initially decays quickly at short lags but then decays slowly at longer lags (Andersen and Bollerslev, 1997a,b, 1998). For this reason, effort has been placed on developing both continuous and discrete-time long-memory SV models. In the discrete case, one approach has been to model log volatility as a fractionally integrated process (Harvey, 2002; Breidt, Crato, and De Lima, 1998). In continuous time, the log volatility has been modeled as fractionally integrated Brownian motion (Comte and Renault, 1998), and a square root model driven by fractionally integrated BM (Comte, Coutin, and Renault, 2012). Another important extension includes the addition of jumps in the diffusion of price (Bates, 1996).

Fitting stochastic volatility models is made difficult by the non-linearities of the likelihood as well the large sample space formed by the latent series $\{\sigma_t\}$. Techniques for estimation of SV models include quasi-maximal likelihood approaches as done by Ruiz (1994), the particle filter approach of Sandmann and Koopman (1998), the generalized method of moments (Melino and Turnbull, 1990), and MCMC methods (Jacquier, Polson, and Rossi, 2002; Kim, Shephard, and Chib, 1998; Omori, Chib, Shephard, and Nakajima, 2007).

1.2.3 Time-changed Brownian Motion

A more general approach for representing changing volatility in financial assets, which encompasses both stochastic volatility and GARCH-type models, is time-changed Brownian motion. Time-changed Brownian motion was first used in the paper of Clark (1973), which introduced the concept into financial economics. Under time-changed Brownian motion, the log-price Y of an asset is written as

$$Y_t = W_{\tau_t}, \quad t \geq 0, \tag{1.4}$$

where W is Brownian motion, t is continuous time and τ_t is a non-negative process with non-decreasing sample paths. If W and τ are independent, $Y_t|\tau_t \sim N(0, \tau_t)$. In this case, it is immediately obvious that this formulation allows for the second moments of the price process to change over time. Further, marginally Y_t is a scale mixture of normals, which means that returns are symmetric but can be fat tailed. The time-change process τ_t can be modeled as either a function of some observables, or as a latent process, which is more in line with the modern development of stochastic volatility.

Time-changed Brownian motion has also become a popular tool for building financial models (Geman, 2005). Stochastic processes of this kind can capture asset price jumps, changing return volatility, and correlation between volatility and asset returns (Carr and Wu, 2004). The mathematical treatment of these models usually involves the computation of the characteristic function, which is typically easy to get as long as τ and W are independent. However, even if the characteristic function is available in closed form, the associated density often needs to be computed numerically.

1.3 Time-Varying Volatility in High-Frequency Data

The advent of electronic trading has allowed for the increase in the volume and speed at which transactions on financial markets are performed and recorded. Today, transactions occur on a microsecond timescale (Brogaard, 2010) for certain types of assets, and there has been increasing interest in using this data to better understand volatility in financial markets.

Much of the effort to use high-frequency data to understand the evolution of volatility has focused on *realized volatility*. The concept of realized volatility was introduced in three independent papers by Andersen et al. (2001), Barndorff-Nielsen and Shephard (2002), and Comte and Renault (1998) as a means to perform the estimation of $\int_t^{t+\Delta} \sigma^2(Y_s, s)ds$, where Δ is some finite period of time and y_t is the log-price of the asset following a diffusion process.

Let $\Delta = 1$ day, and suppose there are m observation of the price over the trading period. The i^{th} interday log-return is defined as

$$r_i^{(m)} = y_{i/m} - y_{(i-1)/m}, \qquad i = 1, \ldots, m.$$

The sum of the squared log intraday returns is the realized volatility

$$RV^{(m)} = \sum_{i=1}^{m} r_i^{(m)^2}. \tag{1.5}$$

It can be shown that $RV^{(m)} \xrightarrow{p} \int_t^{t+\Delta} \sigma^2(Y_s, s)ds$ as $m \to \infty$. Realized volatility is a consistent estimator of the integrated volatility (Barndorff-Nielsen, 2002), and it is asymptotically distributed:

$$\frac{\sqrt{m}(RV^{(m)} - \int_t^{t+\Delta} \sigma^2(Y_s, s)ds)}{\sqrt{2\int_t^{t+\Delta} \sigma^4(Y_s, s)ds}} \xrightarrow{d} N(0,1) \tag{1.6}$$

Barndorff-Nielsen and Shephard (2004) also propose a consistent estimator for the power variation $\int_t^{t+\Delta} \sigma^4(y_s, s)ds$, which is needed to construct credible confidence intervals given the result in equation (1.6). There has been further work in the derivation of consistent estimators of higher moments of the integrated volatility, which are used in the estimation of continuous time SV models with higher frequency data (Todorov, 2011).

1.3.1 Microstructure Noise

One challenge with the analysis of high-frequency price data is the presence of microstructure noise at the microsecond timescale. As the sampling period shrinks down to the transaction-by-transaction frequency, irregular spacing between transactions, discreteness in transaction prices, and very short-term temporal dependence become dominant features of the data (Stoll, 2000). For this reason, the foundational diffusion assumption on the log-price of assets no longer holds. Measurements at high frequencies no longer capture only the evolution of prices, but they capture a noisy version of the price.

In the model-free setting of realized volatility estimators, one approach to mitigating the confounding effects of microstructure noise is to combine averaged RV estimates obtained by subsampling the data with a RV estimate using all available data (Zhang et al., 2005). Another approach is to employ a class of kernel-based methods similar to those used for estimating the long-run variance of a stationary time-series in the presence of autocorrelation (Hansen and Lunde, 2006).

In the context of stochastic volatility, microstructure noise could, for example, be accounted for with a model of the form:

$$\begin{aligned} Y_{t+1} &= \log(S_{t+1}) + m_t \\ \log(S_{t+1}) &= \log(S_t) + \mu(S_t, \sigma_t) + \nu(S_t, \sigma_t)\varepsilon_{t+1,1} \\ \log(\sigma_{t+1}) &= \log(\sigma_t) + \alpha(S_t, \sigma_t) + \beta(S_t, \sigma_t)\varepsilon_{t+1,2}. \end{aligned}$$

The distributional assumptions on the noise m_t can be varied, and there can be correlation between the innovation terms $\varepsilon_{t+1,1}$ and $\varepsilon_{t+1,2}$

1.3.2 Stochastic Volatility Models and High-frequency Data

One of the first papers to use realized volatility in the estimation of continuous-time SV models is by Barndorff-Nielsen (2002). The idea of the paper is use the continuous time SV model and the RV estimator to form a state-space model, which can be estimated with a Kalman filter.

Assuming the variance $\sigma^2(t)$ follows the Ornstein-Uhlenbeck (OU) process instead of the log-variance, the autocorrelation function for the volatility is known to be $r(t,t') = \exp(-\lambda|t'-t|)$. The functional form for $r(t)$ allows for the explicit form of the correlation function between observations in the volatility process:

$$corr(\sigma_t^2, \sigma_{t+s}^2) = d\exp(t\{-\lambda\Delta(s-1)\}),$$

where d is a constant in terms of λ which is less than 1, and Δ is the length of time between returns. Hence, the actual volatility has the autocorrelation function of an autoregressive moving average (ARMA) model of order (1,1). Letting $RV_t^{(m)}$ be the realized volatility in interval $(t-\Delta, t)$ computed with m intraperiod observations, $RV_t^{(m)}$ can be decomposed in terms of the true volatility for the interval plus an error term u_t due to microstructure noise

$$RV_t^{(m)} = \sigma_t^2 + u_t. \tag{1.7}$$

Setting $E(\sigma_t^2) = \xi\Delta$, and using the ARMA(1,1) representation of the volatility process, equation (1.7) becomes

$$RV_t^{(m)} = (\xi\Delta + \phi\sigma_{t-1}^2 + \theta\sigma_\sigma\varepsilon_{t-1} + \sigma_\sigma\varepsilon_t) + (\sigma_u\nu_t),$$

where $\varepsilon_{t-1}, \varepsilon_t, \nu_t \sim N(0,1)$. The parameter σ_σ is the standard deviation of the innovations driving the volatility process, σ_u is the standard deviation of the microstructure noise, and (ϕ, θ) are the ARMA parameters. Therefore, the realized volatility observations can be expressed as the state-space model

$$\begin{aligned} RV_t^{(M)} &= \Delta\xi + (1 \quad 0)\alpha_t + \sigma_u\nu_t \\ \alpha_t &= \begin{pmatrix} \phi & 1 \\ 0 & 0 \end{pmatrix}\alpha_{t-1} + \begin{pmatrix} \sigma_\sigma \\ \sigma_\sigma\theta \end{pmatrix}\varepsilon_t, \end{aligned}$$

which can be estimated with a Kalman filter.

An alternative approach for estimating SV models using realized volatility data is based on the generalized method of moments. In this method, higher powers of realized volatility are used as approximations to higher orders of integrated volatility, as done by Bollerslev and Zhou (2002). By matching the sample moments of the realized volatility to the moments of the integrated volatility implied by the continuous-time SV model, a generalized method of moments estimator for the underlying model parameters can be obtained. Some more recent works dealing with fitting stochastic volatility models to high-frequency data include that of Venter and de Jongh (2012) and Shirota et al. (2014).

1.4 Plan for the Advancement

In Chapter 2 we will derive the likelihood for closing prices of financial asset prices that takes into account the highest and lowest prices within a trading period. We also compare maximum likelihood estimates for volatility based on this approach to realized volatility estimates using simulated high-frequency data. In Chapter 3 we extend the model in Chapter 2 to two dimensions and examine the convergence properties of the thus-derived likelihood solution. In Chapter 4 we outline the future direction of this project.

Chapter 2

Range-Based Methods for High-Frequency Data

Wider availability of tick-by-tick/bid-ask financial transaction data has motivated the development of models that utilize this information to better estimate asset volatility. We focus on an approach which collapses high-frequency intraperiod transaction information in the form of recorded maximum and minimum prices. Work in this direction includes that of Alizadeh et al. (2002), which uses only the range of an asset over a sampling period to extract latent stochastic volatility, while Lildholdt (2002) uses open, high, low, and close data to do maximum likelihood estimation in a GARCH framework. Horst et al. (2012) combine the use of open, high, low, and close data with a Bayesian approach to estimation and prediction in a stochastic volatility framework.

2.1 Problem Formulation

Following the approach of Horst et al. (2012), we are interested in finding the joint probability density for the closing, maximum, and minimum prices for an asset during a nominal period $(t-1,t)$. Recall the GBM model discussed in Section 1.1. Assuming that σ and μ are constant over the period of interest, the log-price follows the SDE

$$dY_s = \mu ds + \sigma dW_s, . \tag{2.1}$$

Using the Fokker-Planck Equation, we can take the SDE in (2.1) and express the probability density of the asset price $q(y,s)$ as the PDE

$$\begin{aligned} \frac{\partial}{\partial s}q(y,s) &= -\mu\frac{\partial}{\partial y}q(y,s) + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial y^2}q(y,s), & (2.2)\\ q(y,t-1) &= \delta(y-y_{t-1}), & (2.3) \end{aligned}$$

where y_{t-1} is the price at the beginning of the interval and δ is the Dirac delta function.

The value $q(y,t)$ is the probability that $Y_t = y$, namely that the log-price of the asset is y at the end of the time interval. Without any additional restrictions, therefore, the solution to the initial value problem (2.2) - (2.3) constitutes the likelihood of a closing price, given the parameters μ and σ, at time s.

Now we can incorporate the information for the high and low prices within our model if we restrict the stochastic process so it does not go below or above a certain range:

$$\begin{aligned} dY_s &= \mu ds + \sigma dW_s, & (2.4)\\ a_t < \ Y_s \ &< b_t. & (2.5) \end{aligned}$$

Equivalently, we can consider the initial value/boundary condition problem, where the boundary

conditions correspond to the process having zero probability of going beyond a_t and b_t:

$$\frac{\partial}{\partial s}q(y,s) = -\mu\frac{\partial}{\partial y}q(y,s) + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial y^2}q(y,s), \tag{2.6}$$

$$q(y,t-1) = \delta(y-y_{t-1}), \tag{2.7}$$

$$q(a_t,s) = 0, \qquad q(b_t,s) = 0. \tag{2.8}$$

The problem in (2.6) - (2.8) can be further simplified by a transformation which eliminates the first derivative term, thereby transforming the advection-diffusion problem to a purely diffusion problem. Letting

$$q(y,s) = \exp(ay+bs)p(y,s), \tag{2.9}$$

we apply the spatial differential operator and the time-derivative to the right side of equation (2.9) and equate the two expressions. We thus obtain the differential equation

$$p(y,s)\left[-\mu a + \frac{1}{2}\sigma^2 a^2\right] + \frac{\partial}{\partial y}p(y,s)\left[-\mu + \frac{1}{2}\sigma^2 a\right] + \frac{\partial^2}{\partial y^2}p(y,s)\left[\frac{1}{2}\sigma^2\right] = bp(y,s) + \frac{\partial}{\partial s}p(y,s) \tag{2.10}$$

The system

$$-\mu a + \frac{1}{2}\sigma^2 a^2 = 0 \tag{2.11}$$

$$b = 0 \tag{2.12}$$

defines $a = \sqrt{2\mu}/\sigma$, and $b = 0$ such that the original problem is reduced to the heat equation on a bounded domain, with the same IC/BV conditions:

$$\frac{\partial}{\partial s}p(y,s) = \frac{1}{2}\sigma^2\frac{\partial^2}{\partial y^2}p(y,s) \tag{2.13}$$

$$p(y,s) = 0, \qquad \text{for} \quad y = a_t, b_t \tag{2.14}$$

$$p(y,0) = \delta(y-y_{t-1}) \tag{2.15}$$

2.1.1 Closed-Form Solution to the IV/BV Problem via Fourier Expansion

The Strum-Liouville problem in (2.13) - (2.15) can be solved by recognizing that the solution is separable. By treating the temporal and spatial differential equations independently, the solution to the problem takes on the familiar Fourier series expansion, except the coefficients of the trigonometric terms decay with time

$$p(y,s) = \sum_{n=1}^{\infty} C_n \exp\left(-\lambda_n s\right)\sin\left(\frac{n\pi(y-a_t)}{b_t-a_t}\right). \tag{2.16}$$

There are no cosine terms in the expansion as they do not satisfy the boundary conditions. The expression for λ_n is found by substituting the solution into the governing system of equations (2.13) - (2.15) and, after cancellation,

$$\lambda_n = -\frac{1}{2}\sigma^2\frac{n^2\pi^2}{(b_t-a_t)^2}$$

The coefficients $\{C_n\}_{n=1}^{\infty}$ are found by first recognizing an important property of the sine-series expansion of the solution, the orthogonality of the basis functions with respect to an inner product. Defining $\xi = y - a_t$ and $L = b_t - a_t$, the inner product of two functions $h(\xi)$ and $g(\xi)$ is given by $\int_0^L h(\xi)g(\xi)d\xi$. The orthogonality is demonstrated by the inner product of two basis functions:

$$\int_0^L \sin\left(\xi\frac{n\pi}{L}\right)\sin\left(\xi\frac{m\pi}{L}\right)d\xi = \frac{1}{2}\int_0^L \cos\left(\xi\frac{(n-m)\pi}{L}\right) - \cos\left(\xi\frac{(n+m)\pi}{L}\right)d\xi = \begin{cases} 0 & \text{if } n \neq m \\ L/2 & \text{if } n = m \end{cases}$$

Next, we consider the initial condition with $s = 0$

$$p(\xi,0) = \sum_{n=1}^{\infty} C_n \sin\left(\xi\frac{n\pi}{L}\right) = \delta(\xi - (y_{t-1} - a_t)).$$

Using the orthogonality condition,

$$\int_0^L p(\xi,0)\sin\left(\xi\frac{n'\pi}{L}\right)d\xi = \sum_{n=1}^{\infty} C_n \int_0^L \sin\left(\xi\frac{n'\pi}{L}\right)\sin\left(\xi\frac{n\pi}{L}\right)d\xi = C_{n'}\frac{L}{2}.$$

Substituting the initial condition,

$$\int_0^L p(\xi,0)\sin\left(\xi\frac{n'\pi}{L}\right)d\xi = \int_0^L \delta(\xi-(y_{t-1}-a_t))\sin\left(\xi\frac{n'\pi}{L}\right)d\xi = \sin\left((y_{t-1}-a_t)\frac{n'\pi}{L}\right)$$

From the two conditions, we have a form for the coefficients:

$$C_{n'} = \frac{2}{L}\sin\left((y_{t-1}-a_t)\frac{n'\pi}{L}\right) \tag{2.17}$$

Thus, the final form for p is

$$p(y,s) = \sum_{n=1}^{\infty}\frac{2}{L}\sin\left((y_{t-1}-a_t)\frac{n\pi}{L}\right)\exp\left(-\frac{1}{2}\sigma^2\frac{n^2\pi^2}{L^2}s\right)\sin\left((y-a_t)\frac{n\pi}{L}\right) \tag{2.18}$$

Combining equations (2.9), (2.11)-(2.12), and (2.18), final solution to the IV/BC problem is

$$q(y,s) = \exp\left(\frac{\sqrt{2\mu}}{\sigma}y\right)\sum_{n=1}^{\infty}\frac{2}{L}\sin\left((y_{t-1}-a_t)\frac{n\pi}{L}\right)\exp\left(-\frac{1}{2}\sigma^2\frac{n^2\pi^2}{L^2}s\right)\sin\left((y-a_t)\frac{n\pi}{L}\right) \tag{2.19}$$

2.1.2 Closed-Form Solution to the IC/BV Problem via Method of Images

An alternative way to solve the diffusion problem in equations (2.13) - (2.15) is through the Method of Images. The general idea of the method is to find the fundamental solution to the IC problem in equations (2.13) and (2.15). For convenience of mathematical notation, we study time evolution from t to $t+1$. The initial value is then $y(0) = y_t$. We enforce the boundary conditions through an appropriate series of reflections. This solution can also be derived probabilistically through the reflection principle for Brownian motion, as given by Freedman (1971).

The solution to the initial condition problem in equations (2.13) and (2.15) is the Green's function for the heat equation, which is known to be pdf of the Normal distribution with mean y_{t-1} and standard deviation $\sigma\sqrt{s}$:

$$g_0(y,s) = \frac{1}{\sqrt{2\pi s\sigma^2}}\exp\left(-\frac{1}{2}\frac{(y-y_t)^2}{s\sigma^2}\right) \equiv N(y;y_t,s\sigma^2) \tag{2.20}$$

A single *reflection* of the fundamental solution about the boundary point a_t is defined in terms of a translation of the mean $\zeta(y) = 2a_t - y$ (since the fundamental solution is symmetric around the mean) and a negation of the function value, as shown in Figure (2.1a). The resultant reflection, denoted $r_0(y,s)$, satisfies the heat equation in (2.13):

$$r_0(y,s) = -\frac{1}{\sqrt{2\pi s\sigma^2}}\exp\left(-\frac{1}{2}\frac{(y-\zeta(y_t))^2}{s\sigma^2}\right) = -N(y;\zeta(y_t),s\sigma^2)$$

By construction, both r_0 and g_0 have the same magnitudes at $y = a_t$ but with opposing signs. The sum of the fundamental solution and its reflection, denoted as $g_1(y,s)$ and shown below in (2.21), therefore satisfies the heat equation, as well as the boundary condition $g_1(a_t,s) = 0$. However, $g_1(y,s)$ does not necessarily satisfy the boundary condition $g_1(b_t,s) = 0$, as illustrated in Figure (2.1a).

$$g_1(y,s) = g_0(y,s) + r_0(y,s) = N(y;y_t,s\sigma^2) - N(y;\zeta(y_t),s\sigma^2) \tag{2.21}$$

If we reflect g_1 about b_t via a translation of the mean $\xi(y) = 2b_t - y$ and a negation:

$$r_1(y,s) = -N(y;\xi(y_t),s\sigma^2) + N(y;\xi(\zeta(y_t)),s\sigma^2)$$

and add this reflection to g_1

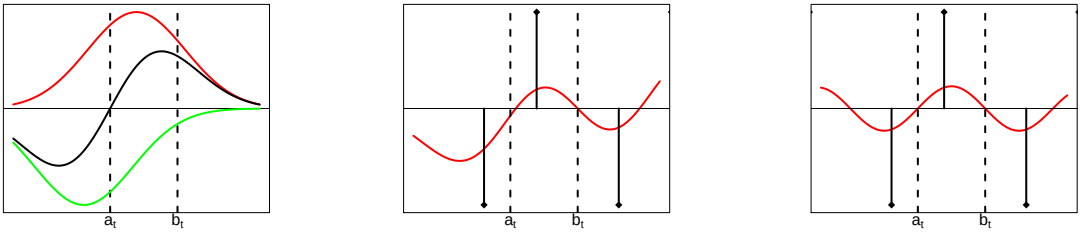
$$g_2(y,s) = g_1(y,s) + r_1(y,s) = N(y;y_t,s\sigma^2) - N(y;\zeta(y_t),s\sigma^2) - N(y;\xi(y_t),s\sigma^2) + N(y;\xi(\zeta(y_t)),s\sigma^2),$$

the resultant function g_2 (Figure (2.1b)) also satisfies the heat equation and it matches the boundary condition at b_t, but the boundary condition at a_t no longer holds. The process can be continued in this manner iteratively, with each reflection enforcing the boundary condition at a_t then b_t. However, it is necessary to ensure that the reflections do not contain images already in the system of images.

Each reflection sends the sources of the reflected fundamental solutions further away from the domain $[a_t, b_t]$, making the maximal factor by which the boundary conditions are violated less and less. Thus, we can write down the solution to the IC/BV problem as the infinite sum of the superposition of reflected fundamental solutions. This can be defined by all possible alternating combinations of ξ and ζ applied to the center of g_0. The expression in (2.22) enumerates all such combinations and is therefore the solution to the diffusion equation under the initial conditions and boundary values.

$$\begin{aligned} p(y,s) = {} & N(y;y_{t-1},s\sigma^2) - N(y;\zeta(y_{t-1}),s\sigma^2) - N(y;\xi(y_{t-1}),s\sigma^2) + \\ & \sum_{n=1}^{\infty} \left[N(y;(\xi\circ\zeta)^n(y_{t-1}),s\sigma^2) + N(y;(\zeta\circ\xi)^n(y_{t-1}),s\sigma^2) \right. \\ & \left. - N(y;(\xi\circ\zeta)^n\circ\xi(y_{t-1}),s\sigma^2) - N(y;(\zeta\circ\xi)^n\circ\zeta(y_{t-1}),s\sigma^2) \right]. \end{aligned} \tag{2.22}$$

The final solution to the original advection-diffusion problem is therefore $p(y,s)$ multiplied by the exponential factor $\exp(ay+bs)$, which is given in equation(2.23). Figure (2.1c) shows the solution from equation (2.22) with $n=1$.



(a) The fundamental solution g_0 (red), its reflection about a_t, r_0 (green), and their sum g_1 (black). Observe that the boundary condition is satisfied at a_t but not b_t.

(b) The solution after a single reflection about a_t, g_1 (red), its reflection about b_t, r_1 (green), and their sum g_2 (black). The boundary condition is satisfied at b_t but is now violated at a_t.

(c) The solution according to equation (2.22) with $n=1$.

Figure 2.1: Construction of the solution using the Method of Images

$$\begin{aligned} q(y,s) = {} & \exp\left(\frac{\sqrt{2\mu}}{\sigma}y\right) \times \\ & \left\{ N(y;y_{t-1},s\sigma^2) - N(y;\zeta(y_{t-1}),s\sigma^2) - N(y;\xi(y_{t-1}),s\sigma^2) + \right. \\ & \sum_{n=1}^{\infty} \left[N(y;(\xi\circ\zeta)^n(y_{t-1}),s\sigma^2) + N(y;(\zeta\circ\xi)^n(y_{t-1}),s\sigma^2) \right. \\ & \left. \left. - N(y;(\xi\circ\zeta)^n\circ\xi(y_{t-1}),s\sigma^2) - N(y;(\zeta\circ\xi)^n\circ\zeta(y_{t-1}),s\sigma^2) \right] \right\}. \end{aligned} \tag{2.23}$$

2.2 Finding the Likelihood for the Closing Price

The full solution $q(y,t)$ is the likelihood that a random process satisfying the stochastic differential equation (2.1) has an infimum greater than or equal to a_t and a supremum less than or equal to M_t over the period $(t-1,t)$. In terms of probabilities

$$q(y,t) = P(m_t \geq a_t, M_t \leq b_t, Y_t = y | Y_{t-1} = y_{t-1}, m_t < M_t).$$

However, to perform the statistical inference we need the density associated with $q(y,t)$, which is derived by differentiating $q(y,t)$ with respect to the boundary conditions:

$$-\frac{\partial^2}{\partial a_t \partial b_t} q(y,t) = P(m_t = a_t, M_t = b_t, Y_t = y | Y_{t-1} = y_{t-1}). \tag{2.24}$$

The analytic expression for the likelihood, whether that in equation (2.19) or (2.23), allows us to perform this differentiation numerically as well as analytically. Once this differentiation is performed, we can use familiar optimization methods to find maximum likelihood estimates for the parameters μ and σ^2, given closing, open, high, and low prices for a given interval $(t-1,t)$.

2.2.1 Evaluating the accuracy of the solutions

We want to compare the convergence of the likelihood (2.24) based on the trigonometric expansion solution (2.19). This convergence is dependent upon the number of terms used in the truncated summation for the expansion, as well as the size of the discrete step when performing numerical differentiation. This analysis sets the stage for our treatment of the bivariate advection-diffusion problem in Chapter 3.

Considered are ten simulations over a nominal period of length 1 based on a forward-Euler discretization scheme of the governing stochastic differential equation:

$$Y_{k+1} = Y_k + \mu \Delta t^* + \sigma \Delta t^* \varepsilon_k, \tag{2.25}$$

with $\Delta t^* = 1/23,400$, and $\varepsilon_k \sim N(0,1)$. This is equivalent to sampling the returns once every second over the length of a trading day. Further,

$$\mu = 1, \qquad \sigma = 1.$$

For each simulation there is a recorded min, max, and closing value, while all log open prices are set to 0. As a test to study the numerical accuracy of the solution, we examine the likelihood value at the closing log return using the observed minima and maxima set to a_t and b_t respectively, as well as the true μ and σ used to generate the data. When performing numerical differentiation, a second-order two-point stencil method is used with a discretization step

$$\Delta x = \frac{1}{2^k} \frac{1}{100} (b_t - a_t).$$

Values of the likelihoods are considered across ranges of N and k and are compared to the likelihood values obtained through analytic differentiation. The results are shown in Tables (2.1) - (2.3). We see that for both numerical differentiation factors $k = 4$ and $k = 8$ likelihood values match with those obtained via analytic differentiation, so either differentiation approach is appropriate for maximum likelihood estimation. Further, we see that likelihood values are the same across all sizes of N. This means that the number of terms needed in the trigonometric expansion for accurate solutions is relatively low, and the Fourier expansion used in our derivation is an efficient representation for the univariate diffusion problem.

high	low	close	$N=4$	$N=8$	$N=16$	$N=32$	$N=64$
1.8562	-0.258	1.5907	1.0229	1.0229	1.0229	1.0229	1.0229
2.2548	-0.193	1.7251	0.0841	0.0841	0.0841	0.0841	0.0841
3.336	-0.0274	3.0904	0.7138	0.7294	0.7294	0.7294	0.7294
2.4832	-0.0806	2.2762	1.5858	1.5858	1.5858	1.5858	1.5858
1.577	-0.1144	1.3633	1.4807	1.4807	1.4807	1.4807	1.4807
1.2175	-0.3971	0.7699	0.1685	0.1685	0.1685	0.1685	0.1685
2.1135	0	2.0993	2.4332	2.4332	2.4332	2.4332	2.4332
0.3217	-0.9688	-0.902	-1.6368	-1.6368	-1.6368	-1.6368	-1.6368
1.0172	-0.3775	-0.2605	-1.4587	-1.4587	-1.4587	-1.4587	-1.4587
0.9197	-0.8386	0.397	-1.5776	-1.5776	-1.5776	-1.5776	-1.5776

Table 2.1: Likelihood values for simulations obtained via numerical differentiation with $k=4$ ($\Delta x = 1/(100\cdot 2^4)\cdot(b_t - a_t)$).

high	low	close	$N=4$	$N=8$	$N=16$	$N=32$	$N=64$
1.8562	-0.258	1.5907	1.0229	1.0229	1.0229	1.0229	1.0229
2.2548	-0.193	1.7251	0.0841	0.0841	0.0841	0.0841	0.0841
3.336	-0.0274	3.0904	0.7137	0.7294	0.7294	0.7294	0.7294
2.4832	-0.0806	2.2762	1.5858	1.5858	1.5858	1.5858	1.5858
1.577	-0.1144	1.3633	1.4807	1.4807	1.4807	1.4807	1.4807
1.2175	-0.3971	0.7699	0.1685	0.1685	0.1685	0.1685	0.1685
2.1135	0	2.0993	2.4332	2.4332	2.4332	2.4332	2.4332
0.3217	-0.9688	-0.902	-1.6368	-1.6368	-1.6368	-1.6368	-1.6368
1.0172	-0.3775	-0.2605	-1.4587	-1.4587	-1.4587	-1.4587	-1.4587
0.9197	-0.8386	0.397	-1.5776	-1.5776	-1.5776	-1.5776	-1.5776

Table 2.2: Likelihood values for simulations obtained via numerical differentiation with $k=8$ ($\Delta x = 1/(100\cdot 2^8)\cdot(b_t - a_t)$).

high	low	close	$N=4$	$N=8$	$N=16$	$N=32$	$N=64$
1.8562	-0.258	1.5907	1.0229	1.0229	1.0229	1.0229	1.0229
2.2548	-0.193	1.7251	0.0841	0.0841	0.0841	0.0841	0.0841
3.336	-0.0274	3.0904	0.7137	0.7294	0.7294	0.7294	0.7294
2.4832	-0.0806	2.2762	1.5858	1.5858	1.5858	1.5858	1.5858
1.577	-0.1144	1.3633	1.4807	1.4807	1.4807	1.4807	1.4807
1.2175	-0.3971	0.7699	0.1685	0.1685	0.1685	0.1685	0.1685
2.1135	0	2.0993	2.4332	2.4332	2.4332	2.4332	2.4332
0.3217	-0.9688	-0.902	-1.6368	-1.6368	-1.6368	-1.6368	-1.6368
1.0172	-0.3775	-0.2605	-1.4587	-1.4587	-1.4587	-1.4587	-1.4587
0.9197	-0.8386	0.397	-1.5776	-1.5776	-1.5776	-1.5776	-1.5776

Table 2.3: Likelihood values for simulations obtained via analytic differentiation.

2.3 Using Close, High, Low, Open Data for High-Frequency Data: a Monte Carlo Simulation Study

For high-frequency data where transactions occur on the microsecond scale, there is a lot of available information even within sampling intervals on the second scale. Realized volatility estimates use all data within sampling intervals, while the estimates based on the likelihood developed above use summaries of intraperiod data in the form of open, close, high, and low (OCHL) recorded prices. In this simulation study, we are interested in seeing how our maximum likelihood estimates for σ^2 based on OCHL data compare with RV estimates.

2.3.1 Maximum Likelihood Estimation

The goal of the simulation study is to understand the trade-offs between using all high-frequency data and, instead, using summaries. Recalling the basic diffusion model in (2.1) with constant μ and σ^2, the simulations use

$$\mu = 7.936508 \cdot 10^{-5}, \qquad \sigma = 0.01259882.$$

The parameters correspond to daily estimates for drift and standard deviation of the returns of the S&P 500 index. Thus, in our simulations Δ corresponds to an idealized 7-hour trading day. The simulated data is obtained by the forward-Euler discretization in equation (2.25), where $\varepsilon_k \sim N(0,1)$, and index k corresponds to the time $k\Delta t^*$, and Δt^* corresponds to a second of a trading day. Generating the data according to equation (2.25) amounts to sampling the diffusion process every second. It should be noted here that the data-generation method used here does not include microstructure noise.

When performing estimation, longer sampling intervals are used. They are

$$\Delta t = 5 \text{ sec}, 10 \text{ sec}, 20 \text{ sec}, 30 \text{ sec}, 1 \text{ min}, 5 \text{ min}, 10 \text{ min}, 20 \text{ min}, 30 \text{ min}, 1\text{hr}, 3.5\text{hr}, 7\text{hr}.$$

We expect the realized volatility estimate to approach the true diffusion parameter, in probability, as $\Delta t \to 0$. As for estimating volatility using OCHL data, each simulated day contains $1/\Delta t$ intervals. Each interval $[k\Delta t, (k+1)\Delta t]$, in turn, contains $\Delta t/\Delta t^* + 1$ data points. For example, if $\Delta t = 5$ sec, there are six data points in the interval $[k\Delta t, (k+1)\Delta t]$: the price observed at time $k\Delta t$, and a price every second up to, and including, time $(k+1)\Delta t$. The price at time $k\Delta t$ is the opening price, the price at time $(k+1)\Delta t$ is the closing price, and the minimum and maximum prices observed in the samples between are used as a_t and b_t respectively for the likelihood in (2.24). Using this data and the analytic expression for the likelihood, we optimize over μ and σ^2 to find the maximum likelihood estimates for the parameters. Thus, for every simulated day for a given Δt, there is an independent set of estimates for σ^2 and μ.

The results for 1000 simulated trading days are shown in Figure (2.2). We see that, for the shortest sampling periods, the realized volatility estimator outperforms the OCHL estimator in terms of both bias and mean-squared error, as expected. However, the variance of the RV estimator is much greater than that of the OCHL estimator. As a result, the OCHL estimator quickly overtakes the RV estimator in terms of MSE performance.

It is notable that the OCHL estimator has a large bias for small sampling periods. As long as the observations of a diffusion process are finite in number, the observed maximum is less than or equal to the true maximum and the observed minimum is greater than or equal to the true minimum. The discrepancy between the true and observed extreme values becomes greater as the number of observations becomes smaller. This artificially smaller range between extreme values biases the estimates for volatility to be below the true value for smaller intervals $[k\Delta t, (k+1)\Delta t]$ since they contain fewer data points.

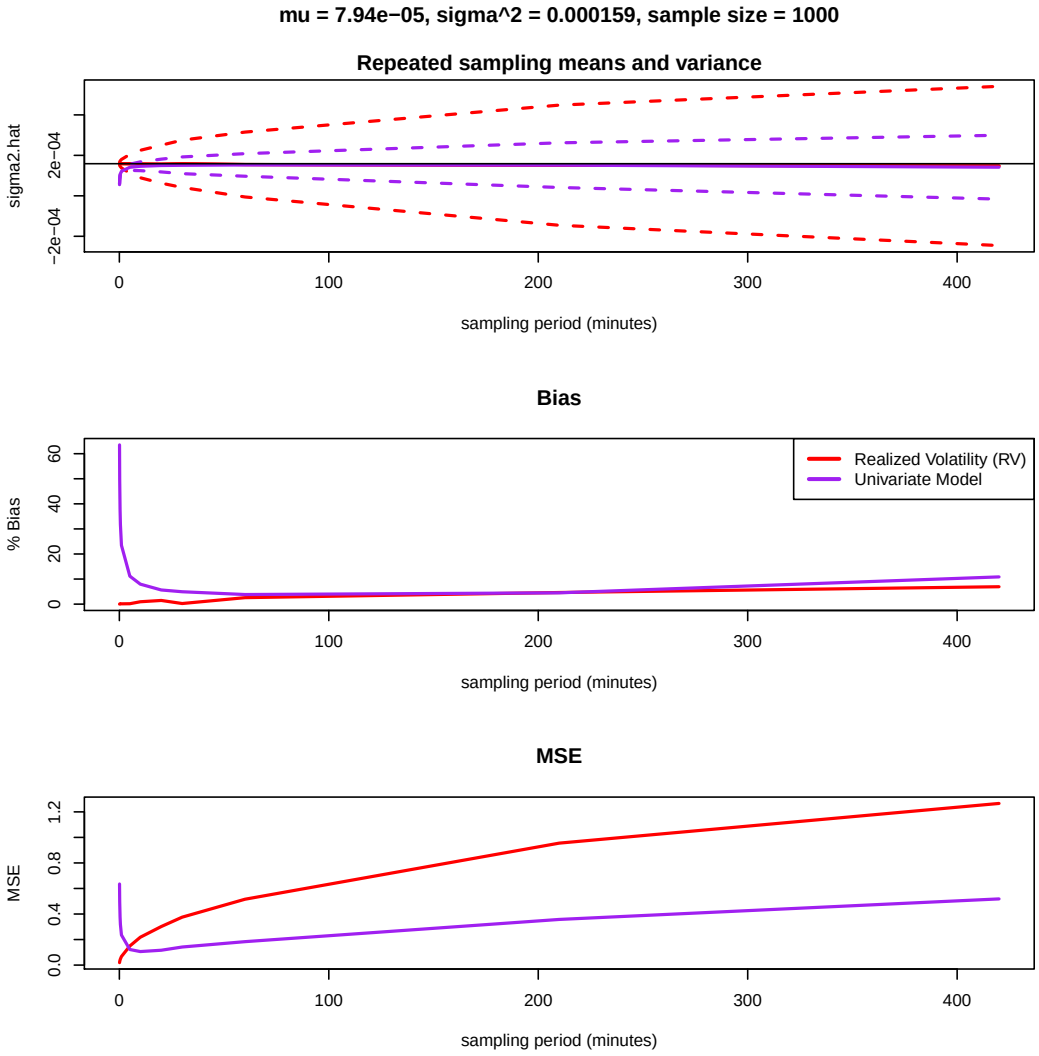


Figure 2.2: Results for 1000 trading days

Chapter 3

Range-Based Methods for High-Frequency Data In Two Dimensions

3.1 Motivation

Just as using the range of an asset, as opposed to only opening and closing values, improves the estimation of drift and volatility, we expect that using OCHL data for two assets can improve estimates for drift and volatility, as well as the estimate of their correlation. Accurate estimation of correlation across assets plays a key role in portfolio theory, as it allows for portfolio diversification to reduce risk.

3.2 Problem Formulation

In the two-dimensional case, we simply extend the geometric Brownian motion model used in Chapter 2 to two dimensions:

$$dY_{1,s} = \mu_1 ds + \sigma_1 dW_{1,s}, \tag{3.1}$$
$$dY_{2,s} = \mu_2 ds + \sigma_2 dW_{2,s}. \tag{3.2}$$

Here we introduce correlation ρ in the log prices of the two assets:

$$Cov(W_1, W_2) = \rho.$$

Again using the Fokker-Planck Equation, we can take the system of SDEs in (3.2) - (3.2) and express the probability density of the asset price $q(y_1, y_2, s)$ as the PDE

$$\begin{aligned} \frac{\partial}{\partial s}q(y,s) &= -\mu_1\frac{\partial}{\partial y_1}q(y_1,y_2,s) - \mu_1\frac{\partial}{\partial y_2}q(y_1,y_2,s) \\ &\quad + \frac{1}{2}\sigma_1^2\frac{\partial^2}{\partial y_1^2}q(y_1,y_2,s) + \frac{1}{2}\sigma_2^2\frac{\partial^2}{\partial^2 y_2}q(y_1,y_2,s) + \rho\sigma_1\sigma_2\frac{\partial^2}{\partial y_1 \partial y_2}q(y_1,y_2,s) \qquad (3.3) \\ q(y_1,y_2,t-1) &= \delta(y_1 - y_{1,t-1})\delta(y_2 - y_{2,t-1}), \qquad (3.4) \end{aligned}$$

where $y_{1,t-1}$ and $y_{2,t-1}$ are the prices at the beginning of the interval, and δ is the Dirac delta function. As in Chapter 2, we introduce the maximum and minimum observed prices of each asset, which establish the boundary conditions for the two-dimensional IC/BV problem

$$\begin{aligned} \frac{\partial}{\partial s}q(y,s) &= -\mu_1\frac{\partial}{\partial y_1}q(y_1,y_2,s) - \mu_1\frac{\partial}{\partial y_2}q(y_1,y_2,s) \\ &\quad + \frac{1}{2}\sigma_1^2\frac{\partial^2}{\partial y_1^2}q(y_1,y_2,s) + \frac{1}{2}\sigma_2^2\frac{\partial^2}{\partial^2 y_2}q(y_1,y_2,s) + \rho\sigma_1\sigma_2\frac{\partial^2}{\partial y_1 \partial y_2}q(y_1,y_2,s) \qquad (3.5) \\ q(y_1,y_2,t-1) &= \delta(y_1 - y_{1,t-1})\delta(y_2 - y_{2,t-1}), \qquad (3.6) \\ q(y_1,y_2,s) &= 0 \quad \text{for} \quad y_1 = a_{1,t}, b_{1,t} \ \text{ or } \ y_2 = a_{2,t}, b_{2,t}. \qquad (3.7) \end{aligned}$$

We use the same treatment as in Chapter 2 to transform the advection-diffusion problem in equations (3.5) - (3.7) into a purely diffusion problem. Letting $\mathbf{a} = (a_1, a_2)^T$ and $\mathbf{y} = (y_1, y_2)^T$, we once again introduce the form

$$q(\mathbf{y}, s) = \exp(\mathbf{a}^T \mathbf{y} + bs) p(\mathbf{y}, s).$$

Substituting the expression into the left and right-hand sides of (3.5) and equating gives us an expression for $\mathbf{a}$ and b which eliminate the zeroeth and first order derivatives of p:

$$b = \mu_1 a_1 + \mu_2 a_2 - \frac{1}{2}\sigma_1^2 a_1^2 - \frac{1}{2}\sigma_2^2 a_2^2 + \rho\sigma_1\sigma_2 a_1 a_2,$$

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}.$$

Under these conditions, the advection-diffusion problem is once again reduced to a pure diffusion problem on a bounded domain:

$$\frac{\partial p(\mathbf{y}, s)}{\partial s} = \frac{1}{2}\sigma_1^2 \frac{\partial^2}{\partial y_1^2} p(\mathbf{y}, s) + \frac{1}{2}\sigma_2^2 \frac{\partial^2}{\partial^2 y_2} p(\mathbf{y}, s) + \rho\sigma_1\sigma_2 \frac{\partial^2}{\partial y_1 \partial y_2} p(\mathbf{y}, s), \tag{3.8}$$

$$p(\mathbf{y}, t-1) = \delta(y_1 - y_{1,t-1})\delta(y_2 - y_{2,t-1}), \tag{3.9}$$

$$p(\mathbf{y}, s) = 0 \quad \text{for} \quad y_1 = a_{1,t}, b_{1,t} \ \text{ or } \ y_2 = a_{2,t}, b_{2,t}. \tag{3.10}$$

3.3 Solution with the Method of Images

Our first attempt to solve the two-dimensional diffusion problem was to apply the Method of Images approach in Chapter 2. The fundamental solution to the diffusion problem on an unbounded domain is the bivariate normal density with mean and covariance matrix

$$\boldsymbol{\mu} = (y_{1,t-1}, y_{2,t-1})^T, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

However, reflection of the fundamental solution across a certain boundary produces an image which satisfies a diffusion equation with a *negated* correlation coefficient ρ (Figure (3.1a)), so that the differential equation satisfied by the reflected function is no longer the original PDE (note the minus sign in front of the mixed derivative term):

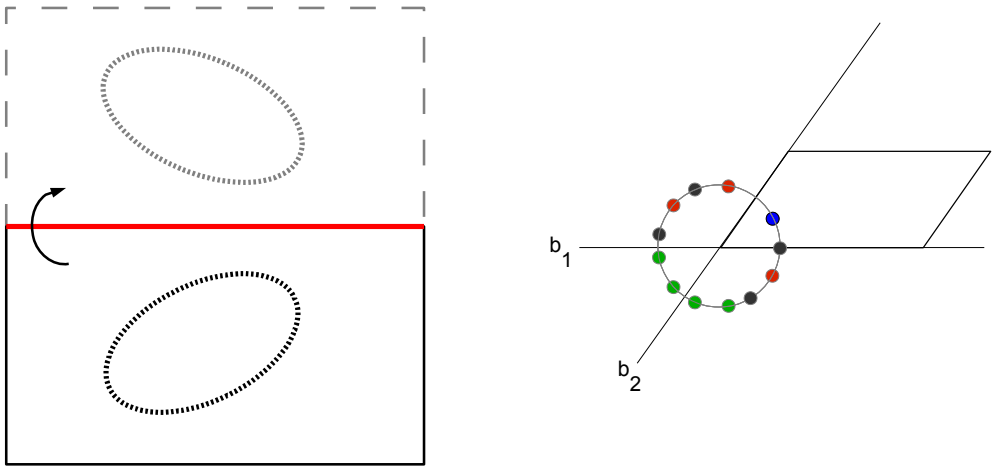
$$\frac{\partial p(\mathbf{z}, s)}{\partial s} = \frac{1}{2}\sigma_1^2 \frac{\partial^2}{\partial z_1^2} p(\mathbf{z}, s) + \frac{1}{2}\sigma_2^2 \frac{\partial^2}{\partial^2 z_2} p(\mathbf{z}, s) - \rho\sigma_1\sigma_2 \frac{\partial^2}{\partial z_1 \partial z_2} p(\mathbf{z}, s)$$

Hence, the summation of reflected images, although satisfying the boundary conditions in the limiting case, does not satisfy the original PDE.

One can eliminate the correlation term ρ through a change of coordinates consisting of a rotation, given by the eigenvalue decomposition of the covariance matrix, and a translation, such that the diffusion problem is reduced even further to

$$\frac{\partial p(\boldsymbol{\zeta}, s)}{\partial s} = \frac{1}{2}\tilde{\sigma}_1^2 \frac{\partial^2}{\partial \zeta_1^2} p(\boldsymbol{\zeta}, s) + \frac{1}{2}\tilde{\sigma}_2^2 \frac{\partial^2}{\partial^2 \zeta_2} p(\boldsymbol{\zeta}, s).$$

However, the boundary where the solution is zero is no longer a rectangle but is instead a parallelogram. The non-orthogonality of the boundaries makes the process of reflecting the fundamental solution about the boundaries problematic. Particularly, if the angles between intersecting sides are not factors of π, repeated reflection places images within the computational domain of the problem, thereby violating the initial condition (see Figure (3.1b)).



(a) When the fundamental solution (covariance ellipsoid in black) is reflected across the red boundary, the reflection has a covariance ellipsoid (gray) with a negative slope, meaning that the reflection satisfies a PDE where the correlation coefficient is negative. The sum of the fundamental solution and the reflection satisfies the boundary condition over the red line but does not satisfy the governing PDE.

(b) The fundamental solution, centered on the initial condition and denoted by the blue-inscribed red disk, is reflected alternatively around the boundaries b_1 and b_2. The displayed system of images is the results of three successive pairs of reflections around b_1 then b_2, with red corresponding to the first, green to the second, and black to the third set of reflections. We see that the system includes a fundamental solution in the computational domain in additional to the original, thereby violating the initial conditions of the problem.

Figure 3.1: Construction of the solution using the Method of Images

3.4 Solution by Fourier Expansion

Similarly to the approach in Chapter 2, equation (2.16), we offer a solution to the diffusion problem in terms of a trigonometric expansion of sine functions, since the basis functions must satisfy the boundary conditions:

$$p(\mathbf{y},s) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{m,n}(s)\sin\left(\frac{m\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\sin\left(\frac{n\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right) \tag{3.11}$$

We substitute the above form into equation (3.8) in order to obtain expressions for $C_{m,n}(t)$:

$$\begin{aligned}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{dC_{m,n}(s)}{ds}\sin\left(\frac{m\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\sin\left(\frac{n\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right) = \\ \frac{1}{2}\sigma_1^2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}C_{m,n}(s)\left(\frac{m\pi}{b_{1,t}-a_{1,t}}\right)^2\sin\left(\frac{m\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\sin\left(\frac{n\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right) \\ + \frac{1}{2}\sigma_2^2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}C_{m,n}(s)\left(\frac{n\pi}{b_{2,t}-a_{2,t}}\right)^2\sin\left(\frac{n\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\sin\left(\frac{n\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right) \\ + \rho\sigma_1\sigma_2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}C_{m,n}(s)\left(\frac{m\pi}{b_{1,t}-a_{1,t}}\right)\left(\frac{n\pi}{b_{2,t}-a_{2,t}}\right)\cos\left(\frac{m\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\cos\left(\frac{n\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right)\end{aligned} \tag{3.12}$$

To deal with the cosine terms, we expand them in terms of sine functions as well, obtaining

$$\cos\left(\frac{m\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\cos\left(\frac{n\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right) = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}d_k^{(m)}d_j^{(n)}\sin\left(\frac{k\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\sin\left(\frac{j\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right).$$

Using the orthogonality of the sine and cosine series with respect to the L^1 norm

$$d_k^{(m)} = \begin{cases} 0 & \text{if } k-m = \text{ even} \\ \frac{4k}{\pi(k^2-m^2)} & \text{if } k-m = \text{ odd.}\end{cases}$$

Focusing on the last line with the cosine terms

$$\begin{aligned}&\rho\sigma_1\sigma_2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}C_{m,n}(s)\left(\frac{m\pi}{b_{1,t}-a_{1,t}}\right)\left(\frac{n\pi}{b_{2,t}-a_{2,t}}\right)\cos\left(\frac{m\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\cos\left(\frac{n\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right) \\ &= \rho\sigma_1\sigma_2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}C_{m,n}(s)\left(\frac{m\pi}{b_{1,t}-a_{1,t}}\right)\left(\frac{n\pi}{b_{2,t}-a_{2,t}}\right)d_k^{(m)}d_j^{(n)}\sin\left(\frac{k\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\sin\left(\frac{j\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right)\end{aligned} \tag{3.13}$$

We want the above sine terms to be indexed by m and n so that the entire expression may be in the same form as the other two terms in the right-hand side of (3.12). This allows us to develop a system of ordinary differential equations for $C_{m,t}(t)$, whose solutions give $p(\mathbf{y},t)$. By Fubini's Theorem, we can exchange the order of summation in (3.13), relabel indexes, and group terms, obtaining

$$\begin{aligned}&\rho\sigma_1\sigma_2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}C_{m,n}(s)\left(\frac{m\pi}{b_{1,t}-a_{1,t}}\right)\left(\frac{n\pi}{b_{2,t}-a_{2,t}}\right)d_k^{(m)}d_j^{(n)}\sin\left(\frac{k\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\sin\left(\frac{j\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right) \\ &= \rho\sigma_1\sigma_2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}C_{k,j}(s)\left(\frac{k\pi}{b_{1,t}-a_{1,t}}\right)\left(\frac{j\pi}{b_{2,t}-a_{2,t}}\right)d_m^{(k)}d_n^{(j)}\right]\sin\left(\frac{m\pi(y_1-a_{1,t})}{b_{1,t}-a_{1,t}}\right)\sin\left(\frac{n\pi(y_2-a_{2,t})}{b_{2,t}-a_{2,t}}\right)\end{aligned} \tag{3.14}$$

Now plugging (3.14) into (3.12) we can obtain the following matrix representation for the system of ODEs for the coefficients $C_{m,n}(s)$. Truncating the expansion to a finite M and N terms over the indexes

m and n respectively, the system has the form

$$\frac{d}{ds}\begin{pmatrix} C_{1,1}(s) \\ C_{1,2}(s) \\ \vdots \\ C_{1,N}(s) \\ C_{2,1}(s) \\ \vdots \\ C_{2,N}(s) \\ \vdots \\ C_{M,N}(s) \end{pmatrix} = \frac{1}{2}\sigma_1^2\mathbf{A}_1 + \frac{1}{2}\sigma_2^2\mathbf{A}_2 + \rho\sigma_1\sigma_2\mathbf{B}\begin{pmatrix} C_{1,1}(s) \\ C_{1,2}(s) \\ \vdots \\ C_{1,N}(s) \\ C_{2,1}(s) \\ \vdots \\ C_{2,N}(s) \\ \vdots \\ C_{M,N}(s) \end{pmatrix} \tag{3.15}$$

where $\mathbf{A}_1, \mathbf{A}_2$, and $\mathbf{B}$ are matrices of size $M \times N$. We can map from the tuple (m,n) to r, the index of the column vector containing the $C_{m,n}(t)$ terms, via the transformation $r = (m-1)N + n$. The inverse of the transformation is $m = \xi(r) = \text{ceil}(r/N)$ and $n = \nu(r) = \text{mod}(r-1,N)+1$. Both $\mathbf{A}_1$ and $\mathbf{A}_2$ are diagonal matrices with entries

$$[\mathbf{A}_1]_{r,r} = \left(\frac{\xi(r)\pi}{b_{1,t} - a_{1,t}}\right)^2,$$

$$[\mathbf{A}_1]_{r,r} = \left(\frac{\nu(r)\pi}{b_{2,t} - a_{2,t}}\right)^2.$$

The matrix $\mathbf{B}$ is not diagonal and has entries

$$[\mathbf{B}]_{r,s} = \left(\frac{\xi(s)\pi}{b_{1,t} - a_{1,t}}\right)\left(\frac{\nu(s)\pi}{b_{2,t} - a_{2,t}}\right) d_{\xi(r)}^{(\xi(s))} d_{\nu(r)}^{(\nu(s))}.$$

Letting

$$\mathbf{A} = \frac{1}{2}\sigma_1^2\mathbf{A}_1 + \frac{1}{2}\sigma_2^2\mathbf{A}_2 + \rho\sigma_1\sigma_2\mathbf{B} \tag{3.16}$$

$$\mathbf{C}(s) = (C_{1,1}(s), C_{1,2}(s), \ldots, C_{1,N}(s), \ldots, C_{M,N}(s))^T$$

the system can be written simply as

$$\frac{d}{ds}\mathbf{C}(s) = \mathbf{A}\mathbf{C}(s).$$

Due to the fact that $\mathbf{B}$ is a symmetric matrix, $\mathbf{A}$ is invertible, so that its eigenvalue decomposition exists. Denoting the eigenvector and eigenvalue matrices as $\mathbf{V}^T$ and $\mathbf{\Lambda}$, respectively, the solution to the system is given by

$$\mathbf{C}(t) = \exp(\mathbf{A}t)\mathbf{C}(0) = \left(\mathbf{V}\exp(\mathbf{\Lambda}t)\mathbf{V}^T\right)\mathbf{C}(0).$$

The vector $\mathbf{C}(0)$ is found by expanding the initial condition to the diffusion problem with sine basis functions.

It should be noted here that if $\rho = 0$, the third term in the right-hand side of equation (3.15) drops. Hence, the problem is reduced to two independent single-dimensional problems, whose solutions were found in Chapter 2. Because $\mathbf{A}_1$ and $\mathbf{A}_2$ are diagonal, under this scenario the number of terms N and M needed to obtain good accuracy is relatively small. However, for nonzero ρ, $\mathbf{B}$ is non-diagonal and dense.

3.5 Finding the Likelihood for the Closing Prices

Using the shorthand

$$s_{m,n} = \sin\left(\frac{m\pi(y_1 - a_{1,t})}{b_{1,t} - a_{1,t}}\right)\sin\left(\frac{n\pi(y_2 - a_{2,t})}{b_{2,t} - a_{2,t}}\right),$$

the vector

$$\mathbf{s} = (s_{1,1}, s_{1,2}, \ldots, s_{1,N}, s_{2,1}, \ldots s_{M,N})^T$$

contains the bases for the trigonometric expansion in the problem. The solution to the advection-diffusions problem is therefore

$$q(\mathbf{y}, s) = \exp(\mathbf{a}^T\mathbf{y} + bs)(\exp(\mathbf{A}s)\mathbf{C}(0))^T\mathbf{s}.$$

To obtain the likelihood for the OCHL data in this model, we need to once again differentiate q with respect to the boundary values. Since there are four boundaries, this will be a fourth-order derivative.

$$P(m_{1,t} = a_{1,t}, m_{2,t} = a_{2,t}, M_{1,t} = b_{1,t}, M_{2,t} = b_{2,t}, Y_{1,t} = y_1, Y_{2,t} = y_2 | Y_{1,t-1} = y_{1,t-1}, Y_{2,t-1} = y_{2,t-1}) =$$

$$\frac{\partial^4 q(\mathbf{y}, t)}{\partial a_{1,t} \partial a_{2,t} \partial b_{1,t} \partial b_{2,t}} \tag{3.17}$$

The simplest approach to computing the above derivative is by using a numerical differentiation scheme. Hence, the likelihood in equation (3.17) is initially computed using a second-order, four-point sentcil numerical differentiation scheme in line with that used in Chapter 2. A major limitation of numerical differentiation of higher orders is the finite precision of number representation systems in digital computers where the attributed round-off error is of order 10^{-16}. When $O(\Delta x^4) < 10^{-16}$, the effect of the round-off error contributes $O(1)$ to the error in the numerical derivative. The limitations of using a purely numerical scheme to compute the likelihood is evident in our simulations.

An alternative approach is to compute the derivatives analytically. The high-order derivatives of the vectors $\mathbf{C}(0)$ and $\mathbf{s}$ can be readily found. However, this is not the case for the matrix $e^{\mathbf{A}s}$. The application of a function to a matrix is defined in terms of its eigenvalue decomposition

$$\exp(\mathbf{A}s) = \mathbf{V}\exp(\mathbf{\Lambda}t)\mathbf{V}^T.$$

Since the eigenvalues and eigenvectors of the system matrix are dependent upon the boundary values, and we have no closed form expression for the eigenvalues and eigenvectors, we must define the analytic derivatives of $\exp(\mathbf{A}s)$ in terms of $\mathbf{V}$ and $\mathbf{\Lambda}$. Wilcox (1967) provides the analytic expressions for second-order derivatives of functions of matrices. In general, if a function f is analytic and a matrix $\mathbf{H}$ is positive-definite,

$$\mathbf{v}_i^T \frac{\partial^2 f(\mathbf{H})}{\partial a_t \partial b_t} \mathbf{v}_j = \mathbf{v}_i^T (\mathbf{A}_{a_t b_t} + \mathbf{A}_{b_t a_t} + \mathbf{B}_{a_t b_t}) \mathbf{v}_j,$$

where

$$\mathbf{v}_i^T \mathbf{A}_{a_t b_t} \mathbf{v}_j = \sum_l \mathbf{v}_i^T \frac{\partial \mathbf{H}}{\partial a_t} \mathbf{v}_l \mathbf{v}_l^T \frac{\partial \mathbf{H}}{\partial b_t} \mathbf{v}_j S_{ilj} \tag{3.18}$$

and

$$\mathbf{v}_i^T \mathbf{B}_{a_t b_t} \mathbf{v}_j = \mathbf{v}_i^T \frac{\partial^2 \mathbf{H}}{\partial a_t \partial b_t} \mathbf{v}_j T_{ij}. \tag{3.19}$$

In equations (3.18) and (3.19), S_{ilj} and T_{ij} are defined by

$$\begin{aligned} S_{ilj} &= \frac{1}{2}\Delta_{il}\Delta_{lj} f''(\lambda_i) + U_{ilj} + U_{jil} + U_{lji}, \\ T_{ij} &= \Delta_{ij} f'(\lambda_i) + \rho_{ij} \frac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j}, \end{aligned}$$

where

$$\Delta_{ij} \equiv 1 - \rho_{ij} \equiv \begin{cases} 1, & \lambda_i = \lambda_j, \\ 0, & \lambda_i \neq \lambda_j, \end{cases} \tag{3.20}$$

$$U_{ilj} \equiv \frac{\rho_{il}\rho_{lj}\rho_{ji} f(\lambda_i)}{(\lambda_i - \lambda_l)(\lambda_i - \lambda_j)} + \Delta_{il}\rho_{lj}\left[\frac{f(\lambda_j) - f(\lambda_i)}{(\lambda_j - \lambda_i)^2} - \frac{f'(\lambda_i)}{\lambda_j - \lambda_i}\right]. \tag{3.21}$$

The same technique as in Wilcox (1967) can be used to derive analytic expressions for higher-order derivatives. However, these expressions proved to be too cumbersome to be used effectively.

With the above expressions, the derivative in equation (3.17) is computed by first finding second-order derivatives analytically, with respect to $a_{1,t}$ and $b_{1,t}$, while the remaining differentiation is completed numerically using the same scheme as in Chapter 2, leading to an error of $O(\Delta x^2)$. Particularly, we first compute using the analytic expressions above

$$\frac{\partial^2 q(\mathbf{y},t)}{\partial a_{1,t}\partial b_{1,t}} \equiv g(a_{2,t},b_{2,t}),$$

after which the final derivative is found numerically

$$\frac{\partial^4 q(\mathbf{y},t)}{\partial a_{1,t}\partial b_{1,t}\partial a_{2,t}\partial b_{2,t}} = \frac{g(a_{2,t}+\Delta x,b_{2,t}+\Delta x)-g(a_{2,t}-\Delta x,b_{2,t}+\Delta x)-g(a_{2,t}+\Delta x,b_{2,t}-\Delta x)+g(a_{2,t}-\Delta x,b_{2,t}-\Delta x)}{4\Delta x^2}.$$

To understand the accuracy of the purely numerical and mixed approaches to computing the likelihood, we considered are five simulations over a nominal period of length 1 based on a forward-Euler discretization scheme of the governing stochastic differential equations:

$$\begin{aligned} Y_{1,k+1} &= Y_{1,k}+\mu_1\Delta t^*+\sigma_1\Delta t^*\varepsilon_{1,k}, \\ Y_{2,k+1} &= Y_{2,k}+\mu_2\Delta t^*+\sigma_2\Delta t^*\varepsilon_{1,k}. \end{aligned}$$

with $\Delta t^* = 1/23,400$, $\varepsilon_{1,k},\varepsilon_{2,k} \sim N(0,1)$, and the correlation between the two innovations is ρ. This is equivalent to sampling the returns once every second over the length of a trading day. Further,

$$\mu_{1,2} = 0, \qquad \sigma_{1,2} = 1, \qquad \rho = (0,0.3,0.5)$$

As in Chapter 2, for each simulation there is a recorded min, max, and closing value, while all open log prices are set to 0. The likelihood is evaluated at the closing return using the observed minima and maxima, as well as the true μ,σ and ρ used to generate the data. When performing numerical differentiation, the numerical discretization method is used with a discretization step

$$\Delta x = \min\left\{\frac{1}{2^k}\frac{1}{100}(b_{1,t}-a_{1,t}),\frac{1}{2^k}\frac{1}{100}(b_{2,t}-a_{2,t})\right\}$$

Values of the likelihoods are considered across ranges of the number of expansion terms $N = M$ and k. Numerical and mixed differentiation results are compared. The results are shown in Tables (3.1) - (3.6).

Table (3.3), where likelihoods are found purely numerically, contains $-\infty$ entries, which is an indication that the numerical differentiation failed. This is corroborated by Table (3.4), where likelihoods are computed using the mixed approach for the same data, and the table contains no such entries. The same derivatives are computed in Tables (3.8) and (3.7), where the Δx factor is now greater since $k = 6$ is used instead of $k = 8$. We see no $-\infty$ entries, and further the values for mixed versus purely numerical differentiation are closer than those obtained with a smaller differentiation step, suggesting that a factor of $k = 6$ strikes the balance between numerical accuracy without running into issues related to round-off error. Further, although small, the discrepancies between derivatives obtained through the mixed method and those obtained with a purely numerical scheme suggest the utility of finding and using the expressions for the analytic fourth-order derivatives of $q(\mathbf{y},t)$.

Finally, for $\rho = 0$, (Tables (3.5) and (3.6)), we see a clear convergence of likelihood values across $N = M$. This is expected, as in this case we are solving two independent univariate problems that inherit the efficiency of the method in Chapter 2. For moderate $\rho = 0.3$ (Tables (3.3) and (3.4)) and higher $\rho = 0.5$ (Tables (3.1) and (3.2)), we do not see a clear convergence of the likelihood values as N increases, meaning that the expansions used in the computations need to be greater. However, the system matrix size for $N = 64$ is already prohibitively large being 4096×4096 in terms of eigenvalue decomposition schemes for dense matrices. This in turn suggests the need for a more efficient basis expansion of the solution to the diffusion problem.

$N=M=4$	$N=M=8$	$N=M=16$	$N=M=32$	$N=M=64$
-3.3403	-3.4383	-3.871	-3.9307	-3.9127
-1.9465	-2.121	-2.0546	-2.0336	-2.0455
-1.9417	-1.7274	-1.5652	-1.6451	-1.6099
-5.4395	-5.0203	-5.4614	-4.8342	-5.7621
-0.6593	-0.6242	-0.6136	-0.5658	-0.6366

Table 3.1: Likelihood values for simulations obtained via numerical differentiation with $k = 8$ ($\Delta x = 1/(100 \cdot 2^8) \cdot (b_t - a_t)$) and $\rho = 0.5$

$N=M=4$	$N=M=8$	$N=M=16$	$N=M=32$	$N=M=64$
-3.3393	-3.4375	-3.8738	-3.9349	-3.8985
-1.947	-2.1215	-2.0548	-2.0339	-2.0457
-1.9413	-1.7269	-1.5639	-1.6444	-1.6007
-5.4417	-5.0203	-5.4664	-4.7916	-6.5573
-0.6592	-0.6241	-0.6134	-0.5639	-0.6359

Table 3.2: Likelihood values for simulations obtained via mixed differentiation with $k = 8$ ($\Delta x = 1/(100 \cdot 2^8) \cdot (b_t - a_t)$) and $\rho = 0.5$

$N=M=4$	$N=M=8$	$N=M=16$	$N=M=32$	$N=M=64$
-1.8564	-1.7292	-1.8217	-1.4718	-Inf
-0.8215	-0.8457	-0.8811	-1.1449	-0.0892
-2.1068	-Inf	-Inf	0.4876	2.3472
-1.3503	-1.3519	-1.2759	-1.2693	-1.3897
-1.7123	-1.7156	-1.675	-1.8193	-1.5801

Table 3.3: Likelihood values for simulations obtained via numerical differentiation with $k = 8$ ($\Delta x = 1/(100 \cdot 2^8) \cdot (b_t - a_t)$) and $\rho = 0.3$

$N=M=4$	$N=M=8$	$N=M=16$	$N=M=32$	$N=M=64$
-1.83	-1.7475	-1.7601	-1.7416	-1.6365
-0.8209	-0.8642	-0.8486	-0.8423	-0.8441
-2.2024	-2.3046	-2.2768	-2.2912	-2.2828
-1.3503	-1.352	-1.2759	-1.2693	-1.3897
-1.7109	-1.7149	-1.7272	-1.7209	-1.7199

Table 3.4: Likelihood values for simulations obtained via mixed differentiation with $k = 8$ ($\Delta x = 1/(100 \cdot 2^8) \cdot (b_t - a_t)$) and $\rho = 0.3$.

$N=M=4$	$N=M=8$	$N=M=16$	$N=M=32$	$N=M=64$
-2.1731	-2.1731	-2.1731	-2.1731	-2.1731
-2.2388	-2.2388	-2.2388	-2.2388	-2.2388
-1.6973	-1.6973	-1.6973	-1.6973	-1.6973
-0.7659	-0.7659	-0.7659	-0.7659	-0.7659
-4.7948	-4.5942	-4.5942	-4.5942	-4.5942

Table 3.5: Likelihood values for simulations obtained via numerical differentiation with $k = 8$ ($\Delta x = 1/(100 \cdot 2^8) \cdot (b_t - a_t)$) and $\rho = 0$

$N = M = 4$	$N = M = 8$	$N = M = 16$	$N = M = 32$	$N = M = 64$
-2.1731	-2.1731	-2.1731	-2.1731	-2.1731
-2.2926	-2.2926	-2.2926	-2.2926	-2.2926
-1.6972	-1.6972	-1.6972	-1.6972	-1.6972
-0.7693	-0.7693	-0.7693	-0.7693	-0.7693
-4.3054	-4.3054	-4.3054	-4.3054	-4.3054

Table 3.6: Likelihood values for simulations obtained via mixed differentiation with $k = 8$ ($\Delta x = 1/(100 \cdot 2^8) \cdot (b_t - a_t)$) and $\rho = 0$.

$N = M = 4$	$N = M = 8$	$N = M = 16$	$N = M = 32$	$N = M = 64$
-1.83	-1.7475	-1.7601	-1.7421	-1.6453
-0.8209	-0.8642	-0.8486	-0.8425	-0.8356
-2.2036	-2.3077	-2.2476	-2.3117	-1.7568
-1.3503	-1.352	-1.2759	-1.2694	-1.3897
-1.711	-1.7149	-1.7271	-1.7207	-1.7205

Table 3.7: Likelihood values for simulations obtained via numerical differentiation with $k = 6$ ($\Delta x = 1/(100 \cdot 2^6) \cdot (b_t - a_t)$) and $\rho = 0.3$

$N = M = 4$	$N = M = 8$	$N = M = 16$	$N = M = 32$	$N = M = 64$
-1.83	-1.7475	-1.7601	-1.7416	-1.6365
-0.8209	-0.8642	-0.8486	-0.8423	-0.8441
-2.2024	-2.3046	-2.2768	-2.2912	-2.2828
-1.3503	-1.352	-1.2759	-1.2693	-1.3897
-1.7109	-1.7149	-1.7272	-1.7209	-1.7199

Table 3.8: Likelihood values for simulations obtained via mixed differentiation with $k = 6$ ($\Delta x = 1/(100 \cdot 2^6) \cdot (b_t - a_t)$) and $\rho = 0.3$.

Chapter 4

Future Work

4.1 Stochastic Volatility Estimation With Microstructure Noise

The first part of our proposed future work is to account for microstructure noise in the context of stochastic volatility models for high-frequency data. To our knowledge, currently stochastic volatility models are fit to high-frequency data by matching realized volatility estimates or range-based estimates to state-space representations of volatility derived from the stochastic volatility models; or by matching higher moments of the model to higher-order realized volatility estimates, as described in Section 1.3.2. Our proposed approach is to simply include observational noise in the discrete-time version of the Orstein-Uhlenbeck process:

$$\begin{aligned} Y_{t+1} &= \log(S_{t+1}) + \xi\gamma_t \\ \log(S_{t+1}) &= \log(S_t) + \mu + \sigma_{t+1}\varepsilon_{t+1,1} \\ \log(\sigma_{t+1}) &= \alpha + \phi[\log(\sigma_t) - \alpha] + \beta\varepsilon_{t+2,2}. \end{aligned}$$

with $\gamma_t \sim N(0,1)$, and the tuple $(\varepsilon_{t+1,1}, \varepsilon_{t+1,2})$ following a bivariate normal distribution with $E(\varepsilon_{t,1}) = E(\varepsilon_{t,2}) = 0$, $\text{Var}(\varepsilon_{t,1}) = \text{Var}(\varepsilon_{t,2}) = 1$ and $\text{Cor}(\varepsilon_{t,1}, \varepsilon_{t,2}) = \rho$. By allowing a nonzero correlation between these innovations, the model incorporates possible leverage effects.

We determine the variance of microstructure noise ξ by considering that this effect arises from the discretization of price quotation and the spread between bid and ask prices on the exchanges. Let D be the maximum of the of the discrete price size and the average bid-ask spread of the considered asset price and let Q be the average price level of the asset. We derive an approximation of ξ by considering a first-order Taylor approximation of $\log(S_{t+1} + \xi^*\gamma_{t+1})$, the log of the true price S_{t+1} plus noise with a standard deviation $\xi = D/2$. In this way most measurement errors are within an interval D of the true price. Approximating $1/S_{t+1}$ with $1/Q$, we have

$$\log(S_{t+1} + \xi^*\gamma_{t+1}) \approx \log(S_{t+1}) + \frac{\xi^*}{Q}\gamma_{t+1}.$$

In this way, as a point estimate, $\xi = D/2Q$. In terms of a Bayesian estimation of the model, we propose a strong prior for ξ:

$$\xi \sim \text{Gamma}\left(\frac{D^2}{4Q^2\delta}, \frac{D}{2Q\delta}\right),$$

such that $E(\xi) = D/2Q$ and $\text{Var}(\xi) = \delta$, where δ is small. The standard deviation ξ of the noise process needs a strong specification; otherwise the model is not identifiable since the effects of noise and the latent volatility process $\{\log(\sigma_t)\}$ cannot be distinguished.

We will fit the above model using an MCMC algorithm. Conditioning on the volatility series and ξ, we can sample $\{S_t\}$ using forward-filtering backward-sampling. In turn, we can estimate the series $\{\log(\sigma_t)\}$ using the algorithm of Omori et al. (2007). In this manner we are resolving the additional level of latency in our model by using the DLM structure of the log-return series.

Once we have estimated the series $\{\log(\sigma_t)\}$, we can recover the posterior distribution of the integrated volatility by computing the sums $IV^{(b)} = \sum_{t=1}^{T} \exp(2\log(\sigma_t^{(b)}))$. The thus derived estimates

for integrated volatility will be compared to Realized Volatility estimates which also account for microstructure noise to measure the utility of our model.

4.1.1 Using maximum and minimum prices for estimating volatility

For very liquid markets with multiple transactions every millisecond, the approach in Section 4.1 can be computationally intensive. By working with longer intervals (of the order of 1 second, for example) we can reduce this computational cost, but instead of discarding all observations within sampling intervals we can include the intra-period maximum and minimum price to improve inference. However, data at this frequency is bound to be contaminated by microstructure noise, so the model from Chapter 2 is not directly applicable in this setting.

Instead, we need the likelihood for the opening, closing, high, and low *observed* returns. The difficulty in deriving this likelihood comes from the fact that microstructure noise does not scale with the observed sampling period length, so there is no immediate stochastic differential form for this type of process, and the Fokker-Planck equation cannot be invoked. One possible starting point in considering this problem is geometric Brownian motion with jumps, where jumps are observed at every sampling point. However, in this setting jumps drive the process, which is equivalent to microstructure noise affecting the evolution of asset prices.

4.2 Bernstein Polynomials

In this part of our work, we aim to improve the computational efficiency of the solution to the two-dimensional problem in Chapter 3. As the simulations in Section 3.5 demonstrated, for moderate to large ρ the trigonometric expansion needs a large number of terms to get a good approximation of the solution. This feature of the Fourier expansion, which is due to the separability of the basis functions, coupled with the fact that the system matrix $\mathbf{A}$ is dense, make accurate numerical solutions very expensive. Hence, we are motivated to seek a better basis representation for our problem.

In this case, a natural alternative to Fourier bases are the Bernstein polynomials, where the $n+1$ basis polynomials of degree n are defined as

$$B_{\nu,n}(x) = \binom{n}{\nu} x^{\nu}(1-x)^{n-\nu}, \quad \nu = 0,\ldots,n, \tag{4.1}$$

and the solution has the form

$$q(y_1,y_2,t) = \sum_{n=1}^{N}\sum_{m=1}^{M} C_{m,n}(t) B_{n,N}(y_1) B_{m,M}(y_2).$$

In addition to having the smoothness properties needed to represent the solution to the diffusion problem, these basis functions are defined on a bounded domain. Further, the derivatives of each basis element can be written as the sum of three other basis functions

$$\begin{aligned} B'_{\nu,n}(x) &= n(B_{\nu-1,n-1}(x) - B_{\nu,n-1}(x)) \\ &= n\left(\frac{n-(\nu-1)}{n}B_{\nu-1,n}(x) + \frac{(\nu-1)+1}{n}B_{\nu,n}(x) + \frac{n-\nu}{n}B_{\nu,n}(x) + \frac{\nu+1}{n}B_{\nu+1,n}(x)\right) \\ &= n\left(\frac{n-(\nu-1)}{n}B_{\nu-1,n}(x) + B_{\nu,n}(x) + \frac{\nu+1}{n}B_{\nu+1,n}(x)\right), \end{aligned}$$

so that the second-order derivatives can be written as the sum of nine other basis functions. The resultant system matrix would therefore be banded. Even though the system matrix will not be necessarily sparse, it will have a known structure that will allow for more efficient eigenvalue decomposition. Additionally, Bernstein polynomials are the kernels of the Beta distribution and so they are characteristically similar to the solution to the heat equation on a bounded domain (Figure (2.1c)). As such, they should be a more efficient method of expansion of the solution of the heat equation.

As described, the solution with Bernstein polynomials should make computing the solution to diffusion problem more efficient, but it still lacks correlation structure. We can potentially introduce correlation structure by writing the expansion as

$$q(y_1, y_2, s) = \sum_{n=1}^{N} \sum_{m=1}^{M} C_{m,n}(s) \mathcal{K}\Big(B_{n,N}(y_1), B_{m,M}(y_2) \Big).$$

A natural starting place where we can look for functions $\mathcal{K}$ are families of copulas, which are functions on marginal CDFs whose effect is introducing correlation between continuous random variables.

4.3 Bivariate Stochastic Volatility Model using Open, High, Low, and Closing Data

We plan to formulate and estimate a birvariate stochastic volatility model that uses results from Chapter 3 and Section 4.2. For this, we need to specify a two-dimensional stochastic volatility model. In keeping with previous formulation, we consider the discrete-time setting

$$\begin{pmatrix} \log(S_{t+1,1}) \\ \log(S_{t+1,2}) \end{pmatrix} = \begin{pmatrix} \log(S_{t,1}) \\ \log(S_{t,2}) \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \mathbf{V_{t+1}} \begin{pmatrix} w_{t+1,1} \\ w_{t+1,2} \end{pmatrix} \tag{4.2}$$

so that

$$\begin{pmatrix} r_{t+1,1} \\ r_{t+1,2} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \mathbf{\Sigma}_{t+1} \right), \qquad \mathbf{\Sigma}_{t+1} = \mathbf{V}'_{t+1} \mathbf{V}_{t+1} \tag{4.3}$$

where $\mathbf{V}_{t+1}$ is the upper triangular Cholesky component of the covariance matrix. The interesting thing here is specifying how the covariance matrix $\mathbf{\Sigma}_{t+1}$ evolves over time, with no shortage of suggested methods present in the literature. One method is the Wishart random walk approach first suggested by Quintana and West (1987). Starting with the Inverse Wishart prior for the covariance matrix, $\mathbf{\Sigma}_{t+1}$ is allowed to slowly change over time through multiplication by a discount factor

$$(\mathbf{\Sigma}_t | \mathbf{D}_t) \sim IW(h_t, \mathbf{D}_t) \quad \rightarrow \quad (\mathbf{\Sigma}_{t+1} | \mathbf{D}_t) \sim IW(1/\beta h_t, 1/\beta \mathbf{D}_t).$$

A more general version of the above procedure, due to Uhlig (1997), proposes sampling $\mathbf{\Gamma}_{t+1}$ from matrix beta distribution with a discount factor β

$$(\mathbf{\Gamma}_{t+1} | \mathcal{D}_t) \sim Be(\beta h_t/2, (1-\beta) h_t/2),$$

then setting

$$\mathbf{\Phi}_{t+1} = \mathbf{U}'_t \mathbf{\Gamma}_{t+1} \mathbf{U}_t,$$

where $\mathbf{\Phi}_t = \mathbf{\Sigma}_t^{-1}$ and $\mathbf{U}_t$ is the upper triangular Cholesky component of the precision matrix $\mathbf{\Phi}_t$. In both approaches, however, the models do not have much flexibility with respect to correlation structure.

A more intuitive approach in the bivariate case is to consider the evolution of $\mathbf{\Sigma}_t$ component-wise. Writing down

$$\mathbf{\Sigma}_t = \begin{pmatrix} \sigma_{t,1}^2 & \rho_t \sigma_{t,1} \sigma_{t,2} \\ \rho_t \sigma_{t,1} \sigma_{t,2} & \sigma_{t,2}^2 \end{pmatrix},$$

we can define the updating process for the covariance matrix through a linear, autoregressive process

$$\begin{pmatrix} \log(\sigma_{t+1,1}) \\ \log(\sigma_{t+1,2}) \\ \text{logit}(\rho_{t+1}) \end{pmatrix} = \begin{pmatrix} \mu_{\sigma_1} \\ \mu_{\sigma_2} \\ \mu_\rho \end{pmatrix} + \mathbf{\Psi} \left(\begin{pmatrix} \log(\sigma_{t,1}) \\ \log(\sigma_{t,2}) \\ \text{logit}(\rho_t) \end{pmatrix} - \begin{pmatrix} \mu_{\sigma_1} \\ \mu_{\sigma_2} \\ \mu_\rho \end{pmatrix} \right) + \boldsymbol{\nu}, \qquad \boldsymbol{\nu} \sim N_3(\mathbf{0}, \mathbf{\Sigma}_\nu)$$

Many different models for the evolution and correlation of both returns and volatilities can be proposed by specifying the entries in $\mathbf{\Psi}$ and $\mathbf{\Sigma}_\nu$. For example, similarly to Bollerslev (1990), we

can impose independent evolution of $\sigma_{1,t}$ and $\sigma_{2,t}$ with no time-dependence of ρ_t by setting $\boldsymbol{\Psi} = \text{diag}(\psi_{11}, \psi_{22}, 1)$ and $\boldsymbol{\Sigma}_\nu = \text{diag}(\sigma^2_{1\nu}, \sigma^2_{2\nu}, 0)$. Alternatively, we can introduce correlation in volatility and a constant correlation in price (Yu and Meyer, 2006) by setting

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{11} & 0 & 0 \\ \psi_{21} & \psi_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_\nu = \text{diag}(\sigma^2_{1\nu}, \sigma^2_{2\nu}, 0).$$

In our formulation, we will use a fully specified matrix $\boldsymbol{\Psi}$, with independent innovations driving the latent process so that $\boldsymbol{\Sigma}_\nu = \text{diag}(\sigma^2_{1\nu}, \sigma^2_{2\nu}, \sigma^2_{\rho\nu})$. We can also implement this specification by imposing the autoregressive process on the three entries in the lower-triangular component of Cholesky factorization of $\boldsymbol{\Sigma}_{t+1}$,

$$\begin{pmatrix} \log(V_{t+1,11}) \\ \log(V_{t+1,22}) \\ V_{t+1,21} \end{pmatrix} = \begin{pmatrix} \mu_{V_{11}} \\ \mu_{V_{22}} \\ \mu_{V_{21}} \end{pmatrix} + \Psi \left(\begin{pmatrix} \log(V_{t,11}) \\ \log(V_{t,22}) \\ V_{t,21} \end{pmatrix} - \begin{pmatrix} \mu_{V_{11}} \\ \mu_{V_{22}} \\ \mu_{V_{21}} \end{pmatrix} \right) + \boldsymbol{\nu}, \qquad \boldsymbol{\nu} \sim N_3(\mathbf{0}, \boldsymbol{\Sigma}_\nu) \tag{4.4}$$

where

$$\mathbf{V}_t = \begin{pmatrix} V_{t,11} & 0 \\ V_{t,21} & V_{t,22} \end{pmatrix}, \qquad \boldsymbol{\Sigma}_t = \mathbf{V}_t' \mathbf{V}_t.$$

The model formulated by combining (4.2) and (4.4) will be fitted using Markov Chain Monte Carlo algorithms, where samples of the volatility and correlation processes will be obtained using a forward-filtering backward-sampling scheme. Our contribution will be the inclusion of results in Chapter 3 and the future work outlined in Section 4.2 to allow us to improve our estimates by including maximum and minimum intraperiod prices within the full likelihood for the model.

4.4 Timeline for Future Work

The timeline for the future projects and writing is included in the table below.

	Fall 2014	Winter 2015	Spring 2015	Summer 2015	Fall 2015	Winter 2016
Stochastic Volatility with microstructure noise:						
Univariate Likelihood with microstructure noise:						
Bivariate Likelihood with Bernstein polynomials:						
Bivariate Stochastic Volatility model:						
Thesis writing:						

Bibliography

Sassan Alizadeh, Michael W Brandt, and Francis X Diebold. Range-based estimation of stochastic volatility models. *The Journal of Finance*, 57(3):1047–1091, 2002.

Torben G Andersen and Tim Bollerslev. Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The journal of Finance*, 52(3):975–1005, 1997a.

Torben G Andersen and Tim Bollerslev. Intraday periodicity and volatility persistence in financial markets. *Journal of empirical finance*, 4(2):115–158, 1997b.

Torben G Andersen and Tim Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, pages 885–905, 1998.

Torben G Andersen, Tim Bollerslev, Francis X Diebold, and Paul Labys. The distribution of realized exchange rate volatility. *Journal of the American statistical association*, 96(453):42–55, 2001.

Richard T Baillie and Robert J Myers. Bivariate garch estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics*, 6(2):109–124, 1991.

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

Ole E Barndorff-Nielsen and Neil Shephard. Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17(5):457–477, 2002.

Ole E Barndorff-Nielsen and Neil Shephard. Power and bipower variation with stochastic volatility and jumps. *Journal of financial econometrics*, 2(1):1–37, 2004.

David S Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of financial studies*, 9(1):69–107, 1996.

Luc Bauwens and Michel Lubrano. Bayesian inference on garch models using the gibbs sampler. *The Econometrics Journal*, 1(1):23–46, 1998.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31 (3):307–327, 1986.

Tim Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, pages 498–505, 1990.

Tim Bollerslev and Benjamin YB Zhang. Measuring and modeling systematic risk in factor pricing models using high-frequency data. *Journal of Empirical Finance*, 10(5):533–558, 2003.

Tim Bollerslev and Hao Zhou. Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, 109(1):33–65, 2002.

F Jay Breidt, Nuno Crato, and Pedro De Lima. The detection and estimation of long memory in stochastic volatility. *Journal of econometrics*, 83(1):325–348, 1998.

Jonathan Brogaard. High frequency trading and its impact on market quality. *Northwestern University Kellogg School of Management Working Paper*, page 66, 2010.

Peter Carr and Liuren Wu. Time-changed lévy processes and option pricing. *Journal of Financial Economics*, 71(1):113–141, 2004.

Peter K Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica: journal of the Econometric Society*, pages 135–155, 1973.

Fabienne Comte and Eric Renault. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8(4):291–323, 1998.

Fabienne Comte, Laure Coutin, and Éric Renault. Affine fractional stochastic volatility models. *Annals of Finance*, 8(2-3):337–378, 2012.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

Eugene F Fama. The behavior of stock-market prices. *Journal of business*, pages 34–105, 1965.

David Freedman. *Brownian motion and diffusion*. Springer, 1971.

Hélyette Geman. From measure changes to time changes in asset pricing. *Journal of Banking & Finance*, 29(11):2701–2722, 2005.

Peter R Hansen and Asger Lunde. Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2):127–161, 2006.

Andrew C Harvey. Long memory in stochastic volatility. *Forecasting volatility in the financial markets*, page 307, 2002.

Enrique Ter Horst, Abel Rodriguez, Henryk Gzyl, and German Molina. Stochastic volatility models including open, close, high and low prices. *Quantitative Finance*, 12(2):199–212, 2012.

John Hull and Alan White. The pricing of options on assets with stochastic volatilities. *The journal of finance*, 42(2):281–300, 1987.

Eric Jacquier, Nicholas G Polson, and Peter E Rossi. Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, 20(1):69–87, 2002.

Sangjoon Kim, Neil Shephard, and Siddhartha Chib. Stochastic volatility: likelihood inference and comparison with arch models. *The Review of Economic Studies*, 65(3):361–393, 1998.

Peter Lildholdt. *Estimation of GARCH models based on open, close, high, and low prices*. CAF, Centre for Analytical Finance, 2002.

Benoit Mandelbrot. The variation of some other speculative prices. *Journal of Business*, pages 393–413, 1967.

Angelo Melino and Stuart M Turnbull. Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, 45(1):239–265, 1990.

Teruo Nakatsuma. Bayesian analysis of arma–garch models: A markov chain sampling approach. *Journal of Econometrics*, 95(1):57–69, 2000.

Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370, 1991.

Yasuhiro Omori, Siddhartha Chib, Neil Shephard, and Jouchi Nakajima. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2):425–449, 2007.

José M Quintana and Mike West. An analysis of international exchange rates using multivariate dlm's. *The Statistician*, pages 275–281, 1987.

Esther Ruiz. Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of econometrics*, 63(1):289–306, 1994.

Gleb Sandmann and Siem Jan Koopman. Estimation of stochastic volatility models via monte carlo maximum likelihood. *Journal of Econometrics*, 87(2):271–301, 1998.

Neil Shephard. *Stochastic Volatility: Selected Readings*. Oxford University Press, UK, 2005.

Shinichiro Shirota, Takayuki Hizu, and Yasuhiro Omori. Realized stochastic volatility with leverage and long memory. *Computational Statistics & Data Analysis*, 76:618–641, 2014.

Hans R Stoll. Presidential address: friction. *The Journal of Finance*, 55(4):1479–1514, 2000.

Viktor Todorov. Econometric analysis of jump-driven stochastic volatility models. *Journal of Econometrics*, 160(1):12–21, 2011.

Harald Uhlig. Bayesian vector autoregressions with stochastic volatility. *Econometrica: Journal of the Econometric Society*, pages 59–73, 1997.

JH Venter and PJ de Jongh. Extended stochastic volatility models incorporating realised measures. *Computational Statistics & Data Analysis*, 2012.

Ioannis D Vrontos, Petros Dellaportas, and Dimitris N Politis. Full bayesian inference for garch and egarch models. *Journal of Business & Economic Statistics*, 18(2):187–198, 2000.

RM Wilcox. Exponential operators and parameter differentiation in quantum physics. *Journal of Mathematical Physics*, 8(4):962–982, 1967.

Jun Yu. On leverage in a stochastic volatility model. *Journal of Econometrics*, 127(2):165–178, 2005.

Jun Yu and Renate Meyer. Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews*, 25(2-3):361–384, 2006.

Lan Zhang, Per A Mykland, and Yacine Aït-Sahalia. A tale of two time scales. *Journal of the American Statistical Association*, 100(472), 2005.