Volatility: Estimating Quadratic Variation using Realized Variance

Master's Thesis in Mathematical Statistics

Anton Körkkö

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg Sweden, 2013

Abstract

We present results, using modeling, in estimating volatility in the three different models: Geometrical Brownian motion (GBM), Cox-Ingerssoll-Ross (CIR) and Heston. We compare estimations, using the quadratic variation estimator realized variance, with estimations using maximum likelihood methods. In the Heston model we use our approximation of the volatility process to estimate the volatility of volatility parameter, by again applying realized variance.

In the GBM model we show that realized variance is useful even when data is low-frequent. The effects on bias when using high-frequency discrete data is studied, and methods of removing bias is discussed. In the CIR-model we use realized variance in two different ways. One of the two estimators is biased, but we successfully remove it. In the Heston model problem arises when using realized variance, and estimation precision decreases, but under certain conditions the long run volatility-parameter, the volatility processes and the volatility of volatility-parameter can be approximated fairly well.

Contents

Acknowledgements

The author whish to thank all those people that contributed to this project. In particular I would like to thank Emma Wijkmark for important discussions and helful comments. Further I like to thank my supervisor Prof. Patrik Albin for letting me do this project, and his guidence and support during the work.

1 Introduction

In finance, mathematical models are widely used. By creating models we can make conclusion of the current and past financial situation. Ideally, these models can help us make predictions of the future.

When modeling financial assets, stochastic differential equations of diffusion type is widely used. The most famous model by Black and Scholes, which describe the price movements as a Geometrical Brownian motion, from 1973 [4] is such a model and since then several other models of diffusion type has been developed. Jointly, these models take the volatility of asset prices into account. When applying these models, parameters must be estimated and a method must be selected. A possible method when estimating volatility parameters is the use of realized variance, that approximates quadratic variation.

There are several methods for estimating volatility in a model. Maximum likelihood methods have advantages and are useful under some conditions. A problem with this method is that the optimization problem that has to be solved can have a high computational burden and if one needs up to date estimates this might be a problem. In the Heston model, maximum likelihood estimation is a non-trivial optimization problem, with high computational cost, which has to be solved by carefully selected numerical methods. This is done by Mariani, Pacceli and Zirilli (2008, [14]), using both stock and option data.

Using realized variance as a estimator for quadratic variation might be a possible method to get up to date estimations that are accurate, this method will be our main focus. There is good access to earlier work concerning the use of quadratic variation, see Barndoff-Nielsen and Shepard (2002, [3]) and Hautch (2012, [11]), among others.

In theory, under some assumptions discussed later on, the realized variance is a consistent estimator for quadratic variation as the number of observations in an interval tend to infininty. We ask the question of how frequent the data have to be for realized variance to be a god estimator, and also investigate the impact of discreetness when dealing with high frequency data. We show in this thesis that one has to be careful, estimations might, as we show, be highly biased under this condition.

Data is in general affected by noise due to several different microeconomic effects. One of the effects is the discreteness, other effects, such as the bid-ask bounce effect, is important and discussed by Greaham, Michealy and Roberts(2003, [10]) and Rhee and Wang (1997, [16]). Several authors investigates the effects of microeconomic noise, see for example Aït-Sahalia and Macnini(2008, [1]) and Zhou (1996, [19]).

We will consider the popular and simple Black-Scholes model describing the asset price as a geometrical Brownian motion (GBM). Further we study the CIR-model, that takes the mean reverting effect in consideration, and finally we apply realized variance on the Heston model by Steven Heston (1993, [12]), which is a mix of the Black-Scholes and the CIR model.

The volatility and the volatility parameters can be estimated with various precision for the different models, therefore this must be a concern when choosing model, and knowledge of how exact these estimations can be done is of interest.

The intention we have is to evaluate the advantages and disadvantages of using realized variance as a method under different assumptions. The approach taken to make conclusion is the use of modeling done in R and comparing parameter estimations using realized variance comapared with maximum or quasi-maximum likelihood methods. We also investigate the possibilities to remove bias, when it is neccecary, for our estimations.

We conclude that the use of realized variance as a method is overall a good method for volatility estimation. We begin this thesis with introducing basic theory of diffusion models, quadratic variation and an answer the question when realized variance is an consistent estimator of quadratic variation. We then, in turn, introduce and a and present results from the modeling for the GBM model, the CIR model, and the Heston model.

2 Theory

There are different definitions of volatility[1], therefore it is convenient for us to define it. Let S_t be the price att time t and $s_t = \log(S_t)$.

Definition 2.1. Quadratic variation
We define quadratic variation (QV_T) of a stochastic process h_t over [0,T]

$$QV_T(h) = \lim_{\|M\|\to 0} \sum_{k=1}^{n} (h_{t_k} - h_{t_{k-1}})^2 \tag{1}$$

With M=$\max_{1\leq k\leq n}(t_k - t_{k-1})$
where $0 = t_0 \leq t_1 \leq t_2 \leq ... \leq t_n = T$

Definition 2.2. Volatility
Volatility υ_t of a process S_t over time [0, T] is defined from

$$\upsilon_t = \sqrt{(QV_T(s))}$$

Definition 2.3. Realized variance
The realized variance (RV) over [0,T] of h_t is given by

$$RV_T(h) = \sum_{k=1}^{n} (h_{tk} - h_{t_{k-1}})^2$$

Definition 2.4. Realized volatility
The realized volatility (Rv_T) of S_t over [0,T] is given by

$$Rv_T = \sqrt{(RV_T(s))}$$

The quadratic variation can under certain conditions for small values of $\max \|(t_i - t_{i-1})\|$ be approximated by the realized variance, that is the sum of the quadratic returns.

[1] volatility is sometimes defined using data S_t instead of logarithmic data s_t. Noticeable is though that if $S_{t+1} = S_t(1+\delta)$ with δ small, then by Taylor expansion $log(S_{t+1}) - log(S_t) \approx \log(S_t) + \delta - \log(S_t) = \delta = S_{t+1} - S_t$ This results in, that when we have moderate volatilities and high frequency data, the realized variance will be approximately equal, no matter if we use data or logarithmic data.

Definition 2.5. Standard Brownian motion
A process $B = B_t, t \in [0, \infty)$ *is a standard Brownian motion if*

1. *The paths are continuous,*

2. *It has stationary, independent increments and*

3. $B_t \sim N(0, t), \forall t > 0$

Values of financial assets are often described by the continuous-time stochastic differential equation (2).

$$dX_t = \mu(X_t)dt + \sigma(X_t)X_t dB_t, 0 \leq t \leq T \tag{2}$$

where $\mu(X_t)$ is the drift coefficient, $\sigma(X_t)$ is the diffusion (volatility) coefficient. $\sigma(X_t)$ is the spot volatility and is the mean volatility over an infitesimal time interval. dB_t is a Brownian motion and $\sigma(X_t)$ can be dependent of unknown variables, and we let the drift coefficient be linear, that is of the form $\alpha + \gamma X(t)$.

Theorem 2.1. Miao(2004,[15])*: Given* $dX_t = \mu(X_t)dt + \sigma g(X_t)dB_t$ *and* X_t *has a strong solution, and* τ *is a partition of [0,T] so that* $0 = t_0 \leq t_1 \leq ... \leq t_n = T$ *then*

$$\frac{\sum_\tau |X_{t_{j+1}} - X_{t_j}|^2}{\sum_\tau g^2(X_{t_j})(t_{j+1} - t_j)} \xrightarrow{P} \sigma^2 \tag{3}$$

where $\max_{1 \leq j \leq n}(t_j - t_{j-1}) \to 0$

2.1 About models

In financial models we have variables, often time dependent or dependent of for example the value of the financial assets studied. Volatility can for example be dependent of the value of the underlying assets. Depending on what we are studying, different models are suitable to apply. In the models we also have parameters, these are constant and unique for the underlying asset that is studied. These parameters can be estimated in different ways. One way of estimating spot volatility or a volatility parameter is the use of realized variance. This method is suitable for many models, both for estimating spot volatility and the value of the underlying volatility parameter. One thing we have to remember when applying a model to an assets is that the parameter values can change value over time. This is due to that the parameter is dependent of the unique characteristic and structures of the underlying assets. If the Characteristics or structures that give rise to the parameter values are changed one has to examine if the parameter values are changed. Some structural changes might even make the model applied invalid and other models are then needed to be considered.

2.2 Micro-economic effects

There are several microeconomic effects, for example the so called bid-ask bounce effect, the effect of differences of trade sizes and the effect of discrete data. These microeconomic effects have impact when analyzing high frequency data. This means that if X_t is the value of the underlying process and $\hat{X}_t$ is the observed value then $X_t = \hat{X}_t + \epsilon_t$. Here ϵ_t is the so called error term. This error term can become dominant for short time intervals, making the error big. If we for example assume that ϵ_t is normally distributed with variance σ^2 and mean 0 then

$$X_{t_j} - X_{t_k} = \hat{X}_{t_j} + \epsilon_{t_j} - \hat{X}_{t_k} - \epsilon_{t_k} = \hat{X}_{t_j} - \hat{X}_{t_k} + \epsilon$$

where ϵ is normally distributed with variance $2\sigma^2$ and

$$\mathbf{E}\{\sum_{k=1}^{n}(X_{t_k} - X_{t_{k-1}} - \epsilon_k - \epsilon_{k-1})^2\} = \mathbf{E}\{\sum_{k=1}^{n}(X_{t_k} - X_{t_{k-1}})^2\} + 2n\sigma^2 = RV_T(X) + 2n\sigma^2.$$

Here the total error $2n\sigma^2$ is time independent but $RV_T(X)$ decreases as T decreases, making the error dominant for small values of T and big values of n.

In our simulations we don't make any assumptions of the distribution of ϵ_t, but as we'll see the discreteness assumption of data makes the error dominant for high frequency data.

2.2.1 Discrete data

When analyzing data from the stock exchanges both in Sweden and the USA prices are commonly given, as we find, in cents per dollar in the USA and in

Sweden 1/10 of a Swedish krona, the prices lies in the interval 0.5 to 500, the daily volatility in the interval 0.5-10% and it is not common to have sampling frequency that is higher than 1 data point per second. A typical stock have a daily volatility of 0.5-5% in average and the data frequency from a stock is often not found higher than 1 trade/ 10 seconds. This is with some exceptions, like Apple Inc that is often one of the most traded stocks. Of course the frequency of data found is bounded by the trading frequency. For some stocks, like Google Inc., that, at the time of writing, has a price around 900 dollars or Apple Inc., which has a price around 400 dollars, it is not possible to find any effects of the discreetness of prices when using realized variance for the models studied. This is due to that a cent is only $\frac{1}{900}$% of the stock value, but for a stock of value of about one dollar the effect of discreteness can be significant, even with high volatilities and quite low frequent data.

These conclusions is found from simulations, and the applicability of realized variance as method of estimating volatility is particulary studied in the Geometrical Brownian motion model.

3 Geometrical Brownian motion (GBM)

In a simple model we assume that the value of our financial asset is the value of a simple stochastic differential equation. Where the change in value is dependent on a mean interest rate and a random term. We state the equation:

Definition 3.1. A process is said to be a Geometrical Brownian motion (GBM) if it satisfies

$$dS_t = \mu S_t dt + \sigma S_t dB_t \tag{4}$$

This is the Black-Scholes model, see: Black and Scholes(1973, [4]). μ is the interest rate, σ the volatility and B_t is a Brownian motion. The solution to the equation is $S_t = S_0 \exp(\hat{\mu} t + \sigma B_t)$ where $\hat{\mu} = \mu - \sigma^2/2$. We recognize $s_t = log(S_t)$ as a Brownian motion with standard deviation σ. Recalling our definition of volatility (2.4) gives that the volatility is the square-rot of the quadratic variation of the Brownian motion. The quadratic variation of a Brownian motion over an interval $[0, T]$ is is $\sigma^2 T$ and the volatility parameter is constant and equal to the constant standard deviation σ.

3.1 Volatility estimation

When estimating the volatility we use two different estimators. Firstly an estimator where we use realized variance, and secondly we bias correct the maximum likelihood estimator sample variance, and use that.

3.1.1 Using realized variance as estimator

As mentioned we can take the logarithm of data to get a brownian motion from the GBM. The quadratic variation of the Brownian motion over the interval [0,T] is $\sigma^2 T$. The use of realized variance gives us

$$\hat{\sigma^2} = \frac{1}{T} \sum_{i=1}^{N} (x_i - x_{i-1})^2 \tag{5}$$

If we instead use formula 2.1 we arive at the same estimator. Formula 2.1 could also be used on the original data, without taking the logarithms but that is not done here.

3.1.2 Realized variance and sample variance

Taking the sample variance of the data set $x_2 - x_1, x3 - x_2, ..., x_N - x_{N-1}$ (the ml-estimator) gives us

$$\begin{aligned}\widetilde{\sigma^2} = \frac{1}{T(N-1)}(\sum_{i=2}^{N}(x_i - x_{i-1})^2) - \frac{N}{T(N-1)}(\frac{1}{N}\sum_{i=2}^{N}(x_i - x_{i-1}))^2 \\ = \frac{1}{T(N-1)}\sum_{i=1}^{N}(x_i - x_{i-1})^2 - \frac{(x_N - x_1)^2}{TN(N-1)} \rightarrow \hat{\sigma^2}\end{aligned} \tag{6}$$

as $N \rightarrow \infty$ ($dt \rightarrow 0$). We've noticed that, for $(\mu - \frac{\sigma^2}{2})\frac{T}{N}$ small and N large

$$\mathbf{E}\{\frac{(x_N - x_1)^2}{TN(N-1)}\} \approx \frac{\sigma^2}{N(N-1)} \approx 0 \tag{7}$$

where dt=T/N, and keeping N large

$$\frac{1}{N-1}\sum_{i=1}^{N}(x_i - x_{i-1})^2 \approx \frac{1}{N}\sum_{i=1}^{N}(x_i - x_{i-1})^2 \tag{8}$$

The maximum-likelihood estimator is biased by (N-1)/(N-2), we scale the estimator to be unbised but still call it the ml-estimator. This is due to that in most cases we use very big values of N, giving that the true maximum likelihood estimator is very close to being unbiased, and the scaling is not needed in those cases.

In this section we first examine how good the estimators are for low frequent data and then we examine the influence on estimations when frequency is high and data is given with low precision.

3.2 Low frequent data

We show that low frequent data such as one data point per day gives unbiased estimations ($|\text{bias}| < 1\%$) of σ, given volatilities and interest rates that are reasonable for stocks. For low volatilities such as 1% and interest rates that are less than 25% per year gives bias less than 1%. For higher interest rates, higher volatilities are necessarily to get a small bias. One should be clear that a long term interest rate of 25% per year might unreasonable for such a safe investment as a stock with volatility of 1% is.

3.2.1 Results for low frequent data

Comparing the maximum likelihood estimator with realized variance shows that the maximum likelihood estimator (mle) is unbiased but realized variance only gives approximately unbiased estimates. Further we see that the higher the frequency of data, the lower the bias. The variance of the estimates are equal. Due to the fact that the weighted mle we use is unbiased and that the two estimators have estimations that are almost equal from a variance point of view, we only present results using realized variance.

In table 3.1 we have drift coefficients μ of 0.05%, 0.1% and 1% per day, that is about 13%, 28% and 1200% respectively per year. So the two last rates are quite high but for rates such as 13% or less per year the bias using realized variance is small, even though we only have one data point per day and low volatilities. Bias gets bigger when data becomes more low frequent, volatilities are low and interest rates are high. The variance of the estimations are approximately the same no matter if one uses maximum likelihood or realized variance. Though, by using realized variance we get biased estimations.

1 Data point per day				
σ	μ	$\hat{\sigma}$	sd($\overline{\hat{\sigma}}$)	bias/σ
0.005	0.05%	0.00502	0.023%	0.513%
0.005	0.10%	0.00510	0.023%	1.92%
0.005	1%	0.0112	0.030%	123%
0.01	0.10%	0.0100	0.022%	0.439%
0.01	1%	0.0141	0.027%	41.0%
0.03	1%	0.0315	0.023%	4.92%
0.05	1%	0.0501	0.028%	0.51%
1 Data point per hour				
σ	μ	$\hat{\sigma}$	sd($\overline{\hat{\sigma}}$)	bias/σ
0.005	0.10%	0.00501	0.027%	0.32%
0.005	1%	0.00634	0.032%	27.0%
0.01	0.10%	0.0100	0.028%	0.028%
0.01	1%	0.0107	0.029%	7.30%
0.03	1%	0.0302	0.028%	0.78%
0.05	1%	0.0501	0.028%	0.24%
4 Data points per hour				
σ	μ	$\hat{\sigma}$	sd($\overline{\hat{\sigma}}$)	bias/σ
0.005	0.10%	0.00500	0.028%	0.05%
0.005	1%	0.00537	0.030%	7.38%
0.01	1%	0.0102	0.029%	1.87%
0.03	1%	0.0301	0.028%	0.20%

Table 3.1: Using RV for low frequent data.

3.3 High frequency data

To minimize the the variance of the volatility estimations we want as many data points as possible for the time interval studied. This means that we need as high frequent data as possible. We will show that one has to be careful when using descrete, high frequency data as this leads to biased estimations.

When doing the analysis data of a volatility of 1-10% per day and interest rate of 0-20% per year is used. These number are selected to be as realistic as possible. When studying high frequency we consider data in a frequency interval of 1 data point per second to one data point each 15 minute. We will use 30 data points. This to compare variances of the estimations. Further we assume that stocks are not traded in higher precision than cents per dollar.

3.3.1 Results for high frequency data

The result part is split in to different parts, we have to take frequency, interest rate, volatility and discreteness into account. In table 3.2 and table 3.3 we show results using realized variance, with $\mu = 0.05\%$ per day.

In table 3.4 we compare results from using maximum likelihood and realized variance for high frequency, discrete data, with initial value x_0. We notice that the estimators are approximately equally good both in term of variance and bias. This is true for all moderate values of dt.

3.3.2 Bias-correction

By analyzing the dependence of the bias of volatility estimations due to discrete data, the bias might be possible to remove.

We find, by analyzing data (see figure (2)-(4)) , that if $C = \sigma^2 {x_0}^2/N$ is kept constant, with N being the number of observation during on day (6.5 hours), then there seems to be a bias that is unique and constant paired with each value of C. This relationship have been studied for many values of C

One day, 1 point per 1s					One day, 1 data point per 5s.				
x_0	σ	$\hat{\sigma}$	bias/σ	sd($\overline{\hat{\sigma}}$)	x_0	σ	$\hat{\sigma}$	bias/σ	sd($\overline{\hat{\sigma}}$)
1	0.01	0.105	952%	3.58%	1	0.01	0.0708	608%	0.46%
1	0.04	0.221	452%	1.09%	1	0.04	0.147	268%	0.11%
1	0.20	0.500	149%	0.67%	1	0.20	0.333	66.6%	0.09%
5	0.01	0.0494	394%	3.26%	5	0.01	0.0330	230%	0.08%
5	0.04	0.0999	147%	0.21%	5	0.04	0.0661	65.4%	0.03%
5	0.20	0.238	18.8%	0.39%	5	0.20	0.2086	4.05%	0.02%
30	0.01	0.0202	102%	0.11%	30	0.01	0.0136	36.2%	0.016%
30	0.04	0.0451	12.7%	0.10%	30	0.04	0.0411	2.66%	0.015%
30	0.2	0.201	0.57 %	0.06%	30	0.1	0.100	0.43%	1.05%
100	0.01	0.0118	17.9%	0.49%	100	0.005	0.00573	14.5%	0.015%
100	0.04	0.0405	1.21 %	0.05%	100	0.04	0.0401	0.243%	0.016%
300	0.005	0.0054	8.31%	0.05%	300	0.005	0.0051	1.74%	0.016%

Table 3.2: Bias and variance of estimated mean volatility.

One day, 1 data point per min					One day, 2 data point per min				
x_0	σ	$\hat{\sigma}$	bias/σ	sd($\overline{\hat{\sigma}}$)	x_0	σ	$\hat{\sigma}$	bias/σ	sd($\overline{\hat{\sigma}}$)
1	0.01	0.0379	279%	1.61%	1	0.03	0.0815	172%	0.088%
1	0.04	0.0793	98.2%	0.37%	1	0.10	0.150	49.9%	0.037%
1	0.20	0.217	8.35%	0.20%	1	1.00	1.035	3.64%	0.067%
5	0.01	0.0177	77.4%	0.30%	5	0.02	0.0299	49.6%	0.029 %
5	0.04	0.0432	7.88%	0.18%	5	0.04	0.0460	15.1%	0.024%
5	0.10	0.101	1.35%	0.18%	5	0.10	0.103	2.67%	0.023%
30	0.005	0.00568	13.7%	0.18%	5	0.30	0.301	0.362%	0.023%
30	0.005	0.00568	13.7%	0.18%	30	0.01	0.0107	7.01%	0.023%
30	0.03	0.0301	0.461%	0.18%	30	0.03	0.0303	0.84%	0.023%
100	0.01	0.0100	0.398%	0.18%	100	0.005	0.0051	2.60%	0.023%

Table 3.3: Bias and variance of estimated mean volatility.

One day, 1 data point per 5 min							
x_0	σ	$\hat{\sigma}_{\mathrm{RV}}$	$\mathrm{bias}_{\mathrm{RV}}/\sigma$	$\mathrm{sd}(\hat{\sigma}_{\mathrm{RV}})$	$\hat{\sigma}_{\mathrm{mle}}$	$\mathrm{bias}_{\mathrm{mle}}/\sigma$	$\mathrm{sd}(\hat{\sigma}_{\mathrm{mle}})$
1	0.01	0.0253	153%	80.0%	0.0252	152%	80.3%
1	0.03	0.0457	52.4%	15.5%	0.0453	52.3%	15.7%
1	0.10	0.106	6.17%	8.77%	0.105	6.17%	8.70%
1	0.30	0.302	0.50%	8.09%	0.297	0.48%	8.03%
5	0.01	0.0122	22.7%	9.87%	0.0121	22.7%	9.83%
5	0.03	0.307	2.55%	8.31%	0.0304	2.54%	8.35%
5	0.1	0.0998	-0.19%	8.01%	0.0985	-0.20%	7.97%
30	0.005	0.00513	2.51%	8.25%	0.00506	2.48%	8.20%
30	0.02	0.0200	-0.16%	7.98%	0.0200	-0.16%	7.94%
100	0.005	0.0499	-0.13%	8.06%	0.0499	-0.13%	8.00%

Table 3.4: Comparing realized variance with mle for discrete hf-data.

and seems accurate.

There is another standard approach to avoid bias, which is applied in section 5.4.2, when estimating the volatility of the volatility in the Heston model. For the ease, assume that we have a high frequency time serie T with n·m data points $X_1, X_2, ..., X_{n \cdot m}$. Then instead of taking realized variance in the previous done way, we split this time serie into n time series with m data points in each time serie. We denote time serie i by $\hat{T}_i$.

The time serie is split up such that $\hat{T}_i = \{X_i, X_{i+n}, ..., X_{i+n(m-1)}\}$, $i = 1, 2, ..., n$. By estimating volatility from each one of these time series and then take the mean value as estimator of the volatility we decrease bias, but the variance of the estimation is still low as we still use all data we have.

In figure 1 we see a log-log plot of the bias. For $log(x) > 1$ there seems to be a linear relationship in the plot, indicating that the bias over the interval might have the form $bias = aX^b$.

In figure 2-4 we have that $\frac{N}{\sigma^2 x^2} \approx 156000$, and N,$x_0$,$\sigma$ is varied. As we can see the bias is close to 77-78% as long as $\frac{N}{\sigma^2 x^2}$ is kept constant. Algorithm 1 is a possible algorithm for bias correction.

Algorithm 1. *Finding* σ

0. Choose $\tilde{\sigma}$, *for example* $\tilde{\sigma} = \hat{\sigma}$

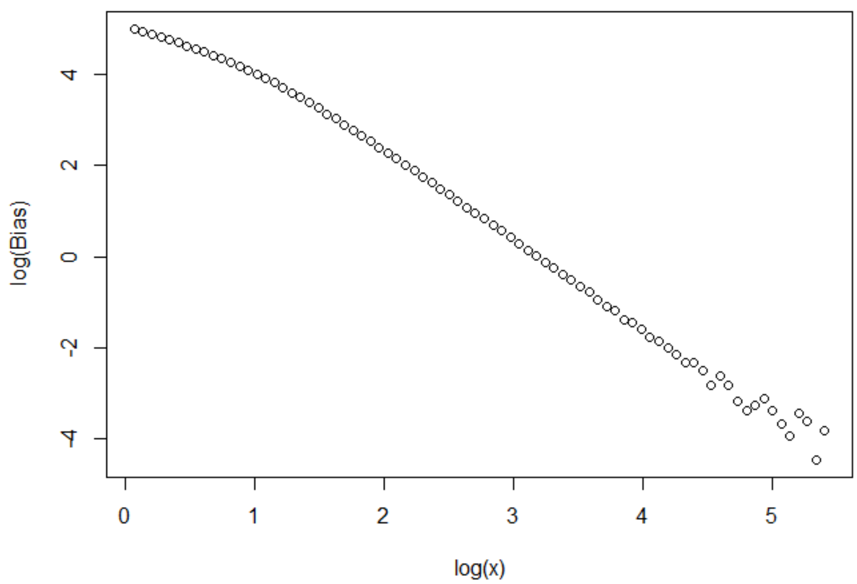


Figure 1: log-log-plot Bias(%) vs. x.

Bias correction		
σ_{biased}	σ_{corr}	bias
0.03032	0.0250	21.3%
0.0251	0.0220	14.1%
0.0236	0.0215	9.8%%
0.0229	0.0212	8.0%
0.0225	0.0211	6.6%
0.0224	0.0212	5.7%
0.0224	0.0215	4.2%
0.0225	0.0217	3.7%

Table 3.5: Bias-corrected volatility.

1. *let* $\tilde{\sigma} = \hat{\sigma}/(1+b)$ *and go to step 1 or stop.* $\tilde{\sigma}$ *is our new estimator of* σ*.*

In figure 5 we see estimations of volatility using the first method. In accordance with the figure and the tables there is an error due to the discreteness for small dt. In figure 6b we see an example of volatility estimation before and after bias removal. In this example the removal seems to be overcompensated.

In table (3.5) we have bias corrected volatility estimations for the eight first points in figure 5. The corrected values seems reasonable when comparing the other volatility estimations in figure 5, which have a minimum estimation of 0.0214 and a mean of 0.0225. It is also worth noticing that the first point probably isn't corrected enough. This can be due to that the corrector isn't correcting for the discreteness enough. But it can also be a result of data being affected by other microeconomic effects aswell, these are not corrected for. Further the corrector is based on data from GBM, an assumption that isn't completely true in reality. But we do notice that the interval length for which the estimations of σ lies within have decreased by about 50%.

In figure 7 we see bias in % as a function of $\frac{N}{\sigma^2 x^2}$

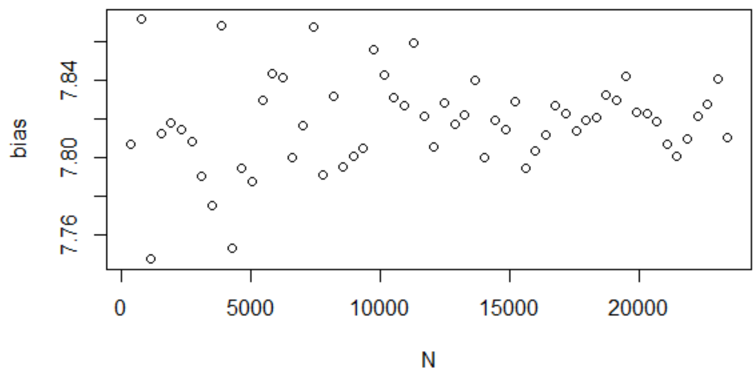


Figure 2: N/x^2=6240, σ=0.02

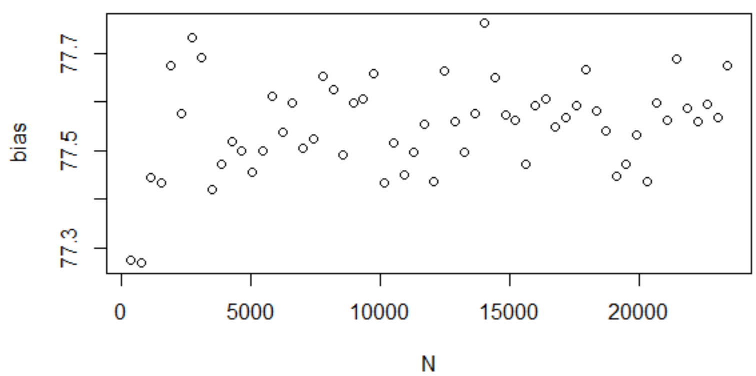


Figure 3: N/σ^2 = 15600000 , x=10

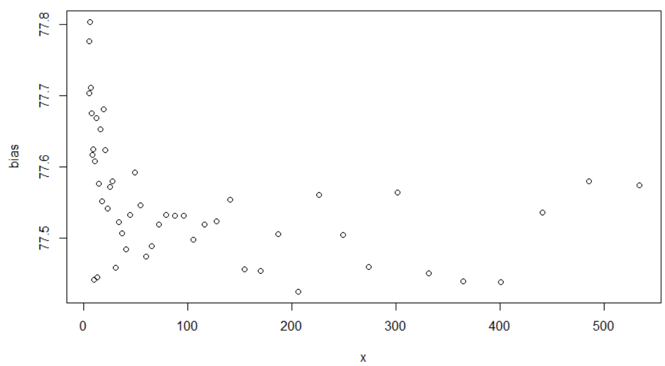


Figure 4: $x^2\sigma^2 = 0.09$, N=14040

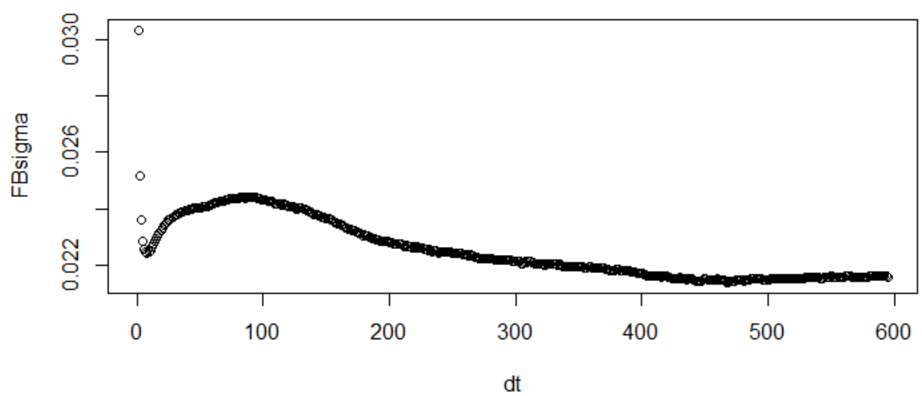
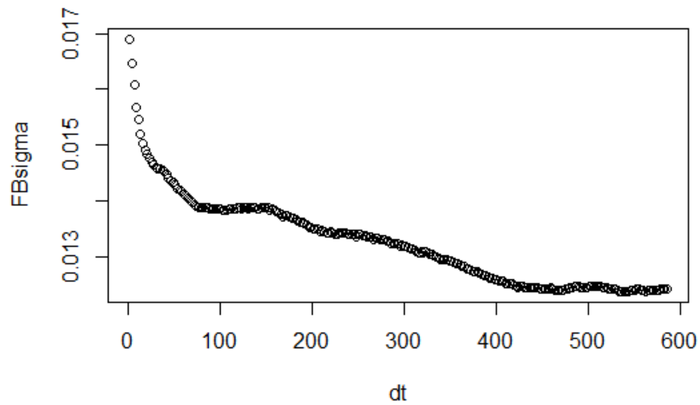
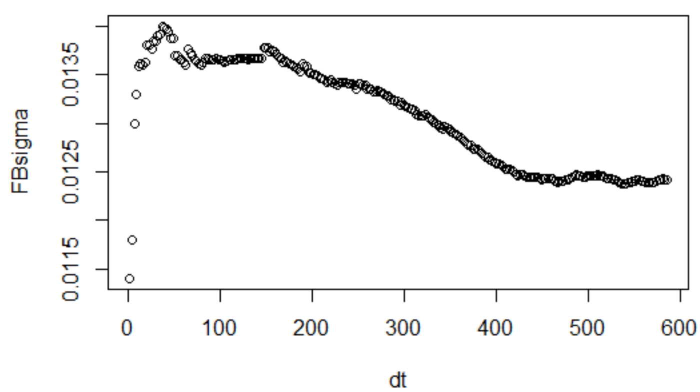


Figure 5: Estimated volatility Facebook 2013-07-26, using different values of dt.



(a) Facebook 2013-07-24.



(b) Facebook 2013-07-24 bias-corrected

Figure 6: Volatility before and after bias-removal.

bias vs. N/($\sigma^2 {x_0}^2$)							
N/($\sigma^2 {x_0}^2$)	bias	N/($\sigma^2 {x_0}^2$)	bias	N/($\sigma^2 x_0^2$)	bias	N/($\sigma^2 {x_0}^2$)	bias
1244	1.86%	2489	2.54%	3733	3.41%	4978	4.28%
6222	5.21%	7467	6.31%	8711	7.21%	9956	8.27%
11200	9.04%	12444	10.0%	13689	10.95%	14933	11.94%
16178	12.85%	17422	13.74%	18667	14.66%	19911	15.55%
21156	16,42%	22400	17.30%	23644	18.23%	24889	19.05%
26133	19.93%	27378	20.81%	28622	21.61%	29867	22.48%
31111	23.34%	32356	24.12%	33600	24.94%	34844	25.77%
36089	26.55%	37333	27.35%	38578	28.17%	39822	28.97%
41067	29.73%	42311	30.58%	43556	31.25%	44800	32.05%
46044	32.80%	47289	33.58%	48533	34.30%	49778	34.93%
51022	35.69%	52267	36.39%	53511	37.13%	54756	37.81%
56000	38.55%	57244	39.23%	58489	39.88%	59733	40.51%
60978	41.17%	62222	41.90%	63467	42.49%	64711	43.19%
65956	43.81%	67200	44.40%	68444	45.04%	69689	45.67%
70933	46.31%	72178	46.88%	73422	47.57%	74667	48.10%
75911	48.68%	77156	49.35%	78400	49.85%	79644	50.56%
80889	50.85%	82133	51.68%	83378	52.06%	84622	52.47%
85867	53.11%	87111	53.72%	88356	54.18%	89600	54.80%
90844	55.32%	92089	55.79%	93333	56.29%	94578	56.79%
95822	57.34%	97067	57.59%	98311	58.37%	99556	58.82%
100800	59.16%	102044	59.55%	103289	60.25%	104533	60.67%
105778	61.11%	107022	61.71%	108267	62.25%	109511	62.63%
110756	63.12%	112000	63.58%	113244	63.96%	114489	64.26%
115733	65.04%	116978	65.37%	118222	65.64%	119467	66.38%
120711	66.64%	121956	67.09%	123200	67.33%	124444	67.78%
125689	68.21%	126933	68.62%	128178	69.23%	129422	69.49%
130667	70.19%	131911	70.53%	133156	70.70%	134400	71.30%
135644	71.48%	136889	71.94%	138133	72.26%	139378	72.72%
140622	73.09%	141867	73.54%	143111	73.80%	144356	73.98%
145600	74.69%	146844	74.92%	148089	75.49%	149333	75.61%
150578	75.91%	151822	76.60 %	153067	76.79%	154311	77.12%
155556	77.15%	156800	78.06%	158044	78.09%	159289	78.61%
160533	79.14%	161778	79.21%	163022	79.45%	164267	79.79%
165511	80.35%	166756	80.69%	168000	81.06%	169244	81.27%
170489	81.53%	171733	81.81%	172978	82.36%	174222	82.10 %
175467	83.09%	176711	82.90%	177956	83.41%	179200	83.96%
180444	84.00%	181689	84.41%	182933	84.39%	184178	85.03%

Table 3.6: Expected bias as function of $N/(\sigma^2 {x_0}^2)$

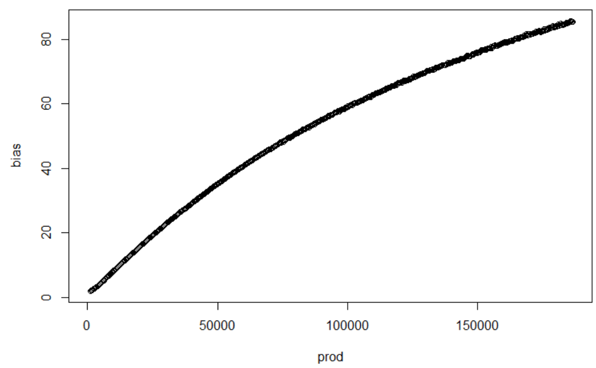


Figure 7: Bias(%) vs. $\frac{N}{\sigma^2 x^2}$

3.4 Bias for both discreteness and big values of dt

We have now showed that for small values of dt, there is bias due to discreteness, for big values of dt there is bias due to realized variance isn't approximating quadratic variation good enough. In figure 8 we see results of data being discrete and varying dt. In this example $\sigma = 0.01, \mu = 0.005, x_0 = 30$ dt is given in seconds and the bias is given in (bias/sigma)%. In this example we have the lowest bias when dt is about five minutes.

4 CIR-model

4.1 Backgrond

The Cox-Ingersoll-Ross model (CIR model) Cox, Ingersoll and Ross (1985, [7]) was originally used to model short time interest ratesand is used in valuation of interest rate derivatives. In 1993 [12] Heston used the model to model stochastic volatility for asset prices. In the CIR model the variable Y is given by the stochastic differential equation

$$dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dB_t \tag{9}$$

where κ is strictly positive and by adding the condition $2\kappa\theta \geq \sigma^2$ we ensure that Y_t is strictly positive, negative Y_t does not make much sense when used to model assets or volatilities.

The process is also known as the mean-reverting square-root process. Mean reverting means that the variable will tend to a long run mean value, this value is θ, see Cox, Ingersoll and Ross (1985, [7]) for details. The process is taking the leverage effect into account, that means that volatility of a variable increases as the value of the variable falls and decreases when the variable value rises.

We mainly introduce this model because it's a part of The Heston model and we only do analysis for simulated data.

We here introduce a theorem that we will use for motivating our second estimator.

Theorem 4.1. Duffie and Kan(1996, [8])*: There exists an unique strong solution to the CIR-process*

$$dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dB_t \tag{10}$$

if $2\kappa\theta \geq \sigma^2$

4.2 Quadratic variation of the CIR-process

The quadratic variation of the CIR-process over [0,T] is

$$QV_T(Y_t) = \sigma^2 \int_0^T (\sqrt{Y_t})^2 dt = \sigma^2 \int_0^T Y_t dt \tag{11}$$

Y_t is known at discrete points, and if σ can be estimated, we can estimate the quadratic variation.

4.3 Volatility

From our definition of volatility we see that the spot volatility is $\sigma/\sqrt{Y_t}$. Y_t is as mentioned known at discrete points, giving that to find $\sigma(Y_t)$ our biggest challange is to estimate the parameter σ. In our first approach in estimating the volatility we first make the assumption that Y_{t_i} is close to $Y_{t_{i+1}}$ over the interval $[t_{i-1}, t_i]$. We have information about Y_t so conclusions about this assumptions can be done. By doing so we can approximate $\sqrt{Y_t}$ over the interval $[t_0, T]$ by $Y_t/\sqrt{Y_{t_0}}$ and $Y_t/Y_{t_0} \approx 1$ giving
$\kappa(\theta - Y_t) \approx \kappa(\theta Y_t/Y_{t_0} - Y_t) = Y_t\kappa(\theta/Y_{t_0} - 1)$

With these approximations we get

$$dY_t \approx \kappa\Big(\frac{\theta}{Y_{t_0}} - 1\Big)Y_t dt + \sigma\frac{Y_t}{\sqrt{Y_{t_0}}}dB_t \tag{12}$$

We recognize this as the equation for the GBM with volatility $\sigma/\sqrt{Y_{t_0}}$ and the volatility dependence of the price becomes obvious.

This gives the estimator

$$\hat{\sigma}^2 = \sum_{i=0}^{n-1}(y_{t_{i+1}} - y_{t_i})^2 Y_{t_0}/\delta \tag{13}$$

As a second estimator we apply lemma 2.1 and theorem 4.1 and get

$$\tilde{\sigma^2} = \frac{\sum_{i=0}^{N-1}(Y_{t_{i+1}} - Y_{t_i})^2}{\sum_{i=0}^{N-1} Y_{t_i} dt} \tag{14}$$

This estimator is also natural when observing the integral for the quadratic variation.

Then we use a quasi-maximum likelihood method (15). This estimator is biased-corrected. For more justification of the estimator see Tanh and Chen (2009, [17]).

The estimator for the quasi-ml method is :

$$\hat{\sigma}^2 = \frac{2\hat{\kappa}\hat{\beta}_3}{1 - \hat{\beta}_1^2} \tag{15}$$

with:

$$\hat{\beta}_1 = \frac{n^{-2}(\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} X_{i-1}^{-1}) - n^{-1}\sum_{i=1}^{n}(X_i X_{i-1}^{-1})}{n^{-2}(\sum_{i=1}^{n} X_{i-1})(\sum_{i=1}^{n} X_{i-1}^{-1}) - 1}$$

$$\hat{\beta}_2 = \frac{n^{-1}\sum_{i}^{n}(X_i X_{i-1}^{-1}) - \hat{\beta}_1}{(1-\hat{\beta}_1)n^{-1}\sum_{i=1}^{n} X_{i-1}^{-1}}$$

$$\hat{\beta}_3 = n^{-1}\sum_{i=1}^{n}((X_i - X_{i-1}\hat{\beta}_1 - \hat{\beta}_2(1-\hat{\beta}_1))^2 X_i^{-1})$$

$$\hat{\kappa} = -\delta^{-1}\log(\hat{\beta}_1)$$

$$\delta = t/T$$

4.4 Modeling

When simulating the CIR-process there are several approaches to take. One simple approach is the use of Euler Maruyama Scheme which gives

$$Y_{t_{i+1}} = Y_{t_i} + \Delta_i(\theta - \kappa Y_{t_i}) + \sigma\sqrt{Y_{t_i}}(B_{t_{i+1}} - B_{t_i}) \tag{16}$$

Here $\Delta_i = (t_{i+1} - t_i)$ This scheme is though not well defined, this is due to we might get negative values but the CIR-process is strictly positive. There are simple ways to get around this, but we us an implicit scheme that has the positivity property. This scheme is presented by Brigo and Alfonsi(2005, [5]), they calls this "The Euler Implicit Positive-Preserving Scheme". They find following formula for $Y_{t_{i+1}}$

$$Y_{t_{i+1}} =$$

$$= \left(\frac{\sigma(B_{t_{i+1}} - B_{t_i}) + \sqrt{(\sigma^2(B_{t_{i+1}} - B_{t_i})^2 + 4(Y_{t_i} + (\kappa\theta - \frac{\sigma^2}{2(1+\kappa\Delta_i)})(\Delta_i(1+\kappa\Delta_i))}}{2(1+\kappa\Delta_i)}\right)^2$$

4.5 Results

Simulations are made with varying values of the parameters. Results show that the first estimator, built on realized variance is heavily biased, but this seems not to be a big problem. Simulations indicates that the bias is about

-20% independent of κ, θ, σ, n and δ. This gives us a new estimator $\tilde{\sigma} = c * \hat{\sigma}$ where we found that $c \approx 1.25$. We use the value c=1.2538 for our bias correction.

In table 4.1 we see a comparison of results from using the three different methods: The likelihood estimator σ_{ml} from equation (15) with the bias corrected quadratic method σ_{qv1} and the method based on lemma 2.1 (σ_{qv2}). We notice that all the methods gives good estimations of the volatility parameter σ, The second quadratic variation estimator and the maximum likelihood estimator seems to be equally good in estimating σ in a variance point of view. In general the first method gives estimates that have a standard deviation of 5-10% higher than the other two estimators. We also notice that the quasi-mle method gives better estimations when observing $|bias|$, than the other two methods which seems to be equally good from this point of view.

These results indicates that the quasi-maximum likelihood estimator is better, but not by that much. By optimizing the constant C we can expect a lower bias for the first method, compared to the current bias. In table 4.1 all parameters values but dt are varied, dt is constant equal to 1 minute. The results of varying dt with one set of parameter values is presented in table 4.2.

These result indicates that varying dt only have a small influence in the results, and these influence is only noticeably for quite large values of dt. In table 4.2 we see effects of varying dt.

One day 1 data point per min								
θ	κ	σ	$\mathrm{sd}(\bar{\hat{\sigma}}_{\mathrm{qv1}})$	$\mathrm{sd}(\bar{\hat{\sigma}}_{\mathrm{qv2}})$	$\mathrm{sd}(\bar{\hat{\sigma}}_{\mathrm{ml}})$	$\mathrm{bias}_{\mathrm{qv1}}$	$\mathrm{bias}_{\mathrm{qv2}}$	$\mathrm{bias}_{\mathrm{ml}}$
0.01	0.01	0.02	0.054%	0.051%	0.051%	-0.01%	-0.08%	-0.03%
0.01	1	0.02	0.054%	0.051%	0.050%	-0.21%	-0.26%	-0.10%
0.1	0.01	0.02	0.052%	0.048%	0.048%	0.07%	-0.01%	0.05%
1	0.01	0.02	0.054%	0.051%	0.050%	0.01%	-0.08%	-0.02%
1	1	0.02	0.054%	0.051%	0.051%	-0.16%	-0.22%	-0.08%
0.03	0.1	0.05	0.055%	0.053%	0.052%	0.04%	-0.03%	0.06%
0.03	10	0.05	0.054%	0.050%	0.050%	-1.86%	-1.93%	-1.23%
0.1	0.1	0.05	0.052%	0.048%	0.049%	-0.14%	-0.18%	-0.11%
0.1	10	0.05	0.053%	0.050%	0.050%	-1.90%	-1.95%	-1.26%
0.3	0.1	0.05	0.053%	0.049%	0.049%	-0.06%	-0.10%	-0.02%
0.3	10	0.05	0.055%	0.051%	0.051%	-1.77%	-1.86%	-1.28%
1	0.1	0.05	0.055%	0.051%	0.051%	0.02%	-0.02%	0.05%
1	10	0.05	0.055%	0.052%	0.051%	-1.89%	-1.97%	-1.26%
0.1	0.3	0.15	0.053%	0.053%	0.050%	0.05%	-0.06%	0.067%
0.1	10	0.2	0.053%	0.051%	0.050%	-1.78%	-1.83%	-1.27%
0.3	0.1	0.2	0.053%	0.054%	0.051%	0.08%	0.06%	0.11%
1	0.1	0.2	0.054%	0.052%	0.051%	0.10%	0.02%	0.10%
1	10	0.2	0.052%	0.049%	0.049%	-1.85%	-1.92%	-1.22%
10	10	0.2	0.052%	0.049%	0.049%	-1.91%	-1.98%	-1.27%

Table 4.1: Estimating σ in the CIR-model using three different methods.

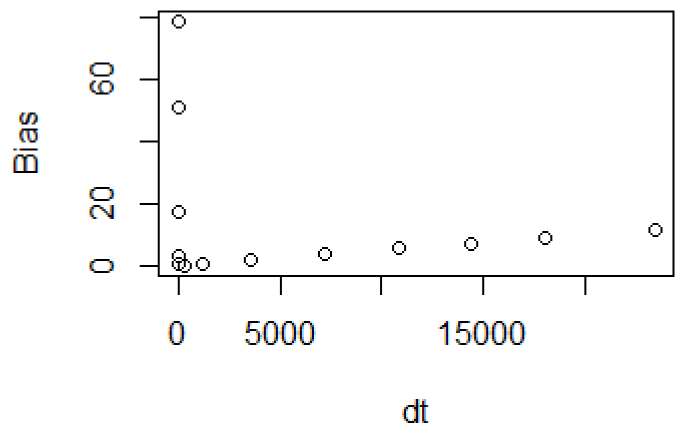


Figure 8: Bias in % for different dt.

$\kappa = 0.1$, $\theta = 1$, $\sigma = 0.1$				
dt	$\hat{c}$	$\sigma_{\mathrm{qv1}}/\sigma$	$\sigma_{\mathrm{qv2}}/\sigma$	$\sigma_{\mathrm{ml}}/\sigma$
0.1s	1.254	1.000	0.9991	1.000
1s	1.254	1.000	0.9992	1.000
5s	1.253	1.000	0.9990	1.000
15s	1.253	1.001	0.9995	1.001
1m	1.254	0.9998	0.9988	1.000
5m	1.255	0.9992	0.9983	1.000
15m	1.258	0.9969	0.9959	0.9986
30m	1.260	0.9948	0.9937	0.9975
1h	1.268	0.9887	0.9879	0.9934
2h	1.281	0.9786	0.9778	0.9869
4h	1.310	0.9572	0.9573	0.9731
6.5h	1.345	0.9322	0.9329	0.9565
13h	1.436	0.8732	0.8755	0.9157

Table 4.2: Time dependence of bias.

5 The Heston model

The Heston model is named after Steven Heston who 1993 introduced the model, Heston(1993, [12]). The model is a two factor model. It is a diffusion model with stochastic volatility, where the volatility follows a CIR-process. The importance of stochastic volatility is well known, and this is a strenght of the model.

5.1 The model

In the Heston model we have that:

$$dS_t = \mu S_t dt + \sqrt{\upsilon_t} S_t dB_t^S \tag{17}$$

$$d\upsilon_t = \kappa(\theta - \upsilon_t)dt + \xi\sqrt{\upsilon_t} dB_t^\upsilon \tag{18}$$

Where

$$\mathrm{corr}(dB_t^S, dB_t^\upsilon) = \rho \tag{19}$$

in Mariani,Pacelli and Zirilli (2008, [14]) we find

$$ds_t = (\mu - \frac{1}{2}\upsilon_t)dt + \sqrt{\upsilon_t} dB_t^S \tag{20}$$

$$d\upsilon_t = \kappa(\theta - \upsilon_t)dt + \xi\sqrt{\upsilon_t} dB_t^\upsilon \tag{21}$$

with $s_t = log(S_t/S_0)$, $s_0 = 0$. That is $S_0 = 1$ which is obtained by scaling.

We have that

- θ is the long run mean variance
- ξ is the volatility of the volatility
- κ is the mean reverting rate of υ_t
- μ is the long run rate of return

Further ρ gives us a relationship between volatility and return rate. In reality it is observed that when stock rises in value the volatility decreases and as the value of stocks decreases volatility rises. This phenomena is taken into account by letting ρ be negative. As earlier, when studying the CIR-process, we introduce the condition that $2\kappa\theta/\xi^2 > 1$ to ensure $\upsilon_t > 0$.

Here $\sqrt{\upsilon_t}$ is the volatility process and the quantity that we are interested in approximating. The volatility is stochastic and driven by the CIR-process (18).

Thus the volatility is mean reverting, this phenomena is observed by Fouqe, Papanicolao and Sircar(2000, [9]), when they studied the S&P 500 index. We have that θ and ξ are the two parameters that we are interested in studying.

We are also interested in finding the spot volatilities υ_t. By estimating the spot volatilities we might get a time serie that approximates the volatility process (18). If the approximation shares volatility properties with the process (17) it might be possible to estimate the parameter ξ.

There is an apparent connection between the Heston model and the geometrical Brownian motion, and if $\xi = 0$ in the volatility process we again have a non-stochastic volatility model that is the GBM (using $\upsilon_0 = \theta$). The bigger ξ the bigger deviations from the GBM we expect.

5.2 Modeling

When modeling we use the discretization proposed by Andersen (2008, [2])

$$\ln \hat{S}_{t_{i+1}} = \ln \hat{S}_t - \frac{1}{2}\hat{\upsilon}_t{}^+\Delta_i + \sqrt{\hat{\upsilon}_t{}^+}Z_S\sqrt{\Delta_i} \tag{22}$$

$$\hat{\upsilon}_{t_{i+1}} = \hat{\upsilon}_{t_i} + \kappa(\theta - \hat{\upsilon}_t{}^+)\Delta_i + \xi\sqrt{\hat{\upsilon}_{t_i}^+}Z_\upsilon\sqrt{\Delta_i} \tag{23}$$

we have that $\hat{\upsilon}^+ = \max(\upsilon, 0)$. Further Z_S and Z_υ have correlation ρ. This can be modeled by letting Z_1, Z_2 be independent variables from the standard normal distribution and the letting $Z_\upsilon = Z_1$ and $Z_S = \rho Z_1 + \sqrt{1-\rho^2}Z_2$.We

5.3 Volatility Heston

If we assume υ_t is constant equal to υ_{t_0} over the interval $[t_0, t_n]$ then the process is approximated as a GBM over each interval. We make this assumption, dividing the the time interval studied into small subintervals and use moving average.

Using the same approach as in GBM with quadratic variation we can estimate υ_{t_n} with the estimator

$$\hat{\upsilon}_{t_n} = \frac{1}{ndt}\sum_{i=1}^{n}(s_{t_{i+1}} - s_{t_i})^2 \tag{24}$$

Theorem 5.1. Castilla(2006, [6])*: There is a strong solution for the Heston model.*

From equation (20) we see that dx is uniquely defined from υ_t. We can apply lemma 2.1, this gives

$$\hat{\upsilon}_t = \frac{1}{ndt}\sum_{i=1}^{n}(s_{t_{i+1}} - s_{t_i})^2 \tag{25}$$

We find that these two estimators are the same! Applying theorem 5.1 gives that this is an consistent estimator of the spot volatility.

Using this method we can calculate the volatility over the intervals studied, but we don't get a value of the volatility parameter θ. We have that as $t \to \infty$ $(v_1 + \ldots + v_{n-1} + v_n)/n- > \theta$. This gives that θ can be approximated using big values of n and T.

The long run mean of υ_t is θ, though simulations show that this might converge slowly and intraday observations might not be enough to find good estimations of θ.

Another way to motivate the approach taken is that the expected quadratic variation of the Heston model, details given by Itkin and Carr(2010, [13]), is

$$E(\frac{1}{T}\int_0^T \upsilon_t dt) = \theta + (\upsilon_0 - \theta)\frac{1-e^{-\kappa T}}{\kappa T} \tag{26}$$

with:

$$\lim_{T\to 0} \theta + (\upsilon_0 - \theta)\frac{1-e^{-\kappa T}}{\kappa T} = \upsilon_0 \tag{27}$$

and:

$$\lim_{T\to \infty} \theta + (\upsilon_0 - \theta)\frac{1-e^{-\kappa T}}{\kappa T} = \theta \tag{28}$$

and further

$$E[v_t] = \theta \tag{29}$$

The second limit implies that we can estimate θ by looking at as long time frames as possible and the first limit motivates us to look at short time intervals when estimating spot volatility, that is we use the method moving average. From equation (26) and (27) we draw the conclusion that the possibilities to estimate θ increases as κ increases. We also expect the bigger ξ, the harder it is to estimate the spot volatility.

We do not find a maximum likelihood estimator for θ, the use of maximum likelihood is a hard optimization problem and in the literature we only find maximum likelihood methods that is applied on option data or a mix of option and stock data instead of just stock data.

Another concern might be to approximate the volatility parameter in the underlying CIR-process, that is the volatility of the volatility. From the simulations for the CIR-process we know that this parameter can be estimated with good precision if we have data from the process. The problem is though that we don't have data from the volatility process. So to take the same approach as earlier we first have to use our data to find approximate data for the underlying CIR-process and then use these data for further approximations.

5.4 Results

5.4.1 Estimating θ and approximating the volatility process

In table 5.1 we compare the possibilities of approximating v_t. First we use the mean value $\overline{v}$ as estimator and secondly we use moving average and realized variance as estimator ($\hat{v}$).

In table 5.2 we see an example of the theoretical possibilities to estimate θ for different parameter values. First, we by long simulations find a distribution of v_0. The expected variance of the volatility is, as time tends to infinity : $\theta\xi^2/(2\kappa)$, so the variance isn't diverging to infinity as in the GBM case, this is due to the mean-reverting property. This property makes it reasonable that we can find a good approximate distribution of v, using long simulations.

Then the expected value of $(\frac{1}{T}\int_0^T v_t dt)$ using the distribution of v_0 found

Mean deviation from υ_t for two approximators						
mean$\lvert\upsilon - \overline{\upsilon}\rvert$	mean$\lvert\upsilon - \hat{\upsilon}\rvert$	ξ	κ	θ	T	dt
0.547	0.185	0.01	1	0.01^2	10m	15s
0.113	0.178	0.002	1	0.01^2	10m	15s
0.0569	0.174	0.002	1	0.02^2	10m	15s
0.0208	0.176	0.002	1	0.05^2	10m	15s
0.0101	0.176	0.002	1	0.1^2	10m	15s
0.484	0.180	0.02	1	0.02^2	10m	15s
0.208	0.176	0.02	1	0.05^2	10m	15s
0.102	0.175	0.02	1	0.1^2	10m	15s
0.491	0.179	0.1	1	0.1^2	10m	15s
0.036	0.176	0.002	10	0.01^2	10m	15s
0.0176	0.175	0.002	10	0.02^2	10m	15s
0.0071	0.176	0.002	10	0.05^2	10m	15s
0.00354	0.175	0.002	10	0.1^2	10m	15s
0.349	0.192	0.02	10	0.01^2	10m	15s
0.176	0.180	0.02	10	0.02^2	10m	15s
0.071	0.176	0.02	10	0.05^2	10m	15s
0.035	0.175	0.02	10	0.1^2	10m	15s
0.347	0.193	0.1	10	0.05^2	10m	15s
0.175	0.179	0.1	10	0.1^2	10m	15s

Table 5.1: Approximate υ_t comparing moving average and mean of process.

Estimating θ				
$\mathrm{E}\{\lvert\theta - \theta_v\rvert\}/\theta$	$\mathrm{E}\{\lvert\theta - \theta_{RV}\rvert\}/\theta$	κ	θ	ξ
0	0.147	0.5	0.05^2	0
0.0793	0.162	0.5	0.05^2	0.005
0.160	0.209	0.5	0.05^2	0.01
0.308	0.339	0.5	0.05^2	0.02
0.441	0.468	0.5	0.05^2	0.03
0.608	0.618	0.5	0.05^2	0.04
0	0.143	2	0.05^2	0
0.0779	0.165	2	0.05^2	0.01
0.230	0.270	2	0.05^2	0.03
0.372	0.398	2	0.05^2	0.05
0.534	0.548	2	0.05^2	0.07
0.639	0.657	2	0.05^2	0.09
0	0.142	0.5	0.02^2	0
0.194	0.242	0.5	0.02^2	0.005
0.387	0.407	0.5	0.02^2	0.01
0.552	0.566	0.5	0.02^2	0.015
0	0.144	2	0.02^2	0
0.198	0.244	2	0.02^2	0.01
0.380	0.409	2	0.02^2	0.02
0.551	0.577	2	0.02^2	0.02

Table 5.2: Estimating θ for 1 hour of observations and dt=1 min.

from simulations and equation (26) is estimated. The expected absolute value of the bias is presented in table 5.2 together with the bias/θ and the quadratic error of the spot volatility using first quadratic variation as estimator and secondly the calculated mean value of the volatility. As expected, a larger ξ makes the first method more beneficial for calculating spot volatility. We also notice that smaller values of ξ gives smaller errors when estimating θ. In the table we can see theoretical values of how good θ can be estimated using quadratic variation, and one hour of observations.

5.4.2 Estimating ξ for simulated data

We here use $dt = 1s$, this to improve the estimations, and five minutes of observation to approximate υ_t and then use every 4 value in υ_t to find ξ, which gives us 4 time series. This is done to reduce the influence of the random deviations for our estimated time serie of υ_t. This bias reducing method is mentioned in subsection 3.3.2.

. Remember that the influence of the random error decreases as the interval used increases, and even though the approximated process is closed to the true process, in mean, this does not neccecarily give that the approximate process have the same variance properties as the real process.

In table 5.3 we see the results from estimations of ξ. As we can se the method used and choices of dt and interval length is suitable when κ and ξ is large, else we have heavily biased estimations. If ξ is small, it's necessarily for θ to be small aswell for the estimations to have small bias. As we can see estimations of ξ, with the parameter values chosen, are in general not very satisfactory.

5.4.3 Estimating ξ for stock data

The data we use when estimating ξ for stocks are not equally apart in time and the mean value of dt is bigger than 1.5 seconds for the studied time-series. We have studied 10 different time-series of length one day. The estimated values of ξ for the different time series lays within the interval $[0.0005, 0.005]$ and for these series θ is within the interval $[0.01, 0.025]$. From the results found from simulations we see that for θ and ξ in the given intervals the estimations are likely to be heavily biased, with estimations being too high. Another unfortunate result is that the estimations of ξ vary significantly

$\sqrt{\theta}$	κ	ξ	$\hat{\xi}$	sd$(\hat{\xi})/\xi$	(bias)$/\xi$	$\sqrt{\theta}$	κ	ξ	$\hat{\xi}$	sd$(\hat{\xi})/\xi$	(bias)$/\xi$
0.01	50	0.001	0.0070	0.155	602%	0.01	5	0.03	0.0362	0.0385	20.6%
0.01	50	0.005	0.0079	0.0447	58.1%	0.03	5	0.001	0.0218	0.472	366%
0.01	50	0.01	0.0102	0.0217	1.88%	0.03	5	0.005	0.0233	0.116	208%
0.01	50	0.02	0.0156	0.0244	-22.1%	0.03	5	0.01	0.0242	0.054	142%
0.01	50	0.03	0.0232	0.0169	-22.4%	0.03	5	0.03	0.0435	0.0351	45.1%
0.01	50	0.04	0.0295	0.0326	-26.1%	0.03	5	0.05	0.0686	0.0341	37.2%
0.01	50	0.05	0.0396	0.0258	-26.1%	0.03	5	0.08	0.109	0.230	36.4%
0.01	50	0.08	0.0530	0.0216	-33.7%	0.05	5	0.005	0.035	0.161	602%
0.03	50	0.005	0.0215	0.0913	331%	0.05	5	0.01	0.0390	0.104	290%
0.03	50	0.01	0.0224	0.0378	124%	0.05	5	0.03	0.0514	0.0414	71.4%
0.03	50	0.03	0.0314	0.0231	4.82%	0.05	5	0.05	0.0710	0.0316	41.9%
0.03	50	0.05	0.0441	0.0240	-11.7%	0.05	5	0.08	0.111	0.0503	38.4%
0.03	50	0.08	0.0615	0.0191	-23.2%	0.05	5	0.15	0.187	0.0420	24.9%
0.03	50	0.15	0.118	0.0263	-20.8%	0.01	0.5	0.001	0.0072	0.132	616%
0.05	50	0.01	0.0363	0.0664	263%	0.01	0.5	0.005	0.0096	0.0487	92.1%
0.05	50	0.03	0.0418	0.0299	39.2%	0.01	0.5	0.009	0.0137	0.0679	52.1%
0.05	50	0.05	0.0517	0.0232	3.33%	0.03	0.5	0.001	0.0208	19.8	1950%
0.05	50	0.08	0.0705	0.0131	-11.8%	0.03	0.5	0.01	0.0261	0.111	161%
0.05	50	0.15	0.117	0.0192	-22.2%	0.03	0.5	0.02	0.0333	0.0660	66.8%
0.01	5	0.001	0.0075	0.148	647%	0.05	0.5	0.005	0.0344	0.109	589%
0.01	5	0.005	0.0088	0.0539	76.5%	0.05	0.5	0.01	0.0374	0.139	274%
0.01	5	0.01	0.0135	0.0370	35.3%	0.05	0.5	0.03	0.0544	0.0699	81.4%
0.01	5	0.02	0.0250	0.0409	24.4%	0.05	0.5	0.045	0.0689	0.0501	53.1%

Table 5.3: Estimating ξ (volatility of volatility)

when we change the number of values used when applying moving average. ξ might appear small, but we have to remember that the spot volatility of the volatility process is not ξ, but $\xi/\sqrt{\upsilon_t}$. A typical value of $\sqrt{\upsilon_t}$ is 0.02. If ξ is in the interval values of ξ for the different time series lays within the interval $[0.0005, 0.005]$, as our estimations indicates, this gives a volatility of volatility in the interval $[0.0005/0.02, 0.005/0.02] = [0.025, 0.25]$. Futher, this indicates a high volatility of the volatility, even though these values probably, from results of simulations, are big overestimations.

The results from estimating ξ are also compared with results of estimating ξ for simulated geometrical Brownian motions ($\xi = 0$), using matching values of observation times and of parameter values for the moving average, with volatilities being equal to the volatilities (θ) estimated for the real time-series used. The estimations of ξ for these simulated geometrical Brownian motions are often only $2 - 20\%$ of the estimations for the corresponding time-series for the stocks. This indicates that we don't have a constant volatility, which contradicts the assumption in the GBM model of constant volatility.

In figure 9 we see an example of estimated volatility process for real stock data. The corresponding time serie is used to estimate the volatility of volatility.

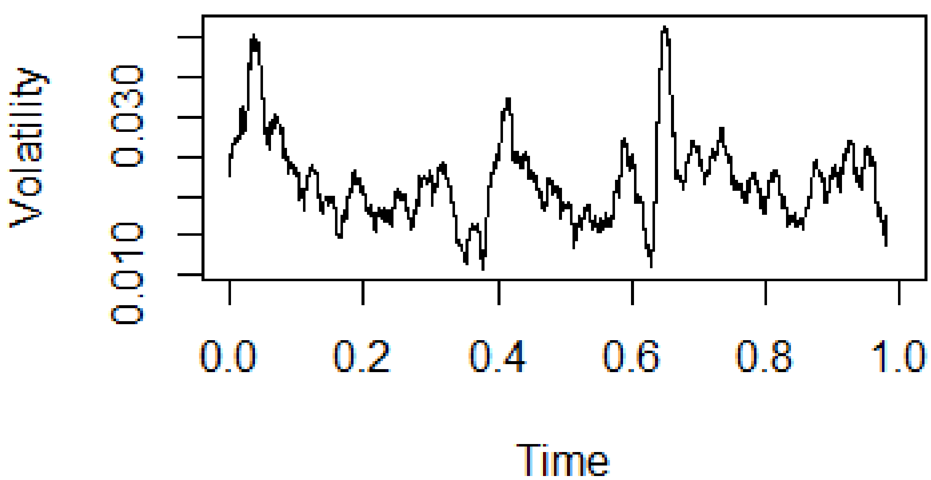


Figure 9: Estimated volatility of Facebook 30th October 2013.

6 Discussion

As we expected realized variance is useful when estimating volatility parameters, but if one wants to approximate the volatility process efficiently one must find a method to choose timescale wisely.

When analyzing high frequency data from geometrical Brownian motions we find effects on bias for discrete data. In our simulations we don't take into account is that in reality stocks aren't traded with constant frequency. Making time steps random is one way to make further analysis, and also take into account that it's not only the value of stocks that are discrete for real time serie data but also the trading times. Further, if time steps are random then we can not use bias-correction as proposed in algorithm 1.

In the CIR-model we found a bias of our first estimator the we could not explain, to do this is a future challenge. In the Heston model, when estimating the volatility of volatility-parameter, there are great possibilities to refine the methods we use. The time scale we use in our moving average could be customized and changed, based on estimations and the the way we use the pre-bias removal should also be refined.

7 Conclusions

Firstly we found that realized variance is useful even when data is low frequent. After that we analyzed the micro economic effects of data being discrete and also found a way to remove the bias, under the assumption of data being equally apart in time. We found that realized variance and maximum likelihood as methods were almost equally good, both in respect for variance of the estimations and bias.

When approximating the volatility process for the Cir-process we saw that the spot volatility is only dependent on the volatility parameter and the value of the process. The value of the process is known at discrete times, which gave that estimating the volatility parameter was the challange we had. We used three different approaches to estimate the volatility parameter, the realized variance method, corresponding to the realized variance method used in the Brownian motion case gave a bias, but we could successfully remove it. We also used another approach, with a weighted realized variance that was about as good, but with lower bias. The best method was the quasi-maximum likelihood method, but the realized variance method was almost as good.

In the Heston model we had three different challenge's: to approximate the volatility process, to estimate the volatility parameter and finally to estimate the volatility of the volatility parmater. We found that, with data being high frequent and with a good choice of time scale when using moving average, our estimator of the process was better than using the mean as estimator, in a squared error sense. Secondly we found that estimating the volatility parameter with realized variance, is possible, if we study the process over a long time. The volatility of the volatility was even harder and under most circumstances our estimations were not very good. We found that for small values of the volatility of volatility parameter our estimations would be to high. When analyzing real time serie data our estimations gave small values of the volatiity of volatility parameter. This gave that the estimator used was not good enough to estimate the parameter, but the estimations could be used to find an upper limit of the parameter value. We also draw the conclusion that the volatility were in fact varying and not constant as in our first model.

References

[1] Aït-Sahalia,Y and Mancini,L. (2008) Out of sample forecasts of quadratic variation *Journal of Econometrics*, Volume 147, Issue 1, pp. 17–33 URL: http://dx.doi.org/10.1016/j.jeconom.2008.09.015

[2] Andersen, L. (2008) Simple and efficient simulation of the Heston stochastic volatility model *Journal of Computational Finance* (1–42) Volume 11/Number 3

[3] Barndorff-Nielsen,O.E and Shephard, N (2002) Estimating Quadratic Variation Using Realized Variance *Journal of Applied Econometrics* , Vol. 17, No. 5, Special Issue: modeling and Forecasting Financial Volatility pp. 457-477 URL: http://www.jstor.org/stable/4129267

[4] Black,F and Scholes,M (1973) *The Pricing of Options and Corporate Liabilities*
Journal of Political Economy, Vol. 81, No. 3 , pp. 637-654 URL: http://www.jstor.org/stable/1831029 .

[5] Brigo,D and Alfonsi,A (2005) Credit Default Swaps Calibration and Option Pricing with the SSRD Stochastic Intensity and Interest-Rate Model emphFinance and Stochastics, 9(1), pp. 29-42. doi:http://dx.doi.org/10.1007/s00780-004-0131-x

[6] Castilla,A.F (2012) Stochastic calculus and analytic chatacteristic functions: Applications to Finance *Tesis doctorals* Volume 145 Universitat Autonoma de Barcelona: Department of mathematics

[7] Cox,J.C, Ingersoll,J.E and Ross S.A (1985) A Theory of the Term Structure of Interest Rates *Econometrica*, Vol. 53, No. 2 pp. 385-407

[8] Duffie,D and Kan,R (1996) yield-factor model of interest rates *Mathematical Finance*, vol 6. no 4 :379-406, URL: http://www.stanford.edu/ duffie/YieldFactorModel.pdf

[9] Fouque,J., Papanicolaou, G., and Sircar, K. (2000) Mean-Reverting Stochastic Volatility *International Journal Of Theoretical & Applied Finance*, 3(1), 101 ISSN 0219-0249

[10] Graham, J.R, Michealy,R and Roberts,M.R (2003) Do Price Discreteness andTransactions Costs Affect Stock Returns? Comparing Ex-Dividend Pricing Before and After Decimalization *The Journal of finance Vol. LVIII, No. 6* URL: http://finance.wharton.upenn.edu/published_files/JOFI_05806012.PDF

[11] Hautsch,N (2012) Econometrics of Financial High-Frequency Data *Springer Berlin Heidelberg* E-book Online ISBN $978-3-642-21925-2$

[12] Heston,S.L (1993) A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options *The Review of Financial Studies* Vol. 6, No. 2 , pp. 327-343 URL: http://www.jstor.org.proxy.lib.chalmers.se/stable/2962057

[13] Itkin,A and Carr,P. (2010) Pricing swaps and options on quadratic variation under stochastic time change models–discrete observations case. *Review of Derivatives Research*, 13(2), pp 141-176. doi:http://dx.doi.org/10.1007/s11147-009-9048-z

[14] Mariani, F,Pacelli,G and Zirilli,F. (2008) Maximum likelihood estimation of the Heston stochastic volatility model using asset and option prices: an application of nonlinear filtering theory *Optimization Letters* Volume 2, Issue 2, pp 177-222 URL: http://dx.doi.org/10.1007/s11590-007-0052-7 Springer-Verlag

[15] Miao,WC. (2004) Quadratic variation estimators for diffusion models in finance *University of Southern California, ProQuest, UMI Dissertations* URL: http://search.proquest.com/docview/305125642?accountid=10041

[16] S.Ghon Rhee, Chi-Jeng Wang Rhee,S.G, Wang,CJ. (1997) The bid-ask bounce effect and the spread size effect: Evidence from the Taiwan stock market *Pacific-Basin Finance Journal* No. 5 pp. 231-258 URL: http://www.sciencedirect.com/science/article/pii/S0927538X97000140

[17] Tang,C.Y and Chen,S.X. (2009) Parameter estimation and bias correction for diffusion processes *Journal of Econometrics*, Volume 149, Issue 1 pp. 65-81 URL: http://songchen.public.iastate.edu/Tang-Chen-09-JoE.pdf

[18] Yamada,T and Watanabe,S. (1971) On the uniqueness of solutions of stochastic differential equations *J. Math. Kyoto Univ. (JMKYAZ)* No. 11-1 pp. 155-167

[19] Zhou,B. (1996) High-frequency data and volatility in foreign-exchange rates *Journal of Business Economic Statistics* , Vol. 14, No. 1, pp. 45-52 URL: http://www.jstor.org.proxy.lib.chalmers.se/stable/1392098 Published by: American Statistical Association Article Stable URL: http://www.jstor.org/stable/1392098

[20] Zubchenko,V.P. and Mishura,YU.S. (2010) Existence and uniqueness of solutions of stochastic differential equations with nonLlipschitz diffusion and Poisson measure *Theor. Probability and Math. Statist.* translated by N. N. Semenov Vip. 80, 2009 No. 80, pp. 47–59

List of Tables

List of Figures