Practical and Theoretical Aspects of Volatility Modelling and Trading

Artur Sepp
artur.sepp@juliusbaer.com

Julius Baer

ETH Practitioner Seminar

October 21, 2016
Content

1. Option replication and trading: the theory vs the real world

2. Derivatives industry and applications of valuation models

3. Volatility modeling and steady-state analysis of stochastic volatility models

4. Volatility trading in practice: the convexity vs the concavity and the volatility risk-premium
It is not true that quantitative/mathematical methods are recent developments in trading applications

Some quotes from a nice book ”Reminiscences of a Stock Operator” (1923) about the biography of a legendary trader Jesse Livermore:

• ”Wall Street makes its money on a mathematical basis, I mean, it makes its money by dealing with facts and figures”

• ”He (the trader) must bet always on probabilities - that is, try to anticipate them”

• ”The game of speculation isn’t all mathematics or set rules, however rigid the main laws may be”

Livermore was trading spot/futures markets making big bets on trends

As for option trading:

• Options are non-linear securities on underlying prices

• Trading and valuation of derivatives can only be possible using quantitative models and tools
Probability and Volatility

Options valuation includes estimation of probabilities of asset price changes.

Volatility is a measure of a likelihood of given price changes.

Figure: empirical tail probabilities of weekly returns on the S&P 500 index (high volatility) and 2year US bond ETFs (low volatility)
Volatility is not the ultimate measure of the risk

Figure: empirical tail probabilities of weekly normalized returns on the S&P 500 index and 2-year US bond ETFs

Risk-parity funds: leverage up low volatility assets to a target volatility

Empirical Tail Probabilities for Standardized Weekly Returns

- S&P 500 Index
- 2y US bond ETF
Volatility is clustered

Figure: time series of hitting indicator when absolute returns exceed one standard deviation
Co-dependence is between asset classes ic clustered

Figure: time series of the joint hitting indicator the S&P 500 index and 2year US bond ETFs
Vanilla Put and Call options are primary derivative instruments traded on exchanges

Values and prices of option contracts are derived from the probability of return distributions

Options enable to create strategies related to statistical and market implied probabilities / volatilities

**European call** option gives the holder the right to buy the asset at maturity time $T$ at strike price $K$:

$$u(S(T)) = (S(T) - K)^+$$

**Put option** gives the right to sell:

$$u(S(T)) = (K - S(T))^+$$

Put and call options on major asset classes and stocks represent the bulk of exchanged traded derivative contracts

Any payoff function $u(s)$ on $S(T)$ can be linearly approximated with put and calls
Option pricing in industry (using Oscar Wilde)

A mathematically-oriented quant = "a man who knows the price of everything and the value of nothing" (?

An empirical quant = "a man who sees an absurd value in everything and doesn’t know the market price of any single thing" (?

For understanding the practicalities of option trading, we need to understand:

1) The theory of option replication

2) Practicalities of options valuation and trading

3) Empirical features of option trading strategies

4) In particular, the interception of risk-neutral valuation measure and the statistical measure
**Fundamental option trading formula** is originated by Black-Scholes-Merton (1973) and extended by Harrison-Pliska (1981)

We can assume a general dynamics for the underlying asset under the statistical measure:

\[ dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t) \]

where \( \mu(t) \) is the drift  
\( \sigma(t) \) is the volatility of asset returns  
\( W(t) \) a is standard Brownian motion

The key result of Black-Scholes-Merton replication framework and risk-neutral valuation is:

**There exists a trading strategy in the underlying asset with the dynamic weight** \( \Delta(s) \) **such that** the terminal payoff \( u(S(T)) \) of the option can be replicated by trading in the underlying for any realization of price path (\(!\)):

\[ u(S(T)) = g(S(t)) + \int_t^T \Delta(S(t'))dS(t') \]
Black-Scholes-Merton framework is an idealization of real market conditions

BSM assumptions vs real trading conditions:

- Continuous trading in diffusion-uncertainty market vs discrete trading with gaps (jumps)
- No transaction costs vs transaction costs and market impact costs
- Unlimited borrowing & lending ability at the same risk-free rate vs limited capacity to borrow funds and finance short & long position at different rates
- No exogenous risk factors vs the risk of changes in the volatility, interest rates, dividends, etc
- Instantaneous price discovery vs wide bid-ask spreads and illiquidity
- Flat/zero "end-of-day" risk vs the illusion of daily mark-to-market replication
**Black-Scholes-Merton implied volatility**

How trading imperfection do affect realized profit & loss?

Black-Scholes-Merton model is based on the log-normal price dynamics under the valuation (risk-neutral, martingale) measure:

\[
dS(t) = \sigma_{BSM} S(t) dW(t)
\]

where \( \sigma_{BSM} \) is the constant volatility under the valuation measure.

Option value \( U(t, S) \) solves the BSM PDE (assuming zero borrowing/lending costs):

\[
\partial_t U + \frac{1}{2} \sigma_{BSM}^2 S^2 \partial_{SS} U = 0 \iff U(T, S) = u(S)
\]

Given market price of an option we can solve the inverse problem to find the BSM implied volatility \( \sigma_{BSM} \) (equate BSM model value to the market price).
Continuous-time Delta-Hedging P&L is the spread between implied and realized volatilities

Delta-hedging portfolio $\Pi(t)$ for hedging a short position in option $U(t, S)$:

$$\Pi(t) = \Delta(t, S)S(t) - U(t, S)$$

Over the infinitesimal time $dt$, using the BSM PDE, the delta-hedging P&L is

$$d\Pi(t) = \frac{1}{2} \left\{ \sigma_{BSM}^2 dt - R(t) \right\} S^2(t) \Gamma(t, S)$$

where $\Gamma(t, S)$ is option gamma $\Gamma(t, S) = \partial_{SS}U(t, S)$

$R(t)$ is the return squared under the statistical measure (!):

$$R(t) = \left( \frac{dS(t)}{S(t)} \right)^2$$

In the limit, $R(t) \to \sigma_{STAT}^2 dt$ where $\sigma_{STAT}$ is returns volatility under the statistical measure

The delta-hedging P&L is zero only if the implied BSM volatility equals to the statistical volatility:

$$\sigma_{BSM} = \sigma_{STAT}$$
The Fundamental Equation relating Implied volatility vs Realized volatility

Real-world imperfections result in the spread between the statistical volatility of returns, $\sigma_{STAT}$, and the BSM volatility implied by market prices of options, $\sigma_{BSM}$

Fundamental equation for the final P&L of delta-hedging strategy (El Karoui-Jeanblanc-Shreve (1998)):

$$\Pi(T) = \frac{1}{2} \int_0^T \left\{ \sigma_{BSM}^2 - \sigma_{STAT}^2 \right\} S^2(t') \Gamma(t, S') dt'$$

Even in the ideal conditions with continuous trading in diffusive uncertainty and no trading costs, this result is fundamental because:

1. If implied BSM and statistical volatilities are different, option trading strategies can be designed to take advantage of this spread

2. These strategies still have little dependence on the real-world drift of the underlying asset

This result holds for price dynamics with stochastic volatility and jumps
The spread between the statistical realized volatility and the implied volatility is significant and persistent

Volatility Risk-premium = Implied volatility – Realized volatility

Figure:

Proxy Volatility Risk-premium = VIX at month start

– Realized volatility of S&P500 in this month

t-statistic is 8.20
Theory vs The Real World

In theory: BSM framework assumes that a derivative security is redundant because it can be replicated and, as a result, it adds no utility to investors' portfolios

In practice:

1. Retail/institutional investors are not able to delta-hedge and replicate derivatives (no infrastructure, little capital for margin, expensive trading costs)

2. A derivative security adds utility to investors' portfolios:
   - Upside speculation (out-of-the-money calls)
   - Downside protection (out-of-the-money puts)
   - Carry strategies (selling options without hedging to generate income)

3. Hedge funds typically use derivatives for tactical discretionary views

4. Dealers (investment banks) and options market makers stand on the other side of transactions with the goal to generate profits on their capital at risk
Derivatives Industry

The impossibility of replication and the spread between implied and realized volatilities (return distributions) give rise to trading and business opportunities which utilize quantitative models and methods with various levels of complexity.

1. Structured derivatives business at investment banks

2. Prime brokerage and exchanges (for clearing and margining)

3. Options market makers

4. Proprietary trading at hedge funds
Structured Derivatives Business employs the classic applications of derivatives pricing models and tools

1. Broker-dealer sells to a client a structured product

2. Risk of this product is computed using a market consistent model

3. The first order risk, delta and vega, are hedged by trading in exchange traded derivatives

4. Flow driven business

The dealer has the advantage:

1. The client sells volatility to the dealer cheaply so the dealer buys cheap volatility and hedges himself by selling volatility more expensively in the market

2. The client buys volatility from the dealer (by buying principal protected note) at expensive levels, the dealer hedges by buying cheaper protection in the market
Structured Derivatives Business - modeling tools

1. A model to compute and interpolate implied BSM volatility from traded option market prices

2. A model for implied volatility surfaces

3. Local and stochastic volatility models, calibrated to implied volatilities, to value and risk-manage structured products

4. Consistency with the statistical dynamics are note relevant as dealers seek to eliminate the first order risks (delta and vega) being compensated by higher spreads from structured products
Prime Brokers, Exchanges and Risk management

Provide clearing, funding and risk-management for exchange traded and OTC derivatives for institutional investors, hedge funds, propriotory traders

Risk management sets trading budgets for trading desks

Objective is to aggregate risk of different instruments by strikes, maturities, underlyings and to provide a "fair" margin for clients

Require the consistency with historical data (both recent data and stress sase data)

Employ time series analysis (PCA) and simple pricing models

Value-at-risk is computed using EWMA and Garch time models to predict the short-term volatility
Option Market Makers

Provide bid-ask quotes for exchange traded options

Primarily apply the BSM model with a function for the implied volatility

Intraday pattern of volatility

Co-dependence with spot price and volatility
Proprietary/systematic trading

Estimate and predict realized volatility

Generate signals by screening cheap/expensive volatility in the market
Volatility models in details

1. Models for Implied Volatility

2. Local Volatility Models

3. Stochastic Volatility Models
Implied Volatility models are applied to interpolate and extrapolate discrete options data

Figure: Snapshot of data for options quotes on Apple stock
**BSM implied volatility**

Figure: Implied BSM volatility as function of strike

Implied volatilities for out-of-the-money puts and calls are expensive
Arbitrage-free implied volatility function is a key input for computing risks and calibration of more advanced models

Key challenges:

- Data is discrete across strikes and maturities
- Bid-Ask spreads are wide for out-of-the-money options

Typical Approaches:

- Parametric form (SVI, SABR)
- Non-parametric (splines)
Non-parametric local volatility model linkes implied volatility into implied distributions

Breeden-Litzenberger (1978) formula relates market prices $C^{\text{market}}$ (implied volatilities) into implied terminal distribution under the valuation measure:

$$
\mathbb{P}[S(T) = K] = \partial_{KK} C^{\text{market}}(T, K)
$$

where $T$ is the maturity time and $K$ is the strike.

Local volatility model specifies a function $\sigma_{\text{loc,dif}}(t, S)$ so that price dynamics are consistent with the implied terminal distribution above:

$$
dS(t) = \sigma_{\text{loc,dif}}(t, S(t)) S(t) dW(t)
$$

Local volatility is computed using Dupire formula (1994):

$$
\sigma^2_{\text{loc,dif}}(T, K) = \frac{C_T(T, K)}{\frac{1}{2} K^2 C_{KK}(T, K)}
$$

A continuum of options market prices is calibrated perfectly

Problem: options market data are discrete
Parametric local volatility models specify parametric functions

This models can be applied to calibrate a small number of market quotes at one maturity (they have a too small number of parameters to fit the whole implied volatility surface)

The classic example is the CEV process (Cox (1975)):

$$dS(t) = \sigma \left( \frac{S(t)}{S(0)} \right)^\beta dW(t)$$

Parameter $\beta$ is the leverage coefficient that allows to calibrate the implied volatility skew

![Implied BSM volatility in the CEV model](image)
Lipton-Sepp (2011) local volatility model has as many parameters as market quotes

Given a discrete set of market prices $C_{mrkt}(T_i, K_j)$, $0 \leq i \leq I$, $0 \leq j \leq J_i$

Introduce a tiled local volatility $\sigma_{loc}(T, K)$:

$$\sigma_{loc}(T, S) = \sigma_{ij}, \quad T_{i-1} < T \leq T_i, \quad K_{j-1} < S \leq K_j, \quad 1 \leq i \leq I, \quad 0 \leq j \leq J_i$$

By construction, for every $T_i$, $\sigma_{loc}(T_i, K)$ depends on as many parameters as there are market quotes.

Semi-analytic model using Laplace transform and recursive solution to Sturm-Liouville problem with least-square calibration.

Illustration (vs CEV model) using the two-tiled case:

$$\sigma(S) = \begin{cases} 
\sigma_0, & S \leq S_0, \\
\sigma_1, & S > S_0.
\end{cases}$$

![Implied BSM volatility in TwoTile local volatility vs CEV model](image.png)
Local Volatility Models

Local volatility models are the most widely used by dealers as interpolation tools from market prices of vanilla products into implied distributions for pricing structured products.

Local volatility model serves only as risk-management tool.

These models lack dynamical properties, in particular, the mean-reverting features of implied volatilities.

![Graph](image-url)

**One month implied and realized vol vs performance of S&P500 index**

- 1 month ATM implied vol (lhs)
- 1 month realized vol (lhs)
- S&P 500 price performance (rhs)
**Stochastic Volatility Models**

Introduce diffusive uncertainty for the log-price $S(t)$ and variance $V(t)$ dynamics with correlated Brownian motions $W^{(0)}(t)$ and $W^{(1)}(t)$:

\[
\frac{dS(t)}{S(t)} = \mu(t)dt + \sqrt{V(t)}dW^{(0)}(t) \\
V(t) = a(V)dt + b(V)dW^{(1)}(t)
\]

In practice and literature, the following concepts are studied:

- The instantaneous variance of price returns
  \[
  \var[r_t] = V(t)dt \quad r_t = \log(S(t)/S(0))
  \]

- Quadratic Variance: the integrated instantaneous variance
  \[
  QV(t) = \int_0^t V(t')dt'
  \]

- Realized Variance: the discrete approximation of the quadratic variance computed over discrete time grid $\{t_k\}$
  \[
  DV(t) = \sum_{t_k \in [0,t]} r_{t_k}^2
  \]
Stochastic Variance Models

The SDE for the price process $S(t)$ with the stochastic variance $V(t)$:

$$
\frac{dS(t)}{S(t)} = \mu(t)dt + \sqrt{V(t)}dW^{(0)}(t), \ S(0) = S
$$

Classical analytically tractable models:

- Heston model (1993):
  $$
  dV(t) = \kappa(\theta^2 - V(t))dt + \varepsilon\sqrt{V(t)}dW^{(1)}(t)
  $$

- $3/2$ SV model (Lewis 2002):
  $$
  dV(t) = \kappa V(t)(\theta^2 - V(t))dt + \varepsilon(V(t))^{3/2}dW^{(1)}(t)
  $$
**Stochastic Volatility Models**

Price dynamics $S(t)$ with the stochastic volatility process $\sigma(t)$:
\[
dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW^{(0)}(t)
\]

The classic models:

- **Stein-Stein model:**
\[
d\sigma(t) = \kappa(\theta - \sigma(t))dt + \varepsilon dW^{(1)}(t)
\]
Volatility is normally distributed: not a good feature

- **SABR Model (Hagan (2003)):**
\[
d\sigma(t) = \varepsilon \sigma(t)dW^{(1)}(t)
\]
The volatility is not mean reverting and explosive: practitioners only use the function for approximation of the model implied volatility

- **Log-normal model:**
\[
d\sigma(t) = \kappa(\theta - \sigma(t))dt + \varepsilon \sigma(t)dW^{(1)}(t)
\]
The model has a very strong empirical evidence (Christoffersen-Jacobs-Mimouni (2010)) but it is not analytic
The distribution of realized volatility is close to log-normal

**Left figure:** empirical frequency of the VIX for last 20 years: it is definitely not normal

**Right figure:** frequency of the logarithm of the VIX: it is close to the normal density (especially the right tail)
The distribution of implied volatility is also close to log-normal

**Left figure:** frequency of realized vol - it is definitely not normal

**Right figure:** frequency of the logarithm of realized vol - again it does look like the normal density (especially for the right tail)
Factor model for changes in volatility (realized or implied) \( \sigma(t_n) \) predicted by returns in price \( S(t_n) \) is simpler to interpret and estimate:

\[
s(t_n) - s(t_{n-1}) = \beta \left[ \frac{s(t_n) - s(t_{n-1})}{s(t_{n-1})} \right] + \sigma(t_{n-1}) \epsilon_n
\]

IID normal residuals \( \epsilon_n \) are scaled by \( \sigma(t_{n-1}) \) due to log-normality.

Left figure: scatter plot of daily changes in the VIX vs returns on S&P 500 for past 14 years: Volatility beta \( \beta \approx -1.0 \) with \( R^2 = 80\% \).

Right: time series of residuals \( \epsilon_n \) does not exhibit any systemic patterns.
More evidence on log-normal dynamics of volatility using high frequency data: independence of regression parameters on level of ATM volatility

Left figure: test $\tilde{\beta}(V) = \beta V^\alpha$ by regression model: $\ln |\tilde{\beta}(V)| = \alpha \ln V + c$

Right: test $\tilde{\varepsilon}(V) = \varepsilon V^{1+\alpha}$ by regression model: $\ln |\tilde{\varepsilon}(V)| = (1+\alpha) \ln V + c$

The estimated value of elasticity $\alpha$ is small and statistically insignificant.

$y = 0.15x + 0.14$
$R^2 = 2\%$

$y = 0.14x - 0.45$
$R^2 = 4\%$
Beta stochastic volatility model (Karasinski-Sepp 2012):

\[
dS(t) = \sigma(t)S(t)dW^{(0)}(t)
\]

\[
d\sigma(t) = \kappa(\theta - \sigma(t))dt + \beta\frac{dS(t)}{S(t)} + \varepsilon\sigma(t)dW^{(1)}(t)
\]

\[
= \kappa(\theta - \sigma(t))dt + \beta\sigma(t)dW^{(0)}(t) + \varepsilon\sigma(t)dW^{(1)}(t)
\]

\(\sigma(t)\) is either returns volatility or short-term ATM implied volatility.

\(W^{(0)}(t)\) and \(W^{(1)}(t)\) are independent Brownian motions.

\(\beta\) is volatility beta - sensitivity of volatility to changes in price.

\(\varepsilon\) is residual volatility-of-volatility - standard deviation of residual changes in vol.

Mean-reversion rate \(\kappa\) and volatility mean rate \(\theta\) are incorporated for the mean-reverting feature and the stationarity of volatility.
Semi-analytic solution of log-normal SV model (Sepp 2015)

Introduce the mean-adjusted volatility:

\[ Y(t) = \sigma(t) - \theta, \quad Y(0) = Y = \sigma(0) - \theta \]

The dynamics for log-price \( X(t) = \ln(S(t)) \) and quadratic variance \( I(t) \):

\[
dX(t) = -\frac{1}{2} (Y(t) + \theta)^2 dt + (Y(t) + \theta) dW^{(0)}(t)
\]

\[
dY(t) = -\kappa Y(t) dt + \beta (Y(t) + \theta) dW^{(0)}(t) + \varepsilon (Y(t) + \theta) dW^{(1)}(t)
\]

\[
dI(t) = (Y(t) + \theta)^2 dt
\]

The valuation PDE is given on the domain \([0, T] \times \mathbb{R} \times \mathbb{R}_+ \times (-\theta, \infty)\):

\[
-U_\tau + \left( L^{(Y)} + L^{(XI)} \right) U = 0
\]

\[
U(0, X, I, Y) = u(X, I)
\]

where the diffusive operators \( L^{(Y)} \) and \( L^{(XI)} \) are defined on the domain \([0, T] \times \mathbb{R} \times \mathbb{R}_+ \times (-\theta, \infty)\):

\[
L^{(Y)} U = \frac{1}{2} \vartheta^2 (Y + \theta)^2 U_{YY} - \kappa Y U_Y + \beta (Y + \theta)^2 U_{XY}
\]

\[
L^{(XI)} U = (Y + \theta)^2 \left[ \frac{1}{2} (U_{XX} - U_X) + \beta U_{XY} + U_I \right]
\]
Affine decomposition for log-normal SV model

The moment generation function (MGF) of the log-price $X(\tau)$ and the QV $I(\tau)$ with transform variables $\Phi, \Psi \in \mathbb{C}$:

$$ G(\tau, X, I, Y; \Phi, \Psi) = \mathbb{E}[e^{-\Phi X(\tau)} - \Psi I(\tau)] $$

MGF $G$ solves the PDE:

$$ -G_\tau + \left( \mathcal{L}^{(Y)} + \mathcal{L}^{(XI)} \right) G = 0, \quad G(0, X, I, Y; \Phi, \Psi) = e^{-\Phi X - \Psi I}. $$

**Theorem.** The MGF function can be decomposed into the leading term $E^{[2]}$ and the remainder term $R^{[2]}$:

$$ G(\tau, X, I, Y; \Phi, \Psi) = E^{[2]}(\tau, X, I, Y; \Phi, \Psi) + R^{[2]}(\tau, X, I, Y; \Phi, \Psi), $$

The leading term $E^{[2]}$ is given by the exponential-affine form:

$$ E^{[2]}(\tau, X, I, Y; \Phi, \Psi) = \exp \left\{ -\Phi X - \Psi I + \sum_{k=0}^{4} A^{(k)}(\tau; \Phi, \Psi) Y^k \right\}, $$

where the functions $A^{(k)}$ solve the system of ODEs as functions of $\tau$.

The remainder term $R^{[2]}(\tau, X, I, Y; \Phi, \Psi)$ solves the following problem:

$$ -R^{[2]}_\tau + \left( \mathcal{L}^{(Y)} + \mathcal{L}^{(XI)} \right) R^{[2]} = -F^{[2]}(Y, A^{(1)}, A^{(2)}, A^{(3)}, A^{(4)}) E^{[2]}(\tau, X, I, Y; \Phi, \Psi), $$

$$ R^{[2]}(0, X, I, Y; \Phi, \Psi) = 0, $$

where the source term $F^{[2]}$ is a polynomial function in $Y$. 


Second-order Affine Decomposition

**Corollary.** The second-order approximation for the MGF $G$ is obtained by the leading affine term $E^{[2]}$:

$$G(\tau, X, I, Y; \Phi, \Psi) = E^{[2]}(\tau, X, I, Y; \Phi, \Psi),$$

with accuracy given by the estimate for the remainder term $R^{[2]}$:

$$\left| R^{[2]}(\tau, X, I, Y; \Phi, \Psi) \right| \leq \sum_{n=5}^{8} C_n(\tau; \Phi, \Psi) \times M^{(n)}_\sigma,$$

where $M^{(n)}_\sigma$ is the $n$-th central moment of the steady-state volatility, and $C_n(\tau; \Phi, \Psi), n = 5, 6, 7, 8$, are real-valued constants depending on $\tau$ and the transform variables $\Phi$ and $\Psi$.

**Proposition.** There exists a unique and continuous solution for coefficients $A^{(k)}(\cdot, \Phi, \Psi), k = 0, \ldots, 4$ in the second-order affine decomposition.

**Proposition.** The second-order leading affine term satisfies the martingale condition:

$$E^{[2]}(\tau, X, I, Y; \Phi = 0, \Psi = 0) = 1,$$

**Proposition.** The second-order leading affine term is consistent with the expected value, variance, and covariance of the log-price and of the QV.
Applications for pricing options under the log-normal stochastic volatility model: consistency across different maturities and strikes (using S&P 500 index options)
Equilibrium / Steady-State Analysis of SV models

Different SV models have apparently different dynamics and distributions of returns: how is about their limiting behaviour?

- The steady-state distribution of the volatility $\sigma$

- The theoretical distribution of volatility-conditional returns:
  \[ X \mid \sigma \overset{d}{=} n \left( 0, \sqrt{c\sigma} \right) \]
  where $c$ is the scaling factor, $c = 1/252$ for daily returns.

- The empirical distribution of volatility-normalized returns:
  \[ \tilde{X}_n = \frac{1}{\hat{\sigma}_{n-1}} \ln \left( \frac{S_n}{S_{n-1}} \right) \]
  where $\hat{\sigma}_{n-1}$ is the empirical estimate of volatility at time $t_{n-1}$
Empirical distribution of volatility-normalized returns is close to normal distribution

**Right:** QQ-plot of monthly returns on the S&P 500 index

**Left:** QQ-plot of monthly returns normalized by the realized historic volatility of daily returns within given month

Anderson-Darling test for normality of returns (H0 hypothesis)

<table>
<thead>
<tr>
<th></th>
<th>Returns</th>
<th>Normalized Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject H0</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0005</td>
<td>0.9515</td>
</tr>
</tbody>
</table>

We have strong presumption against normality of returns

However we cannot reject the hypothesis that volatility-normalized returns

Non-normality of returns can be explained by regimes in the volatility
Hitting times of volatility normalized returns are not clustered unlike those for absolute returns.
Steady-state distribution of volatility

The steady-state distribution of the volatility is obtained by letting the time variable to the infinity.

For the log-normal model the steady-state distribution solves the ODE:

\[
\frac{1}{2} \theta^2 \sigma^2 G_{\sigma \sigma} - [\kappa(\theta - \sigma) G]_{\sigma} = 0
\]
Steady-state distribution of the volatility

For log-normal model:

\[ G^{(LG)}(\sigma) = \frac{\nu^\nu}{\Gamma(\nu)} \frac{\exp\left\{-\frac{\nu}{\sigma}\right\}}{\sigma^{1+\nu}}, \quad \nu = 1 + \frac{2\kappa}{\varepsilon^2}, \quad \nu = \frac{2\kappa\theta^2}{\varepsilon^2} \]

Inverse Gamma distribution with shape \( \alpha = \nu \) and scale \( \beta = \nu \)

For 3/2 model:

\[ G^{(3/2)}(\sigma) = \frac{\nu^\nu}{\Gamma(\nu)} \frac{\exp\left\{-\frac{\nu}{\sigma}\right\}}{\sigma^{1+\nu}}, \quad \nu = 2 + \frac{2\kappa}{\varepsilon^2}, \quad \nu = \frac{2\kappa\theta^2}{\varepsilon^2} \]

Inverse Gamma distribution with shape \( \alpha = \nu \) and scale \( \beta = \nu \)

Structurally the 3/2 model and the log-normal model are similar

For Heston model:

\[ G^{(H)}(\sigma) = \frac{\nu^{-\nu}}{\Gamma(\nu)} \frac{\exp\left\{-\frac{\sigma}{\nu}\right\}}{\sigma^{1-\nu}}, \quad \nu = \left(\frac{2\kappa\theta^2}{\varepsilon^2}\right)^{-1}, \quad \nu = \frac{2\kappa}{\varepsilon^2} \]

Gamma distribution with shape \( \alpha = \nu \) and scale \( 1/\nu \)
Illustration of the steady state density of volatility

Steady-State Distribution of Volatility

0% 5% 10% 15% 20% 25% 30% 35% 40% 45%
Illustration of the steady state density of volatility: log-normal model implies heavy

For the logarithm of the volatility $L = \log(\sigma)$ under the lognormal model, the distribution is given the extreme-value type PDF:

$$G^{(L)}(L) = \frac{\nu^\nu}{\Gamma(\nu)} \exp \left\{ - \frac{\exp \{-L\}}{\nu} + \nu L \right\}$$
The Distribution of Conditional Returns

Consider the distribution of returns conditional on the steady-state volatility:
\[ X \mid \sigma^\infty \sim \mathcal{N}(0, \sqrt{c}\sigma^\infty) \]

where \( c \) is the scaling factor

The unconditional PDF is obtained by the integral:
\[
G(X)(X) = \int_0^\infty \frac{1}{\sqrt{2\pi c}\sigma^\infty} \exp\left\{ -\frac{1}{2} \frac{X^2}{c(\sigma^\infty)^2} \right\} G(\sigma^\infty)(\sigma^\infty) d\sigma^\infty
\]

For the lognormal SV model:
\[
G(X)(X) = \frac{\Gamma\left(\nu + \frac{1}{2}\right)}{\sqrt{2\pi cv}\Gamma\left(\nu\right)} \left(1 + \frac{X^2}{2cv}\right)^{-\frac{1}{2}(2\nu+1)}
\]

This is the Student t-distribution with \( 2\nu \) degrees of freedom

For 3/2 SV model, the Student t-distribution with \( 2(\nu + 1) \) degrees

For Heston model:
\[
G(X)(X) = \frac{2 (2cv)^{-\frac{1}{2}(\frac{1}{2}+\nu)} |X|^{-\frac{1}{2}+\nu}}{\sqrt{\pi}\Gamma(\nu)} K_{\nu-\frac{1}{2}}\left(2 \frac{|X|}{\sqrt{2cv}}\right)
\]

\( K_{\nu}(x) \) is modified Bessel function of the second kind with index \( \nu \)
Illustration: the tail of unconditional distribution of returns is heavier under the lognormal SV model.
Summary of stochastic volatility models: applications for trading volatility risk-premium

The strength of SV models:

1. Calibration to market options prices for estimation of the implied distributions
2. Forecasting of the expected volatility and its probabilistic range conditional on the current observables
3. Steady-state analysis of systematic volatility trading strategies

Figure: expected vs realized volatility risk-premium computed using the log-normal SV model
Volatility trading in practice: Usage of Options and Structured Products

- Hedging (controversial)

- Investment products with limited downside:
  1. To get a convex pay-off (limited downside with large upside)
  2. The key is the volatility premium
The convexity profile without costs is appealing

Figure: the realized convexity of the straddle (long at-the-money call and put options) rolled monthly on S&P 500 index from 2005 up to 2016

The convexity of returns is a very attractive profile for any investment strategy

Taleb: you should own the convexity

\[ y = 573994x^2 - 1669.6x + 3465.5 \]
\[ R^2 = 0.652 \]
The convexity profile accounting to costs is not appealing

Figure: the convexity of straddle rolled monthly on S&P 500 index from 2005 up to 2016 adjusted to market price of straddle

Why Taleb’s advice does not work in practice: the convexity is overpriced

Link to the behavior science (GMO LLC: "What the Beta Puzzle Tells Us about Investing") and preference for "lottery" payoffs

Empirical evidence: the concavity profile (short options) provides higher returns than the convexity profile (long options) in the long term
Empirical evidence: systematic short convexity strategies out-perform the benchmark with smaller risk

[Graph showing Total Return Fund starting with $1 for S&P 500 index, CBOE Put Index, and CBOE Call index]

[Graph showing Running % Drawdown for S&P 500 index, CBOE Put Index, and CBOE Call index]
The cyclicality of the volatility risk-premium makes trend-following with options prohibitive

Two major investment approaches:

1. Follow the trend: buy high and sell higher
2. Contrarian: bet on reversions or range bounds

Trend following using options is prohibitive as the volatility risk-premium is cyclical and the option value decays the fastest at a high volatility

Figure: prior month return on the S&P 500 index and option premium for monthly straddles at the third Fridays
Convexity is equivalent to volatility

Given monthly returns, consider two estimators of the volatility:

- The convexity estimator using the monthly return (equivalent to P&L on the straddle):
  \[
  \hat{\sigma}^{(\text{conv})}_n = \sqrt{12} \times \sqrt{\frac{\pi}{2}} \left| \frac{S(t_n)}{S(t_{n-1})} - 1 \right|
  \]

- The volatility estimator using daily returns within the month (equivalent to P&L on the straddle delta-hedged daily):
  \[
  \hat{\sigma}^{(\text{vol})}_n = \sqrt{12} \times \left( \frac{1}{N} \sum_{t_k \in (t_{n-1}, t_n]} \left( \frac{S(t_k)}{S(t_{k-1})} - 1 \right)^2 \right)
  \]

Figure: Convexity and volatility have equivalent right tails
CTAs (commodity trading advisors) are able to create convex return profiles by applying quant strategies for trend-following

Figure: Monthly returns on SC CTA index (tracking 20 largest CTAs) from 2001 to 2016 vs monthly returns on the S&P 500 index

CTAs attempt to replicate option pay-off without actually buying options (long convexity but short volatility similar to trading with a stop-loss)

CTAs seek to rank trends by volatility (strong trends with small volatility)

CTAs attract much more investments than volatility funds

\[ y = 0.3398x^2 - 0.0733x + 0.0038 \]

\[ R^2 = 0.0214 \]
CTAs vs volatility strategies

CTAs derive its convex pay-off from a positive auto-correlation (Bouchaud et al (2016))

Replication costs of CTAs are linked to the short-term realized volatility

Long options strategy can create a similar convexity profile but its replication costs are derived from implied volatilities, which are expensive

Short options strategy works well when auto-correlations are negative (no trend) and the implied volatility is expensive

Combination of CTAs with short volatility strategies can create a more desirable risk profile than each of the alone
Why it is so difficult to make profits being long volatility

Being long volatility requires for a trader to make an intelligent assessment about the trend of the underlying:

- In a strong trending market, the trader should hedge infrequently (let the delta-risk to accumulate)
- In a choppy range-bound market, the trader should hedge very frequently (reduce the delta-risk fast)

It is not only enough to estimate the expected realized volatility

If options are purchased on the buy-to-hold basis without delta-hedging, the trade can make money only for a strong trend in the underlying:

- Recall the cyclicality of the volatility risk-premium: the timing ability is crucial
- Longer-dated options to reduce the timing risk along with pre-defined profit taking
The convexity profile of returns can also be created by using the statistical volatility as a risk control

Figure: the convexity profile of a proprietary risk-parity strategy

1. Statistical volatility is typically negatively correlated to expected return: apply the estimated volatility for asset allocation
2. Key idea behind the risk-parity and minimum volatility funds
3. These strategies tend to outperform over long-term
4. Very strong interest and inflows by the investment community

\[ y = 1.2332x^2 + 0.2084x + 0.0015 \]
\[ R^2 = 0.299 \]
Conclusions: the cyclicality of markets dynamics

The classic derivative theory deals with option replication assuming ideal conditions and using models that are static in nature.


Classic quantitative investment strategies (trend-following and volatility trading) seek for a statistical arbitrage of these anomalies.

The volatility risk-premium is part of a factor-based investment approach.

Quantitative strategies using statistical volatility as a risk-control can also generate the convexity profile of long option strategies:

Strong investor demand for minimum volatility and risk-parity products.


Disclaimer

The views represented herein are the author own views and do not necessarily represent the views of Julius Baer or its affiliates.