Volatility

BUSS386. Futures and Options

Prof. Ji-Woong Chung

Lecture Outline

- σ is the only unobservable BSM input
- Estimating vol from history: realized, EWMA, GARCH
- Implied vol (IV); VIX and VKOSPI; variance swaps
- The volatility risk premium (bridge to Lec 11)
- Smile, skew, and term structure
- Volatility surface; practitioner BSM
- Bonus: what the smile tells you about the RN distribution

σ Is the Only Unobservable BSM Input

$$V_{\text{BSM}}(S_0, K, T, r, \sigma): \quad \text{5 inputs, 4 observed.}$$

- S_0, K, T, r (and q) all read off a screen.
- σ has to be **estimated** or **implied** — two different things.

Also: BSM assumes σ is constant. Reality:

- **Time-varying.** Volatility clusters (high follows high), mean-reverts, and reacts asymmetrically to price moves (leverage effect).
- **Strike-dependent.** Different IVs across strikes ⇒ the smile/skew.
- **Maturity-dependent.** Term structure of IV.

Two roads.

1. *Estimate* σ from past returns: historical, EWMA, GARCH.
2. *Imply* σ from market option prices: IV, smile, surface.

Stylized Facts About Volatility

- **Not constant.** Vol fluctuates substantially; KOSPI 200 realized vol has ranged from $\sim$10% (2017) to $\sim$45% (2008, 2020 Mar, 2024 Aug intraday).
- **Clustering.** Big move today $\Rightarrow$ probability of big move tomorrow is higher. (Mandelbrot 1963; the GARCH motivation.)
- **Mean reversion.** Extreme highs and lows pull back to a long-run average over months.
- **Leverage effect.** Equity vol *rises* when prices fall and *falls* when prices rise. Reverse pattern for some commodities.
- **Forecastability.** 1-day vol is very noisy; 1–3 month vol is much more predictable (mean reversion).

Each fact has a model. EWMA captures clustering crudely; GARCH adds mean reversion; SV models add the leverage effect; jump models add tail risk.

Historical (Realized) Volatility

Daily log returns from past prices:

$$R_t = \ln(S_t/S_{t-1}).$$

Daily variance estimator, assuming zero mean (sample mean is too noisy to subtract):

$$\hat{\sigma}^2_{\text{daily}} = \frac{1}{K}\sum_{i=1}^{K} R^2_{t-i}.$$

Annualize:

$$\hat{\sigma}_{\text{annual}} = \hat{\sigma}_{\text{daily}}\sqrt{h}, \quad h = 252 \text{ (daily)},\ 52 \text{ (weekly)}.$$

- **Assumes:** constant vol over the look-back window. Cleanest with a window short enough that the regime didn't change but long enough to average noise.
- **Treats all days equally.** A return 60 days ago counts as much as yesterday's — a problem given clustering.

Example — Historical Volatility

Q. Daily closing prices: 100, 102, 101, 105, 103, 106. Compute the annual realized vol.

A. Daily log returns and their squares:

t	1	2	3	4	5
R_t	0.01980	-0.00985	0.03884	-0.01923	0.02871
R_t^2	0.000392	0.000097	0.001509	0.000370	0.000824

$$\sum R_t^2 = 0.003192 \quad \Rightarrow \quad \hat{\sigma}^2_{\text{daily}} = 0.003192/5 = 0.000638.$$

$$\hat{\sigma}_{\text{daily}} = \sqrt{0.000638} = 0.02527 \quad \Rightarrow \quad \hat{\sigma}_{\text{annual}} = 0.02527\sqrt{252} = \boxed{40.1\%}.$$

EWMA — Weighting Recent Data More

Exponentially decaying weights w^k with $0 < w < 1$:

$$\hat{\sigma}^2_{\text{daily, EWMA}} = \frac{\sum_{k=0}^{K} w^k R^2_{t-k}}{\sum_{k=0}^{K} w^k}.$$

RiskMetrics convention: $w = 0.94$ for daily data

Example. Same 5 returns, indexed from most recent to oldest:

k	0	1	2	3	4
R^2_{t-k}	0.000824	0.000370	0.001509	0.000097	0.000392
w^k	1.0000	0.9400	0.8836	0.8306	0.7807

$$\sum w^k = 4.4349, \quad \sum w^k R^2 = 0.002892 \;\Rightarrow\; \hat{\sigma}^2_{\text{daily}} = 0.000652.$$

$$\hat{\sigma}_{\text{annual}} = \sqrt{0.000652 \times 252} = \boxed{40.5\%}.$$

Close to historical (the most recent return happens to be average); the difference grows when there's a recent spike.

GARCH(1,1) — Capturing Clustering + Mean Reversion

Engle 1982, Bollerslev 1986.

$$r_t = \mu + \epsilon_t, \quad \epsilon_t \sim (0, \sigma_t^2), \qquad \sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2.$$

Three parameters:

- $\alpha =$ reaction to new shocks (*volatility jumps*).
- $\beta =$ persistence (*volatility clusters*).
- $\omega =$ baseline; long-run variance $\sigma_\infty^2 = \omega/(1 - \alpha - \beta)$.

Typical empirical values: $\alpha \approx 0.05$–0.10, $\beta \approx 0.90$–0.95, $\alpha + \beta < 1$ (mean reversion).

Why it beats EWMA. EWMA has *no long-run mean* — vol forecasts follow the latest data forever. GARCH has σ_∞, so forecasts *revert*. This matters at horizons of weeks to months.

Forecast formula. *h*-step-ahead variance: $\sigma_{t+h}^2 = \sigma_\infty^2 + (\alpha + \beta)^h(\sigma_t^2 - \sigma_\infty^2)$. Decays geometrically toward σ_∞^2.

GARCH(1,1) — Example

Setup. $\mu = 0$, $\alpha = 0.05$, $\beta = 0.94$, $\sigma^2_\infty = 0.000638$ (matched to historical) $\Rightarrow$ $\omega = (1 - 0.99) \times 0.000638 = 6.38 \times 10^{-6}$.

Start at $\sigma^2_1 = \sigma^2_\infty = 0.000638$. Take $\epsilon_t = R_t$ (mean zero).

Step 1. $\epsilon_1 = 0.01980$, $\epsilon^2_1 = 0.000392$.

$$\sigma^2_2 = 6.38\times10^{-6} + 0.05(0.000392) + 0.94(0.000638) = \boxed{0.000626}.$$

$$\sigma_2 = \sqrt{0.000626} = \boxed{2.50\%/\text{day}}.$$

Step 2. $\epsilon_2 = -0.00985$, $\epsilon^2_2 = 0.000097$.

$$\sigma^2_3 = 6.38\times10^{-6} + 0.05(0.000097) + 0.94(0.000626) = \boxed{0.000600}.$$

$$\sigma_3 = \sqrt{0.000600} = \boxed{2.45\%/\text{day}}.$$

Reading. A small shock $\Rightarrow$ vol barely moved. A big shock would push σ^2_{t+1} noticeably up, then it would mean-revert at rate $(\alpha + \beta) = 0.99/\text{day}$.

Comparing the Three Methods

	Historical	**EWMA**	**GARCH(1,1)**
Idea	Sample variance of past returns	Recency-weighted variance	Recursive update with persistence
Captures clustering?	No	Yes (crudely)	Yes
Mean-reverting forecast?	No	No	Yes
Param. to choose	Window K	Decay w	Estimate α, β, ω

Bottom line.

- For **long-window** forecasts: simple historical is hard to beat.
- For **1–10 day** forecasts: GARCH (or EWMA) typically wins.
- All three are used in **risk management** (VaR, ES, stress).

Practical Issues in Volatility Estimation

Question	Standard practice
Daily / weekly / intraday?	Daily is the workhorse; intraday introduces microstructure noise.
Subtract the mean?	No — assume mean zero (sample mean is noisier than helpful).
How long a window?	30–250 days; longer for stability, shorter for regime changes.
Handle outliers (1987, 2020, 2024)?	No universal rule. Examine sensitivity; report with/without.
Out-of-sample evaluation?	Required. In-sample fit is not predictive of future performance.

Simple historical over a long window performs about as well as complex alternatives *at long horizons*, and is far more robust to model error.

Implied Volatility — The Pricing Workhorse

Definition. Implied volatility is the σ that makes BSM match the market price:

$$C_{\text{mkt}} = C_{\text{BSM}}(S_0, K, T, r, \sigma_{\text{imp}}).$$

Solve by root-finding (Newton-Raphson on σ).

- One-to-one mapping price $\leftrightarrow$ IV. Higher price $\Rightarrow$ higher IV.
- In equity index / FX markets, traders **quote vol, not prices.** "the 25-delta put trades at 18 vol" is the dealer language.
- IV embeds expectations *plus* risk premia — it is not an unbiased forecast of future realized vol (see next).

IV is the same for European calls and puts at the same (K, T).

The Volatility Risk Premium — Bridge to Lec 11

Empirical fact. Implied vol *systematically exceeds* realized vol on equity indices, on average.

Underlying	Avg IV (1-mo)	Avg realized (1-mo)	Premium
S&P 500 (2000–2024)	$\sim$19%	$\sim$15%	$\sim$4 pp
KOSPI 200 (2010–2024)	$\sim$18%	$\sim$14%	$\sim$4 pp

Why? Investors pay a premium to *insure* against drawdowns. The seller (a market maker, or a covered-call writer) earns the premium in exchange for taking on tail risk.

Tying back to Lec 11. A long-gamma delta-hedged position earns

$$\text{P\&L} \;\approx\; \tfrac{1}{2}\Gamma \sum_i (\Delta S_i)^2 - |\Theta|\, T \;\propto\; (\sigma_{\text{real}}^2 - \sigma_{\text{imp}}^2)\, T.$$

Long-gamma loses, on average, because realized $<$ implied. Short-gamma (option sellers, ELS issuers) earns the premium — and pays the price on big moves.

VIX and VKOSPI — Vol Indices and Their Derivatives

Definition. A volatility index is the market's expectation of 30-day annualized vol, extracted from a strip of options.

	VIX (CBOE)	VKOSPI (KRX)	VSTOXX (Eurex)
Underlying	SPX options	KOSPI 200 options	Euro Stoxx 50
Launched	1993 (orig.), 2003 (current)	2009 (idx), 2014 (fut)	2005
Futures?	Yes (2004)	Yes (2014, KRX)	Yes

Key features of vol-index derivatives.

- VIX (or VKOSPI) is *not a tradable asset* — you can't store it. No cost-of-carry pricing.
- Futures price = market's expectation of where the index will be at expiry
- Options on VIX/VKOSPI are options on the *futures*, not on the index level. (Spot would be impossible to hedge.)

The VIX Index

Variance Swaps — Trading Volatility

Payoff at T:

$$\text{Payoff} = \text{Notional} \times \big(\text{Realized variance} - K_{\text{var}}\big).$$

Realized variance $= \frac{252}{N}\sum_{t=1}^{N} R_t^2$ (annualized, daily sampling).

- Strike K_{var} quoted as a variance (e.g., $0.20^2 = 0.04$).
- Buyer gains if realized vol exceeds the strike — a long-vol position.

Example. Notional \$1M, $K_{\text{var}} = 0.04$ (20% vol), realized variance ends at 0.06 (24.5% vol):

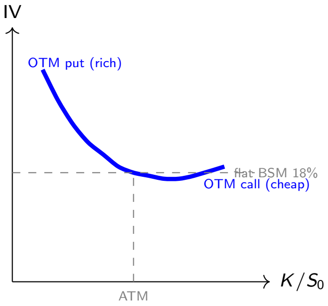
$$\text{Payoff} = \$1\text{M} \times (0.06 - 0.04) = \boxed{+\$20{,}000}.$$

Volatility swap. Same idea, but payoff in vol units, not variance. More intuitive to investors, less clean to hedge — not perfectly replicable because $\sqrt{\cdot}$ is concave. Quoted by dealers but less liquid than variance swaps.

BSM vs Market — The Smile / Skew

If BSM were right, all options on the same underlying would imply the *same* σ.

Equity index reality. Plotting market IV against strike:



Reading. Low-strike (OTM puts, ITM calls) have *higher* IV than ATM; high-strike (OTM calls) have *lower*. This is the **equity skew** (also "smirk"). KOSPI 200, S&P 500, HSCEI: all the same shape.

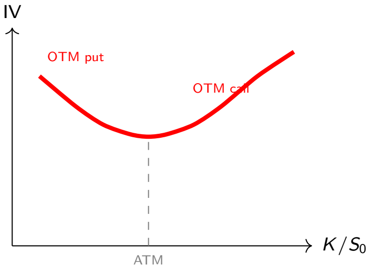
Translation: OTM puts are expensive relative to flat BSM.

Why the Equity Skew Exists

Three complementary explanations, all true at once:

1. **Leverage effect.** As stock prices fall, leverage rises, and realized vol rises. The market prices this into IV: low strikes (where stress would push us) carry higher IV.
2. **1987 crash changed pricing.** Pre-1987 equity options showed little skew. Black Monday's -22% one-day drop forced traders to demand more for OTM puts. The skew became permanent and steep.
3. **Hedging demand.** Pensions and mutual funds buy OTM puts as portfolio insurance. Steady demand $\Rightarrow$ persistent richness in the low-strike wing.

FX — A Symmetric Smile, Not a Skew



Stylized USD/KRW smile.

FX options typically show a **symmetric smile**, not a downward-sloping skew.

- No natural "crash direction" for a currency pair — both wings can move sharply.
- Smile usually slightly tilted depending on rate-differential expectations.
- Quoting convention: "25-delta risk reversal" = IV(call) − IV(put); positive = call wing richer (e.g., USD/JPY in safe-haven flows).

Korean case. USD/KRW exhibits a smile with the OTM put (KRW strengthening) sometimes slightly higher than OTM call — reflects exporter hedging flows.

Implied Volatility for Pricing

- **The volatility smile reveals a limitation of the Black–Scholes model.** If BSM were correct, all options would share the *same* volatility. In reality, implied volatility varies by strike and maturity.
  - This means implied volatility contains market information about future uncertainty, but it is not an *unbiased* forecast of future realized volatility.
  - IV reflects both expectations *and* risk premia demanded by investors.
- **More sophisticated models can generate a smile.** Examples include: stochastic volatility models (Heston), local volatility models (Dupire), models with jumps or stochastic interest rates (see Appendix for discussion).
- **In practice: market makers use "practitioner Black–Scholes."**
  - They keep the BSM formula but replace the constant volatility assumption.
  - Each option gets its own volatility input taken from the **implied volatility surface**.
  - This ensures that model prices match observable market prices.

Practitioner BSM — Plug a Different σ at Each Strike

Workflow on an options desk:

1. Observe liquid option prices $\rightarrow$ invert BSM to get IV at each strike.
2. Fit a smooth IV surface across strikes and maturities.
3. Price illiquid/OTC/exotic options using the surface, not a single σ.
4. Compute Greeks from the BSM formula *with the smile-adjusted* σ.

Crucially: this is *not* a coherent model — it's a curve-fitting device. The IV surface depends on S_t and t; a true model would specify *joint* dynamics. Stochastic-vol (Heston) and local-vol (Dupire) are attempts at coherence. "Practitioner BSM" is the everyday tool desks actually use.

Trade-off: matches market prices exactly today; doesn't tell you how the smile will move tomorrow.

Example — Smile-Adjusted Equity Put Price

Q. Equity $S_0 = 100$, $T = 0.25$, $r = 2\%$, no dividend. Observed equity skew:

K	80	90	100	110	120
IV	28%	24%	20%	18%	17%

Price a 90-strike put two ways: (i) flat ATM IV = 20%, (ii) smile IV = 24%.

Flat 20%: $d_1 = (\ln(100/90) + 0.04 \cdot 0.25)/(0.20\sqrt{0.25}) = 0.1176/0.10 = 1.176$; $d_2 = 1.076$. $N(-d_1) = 0.120$, $N(-d_2) = 0.141$.

$p_{\text{flat}} = 90e^{-0.005}(0.141) - 100(0.120) = \boxed{\$0.63}$.

Smile 24%: $d_1 = 0.1176/0.12 = 0.980$; $d_2 = 0.860$. $N(-d_1) = 0.163$, $N(-d_2) = 0.195$.

$p_{\text{smile}} = 90e^{-0.005}(0.195) - 100(0.163) = \boxed{\$1.16}$.

A More Sophisticated Practitioners' Approach

- **Goal:** Compare the relative value of many options at once (different strikes and maturities).
- **Problem:** Raw option prices cannot be compared directly.
  - Different strikes $\rightarrow$ different intrinsic values.
  - Different maturities $\rightarrow$ different time value, interest rates, and uncertainty.
  - Price differences alone do not indicate whether one option is "expensive" or "cheap."
- **Solution:** Convert all option prices into *implied volatilities*. IV normalizes price differences by expressing everything in volatility units:

$$\text{IV} = \text{“how much volatility is the market implying?”}$$

This allows apples-to-apples comparison across all strikes and maturities.

A More Sophisticated Practitioners' Approach

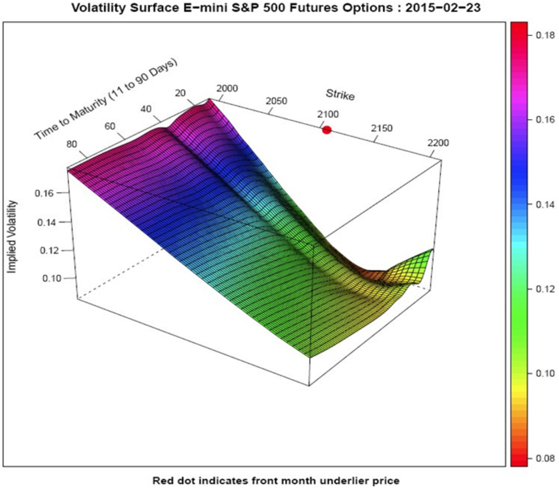
Volatility Surface

- Once IVs are computed across all strikes and maturities, practitioners arrange them into a **volatility surface**:

$$\sigma = \sigma\left(\frac{K}{S_0}, T\right)$$

- This surface summarizes:
  - the skew/smile (variation across strike), and
  - the term structure (variation across maturity).
- The volatility surface becomes the central tool for:
  - pricing OTC and exotic options,
  - computing Greeks,
  - fitting models (local vol, stochastic vol),
  - relative value trading (identifying cheap/expensive options).

The Volatility Surface — Smiles Across Maturities



A More Sophisticated Practitioners' Approach

Examples

- **To value a new option, locate its position on the volatility surface.**

$$\sigma^* = \sigma\left(\frac{K}{S_0}, T\right)$$

Table 20.2 Volatility surface.

	K/S_0				
	0.90	*0.95*	*1.00*	*1.05*	*1.10*
1 month	14.2	13.0	12.0	13.1	14.5
3 month	14.0	13.0	12.0	13.1	14.2
6 month	14.1	13.3	12.5	13.4	14.3
1 year	14.7	14.0	13.5	14.0	14.8
2 year	15.0	14.4	14.0	14.5	15.1
5 year	14.8	14.6	14.4	14.7	15.0

A More Sophisticated Practitioners' Approach

Examples

- **Example 1: 9-month, $K/S_0 = 1.05$ (OTM call).**
  - Table gives IVs: 13.4% (0.5 year) and 14.0% (1.0 year).
  - Interpolate along maturity to get: $\sigma^* \approx 13.7\%$
  - Use σ^* in BSM or a binomial tree to compute the option value.
- **Example 2: 1.5-year, $K/S_0 = 0.925$ (ITM put).**
  - Strike dimension → interpolate across K/S_0.
  - Maturity dimension → interpolate across T.
  - Apply **bilinear interpolation** (2D). Result:$\sigma^* \approx 14.525\%$
  - Use this IV to price the option.
- **Alternative approach: fit a regression-based IV surface.**
  - Estimate:

$$\hat{\sigma} = a + b\left(\frac{K}{S_0}\right) + cT + d\left(\frac{K}{S_0}\right)^2 + e\left(\frac{K}{S_0}T\right) + fT^2.$$

  - Smooths noisy market IVs and avoids jumps in the surface.

Summary

- Despite its inaccuracies BSM serves as a useful benchmark.
  - Gives decent approximation to prices close to the money.

- It also works reasonably well to hedge options positions against changes in stock prices using delta or delta-gamma hedging.

- Models have been proposed to correct some of the shortcomings.
  - Stochastic volatility
  - Jumps
  - Fat tails

- All of these models are consistent with the idea that OTM puts are expensive relative to BSM prices because investors seeking protection from large losses (e.g., jumps down) must pay a higher (insurance) premium

Appendix: Volatility Models
(Chapter 27)

Local Volatility Models

- **Idea:** The Black–Scholes model assumes volatility is constant. Local volatility models relax this by allowing volatility to depend on the **current stock price** and **time**:

$$dS_t = \mu S_t dt \; + \; \sigma(S_t, t) \, S_t dW_t.$$

- One popular example is the **Constant Elasticity of Variance (CEV)** model:

$$dS_t = \mu S_t dt + (\sigma S_t^{\gamma-1}) S_t dW_t.$$

- **Interpretation of γ:**
  - $\gamma = 1$: reduces to the Black–Scholes model (constant volatility).
  - $\gamma < 1$:
    - When S_t falls, S_t^{γ} increases relative to S_t,
    - $\Rightarrow$ volatility rises at low prices,
    - $\Rightarrow$ OTM puts become more valuable,
    - $\Rightarrow$ **volatility smirk (equity skew)** appears.
  - $\gamma > 1$: volatility rises with price (sometimes seen in futures/options on commodities).
- **Takeaway:** Local volatility models can generate a smile or smirk because volatility changes depending on where the stock price is.

Stochastic Volatility Models

- **Idea:** Volatility itself moves randomly over time, rather than being constant. This helps explain why options imply different volatilities at different strikes.
- **The Heston Model** is a leading example:

$$dS_t = \mu S_t dt + \sqrt{v_t}\, S_t dW_t,$$
$$dv_t = \theta(\omega - v_t)dt + \xi\sqrt{v_t}\, dB_t.$$

- **Meaning of parameters:**
  - ω: long-run average variance.
  - θ: speed of mean reversion (how fast variance returns to ω).
  - ξ: "vol of vol" – how much variance itself fluctuates.
  - ρ: correlation between stock returns (dW_t) and volatility shocks (dB_t).
- **Why Heston explains the smirk:**
  - If $\rho < 0$ (empirically true for equities): bad market moves $\rightarrow$ higher volatility.
  - Higher volatility $\rightarrow$ higher crash probability.
  - This makes OTM put options relatively expensive.
  - $\Rightarrow$ **downward-sloping volatility skew**.
- **Intuition:** A drop in price increases volatility, which increases the chance of an even bigger drop.

Jumps in Stock Prices

- **Idea:** Stock prices occasionally experience sudden large jumps (e.g., 1987 crash, 2020 COVID crash). The Black–Scholes model cannot capture this.
- A jump–diffusion model adds a jump component to price movements:

$$dS_t = \mu S_t dt + \sigma S_t dW_t + J(Q)\, S_t\, dP_t.$$

- **Interpretation:**
  - dP_t is usually 0, but equals 1 with a small probability $\rightarrow$ a jump occurs.
  - $J(Q)$ is the jump size (can be random or fixed).
  - If $J(Q) < 0$, the jump is downward $\rightarrow$ a crash.
- **Impact on options:**
  - If downward jumps are possible, tail risk increases.
  - OTM puts become much more valuable.
  - $\Rightarrow$ **steeper volatility smirk**.
- **Downside jump risk = expensive insurance.**
- **Pricing jumps is harder**, because Black–Scholes cannot be used directly. More advanced numerical methods are required (Monte Carlo, Fourier transforms, etc).

Implied Tree Models

- **Idea:** Normally, we start with a model (e.g., binomial tree) $\rightarrow$ compute option prices. With **implied trees**, we reverse the process:

$$\text{Use observed option prices} \;\Rightarrow\; \text{infer the stock price tree.}$$

- Once calibrated, the implied tree:
  - matches market option prices exactly,
  - produces a local volatility surface implicitly,
  - can be used to price other options consistently.

Implied Tree Models

- **Example:** Given:

$$S_0 = 1502.39,\ K = 1500,\ \sigma = 12.36\%,\ r = 4.713\%,\ \delta = 1.91\%,\ T = 0.12.$$

  Compute initial binomial parameters:

  - $u = e^{\sigma\sqrt{T}} = 1.0437, \quad d = 1/u = 0.9581.$
  - Risk-neutral $p = \dfrac{e^{(r-\delta)T} - d}{u - d} = 0.5286.$
  - BSM/binomial price: $c = 28.394$.
  - Market price: $c^{mkt} = 20.35$ (model overprices).

- **Implied-tree step:** Adjust σ (and therefore u and d) until model = market:

$$\sigma = 8.24\% \quad (\text{new } p = 0.5446)$$

  Now the tree is calibrated $\rightarrow$ can be used to price other options.

Other Modern Volatility Models

- **Beyond local volatility, stochastic volatility, and jump models,** modern quantitative finance uses several advanced models to better match the observed volatility smile and surface.
- **SABR Model (Hagan, Kumar, Lesniewski, Woodward, 2002)**
  - Widely used in interest-rate and FX markets.
  - Models both the asset price and its volatility as stochastic processes.
  - Flexible enough to generate smiles, skews, and term structure effects.
  - Provides simple formulas that traders can implement quickly.
- **Rough Volatility Models (Gatheral, Jaisson, Rosenbaum, 2018)**
  - Based on the empirical finding that volatility moves "roughly," showing long memory and very jagged paths.
  - Captures the fine structure of volatility better than classical models.
  - Produces realistic short-term smiles and accurate VIX dynamics.
- **Why these models matter in practice:**
  - They fit market implied volatility surfaces more accurately.
  - They improve pricing of exotic options and risk management.
  - They help traders understand how the smile evolves over time.