MAFS 5030 - Quantitative Modeling of Derivative Securities

Topic 4 – Extended option models

4.1 Continuous dividend yield models: Time dependent parameters

4.2 Exchange options: Change of numeraire

4.3 Quanto options – equity options with exchange rate risk exposure

4.4 Implied volatility and volatility smiles

4.5 Volatility exposure generated by delta hedging options

4.6 VIX

4.7 Guaranteed minimum withdrawal benefit

4.8 Transaction costs models

4.1 Continuous dividend yield models

Let q denote the constant continuous dividend yield, that is, the holder receives dividend of amount equal to $qS\ dt$ within the infinitesimal interval $(t, t + \mathrm{d}t)$. The asset price dynamics is assumed to follow the Geometric Brownian motion

$$\frac{dS}{S} = \rho\, dt + \sigma\ dZ^P.$$

We form a riskless hedging portfolio by short selling one unit of the European call and long holding $\triangle$ units of the underlying asset. The differential change of the portfolio value Π is given by

$$\begin{aligned} d\Pi &= -dc + \triangle\ dS + q\triangle S\ dt \\ &= \left(-\frac{\partial c}{\partial t} - \frac{\sigma^2}{2} S^2 \frac{\partial^2 c}{\partial S^2} + q\triangle S\right) dt + \left(\triangle - \frac{\partial c}{\partial S}\right) dS. \end{aligned}$$

The last term $q\triangle S\ dt$ is the wealth added to the portfolio due to the dividend payment received. By choosing $\triangle = \dfrac{\partial c}{\partial S}$, we obtain a riskless hedge for the portfolio. By no-arbitrage argument, the hedged portfolio should earn the riskless interest rate.

We then have

$$d\Pi = \left(-\frac{\partial c}{\partial t} - \frac{\sigma^2}{2}S^2\frac{\partial^2 c}{\partial S^2} + qS\frac{\partial c}{\partial S}\right)dt = r\left(-c + S\frac{\partial c}{\partial S}\right)dt,$$

which leads to

$$\frac{\partial c}{\partial \tau} = \frac{\sigma^2}{2}S^2\frac{\partial^2 c}{\partial S^2} + (r-q)S\frac{\partial c}{\partial S} - rc, \quad \tau = T - t, \quad 0 < S < \infty, \quad \tau > 0.$$

Note that the expected rate of return of S_t under Q is reduced by q while the discount rate remains at r.

Martingale pricing approach

Suppose all the dividend yields received are continuously used to purchase additional units of asset, then the wealth process of holding one unit of the underlying asset initially is given by

$$\widehat{S}_t = e^{qt} S_t,$$

where e^{qt} represents the growth factor in the number of units. The wealth process $\widehat{S}_t$ follows

$$\frac{d\widehat{S}_t}{\widehat{S}_t} = (\rho + q)\, dt + \sigma\, dZ_t^P.$$

We would like to find the equivalent risk neutral measure Q under which the discounted wealth process $\widehat{S}_t^*$ is Q-martingale. We choose $\gamma(t)$ in the Radon-Nikodym derivative to be

$$\gamma(t) = \frac{\rho + q - r}{\sigma}.$$

Now Z_t^Q is a Q-Brownian motion and

$$dZ_t^Q = dZ_t^P + \frac{\rho + q - r}{\sigma}\, dt.$$

Also, $\widehat{S}_t^*$ becomes Q-martingale since

$$\frac{d\widehat{S}_t^*}{\widehat{S}_t^*} = \sigma\, dZ_t^Q.$$

The asset price S_t under the equivalent risk neutral measure Q becomes

$$\frac{dS_t}{S_t} = (r - q)\, dt + \sigma\, dZ_t^Q.$$

Hence, the risk neutral drift rate of S_t is $r - q$.

Analogy with the foreign currency options

The continuous yield model is also applicable to *options on foreign currencies* where the continuous dividend yield can be considered as the yield due to the interest earned by the foreign currency at the foreign interest rate r_f.

Call and put price formulas

The price of a European call option on a continuous dividend paying asset can be obtained by changing S to $Se^{-q\tau}$ in the price formula since the asset price S will be depleted at the rate q due to payment of dividend yield.

Suppose we let $\widetilde{S} = Se^{-q\tau}$, then the option pricing equation becomes

$$\frac{\partial c}{\partial \tau} = \frac{\sigma^2}{2}\widetilde{S}^2\frac{\partial^2 c}{\partial \widetilde{S}^2} + r\widetilde{S}\frac{\partial c}{\partial \widetilde{S}} - rc.$$

We replace S by $\widetilde{S} = e^{-r\tau}S$ in the usual Black-Scholes formula and observe that $\ln\frac{\widetilde{S}}{X}$ becomes $\ln\frac{S}{X} - q\tau$. The European call price formula with continuous dividend yield q is

$$c(S,\tau) = Se^{-q\tau}N(\widehat{d_1}) - Xe^{-r\tau}N(\widehat{d_2}),$$

where

$$\widehat{d_1} = \frac{\ln\frac{S}{X} + (r - q + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}, \quad \widehat{d_2} = \widehat{d_1} - \sigma\sqrt{\tau}.$$

Alternatively, knowing that the expected rate of return of S_t under Q is $\delta_S = r - q$, we can deduce that

$$c(S,\tau) = e^{-r\tau}[Se^{(r-q)\tau}N(\widehat{d_1}) - XN(\widehat{d_2})].$$

More explicitly, we have

$$c(S,\tau) = e^{-r\tau}\left\{E_Q^t\left[S_T\mathbf{1}_{\{S_T>X\}}\right] - XE_Q^t\left[\mathbf{1}_{\{S_T>X\}}\right]\right\}$$

so that

$$E_Q^t\left[S_T\mathbf{1}_{\{S_T>X\}}\right] = Se^{\delta_S\tau}N(\widehat{d_1})$$
$$E_Q^t\left[\mathbf{1}_{\{S_T>X\}}\right] = N(\widehat{d_2}).$$

Similarly, the European put formula with continuous dividend yield q can be deduced from the Black-Scholes put price formula to be

$$p = Xe^{-r\tau}N(-\widehat{d_2}) - Se^{-q\tau}N(-\widehat{d_1}).$$

In a similar manner, we deduce that

$$E_Q^t\left[S_T\mathbf{1}_{\{S_T<X\}}\right] = Se^{\delta_S\tau}N(-\widehat{d_1})$$
$$E_Q^t\left[\mathbf{1}_{\{S_T<X\}}\right] = N(-\widehat{d_2}).$$

Put-call parity and put-call symmetry

The new put and call prices satisfy the *put-call parity relation*

$$p = c - Se^{-q\tau} + Xe^{-r\tau}.$$

Furthermore, the following *put-call symmetry relation* can also be deduced from the above call and put price formulas

$$c(S, \tau; X, r, q) = p(X, \tau; S, q, r).$$

That is, the put price formula can be obtained from the corresponding call price formula by interchanging S with X and r with q in the two formulas.

- Recall that a call option entitles its holder the right to exchange the riskless asset for the risky asset, and vice versa for a put option. The dividend yield earned from the risky asset is q while that from the riskless asset is r.
- If we interchange the roles of the riskless asset and risky asset in a call option, the call becomes a put option, thus giving the justification for the put-call symmetry relation.

As a verification, consider

$$\begin{aligned}
p(X,\tau;S,q,r) = \; & Se^{-q\tau}N\left(-\frac{\ln\frac{X}{S}+\left(q-r-\frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}\right) \\
& - Xe^{-r\tau}N\left(-\frac{\ln\frac{X}{S}+\left(q-r+\frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}\right) \\
= \; & Se^{-q\tau}N(\widehat{d}_1) - Xe^{-r\tau}N(\widehat{d}_2) \\
= \; & c(S,\tau;X,r,q).
\end{aligned}$$

Time dependent parameters

Suppose the model parameters become time dependent functions, the Black-Scholes equation has to be modified as follows

$$\frac{\partial V}{\partial \tau} = \frac{\sigma^2(\tau)}{2} S^2 \frac{\partial^2 V}{\partial S^2} + [r(\tau) - q(\tau)]\, S \frac{\partial V}{\partial S} - r(\tau) V, \quad 0 < S < \infty, \quad \tau > 0,$$

where V is the price of the derivative security.

When we apply the following transformations: $y = \ln S$ and $w = e^{\int_0^\tau r(u)\, du} V$, then

$$\frac{\partial w}{\partial \tau} = \frac{\sigma^2(\tau)}{2} \frac{\partial^2 w}{\partial y^2} + \left[r(\tau) - q(\tau) - \frac{\sigma^2(\tau)}{2} \right] \frac{\partial w}{\partial y}.$$

Consider the following form of the fundamental solution

$$f(y, \tau) = \frac{1}{\sqrt{2\pi s(\tau)}} \exp\left(-\frac{[y + e(\tau)]^2}{2s(\tau)} \right),$$

which satisfies the initial condition: $f(y, 0^+) = \delta(y)$.

By direct differentiation, it can be shown that $f(y, \tau)$ satisfies the parabolic equation

$$\frac{\partial f}{\partial \tau} = \frac{1}{2} s'(\tau) \frac{\partial^2 f}{\partial y^2} + e'(\tau) \frac{\partial f}{\partial y}.$$

Suppose we let

$$\begin{aligned} s(\tau) &= \int_0^\tau \sigma^2(u)\, du \\ e(\tau) &= \int_0^\tau [r(u) - q(u)]\, du - \frac{s(\tau)}{2}, \end{aligned}$$

then f satisfies the same differential equation as that for $w(y, \tau)$. One can deduce that the fundamental solution is given by

$$\phi(y, \tau) = \frac{1}{\sqrt{2\pi \int_0^\tau \sigma^2(u)\, du}} \; \exp\left(-\frac{\{y + \int_0^\tau [r(u) - q(u) - \frac{\sigma^2(u)}{2}]\, du\}^2}{2 \int_0^\tau \sigma^2(u)\, du} \right).$$

Given the initial condition $w(y, 0)$, the solution can be expressed as

$$w(y, \tau) = \int_{-\infty}^{\infty} w(\xi, 0)\, \phi(y - \xi, \tau)\, d\xi.$$

Note that the time dependency of the coefficients $r(\tau), q(\tau)$ and $\sigma^2(\tau)$ will not affect the spatial integration with respect to ξ. We may simply make the following substitutions in the option price formulas

$$\begin{aligned}
r \text{ is replaced by } & \frac{1}{\tau}\int_0^\tau r(u)\,du \\
q \text{ is replaced by } & \frac{1}{\tau}\int_0^\tau q(u)\,du \\
\sigma^2 \text{ is replaced by } & \frac{1}{\tau}\int_0^\tau \sigma^2(u)\,du.
\end{aligned}$$

For example, the European call price formula with time dependent parameters is modified as follows:

$$c = Se^{-\int_0^\tau q(u)\,du}\,N(\widetilde{d_1}) - Xe^{-\int_0^\tau r(u)\,du}N(\widetilde{d_2})$$

where

$$\widetilde{d_1} = \frac{\ln\frac{S}{X} + \int_0^\tau [r(u) - q(u) + \frac{\sigma^2(u)}{2}]\,du}{\sqrt{\int_0^\tau \sigma^2(u)\,du}}, \quad \widetilde{d_2} = \widetilde{d_1} - \sqrt{\int_0^\tau \sigma^2(u)\,du}.$$

4.2 Exchange options

- An exchange option is an option that gives the holder the right but not the obligation to exchange one risky asset for another.
- Let X_t and Y_t be the price processes of the two risky assets.
- The terminal payoff of a European exchange option at maturity T of exchanging Y_T for X_T is given by $\max(X_T - Y_T, 0)$.

Under the risk neutral measure Q, let X_t and Y_t be governed by

$$\frac{dX_t}{X_t} = (r - q_X)\,dt + \sigma_X\,dZ^Q_{X,t} \quad \text{and} \quad \frac{dY_t}{Y_t} = (r - q_Y)\,dt + \sigma_Y dZ^Q_{Y,t},$$

where r is the constant riskless interest rate, σ_X and σ_Y are the constant volatility of X_t and Y_t, respectively, q_X and q_Y are the dividend yield of X_t and Y_t, respectively. Also, the two standard Brownian motions are correlated with $dZ^Q_{X,t}\,dZ^Q_{Y,t} = \rho\,dt$, where ρ is correlation coefficient.

Numeraire Invariance Theorem

With the use of the money market account $M(t) = \exp\left(\int_0^t r_u \, \mathrm{d}u\right)$ as the numeraire (accounting unit), the effect of the time value of money with respect to normalized asset values becomes immaterial. Discounted security price S_t/M_t is a martingale under a risk neutral measure Q.

The choice of the money market account $M(t)$ as the numeraire is not unique in order that no-arbitrage pricing principle holds. We may choose the price of a tradable asset $N(t)$ as the numeraire.

A numeraire is any strictly positive $(\mathcal{F}_T)_{t\in R_+}$-adapted stochastic process $(N_t)_{t\in R_+}$ that can be used as a unit of reference.

In order to effect the no-arbitrage pricing approach, one has to determine a probability measure under which $\widehat{S}_t = S_t/N_t$ is a martingale. With the martingale property remains intact (原封不动), there will be absence of arbitrage.

How to construct the new measure Q_N from the risk neutral measure Q such that the deflated prices are Q_N-martingale?

It is observed that $N_t^* = \frac{N_t}{M_t} = e^{-\int_0^T r_u \mathrm{d}u} N_t$ is an $\mathcal{F}_t$-martingale under Q since N_t is a traded asset.

Given $(N_t)_{t \in [0,T]}$, we define the associated measure Q_N via

$$\frac{\mathrm{d}Q_N}{\mathrm{d}Q} = \frac{N_T^*}{N_0^*} = e^{-\int_0^T r_u \, \mathrm{d}u} \, \frac{N_T}{N_0}.$$

This is equivalent to stating that

$$\int_\Omega X(\omega) \, \mathrm{d}Q_N(\omega) = \int_\Omega e^{-\int_0^T r_u \, \mathrm{d}u} \, \frac{N_T}{N_0} X(\omega) \, \mathrm{d}Q(\omega).$$

For any integrable $\mathcal{F}_T$-measurable random variable X, we have

$$E_{Q_N}[X] = E_Q\left[e^{-\int_0^T r_u \, \mathrm{d}u} \, \frac{N_T}{N_0} X\right].$$

From the martingale property of N_t^*, we deduce that

$$E_Q\left[\left.\frac{\mathrm{d}Q_N}{\mathrm{d}Q}\right|\mathcal{F}_t\right] = E_Q\left[\left.e^{-\int_0^T r_u\ \mathrm{d}u}\ \frac{N_T}{N_0}\right|\mathcal{F}_t\right] = \frac{N_t e^{-\int_0^t r_u\ \mathrm{d}u}}{N_0} = \frac{N_t^*}{N_0^*}.$$

By the tower rule, for any $\mathcal{F}_t$-measurable random variable G and integrable random variable X, we obtain

$$\begin{aligned}
E_{Q_N}[GX] &= E_Q\left[GXe^{-\int_0^T r_u\ \mathrm{d}u}\ \frac{N_T}{N_0}\right] \\
&= E_Q\left[G\frac{N_t}{N_0}e^{-\int_0^t r_u\ \mathrm{d}u}E_Q\left[\left.Xe^{-\int_t^T r_u\ \mathrm{d}u}\ \frac{N_T}{N_t}\right|\mathcal{F}_t\right]\right] \\
&= E_Q\left[GE_Q\left[\left.\frac{\mathrm{d}Q_N}{\mathrm{d}Q}\right|\mathcal{F}_t\right]E_Q\left[\left.Xe^{-\int_t^T r_u\ \mathrm{d}u}\ \frac{N_T}{N_t}\right|\mathcal{F}_t\right]\right] \\
&= E_Q\left[G\frac{\mathrm{d}Q_N}{\mathrm{d}Q}E_Q\left[\left.Xe^{-\int_t^T r_u\ \mathrm{d}u}\ \frac{N_T}{N_t}\right|\mathcal{F}_t\right]\right] \\
&= E_{Q_N}\left[GE_Q\left[\left.Xe^{-\int_t^T r_u\ \mathrm{d}u}\ \frac{N_T}{N_t}\right|\mathcal{F}_t\right]\right].
\end{aligned}$$

By comparing with

$$E_{Q_N}[GX] = E_{Q_N}[G E_{Q_N}[X|\mathcal{F}_t]],$$

we deduce that

$$E_{Q_N}[X|\mathcal{F}_t] = E_Q\left[X e^{-\int_t^T r_u \, \mathrm{d}u} \frac{N_T}{N_t}\middle| \mathcal{F}_t\right].$$

Consider an option with claim payoff C, by taking $X = \dfrac{C}{N_T}$, we obtain

$$N_t E_{Q_N}\left[\frac{C}{N_T}\middle| \mathcal{F}_t\right] = E_Q\left[e^{-\int_t^T r_u \, \mathrm{d}u} C\middle| \mathcal{F}_t\right].$$

Let V_t denote the time-t price of the contingent claim, where $V_T = C$. We then deduce that

$$\frac{V_t}{N_t} = E_{Q_N}\left[\frac{V_T}{N_T}\middle| \mathcal{F}_t\right]$$

so that the deflated price V_t/N_t is Q_N-martingale.

Remark One motivation of choosing another tradeable security (discount bond) as the numeraire arises from the pricing of an equity option under stochastic interest rate. Recall $\dfrac{M_t}{M_T} = e^{-\int_t^T r_u \, \mathrm{d}u}$ when the interest rate is stochastic so that

$$V_t = E_Q^t \left[e^{-\int_t^T r_u \, \mathrm{d}u} V(S_T) \right].$$

With the choice of the bond price B_t as the numeraire, we have

$$V_t = B_t E_{Q_T}^t \left[\frac{V_T(S_T)}{B_T} \right] = B_t E_{Q_T}^t [V_T(S_T)] \quad \text{since} \quad B_T = 1.$$

Since V_t/B_t is the time-t forward price of forward delivery of V_T at time T, so Q_T is termed the T-forward measure.

Use of the underlying asset as the numeraire (share measure) and the associated change of measure

Recall that the risk neutral measure uses the money market account as the numeraire. Let the starting time be time zero for notational convenience. Consider the Radon-Nikodym derivative L_t as a stochastic process that is defined by taking the ratio of the asset numeraire and the money market account

$$L_t = \left.\frac{dQ^S}{dQ}\right|_{\mathcal{F}_0} = e^{qt}\frac{S_t}{S_0} \bigg/ \frac{M_t}{M_0}, \quad t \in (0, T],$$

where $M_t = e^{rt}$ is the money market account and q is the dividend yield of the underlying asset.

The inclusion of the factor e^{qt} means one unit of the risky asset initially grows to e^{qt} units after time t if all dividends are invested into the purchase of new units of the risky asset.

Let $\widehat{S}_t = e^{qt} S_t$, then $\widehat{S}_t^* = \widehat{S}_t / e^{rt}$ is a martingale under Q. Note that $\widehat{S}_t$ is chosen as the numeraire asset instead of S_t since we require the discounted numeraire asset is a Q-martingale.

We examine the change of measure from Q to Q^S as effected by L_t.

Symbolically, we write the Radon-Nikodym derivative as

$$L_t = \left.\frac{dQ^S}{dQ}\right|_{\mathcal{F}_0}, \; t \in (0, T].$$

Under the risk neutral measure Q, the dynamics of S_t is governed by

$$\frac{dS_t}{S_t} = (r - q)\,dt + \sigma\,dZ_t^Q, \qquad Z_t^Q \text{ is } Q\text{-Brownian.}$$

The solution to S_t is given by

$$S_t = S_0 e^{\left(r-q-\frac{\sigma^2}{2}\right)t+\sigma Z_t^Q}$$

so that

$$L_t = e^{qt}\frac{S_t}{S_0}\bigg/ e^{rt} = e^{-\frac{\sigma^2}{2}t+\sigma Z_t^Q}, \; t \in (0, T].$$

Recall from p.41, Topic 3 that under the Radon-Nikodym derivative:

$$\frac{d\widetilde{P}}{dP} = e^{-\frac{\gamma^2}{2}t-\gamma Z_t^P},$$

the Brownian motion with drift defined by $\widetilde{Z}_P(t) = Z_P(t) + \gamma t$ becomes $\widetilde{P}$-Brownian. Here, $Z_P(t)$ is P-Brownian.

In the current context, this corresponds to the choice of $\gamma = -\sigma$ in the Radon-Nikodym derivative. We then deduce that

$$Z_t^{Q^S} = Z_t^Q - \sigma t \text{ is a } Q^S\text{-Brownian.}$$

We commonly call Q^S to be the share measure with respect to S_t.

As a check, we recall

$$
\begin{aligned}
V_0 = e^{-rT} E_Q[V_T(S_T)|\mathcal{F}_0] &= e^{-rT} E_{Q^S}\left[V_T(S_T) \left/ \frac{\mathrm{d}Q^S}{\mathrm{d}Q} \right| \mathcal{F}_0 \right] \\
&= e^{-rT} E_{Q^S}\left[V_T(S_T) e^{rT} S_0 \left/ e^{qT} S_T \right| \mathcal{F}_0 \right] \\
&= S_0 E_{Q^S}\left[\left. \frac{V_T(S_T)}{S_T e^{qT}} \right| \mathcal{F}_0 \right],
\end{aligned}
$$

so that

$$
\frac{V_0}{\widehat{S}_0} = E_{Q^S}\left[\frac{V_T(S_T)}{\widehat{S}_T} \right], \qquad \text{where } \widehat{S}_T = e^{qT} S_T \text{ and } \widehat{S}_0 = S_0.
$$

This verifies that $V_t/\widehat{S}_t$ is Q^S-martingale.

Write Q^X as the share measure with respect to the asset price process X_t. We have shown that

$$Z_{X,t}^{Q^X} = Z_{X,t}^{Q} - \sigma_X t$$

is Q^X-Brownian.

Since X_t and Y_t have correlation coefficient ρ, we expect that subtracting $\rho\sigma_X t$ from $Z_{Y,t}^{Q}$ would make $Z_{Y,t}^{Q} - \rho\sigma_X t$ to be Q^X–Brownian.

How to verify that $Z_{Y,t}^{Q^X} = Z_{Y,t}^{Q} - \rho\sigma_X t$ is Q^X-Brownian?

By considering the moment generating function, a random variable U is normal with mean m and variance σ^2 under Q^X if and only if

$$E_{Q^X}[\exp(\alpha U)] = \exp\left(\alpha m + \frac{\alpha^2}{2}\sigma^2\right).$$

It suffices to show

$$E_{Q^X}[\exp(\alpha Z_Y^{Q^X}(T))] = E_{Q^X}[\exp(\alpha Z_Y^Q(T) - \alpha\rho\sigma_X T)] = \exp\left(\frac{\alpha^2}{2}T\right).$$

Recall the Radon-Nikodym derivative, where

$$L_T = \frac{dQ^X}{dQ} = \exp\left(-\frac{\sigma_X^2}{2}T + \sigma_X Z_X^Q(T)\right), \text{ so}$$

$$\begin{aligned}
& E_{Q^X}[\exp(\alpha Z_Y^{Q^X}(T))] \\
= \ & E_Q\left[\exp(\alpha Z_Y^Q(T) - \alpha\rho\sigma_X T)\exp\left(-\frac{\sigma_X^2}{2}T + \sigma_X Z_X^Q(T)\right)\right] \\
= \ & \exp\left(-\frac{\sigma_X^2}{2}T - \alpha\rho\sigma_X T\right) E_Q[\exp\left(\alpha Z_Y^Q(T) + \sigma_X Z_X^Q(T)\right)] \\
= \ & \exp\left(-\frac{\sigma_X^2}{2}T - \alpha\rho\sigma_X T\right)\exp\left(\frac{\alpha^2 + 2\rho\alpha\sigma_X + \sigma_X^2}{2}T\right) = \exp\left(\frac{\alpha^2}{2}T\right).
\end{aligned}$$

Derivation of the price formula of an exchange option

Suppose we choose $e^{q_X t}X_t$ as the numeraire, and $M_T/M_0 = e^{rT}$, the corresponding Radon-Nikodym derivative that effects the change from Q to Q^X is given by

$$L_T = e^{(q_X - r)T}\frac{X_T}{X_0}.$$

The price function of the exchange option with maturity T and initial asset values X_0 and Y_0 is given by

$$\begin{aligned} V(X_0, Y_0; T) &= e^{-rT}E_Q[\max(X_T - Y_T, 0)] \\ &= e^{-rT}E_{Q^X}\left[\frac{X_0 e^{(r-q_X)T}}{X_T} X_T \left(1 - \frac{Y_T}{X_T}\right)\mathbf{1}_{\{Y_T/X_T<1\}}\right]. \end{aligned}$$

Note that X_T is canceled. Setting $W_T = Y_T/X_T$, then

$$\frac{V(X_0, Y_0; T)}{X_0} = e^{-q_X T}E_{Q^X}[(1 - W_T)\mathbf{1}_{\{W_T<1\}}].$$

This nice feature of dimension reduction of the option model does not work if the terminal payoff becomes $\max(X_T - Y_T - K, 0)$.

From Ito's lemma, the dynamics of W_t under Q is given by

$$\frac{dW_t}{W_t} = [(r - q_Y) - (r - q_X) - \rho\sigma_X\sigma_Y + \sigma_X^2]\,dt + \sigma_Y\,dZ_{Y,t}^Q - \sigma_X\,dZ_{X,t}^Q.$$

We observe that $Z_{X,t}^{Q^X}$ and $Z_{Y,t}^{Q^X}$ as defined by

$$dZ_{X,t}^{Q^X} = dZ_{X,t}^Q - \sigma_X\,dt \quad \text{and} \quad dZ_{Y,t}^{Q^X} = dZ_{Y,t}^Q - \rho\sigma_X\,dt$$

are Q^X-Brownian motions. The dynamics of W_t under Q^X becomes

$$\frac{dW_t}{W_t} = (q_X - q_Y)\,dt + \sigma_Y\,dZ_{Y,t}^{Q^X} - \sigma_X\,dZ_{X,t}^{Q^X}.$$

We deduce that W_t remains to be a Geometric Brownian motion, and $\sigma_W^2 = \sigma_Y^2 - 2\rho\sigma_X\sigma_Y + \sigma_X^2$ and $\mu_W = q_X - q_Y = (r - q_Y) - (r - q_X)$ under Q^X.

We may write

$$\frac{dW_t}{W_t} = (q_X - q_Y)dt + \sigma_W dZ_{W,t}^{Q^X},$$

where $Z_{W,t}^{Q^X}$ is Q_X-Brownian.

We retain the nice analytical tractability for Y_t/X_t since the ratio of the two lognormal distributions X_t and Y_t remains to be lognormal. However, the difference $X_t - Y_t$ has no nice analytic form of joint distribution function. This explains why we choose to normalize the payoff function by X_T instead of choosing the apparently more obvious choice of $\widetilde{W}_T = X_T - Y_T$.

The payoff $(1 - W_T)\mathbf{1}_{\{W_T<1\}}$ resembles a put payoff with unit strike and underlying W_t. Using the put price formula, we deduce

$$E_{Q_X}[(1 - W_T)\mathbf{1}_{\{W_T<1\}}] = N(d_X) - W_0 e^{(q_X - q_Y)T} N(d_Y), \quad W_0 = \frac{Y_0}{X_0},$$

where

$$d_X = \frac{\ln\frac{X_0}{Y_0} + (q_Y - q_X)T + \frac{\sigma_W^2}{2}T}{\sigma_W\sqrt{T}}, d_Y = \frac{\ln\frac{X_0}{Y_0} + (q_Y - q_X)T - \frac{\sigma_W^2}{2}T}{\sigma_W\sqrt{T}}.$$

Finally, the price function of the exchange option is given by

$$\begin{aligned} V(X_0, Y_0; T) &= e^{-q_X T} X_0 N(d_X) - e^{-q_Y T} Y_0 N(d_Y). \\ &= e^{-rT}\left[e^{(r-q_X)T} X_0 N\left(\frac{\ln\frac{X_0}{Y_0} + \left[(r - q_X) - (r - q_Y) + \frac{\sigma_W^2}{2}\right]T}{\sigma_W\sqrt{T}}\right)\right. \\ &\quad \left. - \, e^{(r-q_Y)T} Y_0 N\left(-\frac{\ln\frac{Y_0}{X_0} + \left[(r - q_Y) - (r - q_X) + \frac{\sigma_W^2}{2}\right]T}{\sigma_W\sqrt{T}}\right)\right]. \end{aligned}$$

Suppose we take Y_t to be the fixed strike price K. By setting $Y_0 = K$, $q_Y = r$ (so the drift rate $r - q_Y$ of Y_t becomes zero) and $\sigma_W^2 = \sigma_X^2$ (since $\sigma_Y = 0$), we recover the usual call price formula.

4.3 Quanto option – equity options with exchange rate risk exposure

- A quanto option is an option on a foreign currency donominated asset but the payoff is in domestic currency.
- The holder of a quanto option is exposed to both exchange rate risk and equity risk. Essentially, quanto option pricing models are two-dimensional with exchange rate and asset price as the pair of state variables.

Some examples of quanto call options are listed below:

1. Foreign equity call struck in foreign currency

$$c_1(S_T, F_T, T) = F_T \max(S_T - X_f, 0).$$

Here, F_T is the terminal exchange rate, S_T is the terminal price of the underlying foreign currency denominated asset and X_f is the strike price in foreign currency.

2. Foreign equity call struck in domestic currency

$$c_2(S_T, T) = \max(F_T S_T - X_d, 0)$$

Here, X_d is the strike price in domestic currency.

3. Fixed exchange rate foreign equity call

$$c_3(S_T, T) = F_0 \max(S_T - X_f, 0)$$

Here, F_0 is some predetermined fixed exchange rate.

4. Equity-linked foreign exchange call

$$c_4(S_T, T) = S_T \max(F_T - X_F, 0).$$

Here, X_F is the strike price on the exchange rate. The holder plans to purchase the foreign asset any way but wishes to place a floor value X_F on the exchange rate. If it happens that the terminal exchange rate F_T shoots above X_F, she receives compensation in exchange rate exposure by benefiting from the positive payoff received through holding the foreign exchange call.

Quanto prewashing technique

Let S_t and F_t denote the stochastic process of the foreign asset price and exchange rate, respectively.

Define $S_t^* = F_t S_t$, which is the foreign asset price in domestic currency.

Let r_d and r_f denote the constant domestic and foreign interest rate, respectively, and let q denote the dividend yield of the foreign asset. Note that the dividend yield is the same in both currency worlds.

We assume that both S_t and F_t follow the Geometric Brownian motion. We attempt to achieve dimension reduction in the pricing model of a quanto option via the quanto prewashing technique. Essentially, it involves adjustment of drift rates under two different currency worlds.

- Under the domestic risk neutral measure Q_d, the drift rate of S^* and F are

$$\delta_{S^*}^d = r_d - q \quad \text{and} \quad \delta_F^d = r_d - r_f.$$

- The reciprocal of F can be considered as the foreign currency price of one unit of domestic currency.

- The drift rate of S and $1/F$ under the foreign risk neutral measure Q_f are given by

$$\delta_S^f = r_f - q \quad \text{and} \quad \delta_{1/F}^f = r_f - r_d,$$

respectively. Note that the dividend yield is the same for the foreign asset in the two-currency world.

- "Quanto prewashing" means finding δ_S^d, that is, the drift rate in the stochastic price process of the foreign currency denominated asset S under the domestic risk neutral measure Q_d.

Quanto-prewashing formula

Let the dynamics of S_t and F_t under Q_d be governed by

$$\begin{aligned}\frac{dS_t}{S_t} &= \delta_S^d\,dt + \sigma_S\,dZ_S^d \\ \frac{dF_t}{F_t} &= \delta_F^d\,dt + \sigma_F\,dZ_F^d,\end{aligned}$$

where $dZ_S^d\,dZ_F^d = \rho\,dt$, σ_S and σ_F are the volatility of S_t and F_t, respectively. Since $S_t^* = F_t S_t$, we obtain from Ito's lemma (see Problem 3 in HW3):

$$\delta_{S^*}^d = \delta_{FS}^d = \delta_F^d + \delta_S^d + \rho\sigma_F\sigma_S.$$

The extra drift rate $\rho\sigma_F\sigma_S$ arises from the correlated diffusion movements of F_t and S_t, where $dZ_S^d\ dZ_F^d = \rho\ dt$. We then obtain

$$\delta_S^d = \delta_{S^*}^d - \delta_F^d - \rho\sigma_F\sigma_S = (r_d - q) - (r_d - r_f) - \rho\sigma_F\sigma_S = r_f - q - \rho\sigma_F\sigma_S.$$

We obtain $\delta_S^d = \delta_S^f - \rho\sigma_F\sigma_S$. It is necessary to add the quanto prewashing term $-\rho\sigma_F\sigma_S$ when we specify the dynamics of S_t changing from Q_f to Q_d.

Siegel's paradox $\delta^d_{1/F} = r_f - r_d + \sigma_F^2 = \delta^f_{1/F} + \sigma_F^2$

Given that the dynamics of F_t under Q_d is

$$\frac{dF_t}{F_t} = (r_d - r_f)\,dt + \sigma_F\,dZ_d,$$

then the dynamics of $1/F_t$ under Q_d is (see Problem 3 in HW3)

$$\frac{d(1/F_t)}{1/F_t} = (r_f - r_d + \sigma_F^2)\,dt - \sigma_F\,dZ_d.$$

To show an alternative proof of the dynamic equation for $1/F_t$, we observe that F_t admits the solution as exponential Brownian motion:

$$F_t = F_0 \exp\left(\left(r_d - r_f - \frac{\sigma_F^2}{2}\right)t + \sigma_F Z_d\right)$$

Taking the reciprocal, we obtain

$$\frac{1}{F_t} = \frac{1}{F_0}\exp\left(\left(r_f - r_d + \sigma_F^2 - \frac{\sigma_F^2}{2}\right)t - \sigma_F Z_d\right).$$

Working in the reverse manner, from the knowledge of the solution, we can deduce the governing stochastic differential equation for $1/F_t$ as given in the above.

This is seen as a puzzle to many people since the risk neutral drift rate for $1/F$ is expected to be $r_f - r_d$ instead of $r_f - r_d + \sigma_F^2$.

We observe directly from the above SDE's for $1/F_t$ that

$$\sigma_F = \sigma_{1/F} \quad \text{and} \quad \rho_{F,1/F} = -1.$$

Note that $\delta^f_{1/F} = r_f - r_d$. This is also consistent with the quanto prewashing technique when it is applied to $1/F$, where the added prewashing term $-\rho\sigma_F\sigma_{1/F}$ becomes $-(-1)\sigma_F^2 = \sigma_F^2$.

From $\delta^d_{1/F} = \delta^f_{1/F} + \sigma_F^2$ and observing $\sigma_F = \sigma_{1/F}$, we deduce that

$$\delta^f_F = \delta^d_F + \sigma_F^2.$$

This result is consistent with the Siegel formula if we interchange the foreign and domestic currency worlds.

Suppose the terminal payoff of an exchange rate option in domestic currency world is $F_T \mathbf{1}_{\{F_T > K\}}$, so that the terminal payoff of the exchange rate option in foreign currency world is $\mathbf{1}_{\{F_T > K\}}$. That is, the option holder receives one unit of foreign currency at maturity when $F_T > K$. Let $V^d(F,t)$ denote the value of the option in the domestic currency world. Obviously, the option value in the foreign currency is $V^f(F_t,t) = V^d(F_t,t)/F_t$.

Considering valuation in different currency world, we obtain

$$\begin{aligned}
V^f(F,t) &= e^{-r_f\tau} E^t_{Q_f}[\mathbf{1}_{\{F_T>K\}} | F_t = F] \\
&= e^{-r_f\tau} N\left(\frac{\ln\frac{F}{K} + (\delta^f_F - \frac{\sigma_F^2}{2})\tau}{\sigma_F\sqrt{\tau}}\right), \\
V^d(F,t) &= e^{-r_d\tau} E^t_{Q_d}[F_T \mathbf{1}_{\{F_T>K\}} | F_t = F] \\
&= e^{-r_d\tau} e^{\delta^d_F\tau} F N\left(\frac{\ln\frac{F}{K} + (\delta^d_F + \frac{\sigma_F^2}{2})\tau}{\sigma_F\sqrt{\tau}}\right).
\end{aligned}$$

Knowing the relations on the drift rates in different currency worlds:

$$\delta_F^d = r_d - r_f \quad \text{and} \quad \delta_F^f = \delta_F^d + \sigma_F^2,$$

we obtain

$$V^d(F,t) = FV^f(F,t) = e^{-r_f\tau}FN(d) = e^{-r_d\tau}e^{\delta_F^d\tau}FN(d)$$

where

$$d = \frac{\ln\frac{F}{K} + \left(\delta_F^f - \frac{\sigma_F^2}{2}\right)\tau}{\sigma_F\sqrt{\tau}} = \frac{\ln\frac{F}{K} + \left(\delta_F^d + \frac{\sigma_F^2}{2}\right)\tau}{\sigma_F\sqrt{\tau}}.$$

These calculations verify that we may perform quanto option valuation in any of the two currency worlds. Be careful with the adjustment of the drift rate in different currency worlds using the quanto prewashing technique.

Price formulas of various quanto options

1. Foreign equity call struck in foreign currency

   Let $c_1^f(S,\tau)$ denote the usual vanilla call option on the foreign currency asset in the foreign currency world. The terminal payoff is

   $$c_1^f(S,0) = \max(S - X_f, 0).$$

   We treat this call as if it is structured in the foreign currency world. Its value can always be converted into domestic currency using the prevailing exchange rate.

$$c_1(S,F,\tau) = Fc_1^f(S,\tau) = F\left[Se^{-q\tau}N(d_1^{(1)}) - X_f e^{-r_f\tau}N(d_2^{(1)})\right],$$

where

$$d_1^{(1)} = \frac{\ln \frac{S}{X_f} + \left(\delta_S^f + \frac{\sigma_S^2}{2}\right)\tau}{\sigma_S\sqrt{\tau}}, \quad d_2^{(1)} = d_1^{(1)} - \sigma_S\sqrt{\tau},$$

$\delta_S^f = r_f - q.$

Note that both the correlation risk ρ and exchange rate risk σ_F do not appear in the price formula! This is reasonable since we allow the exchange rate to float and do not set the exchange rate to some fixed value F_0.

2. Foreign equity call struck in domestic currency

The terminal payoff in domestic currency is

$$c_2(S, F, 0) = \max(S^* - X_d, 0),$$

where $S^* = FS$ is the price of a domestic currency denominated asset. Note that

$$\delta^d_{S^*} = r_d - q \quad \text{and} \quad \sigma^2_{S^*} = \sigma^2_S + 2\rho\sigma_S\sigma_F + \sigma^2_F.$$

The price formula of the foreign equity call is then given by

$$c_2(S, F, \tau) = S^* e^{-q\tau} N(d_1^{(2)}) - X_d e^{-r_d\tau} N(d_2^{(2)}),$$

where

$$d_1^{(2)} = \frac{\ln \frac{S^*}{X_d} + \left(\delta^d_{S^*} + \frac{\sigma^2_{S^*}}{2}\right)\tau}{\sigma_{S^*}\sqrt{\tau}}, \quad d_2^{(2)} = d_1^{(2)} - \sigma_{S^*}\sqrt{\tau}.$$

Note that r_f does not appear since we perform valuation in the domestic currency world and no foreign-denominated asset is involved. However, ρ and σ_F are involved since the volatility of S^* comes into play.

3. Fixed exchange rate foreign equity call

The terminal payoff is denominated in the domestic currency world, so the drift rate δ_S^d of the foreign asset in Q_d should be used. The price function of the fixed exchange rate foreign equity call is given by

$$c_3(S,\tau) = F_0 e^{-r_d\tau}\left[Se^{\delta_S^d\tau}N(d_1^{(3)}) - X_f N(d_2^{(3)})\right],$$

where

$$d_1^{(3)} = \frac{\ln\frac{S}{X_f} + \left(\delta_S^d + \frac{\sigma_S^2}{2}\right)\tau}{\sigma_S\sqrt{\tau}}, \quad d_2^{(3)} = d_1^{(3)} - \sigma_S\sqrt{\tau}.$$

- The price formula does not depend on the exchange rate F since the exchange rate has been chosen to be the fixed value F_0.
- The currency exposure of the call is embedded in the quanto-prewashing term $-\rho\sigma_S\sigma_F$ in δ_S^d. This call option has exposure to both correlation risk and exchange rate risk (in terms of ρ and σ_F, respectively).

4. Equity-linked foreign exchange call

Write the terminal payoff in the form of an exchange option

$$c_4(S, F, 0) = \max(S^* - XS, 0).$$

Taking the two assets to be an exchange XS for S^*, the ratio of the two assets is $\dfrac{S^*}{XS} = \dfrac{F}{X}$ and the difference of the drift rates under Q_d is $\delta^d_{S^*} - \delta^d_S = r_d - q - (r_f - q - \rho\sigma_S\sigma_F) = r_d - r_f + \rho\sigma_F\sigma_S$. By recalling the exchange option price formula, where $\ln\dfrac{X_0}{Y_0}$ becomes $\ln\dfrac{F}{X}$, $r - q_X$ becomes $\delta^d_{S^*}$, $r - q_Y$ becomes δ^d_S, σ_W becomes σ_F, etc., we obtain

$$\begin{aligned} c_4(S, \tau) &= e^{-r_d\tau}\left[S^* e^{\delta^d_{S^*}\tau} N(d_1^{(4)}) - XSe^{\delta^d_S\tau} N(d_2^{(4)})\right] \\ &= Se^{-q\tau}\left[FN(d_1^{(4)}) - Xe^{(r_f - r_d - \rho\sigma_F\sigma_S)\tau} N(d_2^{(4)})\right], \end{aligned}$$

where

$$d_1^{(4)} = \frac{\ln\frac{F}{X} + \left(r_d - r_f + \rho\sigma_F\sigma_S + \frac{\sigma_F^2}{2}\right)\tau}{\sigma_F\sqrt{\tau}}, \quad d_2^{(4)} = d_1^{(4)} - \sigma_F\sqrt{\tau}.$$

Digital quanto option relating 3 currency worlds

$F_{S\backslash U}$ = SGD currency price of one unit of USD currency

$F_{H\backslash S}$ = HKD currency price of one unit of SGD currency

We visualize $F_{S\backslash U}$ as Singaporean currency denominated asset.

Example 1

Digital quanto option payoff: pay one HKD if $F_{S\backslash U}$ is above some strike level K.

Since $F_{S\backslash U}$ can be visualized as a Singaporean asset, the dynamics of $F_{S\backslash U}$ under Q^S is governed by

$$\frac{dF_{S\backslash U}}{F_{S\backslash U}} = (r_{SGD} - r_{USD})\,dt + \sigma_{F_{S\backslash U}}\,dZ^S_{F_{S\backslash U}}.$$

Given $\delta^S_{F_{S\backslash U}} = r_{SGD} - r_{USD}$, how to find $\delta^H_{F_{S\backslash U}}$, the risk neutral drift rate of the SGD asset denominated in Hong Kong dollar?

Taking Hong Kong as the domestic world. Treating $F_{S\backslash U}$ as the foreign asset denominated in Singaporean currency and $F_{H\backslash S}$ as the exchange rate, by the quanto-prewashing technique

$$\delta^H_{F_{S\backslash U}} = \delta^S_{F_{S\backslash U}} - \rho\sigma_{F_{S\backslash U}}\sigma_{F_{H\backslash S}},$$

where $\rho\ dt = dZ^H_{F_{S\backslash U}}\ dZ^H_{F_{H\backslash S}}$. Here, $F_{S\backslash U}$ is visualized as the foreign asset S and $F_{H\backslash S}$ as the exchange rate F in the quanto prewashing formula.

Digital option value $= e^{-r_{HKD}\tau} E^t_{Q^H}\left[\mathbf{1}_{\{F_{S\backslash U} > K\}}\right] = e^{-r_{HKD}\tau} N(d)$

where

$$d = \frac{\ln \frac{F_{S\backslash U}}{K} + \left(\delta^H_{F_{S\backslash U}} - \frac{\sigma^2_{F_{S\backslash U}}}{2}\right)\tau}{\sigma_{F_{S\backslash U}}\sqrt{\tau}}.$$

Example 2

The quanto option pays $F_{H\backslash S}$ Hong Kong dollars when $F_{S\backslash U} > K$. This is equivalent to pay one Singaporean dollar. Value of the quanto option in Singaporean dollar is

$$e^{-r_{SGD}\tau} E^t_{Q^S}\left[\mathbf{1}_{\{F_{S\backslash U}>K\}}\right] = e^{-r_{SGD}\tau} N(\widehat{d})$$

where

$$\widehat{d} = \frac{\ln \frac{F_{S\backslash U}}{K} + \left(\delta^S_{F_{S\backslash U}} - \frac{\sigma^2_{F_{S\backslash U}}}{2}\right)\tau}{\sigma_{F_{S\backslash U}}\sqrt{\tau}}, \quad \delta^S_{F_{S\backslash U}} = r_{SGD} - r_{USD}.$$

This option model is similar to $c_1(S, F, \tau)$, where the option payoff in foreign currency is converted into domestic currency using the prevailing exchange rate at maturity. The most efficient approach is to perform valuation of the option under the foreign currency world. The present value of the quanto option in Hong Kong dollar is $F_{H\backslash S} e^{-r_{SGD}\tau} N(\widehat{d})$.

Example 3

The quanto option pays $F_{H\backslash U}$ Hong Kong dollars when $F_{S\backslash U} > K$. This is equivalent to pay one US dollars.

Method One – Valuation in the Singaporean currency world

Observe that $F_{H\backslash U} = F_{H\backslash S} F_{S\backslash U}$ so that it is like paying $F_{S\backslash U}$ Singaporean dollars (equivalent to one US dollar) when $F_{S\backslash U} > K$. Note that $\delta^S_{F_{S\backslash U}} = r_{SGD} - r_{USD}$.

The present value of the quanto option in Hong Kong dollars is

$$\begin{aligned} F_{H\backslash S} e^{-r_{SGD}\tau} E^t_{Q^S}\left[F_{S\backslash U}\mathbf{1}_{\{F_{S\backslash U} > K\}}\right] &= F_{H\backslash S} e^{-r_{SGD}\tau} e^{(r_{SGD}-r_{USD})\tau} F_{S\backslash U} N(d_1) \\ &= F_{H\backslash U} e^{-r_{USD}\tau} N(d_1) \end{aligned}$$

where

$$d_1 = \frac{\ln \frac{F_{S\backslash U}}{K} + \left(r_{SGD} - r_{USD} + \frac{\sigma^2_{F_{S\backslash U}}}{2} \right)\tau}{\sigma_{F_{S\backslash U}}\sqrt{\tau}}.$$

The discount factor is $e^{-r_{SGD}\tau}$ and multiplication of $F_{H\backslash S}$ converts the Singaporean dollar payoff into Hong Kong dollars payoff.

Method Two – Valuation in the US currency world

The quanto option pays one US dollars when $F_{S\backslash U} > K \Leftrightarrow \frac{1}{K} > \frac{1}{F_{S\backslash U}} = F_{U\backslash S}$. Later, we multiply the option value in US currency by the exchange rate $F_{H\backslash U}$ to convert into Hong Kong dollars.

The present value of the quanto option in Hong Kong dollars is

$$F_{H\backslash U} e^{-r_{USD}\tau} E^t_{Q^U}\left[\mathbf{1}_{\left\{F_{U\backslash S} < \frac{1}{K}\right\}}\right] = F_{H\backslash U} e^{-r_{UDS}\tau} N(-d_2),$$

where

$$d_2 = \frac{\ln \frac{F_{U\backslash S}}{1/K} + \left(r_{USD} - r_{SGD} - \frac{\sigma^2_{F_{U\backslash S}}}{2}\right)\tau}{\sigma_{F_{U\backslash S}}\sqrt{\tau}} = -d_1.$$

By noting $\sigma_{F_{U\backslash S}} = \sigma_{F_{S\backslash U}}$, we can check easily that the quanto option value in Hong Kong dollars using the two approaches agree with each other.

4.4 Implied volatilities and volatility smiles

The difficulties of setting volatility value in the option price formulas lie in the fact that the input value should be the forecast volatility value over the remaining life of the option rather than an estimated volatility value from the past market data of the asset price *(historical volatility)*.

The Black-Scholes model assumes a lognormal probability distribution of the asset price at all future times. Since volatility is the only unobservable parameter in the Black-Scholes model, the model gives the option price as a function of volatility. The Black-Scholes implied volatility $\sigma_{imp}(X,T)$ is the unique solution to

$$V_{market}(X,T) = V^{BS}(S,t;X,T,\sigma_{imp}(X,T)).$$

The above equation is an answer to: What volatility is implied in the observed option prices, if the Black-Scholes model is a valid description of the market conditions?

Remark

Implied volatility is derived based on the Black-Scholes model, and it translates into an option price through the Black-Scholes equation. It is similar to the yield to maturity Y of a bond, where the implied yield is implied from the bond price through the bond price formula. Recall that

$$\text{bond price} = e^{-Y(T-t)}.$$

There is one-to-one correspondence between the implied volatility and option price, same as the correspondence between implied yield and bond price.

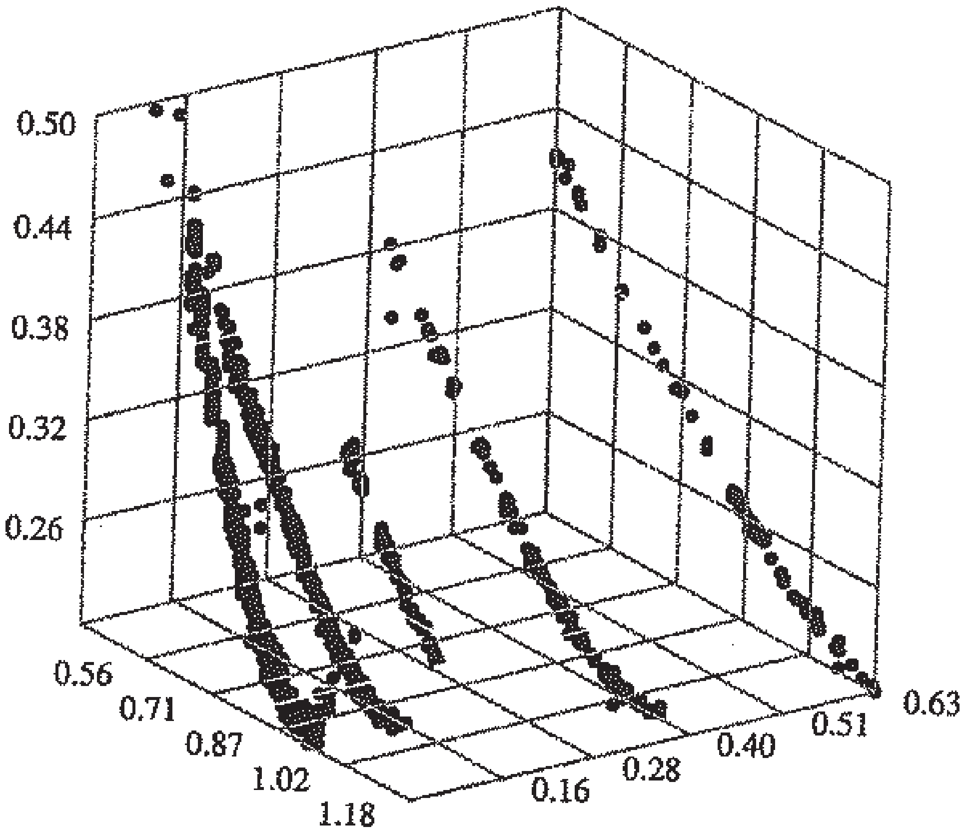
Volatility smiles and volatility term structures

- In financial markets, it becomes a common practice for traders to quote an option's market price in terms of implied volatility σ_{imp}.

- In particular, several implied volatility values obtained simultaneously from different options with varying maturities and strike prices on the same underlying asset provide an extensive market view about the volatility at varying strikes and maturities.

- The Black-Scholes (BS) implied volatility computed from the market option price by inverting the BS price formula varies with the strike price and time to expiration – volatility smile (skew) and volatility term structure, respectively. The plot of the implied volatilities against moneyness (S/K) and time to expiration $T - t$ generates the *implied volatility surface*.

Implied volatility surface

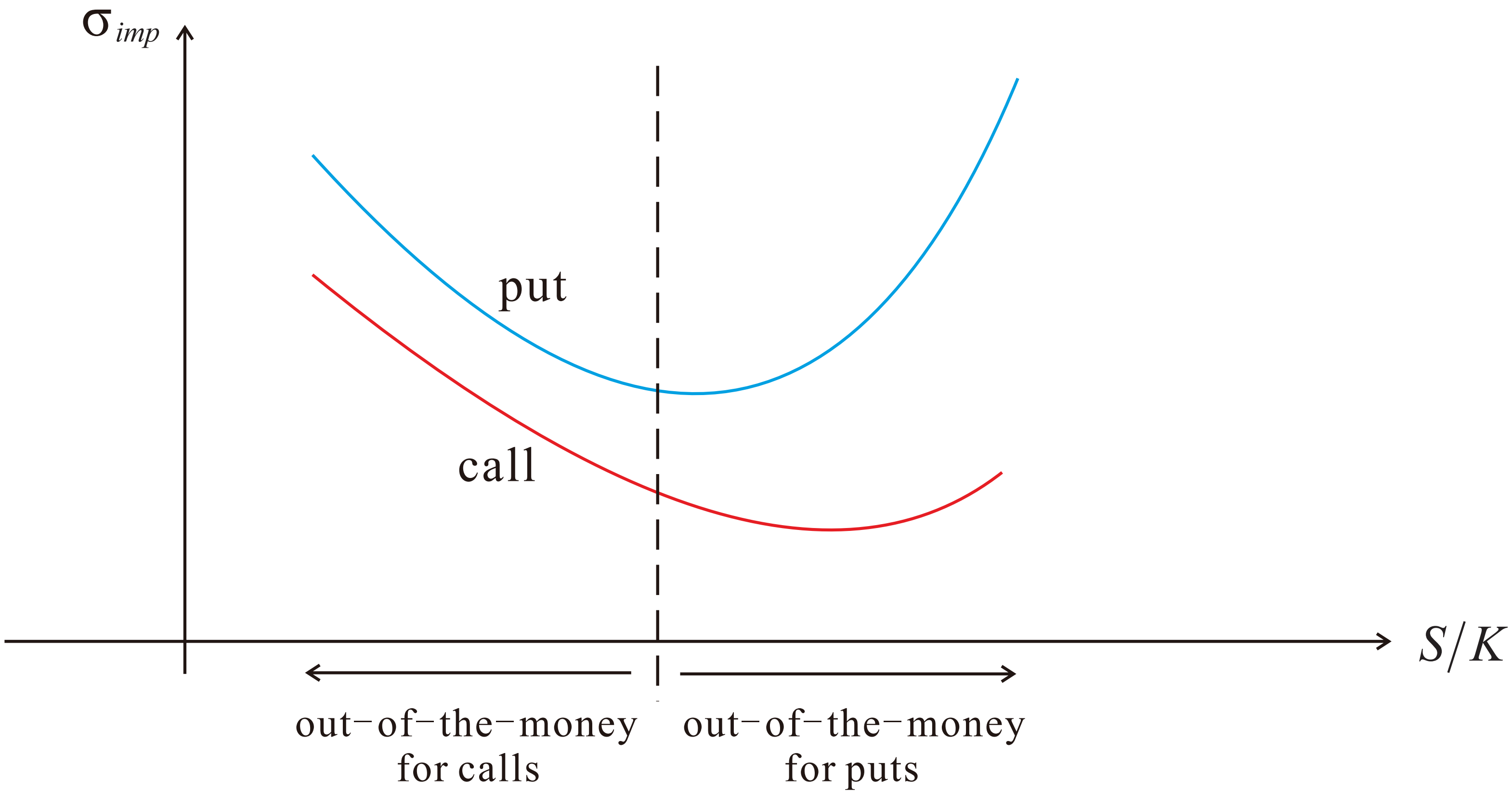


DAX option implied volatilities (as black dots) on 2000/05/02. The lower left axis is moneyness S/K and the right axis is time to expiration measured in years.

- $\sigma_{imp}(X, T)$ is non-linear in strikes and time to expiration; and if observed over in calendar time, it is also time-dependent.

Volatility skew

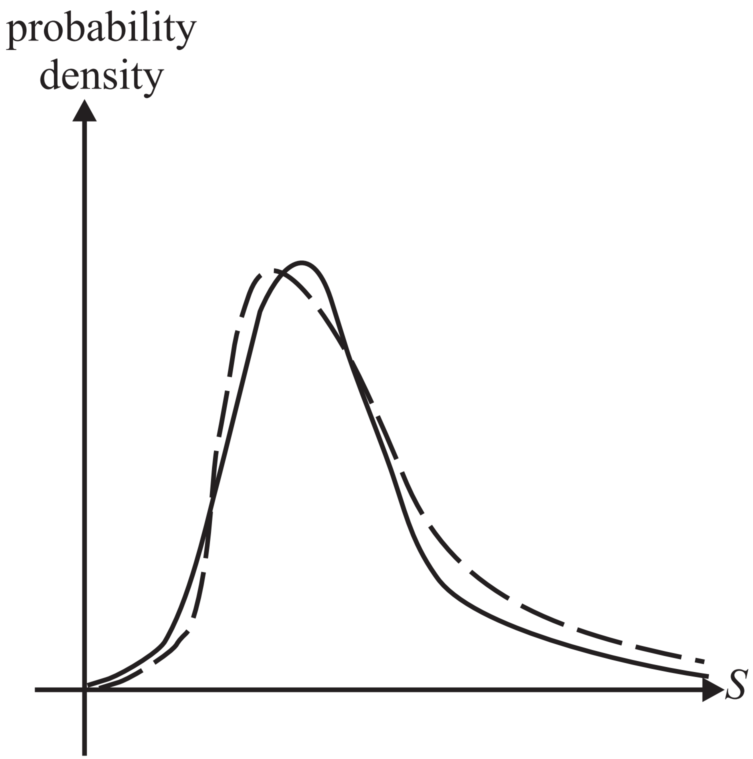
The plots of the implied volatility σ_{imp} against moneyness S/K for traded call and put options with the same maturity date typically show the skew shape as shown in the following figure.



- The volatility skew across moneyness occurs since the option prices for deep out-of-the-money options are bid up higher than around-the-money counterparts. The deep out-of-the-money options have stronger price action since potential gains can be substantial if the price moves from the out-of-the-money region to the in-the-money region.

- Since the stock markets tend to crash downward faster than they move upward, option buyers are willing to bid at higher prices for puts than calls, so σ_{imp} of puts are typically higher than σ_{imp} of calls.

- Investors have stronger motive to use deep out-of-the-money puts to hedge against drastic market decline, so out-of-the-money puts are trading more expensively than out-of-the-money calls for the same level of out-of-the-moneyness (say, $S/K = 0.8$ for a call and $S/K = 1.25$ for a put).



Comparison of the risk neutral probability density of asset price (solid curve) implied from market data and the theoretical lognormal distribution (dotted curve). The risk neutral probability density is thicker at the left tail and thinner at the right tail, indicating that there is a higher change of more acute drop when S is low and a lower chance of further increase when S is high.

Negative correlation between stock price process and volatility process

In real market situation, it is a common occurrence that when the asset price is high, volatility tends to decrease, making it less probable for a higher asset price to be realized. When the asset price is low, volatility tends to increase, that is, it is more probable that the asset price plummets further down. In other words, stock price process and volatility process are in general negatively correlated.

Extreme events in stock price movements

Probability distributions of stock market returns have typically been estimated from historical time series. Unfortunately, common hypotheses may not capture the probability of extreme events. The crash events are rare and may not be present in the historical record.

Examples

1. On October 19, 1987, the two-month S&P 500 futures price fell 29%. Under the lognormal hypothesis of annualized volatility of 20%, this is a -27 standard deviation event with probability 10^{-160} (virtually impossible).

2. On October 13, 1989, the S&P 500 index fell about 6%, a -5 standard deviation event. Under the maintained hypothesis, this should occur only once in $14,756$ years.

Implied volatilities across time

Supply and demand

When markets are very quiet, the implied volatilities of the near month options are generally lower than those of the far month. When markets are very volatile, the reverse is generally true. Why?

- Recall that gamma is the second order derivative of the option price function with respect to the stock price, which is highly dependent on volatility. In very volatile markets, everyone wants or needs to load with gamma so that the portfolio increases in value when volatility increases. Near-dated options provide the most gamma (delta changes most rapidly when the option is around-the-money) and the resulting buying pressure will have the effect of pushing prices up.

- In quiet markets, no one wants a portfolio long of near dated options since the loss of over time over the passage of time is higher for short-lived options.

Implied volatilities for stock options and commodity options

1. *Stock options*

   - In a falling market, everyone needs out-of-the-money puts for insurance and pushes a higher price for the lower strike options due to good demand.
   - Equity fund managers are long billions of dollars worth of stock and writing out-of-the-money call options against their holdings as a way of generating extra income. This pushes the value of out-of-the-money call options down due to good supply.

2. *Commodity options*

   - Government intervention – no worry about a large price fall. Speculators are tempted to sell puts aggressively.
   - Risk of shortages – no upper limit on the price. Demand for higher strike price options.

Term structure of volatility

The Black-Scholes formulas remain valid under time dependent volatility except that $\sqrt{\frac{1}{T-t}\int_t^T \sigma(\tau)^2\,d\tau}$ is used to replace σ. For an option at time t^* with time to expiry $t - t^*$, the substitution of the implied volatility $\sigma_{imp}(t^*, t)$ into the standard Black-Scholes formula under constant volatility gives the option price.

The equivalence of giving the same observed option price by adopting the two different forms of volatility in the two separate option price formulas leads to

$$\int_{t^*}^{t} \sigma(u)^2\,du = \sigma_{imp}^2(t^*, t)(t - t^*).$$

The left hand side is the assumption of time dependent volatility $\sigma(t)$ and the right hand side is the application of the implied volatility formula based on constant volatility in the Black-Scholes pricing formula.

Remark

The implied volatility $\sigma_{imp}(t^*, t)$ may be obtained by averaging the implied volatility values for options at different strikes but with same maturity t. This treatment is acceptable since we focus on the consideration of the term structure of volatility.

How to obtain the term structure of volatility $\sigma(t)$ given the implied volatility measured at time t^* of a European option expiring at time t?

Differentiating with respect to t, we obtain the term structure of volatility in terms of the term structure of implied volatility

$$\sigma(t) = \sqrt{\sigma_{imp}(t^*, t)^2 + 2(t - t^*)\sigma_{imp}(t^*, t)\frac{\partial \sigma_{imp}(t^*, t)}{\partial t}}.$$

It is not easy to compute $\dfrac{\partial \sigma_{imp}(t^*, t)}{\partial t}$ effectively.

Approximation of $\sigma(t)$ as a piecewise constant function

Practically, we do not have a continuous differentiable implied volatility function $\sigma_{imp}(t^*, t)$, but rather implied volatilities are available at discrete instants t_i, $i = 1, 2, \ldots, n$. Suppose we assume $\sigma(t)$ to be piecewise constant over (t_{i-1}, t_i), where $\sigma(t) = \sigma_i$, $t_{i-1} < t < t_i$, $i = 1, 2, \ldots, n$. We then have

$$\begin{aligned}
&\int_{t^*}^{t_i} \sigma^2(\tau)\, d\tau - \int_{t^*}^{t_{i-1}} \sigma^2(\tau)\, d\tau \\
= \;& (t_i - t^*)\sigma_{imp}^2(t^*, t_i) - (t_{i-1} - t^*)\sigma_{imp}^2(t^*, t_{i-1}) \\
= \;& \int_{t_{i-1}}^{t_i} \sigma^2(\tau)\, d\tau = \sigma_i^2(t_i - t_{i-1}), \quad t_{i-1} < t < t_i,
\end{aligned}$$

giving

$$\sigma_i = \sqrt{\frac{(t_i - t^*)\sigma_{imp}^2(t^*, t_i) - (t_{i-1} - t^*)\sigma_{imp}^2(t^*, t_{i-1})}{t_i - t_{i-1}}}, \quad t_{i-1} < t < t_i.$$

Risk neutral density function

- Let $\psi(S_T, T; S_t, t)$ denote the transition density function of the asset price. The time-t price of a European call with maturity date T and strike price X is given by

$$c(S_t, t; X, T) = e^{-r(T-t)} \int_X^\infty (S_T - X)\psi(S_T, T; S_t, t)\ dS_T.$$

- If we differentiate c with respect to X, we obtain

$$\frac{\partial c}{\partial X} = -e^{-r(T-t)} \int_X^\infty \psi(S_T, T; S_t, t)\ dS_T;$$

  and differentiate once more, we have

$$\psi(X, T; S_t, t) = e^{r(T-t)} \frac{\partial^2 c}{\partial X^2}.$$

- Suppose that market European option prices at all strikes are available, the risk neutral density function can be inferred completely from the market prices of options with the same maturity and different strikes, without knowing the volatility function.

Dupire equation and local volatility function

Assuming that the asset price dynamics under the risk neutral measure is governed by

$$\frac{dS_t}{S_t} = (r - q)dt + \sigma(S_t, t)dZ_t,$$

where the local volatility function $\sigma(S_t, t)$ is assumed to have both state and time dependence. This assumption on volatility goes beyond the time dependent volatility.

Suppose we visualize the call price function as a function of X and T, where $c = c(X, T)$, the Dupire equation takes the form

$$\frac{\partial c}{\partial T} = -qc - (r - q)X\frac{\partial c}{\partial X} + \frac{\sigma^2(X, T)}{2}X^2\frac{\partial^2 c}{\partial X^2}.$$

The forward asset price S_T in $\sigma(S_T, T)$ is replaced by X, in par with treating X as an independent variable in $c(X, T)$. The Dupire equation is seen to be the adjoint equation of the Black-Scholes equation.

Proof

We differentiate $\psi(X, T; S_t, t)$ with respect to T to obtain

$$\frac{\partial \psi}{\partial T} = e^{r(T-t)} \left(r \frac{\partial^2 c}{\partial X^2} + \frac{\partial^2}{\partial X^2} \frac{\partial c}{\partial T} \right),$$

and $\psi(X, T; S, t)$ satisfies the forward Fokker-Planck equation, where

$$\begin{aligned} \frac{\partial \psi}{\partial T} &= \frac{\partial^2}{\partial X^2} \left[\frac{\sigma^2(X,T)}{2} X^2 \psi \right] - \frac{\partial}{\partial X} [(r-q) X \psi] \\ &= e^{r(T-t)} \left\{ \frac{\partial^2}{\partial X^2} \left[\frac{\sigma^2(X,T)}{2} X^2 \frac{\partial^2 c}{\partial X^2} \right] - \frac{\partial}{\partial X} \left[(r-q) X \frac{\partial^2 c}{\partial X^2} \right] \right\}. \end{aligned}$$

See Problem 3.8 on P.166 in Kwok's text for a proof of the forward Fokker-Planck equation.

Combining the above equations and eliminating the common factor $e^{r(T-t)}$, we have

$$\begin{aligned} & r\frac{\partial^2 c}{\partial X^2} + \frac{\partial^2}{\partial X^2}\frac{\partial c}{\partial T} \\ = \quad & \frac{\partial^2}{\partial X^2}\left[\frac{\sigma^2(X,T)}{2}X^2\frac{\partial^2 c}{\partial X^2}\right] - \frac{\partial}{\partial X}\left[(r-q)X\frac{\partial^2 c}{\partial X^2}\right]. \end{aligned}$$

Integrating the above equation with respect to X twice, we obtain

$$\begin{aligned} & \frac{\partial c}{\partial T} + rc + (r-q)\left(X\frac{\partial c}{\partial X} - c\right) \\ = \quad & \frac{\sigma^2(X,T)}{2}X^2\frac{\partial^2 c}{\partial X^2} + \alpha(T)X + \beta(T), \end{aligned}$$

where $\alpha(T)$ and $\beta(T)$ are arbitrary functions of T.

Since all functions involving c in the above equation vanish as X tends to infinity, hence $\alpha(T)$ and $\beta(T)$ must be zero (one can check $\lim_{X\to\infty} X\frac{\partial c}{\partial X} = 0$ and $\lim_{X\to\infty} X^2\frac{\partial^2 c}{\partial X^2} = 0$). Grouping the remaining terms in the equation, we obtain the Dupire equation.

From the Dupire equation, we may express the local volatility $\sigma(X, T)$ explicitly in terms of the call price function and its derivatives, where

$$\sigma^2(X,T) = \frac{2\left[\frac{\partial c}{\partial T} + qc + (r-q)X\frac{\partial c}{\partial X}\right]}{X^2\frac{\partial^2 c}{\partial X^2}}. \tag{A}$$

- Suppose a sufficiently large number of market option prices are available at many maturities and strikes, we can estimate the local volatility from the above equation by approximating the derivatives of c with respect to X and T using the market data.

- In real market conditions, market prices of options are available only at limited number of maturities and strikes.

Financial interpretation of local volatility

The local volatility $\sigma(S,T)$ of an asset price process at some future market level S and time T is the future volatility the asset must have at that market level and time in order to make the current option prices fair. Equation (A) dictates the relation between $\sigma(S,T)$ and observable option prices at different future times and market levels.

Relationship between local volatility and implied volatility

Dupire's equation shows how to compute $\sigma_{loc}(X,T)$ from market prices of European options. On the other hand, the market quote prices for European options are in terms of their implied volatilities. One may want to relate $\sigma_{loc}(X,T)$ with $\sigma_{imp}(X,T)$.

We have (see Problem 11 in HW4) the following relation between σ_{loc}^2 and σ_{imp}.

$$\sigma_{loc}^2(X,T) = \frac{\sigma_{imp}^2 + 2T\sigma_{imp}\frac{\partial \sigma_{imp}}{\partial T} + 2(r-q)XT\sigma_{imp}\frac{\partial \sigma_{imp}}{\partial X}}{\left(1 + Xd_1\sqrt{T}\frac{\partial \sigma_{imp}}{\partial X}\right)^2 + X^2 T\sigma_{imp}\left[\frac{\partial^2 \sigma_{imp}}{\partial X^2} - d_1\sqrt{T}\left(\frac{\partial \sigma_{imp}}{\partial X}\right)^2\right]},$$

where

$$d_1 = \frac{\ln\frac{S}{X} + \left[r - q + \frac{\sigma_{imp}^2(X,T)}{2}\right]T}{\sigma_{imp}(X,T)\sqrt{T}}.$$

4.5 Volatility exposure generated by delta hedging options

Delta hedging an option based on some chosen time dependent hedge volatility generates a profit and loss (P&L) that is related to the realized variance and the cash gamma position (defined as product of option gamma and square of asset price).

Consider the underlying asset price which follows an Ito process under a risk neutral measure Q as specified by

$$\frac{\mathrm{d}S_t}{S_t} = (r-q)\mathrm{d}t + \sigma_t \, \mathrm{d}W_t, \tag{1}$$

where r and q are constant riskfree rate and dividend yield, respectively, σ_t is the instantaneous volatility process, and W_t is a standard Brownian motion under Q. The assumed dynamics of S_t allows no jump.

Let σ_t^i be the time dependent implied volatility derived from traded option prices at varying times. We write the time-t option price as $V_t^i = V(S_t, t; \sigma_t^i)$ with reference to implied volatility σ_t^i. Suppose an option trader sells an option at time zero priced at the current market implied volatility σ_0^i, the option price is given by $V_0^i = V(S_0, 0; \sigma_0^i)$.

The seller's short position in the option is delta hedged at some chosen time dependent hedge volatility σ_t^h for the remaining life of the option, that is, by holding $\Delta_t^h = \frac{\partial}{\partial S} V(S_t, t; \sigma_t^h)$ units of the underlying asset at time $t \in [0, T]$, plus a money market account M_t worth $V_t^i - \Delta_t^h S_t$. The net cash amount $V_t^i - \Delta_t^h S_t$ is put into money market account M_t so that $V_t^i = \Delta_t^h S_t + M_t$.

In summary, there are 3 volatilities:

(i) σ_t - Mother Nature's choice

(ii) σ_t^i - market's choice

(iii) σ_t^h - hedger's choice

After selling one unit of the option, the hedger replicates the option by holding Δ_t^h units of stock and money market account. These three instruments constitute the delta hedged portfolio that hedges against the stock price risk. However, variance risk remains since only stock is used as the hedging instrument.

The P&L of the delta hedged portfolio over the infinitesimal time interval $[t, t + \mathrm{d}t]$ consists of the following three components:

- change of the option value from shorting one unit of option: $-\mathrm{d}V_t^i$;
- P&L resulted from the dynamic position of the underlying asset and dividend income: $\Delta_t^h(\mathrm{d}S_t + qS_t\ \mathrm{d}t)$;
- riskfree interest income earned from the money market account: $r(V_t^i - \Delta_t^h S_t)\mathrm{d}t$.

At time t, the hedger adopts his own hedging volatility σ_t^h. We allow the flexibility that the hedger may not choose the implied volatility σ_t^i from the market option price. The hedger computes $V_t^h = V(S_t, t; \sigma_t^h)$ based on the choice of the volatility parameter value σ_t^h in the Black-Scholes option price formula and use V_t^h to compute the hedge ratio $\Delta_t^h = \frac{\partial V_t^h}{\partial S}$. The various Greek parameters $\Theta_t^h = \frac{\partial V_t^h}{\partial t}$, $\Delta_t^h = \frac{\partial V_t^h}{\partial S}$ and $\Gamma_t^h = \frac{\partial^2 V_t^h}{\partial S^2}$ are related by (based on the Black-Scholes equation)

$$\Theta_t^h = -\frac{\Gamma_t^h S_t^2}{2}(\sigma_t^h)^2 + rV_t^h - (r-q)\Delta_t^h S_t. \tag{2a}$$

By applying Itô's lemma to the price function $V_t^h = V(S_t, t; \sigma_t^h)$, we have

$$\begin{aligned} \mathrm{d}V_t^h = &\ \Theta_t^h \mathrm{d}t + \Delta_t^h \mathrm{d}S_t + \frac{1}{2}\frac{\partial^2 V_t^h}{\partial S^2}(\mathrm{d}S_t)^2 \\ = &\ \Delta_t^h \mathrm{d}S_t + \left(\Theta_t^h + \frac{\Gamma_t^h S_t^2}{2}\sigma_t^2\right)\mathrm{d}t, \end{aligned} \tag{2b}$$

where $\mathrm{d}S_t^2 = \sigma_t^2 S_t^2\, \mathrm{d}t$. The diffusion term in (2b) involves σ_t^2 instead of $(\sigma_t^h)^2$ since S_t follows the dynamics as specified in (1).

Substituting Θ_t^h in (2a) into (2b), we obtain

$$\mathrm{d}V_t^h = \Delta_t^h \mathrm{d}S_t + \frac{\Gamma_t^h S_t^2}{2}[\sigma_t^2 - (\sigma_t^h)^2]\mathrm{d}t + r(V_t^h - \Delta_t^h S_t)\mathrm{d}t + q\Delta_t^h S_t \ \mathrm{d}t. \tag{2c}$$

Let Π_t denote the time-t value of the P&L, then $\mathrm{d}\Pi_t$ over $[t, t+\mathrm{d}t]$ is given by the sum of the three components:

$$\mathrm{d}\Pi_t = \Delta_t^h(\mathrm{d}S_t + qS_t \ \mathrm{d}t) + r(V_t^i - \Delta_t^h S_t)\mathrm{d}t - \mathrm{d}V_t^i.$$

We eliminate the hedge ratio term by substituting (2c) into the above equation to obtain

$$\begin{aligned}\mathrm{d}\Pi_t &= (\mathrm{d}V_t^h - \mathrm{d}V_t^i) - r(V_t^h - V_t^i)\mathrm{d}t - \frac{\Gamma_t^h S_t^2}{2}[\sigma_t^2 - (\sigma_t^h)^2]\mathrm{d}t \\ &= e^{-r(T-t)}\frac{\mathrm{d}}{\mathrm{d}t}\left[e^{r(T-t)}(V_t^h - V_t^i)\right]\mathrm{d}t - \frac{\Gamma_t^h S_t^2}{2}[\sigma_t^2 - (\sigma_t^h)^2]\mathrm{d}t.\end{aligned}$$

The total P&L at maturity T is given by the accumulated sum of the forward value of the differential P&L, where

$$\begin{aligned}\Pi_T &= \int_0^T e^{r(T-t)}\,\mathrm{d}\Pi_t \\ &= (V_T^h - V_T^i) - e^{rT}(V_0^h - V_0^i) - \int_0^T e^{r(T-t)}\frac{\Gamma_t^h S_t^2}{2}[\sigma_t^2 - (\sigma_t^h)^2]\,\mathrm{d}t.\end{aligned}$$

At maturity T, both V_T^h and V_T^i become the terminal payoff, where

$$V_T^h = V_T^i = V(S_T).$$

The total P&L at maturity T is

$$\Pi_T = e^{rT}[V(S_0, 0; \sigma_0^i) - V(S_0, 0; \sigma_0^h)] + \int_0^T e^{r(T-t)}\frac{\Gamma_t^h S_t^2}{2}[(\sigma_t^h)^2 - \sigma_t^2]\,\mathrm{d}t.$$

Since the factor $\Gamma_t^h S_t^2$ appears as cash term in Π_T, it is commonly called the *cash gamma* or *dollar gamma* since it is expressible in dollar value.

Remarks

1. The total P&L generated by delta hedging an option at hedge volatility σ_t^h can be decomposed into two parts.

    - The first component is the future value of the time-0 price difference of the two options priced at the implied volatility and the hedge volatility, respectively.

    - The second component arises since the option is hedged at the hedge volatility instead of the realized volatility. It is equal to the future value of the weighted difference of the variance with reference to the hedge volatility and realized variance, and the weight is half of the cash gamma, $\frac{\Gamma_t^h S_t^2}{2}$.

2. Suppose the trader is delta hedging the option at the implied volatility, where $\sigma_t^h = \sigma_t^i$, then the total P&L becomes

$$\Pi_T = \int_0^T e^{r(T-t)} \frac{\Gamma_t^i S_t^2}{2} [(\sigma_t^i)^2 - \sigma_t^2] \, \mathrm{d}t.$$

In this case, the total P&L is equal to the future value of the weighted sum of the difference between the implied variance and realized variance, where the weight factor is half of the cash gamma, $\frac{\Gamma_t^i S_t^2}{2}$. Usually, options are over priced, which means $\sigma_t^i > \sigma_t$. As most options observe the property $\Gamma_t^i > 0$, we then deduce that delta hedging strategy mostly generates positive P&L to the option seller.

Note that the variance exposure associated with the delta hedged option is also dependent on the realized path of S_t. It is still possible to obtain

$$\int_0^T [(\sigma_t^i)^2 - \sigma_t^2] \, \mathrm{d}t > 0,$$

while the P&L can be negative due to the path dependent factor $\frac{\Gamma_t^i S_t^2}{2}$.

4.6 VIX

Characteristics of volatility (hidden stochastic process)

- Volatility is likely to grow when uncertainty and risk increase. May serve as a proxy for market confidence – fear gauge.

- Volatilities appear to revert to the mean (non-linear drift).
  - After a large volatility spike, the volatility can potentially decrease rapidly.
  - After a low volatility period, it may start to increase slowly.

- Volatility is often negatively correlated with stock or index level, and tends to stay high after large downward moves.

- Stock options are impure in terms of volatility exposure. They provide exposure to both direction of the stock price and its volatility. If one hedges an option according to the Black-Scholes prescription, then she can remove the exposure to the stock price. Volatility exposure remains under delta hedging procedure.

Mathematical derivation of VIX

VIX expresses volatility in percentage points. It is calculated as 100 times the square root of the expected 30-day variance of the rate of return of the forward price of the S&P 500 index.

$$\text{VIX} = 100\sqrt{\text{forward price of realized cumulative variance}}$$

Suppose the forward price F_t of the S&P index under Q follows

$$\frac{dF_t}{F_t} = \sigma_t\ dW_t \text{ so that } d\ln F_t = -\frac{\sigma_t^2}{2}\ dt + \sigma_t\ dW_t.$$

Here, the volatility function σ_t is assumed to be stochastic. Note that F_t has zero drift rate under Q since $F_t = e^{(r-q)(T-t)}S_t$ and S_t has the drift rate equals $r - q$ under Q.

The drift term $-\frac{\sigma_t^2}{2}\ dt$ in $d\ln F_t$ arises from the Ito lemma, where

$$\frac{1}{2}\sigma_t^2 F_t^2 (dW_t)^2 \frac{\partial^2}{\partial F_t^2} \ln F_t = -\frac{\sigma_t^2}{2}\ dt.$$

Subtracting the two equations, we obtain the cumulative variance over $[0, T]$ under continuous time model as follows:

$$\frac{dF_t}{F_t} - d\ln F_t = \frac{\sigma_t^2}{2}\ dt, \text{ so } \int_0^T \sigma_t^2\ dt = 2\left[\int_0^T \frac{dF_t}{F_t} - \ln\frac{F_T}{F_0}\right].$$

Our goal is to find the mathematical formula for $E_Q\left[\int_0^T \sigma_t^2\ dt\right]$, visualized as the forward price of the realized cumulative variance over $[0, T]$. Recall that the T-maturity forward price of a risky asset S_t is given by $E_Q[S_T]$. In the current context, the underlying is the random cumulative variance $\int_0^T \sigma_t^2\ dt$.

Note that

$$E\left[\int_0^T \frac{dF_t}{F_t}\right] = E_Q\left[\int_0^T \sigma_t\ \mathrm{d}W_t\right] = 0$$

so that

$$E_Q\left[\int_0^T \sigma_t^2\ \mathrm{d}t\right] = -2E_Q\left[\ln\frac{F_T}{F_0}\right].$$

The log contract with terminal payoff $\ln\frac{F_T}{F_0}$ appears naturally. How to relate the log contract with the usual call and put options?

Technical result For any twice-differentiable function $f\colon \mathbb{R} \to \mathbb{R}$, and any non-negative S_*, we have

$$f(S_T) = f(S_*) + f'(S_*)(S_T - S*) + \int_0^{S_*} f''(K)(K - S_T)^+ \, dK + \int_{S_*}^{\infty} f''(K)(S_T - K)^+ \, dK.$$

The sum of the two integrals is the integral representation of the remainder term in the Taylor expansion of $f(S_T)$ up to the first power term.

Proof

Assuming $x_0 > 0$, we have

$$\int_0^{\xi} \delta(x - x_0) \, dx = \begin{cases} 0 & \text{if } \xi < x_0 \\ 1 & \text{if } \xi > x_0 \end{cases} = \mathbf{1}_{\{x_0 < \xi\}};$$

$$\int_{\xi}^{\infty} \delta(x - x_0) \, dx = \begin{cases} 1 & \text{if } \xi < x_0 \\ 0 & \text{if } \xi > x_0 \end{cases} = \mathbf{1}_{\{x_0 > \xi\}}.$$

To establish the technical results, we perform repeated integration by parts in order to generate the option payoff terms: $(S_T - K)^+$ and $(K - S_T)^+$. For any choice of S_*, we have

$$
\begin{aligned}
f(S_T) =& \int_0^{S_*} f(K)\delta(K - S_T)\ dK + \int_{S_*}^{\infty} f(K)\delta(K - S_T)\ dK \\
=& f(S_*)\mathbf{1}_{\{S_T < S_*\}} - \int_0^{S_*} f'(K)\mathbf{1}_{\{S_T < K\}}\ dK \\
& + f(S_*)\mathbf{1}_{\{S_T \geq S_*\}} + \int_{S_*}^{\infty} f'(K)\mathbf{1}_{\{S_T \geq K\}}\ dK \\
=& f(S_*)\mathbf{1}_{\{S_T < S_*\}} - \left[f'(K)(K - S_T)^+\right]_0^{S_*} + \int_0^{S_*} f''(K)(K - S_T)^+\ dK \\
& + f(S_*)\mathbf{1}_{\{S_T \geq S_*\}} - \left[f'(K)(S_T - K)^+\right]_{S_*}^{\infty} + \int_{S_*}^{\infty} f''(K)(S_T - K)^+\ dK \\
=& f(S_*) + f'(S_*)[(S_T - S_*)^+ - (S_* - S_T)^+] \\
& + \int_0^{S_*} f''(K)(K - S_T)^+\ dK + \int_{S_*}^{\infty} f''(K)(S_T - K)^+\ dK
\end{aligned}
$$

Rearranging the terms and rewriting F_T and F_0 for S_T and S^*, respectively, we have

$$f(F_T) - f(F_0) = f'(F_0)(F_T - F_0) + \int_0^{F_0} f''(K)(K - F_T)^+ dK + \int_{F_0}^{\infty} f''(K)(F_T - K)^+ dK.$$

Here, F_0 is the time-0 forward price of the S&P index, an observable known quantity. Taking $f(F_T) = \ln F_T$, we have

$$\ln \frac{F_T}{F_0} = \frac{F_T - F_0}{F_0} - \int_0^{F_0} \frac{(K - F_T)^+}{K^2}\, dK - \int_{F_0}^{\infty} \frac{(F_T - K)^+}{K^2}\, dK.$$

Recall $E_Q[F_T] = F_0$ so that $E_Q\left[\frac{F_T}{F_0} - 1\right] = 0$. We then have

$$E_Q\left[\int_0^T \sigma_t^2\, dt\right] = 2E_Q\left[\int_0^{F_0} \frac{(K - F_T)^+}{K^2}\, dK + \int_{F_0}^{\infty} \frac{(F_T - K)^+}{K^2}\, dK\right].$$

Note that

$$E_Q[(K - F_T)^+] = e^{rT} \text{ put}_K,$$

where put_K is the time-0 price of a put on the S&P index with strike K. At T, forward price and the underlying index coincide in value.

We then obtain

$$E_Q\left[\int_0^T \sigma_t^2 \text{ d}t\right] = 2e^{rT}\left[\int_0^{F_0} \frac{\text{put}_K}{K^2} \text{ d}K + \int_{F_0}^{\infty} \frac{\text{call}_K}{K^2} \text{ d}K\right].$$

The two terms represent continuum of puts whose strikes are below F_0 and calls whose strikes are above F_0, respectively. They represent out-of-the-money options with respect to the current forward price F_0. The CBOE's choice is sensible since out-of-the-money options tend to be more liquid contracts.

Actual implementation of VIX formula

In the actual implementation of replication formula, one has to face with availability of options with discrete number of strikes. Also, only a finite range of strikes of traded options are available. In addition, the option with strike that exactly equals F_0 is not available in general. In the calculation formula for VIX, the CBOE procedure takes the out-of-the-money options within the bounded interval $[K_L, K_U]$, and choose K_0 to be the closest listed strike below F_0. The out-of-the-money options include all listed put options with strikes at or below K_0, and all listed call options with strike at or above K_0.

We remark that F_0 is chosen for mathematical simplicity while K_0 is chosen due to practical implementation.

When K_0 is chosen, we have

$$\ln \frac{F_T}{K_0} = \frac{F_T - K_0}{K_0} - \int_0^{K_0} \frac{(K - F_T)^+}{K^2} \, \mathrm{d}K - \int_{K_0}^{\infty} \frac{(F_T - K)^+}{K^2} \, \mathrm{d}K$$

so that the corresponding expectation admits two representations:

$$\begin{aligned} E_Q\left[\ln \frac{F_T}{K_0}\right] &= E_Q\left[\ln \frac{F_T}{F_0}\right] + \ln \frac{F_0}{K_0} \\ &= \frac{F_0}{K_0} - 1 - \int_0^{K_0} \frac{e^{rT}\mathrm{put}_K}{K^2} \, \mathrm{d}K - \int_{K_0}^{\infty} \frac{e^{rT}\mathrm{call}_K}{K^2} \, \mathrm{d}K. \end{aligned}$$

We use the approximation: $\ln(1+x) \approx x - \frac{x^2}{2} + \cdots$, so that

$$\ln \frac{F_0}{K_0} \approx \left(\frac{F_0}{K_0} - 1\right) - \frac{1}{2}\left(\frac{F_0}{K_0} - 1\right)^2.$$

The difference between the choice of F_0 or K_0 amounts to

$$2e^{rT}\left\{\left[\int_0^{F_0}\frac{\text{put}_K}{K^2}\;\mathrm{d}K+\int_{F_0}^{\infty}\frac{\text{call}_K}{K^2}\;\mathrm{d}K\right]\right.$$
$$\left.-\left[\int_0^{K_0}\frac{\text{put}_K}{K^2}\;\mathrm{d}K+\int_{K_0}^{\infty}\frac{\text{call}_K}{K^2}\;\mathrm{d}K\right]\right\}$$
$$=2\left[\ln\frac{F_0}{K_0}-\left(\frac{F_0}{K_0}-1\right)\right]\approx-\left(\frac{F_0}{K_0}-1\right)^2.$$

Finally, we obtain

$$E_Q\left[\int_0^T\;\sigma_t^2\;\mathrm{d}t\right]$$
$$\approx\;2e^{rT}\left[\int_0^{K_0}\frac{\text{put}_K}{K^2}\;\mathrm{d}K+\int_{K_0}^{\infty}\frac{\text{call}_K}{K^2}\;\mathrm{d}K\right]-\left(\frac{F_0}{K_0}-1\right)^2.$$

Finally, we multiply the above expected realized cumulative variance by the product of the annualization conversion factor $\frac{365}{30}$ and percentage point factor 100 to obtain

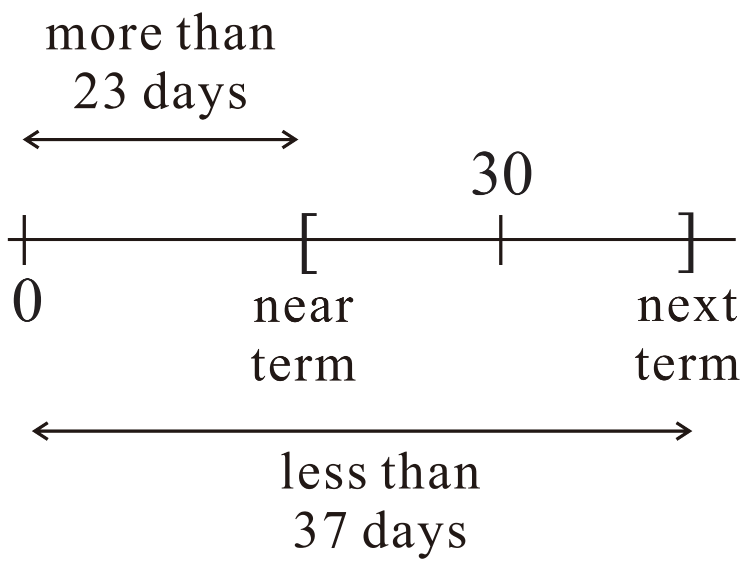
$$\begin{aligned} & \text{VIX}_t^2 \\ = \; & 100^2 \left\{ \frac{2}{30/365} \sum_i \frac{\Delta K_i}{K_i^2} e^{r(30/365)} Q(K_i) - \frac{1}{30/365} \left(\frac{F_0}{K_0} - 1 \right)^2 \right\}. \end{aligned}$$

Here, K_0 is the first strike below the forward index level F_0, $Q(K_i)$ is the time-t out-of-the-money option price with strike K_i.

Implicitly, put (call) options are chosen as out-of-the-money options when the strike price is below (above) K_0.

Linear maturity interpolation

The linear maturity interpolation is another source of approximation error. The VIX is calculated based on options with a fixed 30-day maturity. However, there are generally no options that expire exactly on 30-day maturity. The CBOE calculation procedure finds two maturities: near-term T_1 and next-term T_2 that are closest to the required 30-day maturity T_0. According to the CBOE calculation procedures, the near-term maturity and next-term maturity are set to be more than 23 days and not more than 37 days to expiration, respectively.



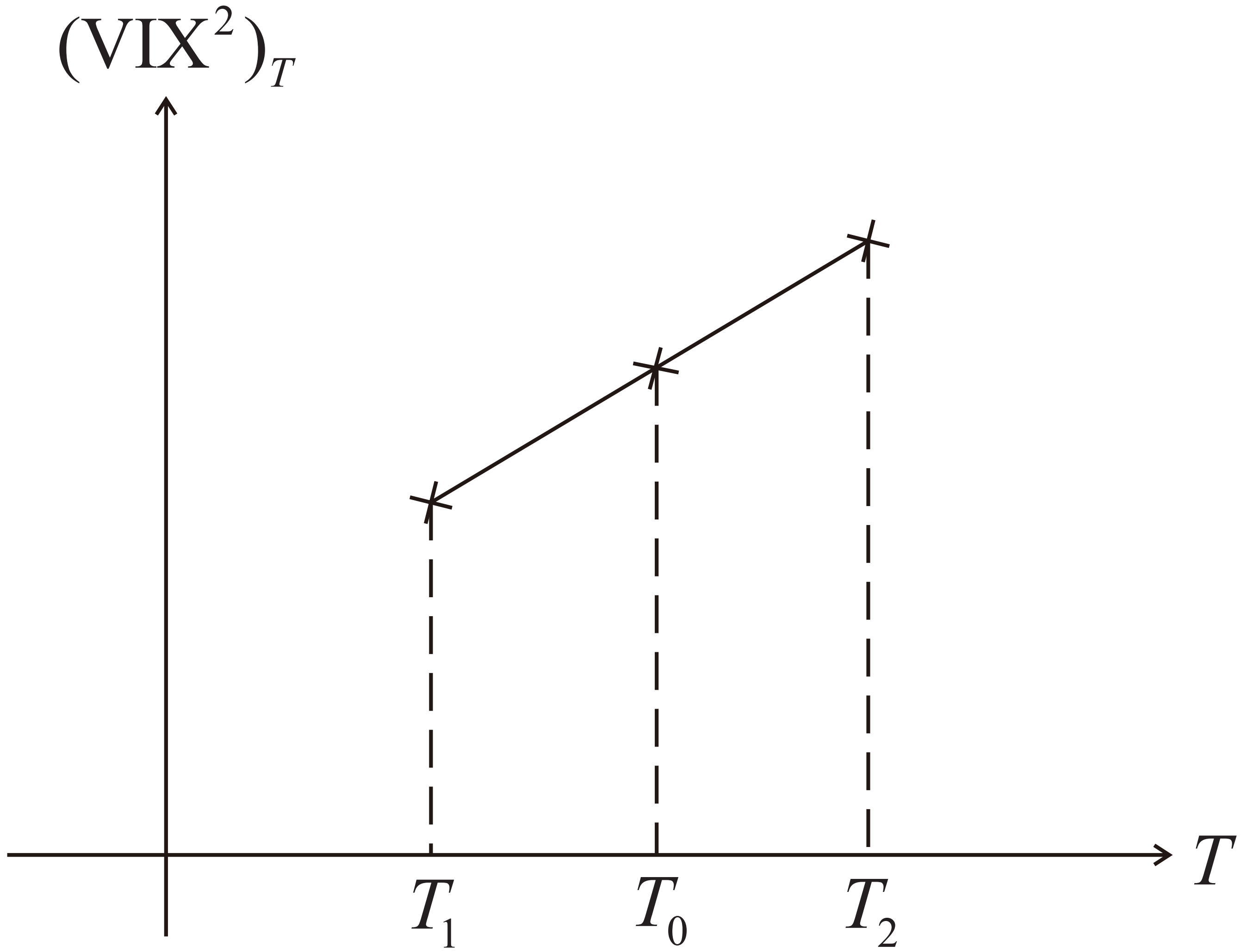
We label the variance measures for maturities T_1 and T_2 as $\widehat{\text{VIX}}^2_{T_1}$ and $\widehat{\text{VIX}}^2_{T_2}$, respectively. The VIX^2 is computed by applying the linear maturity interpolation as follows:

$$\text{VIX}^2 = \frac{1}{30}\left[\alpha T_1 \widehat{\text{VIX}}^2_{T_1} + (1-\alpha) T_2 \widehat{\text{VIX}}^2_{T_2}\right],$$

where the weight α is given by

$$\alpha = \frac{T_2 - T_0}{T_2 - T_1}.$$

Note that $30(\text{VIX})^2$ is the expectation of variance over 30 days, which is approximated by the weighted average of $T_1(\text{VIX})^2$ and $T_2(\text{VIX})^2$.



Say, take $T_0 = 30$, $T_1 = 27$ and $T_2 = 34$. We then have

$$\left(\mathrm{VIX}^2\right)_{T_0}(30) \approx \left(\mathrm{VIX}^2\right)_{T_1}\left(27 \times \frac{34-30}{34-27}\right) + \left(\mathrm{VIX}^2\right)_{T_2}\left(34 \times \frac{30-27}{34-27}\right).$$

Summary: Sources of errors

1. Truncation error

   We choose a bounded truncation interval $[K_L, K_U]$ of strikes instead of the theoretical interval $[0, \infty]$. Note that the CBOE may add new strikes as the underlying S&P 500 index moves. The added strikes expand the truncation interval. During the period of frequent spikes, the expansion of interval can be frequent and significant.

2. Discretization error

   The continuous integration of option prices with respect to continuum of strikes is approximated by the sum of weighted out-of-the-money option prices.

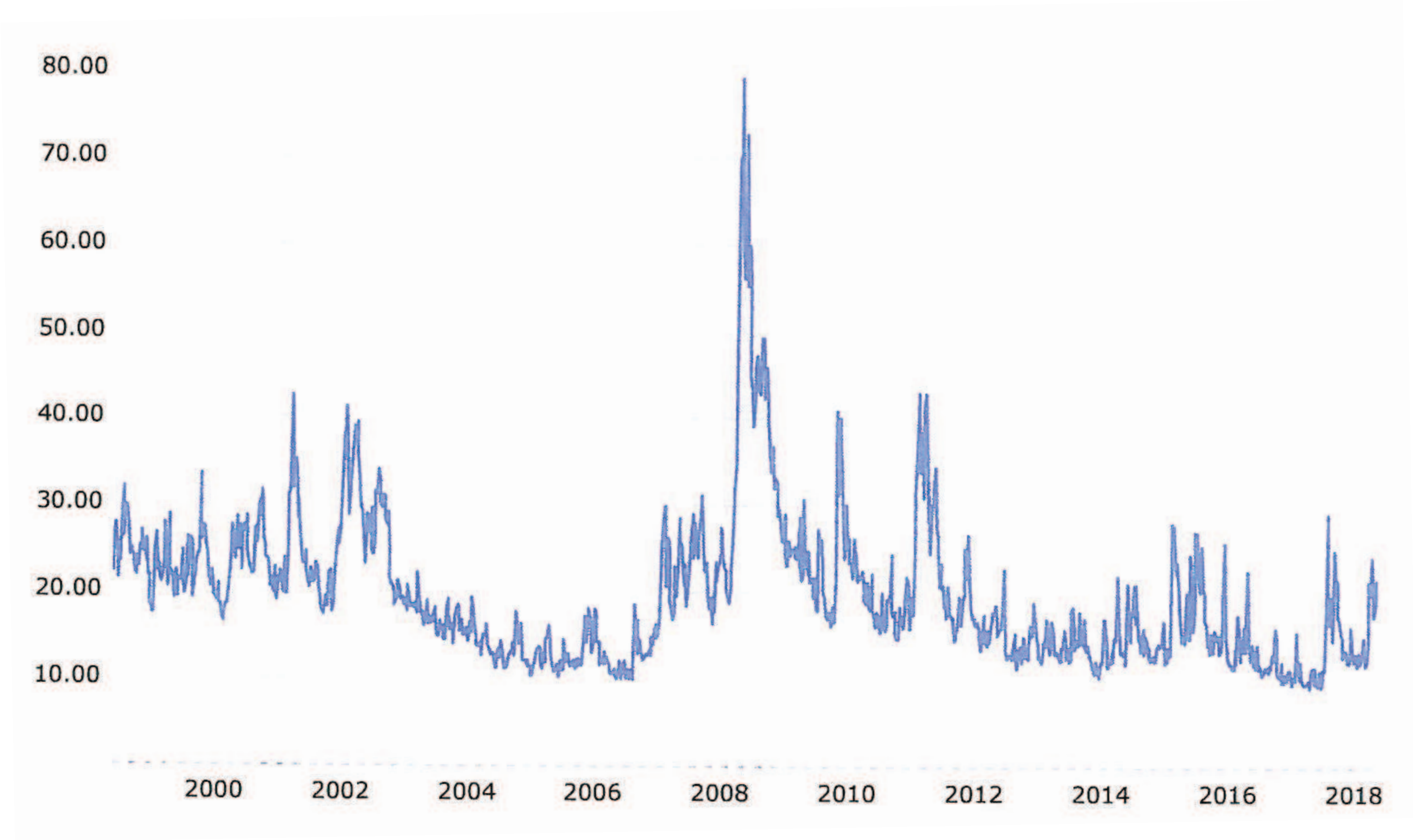
3. Approximation error arising from replacing F_0 by K_0

   The logarithm term $\ln \frac{F_0}{K_0}$ is approximated by the Taylor expansion in powers of $\frac{F_0}{K_0} - 1$ up to the quadratic term.

4. Linear maturity interpolation

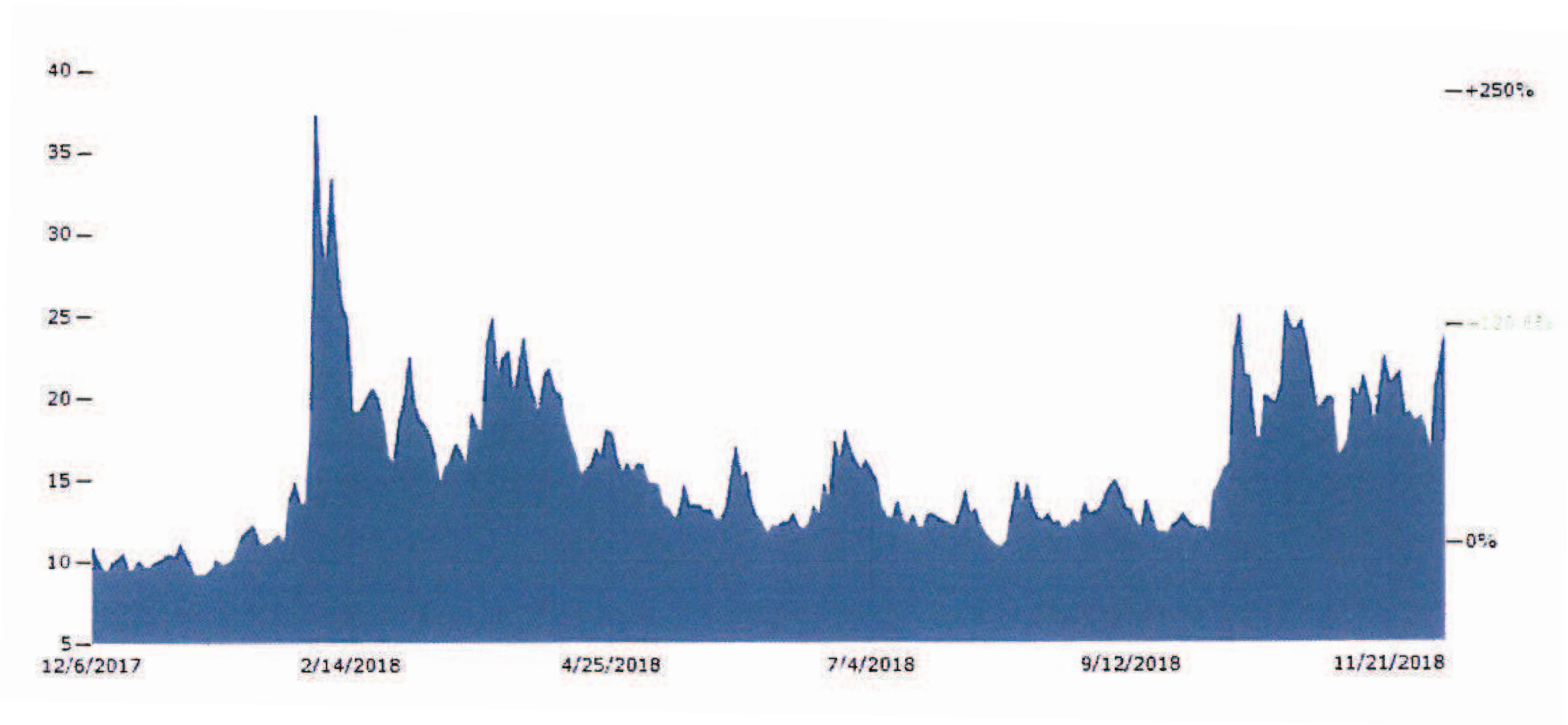
   The VIX is calculated based on options with a fixed 30-day maturity. The CBOE calculation procedure finds two maturities T_1 and T_2 that bracket the required 30-day maturity. The VIX^2 based on 30-day maturity is computed by applying the linear maturity interpolation. Errors are introduced since the model free VIX^2 is a nonlinear function of maturity.

Plot of VIX for the past two decades



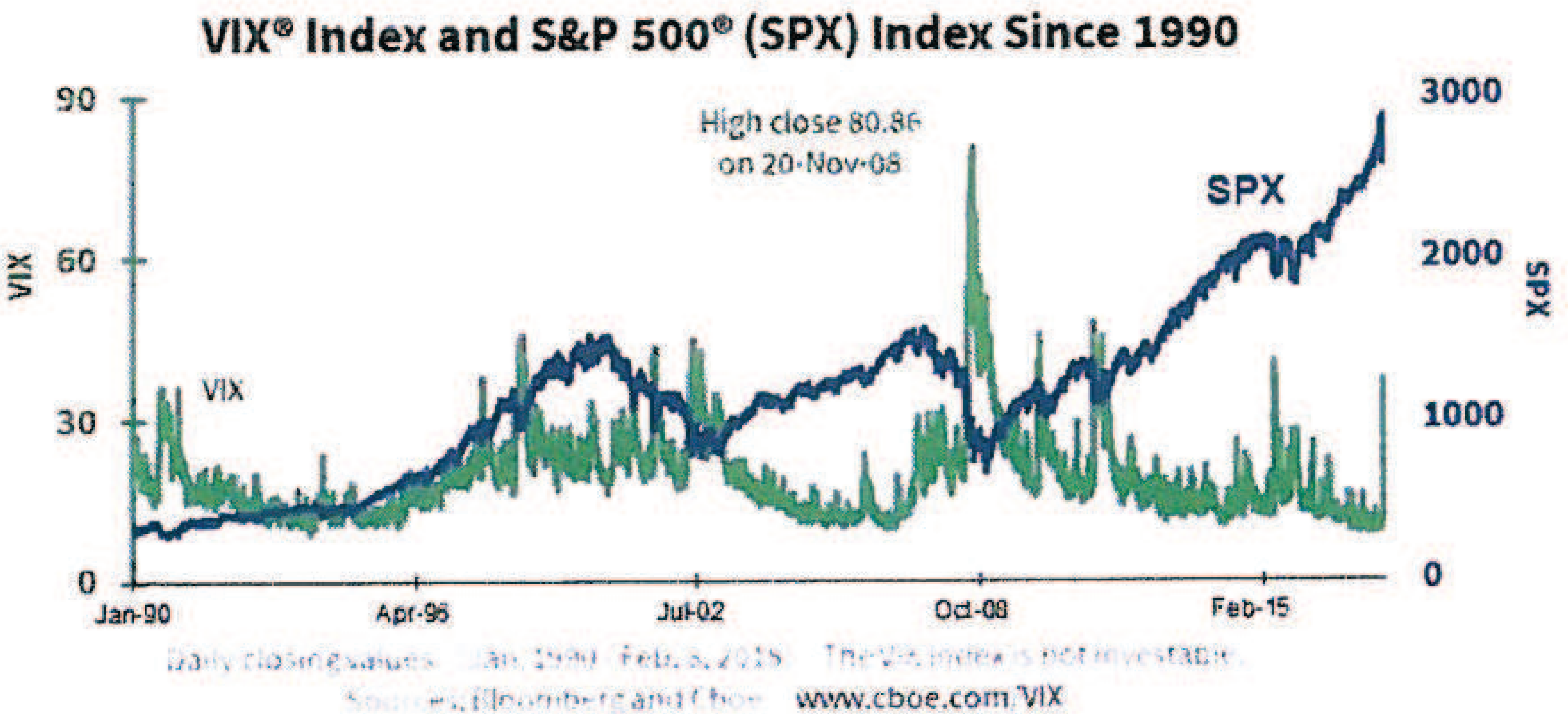
Volatility is bursty in nature: tendency of high volatility to come in bursts. VIX spiked in 2008: period of financial tsunami.

Plot of VIX in 2018



Volatility is negatively correlated to the S&P index return. Fear of inflation that might prompt the Federal Reserve to raise interest rate in February. Deepening of the trade war between China and US pushes up the VIX since September. Reversion of the long-term and short-term yield curves shocked the market on December 4, 2018.

VIX index and SPX



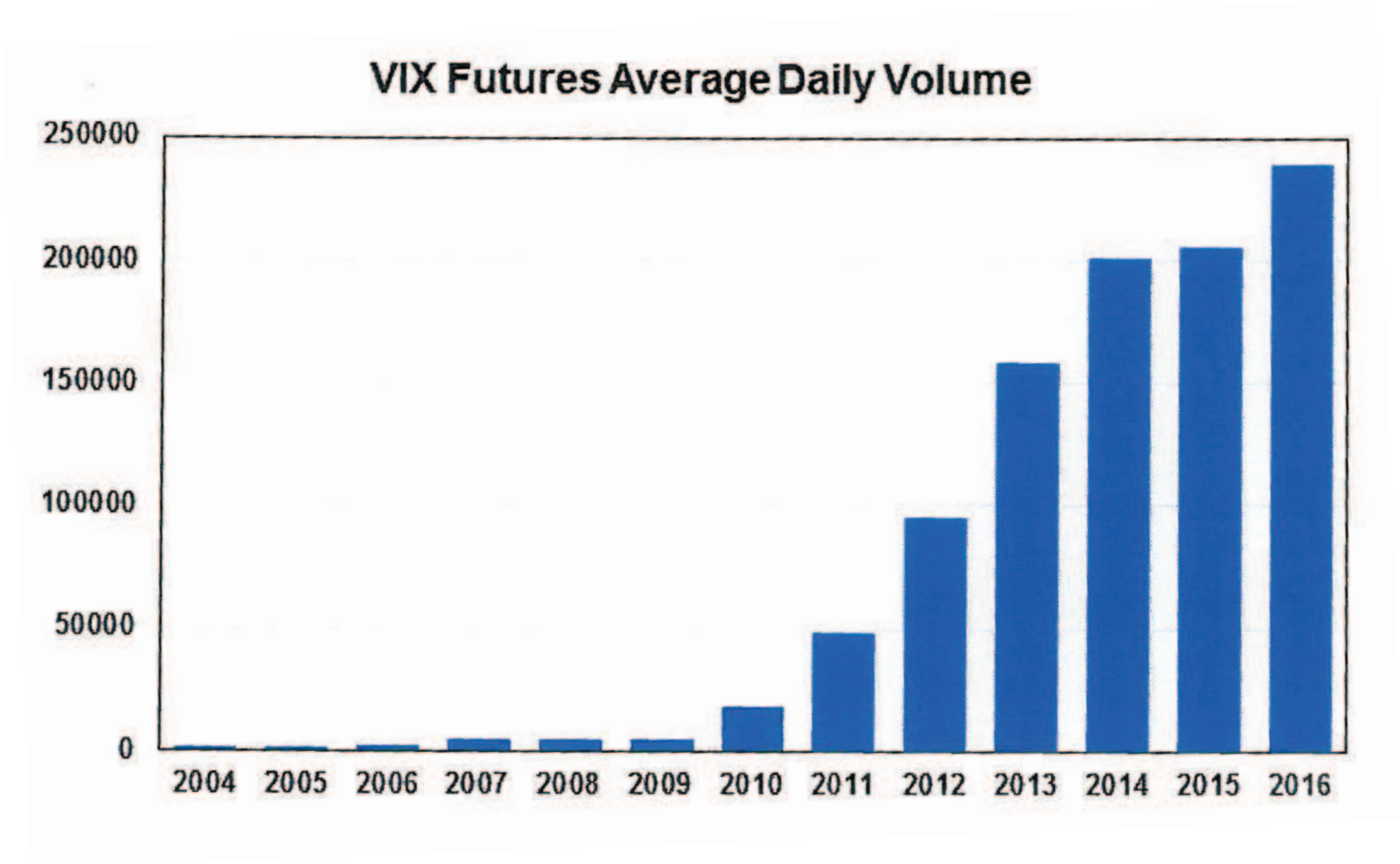
S&P 500 index and VIX are negatively correlated. Growth of SPX over the last 10 years since 2008 financial tsunami with low level of VIX.

VIX derivatives

Investors directly invest in volatility as an asset class by mean of VIX derivatives. CBOE began trading in futures on VIX on March 26, 2004 and European options on February 24, 2006. Both are cash settled. The contract multiplier for each VIX futures contract is \$1000, while that of VIX option is \$100.

VIX derivatives can be used to hedge the risks of investments in the S&P 500 index and/or achieve exposure to S&P 500 volatility without having to delta hedge their S&P 500 option positions with the stock index.

Trading volumes of VIX futures



On June 24, 2016, in reaction to the Brexit referendum, over 721,000 VIX futures contracts changed hands and on November 9, 2016, the volume was 644,892 in reaction to the surprise outcome of the US election.

Increasing popularity of VIX futures and options

Year-to-date through the end of July, 2017, average daily volume in VIX futures was 283,342 contracts, 20 percent ahead of the same period a year ago.

In VIX options at CBOE, a reported 2,562,477 contracts traded on August 10, 2017 , surpassing the previous single-day record of 2,382,752 contracts on February 3, 2014.

Year-to-date through the end of July, 2017, average daily volume in VIX options was 687,181 contracts, 11 percent ahead of the same period a year ago.

4.7 Guaranteed minimum withdrawal benefit

Product Nature

- Variable annuities — deferred annuities that are fund-linked.
- The single lump sum paid by the policyholder at initiation is invested in a portfolio of funds chosen by the policyholder — equity participation.
- The GMWB allows the policyholder to withdraw funds on an annual or semi-annual basis until the entire principal is returned.
- In 2004, 69% of all variable annuity contracts sold in the US included the GMWB option.

Numerical example

- Let the initial fund value be $\$100,000$ and the withdrawal rate be fixed at 7% per annum. Suppose the investment account earns 10% in the first two years but earns returns of -60% in each of the next three years.

Year	Rate of return during the year	Fund value before withdrawals	Amount withdrawn	Fund value after withdrawals	Guaranteed withdrawals remaining balance
1	10%	$110,000$	$7,000$	$103,000$	$93,000$
2	10%	$113,300$	$7,000$	$106,300$	$86,000$
3	-60%	$42,520$	$7,000$	$35,520$	$79,000$
4	-60%	$14,208$	$7,000$	$7,208$	$72,000$
5	-60%	$2,883$	$7,000$	0	$65,000$

- At the end of year five before any withdrawal, the fund value $\$2,883$ is not enough to cover the annual withdrawal payment of $\$7,000$.

The guarantee kicks in when the fund is non-performing

The value of the fund is set to be zero and the policyholder's 10 remaining withdrawal payments are financed under the writer's guarantee. The policyholder's income stream of annual withdrawals is protected irrespective of the market performance of the underlying fund. The investment account balance may have shrunk to zero before the principal is repaid and will remain there once the account balance hits zero.

Good performance of the fund

If the market does well, then there will be funds left at policy's maturity. The residual fund value will be paid back to the policyholder.

Numerical example revisited

Suppose the initial lump sum investment of $\$100,000$ is used to purchase 100 units of the mutual fund, so each unit worths $\$1,000$.

- After the first year, the rate of return is 10% so each unit is $\$1,100$. The annual guaranteed withdrawal of $\$7,000$ represents $\$7,000/\$1,100 = 6.364$ units. The remaining number of units of the mutual fund is $100 - 6.364 = 93.636$ units.

- After the second year, there is another rate of return of 10%, so each unit of the mutual fund worths $\$1,210$. The withdrawal of $\$7,000$ represents $\$7,000/\$1,210 = 5.785$ units, so the remaining number of units $= 87.851$.

- There is a negative rate of return of 60% in the third year, so each unit of the mutual fund worths \$484. The withdrawal of \$7,000 represents $\$7,000/\$484 = 14.463$ units, so the remaining number of units $= 73.388$.

- Depending on the performance of the mutual fund, there may be certain number of units remaining if the fund is performing or perhaps no unit is left if it comes to the worst senario.
  - In the former case, the holder receives the guaranteed total withdrawal amount of \$100,000 (neglecting time value) plus the remaining units of mutual funds held at maturity.
  - Even when the mutual fund is non-performing and the account balance falls to zero, the total withdrawal amount of \$100,000 over the whole policy life is guaranteed.

How is the benefit funded?

- Proportional fee on the investment account value
  - for a contract with a 7% withdrawal allowance, a typical charge is around 40 to 50 basis points of proportional fee on the investment account value.
- GMWB can also be seen as a guaranteed stream of 7% per annum plus a call option on the terminal investment account value W_T, $W_T \geq 0$. The strike price of the call is zero.

Static withdrawal model – continuous time model

- The withdrawal rate G (dollar per annum) is fixed throughout the life of the policy.
- When the investment account value W_t ever reaches 0, it stays at this value thereafter (absorbing barrier).

$$\tau = \inf\{t : W_t = 0\}, \quad \tau \text{ is the first passage time of hitting 0.}$$

Under the risk neutral measure Q, starting with $W_0 = w_0$, the dynamics of W_t is governed by

$$\begin{aligned} dW_t &= (r-\alpha)W_t\,dt + \sigma W_t\,dB_t - G\,dt, && t < \tau \\ W_t &= 0, && t \geq \tau \end{aligned}$$

where α is the proportional annual fee charge on the investment account as the withdrawal allowance. When W_t is at low value, small fees are received by the insurer while the value of the guarantee becomes higher. This poses challenges for hedging.

$$\text{policy value} = E_Q\left[\int_0^T Ge^{-ru}\,du\right] + E_Q[e^{-rT}W_T].$$

Surrogate unrestricted process

To enhance analytic tractability, the restricted account value process W_t is replaced by a surrogate unrestricted process $\widetilde{W}_t$ at the expense of introducing optionality in the terminal payoff (zero strike call payoff). Consider the modified unrestricted stochastic process:

$$\begin{aligned} d\widetilde{W}_t &= (r-\alpha)\widetilde{W}_t\,dt - G\,dt + \sigma\widetilde{W}_t\,dB_t, \quad t>0, \\ \widetilde{W}_0 &= w_0. \end{aligned}$$

Solving for $\widetilde{W}_t$, we obtain

$$\widetilde{W}_t = X_t\left(w_0 - G\int_0^t \frac{1}{X_u}\,du\right)$$

where

$$X_t = e^{\left(r-\alpha-\frac{\sigma^2}{2}\right)t+\sigma B_t}.$$

The solution is the unit exponential Brownian motion X_t multiplied by the number of units remaining after depletion by withdrawals.

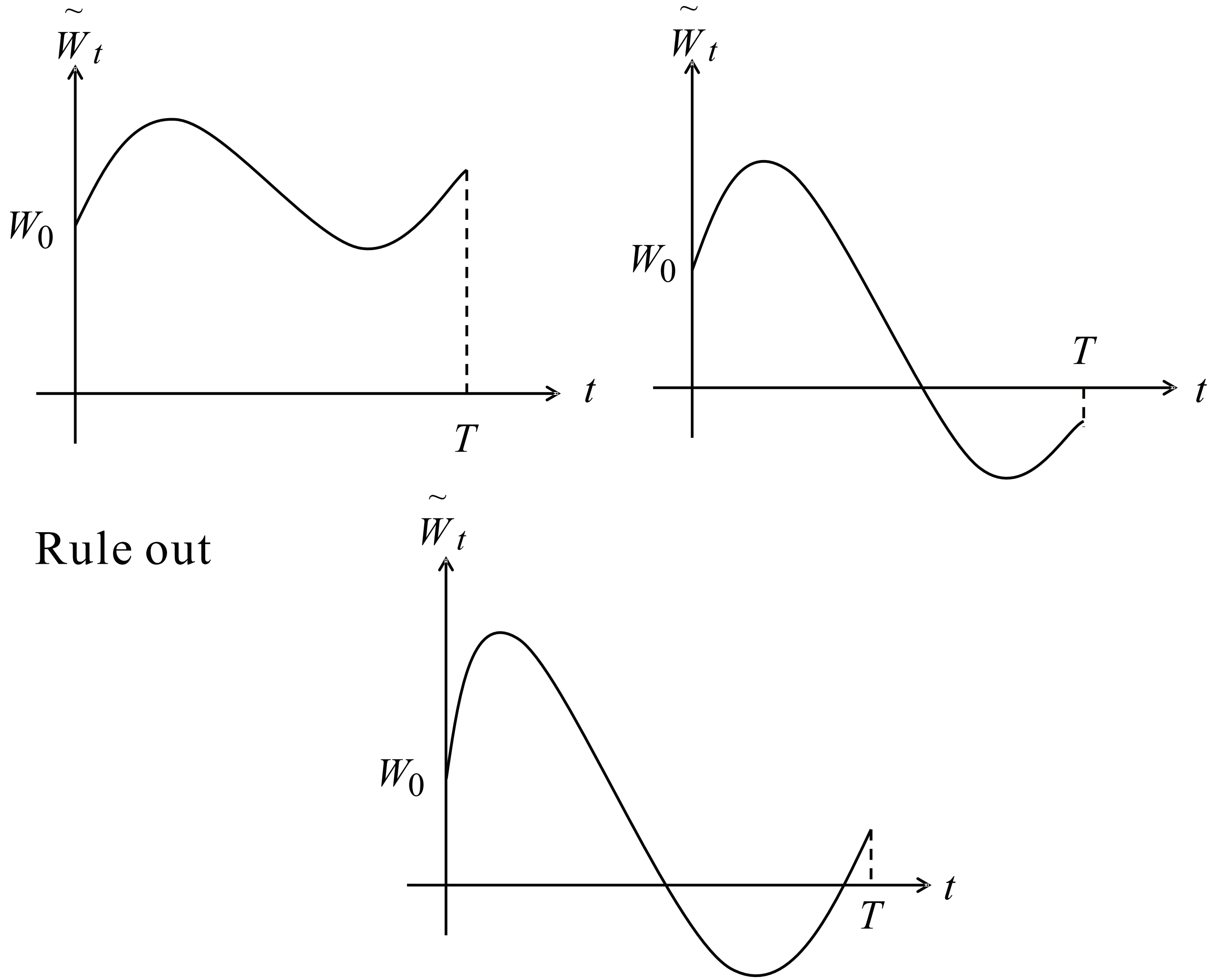
Financial interpretation

Take the initial value of one unit of the fund to be unity for convenience. Here, X_t represents the corresponding fund value process with $X_0 = 1$.

- The number of units acquired at initiation is w_0. Over $(u, u+\mathrm{d}u)$, the number of units withdrawal is $G\ \mathrm{d}u/X_u$. The total number of units withdrawn over $(0, t]$ is given by $G \int_0^t \frac{1}{X_u}\, du$.

- Under the unrestricted process assumption, $\widetilde{W}_t$ may become negative when the number of units withdrawn exceeds w_0. However, in the actual case, W_t stays at the absorbing state of zero value once the number of unit withdrawn hits w_0.

Either $\widetilde{W}_t > 0$ for $t \leq T$ or $\widetilde{W}_T$ remains negative once W_t reaches the negative region at some earlier time prior to T.



Lemma $\tau_0 > T$ if and only if $\widetilde{W}_T > 0$.

$\Longrightarrow$ part. Suppose $\tau_0 > T$, then by the definition of the first passage time, we have $\widetilde{W}_T > 0$.

$\Longleftarrow$ part. Recall that

$$\widetilde{W}_t = X_t \left(w_0 - \int_0^t \frac{G}{X_u}\, du \right)$$

so that

$$\widetilde{W}_t > 0 \quad \text{if and only if} \quad \int_0^t \frac{G}{X_u}\, du < w_0.$$

Suppose $\widetilde{W}_T > 0$, this implies that the number of units withdrawn by time $T = \int_0^T \frac{G}{X_u}\, du < w_0$. Since $X_u \geq 0$, for any $t < T$, we have

$$\text{number of units withdrawn by time } t = \int_0^t \frac{G}{X_u}\, du \leq \int_0^T \frac{G}{X_u}\, du < w_0.$$

Hence, if $\widetilde{W}_T > 0$, then $\widetilde{W}_t > 0$ for any $t < T$.

Intuition of the dynamics

Once the process $\widetilde{W}_t$ becomes negative, it will never return to the positive region. This is because when $\widetilde{W}_t$ increases from below back to the zero level, only the drift term $-G\,dt$ survives. This always pulls $\widetilde{W}_t$ back into the negative region.

Relation between W_T and $\widetilde{W}_T$

Note that $\tau_0 > T \Leftrightarrow \widetilde{W}_T > 0$. We then have

$$W_T = \widetilde{W}_T \mathbf{1}_{\{\tau_0 > T\}} = \widetilde{W}_T \mathbf{1}_{\{\widetilde{W}_T > 0\}} = \max(\widetilde{W}_T, 0).$$

Optionality in the terminal payoff

The terminal payoff from the investment account becomes

$$\max(\widetilde{W}_T, 0) = GX_T \left(\frac{w_0}{G} - \int_0^T \frac{1}{X_u}\,dx \right)^+, \quad x^+ = \max(x, 0).$$

Defining $U_t = \frac{G}{w_0}\int_0^t \frac{1}{X_u}\,du$, which represents the fraction of units withdrawn up to time t. This captures the path dependence of the depletion process of the investment account due to the continuous withdrawal process. We obtain

$$E_Q[e^{-rT}\widetilde{W}_T^+] = w_0 E_Q[e^{-rT} X_T (1 - U_T)^+].$$

Lastly, we have

$$\text{policy value} = E_Q\left[\int_0^T G e^{-ru} du\right] + w_0 E_Q\left[e^{-rT} X_T (1 - U_T)^+\right].$$

The pricing issue is to find the fair value for the participating fee rate α such that the initial policy value equals the lump sum paid upfront by the policyholder so that the policy contract is fair to both counterparties.

4.8 Transaction costs models

How to construct the hedging strategy that best replicates the payoff of a derivative security in the presence of transaction costs?

Recall that one can create a portfolio containing Δ units of the underlying asset and money market account which replicates the payoff of the option. By the portfolio replication argument, the value of an option is equal to the initial cost of setting up the replicating portfolio which mimics the payoff of the option.

Leland proposes a modification to the Black-Scholes model where the portfolio is adjusted at regular time intervals. His model assumes proportional transaction costs where the costs in buying and selling the asset are proportional to the monetary value of the transaction.

Let k denote the round trip transaction cost per unit dollar of transaction. Suppose α units of assets are bought ($\alpha > 0$) or sold ($\alpha < 0$) at the price S, then the transaction cost is given by $\frac{k}{2}|\alpha|S$ in either buying or selling.

We consider a hedged portfolio of the writer of the option, where he is shorting one unit of option and long holding Δ units of the underlying asset. The value of this hedged portfolio at time t is given by

$$\Pi(t) = -V(S,t) + \Delta S,$$

where $V(S,t)$ is the value of the option and S is the asset price at time t. Let δt denote the fixed and small finite time interval between successive rebalancing of the portfolio.

After the small time interval δt, the change in value of the portfolio is

$$\delta \Pi = -\delta V + \Delta \ \delta S - \frac{k}{2}|\delta \Delta| S,$$

where δS is the change in asset price and $\delta \Delta$ is the change in the number of units of asset held in the portfolio.

A cautious reader may doubt why the proportional transaction cost term $-\frac{k}{2}|\delta \Delta| S$ appears in $\delta \Pi$ while the term $S\,\delta \Delta$ is missing.

- The transaction cost term represents the single trip transaction cost paid due to rebalancing of the position in the underlying asset.

- By following the "pragmatic" approach used by Black and Scholes (1973), the number of units Δ is taken to be instantaneously constant.

By Ito's lemma, the change in option value in time δt to leading orders is given by

$$\delta V \approx \frac{\partial V}{\partial S}\,\delta S + \left(\frac{\partial V}{\partial t} + \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2}\right)\delta t.$$

In order to cancel the stochastic terms, one chooses $\Delta = \dfrac{\partial V}{\partial S}$. The change in the number of units of asset in time δt is given by

$$\delta\Delta = \frac{\partial V}{\partial S}(S+\delta S, t+\delta t) - \frac{\partial V}{\partial S}(S,t).$$

From the dynamics of S_t, we observe $\delta S \approx \rho S\delta t + \sigma S\delta Z$. Note that $\delta Z \approx O(\sqrt{\delta t})$ is dominant over $\rho S\delta t$, so the leading order of $|\delta\Delta|$ is found to be

$$|\delta\Delta| \approx \left|\frac{\partial^2 V}{\partial S^2}\right| |\delta S| \approx \sigma S\left|\frac{\partial^2 V}{\partial S^2}\right| |\delta Z|.$$

Formally, we may treat δZ as $\widetilde{x}\sqrt{\delta t}$, where $\widetilde{x}$ is the standard normal variable. The expectation of the reflected Brownian motion $|\delta Z|$ is given by

$$\begin{aligned} E(|\delta Z|) &= 2\left(\int_0^\infty t\frac{1}{\sqrt{2\pi}}\ e^{-t^2/2}\,dt\right)\sqrt{\delta t} \\ &= 2\left(\int_0^\infty \frac{1}{\sqrt{2\pi}}\ e^{-u}\,du\right)\sqrt{\delta t}, \quad u = t^2/2, \\ &= \sqrt{\frac{2}{\pi}}\sqrt{\delta t}. \end{aligned}$$

The risk of loss associated with transaction costs is investor-specific, so it should not be compensated. By the Capital Asset Pricing Model, the hedged portfolio should earn an expected rate of return same as that of a riskless asset. This gives

$$\begin{aligned} E[\delta\Pi] &= \left(-\frac{\partial V}{\partial t} - \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2}\right)\delta t - \frac{k}{2}\sigma S^2\sqrt{\frac{2}{\pi}}\left|\frac{\partial^2 V}{\partial S^2}\right|\sqrt{\delta t} \\ &= r\left(-V + \frac{\partial V}{\partial S}S\right)\delta t. \end{aligned}$$

By putting all the above results together, the above equation can be rewritten as

$$\left(-\frac{\partial V}{\partial t} - \frac{\sigma^2}{2}S^2\frac{\partial^2 V}{\partial S^2} - \frac{\sigma^2}{2}S^2\sqrt{\frac{2}{\pi}}\frac{k}{\sigma\sqrt{\delta t}}\left|\frac{\partial^2 V}{\partial S^2}\right|\right)\delta t = r\left(-V + \frac{\partial V}{\partial S}S\right)\delta t.$$

If we define the Leland number to be $Le = \sqrt{\frac{2}{\pi}}\left(\frac{k}{\sigma\sqrt{\delta t}}\right)$, we obtain

$$\frac{\partial V}{\partial t} + \frac{\sigma^2}{2}\,S^2\frac{\partial^2 V}{\partial S^2} + \frac{\sigma^2}{2}Le\,S^2\left|\frac{\partial^2 V}{\partial S^2}\right| + rS\,\frac{\partial V}{\partial S} - rV = 0.$$

The Leland number is related to the ratio of k and standard deviation of the asset price process over the rebalancing time interval δt.

In the *proportional transaction costs model*, the term $\frac{\sigma^2}{2} Le\, S^2 \left|\frac{\partial^2 V}{\partial S^2}\right|$ is in general non-linear, except when the comparative static $\Gamma = \frac{\partial^2 V}{\partial S^2}$ does not change sign for all S. The transaction cost term is dependent on Γ, where $\Gamma = \frac{\partial \Delta}{\partial S}$ measures the sensitivity of the hedge ratio Δ to the underlying asset price S.

One may rewrite the equation into the form that resembles the Black-Scholes equation

$$\frac{\partial V}{\partial t} + \frac{\tilde{\sigma}^2}{2} S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0,$$

where the modified volatility under transaction costs is given by

$$\tilde{\sigma}^2 = \sigma^2 [1 + Le \, \mathrm{sign}(\Gamma)].$$

The governing equation becomes mathematically ill-posed when $\tilde{\sigma}^2$ becomes negative. This occurs when $\Gamma < 0$ and $Le > 1$.

Modified volatility

It is known that Γ is always positive for the vanilla European call and put options in the absence of transaction costs. If we postulate the same sign behavior for Γ in the presence of transaction costs, then $\widetilde{\sigma}^2 = \sigma^2(1 + Le) > \sigma^2$.

The governing equation then becomes linear under the above assumption so that the Black-Scholes formulas become applicable except that the modified volatility $\widetilde{\sigma}$ is now used as the volatility parameter.

We can deduce $V(S, t)$ to be an increasing function of Le since we expect a higher option value for a high value of modified volatility. Financially speaking, the more frequent the rebalancing (smaller δt) the higher the transaction costs and so the writer of an option should charge higher for the price of the option.

Let $V(S,t;\widetilde{\sigma})$ and $V(S,t;\sigma)$ denote the option values obtained from the Black-Scholes formula with volatility values $\widetilde{\sigma}$ and σ, respectively. The total transaction costs associated with the replicating strategy is then given by

$$\mathcal{T} = V(S,t;\widetilde{\sigma}) - V(S,t;\sigma).$$

When Le is small, $\mathcal{T}$ can be approximated by

$$\mathcal{T} \approx \frac{\partial V}{\partial \sigma}(\widetilde{\sigma} - \sigma).$$

Since $\widetilde{\sigma} = \sigma[1 + Le \, \text{sign}(\Gamma)]^{1/2} \approx \sigma \left[1 + \frac{Le}{2}\text{sign}(\Gamma)\right]$, so $\widetilde{\sigma} - \sigma \approx \frac{k}{\sqrt{2\pi\delta t}}$.

Note that $\frac{\partial V}{\partial \sigma} = \frac{S\sqrt{T-t}e^{-\frac{d_1^2}{2}}}{\sqrt{2\pi}}$ is the same for both call and put options. For $Le \ll 1$, the total transaction costs for either a call or a put is approximately given by

$$\mathcal{T} \approx \frac{kSe^{-\frac{d_1^2}{2}}}{2\pi}\sqrt{\frac{T-t}{\delta t}}.$$