Volatility Modeling: The Measurement Paradox

Measuring risk changes risk – and the map is never the territory

Digital Finance

Why Do Markets Alternate Between Long Calms and Sudden Storms?

The Clustering Puzzle

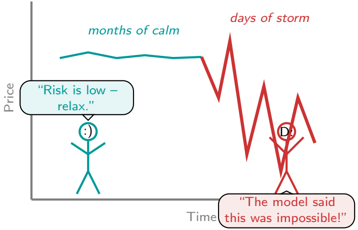
Markets spend months barely moving – then lose years of gains in days. This is not random. Volatility *clusters*: large moves follow large moves, small moves follow small moves. A model that assumes constant risk will miss both the calm and the storm.

What constant-volatility models assume:

- Tomorrow's risk is the same as today's
- Extreme events are vanishingly rare (normal tails)
- Past turbulence tells you nothing about future turbulence
- A single number (annualized vol) captures all risk

What markets actually do:

- Volatility persists: high-vol days predict high-vol days
- Fat tails: extreme moves are 5–10x more frequent than normal
- Leverage effect: negative returns increase future volatility
- Regime switches: calm-to-storm transitions are abrupt



Risk is not constant.
The calm is the setup for the storm.

Volatility clustering is the most robust stylized fact of financial returns – large moves predict large moves, and constant-risk models miss it entirely.

What Does It Feel Like to Lose a Year of Gains in a Single Afternoon?

Reflection Prompt

Your portfolio gained 8% over twelve steady months. Then, on one afternoon, a volatility spike erased all of it. Your risk model had said there was a 0.01% chance of this happening. Was the model wrong – or is 0.01% just another way of saying "not today, but eventually"?

The honest answer: the model was not wrong about direction – it was wrong about magnitude. Constant-volatility models systematically underestimate how bad "bad" can get.

Three things that feel different from what models predict:

- **Speed:** Losses happen faster than gains. A year of steady appreciation can vanish in hours because volatility spikes are instantaneous, not gradual
- **Asymmetry:** Markets fall faster than they rise. Negative returns amplify future volatility (the leverage effect), creating a self-reinforcing downward spiral
- **Contagion:** When one asset crashes, correlations spike across all assets. Diversification fails exactly when you need it most – during volatility regime switches

Bring to class: Find the worst single-day loss for the S&P 500 in each decade since 1980. How many standard deviations was each? Does the normal distribution explain any of them?

The gap between what risk models predict and what markets deliver is not a bug – it is the central challenge of volatility modeling.

What Is Volatility – and Why Are There At Least Three Ways to Measure It?

Dimension	Historical Vol	Implied Vol	Realized Vol
Source	Past returns	Option prices	Intraday tick data
Window	20–252 days	Forward-looking	Single day (5-min)
Formula	$\sqrt{\frac{1}{T}\sum r_t^2}$	Black-Scholes inversion	$\sqrt{\sum r_i^2}$ (HF)
Updates	Daily	Continuous	Continuous
Captures clustering	Slowly (lagged)	Instantly (priced in)	Instantly (measured)
Primary user	Risk managers	Options traders	Quant researchers

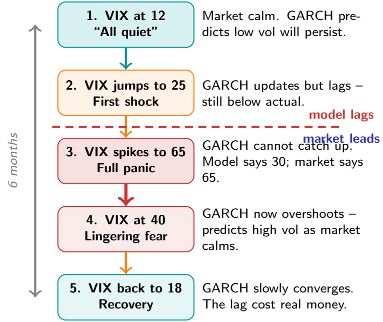
Three lenses on the same phenomenon

- **Historical:** Looks backward. Simple and universal but slow to react. A 20-day window misses intraday regime changes entirely.
- **Implied:** Looks forward. Extracted from option prices, it reflects the market's consensus expectation of future volatility. But it includes a risk premium – implied vol systematically exceeds realized vol.
- **Realized:** Looks at the present with a microscope. Computed from intraday returns (e.g., 5-minute intervals), it captures within-day volatility dynamics. But microstructure noise contaminates the estimate at very high frequencies.

Key insight: No single measure is "correct." Each answers a different question: What happened? What does the market expect? What is happening right now?

Historical, implied, and realized volatility measure the same thing differently – the choice depends on whether you need hindsight, foresight, or real-time measurement.

Follow the VIX Through One Market Crash – and Watch Predicted vs Actual Diverge



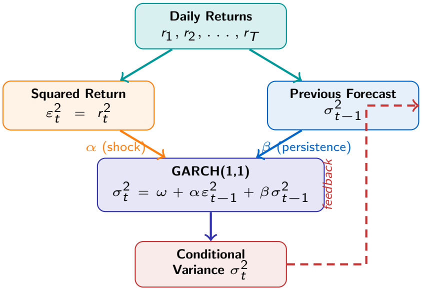
Why GARCH lags during crises

- **GARCH updates incrementally:** Each day's forecast blends yesterday's forecast with today's squared return. During a spike, it takes many days to "catch up" because the blend weight on old data is high.
- **The VIX jumps instantly:** Option prices react within seconds to new information. The VIX (implied vol) moves to the new regime immediately while GARCH is still averaging over the old one.
- **The lag is asymmetric:** GARCH is slow to rise during panic and slow to fall during recovery. It overshoots on both sides.
- **Practical cost:** A risk manager using GARCH alone would have set position limits too loosely before the crash and too tightly during the recovery.

Takeaway: GARCH captures the *persistence* of volatility but misses the *jumps*. Combining it with implied volatility compensates for the lag.

GARCH is a rearview mirror – it tells you where volatility was, not where it is going. During regime switches, it systematically lags the market.

How Does a GARCH Model Capture the Fact That Volatility Breeds Volatility?



Stationarity: $\alpha + \beta < 1$
Persistence: $\alpha + \beta \approx 0.97$ *typical*

Three parameters, one feedback loop

- ω (omega): the long-run baseline variance. When nothing unusual happens, variance reverts here.
- α (alpha): the **shock weight**. How much does today's squared return affect tomorrow's forecast? High α means the model reacts quickly to surprises.
- β (beta): the **persistence weight**. How much does yesterday's forecast carry forward? High β means volatility decays slowly.

Why $\alpha + \beta < 1$ matters:

- If $\alpha + \beta \geq 1$, shocks never die out – the variance process is non-stationary (IGARCH)
- Typical equity estimates: $\alpha \approx 0.05$, $\beta \approx 0.92$, so $\alpha + \beta \approx 0.97$
- The half-life of a volatility shock is $\ln(2)/\ln(\alpha + \beta) \approx 23$ days

The feedback loop is the key: tomorrow's forecast depends on today's forecast, which depended on yesterday's. Volatility breeds volatility – that is clustering.

GARCH captures clustering through a single feedback loop: today's variance forecast feeds into tomorrow's. Three parameters control baseline, reactivity, and persistence.

What Happens When Everyone Uses the Same Volatility Model – and the Model Is Wrong?

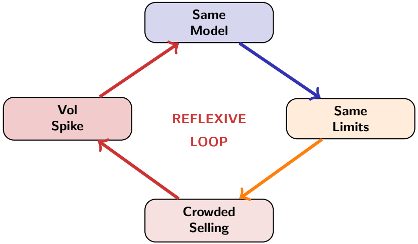
When Models Become Self-Fulfilling
If every risk manager uses the same GARCH model, they all set the same position limits. When volatility spikes, they all reduce exposure simultaneously – amplifying the very crash the model was supposed to predict.

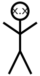
Model risk failure modes:

- **Crowded hedging:** VIX-linked products force delta-hedging that amplifies moves. In Feb 2018, short-vol ETPs lost 90% in one day ("Volmageddon")
- **Correlation breakdown:** GARCH models individual assets. During crises, correlations spike to 1.0 and portfolio-level risk explodes beyond any single-asset model
- **Regime blindness:** GARCH assumes one regime. Markets have at least two (calm and crisis). A model estimated in calm systematically underestimates crisis risk
- **Fat-tail denial:** Standard GARCH with normal innovations assigns probability 10^{-9} to a 5-sigma event. Equity markets produce one roughly every 3 years

Volmageddon (Feb 5, 2018):

- VIX doubled in one afternoon (17 to 37)
- XIV (inverse VIX ETN) lost 96%, was terminated
- Cause: reflexive feedback between vol sellers and delta-hedging algorithms, not a fundamental shock

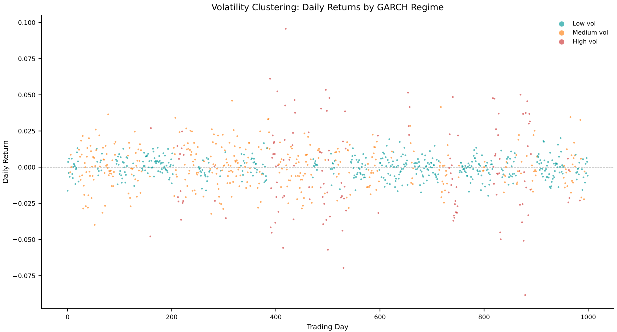




When everyone hedges the same risk the same way, the hedge becomes the risk.

Model risk is not about individual errors – it is about systemic herding. When every institution uses the same model, the model itself becomes a source of instability.

Where Does Volatility Cluster – and What Triggers Regime Switches?



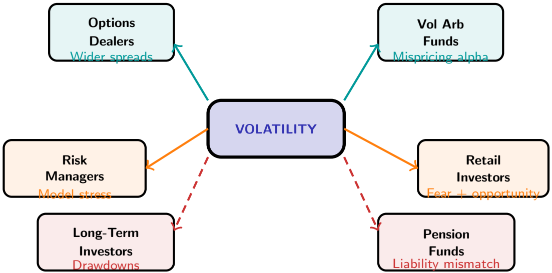
Reading the scatter plot

- Each dot is one trading day's return, colored by its volatility regime (low, medium, high)
- **Clustering is visible:** teal dots (calm) group together in long stretches; red dots (storm) appear in bursts. Regimes persist before switching abruptly
- **Transition asymmetry:** the switch from calm to storm is sudden (a few days), while the return to calm is gradual (weeks to months)
- **Amplitude scaling:** high-regime returns are roughly 4x larger than low-regime returns, matching the ratio of GARCH conditional standard deviations
- **GARCH captures the persistence** within each regime but struggles at the transition boundaries where the regime itself changes

Implication: A risk model that ignores regimes will underestimate risk during storms and overestimate risk during calms – wrong in both directions.

Illustrative regime-switching simulation. Volatility clusters because regimes persist – and the transitions between regimes are the most dangerous moments for risk models.

Who Benefits from Volatility – Dealers, Quants, or Long-Term Investors?



Winners:

- **Options dealers** earn wider bid-ask spreads when implied vol rises; their inventory gains value
- **Vol arb funds** exploit the gap between implied and realized vol for systematic alpha

Losers:

- **Long-term investors** suffer drawdowns that compound asymmetrically (a 50% loss requires 100% gain)
- **Pension funds** face liability mismatches when discount rates move with vol spikes

Mixed:

- **Risk managers** are most needed during vol spikes but their models are least reliable then
- **Retail investors** panic-sell at peaks but contrarians who buy vol dips earn long-run premia

Volatility is a cost for those who hold assets and a revenue source for those who trade it. The winners are those who can price volatility; the losers are those who endure it.

The Volatility Estimation Checklist: When Should You Trust the Model and When Should You Override It?

When using any volatility model for risk management or trading decisions, ask these four questions:

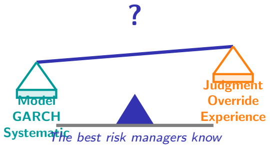
1. Is the regime stable or transitioning?
GARCH works well within a regime but lags at transitions. If the VIX has moved more than 5 points in a week, consider overriding the model with implied vol.

2. Are tails being modeled or assumed away?
Standard GARCH with normal innovations underestimates tail risk by orders of magnitude. Use Student-t or GED innovations, or check realized tail frequencies.

3. Is the model capturing asymmetry?
Symmetric GARCH misses the leverage effect. Use EGARCH or GJR-GARCH if negative returns amplify vol more than positive returns of equal size.

4. What is the model's half-life?
If $\alpha + \beta = 0.97$, shocks decay with a 23-day half-life. If your trading horizon is shorter, the model may be too slow; if longer, it may overreact to noise.

The meta-rule: No single volatility model works in all regimes. The checklist helps you know when to trust the model and when to augment it with judgment.



The best risk managers know when to trust the model – and when to override it.

A volatility model is a tool, not an oracle. The checklist helps you calibrate your trust in the model to the current market regime.

Your Challenge: Estimate GARCH Parameters from Return Data and Forecast 5-Day Volatility

Mini-Challenge (15 minutes)

You are given 500 daily returns for a stock index. The sample variance is 0.0001 (1% daily vol). Yesterday's return was −3% (a 3-sigma event). Your GARCH(1,1) parameters are $\omega = 2 \times 10^{-6}$, $\alpha = 0.06$, $\beta = 0.92$. Forecast conditional variance for the next 5 days.

Step-by-step:

1. **Day 1 forecast:** $\sigma_1^2 = \omega + \alpha \cdot (-0.03)^2 + \beta \cdot 0.0001 = 2 \times 10^{-6} + 0.06 \times 0.0009 + 0.92 \times 0.0001 = 0.000150$ (daily vol = 1.22%)
2. **Days 2–5:** Iterate $\sigma_{t+1}^2 = \omega + (\alpha + \beta) \cdot \sigma_t^2$ (since $E[\varepsilon_t^2] = \sigma_t^2$). The forecast decays toward the unconditional variance $\bar{\sigma}^2 = \omega/(1 - \alpha - \beta) = 0.0001$.
3. **5-day volatility:** $\sigma_{5d} = \sqrt{\sum_{t=1}^{5} \sigma_t^2}$. Compare this to the naive $\sqrt{5} \times 0.01 = 2.24\%$. The GARCH forecast should be higher because the 3-sigma shock elevated near-term risk.

Discussion: How many days until the GARCH forecast returns within 10% of the unconditional variance? What does the half-life $\ln(2)/\ln(0.98) \approx 34$ days tell you about risk management?

This exercise shows GARCH in action: a single large return elevates the variance forecast for weeks, then it slowly decays back to baseline. That decay is volatility clustering.