Derivatives in Financial Markets with Stochastic Volatility — Notes and Extensions on Chapter 1 & 2

April 17, 2020

Contents

1 The Black-Scholes Theory

The aim of this chapter is to review the basic objects, ideas, and results of the classical Black-Scholes theory of derivative pricing. There are lots of methods for BS's theory derivation:

1. derive BS-PDE via replicating strategies and self-financing property
2. derive BS-PDE via hedging
3. obtain pricing formula from BS-PDE by PDE methods, like Feynman-Kac formula or Fourier transform.
4. approximate price by binomial tree or Monte Carlo simulation
5. risk-neutral valuation via martingale theory (finding ELMM)

1.1 Brownian Motion

Following Samuelson, we assume that there are two assets in the market. One is a risk-free asset with price β_t,

$$d\beta_t = r\beta_t dt, \tag{1.1}$$

where $r \geq 0$ is the instantaneous interest rate for lending or borrowing money. Setting $\beta_0 = 1$, we have $\beta_t = e^{rt}$ for $t \geq 0$. The other is the risky asset with price X_t,

$$dX_t = \mu X_t dt + \sigma X_t dW_t, \tag{1.2}$$

where μ is a constant mean return rate (drift rate), $\sigma > 0$ is a constant volatility, and $(W_t)_{t\geq 0}$ is a standard Brownian motion (Wiener process).

Definition 1.1. Brownian Motion is a real-valued stochastic process with continuous trajectories $t \to W_t$ that have independent and stationary increments, i.e.

- $W_0 = 0$;
- for any $0 < t_1 < \cdots < t_n$, the random variables $\left(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}\right)$ are independent;
- for any $0 \leq s < t$, the increment $W_t - W_s \sim N(0, t-s)$. Especially, $W_t \sim N(0, t)$.

Definition 1.2. The probability space is denoted by $(\Omega, \mathcal{F}, P)$, P is the Wiener measure.

Definition 1.3. The increasing family of σ-algebras $\mathcal{F}_t$ generated by $(W_s)_{s\leq t}$ are called the information on W up to time t.

Definition 1.4. All the sets of probability 0 in $\mathcal{F}_t$ is called *the natural filtration* of the BM.

Definition 1.5. $(X_t)_{t\geq 0}$ is *adapted* to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if X_t is $\mathcal{F}_t$-measurable for every t, and X_t is said to be *$\mathcal{F}_t$-adapted.*

Remark 1.1. We can summarize definition 1.1 via *conditional characteristic functions.* For $0 \leq s < t$ and $u \in \mathbb{R}$,

$$E\left[e^{iu(W_t - W_s)} | \mathcal{F}_s\right] = e^{-u^2(t-s)/2}. \tag{1.3}$$

Remark 1.2. The drawback of BM is that the trajectories are not of bounded variation, as followed:

$$E\left[\sum_{i=1}^{n} |W_{t_i} - W_{t_{i-1}}|\right] = nE\left[|W_{\frac{t}{n}}|\right] = n\sqrt{\frac{t}{n}}E\left[|W_1|\right]. \tag{1.4}$$

1.2 Stochastic Integral

Definition 1.6. For T a fixed finite time, let $(X_t)_{0\leq t\leq T}$ be a stochastic process adapted to $(\mathcal{F}_t)_{0\leq t\leq T}$, the filtration of the BM up to time T, such that

$$E\left[\int_0^T (X_t)^2 dt\right] < +\infty \tag{1.5}$$

The *stochastic integral* of X_t with respect to BM is defined as a limit on $L^2(\Omega)$,

$$\int_0^t X_s dW_s = \lim_{n\to+\infty} \sum_{i=1}^n X_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}), \tag{1.6}$$

as the mesh size of the subdivision goes to zero.

Remark 1.3. By iterated conditional expectations and the independent increments property of BM, we note that

$$E\left[\left(\sum_{i=1}^n X_{t_{i-1}}(W_{t_i} - W_{t_{i-1}})\right)^2\right] = E\left[\sum_{i=1}^n (X_{t_{i-1}})^2 (t_i - t_{i-1})\right] \tag{1.7}$$

for $t \leq T$. So, we have that

$$E\left[\left(\int_0^t X_s dW_s\right)^2\right] = E\left[\int_0^t X_s^2 ds\right] < +\infty \tag{1.8}$$

Remark 1.4. Also, stochastic integral in (1.6) has the *martingale property*:

$$E\left[\int_0^t X_u dW_u | \mathcal{F}_s\right] = \int_0^s X_u dW_u \quad (a.s. \quad s \leq t). \tag{1.9}$$

Definition 1.7. The quadratic variation $\langle Y \rangle_t$ of the stochastic integral $Y_t = \int_0^t X_u dW_u$ is

$$\langle Y \rangle_t = \lim_{n\to+\infty} \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 = \int_0^t X_s^2 ds. \tag{1.10}$$

Theorem 1.1. *Y_t is a mean-zero, continuous and square integral martingale. The converse is also true: every mean-zero, continuous, and square integrable martingale is a Brownian stochastic integral.*

1.3 Itô's Formula

Based on Taylor's formula and the fact that W_t is not differentiable, we can use

Theorem 1.2. *(Itô's Lemma)*

$$dg(W_t) = g'(W_t)dW_t + \frac{1}{2} g''(W_t)dt \tag{1.11}$$

Theorem 1.3. *(Itô's Formula) For $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$,*

$$\begin{aligned} dg(t, X_t) &= \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} d\langle X \rangle_t \\ &= \left(\frac{\partial g}{\partial t} + \mu(t, X_t)\frac{\partial g}{\partial x} + \frac{1}{2}\sigma^2(t, X_t)\frac{\partial^2 g}{\partial x^2}\right) dt + \sigma(t, X_t)\frac{\partial g}{\partial x} dW_t, \end{aligned} \tag{1.12}$$

where all the partial derivatives of g are evaluated at (t, X_t).

Corollary 1.3.1.

$$d(X_tY_t) = X_tdY_t + Y_tdX_t + d\langle X, Y\rangle_t, \tag{1.13}$$

where the co-variation (aka. "bracket") of X and Y is given by

$$d\langle X, Y\rangle_t = \sigma_X(t, X_t)\sigma_Y(t, Y_t)dt. \tag{1.14}$$

Remark 1.5. $\langle X, X\rangle_t = \langle X\rangle_t$.

Remark 1.6. By guessing and testing, we find the solution of (1.2) is GBM, and the return X_t/X_0 is *log-normal*, i.e.

$$X_t = X_0 \exp\left[(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\right]. \tag{1.15}$$

Remark 1.7. Notice that bankruptcy (zero stock price) in this model is a permanent state, since it stays at zero for all time thereafter if X_t becomes zero.

Derivative Contracts

- Derivatives are contracts based on the underlying asset price. They are also called *contingent claims*.
- strike price K, *maturity* date T
- *Payoff* $h(X_T)$ is the non-negative value of this contract at maturity.
- *Derivative price* $P(t, x)$ is the value of this contract at time t for a underlying asset price $X_t = x$.
- Exercise time τ is a *stopping time* wrt the filtration $(\mathcal{F}_t)$, i.e. random time s.t. the event $\tau \le t$ belongs to $\mathcal{F}_t$ for any $t \le T$.
- "Exotic option" refers here to not a standard European or American option.

1.4 Replicating Strategies

Definition 1.8. (replicating) This is a way for traders to find the value of an asset by replicating its cash flows or price movement using other assets whose values they already know.

Considering a European-style derivative with payoff $h(X_T)$, we replicate the derivative at maturity:

$$a_TX_T + b_Te^{rT} = h(X_T) \quad a.s. \tag{1.16}$$

This portfolio is to be *self-financing*, i.e. no further funds are required after the initial investment, so

$$d(a_tX_t + b_te^{rt}) = a_tdX_t + rb_te^{rt}dt, \tag{1.17}$$

which implies that

$$X_tda_t + e^{rt}db_t + d\langle a, X\rangle_t = 0, \tag{1.18}$$

and integral form is

$$a_tX_t + b_te^{rt} = a_0X_0 + b_0 + \int_0^t a_sdX_s + \int_0^t rb_se^{rs}ds, \quad 0 \le t \le T. \tag{1.19}$$

Now we assume a pricing function $P(t,x)$ in Section 1.3 exists and is regular enough to apply Itô's formula (1.12). *No-arbitrage* assumption implies that

$$a_t X_t + b_t e^{rt} = P(t, X_t) \qquad \forall\, 0 \le t \le T. \tag{1.20}$$

Differentiating (1.20) and using self-financing property (1.17), Itô's formula (1.12) and (1.2), we obtain

$$(a_t \mu X_t + b_t r e^{rt})dt + a_t \sigma X_t dW_t = \left(\frac{\partial P}{\partial t} + \mu X_t \frac{\partial P}{\partial x} + \frac{1}{2}\sigma^2 X_t^2 \frac{\partial^2 P}{\partial x^2} \right) dt + \sigma X_t \frac{\partial P}{\partial x} dW_t, \tag{1.21}$$

where all the partial derivatives of P are evaluated at (t, X_t). Equating the coefficients of the dW_t terms gives

$$a_t = \frac{\partial P}{\partial x}(t, X_t). \tag{1.22}$$

From (1.20) we obtain

$$b_t = (P(t, X_t) - a_t X_t)e^{-rt}. \tag{1.23}$$

Equating the dt terms in (1.21) gives

$$r\left(P - X_t \frac{\partial P}{\partial x} \right) = \frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 X_t^2 \frac{\partial^2 P}{\partial x^2}, \tag{1.24}$$

Definition 1.9. (Black-Scholes partial differential equation)

$$\mathcal{L}_{BS}(\sigma)P = 0 \tag{1.25}$$

where

$$\mathcal{L}_{BS} = \frac{\partial}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + r\left(x\frac{\partial}{\partial x} - \cdot \right) \tag{1.26}$$

holds in the domain $t \le T$ and $x > 0$.

Remark 1.8. It is to be solved via the final condition $P(T, x) = h(x)$.

Remark 1.9. The rate of return μ does not enter at all in the valuation of this portfolio.

1.5 Hedging

There is another way to derive the Black-Scholes partial differential equation that emphasizes risk elimination, aka. **hedging**.

Let $P_t = P(t, X_t)$ be the price of the option. If we sell N_t options and hold A_t risky asset, then the change in the value of this portfolio is $A_t dX_t - N_t dP_t$ due to being self-financing. This portfolio should be risk-free, so the coefficient of dW_t should be zero, i.e.

$$A_t dX_t - N_t dP_t = r(A_t X_t - N_t P_t)dt. \tag{1.27}$$

Using Itô's formula (1.12) and (1.2), we have

$$\begin{aligned} A_t(\mu X_t dt + \sigma X_t dW_t) - N_t \left\{ \left(\frac{\partial P}{\partial t} + \mu X_t \frac{\partial P}{\partial x} + \frac{1}{2}\sigma^2 X_t^2 \frac{\partial^2 P}{\partial x^2} \right) dt - \sigma X_t \frac{\partial P}{\partial x} dW_t \right\} \\ = r(A_t X_t - N_t P_t)dt. \end{aligned} \tag{1.28}$$

Eliminating the dW_t terms gives

$$A_t = N_t \frac{\partial P}{\partial x}(t, X_t). \tag{1.29}$$

Remark 1.10. The derivation is known as *Delta Hedging*: selling the option and holding a dynamically adjusted amount of the risky asset.

1.6 Black-Scholes Formula

BS-PDE (1.25) is solved with the final condition $h(x) = (x - K)^+$. Let $C_{BS}(t, x)$ be the price of European call option.

Theorem 1.4. *(BS formula for call options)*

$$C_{BS}(t, x) = xN(d_1) - Ke^{-r(T-t)}N(d_2), \tag{1.30}$$

where

$$\begin{aligned} d_1 &= \frac{\log(\frac{x}{K}) + (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}, \\ d_2 &= d_1 - \sigma\sqrt{T - t}, \end{aligned} \tag{1.31}$$

and

$$N(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{y^2}{2}} dy. \tag{1.32}$$

Let $P_{BS}(t, x)$ be the price of a European put option. There is a *model-free* relationship that follows from no-arbitrage arguments at time T.

Theorem 1.5. *(put-call parity)*

$$C_{BS}(t, X_t) - P_{BS}(t, X_t) = X_t - Ke^{-r(T-t)}, \tag{1.33}$$

between put and call options with the same maturity and strike price.

Theorem 1.6. *(BS formula for put option)*

$$P_{BS}(t, x) = Ke^{-r(T-t)}N(-d_2) - xN(-d_1), \tag{1.34}$$

where d_1, d_2, N are as in (1.31) and (1.32), respectively.

Remark 1.11. Through plotting, we find that the pricing function of put option crosses over its terminal payoff for some (small enough) x, which does not happen with the call option function. This fact is important for pricing American options.

1.7 Risk-Neutral Valuation

In the BS framework, there is a *unique* probability measure P^* equivalent to P s.t. under this probability measure:

1. the discounted price $\widetilde{X}_t = e^{-rt}X_t$ is a martingale;
2. the expected value under P^* of the discounted payoff of a derivative gives its no-arbitrage price.

Such a probability measure describing a *risk-neutral* world is called an *equivalent martingale measure* or risk-neutral measure. To find unique EMM of the discounted price, we use

$$d\widetilde{X}_t = \sigma\widetilde{X}_t \left[dW_t + \left(\frac{\mu - r}{\sigma} \right) dt \right]. \tag{1.35}$$

We set

$$\theta = \frac{\mu - r}{\sigma}, \tag{1.36}$$

called the *market price of asset risk*, and define

$$W_t^* = W_t + \int_0^t \theta ds = W_t + \theta t, \tag{1.37}$$

so that

$$d\widetilde{X}_t = \sigma \widetilde{X}_t dW_t^*. \tag{1.38}$$

We try to construct EMM P^*, an equivalent measure to P, meaning that it has the same null sets. According to **Radon-Nikodym theorem**, we can change the measure. Let P^* has the density ξ_T^θ wrt P:

$$dP^* = \xi_T^\theta dP \tag{1.39}$$

where

$$\xi_T^\theta = \exp(-\theta W_T - \frac{1}{2}\theta^2 T). \tag{1.40}$$

Remark 1.12. $E[\xi_T^\theta] = 1$

Remark 1.13. The *Radon-Nikodym* process $(\xi_t^\theta)_{0\le t\le T}$ is a martingale wrt $\mathcal{F}_t$.

Remark 1.14. For any integrable random variable Z, $E^*[Z] = E[\xi_T^\theta Z]$.

Remark 1.15. For any adapted and integrable process (Z_t),[1]

$$E^*[Z_t|\mathcal{F}_t] = \frac{1}{\xi_s^\theta} E[\xi_t^\theta Z_t|\mathcal{F}_s]. \tag{1.41}$$

Remark 1.16. Under the probability P^*, the process (W_t^*) is a standard BM. This result in its full generality, when θ is an adapted stochastic process, is known as **Girsanov's theorem**. [2] In our case (θ is constant), it is easily derived by using the characterization (1.3) and formula (1.41).

Definition 1.10. (Self-financing) Let V_t be the portfolio value at time t

$$V_t = a_t X_t + b_t e^{rt}. \tag{1.42}$$

Then the self-financing means that $\widetilde{V}_t = e^{-rt}V_t$ is a martingale under EMM P^*, i.e.

$$d\widetilde{V}_t = a_t d\widetilde{X}_t = \sigma a_t \widetilde{X}_t dW_t^*. \tag{1.43}$$

Remark 1.17. Due to being self-financing,

$$V_t = E^*[e^{-r(T-t)}H|\mathcal{F}_t] \tag{1.44}$$

where $H = h(X_T)$.

The existence of self-financing portfolio is guaranteed by

Theorem 1.7. *(Martingale Representation Theorem) For* $0 \le t \le T$*,*

$$M_t = E^*[e^{-rT}H|\mathcal{F}_t] \tag{1.45}$$

defines a square integrable martingale under P^* *wrt* $(\mathcal{F}_t)$*, which is also the natural filtration of the BM* W^**.*

[1] The details can be seen in Section 2.3 of coursenotes2003.pdf.

[2] The details and proof can be seen in Section 4 of coursenotes2003.pdf.

Definition 1.11. (Markov Property) X_t is Markovian if conditioning wrt $\mathcal{F}_t$ is the same as conditioning wrt X_t, i.e. $E[\cdot|\mathcal{F}_t] = E[\cdot|X_t]$.

Remark 1.18. Through this method, we can obtain the same pricing formula, i.e.

$$\begin{aligned} P(t,x) &= E^*[e^{-r(T-t)}h(X_T)|X_t = x] \\ &= \frac{1}{\sqrt{2\pi(T-t)}} \int_{-\infty}^{+\infty} e^{-r(T-t)} h(xe^{(r-\frac{\sigma^2}{2})(T-t)+\sigma z}) e^{-\frac{z^2}{2(T-t)}} dz. \end{aligned} \tag{1.46}$$

American Options

Via the theory of *optimal stopping*, it can be shown that the price of an American derivative with payoff function h is obtained by maximizing the expected value of the discounted payoff. In other words,

$$P(t,x) = \sup_{t \le \tau \le T} E^*[e^{-r(T-t)} h(X_\tau^{t,x})], \tag{1.47}$$

where $(X_\tau^{t,x})_{\tau \ge t}$ is the stock price starting at $X_t = x$.

In order to determine the optimal stopping time τ^*, this leads to a *free boundary value problem.* According to a no-arbitrage argument, we have

Theorem 1.8. *For non-negative interest rates and no dividend paid, the price of an American call option is the same as its corresponding European option.*

Remark 1.19. The price of an American put option is in general strictly higher than the price of the corresponding European put option.

Complete Market

In the previous sections we assumed the existence, uniqueness and regularity of the solution of the *parabolic partial differential equation (aka. heat equation)*

$$\frac{\partial u}{\partial t} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 u}{\partial x^2} + \mu(t,x)\frac{\partial u}{\partial x} - ru = 0 \tag{1.48}$$

with the final condition $u(T,x) = h(x)$ in order to apply Itô's formula. A sufficient condition for this is that the coefficients μ and σ are regular enough and that there exists a positive constant A s.t.

$$\sigma^2(t,x) \ge A \ge 0 \quad \forall\, t \ge 0, x \in \mathcal{D} \tag{1.49}$$

where $\mathcal{D}$ is the domain of the process (X_t).

In addition, the model we have analyzed here is an example of a *complete* market model.

Definition 1.12. Complete market is one in which every contingent claim can be replicated by a self-financing trading strategy in the stock and bond.

When looking at stochastic volatility market models, we shall see that the market is *incomplete.* There is a whole family of EMMs, and derivatives securities cannot be perfectly hedged with just the stock and bond.

2 Stochastic Volatility Models

The BS model needs a number of assumptions that are, to some extent, "counterfactual":

1. continuity of the stock-price process (it does not jump),
2. the ability to hedge continuously without transaction costs,
3. independent Gaussian returns,
4. constant volatility.

We shall focus here on relaxing the last assumption by allowing volatility to vary randomly, for the following reason: a well-known discrepancy between BS-predicted European option prices and market-traded options prices, *the smile curve*, can be accounted for by SV models.

2.1 Implied Volatility and the Smile Curve

Definition 2.1. (Implied Volatility) Given an European call option with observed price C^{obs}, strike price K and expiration date T. The implied volatility I is the value of the volatility parameter that must go into BS formula (1.30) to match,

$$C_{BS}(t,x;K,T;I) = C^{obs}. \tag{2.1}$$

Remark 2.1. A unique non-negative IV $I > 0$ can be found if $C^{obs} > C_{BS}(t,x;K,T;0)$ since

$$\frac{\partial C_{BS}}{\partial \sigma} = \frac{xe^{-\frac{d_1^2}{2}}\sqrt{T-t}}{\sqrt{2\pi}} > 0. \tag{2.2}$$

Remark 2.2. The implied volatilities from put and call options of the same K, T are the same due to put-call parity (1.33).

Definition 2.2. (Smile Effect) Implied volatilities of the market prices vary with strike price and the time to maturity of the contract.

Before the 1987 crash, the graph of $I(K)$ for fixed t, x, T from market options prices was often observed to be the ∪-shaped, called the smile curve or *smirk*. Since 1987, the curve is more typically downward sloping at and near the money ($95\% \leq K/x \leq 105\%$) and then curves upwards for far out-of-the-money strikes ($K \gg x$). Other qualitative features of IV from stock options are that it is higher than historical volatility and is often decreasing with time to maturity.

Remark 2.3. Smirk tells us that there is a premium charged for out-of-the-money put options and in-the-money calls above their BS price computed with the at-the-money IV.

Bounds of $\partial I/\partial K$ can be obtained by the BS formula (1.30). Since no-arbitrage assumption, call prices must be decreasing with K. Thus, differentiating (2.1) wrt K gives

$$\begin{aligned} \frac{\partial C^{obs}}{\partial K} &= \frac{\partial C_{BS}}{\partial K} + \frac{\partial C_{BS}}{\partial \sigma}\frac{\partial I}{\partial K} \leq 0, \\ \Rightarrow \frac{\partial I}{\partial K} &\leq -\frac{\partial C_{BS}/\partial K}{\partial C_{BS}/\partial \sigma}. \end{aligned} \tag{2.3}$$

Similarly for put options

$$\begin{aligned} \frac{\partial P^{obs}}{\partial K} &= \frac{\partial P_{BS}}{\partial K} + \frac{\partial P_{BS}}{\partial \sigma}\frac{\partial I}{\partial K} \geq 0, \\ \Rightarrow \frac{\partial I}{\partial K} &\geq -\frac{\partial P_{BS}/\partial K}{\partial P_{BS}/\partial \sigma}. \end{aligned} \tag{2.4}$$

2.2 Implied Deterministic Volatility

One popular way to modify the lognormal model is to suppose that volatility is a deterministic positive function of time and stock price: $\sigma = \sigma(t, X_t)$.

In the special case $\sigma(t,x) = \sigma(t)$, we can solve the SDE

$$dX_t = rX_t dt + \sigma(t) X_t dW_t^* \tag{2.5}$$

under EMM P^* by the logarithmic transformation to obtain

$$X_T = X_t \exp\left(r(T-t) - \frac{1}{2}\int_t^T \sigma^2(s)ds + \int_t^T \sigma(s)dW_s^* \right), \tag{2.6}$$

so that $\log(X_T/X_t) \sim N\left((r - \frac{1}{2}\bar{\sigma^2})(T-t)), (\bar{\sigma^2}(T-t)\right)$, where

$$\bar{\sigma^2} = \frac{1}{T-t}\int_t^T \sigma^2(s)ds. \tag{2.7}$$

Thus, the answer is still the BS formula with volatility parameter $\sqrt{\bar{\sigma^2}}$, the root-mean-square volatility. In this time-dependent case, there is no smile across strike prices. To have a smile across strike prices, we need σ to depend on x and t in this framework, which is the case of level-dependent volatility.

There are many competing ways — parametric or non-parametric — to estimate the volatility surface $\sigma(t,x)$ from traded option prices. This is called "finding the implied deterministic volatility." It has the advantage of preserving a complete market model.

2.3 Modelling Stochastic Volatility

In "pure" SV models, the asset price $(X_t)_{t\geq 0}$ satisfies the SDE

$$dX_t = \mu X_t dt + \sigma_t X_t dW_t, \tag{2.8}$$

where $(\sigma_t)_{t\geq 0}$ is called the volatility process. It must satisfy some regularity conditions for the model to be well-defined, but *it does not have to be an Itô process: it can be a jump process, a Markov chain, etc.*

Definition 2.3. (Mean Reversion) Let us denote

$$\sigma_t = f(Y_t), \tag{2.9}$$

where f is some positive function. Then mean-reverting stochastic volatility means that the SDE for (Y_t) looks like

$$dY_t = \alpha(m - Y_t)dt + \cdots d\hat{Z}_t, \tag{2.10}$$

where $(\hat{Z}_t)_{t\geq 0}$ is a BM correlated with (W_t). Here α is called the rate of mean-reversion and m is the long-run mean level of Y_t.

Definition 2.4. (Ornstein-Uhlenbeck process) A stochastic process satisfies the SDE

$$dY_t = \alpha(m - Y_t)dt + \beta d\hat{Z}_t. \tag{2.11}$$

Remark 2.4. OU process is a Gaussian process explicitly given in terms of its starting value y by

$$Y_t = m + (y - m)e^{-\alpha t} + \beta \int_0^t e^{-\alpha(t-s)} d\hat{Z}_s, \tag{2.12}$$

so that $Y_t \sim N(m + (y - m)e^{-\alpha t}, \frac{\beta^2}{2\alpha}(1 - e^{-2\alpha t}))$. Its *invariant distribution*, obtained as $t \to \infty$, is $N(m, \frac{\beta}{2\alpha})$, which does not depend on y.

Remark 2.5. The second BM $(\hat{Z}_t)$ is typically correlated with the BM (W_t) driving (2.8). We denote by $\rho \in [-1, 1]$ the instantaneous correlation coefficient defined by $d\langle W, \hat{Z}\rangle_t = \rho dt$. It is also convenient to write

$$\hat{Z}_t = \rho W_t + \sqrt{1 - \rho^2} Z_t, \tag{2.13}$$

where (Z_t) is a standard BM independent of (W_t).

Some facts and assumptions are:

- It is often found from financial data that $\rho < 0$, and there are economic arguments for a negative correlation or *leverage effect* between stock price and volatility shocks.
- Notice the fat tails of stock-price distribution due to the random volatility. In particular, the negative correlation causes the tails to be asymmetric: the left tail is fatter.
- We assume ρ a constant from now on for simplicity and because it is taken to be such in most practical situations.
- We still consider the underlying probability space $(\Omega, \mathcal{F}, P)$, where now we can take $\Omega = C([0, \infty) : \mathbb{R}^2)$, the space of all continuous trajectories $(W_t(\omega), Z_t(\omega)) = \omega(t)$ in $\mathbb{R}^2$.
- The filtration $(\mathcal{F}_t)_{t \geq 0}$ represents the information on the two BMs.

Some common driving processes (Y_t) are:

1. Log-Normal (LN) $dY_t = c_1 Y_t dt + c_2 Y_t d\hat{Z}_t$
2. Ornstein-Uhlenbeck (OU) $dY_t = \alpha(m - Y_t)dt + \beta d\hat{Z}_t$
3. Feller or Cox-Ingersoll-Ross (CIR) $dY_t = \kappa(m' - Y_t)dt + \nu\sqrt{Y_t} d\hat{Z}_t$

Some models often studied in the literature are listed in Table 1.

Table 1: Models of Stochastic Volatility

Authors	Correlation	$f(y)$	Process
Hull-White	$\rho = 0$	$\sqrt{y}$	LN
Scott	$\rho = 0$	$\exp(y)$	OU
Stein-Stein	$\rho = 0$	$\lvert y\rvert$	OU
Ball-Roma	$\rho = 0$	$\sqrt{y}$	CIR
Heston	$\rho \neq 0$	$\sqrt{y}$	CIR

2.4 Derivative Pricing

When the volatility is a *Markovian Itô process*, we can find a pricing function for European derivatives of the form $P(t, X_t, Y_t)$ from no-arbitrage arguments, as in the BS case. We will derive the pricing PDE.[3] Assuming that volatility is a function of a mean-reverting OU process:

$$\begin{aligned} dX_t &= \mu X_t dt + \sigma_t X_t dW_t, \\ \sigma_t &= f(Y_t), \\ dY_t &= \alpha(m - Y_t)dt + \beta d\hat{Z}_t, \\ \hat{Z}_t &= \rho W_t + \sqrt{1-\rho^2} Z_t, \end{aligned} \tag{2.14}$$

where W_t and Z_t are independent BMs.

Unlike the BS case, it is not sufficient to hedge solely with the underlying asset, since there are two Wiener processes. Thus, we try to hedge with the underlying asset and *another option* has a different expiration date. Let $P^{(1)}$ be the price of a European derivative with expiration date T_1 and payoff function h, $P^{(2)}$ be the price of a European contract with $T_2 > T_1 > t$ and h. We try to find processes $\{a_t, b_t, c_t\}$ s.t.

$$P^{(1)}(T_1, X_{T_1}, Y_{T_1}) = a_{T_1} X_{T_1} + b_{T_1} \beta_{T_1} + c_{T_1} P^{(2)}(T_1, X_{T_1}, Y_{T_1}), \tag{2.15}$$

where $\beta_t = e^{rt}$ is the price of a risk-free bond under the prevailing short-term constant interest rate r. Also, the portfolio is to be self-financing, so that

$$dP^{(1)} = a_t dX_t + b_t r e^{rt} dt + c_t dP^{(2)}. \tag{2.16}$$

If such a portfolio can be found then, in order for there to be no-arbitrage opportunities, it must be that

$$P^{(1)}(t, X_t, Y_t) = a_t X_t + b_t \beta_t + c_t P^{(2)}(t, X_t, Y_t), \quad \forall\, t < T_1. \tag{2.17}$$

Theorem 2.1. *(Two-Dimensional Version of Itô's Formula)*

$$dg(t, X_t, Y_t) = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX_t + \frac{\partial g}{\partial y} dY_t + \frac{1}{2}\left(\frac{\partial^2 g}{\partial x^2} d\langle X \rangle_t + 2\frac{\partial^2 g}{\partial x \partial y} d\langle X, Y \rangle_t + \frac{\partial^2 g}{\partial y^2} d\langle Y \rangle_t \right). \tag{2.18}$$

Applying this to both sides of (2.16) yields

$$\begin{aligned} &\left(\frac{\partial P^{(1)}}{\partial t} + \mathcal{M}_1 P^{(1)} \right) dt + \frac{\partial P^{(1)}}{\partial x} dX_t + \frac{\partial P^{(1)}}{\partial y} dY_t \\ &= \left(a_t + c_t \frac{\partial P^{(2)}}{\partial x} \right) dX_t + c_t \frac{\partial P^{(2)}}{\partial y} dY_t + \left[c_t \left(\frac{\partial}{\partial t} + \mathcal{M}_1 \right) P^{(2)} + b_t r e^{rt} \right] dt, \end{aligned} \tag{2.19}$$

where

$$\mathcal{M}_1 = \frac{1}{2} f(y)^2 x^2 \frac{\partial^2}{\partial x^2} + \rho \beta x f(y) \frac{\partial^2}{\partial x \partial y} + \frac{1}{2}\beta^2 \frac{\partial^2}{\partial y^2}. \tag{2.20}$$

Equating dY_t, dX_t and using (2.17), which gives

$$\begin{aligned} c_t &= \frac{\partial P^{(1)}/\partial y}{\partial P^{(2)}/\partial y}, \\ a_t &= \frac{\partial P^{(1)}}{\partial x} - c_t \frac{\partial P^{(2)}}{\partial x}, \\ b_t &= (P_t^{(1)} - a_t X_t - c_t P^{(2)}) e^{-rt}. \end{aligned} \tag{2.21}$$

[3] This PDE is not solved here.

Comparing dt terms in (2.19) gives

$$\left(\frac{\partial P^{(1)}}{\partial y}\right)^{-1} \mathcal{M}_2 P^{(1)}(t, X_t, Y_t) = \left(\frac{\partial P^{(2)}}{\partial y}\right)^{-1} \mathcal{M}_2 P^{(2)}(t, X_t, Y_t), \tag{2.22}$$

where

$$\mathcal{M}_2 = \frac{\partial}{\partial t} + \mathcal{M}_1 + r\left(x\frac{\partial}{\partial x} - \cdot\right). \tag{2.23}$$

Now, the left-hand side of (2.22) contains terms depending on T_1 but not T_2 and vice versa for the right-hand side. Thus, both sides must be equal to a function that does not depend on expiration date T_1, T_2. We denote this function by

$$-\left(\alpha(m-y) - \beta\left(\rho\frac{\mu - r}{f(y)} + \gamma(t,x,y)\sqrt{1-\rho^2}\right)\right) \tag{2.24}$$

where $\gamma(t,x,y)$ is an arbitrary function.[4] Thus, the pricing function $P(t,x,y)$ must satisfy

Definition 2.5. (Kolmogorov or Feynman-Kac PDE)

$$\begin{aligned}&\frac{\partial P}{\partial t} + \frac{1}{2}f(y)^2x^2\frac{\partial^2 P}{\partial x^2} + \rho\beta x f(y)\frac{\partial^2 P}{\partial x \partial y} + \frac{1}{2}\beta^2\frac{\partial^2 P}{\partial y^2}\\ &+r\left(x\frac{\partial P}{\partial x} - P\right) + \left[(\alpha(m-y) - \beta\Lambda(t,x,y))\right]\frac{\partial P}{\partial y} = 0\end{aligned} \tag{2.25}$$

where

$$\Lambda(t,x,y) = \rho\frac{\mu - r}{f(y)} + \gamma(t,x,y)\sqrt{1-\rho^2}. \tag{2.26}$$

The terminal condition is $P(T,x,y) = h(x)$.

Remark 2.6. The function γ in (2.26) is the *risk premium factor* from the second source of randomness Z_t in (2.13) that drives the volatility. In the perfectly correlated case $|\rho| = 1$ it does not appear, as expected. The reason for this terminology is the calculation

$$\begin{aligned}dP(t, X_t, Y_t) = &\left[\underline{\frac{\mu - r}{f(y)}}\left(x f(y)\frac{\partial P}{\partial x} + \underline{\beta\rho\frac{\partial P}{\partial y}}\right) + rP + \underline{\gamma\beta\sqrt{1-\rho^2}\frac{\partial P}{\partial y}}\right]dt\\ &+\left(x f(y)\frac{\partial P}{\partial x} + \beta\rho\frac{\partial P}{\partial y}\right)dW_t + \beta\sqrt{1-\rho^2}\frac{\partial P}{\partial y}dZ_t,\end{aligned} \tag{2.27}$$

which is obtained by the Itô's formula (2.18) and the PDE (2.25) satisfied by P. From this expression we see that an infinitesimal fractional increase in the volatility risk β increases the infinitesimal rate of return on the option by γ times that fraction, in addition to the increase from the excess return-to-risk ratio $(\mu - r)/f(y)$.

2.5 Pricing with Equivalent Martingale Measures

We give an alternative derivation of the no-arbitrage derivative price using *risk-neutral* theory, again for the model (2.14), but the procedure is valid for *general models*, including non-Markovian ones.

[4]The original idea was to guess the form.

Suppose that there is a an EMM P^* under which the discounted stock price $\widetilde{X}_t = e^{-rt}X_t$ is a martingale. Thus, we can price any derivative with expiration T and payoff h by

$$V_t = E^*[e^{-r(T-t)}H|\mathcal{F}_t] \quad \forall\, t \le T, \tag{2.28}$$

where $H = h(X_T)$, then there is no-arbitrage opportunity (See Section 1.7).

We absorb the drift term of $\widetilde{X}_t$ in its martingale term by setting

$$W_t^* = W_t + \int_0^t \frac{\mu - r}{f(Y_s)} ds. \tag{2.29}$$

Any shift of the *second independent BM* of the form

$$Z_t^* = Z_t + \int_0^t \gamma_s ds \tag{2.30}$$

will not change the drift of $\widehat{X}_t$. Here (γ_t) is any adapted (and suitably regular) process. By *Girsanov's theorem*[5], (W^*) and (Z^*) are independent standard BMs under a measure $P^{*(\gamma)}$ defined by[6]

$$\begin{aligned}\frac{dP^{*(\gamma)}}{dP} &= \exp\left\{-\int_0^T \theta_t^{(1)} dW_t - \int_0^T \theta_t^{(2)} dZ_t - \frac{1}{2}\int_0^T [(\theta_t^{(1)})^2 + (\theta_t^{(2)})^2] dt\right\}, \\ \theta_t^{(1)} &= \frac{\mu - r}{f(Y_t)}, \quad \theta_t^{(2)} = \gamma_t.\end{aligned} \tag{2.31}$$

Technically, we shall make an assumption on the pair $(\theta_t^{(1)}, \theta_t^{(2)})$ so that $P^{*(\gamma)}$ is well-defined as a probability measure.[7] In particular, this will be the case if f is bounded away from zero and (γ_t) is bounded. Then, under $P^{*(\gamma)}$, the SDEs (2.14) become

$$\begin{aligned} dX_t &= rX_t dt + f(Y_t)X_t dW_t^*, \\ dY_t &= \left[\alpha(m - Y_t) - \beta\left(\rho\frac{\mu - r}{f(Y_t)} + \gamma_t\sqrt{1-\rho^2}\right)\right] dt + \beta d\hat{Z}_t^*, \\ \hat{Z}_t^* &= \rho W_t^* + \sqrt{1-\rho^2} Z_t^*. \end{aligned} \tag{2.32}$$

Any allowable choice of γ leads to an EMM $P^{*(\gamma)}$ and the possible no-arbitrage prices are[8]

$$V_t = E^{*(\gamma)}[e^{-r(T-t)}H|\mathcal{F}_t]. \tag{2.33}$$

The process (γ_t) is called the risk premium factor or the *market price of volatility risk* from the second source of randomness Z that drives the volatility. As a result, the *martingale representation theorem* 1.7 says that under $P^{*(\gamma)}$ martingale $M_t = e^{-rt}V_t$ is a stochastic integral wrt (W^*, Z^*):

$$M_t = M_0 + \int_0^t \eta_s^{(1)} dW_s^* + \int_0^t \eta_s^{(2)} dZ_s^* \tag{2.34}$$

for some adapted and suitably bounded processes $(\eta_t^{(1)})$ and $(\eta_t^{(2)})$. So, we cannot replicate the claim by trading in stock and bond only due to the last integral in (2.34).[9] We can, however, hedge one derivative contract $P^{(1)}$ with the stock and another derivative security $P^{(2)}$. The calculation is like the Markovian case in Section 2.4, but the hedging ratios are *non-unique* since they depend on γ. This procedure is usually unsatisfactory owing to the *higher transaction costs* and *less liquidity* associated with trading the second derivative.

[5] See Remark 1.16.

[6] The first equation is Radon-Nikodym derivative mentioned in Remark 1.13.

[7] For Example, Heston model needs the Feller condition for the existence of ELMM, detailed in Section 3 of Wong2006.pdf and arxiv20191230.pdf.

[8] This expectation can be obtained by (1.41).

[9] Q: How about VIX options? If we trade volatility derivatives, is there any difference?

2.6 Implied Volatility as a Function of Moneyness

Another reason that IV is a particularly useful measure of the performance of a SV model is that IV is a function of a European option contract's *moneyness* K/x. Given any Markovian SV model under which the stock price satisfies

$$dX_s = rX_s ds + \sigma_s X_s dW_s^* \tag{2.35}$$

under some risk-neutral measure $P^{*(\gamma)}$, suppose it is now time t, $X_t = x$ and define $\widetilde{X} = X/x$. Then $(\widetilde{X}_s)_{s\geq t}$ satisfies the same SDE (2.35), with initial value $\widetilde{X}_t = 1$, neither of which depends on x. The call option price

$$\begin{aligned} C &= E^{*(\gamma)}[e^{-r(T-t)}(X_T - K)^+ | X_t = x, \sigma_t] \\ &= E^{*(\gamma)}[e^{-r(T-t)}(x\widetilde{X}_T - K)^+ | \widetilde{X}_t = 1, \sigma_t] \\ &= KE^{*(\gamma)}[e^{-r(T-t)}(\frac{x}{K}\widetilde{X}_T - 1)^+ | \widetilde{X}_t = 1, \sigma_t] \\ &= KQ_1(t, K/x; T) \end{aligned} \tag{2.36}$$

for some function Q_1 depending on K/x but not on x and K separately. The IV I is computed from

$$C = C_{BS}(t, x; K, I), \tag{2.37}$$

and from the BS-formula (1.30) we have

$$C_{BS}(t, x; K, I) = KQ_2(t, K/x; I) \tag{2.38}$$

for some function Q_2 also depending on K/x but not on x and K separately. From the relation $KQ_1(t, K/x; T) = KQ_2(t, K/x; I)$ we see that I must be a function of moneyness K/x but not of K and x separately.

This is useful because it tells us that we can obtain the implied volatility curve predicted by a SV model by solving the PDE (2.25) with terminal condition $h(x) = (x - K)^+$ for *a fixed strike price* K. This is because plotting the resulting implied volatilities as a function of moneyness for different starting values x gives the same curve as if we were varying K.

2.7 Market Price of Volatility Risk and Data

Now, there are more details.

Bound: In Section 2.5, we know that (2.33) is a possible no-arbitrage derivative pricing formula for any EMM $P^{*(\gamma)}$. If the volatility is unbounded, then the range of European call option prices given by (2.33) with $H = (X_T - K)^+$ is between the price of stock and intrinsic value of the contract, $(X_T - K)^+ \leq V_t \leq X_t$, and the extreme values are attained for some EMMs. When volatility is assumed to be bounded with $\sigma_t \in [\sigma_{\min}, \sigma_{\max}]$, the bounds are $C_{BS}(\sigma_{\min}) \leq V_t \leq C_{BS}(\sigma_{\max})$.

Super-hedge: The viewpoint, market selects a *unique* EMM under which derivative contracts are priced, is sometimes called *selecting an approximating complete market*. Although discounted derivative prices cannot be replicated by stock and bond alone, they can be *super-replicated*: for example, buying the stock at time $t < T$ and holding it until expiration *super-hedges* a short call position because $X_T > (X_T - K)^+$; this may yield a profit but never a loss. This (trivial) strategy is very expensive (it costs \$ X_t), so many researches concern finding the cheapest super-hedging strategy.

Parameter Estimation: When we estimate the parameters for our model, we could use econometric methods such as maximum likelihood or method of moments on historical stock-price time-series data to find (α, β, m, ρ) plus the present volatility in the model (2.14). Then we would need some derivative data to estimate γ, assuming for instance that it is constant. The common practice, called *cross-sectional fitting*, is to estimate all the parameters from derivative data, usually *at-the-money* European option prices (or a section of the observed implied volatility surface). This ignores the statistical basis for the modeling but is easier to implement than time-series methods, which suffer because the (σ_t) process is not directly observable. If today is time $t = 0$ and we denote $\vartheta = (\alpha, \beta, m, \rho, \sigma_0, \gamma, \mu)$, then a typical *least-square fit* is to observe call option prices $C^{obs}(K, T)$ for (K, T) in some set $\mathcal{K}$ and to solve

$$\min_{\vartheta} \sum_{(K,T)\in\mathcal{K}} (C(K, T; \vartheta) - C^{obs}(K, T))^2, \tag{2.39}$$

where $C(K, T; \vartheta)$ is the model-predicted call option price (either from solving the PDE (2.25) with $h(x) = (x - K)^+$ or from Monte Carlo simulation). But this process can be very slow and computationally intensive.

3 Heston Stochastic Volatility Model

This section introduces the *Heston model* which is used by Steven L. Heston in 1993.[10] Heston used a new technique to derive a closed-form solution for the price of European call option on an asset with SV. The model allows arbitrary correlation between volatility and spot-asset returns. He also shows that correlation is important for explaining return skewness and strike-price biases in the BS model. The solution technique is based on *characteristic functions*.

3.1 Stochastic Differential Equations

Assuming that the spot asset price S_t satisfies the diffusion equation, the volatility follows the OU process, and Itô's lemma (1.11) shows that the variance v_t follows the CIR process,

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dW_t, \\ d\sqrt{v_t} &= -\beta\sqrt{v_t}dt + \delta d\hat{Z}_t, \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}d\hat{Z}_t, \\ d\langle W, \hat{Z}\rangle_t &= \rho dt. \end{aligned} \tag{3.1}$$

where $W_t, \hat{Z}_t$ are Wiener processes. Assume $\kappa, \theta, \sigma > 0$ and $-1 < \rho < 1$. Also, we assume the variance process satisfies the Feller condition $2\kappa\theta > \sigma^2$, so it is always positive and cannot reach zero.

For simplicity, we assume a constant interest rate r. Therefore, the price at time t of a unit discount bond that matures at time $t + \tau$ is

$$P(t, t+\tau) = e^{-r\tau}. \tag{3.2}$$

3.2 Heston Partial Differential Equation

Standard arbitrage arguments demonstrate that the value of any asset $U(S, v, t)$ must satisfy the PDE[11]

$$\begin{aligned} &\frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \rho\sigma v S\frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 U}{\partial v^2} + rS\frac{\partial U}{\partial S} \\ &\qquad + [\kappa(\theta - v_t) - \lambda(S, v, t)]\frac{\partial U}{\partial v} - rU + \frac{\partial U}{\partial t} = 0. \end{aligned} \tag{3.3}$$

The unspecified term $\lambda(S, v, t)$ represents *the price of volatility risk*, and must be independent of the particular asset. To motivate the choice of $\lambda(S, v, t)$, Heston noted that in Breeden's (1979) consumption-based model,

$$\lambda(S, v, t)dt = \gamma\,\mathrm{Cov}[dv, dC/C], \tag{3.4}$$

where C_t is the consumption rate and γ is the relative-risk aversion of an investor. Consider the consumption process that emerges in the Cox, Ingersoll and Ross (1985) model

$$dC_t = \mu_c v_t C_t dt + \sigma_c\sqrt{v_t}C_t d\hat{Z}_{c,t}, \tag{3.5}$$

where $\hat{Z}_{c,t}$ is another Wiener process with the constant correlations with $W_t, \hat{Z}_t$. This generates a risk premium proportional to v, i.e. $\lambda(S, v, t) = \lambda v$.

[10] All researches mentioned in this section can be found in the reference of Heston1993_RFS.pdf.
[11] This is Feynman-Kac PDE derived in Section 2.4.

A European call option with strike price K and maturing at time T satisfies the PDE (3.3) subject to the following boundary conditions:

$$\begin{aligned} U(S, v, T) &= (S_T - K)^+, \\ U(0, v, t) &= 0, \\ \frac{\partial U}{\partial S}(\infty, v, t) &= 1, \\ \left(rS\frac{\partial U}{\partial S} + \kappa\theta\frac{\partial U}{\partial v} - rU + \frac{\partial U}{\partial t} \right)(S, 0, t) &= 0, \\ U(S, \infty, t) &= S_t. \end{aligned} \tag{3.6}$$

Remark 3.1. Although we will use this form of the risk premium, the pricing results are obtained by arbitrage and do not depend on the other assumptions of the Breeden (1979) or CIR (1985) models.

3.3 The Call Option Price

By analogy with the BS formula, we guess a solution of the form [12]

$$C(S, v, t) = S_t P_1 - K P(t, T) P_2, \tag{3.7}$$

where the first term is the present value of the spot asset upon optimal exercise, and the second term is the present value of the strike-price payment. Both of these terms must satisfy the original PDE (3.3). For notional convenience, let

$$x_t = \log(S_t). \tag{3.8}$$

Substituting the proposed solution (3.7) into the original PDE (3.3) shows that P_1 and P_2 must satisfy the PDEs[13]

$$\begin{aligned} \frac{1}{2}v\frac{\partial^2 P_j}{\partial x^2} + \rho\sigma v\frac{\partial^2 P_j}{\partial x \partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_j}{\partial v^2} + (r + u_j v)\frac{\partial P_j}{\partial x} \\ + (a - b_j v)\frac{\partial P_j}{\partial v} + \frac{\partial P_j}{\partial t} = 0. \end{aligned} \tag{3.9}$$

for $j = 1, 2$, where

$$u_1 = \frac{1}{2}, \quad u_2 = -\frac{1}{2}, \quad a = \kappa\theta, \quad b_1 = \kappa + \lambda - \rho\sigma, \quad b_2 = \kappa + \lambda. \tag{3.10}$$

For the option price to satisfy the terminal conditions in (3.6), these PDEs (3.9) are subject to the terminal condition

$$P_j(x, v, T; K) = 1_{\{x_T \geq \ln K\}}. \tag{3.11}$$

3.4 Risk-Neutral Valuation

Following Heston, we know that $P_{1,2}$ are risk-neutral when x_t, v_t satisfy respectively[14]

$$\begin{aligned} dx_t &= (r + u_j v_t)dt + \sqrt{v_t}dW_t^*, \\ dv_t &= (a - b_j v_t)dt + \sigma\sqrt{v_t}d\hat{Z}_t^*. \end{aligned} \tag{3.12}$$

[12] This is a corollary of Girsanov's Theorem.

[13] Just set $S_t = 1$, $K = 0$ and $S_t = 0$, $K = -1/P(t, T)$.

[14] In fact, we will choose $j = 2$ and set $\lambda = 0$.

Then we obtain

$$P_j(x,v,t;K) = E^*[1_{\{x_T \geq \ln K\}} | x_t = x, v_t = v]. \tag{3.13}$$

They are not immediately available in closed form. However, we know that[15] their characteristic functions satisfy the same PDEs (3.9), subject to the terminal condition under the risk-neutral measure

$$\Phi_j(x,v,T;u) = e^{iux}. \tag{3.14}$$

To obtain the solution, we guess the form of characteristic functions [16]

$$\Phi_j(\tau,u) = E^*\left[e^{iux_T} | x_t = x, v_t = v\right] = \exp[C_j(\tau,u) + D_j(\tau,u)v + iux], \tag{3.15}$$

where $\tau = T - t$. Substituting these into (3.9) shows a *Riccati equation* and a straightforward ODE

$$\begin{aligned} \frac{dD_j}{d\tau} &= \rho\sigma iuD_j - \frac{1}{2}u^2 + \frac{1}{2}\sigma^2 D_j^2 + u_j iu - b_j D_j, \\ \frac{dC_j}{d\tau} &= riu + aD_j. \end{aligned} \tag{3.16}$$

Heston specifies the initial conditions $C_j(0,u) = D_j(0,u) = 0$. Then the results are

$$\begin{aligned} D_j(\tau,u) &= \frac{b_j - \rho\sigma iu + d_j}{\sigma^2}\left(\frac{1 - e^{d_j\tau}}{1 - g_j e^{d_j\tau}}\right), \\ C_j(\tau,u) &= riu\tau + \frac{a}{\sigma^2}\left[(b_j - \rho\sigma iu + d_j)\tau - 2\ln\left(\frac{1 - g_j e^{d_j\tau}}{1 - g_j}\right)\right], \end{aligned} \tag{3.17}$$

where

$$\begin{aligned} d_j &= \sqrt{(\rho\sigma iu - b_j)^2 - \sigma^2(2u_j iu - u^2)}, \\ g_j &= \frac{b_j - \rho\sigma iu + d_j}{b_j - \rho\sigma iu - d_j}. \end{aligned} \tag{3.18}$$

3.5 The Heston Pricing Formula

Through some calculations, we find that

$$P(t,T)\Phi_2(\tau,u) = e^{-r(T-t)}\Phi_2(\tau,u) = e^x\Phi_1(\tau,u+i) = S_t\Phi_1(\tau,u+i). \tag{3.19}$$

Also, one can invert the characteristic functions to get cdf and (3.11) becomes

$$\begin{aligned} P_j = 1 - F_j^*(\ln K) &= \frac{1}{2\pi}\lim_{y\to+\infty}\lim_{L\to\infty}\int_{-L}^{L}\frac{e^{-iu\ln K} - e^{-iuy}}{iu}\Phi_j du \\ &= \frac{1}{2\pi}\lim_{y\to+\infty}\lim_{L\to\infty}\int_{-L}^{L}\frac{1 - e^{-iuy}}{iu}\Phi_j du + \frac{1}{2\pi}\lim_{L\to\infty}\int_{-L}^{L}\frac{e^{-iu\ln K} - 1}{iu}\Phi_j du \\ &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left(\frac{e^{-iu\ln K}\Phi_j}{iu}\right)du. \end{aligned} \tag{3.20}$$

[15] There are details in the Appendix of Heston1993_RFS.pdf.

[16] This exploits the linearity of the coefficients in the PDE (3.9).

The third equal sign adds 1 on the numerators so the integrands can be finite at $u = 0$. The last equal sign is based on characteristic function's properties $\Phi(0) = 1$, $\Phi(u) = \overline{\Phi(-u)}$ and a useful integral

$$\int_0^{+\infty} \frac{\sin \alpha x}{x} dx = \begin{cases} \frac{\pi}{2} & \alpha > 0 \\ 0 & \alpha = 0 \\ -\frac{\pi}{2} & \alpha < 0 \end{cases} \tag{3.21}$$

Finally, we can obtain the European call option price

$$C(S, v, t; K, T, r) = P(t, T) \left[\frac{1}{2}(F_t - K) + \frac{1}{\pi} \int_0^{\infty} (F_t \cdot f_1 - K \cdot f_2)\, du \right] \tag{3.22}$$

where $F_t = \frac{S_t}{P(t,T)} = S_t e^{r(T-t)}$ and

$$\begin{aligned} f_1 &= \Re\left(\frac{e^{-iu \ln K} \varphi(u - i)}{iuF_t} \right), \\ f_2 &= \Re\left(\frac{e^{-iu \ln K} \varphi(u)}{iu} \right), \\ \varphi(u) &= \Phi_2(\tau; u) = \exp(C(\tau, u) + D(\tau, u)v + iu \ln S_t), \\ C(\tau, u) &= \frac{\kappa\theta}{\sigma^2} \left((\kappa - i\rho\sigma u + d(u))\tau - 2 \ln\left(\frac{1 - g(u)e^{d(u)\tau}}{1 - g(u)} \right) \right), \\ D(\tau, u) &= \frac{\kappa - i\rho\sigma u + d(u)}{\sigma^2} \left(\frac{1 - e^{d(u)\tau}}{1 - g(u)e^{d(u)\tau}} \right), \\ d(u) &= \sqrt{(\kappa - i\rho\sigma u)^2 + \sigma^2(u^2 + iu)}, \\ g(u) &= \frac{\kappa - i\rho\sigma u + d(u)}{\kappa - i\rho\sigma u - d(u)}. \end{aligned} \tag{3.23}$$

Remark 3.2. There are mainly two troubles when doing numerical quadrature in (3.22). One is that the Heston factors $f_{1,2}$ are usually oscillatory. The other difficulty is the calculations of complex logarithm and square root, which can result in numerical instability.[17]

Remark 3.3. The put option price can be obtained by put-call parity (1.33).

3.6 Finite Difference Schemes

3.7 Numerical Integration

Gauss-Kronrod quadrature

[17] There are discussions in HestonTrap.pdf and NotSoComplexLogarithmsInTheHestonModel.pdf.

4 Heston-Nandi GARCH Model

This section introduce Heston's further research in 2000.[18] In their paper, Heston and Nandi develops a closed-form option valuation formula for a spot asset whose variance follows a Generalized Auto-regressive Conditional Heteroskedasticity (GARCH) process that can be correlated with the returns of the spot asset. It provides the first readily computed option formula for a random volatility model that can be estimated and implemented solely on the basis of observables. They also show that the improvement, to *ad hoc* BS model, is largely due to the ability of the GARCH model to simultaneously capture the correlation of volatility with spot returns and the path dependence in volatility.

4.1 Time Series Model

Under the physical measure P, the discertized form of SDE (3.1) is a GARCH$(1,1)$ process, which looks like Engle and Ng's NGARCH rather than Duan's classic GARCH. We assume the variance process follows GARCH(p,q):

$$\begin{aligned} r_t &= \ln\left(\frac{S_t}{S_{t-1}}\right) = r + \lambda v_t + \sqrt{v_t} z_t, \\ v_t &= w + \sum_{i=1}^{p} b_i v_{t-i} + \sum_{j=1}^{q} a_j (z_{t-j} - c_j \sqrt{v_{t-j}})^2 \\ &= w + \sum_{i=1}^{p} b_i v_{t-i} + \sum_{j=1}^{q} a_j \frac{(r_{t-j} - r - \lambda v_{t-j} - c_j v_{t-j})^2}{v_{t-j}}, \end{aligned} \tag{4.1}$$

where S_t is the underlying asset price at time t, r is the continuously compounded risk-free rate, v_t is the conditional variance of the log return between $t-1$ and t and is known at time $t-1$, λv_t is *equity risk premium*, z_t is a standard BM.

At this point we cannot value options or other contingent because we do not know the risk-neutral distribution of the spot price. Thus, we rearrange SDE (4.1) in the form

$$\begin{aligned} r_t &= r - \frac{1}{2} v_t + \sqrt{v_t} z_t^*, \\ v_t &= w + \sum_{i=1}^{p} b_i v_{t-i} + \sum_{j=2}^{q} a_j (z_{t-j} - c_j \sqrt{v_{t-j}})^2 + a_1 (z_{t-1}^* - c_1^* \sqrt{v_{t-1}})^2, \end{aligned} \tag{4.2}$$

where $z_t^* = z_t + (\lambda + 1/2)\sqrt{v_t}$ and $c_1^* = c_1 + \lambda + 1/2$.

There is no reason for the risk-neutral distribution of z_t^* to be normal because BS price do not follow absence of arbitrage with discrete-time trading. In order for z_t^* to have a standard normal risk-neutral distribution, we need a assumption: *the value of a call option with one period to expiration obeys the Black-Scholes-Rubinstein formula.*

Remark 4.1. Following Heston and Nandi, we will focus on the first-order, i.e. $p = q = 1$. The first-order process remains stationary with finite mean and variance if $b_1 + a_1 c_1^2 < 1$. In the multiple factor case, one must add a condition that the roots of $x^p - \sum_{i=1}^{p} (b_i + a_i c_i^2) x^{p-i}$ lie inside the unit circle.

[18] All reference in this section can be found in Heston2000_RFS.pdf

4.2 Estimation of Model Parameters

Typically, we use GARCH$(1,1)$, which converges to continuous time Heston's SV model, when we deal with the short term data, like daily data. Thus, we need to estimate five parameters (a, b, c, w, λ). We do this with the Maximum Likelihood Estimation (MLE) used by Bollerslev (1986) and many others. [19] Since $z_t \sim N(0,1)$, the likelihood function is

$$L(a, b, c, w, \lambda; r_t) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi v_t}} \exp\left[-\frac{(r_t - r - \lambda v_t)^2}{2v_t}\right]. \tag{4.3}$$

Then the log-likelihood function is

$$\ln L = -\frac{1}{2}\sum_{t=1}^{T}\left[\ln(2\pi v_t) + \frac{(r_t - r - \lambda v_t)^2}{v_t}\right], \tag{4.4}$$

where

$$v_t = w + b v_{t-1} + a\frac{(r_{t-1} - r - \lambda v_{t-1} - c v_{t-1})^2}{v_{t-1}}. \tag{4.5}$$

Remark 4.2. Following Heston and Nandi, we assume that v_1 equals to the sample variance of r_t. Due to strong mean reversion of volatility, all results should be insensitive to the starting value v_1.

4.3 Call Option Pricing

There are mainly two methods about pricing the call option after we get the parameters in (4.1). The first one is Monte Carlo Simulation under EMM, which will be slow and computationally intensive for empirical work. In contrast, we can use the Heston model's analytic formula introduced in Section 3.

Let $f(u)$ be the moment generating function of the log asset price under the physical measure. Following the derivation in Section 3.4, we can get

$$f(u) = E_t[e^{u \ln S_T}] = E_t[S_T] = S_t^u \exp(A_t + B_t v_{t+1}). \tag{4.6}$$

The two coefficients can be calculated recursively, by working backward from the maturity date and using the terminal conditions [20]

$$\begin{aligned} A_t &= A_{t+1} + ur + B_{t+1}w - \frac{1}{2}\ln(1 - 2aB_{t+1}), \\ B_t &= u(\lambda + c) - \frac{1}{2}c^2 + bB_{t+1} + \frac{(u-c)^2}{2(1-2aB_{t+1})}, \\ A_T &= B_T = 0. \end{aligned} \tag{4.7}$$

We can rewrite the closed-form formula (3.22) as followed:

$$\begin{aligned} C = \frac{1}{2}S_t + \frac{e^{-r(T-t)}}{\pi}\int_0^\infty \Re\left[\frac{K^{-iu} f^*(iu+1)}{iu}\right] du \\ - Ke^{-r(T-t)}\left\{\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left[\frac{K^{-iu} f^*(iu)}{iu}\right] du\right\}, \end{aligned} \tag{4.8}$$

where $f^*(iu)$ is the conditional characteristic function of the log asset price under the EMM.

[19] Q: How to solve? GD or manually?

[20] The details are in the Appendix A of Heston_2000.pdf.

Remark 4.3. Observing the GARCH forms in (4.1) and (4.2), we find the similarity if $p = q = 1$. Since we have the formula of $f(u)$, the calculation of $f^*(iu)$ only needs to replace λ with $-1/2$ and c with $c + \lambda + 1/2$.