# The VIXation on Options

Austin Clime

January 2023


## Abstract

The VIX Index is a financial instrument designed to give investors the opportunity to bet on the expected volatility of the S&P 500 Index. The CBOE has, in subsequent years, expanded on this construction by similarly calculating indexes which gauge expected volatility on other underlying assets. The VIX itself is not an actively traded product, but investors who wish to bet on volatility can do so in the form of futures and option contracts written on the VIX, which are more efficient than prior methods of doing so in that investors can obtain a portfolio with pure exposure to volatility **[5]**. That is to say derivative instruments, such as options, have impure exposure to volatility by containing exposure to both the volatility and direction of returns on the underlying **[6]**. In what follows, I explain the methodologies behind the construction of the VIX. Section I establishes preliminary knowledge necessary for subsequent sections, section II demonstrates how the VIX can be statically replicated with a portfolio of European options and section III demonstrates a framework for mapping the fair value of expected variance to the probabilities of various terminal states of the underlying asset's prices.


## Section I: Preliminaries

We start by assuming markets behave as described in Black and Scholes (1973) in which asset prices evolve as the result of continuous, frictionless trading of infinitely divisible quantities according to the following geometric Brownian motion:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dz_t \qquad (1.1)$$

where $\mu$ and $\sigma$ are a particular asset's drift and volatility coefficients, respectively, and

$$dz_t \sim \Delta z_t = \sqrt{\Delta t}\,\phi(0, 1) \qquad (1.2)$$

in the limit as $\Delta t \to 0$. Prices of European call options maturing at time $T$, paying $C(T, K) = max(S_T - K, 0)$ upon maturity, are given by

$$C(t, K) = S_t N(d_1) - Ke^{-r(T-t)}N(d_2) \qquad (1.3)$$

where

$$d_1 = \frac{log(S_t/K) + (r+\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{T - t} \qquad (1.4,1.5)$$

. European put options of the same maturity paying $P(T, K) = max(K - S_T, 0)$ are similarly priced according to the following formula:

$$P(t,K) = Ke^{-r(T-t)}N(-d_2) - S_tN(-d_1) \qquad (1.6)$$

.

## Section II: Replicating Variance with Options

The value of the VIX, as mentioned before, is designed to reflect market expected volatility, which is proxied by the fair value of the replicating portfolio whose payoff equals realized volatility. Demeterfi, Derman, Kamal and Zou (1999) show that such a replicating portfolio can be constructed via weighted long positions in European call and put options spanned over a continuum of strike prices. The additive property of variance makes it computationally simpler than volatility, so we will be instead focusing on a replicating portfolio for variance (denoted by $\Pi$), whose payoff is equal to realized volatility squared.

We can think of the periodic realized returns variance ($\sigma^2$), from time 0 to time $T$, as

$$\sigma^2 = \frac{1}{T}\int_0^T \sigma_t^2\, dt \qquad (2.1)$$

and the total realized returns variance ($V$) over that same interval as

$$V = T\sigma^2 \qquad (2.2)$$

where

$$\sigma_t^2 = \frac{1}{N}\int_{i=0}^{N} (r_i - E[r_t])^2 \qquad (2.3)$$

. (2.1) is the payoff we wish to replicate, but since variance is not directly observable in financial markets, we need to transform (2.1) into a function of traded instruments whose prices are directly observable. It would be useful to do so by way of some function $f(S_T, S_*)$ of the terminal stock price and some reference price $S_*$, similar to the payoff structure of a typical forward contract. We can do so by representing returns variance as the difference between an asset's arithmetic and geometric mean returns, multiplied by a factor of two[1].

Let an asset's arithmetic mean return be given by

$$\frac{1}{T}\int_{t=0}^{T} \frac{dS_t}{S_t} \qquad (2.4)$$

and its geometric mean counterpart be given by

$$\frac{1}{T}\int_{t=0}^{T} log(\frac{S_{t+1}}{S_t}) = \frac{1}{T}\, log(\frac{S_T}{S_0}) \qquad (2.5)$$

[1] See appendix A for further explanation.

. Subtracting (2.5) from (2.4) gives us half of the periodic returns variance, and the following therefore gives us the realized variance from (2.1):

$$\Pi_T = \frac{2}{T}\left[\int_{t=0}^{T} \frac{dS_t}{S_t} - log\left(\frac{S_T}{S_0}\right)\right] = \frac{2}{T}\left[\frac{S_T - S_0}{S_0} - log\left(\frac{S_T}{S_0}\right)\right] \quad (2.6)$$

which is the exact payoff we wish to replicate.

To find the fair value of $\Pi_\theta$ for some $\theta$ in the interval $[0, T]$, consider again the relationship between the arithmetic and geometric mean returns given by

$$\frac{1}{\theta}\int_{t=0}^{\theta}\left[\frac{dS_t}{S_t} - d\, log(S_t)\right] = \frac{1}{2}\frac{1}{\theta}\int_{t=0}^{\theta} \sigma_t^2\, dt \quad (2.7)$$

, which is given discreetly by

$$\frac{1}{\theta}\sum_{t=0}^{\theta}\left[\frac{S_{t+1} - S_t}{S_t} - log\left(\frac{S_{t+1}}{S_t}\right)\right] = \frac{1}{\theta}\left[\mu\theta - \theta\left(\mu - \frac{1}{2}\sigma^2\right)\right] = \frac{1}{\theta}\frac{1}{2}\sum_{t=0}^{\theta} \sigma_t^2 \Delta t \quad (2.8)$$

. That is,

$$\frac{1}{\theta}\frac{S_\theta - S_0}{S_0} = \frac{1}{\theta}\left[log\left(\frac{S_\theta}{S_0}\right) + \frac{\theta}{2}\sigma^2\right] \quad (2.9)$$

, and by generalizing (2.8), it can be shown that for any time $\theta$, the realized periodic variance from time $0$ to $T$ is a weighted combination of the realized variance from time $0$ to $\theta$ and the realized variance from time $\theta$ to $T$. More formally,

$$\sigma^2 = \frac{V}{T} = \frac{1}{T}\int_{t=0}^{T} \sigma_{0,T}^2 dt = \frac{1}{T}\left[\int_{t=0}^{\theta} \sigma_{0,\theta}^2 dt + \int_{t=\theta}^{T} \sigma_{\theta,T}^2 dt\right]$$

$$= \frac{1}{T}\left[\theta\sigma_{0,\theta}^2 + (T - \theta)\sigma_{\theta,T}^2\right] = \frac{\theta}{T}\sigma_{0,\theta}^2 + \frac{(T-\theta)}{T}\sigma_{\theta,T}^2 \quad (2.10)$$

. The fair value of $\Pi_\theta$ is therefore a weighted combination of realized periodic variance from $0$ to $\theta$ and the expected periodic variance from time $\theta$ to $T$, which can be expressed as

$$\Pi_\theta = \frac{2}{T}\left[\frac{S_\theta - S_0}{S_0} - log\left(\frac{S_\theta}{S_0}\right)\right] + \frac{2}{T}e^{r(T-\theta)}E\left[\frac{S_T - S_\theta}{S_\theta} - log\left(\frac{S_T}{S_\theta}\right)\right]$$

$$= \frac{2}{T}\left[\mu\theta - \left(\mu\theta - \frac{1}{2}\sigma_{0,\theta}^2\theta\right)\right] + \frac{2}{T}e^{r(T-\theta)}E\left[r(T - \theta) - (T - \theta)\left(r - \frac{1}{2}\sigma_{\theta,T}^2\right)\right]$$

$$= \frac{\theta}{T}\sigma_{0,\theta}^2 + e^{r(T-\theta)}\frac{T-\theta}{T}E\left[\sigma_{\theta,T}^2\right] \quad (2.11)$$

, which leads to the conclusion that

$$\Pi_\theta = E\left[\sigma_{0,T}^2 | t = \theta\right] = \frac{2}{T}\left[\frac{S_\theta - S_0}{S_0} - log\left(\frac{S_\theta}{S_0}\right)\right] + e^{r(T-\theta)}\frac{T-\theta}{T}E\left[\sigma_{\theta,T}^2\right] \quad (2.12)$$

.

The second term on the right-hand side of (2.12) can be constructed model independently with $2/T$ units of a portfolio which is long $1/S_*$ forward contracts, with a fair, risk-neutral reference price of $S_* = E[S_T] = e^{r(T-\theta)}S_\theta$ , and short one log contract, paying $log(S_T/S_*)$ **[6]**. The value of the log contract is a nonlinear function of the stock price, and gives the portfolio its curvature around $S_*$ (see figure 1). Holding $2/T$ units of the previously mentioned portfolio gives $\Pi_0$ an initial variance exposure($\frac{\partial\Pi}{\partial V}|_{t=0}$) of $1, which declines linearly as a function of time, independently of the stock price (see figure 2), or

$$\frac{\partial\Pi}{\partial V}|_{t=\theta} = \frac{T-\theta}{T} \qquad (2.13)$$

,which is consistent with (2.12) **[6]**. The positions in forward and log contracts can both be statically constructed, but the log contract, however, is not a tradable instrument, and needs to be replicated using European options, whose payoffs are linear functions of stock price, of the same desired maturity. With a sufficiently full set of available strike prices, we can adequately approximate the payoff given by (2.6).

At time $t = 0$, we can establish a static position that is long $1/S_0$forward contracts paying $S_T - S_0$at $t = T$. Turning our attention to replicating a short position in a log contract, paying $log(S_T/S_0)$ at $t = T$, we can separate this payout into

$$log(\frac{S_T}{S_0}) = log(\frac{S_*}{S_0}) + log(\frac{S_T}{S_*}) \qquad (2.14)$$

. The first term on the right-hand side of (2.14) is a constant term independent of the terminal price, while the second term needs to be replicated with a portfolio of options. From **[4]**, we know that any continuous function twice differentiable is given generally by

$$f(S_T) = f(S_*) + f'(S_*)[(S_T - S_*)^+ - (S_* - S_T)^+] + \int_0^{S_*} f''(K)(K - S_T)^+ dK + \int_{S_*}^{\infty} f''(K)(S_T - K)^+ dK \quad [2] \qquad (2.15)$$

and by substituting $log(S_T/S_*)$ for $f(S_T)$ in (2.15), we get

$$log(\frac{S_T}{S_*}) = \frac{[(S_T-S_*)^+-(S_*-S_T)^+]}{S_*} - \int_0^{S_*}\frac{(K-S_T)^+}{K^2}dK - \int_{S_*}^{\infty}\frac{(S_T-K)^+}{K^2}\,dK$$

$$= \frac{S_T-S_*}{S_*} - \int_0^{S_*}\frac{(K-S_T)^+}{K^2}dK - \int_{S_*}^{\infty}\frac{(S_T-K)^+}{K^2}\,dK \qquad (2.16)$$

---

[2] See appendix B for further explanation.

**Figure 1: Decomposition of the portfolio $\Pi_T$ into separate positions in forward and log contracts**

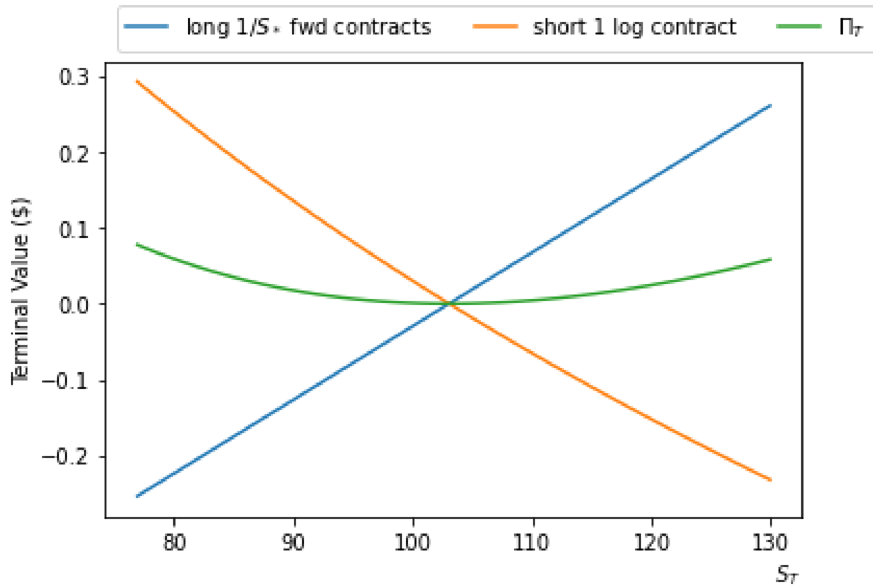
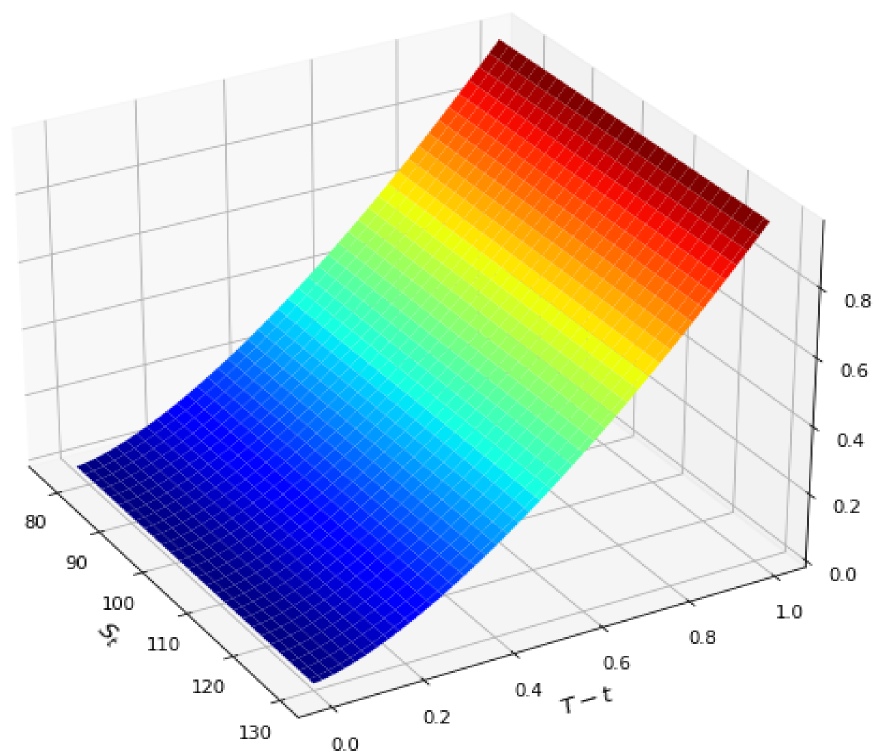



**Figure 2: Variance exposure ($\partial\Pi_t/\partial\sigma^2$) of holding $2/T$ units of a contract paying $log(S_T/S_0)$**



. The first term on the right-hand side of (2.16) represents a position in $1/S_*$ forward contracts with a reference price of $S_*$, while the second term represents a portfolio of short positions in various put options with strikes spanning from 0 to $S_*$, where, for each strike $K$ in $[0, S_*]$, the number of units of $P(T, K)$ to be bought/sold is given by $f''(K) =- 1/K^2$. That is, we need to sell the squared reciprocal of the strike price units of each put option with a strike in $[0, S_*]$. Applying this concept to the third term on the right-hand side of (2.16), it is clear that this position is constructed by shorting $1/K^2$ call options struck at $K$ for all strikes in $[S_*, \infty]$. To see why both portfolios of call and put options are needed, consider that

$$\int_0^{S_*} \frac{(K-S_T)^+}{K^2} dK = \frac{S_T}{S_*} - 1 + log(\frac{S_*}{S_T}) \qquad (2.17)$$

when $S_* > S_T$ and equals 0 otherwise, and in the case where $S_T > S_*$ ,

$$\int_{S_*}^{\infty} \frac{(S_T-K)^+}{K^2} dK = \frac{S_T}{S_*} - \frac{S_T}{max(S_*,S_T)} + log(S_*) - log(max(S_*, S_T))$$

$$= \frac{S_T}{S_*} - 1 + log(\frac{S_*}{S_T}) \qquad (2.18)$$

, and evaluates to 0 otherwise. (2.17) and (2.18) highlight the idea that portfolios of out-of-the-money options are the most cost-effective method for replicating any arbitrary derivative's payoff ($f(S_T)$) and that combining call and put options together guarantees we realize $f(S_T)\ \forall\ S_T \in [0, \infty]$.

Substituting (2.17) and (2.18) into (2.16), we have

$$\frac{S_T-S_*}{S_*} - \int_0^{S_*} \frac{(K-S_T)^+}{K^2} dK - \int_{S_*}^{\infty} \frac{(S_T-K)^+}{K^2} dK = \frac{S_T-S_*}{S_*} - [\frac{S_T}{S_*} - 1 + log(\frac{S_*}{S_T})]$$

$$=- log(\frac{S_*}{S_T}) = log(\frac{S_T}{S_*}) \qquad (2.19)$$

, which is consistent with section II of **[6]**, which states that a short position in a log contract can be constructed by selling $1/S_*$ forward contracts struck at $S_*$ , buying $1/K^2$ put options for all strikes $K$ less than or equal to $S_*$and buying $1/K^2$ call options for all strikes $K$ greater than $S_*$.

Substituting (2.14) and (2.19) into (2.6), we have

$$\Pi_T = \frac{2}{T}[\frac{S_T - S_0}{S_0} - log(\frac{S_*}{S_0}) - log(\frac{S_T}{S_*})]$$

$$= \frac{2}{T}[\frac{S_T - S_0}{S_0} - log(\frac{S_*}{S_0}) - \frac{S_T - S_*}{S_*} + \int_0^{S_*} \frac{(K - S_T)^+}{K^2} dK + \int_{S_*}^{\infty} \frac{(S_T - K)^+}{K^2} dK] \quad (2.20)$$

, which is given more generally, for any $\theta \in [0, T]$ , by

$$\Pi_\theta = E[\sigma^2 | t = \theta] = \frac{2}{T}[\frac{S_\theta - S_0}{S_0} - log(\frac{S_\theta}{S_0})]$$

$$+ e^{r(T-\theta)} \frac{2}{T} E[\frac{S_T}{S_\theta} - 1 - log(\frac{S_*}{S_\theta})]$$

$$- e^{r(T-\theta)} \frac{2}{T} E[\frac{S_T}{S_*} - 1]$$

$$+ \frac{2}{T}[\int_0^{S_*} \frac{P(\theta,K)}{K^2} dK + \int_{S_*}^{\infty} \frac{C(\theta,K)}{K^2} dK] \quad (2.21)$$

. We can simplify (2.21) by the fact that

$$E[\frac{S_T - S_0}{S_0}] = rT = log(\frac{S_0 e^{rT}}{S_0}) \quad (2.22)$$

and using the following Taylor series approximation:

$$log(1 + r) \approx r - \frac{r^2}{2} \quad (2.23)$$

. Doing so gives us

$$E[\Pi_T] = \frac{2}{T}[log(\frac{S_0 e^{rT}}{S_0}) - log(\frac{S_*}{S_0}) - \frac{S_0 e^{rT} - S_*}{S_*} + e^{rT} \int_0^{S_*} \frac{P(0,K)}{K^2} dK + e^{rT} \int_{S_*}^{\infty} \frac{C(0,K)}{K^2} dK]$$

$$= \frac{2}{T}[log(\frac{S_0 e^{rT}}{S_*}) - \frac{S_0 e^{rT} - S_*}{S_*} + e^{rT} \int_0^{S_*} \frac{P(0,K)}{K^2} dK + e^{rT} \int_{S_*}^{\infty} \frac{C(0,K)}{K^2} dK]$$

$$= \frac{2}{T}[log(1 + (\frac{S_0 e^{rT}}{S_*} - 1)) - \frac{S_0 e^{rT}}{S_*} - 1 + e^{rT} \int_0^{S_*} \frac{P(0,K)}{K^2} dK + e^{rT} \int_{S_*}^{\infty} \frac{C(0,K)}{K^2} dK]$$

$$\approx \frac{2}{T}[(\frac{S_0 e^{rT}}{S_*} - 1) - \frac{1}{2}(\frac{S_0 e^{rT}}{S_*} - 1)^2 - (\frac{S_0 e^{rT}}{S_*} - 1)] + \frac{2}{T} e^{rT}[\int_0^{S_*} \frac{P(0,K)}{K^2} dK + \int_{S_*}^{\infty} \frac{C(0,K)}{K^2} dK]$$

$$\approx -\frac{1}{T}[(\frac{S_0 e^{rT}}{S_*} - 1)^2] + \frac{2}{T} e^{rT}[\int_0^{S_*} \frac{P(0,K)}{K^2} dK + \int_{S_*}^{\infty} \frac{C(0,K)}{K^2} dK] \quad (2.24)$$

, which leads to the following general formula for the VIX found in **[5]**:

$$VIX = \sqrt{\frac{2}{T}e^{rT}[\int_0^{S_*} \frac{P(0,K)}{K^2} dK + \int_{S_*}^{\infty} \frac{C(0,K)}{K^2} dK] - \frac{1}{T}[(\frac{S_0 e^{rT}}{S_*} - 1)^2]} \quad (2.25)$$

.

## Section III: Deriving Probabilities from Variance

From Breeden and Litzneberger (1978), we can replicate an elementary derivative that pays $1 if the underlying asset's price at time $T$ ($S_T$) is equal to $K$, and zero otherwise by buying $1/\delta$ units of a butterfly spread which pays $C(T, K - \delta) - 2C(T, K) + C(T, K + \delta)$ in the limit as $\delta \to 0$ **[3]** . More generally, we can replicate any derivative $f$ paying $f(S_T)$ at maturity with the following portfolio:

$$\frac{f(S_T)}{\delta} C(T, S_T - \delta) - 2C(T, S_T) + C(T, S_T + \delta) .$$

. Applying this concept to replicating the payoff given by(2.6), suppose the desired payoff at maturity, given by

$$f(S_T) = \frac{2}{T}[(\frac{S_T - S_0}{S_0}) - log(\frac{S_T}{0})]$$

, is a continuous function of $S_T$ over the terminal price interval $[K(T)_\alpha, K(T)_\beta]$ and let $P = \{K(T)_1, K(T)_2, . . . , K(T)_n\} = \{K(T)_1, K(T)_1 + 2\delta, . . . , K(T)_{n-1} + 2\delta\}$ be a partition of the same interval for some $\delta > 0$. For some $i \in [1, n]$, suppose we construct a butterfly spread, centered at $K(T)_i$ with legs $K(T)_i - \delta$ and $K(T)_i + \delta$, whose payoff at maturity is characterized by

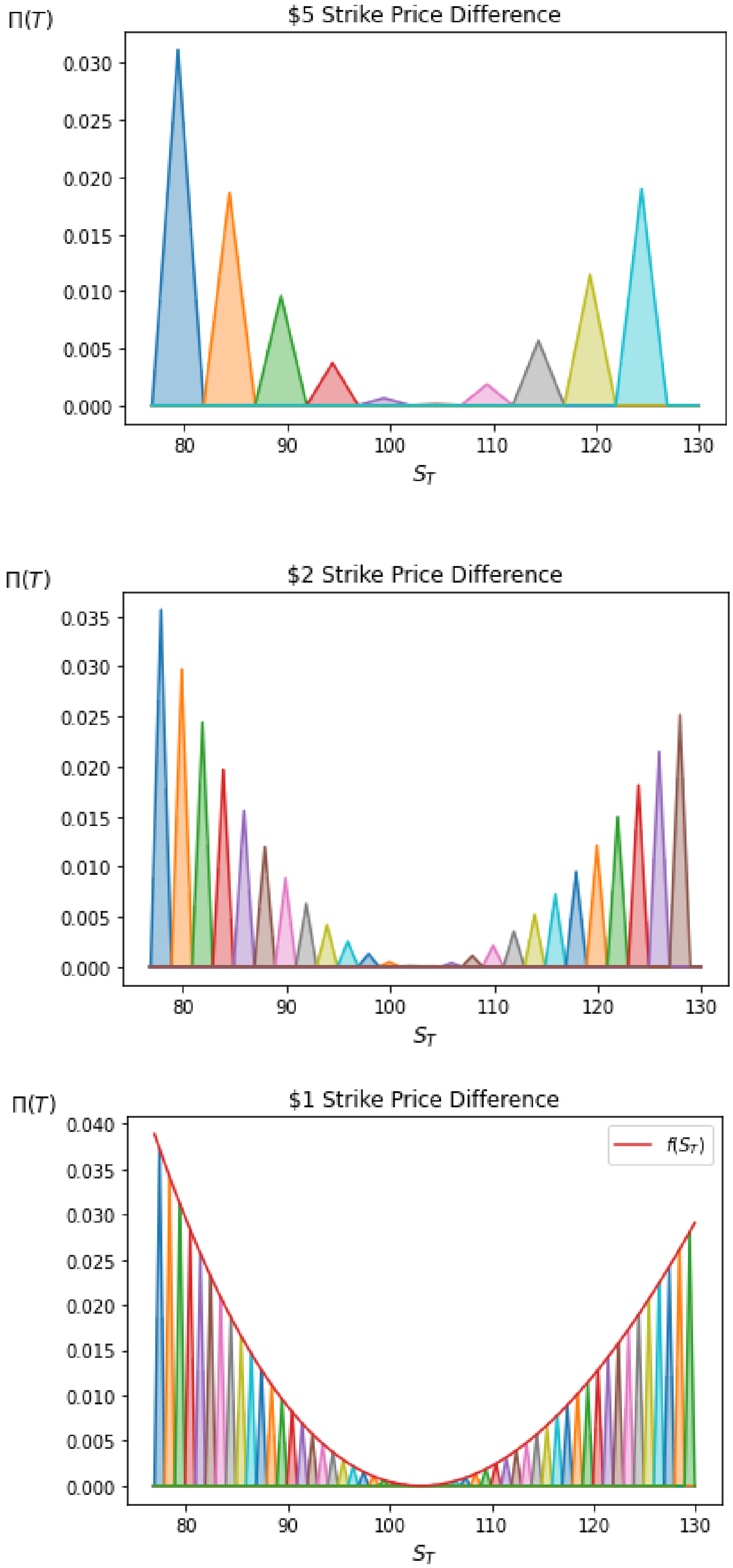
$$B(K(T)_i, \delta) = \sum_{x=-1}^{1} C(T, K(T)_i + \delta x) = \frac{|S_T - K(T)_i + \delta| - 2|S_T - K(T)_i| + |S_T - K(T)_i - \delta|}{2} \quad (3.1)$$

. The area of this portfolio's payoff,between $[K(T)_i - \delta, K(T)_i + \delta]$, is given by that of an isosceles triangle with a base of $2\delta$ and a height of $\delta$ [3], and by choosing a sufficiently small $\delta$ such that the slope of $B(K(T)_i, \delta)$ from $K(T)_i \pm \delta$ to $K(T)_i$ is sufficiently steep, we can replicate $f$ at some price $S_T = K(T)_i$ by buying $\frac{f(K(T)_i)}{\delta}$units of $B(K(T)_i, \delta)$. Iterating over each $i$ in {1, . . . ,$n$}, such that $K(T)_{i+1} - K(T)_i = \Delta K_i = 2\delta$, we can replicate $f(K(T)_i)$ by buying the appropriate number of units of $B(K(T)_i, \delta)$. Doing so in the limit as $\delta \to 0$, we realize the portfolio $\Pi_T$ that sufficiently replicates

---

[3] A butterfly spread centered at K with a width of a 2δ generates a maximum payoff when the underlying asset's price observed at maturity equals K, and said payoff is equal to half the distance between the upper and lower legs (i.e. (K+δ-(K-δ))/2) **[8]**.

**Figure 3: Approximations of $f(S_T)$ with various portfolios of butterfly spreads**



$f(S_T)\ \forall\ S_T \in [K(T)_\alpha,\ K(T)_\beta]$ (see figure 3). The payoff of said portfolio at maturity is given by the Riemann sum

$$\Pi_T = \int_{K(T)_\alpha}^{K(T)_\beta} \frac{\delta}{2}\ \frac{f(K_i)}{\delta}\ dK_i = 2\int_{K(T)_\alpha}^{K(T)_\beta} \frac{\delta^2}{2}\ \frac{f(K_i)}{\delta} = \int_{K(T)_\alpha}^{K(T)_\beta} f(K_i)\ \delta \quad (3.2)$$

[4]

.

Butterfly options not only allow us to replicate the payoff of some derivative $f(K)$ for any given number of states of $S_T$, but the value of this replicating portfolio gives us insight into expectations about the probability of the terminal price ($S_T$) being equal to $K$ (denoted $P(S_T = K)$). Let the probability density function (PDF) $\phi(S_T)$, which exhibits the following property:

$$\phi(K)dK = P(S_T = K) \qquad (3.3)$$

, be approximated by a long position in $\delta$ units of $B(K, \delta)$ in the limit as $\delta \rightarrow 0$ (see figure 4). By taking limits of (3.1), it can be shown that

$$\lim_{\delta\to 0} B(K(T), \delta) = \begin{cases} \delta, & \text{if } S = K \\ 0, & \text{otherwise} \end{cases} \qquad (3.4)$$

. The expected value of (3.4) is therefore given by

$$E[\lim_{\delta \rightarrow 0} B(K, \delta)] = [p\,\delta + 0(1 - p)\,]e^{r(T-t)} = p\delta e^{r(T-t)} \quad (3.5)$$

**[8]** where $p = P(S_T = K)$. Rearranging (3.5), we get

$$\phi\,(K) = \frac{E[\lim_{\delta\rightarrow 0} B(K,\delta)]}{\delta^2} e^{r(T-t)} \qquad (3.6)$$

and

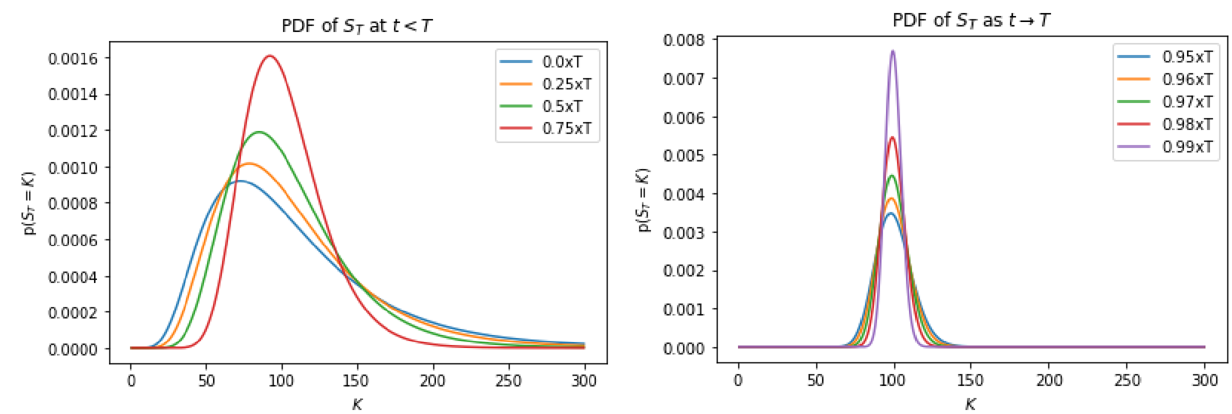
$$p = \frac{E[\lim_{\delta\rightarrow 0} B(K,\delta)]}{\delta} e^{r(T-t)} \qquad (3.7)$$

meaning we can recover the spectrum of probabilities for various terminal states of $S_T$ at any time $t \in [0, T]$, and therefore the cumulative density function (CDF)(denoted $\Phi(K)$), as well from the prices of butterfly spreads. Therefore, given that the function $f(S_t)$ tells us something about the expected returns variance of $S_T$, it also tells us something about expected probabilities of the future state(s) of $S_T$ (i.e., $e^{r(T-t)}E[S_T]$) by way of the butterfly spreads used in replicating $f(S_t)$, which is the principal idea presented in **[3]**.

---

[4] The factor 2 preceding the Riemann sum in equation 3.2 is a correction for the fact a Riemann sum constructed using isosceles triangles as we have done above only achieves half the area under the curve of $f(S_T)$ compared to that of one constructed with similarly thin rectangles.

**Figure 4: PDFs of $S_T$ far from (left) and close to (right) terminal time T**




Suppose at time $t = 0$, we observe that $S_t = 100$ and the riskless rate-of-return ($r$) equals 3%. With this information, we wish to gauge the expected probability of $S_T = 110$ at some terminal point-in-time $T = 2$(years). By (2.6) we have that the expected realized variance is given by

$$f(100) = e^{0.06}\frac{2}{2}\left[\frac{100}{100e^{0.06}} - 1 - log\left(\frac{100}{100e^{0.06}}\right)\right] = 0.001873 \qquad (3.8)$$

. In order replicate $f(110)$, we can buy $\frac{f(110)}{\delta}$ units of the butterfly spread $B(K = 110, \delta = 1)$ with legs at 109 and 111. It is clear to see that buying $\frac{1}{f(110)}$ units of this spread gives us

$$P(S_T = 110) = e^{0.06}\frac{1}{f(110)}\frac{f(110)}{\delta}[C(0, 110 - \delta) - 2C(0, 110) + C(0, 110)]$$

$$= e^{0.06}[C(0, 109) - 2C(0, 110) + C(0, 111)] = 0.008183 \qquad (3.9)$$

. We can therefore conclude $P(S_T = 110) \approx 0.8183\%$. Similarly, suppose we wanted gauge the probability of $S_T$ lying between some interval $[90, 110]$ at time $T$(denoted $P(90 \leq S_T \leq 110)$). Such can be accomplished by simply summing over the exercise from before for all $K \in [90, 110]$, which is analogous to partially recovering the CDF of $S_T$. In doing so, we see that

$$P(90 \leq S_T \leq 110) = \Phi(110) - \Phi(90)$$

$$= \int_{K=90}^{110} e^{0.06}[C(0, K - 1) - 2C(0, K) + C(0, K + 1)] = 0.19586 \quad (3.10)$$

. Therefore, given that the expected realized variance is equal to 0.001873, we can expect an approximate 19.6% probability that $S_T$ will be between \$90 and \$110.

## Appendix A: Extracting variance from arithmetic and geometric returns

Let the arithmetic mean return ($A$) be given by

$$A = \frac{1}{T}\sum_{t=1}^{T} r_t \qquad \text{(A1)}$$

where

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}} \qquad \text{(A2)}$$

and the geometric mean($G$) be given by

$$G = \frac{1}{T}\sum_{t=1}^{T} log(\frac{S_t}{S_{t-1}}) = \frac{1}{T}\sum_{t=1}^{T} log(1 + r_t) \qquad \text{(A3)}$$

. Furthermore, let the realized variance ($V$) be defined as

$$V = \int_{t=1}^{T} \sigma_t^2 dt = \sum_{t=1}^{T} log(1 + r_t)^2 \qquad \text{(A4)}$$

. Additionally, it will be useful to recall the following approximation of the exponential function:

$$e^x = exp(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + ... \approx 1 + x + \frac{1}{2}x^2 \qquad \text{(A5)}$$

. With this in mind, we can rewrite (A1) as follows:

$$\begin{aligned}
A &= \frac{1}{T}\sum_{t=1}^{T} r_t \\
&= -1 + \frac{1}{T}\sum_{t=1}^{T} (1 + r_t) \\
&= -1 + \frac{1}{T}\sum_{t=1}^{T} e^{log(1+r_t)} \\
&= -1 + \frac{1}{T}\sum_{t=1}^{T} [\, 1 + log(1 + r_t) + \frac{1}{2}log(1 + r_t)^2 \,] \\
&= \frac{1}{T}\sum_{t=1}^{T} log(1 + r_t) + \frac{1}{2}\frac{1}{T}\sum_{t=1}^{T} log(1 + r_t)^2 \\
&= G + \frac{1}{2}\sigma^2 \qquad \text{(A6)}
\end{aligned}$$

. The assertion made in (A4) that realized variance can be measured by summing over the series of squared log returns is a well understood concept in econometrics, first introduced by Barndorff-Nielsen and Shephard in **[2]**. It is trivial to show that setting $E[r_t]$ in (2.3) equal to zero yields the result given above, which is practically reasonable as the number of observations becomes large (over a fixed timer interval), but to see why this is the case, assume that log asset prices evolve according to the following arithmetic Brownian motion, whose drift and diffusion coefficients are $\mu$ and $\sigma$, respectively:

$$log(S_T) = log(S_0) + \mu \int_{t=0}^{T} dt + \sigma \int_{t=0}^{T} dz_t \quad \text{(A7)}$$

where $dz_t = \sqrt{dt}\, \phi(0, 1)$. The distribution of $dz_t$ means that

$$E[dz_t^2] - E[dz_t]^2 = 1 \quad \text{(A8)}$$

and $E[dz_t]^2 = 0$, which together imply

$$E[dz_t^2] = 1 \quad \text{(A9)}$$

**[7]**. Taking expectations we have

$$E[d\, log(S_t)^2] = (\mu_t dt)^2 + (\sigma\sqrt{dt})^2 E[dz_t^2] = \mu^2 dt^2 + \sigma^2 dt \quad \text{(A10)}$$

, and the predominance of variance in estimating log returns becomes clear as $dt$ becomes small **[1]**. Finally, by summing over (A10) for all $t \in [0, T]$, we can assert the following:

$$\sum_{t=0}^{T-1} log(\frac{S_{t+1}}{S_t})^2 = \sum_{t=1}^{T} log(1 + r_t)^2 = \sigma^2 \int_{t=0}^{T} dt = V \quad \text{(A11)}$$

.

## Appendix B: Replicating arbitrary derivative payoffs

**This appendix is largely inspired by and adapted from appendix 1 of **[4]** **

Let the dirac delta function, given by,

$$\delta(S-K)=\begin{cases}\infty, & \text{if } S=K\\ 0, & \text{otherwise}\end{cases} \qquad \text{(B1)}$$

satisfy the following by the sifting property:

$$\int_{-\infty}^{\infty} f(K)\,\delta(S-K)dK = f(S) \qquad \text{(B2)}$$

. Let us also define the unit step function $\theta(S-K)$ , whose derivative is the dirac delta function, to be given by

$$\theta(S-K)=\begin{cases}1, & \text{if } S-K>0\\ 0, & \text{otherwise}\end{cases} \qquad \text{(B3)}$$

, which we can recognize as the payout structure of a cash-or-nothing call option. That is to say

$$\theta(S-K) = \frac{\partial\, C(t,K)}{\partial\, K} \qquad \theta(K-S) = \frac{\partial\, P(t,K)}{\partial\, K} \qquad \text{(B4,B5)}$$

. Integrating (B1) over all possible $K \in [0,\infty]$, we get a function paying $1 in all possible of states of the underlying asset's price $S$, or

$$\int_{0}^{\infty} \delta(S-K)\,dK = \theta(S) = 1\ \forall\, S > 0 \qquad \text{(B6)}$$

. Now consider that (B2) can be split into the following(for $S, K \geq 0$):

$$f(S) = \int_{0}^{\infty} f(K)\,\delta(S-K)\,dK = \int_{0}^{S_*} f(K)\,\delta(S-K)\,dK + \int_{S_*}^{\infty} f(K)\,\delta(S-K)\,dK \qquad \text{(B7)}$$

, and in order to practically replicate $f(S)$, we must integrate both terms in the right-hand side of (B7) by parts as there is no traded instrument whose payoff gives the dirac delta function. In doing so we get,

$$\begin{aligned} f(S) &= f(K)\theta(K-S)\Big|_{0}^{S_*} - \int_{0}^{S_*} f'(K)\,\theta(K-S)\,dK \\ &+ f(K)\,\theta(S-K)\Big|_{S_*}^{\infty} - \int_{S_*}^{\infty} f'(K)\,\theta(S-K)\,dK \end{aligned} \qquad \text{(B8)}$$

, which represents

- a long position in $f(S_*)$ cash-or-nothing put options struck at $S_*$
- a short position in $f'(K)$ cash-or-nothing put options struck at $K$ for each $K \in [0, S_*]$

- a short position in $f(S_*)$ cash-or-nothing call options struck at $S_*$
- a short position in $f'(K)$ cash-or-nothing call options struck at $K$ for each $K \in [S_*, \infty]$.

Integrating (B8) by parts once again gives

$$\begin{aligned} f(S) &= f(S_*) \\ &+ f'(S_*)[(S - S_*)^+ - (S_* - S)^+] \\ &+ \int_0^{S_*} f''(K)(K - S)^+ dK + \int_{S_*}^{\infty} f''(K)(S - K)^+ dK \end{aligned} \qquad \text{(B9)}$$

, which is analogous to the idea given in section II of **[4]**: that any general function ($f(S)$) dependent on an underlying asset's price $S$ that is continuously twice differentiable can be replicated by

- a long position in $f(S_*)$ zero-coupon bonds
- a long position in $f'(S_*)$ call options struck at $S_*$
- a short position in $f'(S_*)$ put options struck at $S_*$
- a long position in $f''(K)$ put options struck at $K$ for each $K \in [0, S_*]$
- a long position in $f''(K)$ call options struck at $K$ for each $K \in [S_*, \infty]$.

## References


**[1]** Andersen, T. G., & Benzoni, L. (2008). Realized Volatility, Working Paper 2008-14.

**[2]** Barndorff-Nielsen, O. E., & Shephard, N. (2000). *Econometric analysis of realised volatility and its use in estimating Lévy based non-Gaussian OU type stochastic volatility models*. CAF, Centre for Analytical Finance.

**[3]** Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of business*, 621-651.

**[4]** Carr, P., & Madan, D. (2001). Towards a theory of volatility trading. *Option pricing, interest rates and risk management, handbooks in mathematical finance*, *22*(7), 458-476.

**[5]** Cboe, V. I. X. (2019). White Paper, Cboe Volatility Index®.

**[6]** Demeterfi, B., Derman, E., Kamal, Z., & Zou, J. (1999). More than you ever wanted to know about Variance Swaps. *Goldman Sachs, Quantitative Strategies Research Notes*.

**[7]** Hull, J. C. (2014). *Options, Futures, and Other Derivatives* (9th ed.). Pearson.

**[8]** Martin, R. Andrew. (2020, October 10). *Option-implied probability distributions, part 2 · Reasonable Deviations*. Reasonable Deviations. https://reasonabledeviations.com/2020/10/10/option-implied-pdfs-2/