FINANCE AND RISK ENGINEERING

FRE-GY.7801 PARTIAL DIFFERENTIAL EQUATION FOR FINANCE

SPRING 2022

FINAL WRITTEN PAPER

Heston Model and its Implementation on WTI Options

Jonathan Sun (ws2381)
Simon Chau (cwc425)
Guojun Chen (gc2817)

Supervised By
Prof. Roza Galeeva

1 Heston Model (1993)

Our introduction to the Heston model(1993)[1] follows closely to the Chapter 2 in Gatheral (2012)[2]. Any mistake is the sole responsibility of the authors of this paper.

Consider underlying $S(t)$ and its volatility $v(t)$ satisfy the following SDE:

$$dS(t) = \mu(t)S(t)dt + \sqrt{v(t)}S(t)dW_1(t) \tag{1}$$

$$dv(t) = -\lambda(v(t) - \overline{v})dt + \eta\sqrt{v(t)}dW_2(t) \tag{2}$$

with

$$d[W_1(t), W_2(t)] = \rho dt$$

where $\mu(t)$ is the instantaneous drift of the underlying price returns, η is the volatility of volatility, and ρ is the correlation between the underlying price returns and changes in $v(t)$. $W_1(t)$ and $W_2(t)$ are Brownian motions. The variance $v(t)$ is asymptotically stable if $\lambda > 0$ and always positive if $2\lambda\overline{v} > \eta^2$, the latter also known as Feller's Condition for the square-root process.

Through the arbitrage arguments and stochastic calculus techniques, we can arrive the corresponding PDE satisfied by any derivative $V(S, v, t)$

$$V_t + \frac{1}{2}vS^2V_{ss} + \rho\eta vSV_{sv} + \frac{1}{2}\eta^2 vV_{vv} + rSV_s - rV = (\lambda(v - \overline{v}) + \phi(S, v, t)\sqrt{v})V_v$$

The market incompleteness implied by the structure of Heston Model leads to multiple measures under which the underlying grows at the risk-free rate r. Accordingly, the market price of volatility risk $\phi(S, v, t)$ can be any arbitrary function. In terms of pricing purpose, one can choose the measure with zero market price of volatility. To see this, use the Girsanov theorem to shift the drift term so that the market price of volatility disappears in the PDE. Specifically, if we define the risk-neutral drift of the volatility as $-(\lambda(v - \overline{v}) - \phi(S, v, t)\sqrt{v}$, the resulting PDE under the risk-neutral measure is:

$$V_t + \frac{1}{2}vS^2V_{ss} + \rho\eta vSV_{sv} + \frac{1}{2}\eta^2 vV_{vv} + rSV_s - rV = \lambda(v - \overline{v})V_v$$

The problem can be simplified through changes of variable. Let's first consider

$$\tau = T - t$$
$$x = log(Se^{r\tau}/K)$$

The changes of variable lead to the following relationships

$$V_s = \frac{V_x}{S},\ V_{ss} = \frac{V_{xx}}{S^2} - \frac{V_x}{S^2},\ V_t = -V_\tau - rV_x,\ V_v = V_v,\ V_{sv} = \frac{V_{xv}}{S},\ V_{vv} = V_{vv}$$

Substitute the relationships into the above PDE would yield

$$-V_\tau + \frac{1}{2}vV_{xx} - \frac{1}{2}vV_x + \frac{1}{2}\eta^2 vV_{vv} + \rho\eta vV_{xv} + rV_x = \lambda(v - \overline{v})V_v$$

Let's then consider the future value of the derivative $C(x, v, \tau) = e^{r\tau}V(x, v, t)$. Multiplying $e^{r\tau}$ on both sides of the PDE and using relationship $C_\tau = V_\tau e^{r\tau} + rVe^{r\tau}$ would arrive

$$-C_\tau + \frac{1}{2}vC_{xx} - \frac{1}{2}vC_x + \frac{1}{2}\eta^2 vC_{vv} + \rho\eta vC_{xv} = \lambda(v - \overline{v})C_v$$

According to Duffie, Pan and Singleton (2000)[3], the solution to our problem has the form similar to Black and Scholes (1973)[4]

$$C(x,v,\tau) = K[e^x P^{(1)}(x,v,\tau) - P^{(0)}(x,v,\tau)]$$

where $P^{(j)}, j \in \{0,1\}$ are the prices of each Arrow-Debreu security at the time to maturity. See Bakshi and Madan (2000)[5] for more discussion on the solution form. Moreover, as in the case of Black and Scholes (1973)[4], $P^{(1)}$ is the delta of the derivative, and $P^{(0)}$ is the risk-neutral probability of exercise. This paper is particularly interested in the European-style vanilla options. For the call option, the terminal condition is a pair of Heaviside function of x:

$$\lim_{\tau \to 0_+} P^{(j)}(x,v,\tau) = 1, \forall x > 0, \forall j \in \{0,1\}$$

$$\lim_{\tau \to 0_+} P^{(j)}(x,v,\tau) = 0, \forall x \leq 0, \forall j \in \{0,1\}$$

Substituting the solution form into the PDE yields:

$$-e^x P_\tau^{(1)} + \frac{1}{2} v e^x (P^{(1)} + 2P_x^{(1)} + P_{xx}^{(1)}) - \frac{1}{2} v e^x (P^{(1)} + P_x^{(1)}) + \frac{1}{2}\eta^2 v e^x P_{vv}^{(1)} + \rho\eta v e^x (P_v^{(1)} + P_{xv}^{(1)}) - \lambda(v - \overline{v}) P_v^{(1)}$$

$$= -P_\tau^{(0)} + \frac{1}{2} v P_{xx}^{(0)} - \frac{1}{2} v P_x^{(0)} + \frac{1}{2}\eta^2 v P_{vv}^{(0)} + \rho\eta v P_{xv}^{(0)} - \lambda(v - \overline{v}) P_v^{(0)}$$

Observe that the only way to make the above hold for the entire domain is that the left-hand side and the right-hand side are both equal zero for the entire domain. The two conditions can be compactly expressed as:

$$-P_\tau^{(j)} + \frac{1}{2} v P_{xx}^{(j)} - (\frac{1}{2} - j) v P_x^{(j)} + \frac{1}{2}\eta^2 v P_{vv}^{(j)} + \rho\eta v P_{xv}^{(j)} + (\lambda\overline{v} - \lambda v + j\rho\eta v) P_v^{(j)} = 0, \forall j \in \{0,1\}$$

Conducting the Fourier transformation to the above will arrive

$$-Q_\tau^{(j)} - \frac{1}{2} u^2 v Q^{(j)} - (\frac{1}{2} - j) iuv Q^{(j)} + \frac{1}{2}\eta^2 v Q_{vv}^{(j)} + \rho\eta iuv Q_v^{(j)} + (\lambda\overline{v} - \lambda v + j\rho\eta v) Q_v^{(j)} = 0, \forall j \in \{0,1\}$$

where

$$Q^{(j)}(u,v,\tau) = \int_{-\infty}^{\infty} e^{-iux} P^{(j)}(x,v,\tau) dx$$

Rearranging the terms for the above PDE will get

$$v\{\alpha^{(j)} Q^{(j)} - \beta^{(j)} Q_v^{(j)} + \gamma Q_{vv}^{(j)}\} + (\lambda\overline{v}) Q_v^{(j)} - Q_\tau^{(j)} = 0, \forall j \in \{0,1\}$$

where

$$\alpha^{(j)} = -\frac{u^2}{2} - \frac{iu}{2} + jiu$$
$$\beta^{(j)} = \lambda - j\rho\eta - \rho\eta iu$$
$$\gamma = \frac{\eta^2}{2}$$

Consider solution in the form

$$\forall j \in \{0,1\},\ Q^{(j)}(u,v,\tau) = Q^{(j)}(u,v,0)e^{C^{(j)}(u,\tau)\overline{v}+D^{(j)}(u,\tau)v}$$
$$= \frac{1}{iu}e^{C^{(j)}(u,\tau)\overline{v}+D^{(j)}(u,\tau)v}$$

The last equality follows from the fact that $\lim_{\tau\to 0} P^{(j)}(x,v,\tau)$ is the Heaviside function of x. From now on, we will temporarily drop index j in the coming equations since the structures are the same $\forall j \in \{0,1\}$. The above form would imply the following relationship

$$Q_\tau = Q(\overline{v}C_\tau + vD_\tau),\ Q_v = DQ,\ Q_{vv} = D^2Q$$

Substitute back into the PDE would arrive

$$v\{\alpha - \beta D + \gamma D^2 - D_\tau\} = \overline{v}C_\tau - (\lambda\overline{v})D, \forall j \in \{0,1\}$$

The above equation implies the following ODE:
Define:

$$r_\pm = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\gamma} = \frac{\beta \pm d}{\eta^2}, \forall j \in \{0,1\}$$

$$D_\tau = \alpha - \beta D + \gamma D^2 = \gamma(D - r_+)(D - r_-)$$
$$C_\tau = \lambda D$$

The solution to the above ODE is

$$D = r_-(\frac{1 - e^{-d\tau}}{1 - ge^{-d\tau}})$$
$$C = \lambda(r_-\tau - \frac{2}{\eta^2}log(\frac{1 - ge^{-d\tau}}{1 - g}))$$

where $j \in \{0,1\}$, $g = r_-/r_+$,
<Verify C and D are the solution>
For D,

$$D_\tau = r_-(\frac{de^{-d\tau}(1-g)}{(1-ge^{-d\tau})^2})$$
$$= r_-(\frac{\gamma(r_- - r_+)e^{-d\tau}(g-1)}{(1-ge^{-d\tau})^2})$$
$$= r_-\frac{\gamma(r_- - r_-e^{-d\tau} - r_+ + r_e^{-d\tau})(ge^{-d\tau} - e^{-d\tau})}{(1-ge^{-d\tau})^2}$$
$$= \gamma(r_-(\frac{1-e^{-d\tau}}{1-ge^{-d\tau}}) - r_+)(r_-(\frac{1-e^{-d\tau}}{1-ge^{-d\tau}}) - r_-)$$
$$= \gamma(D - r_+)(D - r_-)$$

For C,

$$\begin{aligned}
C_\tau &= \lambda(r_- - \frac{2}{\eta^2}\frac{dge^{-d\tau}}{1-ge^{-d\tau}}) \\
&= \lambda r_-(1 - \frac{\frac{2de^{-d\tau}}{\eta^2 r_+}}{1-ge^{-d\tau}}) \\
&= \lambda r_-(\frac{1-ge^{-d\tau} - \frac{2de^{-d\tau}}{\eta^2 r_+}}{1-ge^{-d\tau}}) \\
&= \lambda r_-(\frac{1-e^{-d\tau}}{1-ge^{-d\tau}}) \\
&= \lambda D
\end{aligned}$$

<end of verification>
The last step is to take the inverse transform from $Q^{(j)}$ back to $P^{(j)}$

$$\begin{aligned}
P^{(j)}(x,v,\tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{iux}Q^{(j)}(u,v,\tau)du \\
&= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} Re(\frac{e^{C^{(j)}(u,\tau)\overline{v}+D^{(j)}(u,\tau)v+iux}}{iu})du
\end{aligned}$$

The above inverse transform formula can be derived easily by knowing two facts. First, $\forall j \in \{0,1\}$, characteristic function $\varphi_T^{(j)}(u) = -\partial P^{(j)}/\partial k = e^{C^{(j)}(u,\tau)\overline{v}+D^{(j)}(u,\tau)v} = iuQ^{(j)}$. Second, the cumulative probability function $F(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} Re(\frac{\varphi_T^{(j)}(u)e^{-iux}}{iu})du$ (under the assumption of the existence of a mean for the random variable of interest). The second fact is the Theorem 3 in Shephard (1991)[6]; moreover, you can find the inversion theorem in Bakshi and Madan (1999)[5] whose framework is more general than Heston (1993)[1].

<proof for the characteristic function>
Let's first show that

$$\begin{aligned}
Q^{(j)}(u,v,\tau) &= \int_{-\infty}^{\infty} e^{-iux}P^{(j)}(x,v,\tau)dx \\
&= \int_{-\infty}^{\infty} e^{iux}P^{(j)}(-x,v,\tau)dx \\
&= (\frac{e^{iux}P^{(j)}(-x,v,\tau)}{iu})_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\frac{e^{iux}P_1^{(j)}(-x,v,\tau)}{iu}dx \\
&= \int_{-\infty}^{\infty}\frac{e^{iux}P_1^{(j)}(-x,v,\tau)}{iu}dx \\
&= \varphi_T^{(j)}(u)/iu
\end{aligned}$$

The third equality comes from integration by parts, and $P_1^{(j)}$ denotes the partial derivative with respect to the first argument

<end of proof>

2 Numerical Implementation

2.1 Closed-form solution

To summarize the previous section, the call option in the Heston model can be computed using the following equations. Our Python implementation follows exactly the flow from top to down.

$$\tau = T - t,\ \ x = log(Se^{r\tau}/K),\ \ \gamma = \frac{\eta^2}{2}$$
$$\alpha^{(j)} = -\frac{u^2}{2} - \frac{iu}{2} + jiu, \forall j \in \{0,1\}$$
$$\beta^{(j)} = \lambda - j\rho\eta - \rho\eta iu, \forall j \in \{0,1\}$$
$$d^{(j)} = \sqrt{\beta^2 - 4\alpha\gamma}, \forall j \in \{0,1\}$$
$$r_{\pm}^{(j)} = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\gamma} = \frac{\beta \pm d}{\eta^2}, \forall j \in \{0,1\}$$
$$g^{(j)} = r_-/r_+, \forall j \in \{0,1\}$$
$$D^{(j)} = r_-(\frac{1 - e^{-d\tau}}{1 - ge^{-d\tau}}), \forall j \in \{0,1\}$$
$$C^{(j)} = \lambda(r_-\tau - \frac{2}{\eta^2}log(\frac{1 - ge^{-d\tau}}{1 - g})), \forall j \in \{0,1\}$$
$$P^{(j)}(x,v,\tau) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re(\frac{e^{C^{(j)}(u,\tau)\overline{v} + D^{(j)}(u,\tau)v + iux}}{iu})du, \forall j \in \{0,1\}$$
$$V(x,v,\tau) = e^{-r\tau}C(x,v,\tau) = Ke^{-r\tau}[e^x P^{(1)}(x,v,\tau) - P^{(0)}(x,v,\tau)]$$

The challenges for implementation are the complex logarithm in $C^{(j)}$ and the improper integration in $P^{(j)}$. Many authors including Heston (1993)[1] write $C = \lambda(r_+\tau - \frac{2}{\eta^2}log(\frac{e^{d\tau} - g}{1-g}))$. Their solutions would coincide numerically with ours only if C is continuous with respect to u under the chosen complex logarithm. If one restrict the logarithm to its principle branch, the default setting for most software packages, C jumps whenever the imaginary part crosses the negative real axis. Kahl and Jackel (2005) [7] found an ingenious way to overcome the issue: rewrite the solution so that the imaginary part can only cross the positive real axis, as can be shown in our setting.

For improper integration, *integration.quad* in SciPy Python package works robustly and reliably. On the other hand, *integrate.quadrature* in the same package fails when $\tau > 0.4$ and the computed $P^{(j)}$ would both converge to 0.5, indicating that the integration returns 0 when τ is large. We have not found the reason behind the failure, which deserve further research. See Firgure 1 for the computed $P^{(1)}$(the vertical axis) with respect to τ (the horizontal axis) using *integrate.quadrature* in SciPy Python package.

In numerical integration, the infinity upper bound of the improper integration may cause a problem. There are several solutions to handle the infinity upper bound. 1) Transform the integration by logarithm such that the interval is restricted in $(0,1]$. 2) Truncated the interval by a reasonable large number. *integrate.quad* in SciPy uses the transformation technique and performs it efficiently and effectively.

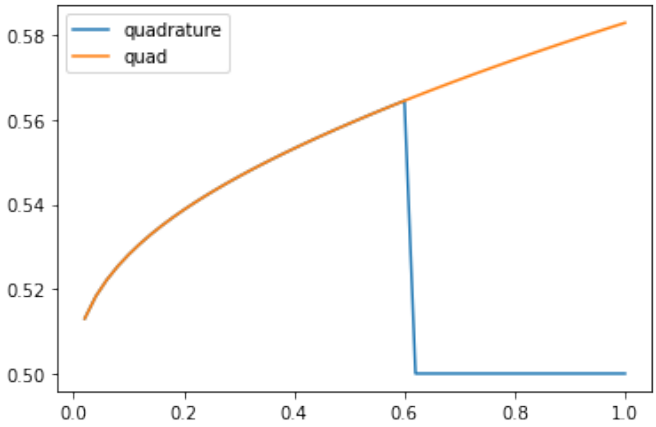

Figure 1: Computed $P^{(1)}(\tau)$ by quad and quadrature

2.2 Monte Carlo simulation with Mixing Theorem

Under the risk-neutral measure, the underlying $S(t)$ and its derivative can be computed using the martingale pricing technique. That is, the European call option on the underlying $V(s,v,t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S(T)-K)^+|\mathcal{F}(t)]$. The risk-adjusted SDE become

$$\begin{aligned} dS(t) &= rS(t)dt + \sqrt{v(t)}S(t)d\tilde{W}_1(t) \\ dv(t) &= -\lambda(v(t)-\overline{v})dt + \eta\sqrt{v(t)}d\tilde{W}_2(t) \end{aligned}$$

with

$$d[\tilde{W}_1(t), \tilde{W}_2(t)] = \rho dt$$

Through the change of variable $X = log(S/K)$ and introducing another Brownian motion, $\tilde{W}_3$, independent to $\tilde{W}_2$, the SDE become:

$$\begin{aligned} d\tilde{W}_1(t) &= \rho d\tilde{W}_2(t) + \sqrt{1-\rho^2}d\tilde{W}_3(t) \\ dv(t) &= -\lambda(v(t)-\overline{v})dt + \eta\sqrt{v(t)}d\tilde{W}_2(t) \end{aligned}$$

Define dX as $d\ln(S(t))$ we have:

$$\begin{aligned} dX = d\ln S(t) =& (r - \frac{1}{2}v(t))dt + \sqrt{v(t)}(\rho d\tilde{W}_2(t) + \sqrt{1-\rho^2}d\tilde{W}_3(t)) \\ =& rdt + (-\frac{1}{2}(1-\rho^2)v(t))dt + \sqrt{1-\rho^2}\sqrt{v(t)}d\tilde{W}_3(t) \\ &+ (-\frac{1}{2}v(t)\rho^2)dt + \rho\sqrt{v(t)}d\tilde{W}_2(t) \\ =& rdt - \frac{1}{2}(1-\rho^2)v(t)dt + \sqrt{1-\rho^2}\sqrt{v(t)}d\tilde{W}_3(t) + dY(t) \\ where,\ dY(t) =& -\frac{1}{2}v(t)\rho^2 dt + \rho\sqrt{v(t)}d\tilde{W}_2(t) \end{aligned}$$

Integrating both equation from 0 to T we have

$$v(T) = v(0) - \int_0^T \lambda(v(t) - \overline{v})dt + \eta \int_0^T \sqrt{v(t)}d\tilde{W}_2(t)$$
$$Y(T) = Y(0) - \int_0^T \frac{1}{2}v(t)\rho^2 dt + \int_0^T \rho\sqrt{v(t)}d\tilde{W}_2(t)$$
$$X(T) = X(0) + \int_0^T dY(t) + \int_0^T (r - \frac{1}{2}(1-\rho^2)v(t))dt + \int_0^T \sqrt{1-\rho^2}\sqrt{v(t)}d\tilde{W}_3(t)$$

Consider a discretized time process $t \in \{0, \Delta t, 2\Delta t,..., T\}$ and draw two independent standard normal variables $\tilde{Z}_2(t)$ and $\tilde{Z}_3(t), \forall t$, starting with $X(0) = log(\frac{S_0}{K}), v(0) = v_0, Y(0) = 0$. Let set $I = \{0, \Delta t, 2\Delta t,..., T - \Delta t\}$

$$v(T) = v(0) - \sum_{t\in I} \lambda(v(t) - \overline{v})\Delta t + \eta \sum_{t\in I} \sqrt{v(t)}\sqrt{\Delta t}\tilde{Z}_2(t)$$
$$Y(T) = -\frac{1}{2}\sum_{t\in I} \rho^2 v(t)\Delta t + \sum_{t\in I} \rho\sqrt{v(t)}\sqrt{\Delta t}\tilde{Z}_2(t)$$
$$X(T) = X(0) + Y(T) + rT - \frac{1}{2}\sum_{t\in I}(1-\rho^2)v(t)\Delta t + \sum_{t\in I}\sqrt{1-\rho^2}\sqrt{v(t)}\sqrt{\Delta t}\tilde{Z}_3(t)$$

Let's define

$$\sigma_T = \sqrt{\frac{\int_0^T (1-\rho^2(t))v(t)dt}{T}}$$

Using the fact $\int_0^T \sqrt{1-\rho^2}\sqrt{v(t)}d\tilde{W}_3(t) \overset{d}{=} \sigma_T \tilde{W}_3(T)$, we can attain the result in Mixing Theorem. To see the fact, it's sufficient to show that $\frac{1}{\sigma_T}\int_0^T \sqrt{1-\rho^2}\sqrt{v(t)}d\tilde{W}_3(t)$ is a standard Brownian motion.

<proof>

It's trivially to see that the process starts at 0. Furthermore, the process is a continuous martingale due to the property of Ito integral. Last, the quadratic variance (through Ito isometry) of the process is

$$\frac{1}{\sigma_T^2}\int_0^T (1-\rho^2)v(t)dt = T$$

By the Levy's characterization of Brownian motion, the process is indeed a standard Brownian motion

<end of proof>

Using the above fact and iterative expectation property:

$$e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[\,(S(T)-K)^+] = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[\,\mathbb{E}^{\mathbb{Q}}[\,(S(T)-K)^+|\,v(t), t\in[0,T]\,\,]\,\,]$$

Conditional on the realization of v, $S(t)$ follows geometric Brownian motion with constant volatility in distribution, so we can use the Black-Scholes call option formula for every realization of v. However, we need to adjust the initial stock price downward by $Y(T)$ which solely depends on the path of $v(t)$. This is the gist of the Mixing theorem. The overall procedure goes as follows: sample v and then calculate S_{eff} and σ_{eff}^2 (defined below). Apply the Black-Scholes call option formula $C_{BS}(S_{eff}, \sigma_{eff}, T)$. Repeat many times to find the sample mean for

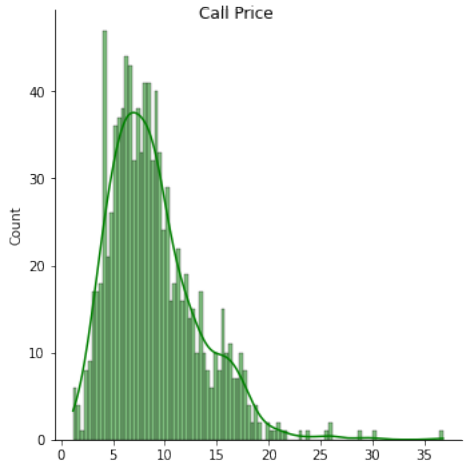

Figure 2: Histogram for call (n = 1,000)

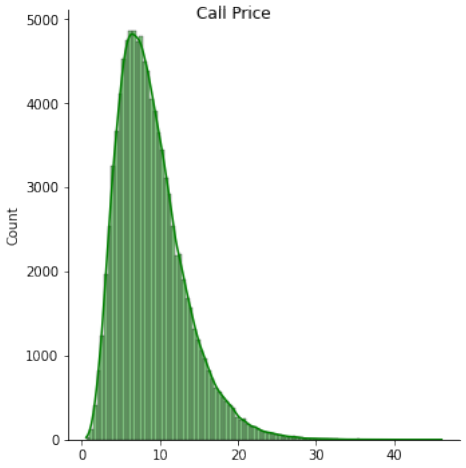

Figure 3: Histogram for call (n = 100,000)

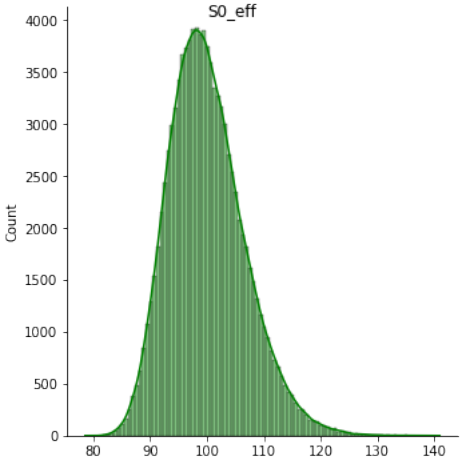

Figure 4: Histogram for S_{eff}

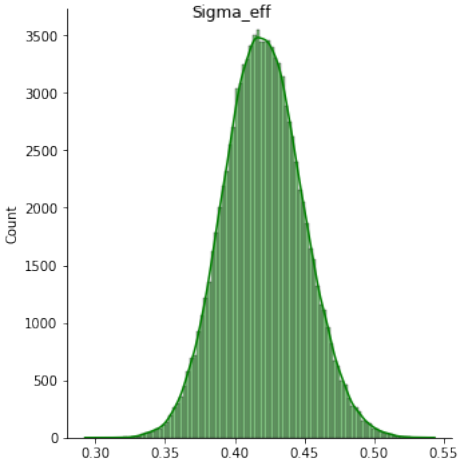

Figure 5: Histogram for σ_{eff}

C_{BS}, which will be our estimate for the call price. The Figure 2 and Figure 3 show the convergence behavior of our Monte Carlo simulation. The simulation on the left repeats 1,000 times whereas it repeats 100,000 times on the right. As suggested by the theory, the standard error is approximately 10 times larger on the left.

$$S_{eff} = S(0)e^{\sum_{t\in I}\frac{1}{2}\rho^2 v(t)\Delta t + \sum_{t\in I}\rho\sqrt{v(t)}\sqrt{\Delta t}\tilde{Z}_2(t)}$$
$$\sigma^2_{eff} = \frac{\sum_{t\in I}(1-\rho^2)v(t)\Delta t}{T}$$

According to our experiment, the decrease in Δ does not exhibit any systematic pattern. However, the theory states that the smaller the Δ, the better the approximation is. Figure 4 and Figure 5 are histograms for S_{eff} and σ_{eff} with repeated times $n = 100,000$.

2.3 Numerical results

To verify our numerical implementation, we compute the call prices with the same set of parameters and show that two results are consistent. Suggested by the current market data, the initial set of parameters is put in the below table.

Param	Value	Notes
$S(0)$	100	Current underlying price
$v(0)$	0.5^2	Current underlying volatility
K	100	Strike Price
λ	5	Mean Reversion Rate
τ	3/12	Time to Maturity
$\overline{v}$	0.35^2	Long term Variance
η	0.3	Vol of Vol
ρ	0.3	Correlation between two Brownian Motion
r	0.014	Risk free rate: 3-Month Labor

Table 1: Intial Parameter List

Our closed-form solution implementation gives 8.9523 whereas Monte Carlo yields 8.9501 with standard error 0.0138. We have also tested other sets of parameters and the corresponding results remain consistent, indicating that our two implementations are both robust and reliable.

We then turn our attention to Greeks, the sensitivity to each market parameter, for the call option.

Greeks	Value		Notes
Delta	0.5429	$\partial V/\partial S$	changes in option value to changes in underlying price
Gamma	0.0180	$\partial^2 V/\partial S^2$	changes in Delta to changes in underlying price
Vega	0.1282	$\partial V/\partial \sigma$	changes in option value to one percent changes in initial Vol
Theta	-0.0582	$\partial V/\partial t$	decline in option value over time
Rho	0.1133	$\partial V/\partial r$	changes in option value to one percent changes in interest rate

Table 2: Value of Greeks for initial parameters

We start with delta which measures the changes in the theoretical option value with respect to changes in the underlying asset's price. More precisely, $\Delta = \partial V/\partial S$. For the Heston model, the delta can be solved analytically to $P^{(1)}$. Our $P^{(1)}$ and delta estimated by finite difference method are both 0.5429. The below Figure 6 shows the relationship between delta (the vertical axis) and the underlying asset's price (the horizontal axis). The S curve is similar to the delta in the Black-Scholes model. As the time to maturity decrese, the curve converges to Heaviside.

<proof $P^{(1)}$ is the delta>

$$V(S_t, K, T) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(S_T - K)^+|\mathcal{F}(t)]$$
$$V(aS_t, aK, T) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(aS_T - aK)^+|\mathcal{F}(t)] = aV(S_t, K, T)$$

Take the derivative with respect to a and equate to the closed-form solution

$$S_t\frac{\partial V}{\partial S_t} + K\frac{\partial V}{\partial K} = V = S_t P^{(1)} + Ke^{-r(T-t)}P^{(0)}$$

By comparison,

$$\frac{\partial V}{\partial S_t} = P^{(1)} = \Delta$$
$$\frac{\partial V}{\partial K} = e^{-r(T-t)}P^{(0)}$$

<end of proof>

Gamma represents the change in Delta relative to the change in price of the underlying, defined as $\partial^2 V/\partial S^2$. This is called second-order (second-derivative) price sensitivity. At-the-money options with less maturity have a greater gamma, and our result behaves like Gamma in Black-Scholes Model. See the Figure 7 for the relationship between gamma (the vertical axis) and the the underlying price (horizontal axis) with different time to maturity.

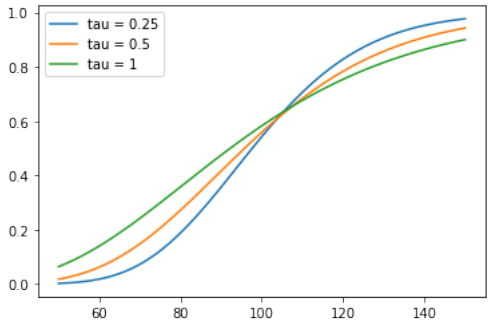

Figure 6: Delta$(S(0), \tau)$

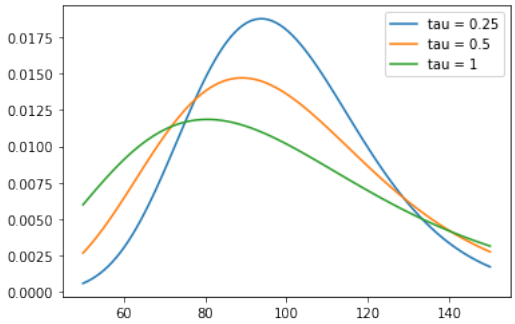

Figure 7: Gamma$(S(0), \tau)$

Vega represents the change in option's value corresponding to one percent change in volatility of the underlying, defined as $\partial V/\partial \sigma$. In our setting, we defined σ as the initial volatility $\sqrt{v(0)}$. That is, the vega here is the option's sensitivity to the initial volatility, indicating the amount an option's price changes given one percent change in the initial volatility. See the Figure 8 for the relationship between vega (the vertical axis) and the the underlying price (horizontal axis) with different time to maturity.

Theta measures decline in the value of an option over time, defined as $\partial V/\partial t$. Theta is generally expressed as a negative number and can be read as the amount by which an option's value declines every day as it moves closer to its maturity. Notice that the theta and gamma is fairly symmetry. At-the-money options with less maturity have a greater magnitude in theta. See the Figure 9 for the relationship between theta (the vertical axis) and the the underlying price (horizontal axis) with different time to maturity.

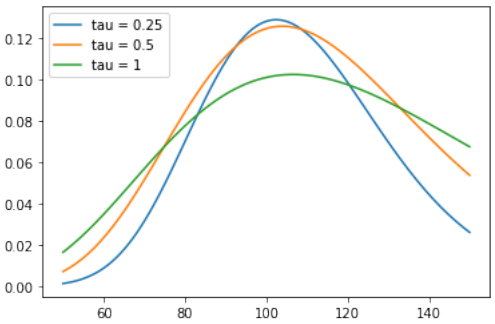

Figure 8: Vega$(S(0), \tau)$

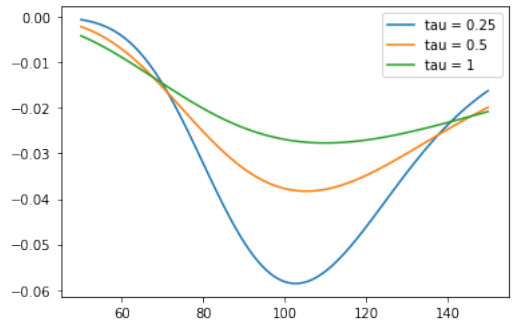

Figure 9: Theta$(S(0), \tau)$

Rho represents the change between an option's value and one percent change in the interest rate, defined as $\partial V/\partial r$. This measures sensitivity to the interest rate. The Figure 13 shows the relationship between rho (the vertical axis) and the underlying asset's price (the horizontal axis) with different time to maturity.

Furthermore, our numerical results converge to Black-Scholes option pricing formula when volatility behaves like a constant. A sufficient condition for $dv(t) = 0\ \forall t$ is to set η close to 0 and eliminate the drift term in volatility, e.g. $\lambda = 0$ or $\overline{v} = v(0)$. Here we change η from 0.5 to 0 and make $\overline{v} = v(0)$. As Figure 11 shows, when η goes to zero, option's prices calculated by

Heston model converge to the Black-Scholes model. After all, Black-Scholes model is a special case of Heston when the volatility is a constant value.

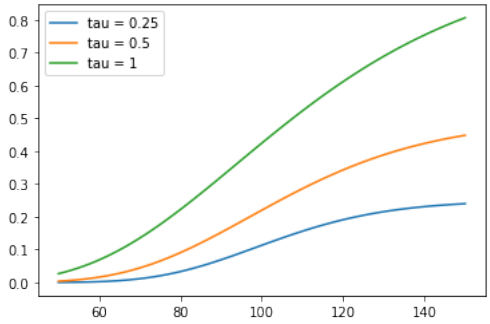

Figure 10: Rho$(S(0), \tau)$

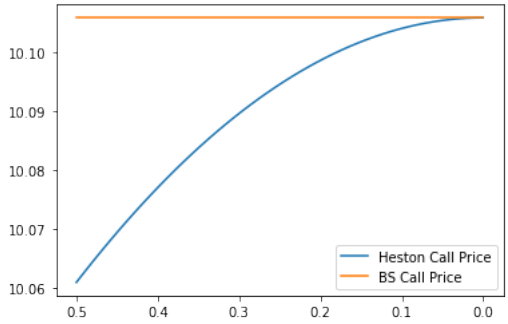

Figure 11: C(η) convergence to BS

3 Data

Our data consists of WTI Call and Put settlement price from 05/02/2022 to 05/04/2022. If market data admits simple arbitrage opportunities, then it would seriously violate the arbitrage-free assumption in our model. Both call and put data need to be examined before calibrating the model. Here we check the three conditions in Carr and Madan (2005)[8] to exclude all static arbitrage opportunities.

3.1 Sufficient condition for no-arbitrage

Carr and Madan (2005) establishes the three sufficient conditions for the existence of a risk-neutral probability measure at each maturity. They denote C_{ij} as the quote for call of strike $K_i, i = 1, ..., \infty$ maturities $T_j, j = 1, ..., M$. Furthermore, they assume interest rates and dividends are zero over the period ending at longest maturity.

Under their settings, the following three conditions suffice to exclude all static arbitrage opportunities: Bull Spread, Butterfly Spread and Calendar Spread conditions.

Bull Spread condition is defined as:

$$BullSpread_{i,j} \equiv \frac{C_{i-1,j} - C_{i,j}}{K_i - K_{i-1}} \geq 0, \forall i, j \geq 0$$

For each $i > 0$, $BullSpread_{i,j}$ is the cost of a bull spread which by definition is long $1/(K_i - K_{i-1})$ calls of strike of K_{i-1} and short $1/(K_i - K_{i-1})$ calls of strike K_i.

Butterfly Spread condition is defined as:

$$ButterflySpread_{i,j} \equiv C_{i-1,j} - \frac{K_{i+1} - K_{i-1}}{K_{i+1} - K_i} C_{i,j} + \frac{K_i - K_{i-1}}{K_{i+1} - K_i} C_{i+1,j} \geq 0, \forall i, j \geq 0$$

For each $i > 0$, $ButterflySpread_{i,j}$ is the cost of a butterfly spread which by definition is long the call struck at K_{i-1}, short $(K_{i+1} - K_{i-1})/(K_{i+1} - K_i)$ calls struck at K_i, and long $(K_i - K_{i-1}/(K_{i+1} - K_i)$ calls struck at K_{i+1}.

Calendar Spread condition is defined as:

$$CalendarSpread_{i,j} \equiv C_{i,j+1} - C_{i,j} \geq 0, \forall i, j \geq 0$$

$CalendarSpread_{i,j}$ consists of long a call of maturity T_{j+1} and short one call of maturity T_j, with both calls stuck at K_i

From the above, we only need to check the vicinity call prices for the three conditions to exclude all static arbitrage opportunities defined in Carr and Madan (2005) [8]. In our dataset, the market prices never violate Bull Spread and Calendar Spread conditions but breach Butterfly condition in some prices. We preprocess the dataset in order to satisfy the Butterfly Spread condition. Our preprocess procedure can be summarized as

$$\min_{C_{i,j}} \sum_{i,j} (C_{i,j} - C_{i,j}^{market})^2$$

$$s.t. \begin{cases} ButterflySpread_{i,j} \geq 0, \forall i, j \\ BullSpread_{i,j} \geq 0, \forall i, j \\ CalendarSpread_{i,j} \geq 0, \forall i, j \end{cases}$$

Since there were just a few points not satisfies the butterfly spread condition, our estimation for $C_{i,j}$ converge quickly and accurately. Newly estimated option prices do not deviate from the market price much, which suggests that WTI options are fairly priced.

4 Calibration

In our Heston model closed-form solution, there are 5 parameters, $\{\lambda, \overline{v}, \eta, \rho, v(0)\}$, that need to be calibrated from the market. However, the optimization process is not easy since the error function is not smooth and convex which leads to many local minima. Sometimes the parameters sticks in an unreasonable local minimum. As a result, it is better to understand the meaning of each parameter such that we can start the optimization process with a reasonable initial guess.

4.1 Parameter sensitivity analysis

First of all, one can see the relationship between η and implied volatility curve from Figure 12. η can be treated as volatility of variance which is similar to the fourth moment or kurtosis. Increasing η results in increasing the thickness of the tail of the distribution. Consequently, price of out-of-money option and the corresponding implied volatility will increase, resulting the curve more convex.

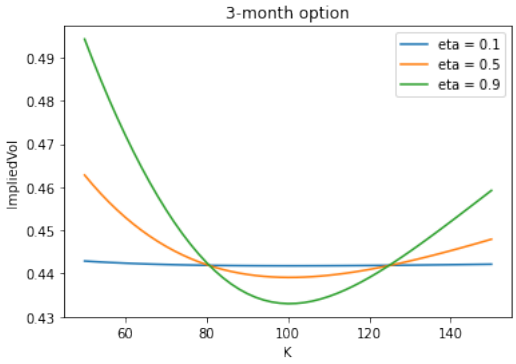

Figure 12: Implied Vol(K, η)

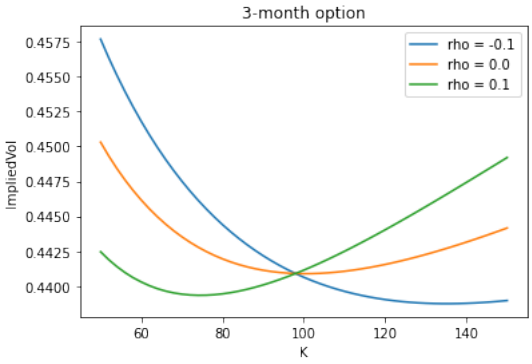

Figure 13: Implied Vol(K, ρ)

Refer to Figure 13, ρ determines the symmetry of the implied volatility curve. The larger magnitude of ρ it is, the more asymmetric the implied volatility curve it is. When ρ is positive, the curve is left-skewed and the minimum attained at lower K whereas when ρ is negative, the curve is right-skewed and the minimum attained at higher K. When $\rho = 0$, the curve is fairly symmetric.

The remaining 3 parameters, $\{\lambda, \overline{v}, v(0)\}$, parallel-shift implied volatility along with the moneyness axis and affect the shape of implied volatility along with the time-to-maturity axis. λ is the mean-reversion rate. If $v(0) > \overline{v}$, then as λ increases, the curve shifts downward as Figure 14. One can see how the degree of convergence to long-term variance changes against λ from

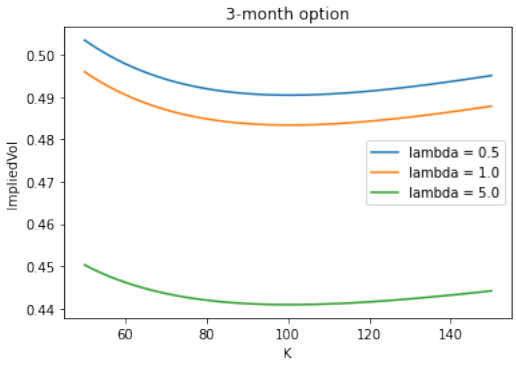

Figure 14: Implied Vol(K, λ)

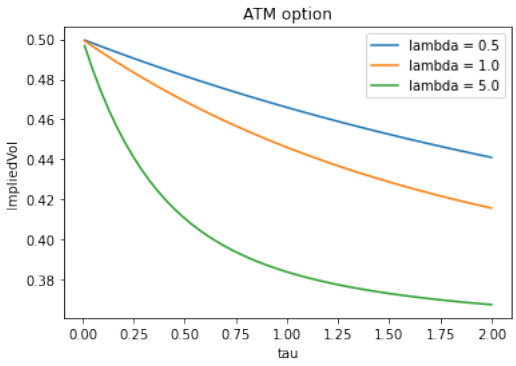

Figure 15: Implied Vol(τ, λ)

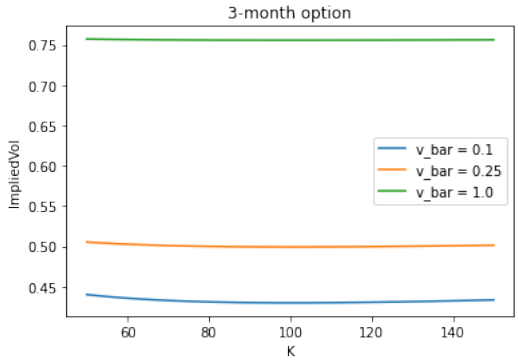

Figure 16: Implied Vol($K, \overline{v}$)

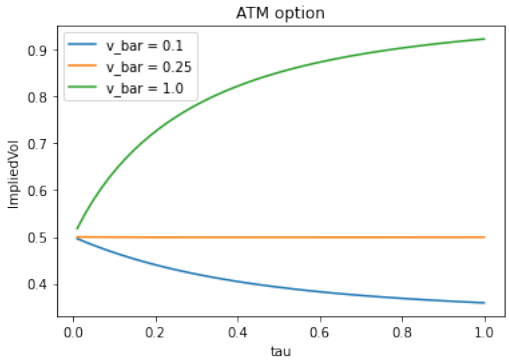

Figure 17: Implied Vol($\tau, \overline{v}$)

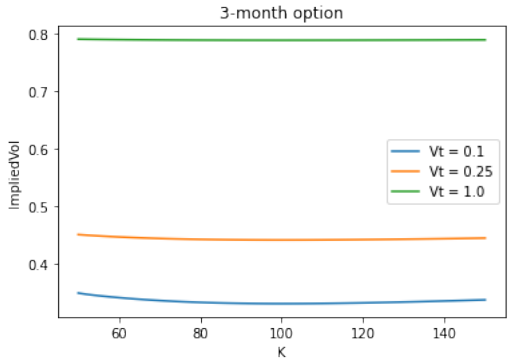

Figure 18: Implied Vol($K, v(0)$)

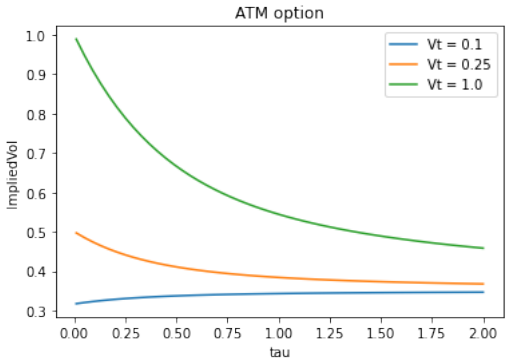

Figure 19: Implied Vol($\tau, v(0)$)

Figure 15. The larger the λ, the faster the convergence rate. $\overline{v}$ represents the long-term variance. It's obvious that as $\overline{v}$ increases, the curve moves up along with the moneyness axis as shown in Figure 16. Refer to Figure 17, the price/implied volatility of a longer time-to-maturity contract is more sensitive to $\overline{v}$. On the other hand, $v(0)$ is the current variance of the underlying asset which gives a large impact on a shorter time-to-maturity contract, see Figure 19. As expected, the higher the $\overline{v}$, the higher the volatility along with the time-to-maturity axis.

4.2 Initial parameters guess

With the help of Bloomberg, we can observe the implied volatility surface easily and have a reasonable initial guess on the parameters through the insights from the previous section. The volatility of the nearest and furthest At-The-Money contract gives a guess on $v(0) \approx 0.53^2$, and $\overline{v} \approx 0.35^2$ respectively. The slight asymmetry of the implied volatility in the sub-graph, Moneyness vs. Volatility, shows a small positive value of ρ. The larger implied volatility of the non-ATM contracts represents a positive value in η. One can also estimate λ by half life = $(\ln 2)/\lambda$. It gives an estimation of λ around 1.3 which is pretty close to our calibration result.

4.3 Calibration Scheme I - all underlying contracts

We performed the calibration with the use of least square error function, with reference to Mikhailov and Nögel (2004)[9]. Define the calibration parameters set $\boldsymbol{\theta} = \{\lambda, \overline{v}, \eta, \rho, v(0)\}$. The error function is composed of two parts, weighted average of squared error and penalty part. First, we assume all At-The-Money options are the most liquid contract and give the most accurate pricing. Thus, a higher weighting is given to all the ATM options. Second, as we mentioned in the previous section, the optimization process is easily stuck at some abnormal

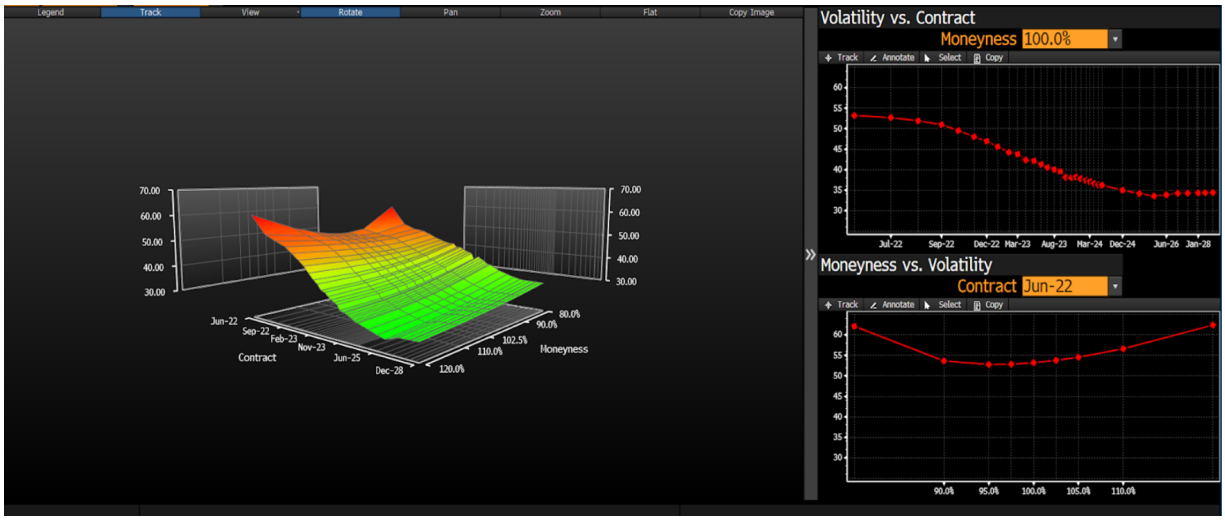

Figure 20: Volatility Surface from Bloomberg Terminal

local minima and the parameters set should not change a lot from the previous day, we imposed a penalty on the deviation of the parameters set:

$$f_t(\boldsymbol{\theta}) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij}(C_t^{market}(X_i,\tau_j) - C_t^{model}(X_i,\tau_j;\boldsymbol{\theta}))^2}{\sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij}} + \boldsymbol{\alpha}||\boldsymbol{\theta} - \boldsymbol{\theta_0}||_2^2$$

where N is number of contracts with different strikes X_i; M is the number of contracts with different time to maturity τ_j; w_{ij} is the weights of different contracts ; $\boldsymbol{\theta_0}$ is the initial guess or the parameters calibrated from previous trading date. We use contracts with different maturity and strikes on a single day to calibrate the model. In this paper, we set $\alpha_\lambda = 0.4, \alpha_{\overline{v}} = 4, \alpha_\eta = 0, \alpha_\rho = 0, \alpha_{v(0)} = 4$. Since we only have an estimation on $\{\lambda, \overline{v}, v(0)\}$ from the market, α is only set on them according to their scale.

4.4 The result for calibration scheme I

Parameters are calibrated using our initial guess or parameters from the previous trading day and then tested the accuracy on the next trading date. The result shows that the pricing is not very robust. Having a penalty will give a more robust model but increasing weight on ATM option does not show a significant difference in result. From the table3, we can see that training without penalty, the parameters optimized are unreasonable, violating our observation from the market.

ATM weight	Train with Penalty	λ	$\overline{v}$	η	ρ	$v(0)$	Mean Training MSE	Mean Prediction MSE
100	TRUE	1.3709	0.0806	0.4701	0.0134	0.2153	0.0117	0.0256
1	TRUE	1.3749	0.0775	0.4616	0.0709	0.2164	0.0088	0.0232
100	FALSE	0.9012	0.0010	0.0433	0.9994	0.2134	0.0062	0.0253

Table 3: Calibration Results for Scheme I

Figures 21 and 22 show the difference between market price and our calibrated model price of contracts with different maturity. Figure21 shows the error of training result while Figure22

shows the error of the prediction result. Comparing the two graphs, the model does not capture the WTI option market structure well. For example, the skewness of the curve changes with time, but in Heston model, these parameters are assumed to be constant. An ideal model should allow the changes in skewness and convexity of the volatility curve at least. However, the Heston model can not accommodate the dynamics we have observed in the market data as we have shown in the parameter sensitivity analysis. According to our experiment, the estimated parameters are not robust. In using the Heston model, the modeler is required to calibrate the model frequently to attain acceptable accuracy, the fact also pointed out by Mikhailov and Nögel (2004) [9]. Our experiment calls for the necessity to extend the model to capture essential structure in the WTI option market. Unfortunately, if we allow the parameters to be functions of time or other variables, then we need to derive the corresponding $C^{(0)}$, $C^{(1)}$, $D^{(0)}$ and $D^{(1)}$. In most cases, we can't find the analytical solution.

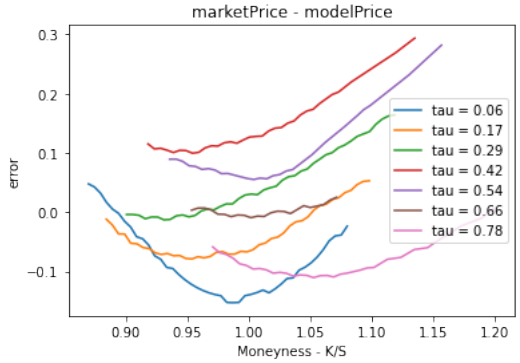

Figure 21: Error on Training

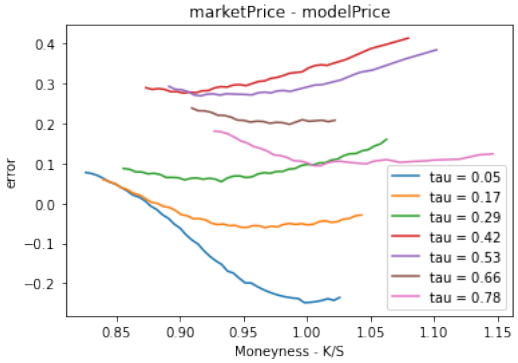

Figure 22: Error on Prediction

See from Figure 21 and 22, we can see the difference between market price and model price have general positive relationship with moneyness on training set, which means that weighting of the in-the-money and out-of-money call option have a huge difference due to the scale of their price. Therefore, we tried to remove the scaling effects and use implied volatility(IV) in Calibration Scheme II.

4.5 Calibration scheme II - one underlying contract

In this section, we made another assumption that each contract with different maturity has different behavior. As the result, we conduct calibration on single underlying contracts, Aug 2022 WTI options, prices with different strikes, i.e. calibrate on options whose underlying contracts are the same. To put it, we now consider WTI contracts with different maturity are independent of each other.

The calibration scheme is different from the previous one with the use of least square error function for implied volatility(IV). Define parameter $\boldsymbol{\theta} = \{\lambda, \overline{v}, \eta, \rho, v(0)\}$. The error function in this section consists of only weighted average of squared error term but no penalty term. Our study finds the result for Scheme II is radically different from the one for Scheme I. The penalty term in Scheme I is to prevent optimization from coming up with unrealistic parameter. In Scheme II, the penalty structure in the error function is not useful as the estimated parameters are within the interior of reasonable parameter set. We are looking for the parameter that minimizes the new error function:

$$f_{jt}(\boldsymbol{\theta}) = \frac{\sum_{i=1}^{N} w_{it}(\sigma_t^{market}(X_i, \tau_j) - \sigma_t^{model}(X_i, \tau_j; \boldsymbol{\theta}))^2}{\sum_{i=1}^{N} w_{it}}$$

where θ_0 is the initial guess or the parameters calibrated from previous trading date. Here we are indifferent with the initial parameters since we remove the penalty term.

Figure 23 and 24 show the errors on implied volatility by price error function(Scheme I) and IV error function (Scheme II). Clearly the sample is better fitted using the IV error function. With IV error function, we also remove ITM and OTM Option price scale discrepancy.

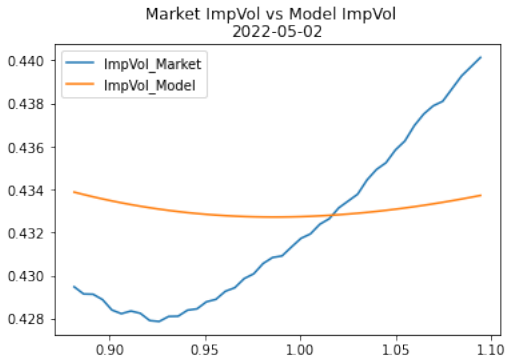

Figure 23: Implied Volatility from price error function

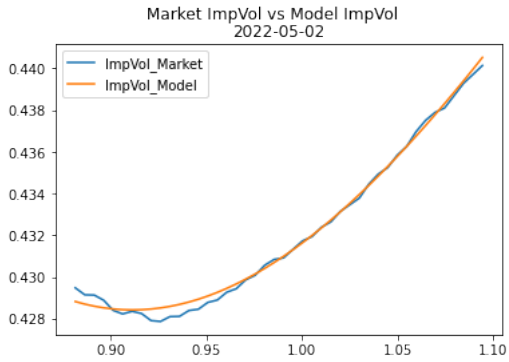

Figure 24: Implied Volatility from IV error function

4.6 The result for calibration scheme II

Table 4 shows the final results of our calibration. All five parameters not change significantly during 2022/05/02 to 2022/05/05, λ is rather large since we chose a shorter life contract. The parameter calibrated each day is very stable. Under this calibration scheme, we can capture the Implied Volatility structure of this contract. However, since we calibrate the contracts separately on each trading date, it is unable to capture the market movement of volatility structure from day to day, i.e. limited dynamic prediction power. One can see the details of the calibration result in the Appendix A.

	Optimized Parameters					Implied Vol error function	
Date	λ	$\overline{v}$	η	ρ	v_0	Mean Prediction RMSE	Std Prediction RMSE
2022-05-02	6.4204	0.2024	1.6120	0.1434	0.2011	/	/
2022-05-03	6.4189	0.2002	1.5062	0.0946	0.2061		
2022-05-04	6.3548	0.2173	1.5598	0.0766	0.2005	0.009372	0.001424
2022-05-05	6.1991	0.2155	1.5838	0.0793	0.2115		
2022-05-06	/	/	/	/	/		

Table 4: Calibration Results for scheme II

5 Conclusion

Heston model (1993)[1] extends the Black and Scholes (1973)][4] model by allowing the changes in volatility and the correlation between the underlying price and volatility. Our paper re-derives the call option analytical solution to the Heston model and implement the closed-form solution. Furthermore, this paper built a Monte Carlo simulation utilizing the Mixing Theorem, which significantly lowers the running-time complexity of the simulation. The two numerical implementations in this paper are consistent. Greeks are also computed and exhibit a similar patterns as Black and Scholes model do, especially when the volatility behave like a constant value.

The second half of the paper mainly concerns about model calibration. There are five parameters required to estimate. This paper first examines the changes in implied volatility surface with respect to each parameter. For example, the correlation governs the symmetry and volatility of volatility affects the convexity. One can find a reasonable initial value simply by inspecting the implied volatility surface. For this paper, some of the market data violates the Butterfly Spread condition, so the data needs to be preprocessed before calibration. We look for the adjusted prices that is closest to the market prices in L^2 norm subject to the three conditions in Carr and Madan (2005)[8]. This papar calibrates the model using the mean squared error function with higher weights on at-the-money call prices. Our study does not show the weights give the model robustness. Mikhailov and Nögel (2004)[9] suggests to put penalty on the deviation from the initial parameters. We find their method make the model more robust. However, the Heston model seems to be insufficient to capture the dynamics of the WTI option market structure. The skewness and the convexity of the volatility curve changes over time, but the Heston model treats the relevant parameters, correlation and volatility of volatility, constant.

From the success implementation of Scheme II, WTI contracts should be consider different things despite their similarity. When calibrating the Heston model, one should not mix all the derivatives of WTI contracts with different maturities together as we did in Scheme I whose result is unsatisfactory. With the use of Implied Volatility (IV) error function, the calibration scheme can capture the volatility structure very well but it still provides limited help on the dynamics of the market structure.

References

[1] S. L. Heston, "The volatility surface: A practitioner's guide," *The Review of Financial Studies*, vol. 6, no. 2, pp. 327–343, 1993.

[2] J. Gatheral, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," 2012.

[3] D. Duffie, J. Pan, and K. Singleton, "Transform analysis and asset pricing for affine jump-diffusions," *Econometrica*, vol. 68, no. 6, pp. 1343–1376, 2000.

[4] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, no. 3, pp. 637–654, 1973.

[5] G. Bakshi and D. Madan, "Spanning and derivative-security valuation," *Journal of Financial Economics*, vol. 55, no. 2, pp. 205–238, 2000.

[6] N. G. Shephard, "From characteristic function to distribution function: A simple framework for the theory," *Econometric Theory*, vol. 7, no. 4, pp. 519–529, 1991.

[7] C. Kahl and P. Jackel, "Not-so-complex logarithms in the heston model," *Wilmott Magazine*, vol. 19, no. 9, pp. 94–103, 2005.

[8] P. Carr and D. B. Madan, "A note on sufficient conditions for no arbitrage," *Finance Research Letters*, vol. 2, pp. 125–130, 2005.

[9] S. Mikhailov and U. Nögel, "Heston's stochastic volatility model implementation, calibration and some extensions," *Wilmott Magazine*, pp. 74–79, 2004.

A Appendix

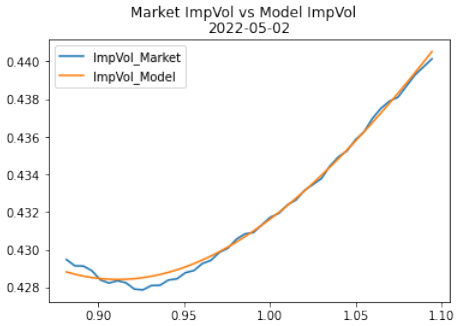

Figure 25: Training set on 2022/05/02

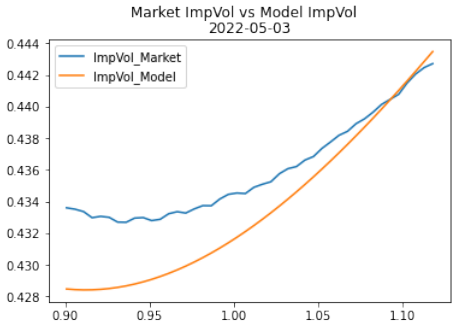

Figure 26: Testing set on 2022/05/03

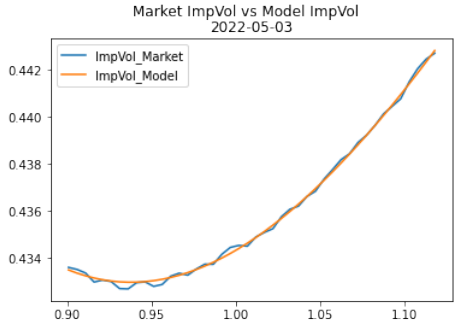

Figure 27: Training set on 2022/05/03

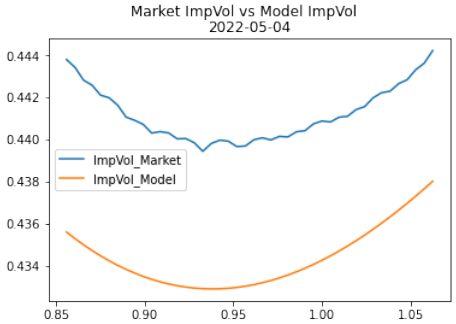

Figure 28: Testing set on 2022/05/04

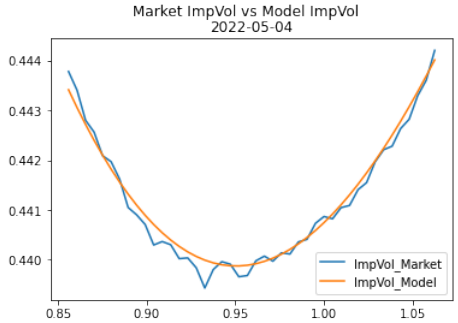

Figure 29: Training set on 2022/05/04

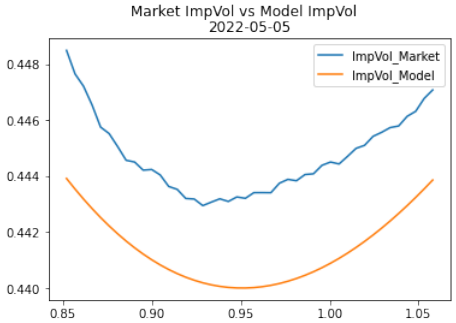

Figure 30: Testing set on 2022/05/05

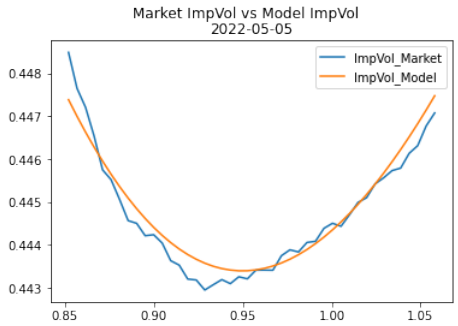

Figure 31: Training set on 2022/05/05

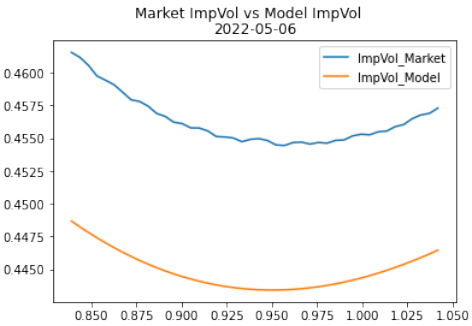

Figure 32: Testing set on 2022/05/06