

# **A Stylized History of Quantitative Finance**

Emanuel Derman  
Columbia University

# Outline

The evolution of a quantitative approach to finance has proceeded through many small but significant steps and occasional large epiphanies.

This talk outlines how, over the past 70 years, financial models have quantified the notion of derivatives, diffusion, risk, volatility, the riskless rate, diversification, hedging, replication, and the principle of no riskless arbitrage.

# Modern Finance in One Sentence

Feynman summarizes physics in one sentence:  
Everything is made out of atoms.

---

IF

you can **hedge** away all correlated risk

AND

you can then **diversify** over all uncorrelated risk

THEN

you should expect to earn only **the riskless rate**

This sentence leads to CAPM, APT, Black-Scholes, ...

# Derivatives as a Method of Understanding

- Euclid axiomatized geometry, starting with **point, lines, planes** and then proving theorems.
- Spinoza tried to axiomatize human emotions by starting with the primitive visceral affects:  
**Desire, Pleasure, Pain.**
- Of Human Bondage — **Somerset Maugham**
- Good = everything that brings pleasure.
- *Evil* = everything that brings pain.
- *Love* = **Pleasure** associated with an external object. (Equity)
- *Hate* = **Pain** associated with an external object.
- *Envy* = **Pain** at another's **Pleasure**. (Equity and Debt)
- **Schadenfreude**
- Cruelty: **Desire** to inflict **Pain** on a someone you **Love**. (Equity, Debt, Credit)
- Three more primitives: Vacillation, Wonder, Contempt. But no real motion, it's static.

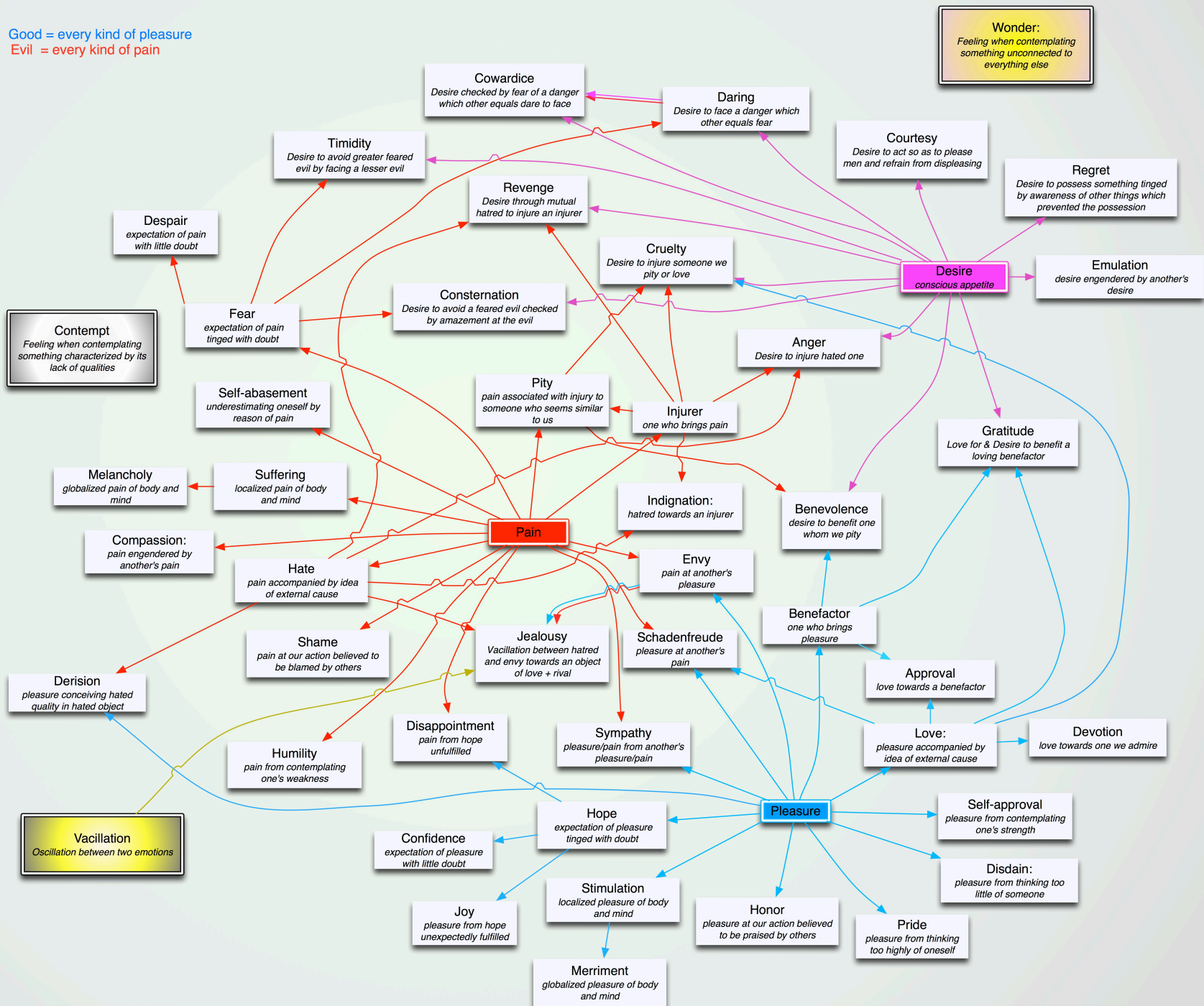


# Pleasure Pain Desire: A Map of the Emotions

Emanuel Derman

\*

Good = every kind of pleasure  
Evil = every kind of pain



# Derivatives *sans* Diffusion

- Spinoza believed that **human behavior follows laws**, that nothing is random. (Behavioral finance)
- But his scheme and definitions are static: there is almost no possibility of motion, except for

**Vacillation:** the cyclic alternation between two different passions.

- *Jealousy is the oscillation between **Hatred** and **Envy** in relation to a Love object and a rival*



*Hatred is **Pain** associated  
with an external Person*

*Envy is **Pain** at another's **Pleasure**.*

- Vacillation involves **volatility** — the more rapidly and intensely one Vacillates, the greater the Jealousy.
- Spinoza has no Anxiety in his system. Various opinions:
  - Anxiety is a vacillation between Hope and Fear.
  - Anxiety is not a Passion.
  - There was no Anxiety in the 17<sup>th</sup> Century.

# Diffusion *sans* Derivatives

- 1831: [Thomas Graham](#)

“...gases... when brought into contact, do not arrange themselves according to their density, ... but they spontaneously diffuse, mutually and equally, through each other, and so remain in the intimate state of mixture for any length of time.”

- 1858 - 1872: [James Clerk Maxwell](#) and [Ludwig Boltzmann](#)

explained the behavior of gases from the statistical behavior of the  $10^{23}$  hypothetical and invisible atoms inside

- Early 20th Century: [Albert Einstein](#), [Marian Smoluchowski](#) and [Jean-Baptiste Perrin](#)

[Confirm](#) the existence of atoms from Brownian motions of pollen, and deduces properties of atoms consistent with chemistry and physics.

- Physicists understood diffusion for underliers/atoms, but not functions of underliers. The exception: [Bachelier](#) in 1900 analyzed the behavior of options that are dependent on stocks that diffuse according to arithmetic Brownian motion. (Ahead of his time, rediscovered in the 1960s)

# Defining Risk as Volatility

- Investors were traditionally interested only in how much return they might earn.
- But return is uncertain.
- 1952: Harry Markowitz emphasized the statistics of returns, and emphasized looking at portfolios rather than individual stocks.
- Look at the relation between risk  $\sigma$  and return  $\mu$ .
  - Risk = the standard deviation of returns. *Volatility*  $\sigma$
  - Suggests finding the portfolio with the most return for a given risk.

# The Question: What is the Relation Between Risk $\sigma$ and Return $\mu$ ?

- Given that returns on stocks are uncertain, what is the appropriate relation between the risk we expose ourselves to and the return we expect?
- **The key question of finance:**  
*What  $\mu$  should one expect to earn, on average, for taking on a particular future risk  $\sigma$ ?*
- To answer this, note that there is one security whose returns has no uncertainty at all:  
the riskless bond, whose return is **guaranteed** to be  $r$  (a T Bill, say)
- This serves as a touchstone for measuring all other returns.
- We can denote every stock by the doublet  $(\mu, \sigma)$  denoting its expected return and its volatility.
- The riskless bond is  $(r, 0)$ .

# A Strategy for Answering the Question: Replicate a Riskless Bond

- 1958: Modigliani and Miller introduced **replication as a strategy for valuation**:
  - To value a security  $(\mu, \sigma)$ , reduce its risk to zero by combining it with other securities into a portfolio P **that has zero risk**.
  - Then P has the risk of a riskless bond.
  - By the *Law of One Price*, P must be guaranteed to earn the riskless rate  $r$ .
  - Imposing this on P leads to relation between  $\mu$  and  $\sigma$  for the security  $(\mu, \sigma)$ .
- This strategy will lead to CAPM, APT, Black-Scholes ...
- To do this one must know how to reduce risk.

# How Can One Reduce Risk?

- Dilution:  
Combine a security with a riskless bond
- Diversification:  
Combine a security with many other uncorrelated securities
- Hedging:  
Combine a security with a correlated security
- Apply this in three successively more realistic toy worlds.

## Simple World 1

A few uncorrelated stocks and a riskless bond:  
All stocks have same Sharpe Ratio

- Dilution: Combine weight  $w$  of a risky stock  $S (\mu, \sigma)$  with a weight  $(1 - w)$  of a riskless bond  $B (r, 0)$  to create a new security with lower risk & return

$$[w\mu + (1 - w)r, w\sigma] = [r + w(\mu - r), w\sigma]$$

- Law of One Price:  
All uncorrelated stocks with risk  $w\sigma$  earn excess return  $w(\mu - r)$
- One parameter fixes everything
- Same Sharpe ratio for all stocks!
- More risk, more return

$$\frac{\mu - r}{\sigma} = \lambda$$



Less Simple World 2:  
Many Uncorrelated Stocks: Diversify!

The Sharpe Ratio is zero  
Every stock is expected to earn the riskless rate  $r$

- Suppose there are countless **uncorrelated** stocks  $(\mu_i, \sigma_i)$
- Put them all in a portfolio with weights:  $P = \sum w_i S_i$
- Then the portfolio risk  $\sigma$  diversifies to zero.

$$\mu - r = \lambda \sigma = 0$$

- Thus the portfolio is riskless:  $\mu = r$
- But the portfolio return is the sum of individual returns:

$$\mu = \sum w_i \mu_i = \sum w_i (r + \lambda \sigma_i) = r + \lambda \sum \sigma_i \quad \text{therefore} \quad \lambda = 0$$

- Thus every stock is expected to earn the riskless rate!

## More Realistic World 3: CAPM

All Stocks are Correlated with the Market:

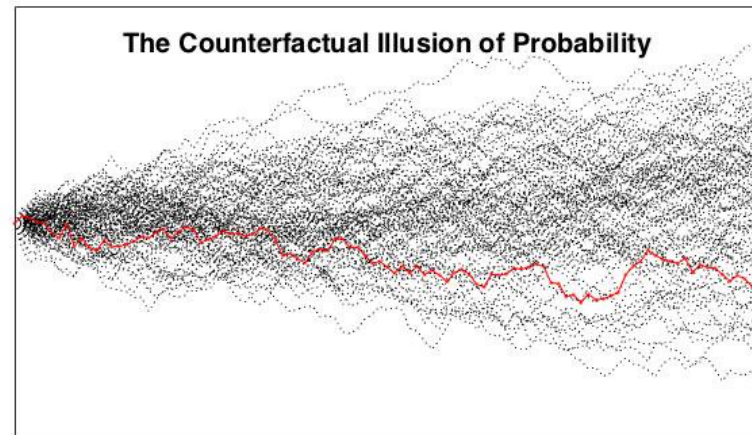
Hedge the Market Risk, then Diversify!

Every market-neutral stock must earn the riskless rate.

- Suppose there are countless stocks  $S_i$  correlated with the market  $M$
- Then the **market-neutral stock**  $S_i^M = S_i - \beta_i \left( \frac{S_i}{M} \right) M$  is uncorrelated with M
- After diversification each market-neutral stock  $S_i^M$  earns the riskless rate.
- Which means  $(\mu_i - r) = \beta_i(\mu_M - r)$
- CAPM just says that if you hedge every stock with the market, and then diversify over all remaining risk, you should earn only the riskless rate.

# Why is CAPM Bad?

- Because Risk is Not Really the Standard Deviation of Returns.
- Because the market M and the stock S are not really stably correlated.
- Markets are not exactly like flipping coins. There isn't a well-defined *a priori* probability of a market crash. Probability is a bit of an illusion



# 1960s: Early Options Models: Diffusion and Volatility but No Replication

- Samuelson, Sprenkel, Ayres, Boness ... value call options actuarially, as the expected discounted payoff of the option under a lognormal distribution with an unknown future growth rate and a known volatility.
- But at what rate does the stock grow?
- And what discount rate to use?

# 1973: Putting everything together

## Black-Scholes-Merton

### Diffusion+Volatility+Hedging+Replication

- Create a portfolio long the call, short  $\Delta$  shares of stock:  $P = C - \Delta S$
- Calculate change in value  $dP = dC - \Delta dS$
- Use diffusion for the move in the underlying stock price  $dS$
- Use stochastic calculus to find the move in the derivative  $dC(S)$
- Choose  $\Delta$  to **eliminate stock risk** in  $dP = dC - \Delta dS$
- Require that hedged portfolio, which is riskless, earns the known riskless rate  $r$ :

$$dC - \Delta dS = r(C - \Delta S)$$

- Then we get the same formula for the price of the call  $C$  as the actuarial one, but where all growth and discount rates are replaced by the riskless rate  $r$ .

# Black-Scholes-Merton

- You can replicate/hedge an option with stock
- Option C and stock S must have same Sharpe ratio

$$\frac{\mu_S - r}{\sigma_S} = \frac{\mu_C - r}{\sigma_C}$$

- Ito's Lemma applied to a call C leads to the Black-Scholes pde:

$$\frac{\partial C}{\partial t} + rS \frac{\partial C}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC$$


- A unified treatment of BSM and CAPM from one principle

# Why is BSM better than CAPM?

- Because you **more realistically** can hedge an option with a stock, because the correlation is really close to 1.
- So even if you don't believe the risk is the standard deviation of returns, the two securities really are connected, unlike the statistical connection between two different stocks.
- **The Caveat: We have assumed that volatility is unchanging! If volatility is random, then the derivative is not really a derivative except at expiration**

# 1970s: Using the BS Equation

- Now, to value an option, instead of forecasting **the return** of the stock, traders must **forecast the volatility** of the stock.

$$\frac{\partial C}{\partial t} + rS \frac{\partial C}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC$$


- Black and Scholes set about using the equation by using historical volatilities to estimate future volatilities. But who knows what future volatility will be?



# 1976: Calibration

## The Invention of Implied Volatility

- Latane and Rendelman suggested fitting option market prices to the Black-Scholes formula and extracting the **implied future volatility** of the stock that fits the option market price. They then suggest calculating hedge ratios from the model using the implied volatility.
- Implied volatility is a parameter, not a statistic.
- **But implied volatilities are unstable.**  
This process must be repeated as the implied volatility keeps changing, so there is *something not quite right*.
- Implied volatilities tell you that, given an option price, if you believe the model, this is what the future must be like. But the future doesn't turn out that way.
- **Nevertheless, from now on everyone calibrates models.**
- Most people don't even realize it was an invention.

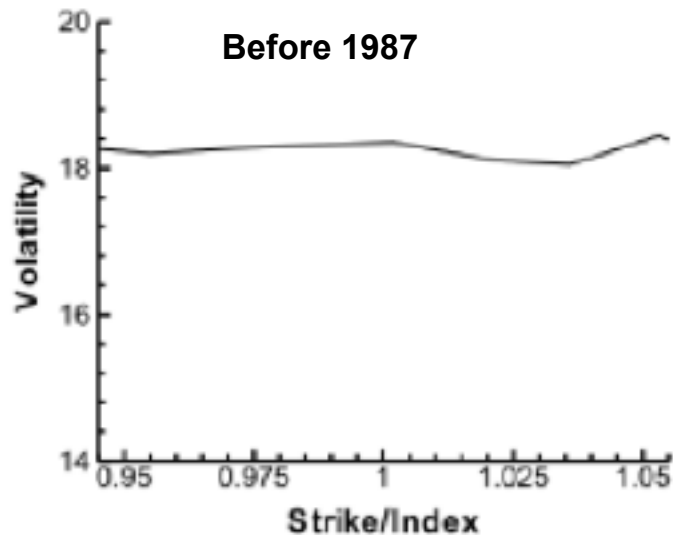
# Because the Model Doesn't Work, Trading Volatility is Now a Possibility

- In fact:
  - Volatility is stochastic.
  - So you can't **really** replicate an option
  - Instead, you can speculate on the **volatility parameter**, using options.
- If you take the model seriously, hedged options now become a way of trading the volatility parameter rather than speculating on the stock price.
- **Volatility as an asset class.**

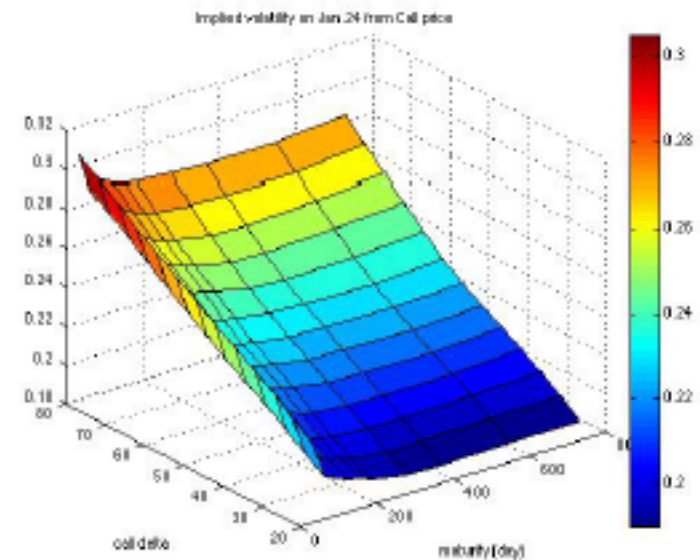
# 1977: Yield Curve Modeling: Parameters

- Modeling the yield curve — an extension of Black-Scholes to bonds rather than stocks, to rates rather than securities. Initiated by Vašíček.
- Focus on **evolution of parameters** (interest rates) rather than securities.
- The difficulty is avoiding future arbitrages in the model when there is more than one security, as with the yield curve.
- You can value options on two stocks independently, but you cannot value options on two different-maturity bonds independently:
  - There are no-arbitrage constraints on bond prices — no negative forward rates.
- And so on ... to other extensions (volatility, default, inflation ...) of the hedging paradigm for 47 years ...

# 1987: The Smile



Jan 24 2008



- When you fit the BS model to different option strikes, each one implies a different future volatility for the underlying stock.
- *Now something is really wrong* — the BS model cannot accommodate different volatilities for the same stock.
- Nevertheless, people keep using the model inconsistently to estimate hedge ratios as they calibrate the model to a particular option price.

# 1994 - present: The Smile

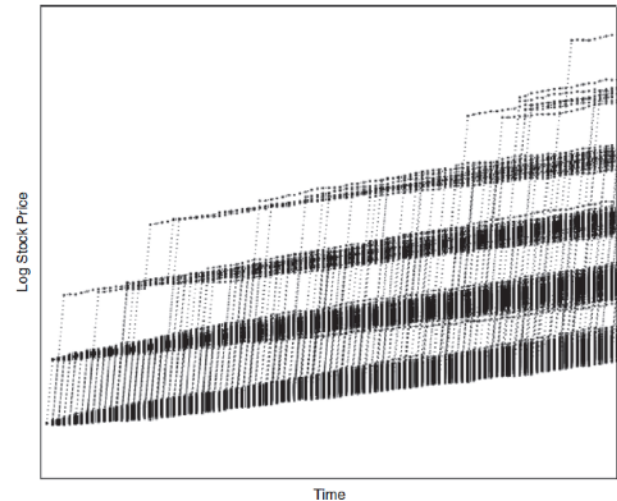
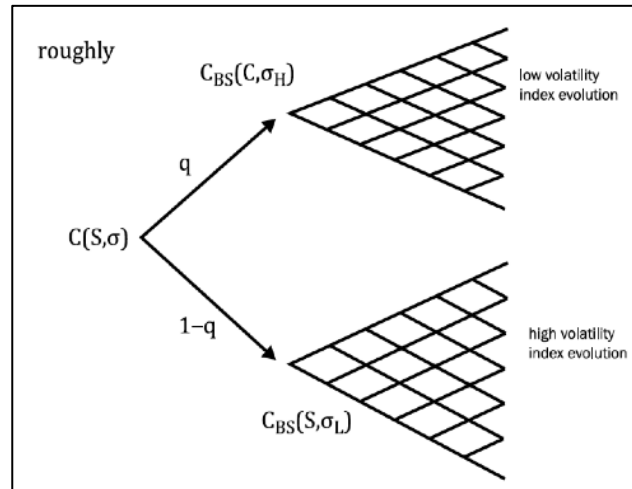
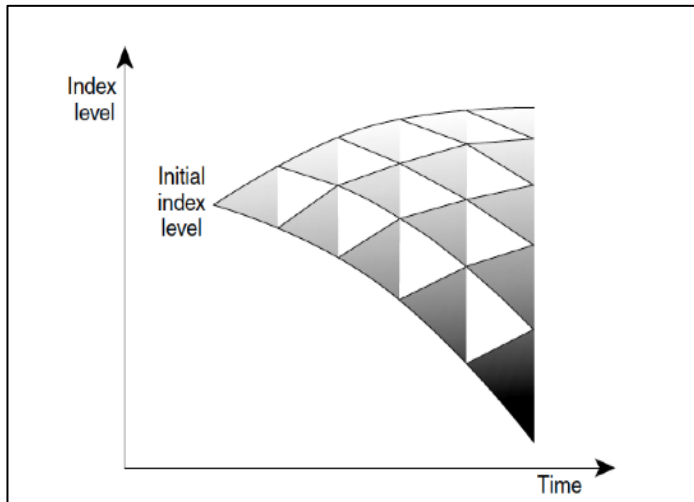


FIGURE 24.3 A Monte Carlo Simulation of the Log Stock Prices in the Jump-Diffusion Model

- BS Extensions: local volatility, stochastic volatility, jumps plus diffusion ...
- More complexity without accurate knowledge of the parameters.
- Use calibration-to-market-prices to imply the parameters — e.g. the volatility of volatility in a calibrated stochastic volatility model.
- But markets change and these implied parameters are themselves unstable and random, and there are now more of them. **So, for example, volatility of volatility is stochastic.**
- **So now, using a “better” model you have a market for trading volatility of volatility**

# Where We Are

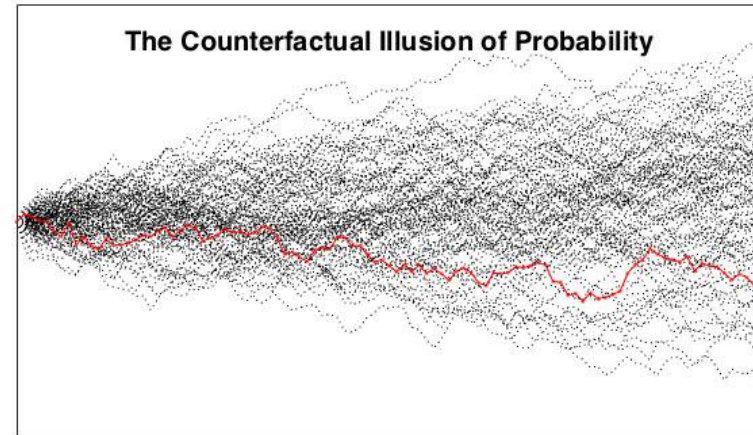
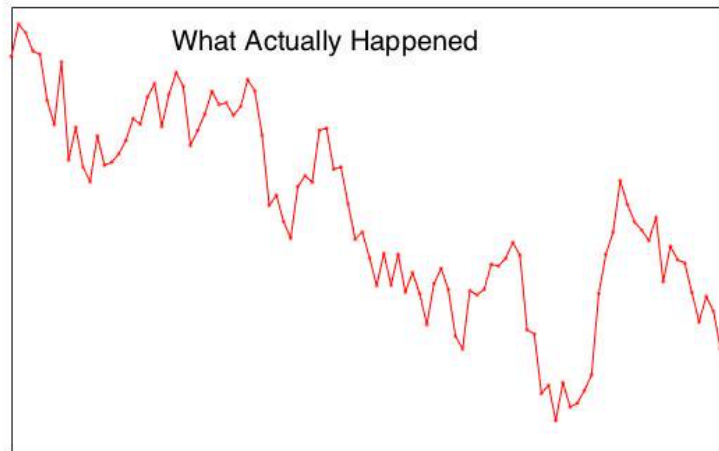
## 1 ...

- The BS Model is genuinely useful yet always inadequate. Derivatives are not truly derivative, except at expiration.
- The analogy with diffusion is inaccurate, but the framework has reified the notion of risk as equivalent to volatility.
- Replication doesn't quite work. Each new model is inadequate, **introduces new parameters**, which, as the model is embraced, become quantities the market can speculate and trade on.
- When a widely embraced reification is proved to be inaccurate by the behavior of the market, systemic market collapses can occur.

# ... Where We Are

## 2

- The probabilistic approach to distributions is a fallacy. They are **implied probabilities** that change each time someone offers a price.
- Probabilistic models ... are only internal episodes attributing a hypothetical probability that we require in order to imperfectly replicating or hedge something (Elie Ayache). Because of this models always have to be recalibrated when new prices are offered.



# One More Remark: The Dangerous Focus on Parameters

- One used to eat food, now one eats nutrients, and thinks of food items as baskets of nutrients.
- One used to buy stocks and portfolios of stocks. Now one thinks of stocks as combinations of factors: momentum, volatility, popularity ...
- One used to trade securities; now one trades model parameters.
  - Rates instead of bond prices.
  - Volatility instead of options prices.
  - Credit, VIX, momentum, smart beta, low vol, popularity ...
- Models turn prices into parameters. This is how models work.
- The danger is that this reification makes it easy for unskilled crowds to trade subtle things that previously took skill.
  - CDS made it too easy to trade credit?
  - VIX derivatives made it easy to trade vol?



# A Plea for Less Formalism

- Gian Carla Rota on Mark Kaç:

“Throughout his life he remained skeptical of abstraction, of techniques, of axiomatics. Instead he inspired the first generation of scientists who learned to think probabilistically. He warned them that axioms will change with the whims of time, but an application is forever.”

- Paul Dirac: “I am not interested in proofs, I am interested only in what nature does.”
- Barenblatt in a book on “Scaling”:

Of special importance is the following fact: the construction of models, like any genuine art, cannot be taught by reading books and/or journal articles (I assume that there could be exceptions, but they are not known to me). The reason is that in articles and especially in books the ‘scaffolding’ is removed, and the presentation of results is shown not in the way that they were actually obtained but in a different, perhaps more elegant way. Therefore it is very difficult, if not impossible, to understand the real ‘strings’ of the work: how the author really came to certain results and how to learn to obtain results on your own.