

# MODELING RISK FACTORS



# INTRODUCTION

- ▶ Risk → Volatility → Measure of dispersion
- ▶ Most used distributions: Normal and Log-Normal
- ▶ Stylized facts → empirical distribution is not normal

# NORMAL DISTRIBUTION

- ▶ It is easy
- ▶ Stable under addition: sum of normal random variables is normal
- ▶ Limiting distribution of the average of independent random variables [Central limit theorem (CTL)]
- ▶ Good first order approximation for many random variables: returns on stocks, exchange rate returns, bond returns, etc.



probability of small movements much higher than probability of large movements

# NORMAL AND LOG-NORMAL DISTRIBUTIONS: RETURNS

*Normal*

$$r \sim \Phi(\mu, \sigma)$$

$$P_1 = P_0(1 + r)$$

$$P_1 \sim P_0 + \Phi(P_0\mu, P_0\sigma)$$

$$P_1 \sim \$100 + \Phi(0, 15)$$

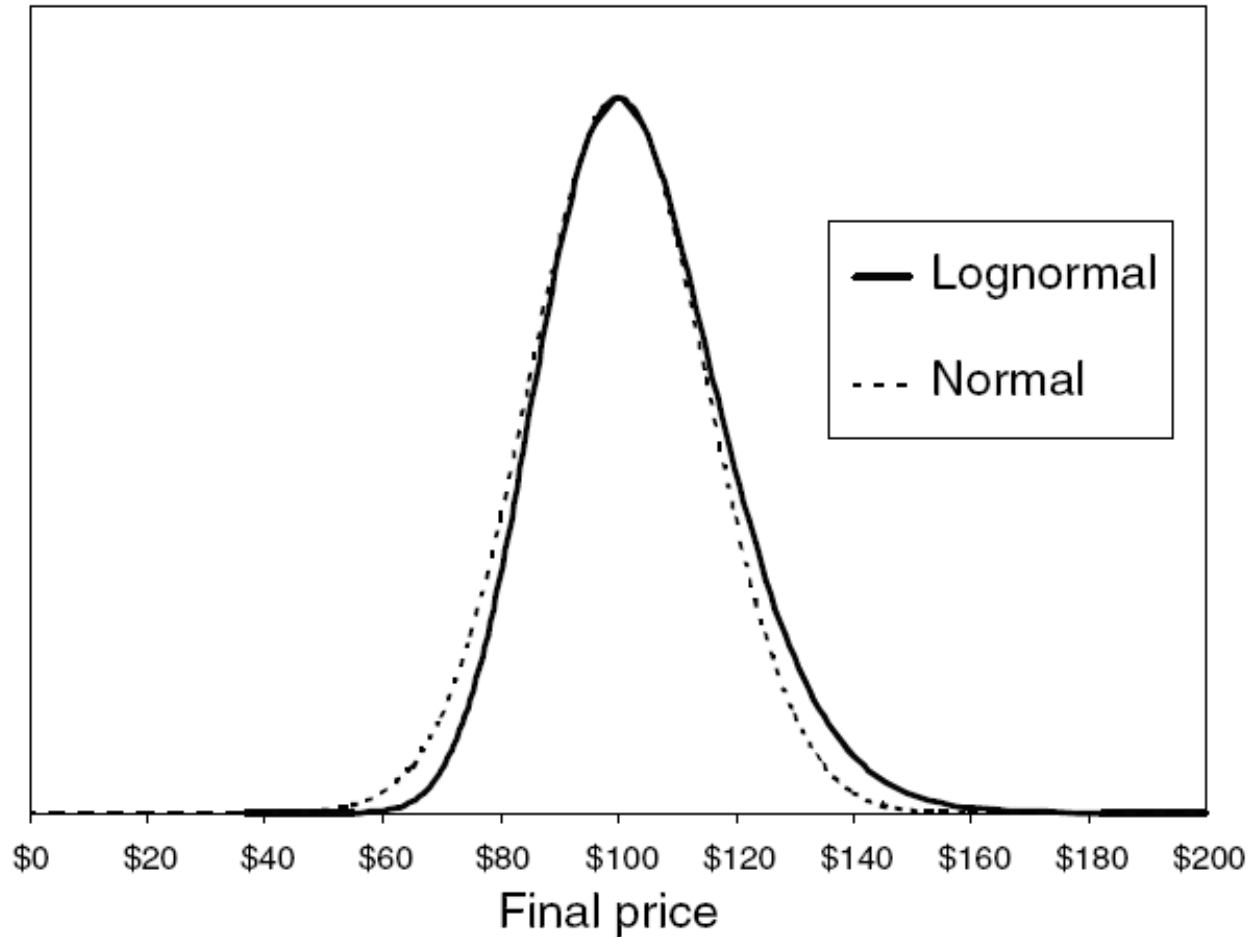
*Log - Normal*

$$R = \ln(P_1 / P_0) \sim \Phi(\mu, \sigma)$$

$$P_1 = P_0 \exp(R)$$

**FIGURE 14-1 Normal and Lognormal Distributions: Annual Horizon**

Probability density function



# NORMAL AND LOG-NORMAL DISTRIBUTIONS

The difference between the two distributions is driven by the size of the volatility parameter → Small volatility values imply that the distributions are virtually identical

This can happen either when the asset is not very risky → annual volatility is small, or when the horizon ( $T$ ) is very short (i.e. daily)

## **Key concept:**

The normal and lognormal distributions are very similar for short horizons or low volatilities

# LOG-NORMAL ASSUMPTION

- ▶ If log-normal assumption and i.i.d. (you should know that:)

$$V[R(0,T)] = V[R(0,1)] + V[R(1,2)] + \dots + V[R(T-1,T)] = V[R(0,1)]T$$

$$E[R(0,T)] = E[R(0,1)] + E[R(1,2)] + \dots + E[R(T-1,T)] = E[R(0,1)]T$$

$$\rightarrow \sigma_T = \sigma T^{0.5} \quad \rightarrow \text{VaR} = \alpha(\sigma T^{0.5})W$$

This holds only if:

- ▶ The distribution is the same at each period
- ▶ Returns are uncorrelated/independent across periods, so that all covariance terms disappear
- ▶ The distribution is the same for one- and  $T$ -period returns, or is stable under addition, such as the normal

# IF RETURNS ARE NOT I.I.D.

- ▶ If returns follow a first-order autoregressive process:

$$R_t = \rho R_{t-1} + u_t$$



$$V[R_t + R_{t-1}] = V[R_t] + V[R_{t-1}] + 2\text{Cov}[R_t, R_{t-1}] = \sigma^2 + \sigma^2 + 2\rho\sigma^2 = 2\sigma^2(1 + \rho)$$

- ▶ If  $\rho > 0$  → trend → longer term volatility increases faster than with the usual square-root-of-time rule
- ▶ If  $\rho < 0$  → mean reversion → longer term volatility increases more slowly than with the usual square-root-of-time rule

# FAT TAILS

- ▶ **Key rule of thumb:**

Every financial market experiences one or more daily price moves of 4 standard deviations or more each year. And in any year, there is usually at least one market that has a daily move greater than 10 standard deviations

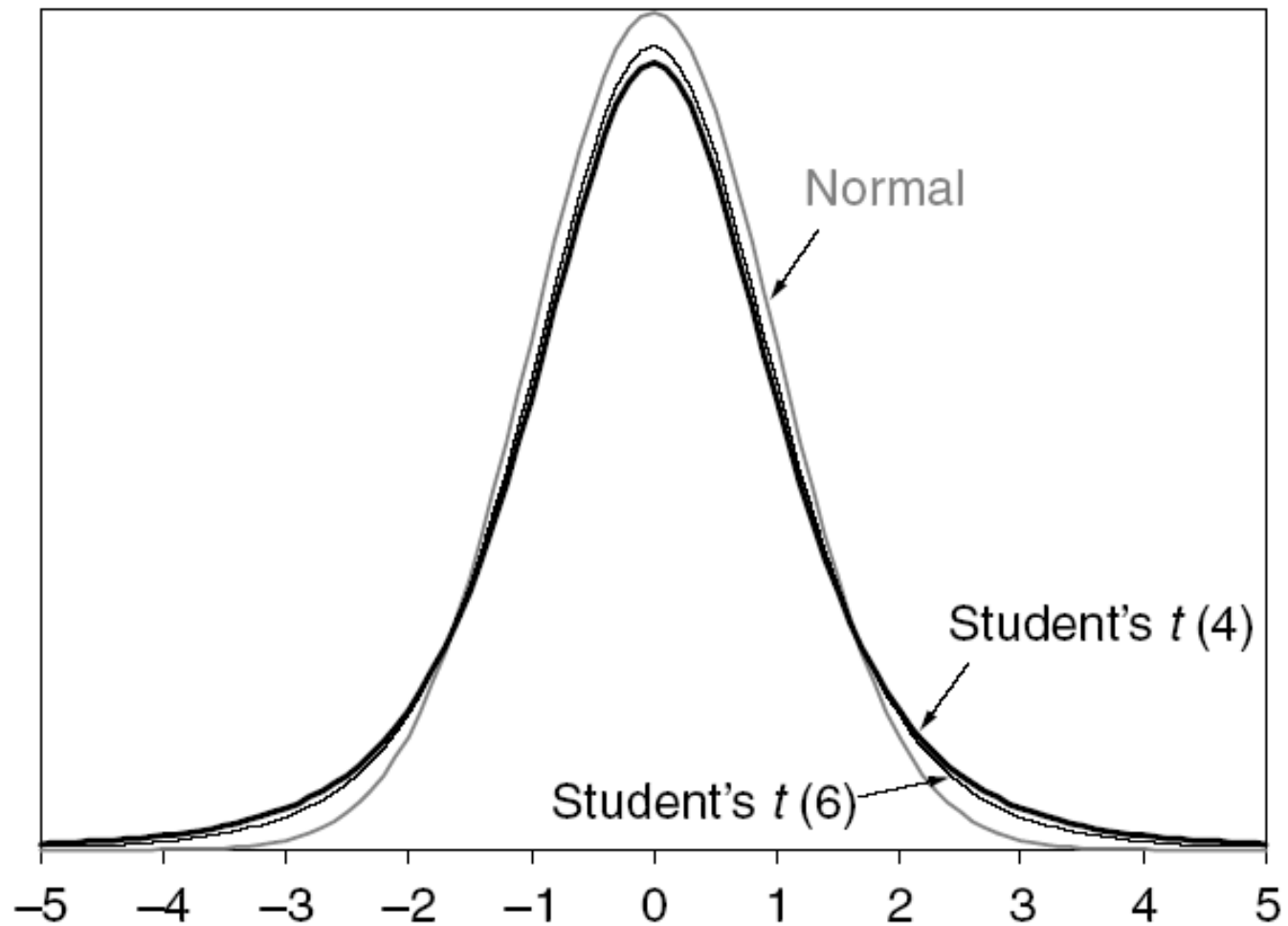
Normal distribution does not work!

This empirical observation can be explained in a number of ways:

- ▶ The true distribution has fatter tails (e.g., Student's  $t$ )
- ▶ The observations are drawn from a mix of distributions (e.g., a mix of two normals, one with low risk and the other with high risk)
- ▶ The distribution is nonstationary

**FIGURE 14-3 Normal and Student Distributions**

Probability density function



**TABLE 14-2 Comparison of the Normal and Student's *t* Distributions**

Deviate	Tail Probability			Expected Number in 250 Days		
	Normal	<i>t</i> df = 6	<i>t</i> df = 4	Normal	<i>t</i> df = 6	<i>t</i> df = 4
-5	0.00000	0.00123	0.00375	0.00	0.31	0.94
-4	0.00003	0.00356	0.00807	0.01	0.89	2.02
-3	0.00135	0.01200	0.01997	0.34	3.00	4.99
-2	0.02275	0.04621	0.05806	5.69	11.55	14.51
-1	0.15866	0.17796	0.18695	39.66	44.49	46.74
				Deviate (alpha)		
Probability = 1%				2.33	3.14	3.75
Ratio to normal				1.00	1.35	1.61

# MEASURING VOLATILITY

- ▶ Indirect Measures:

  - Generalized AutoRegressive Conditional Heteroskedasticity (GARCH),  
Exponential Weighted Moving-Average (EWMA: RiskMetrics)

  - Stochastic Volatility models

- ▶ Direct Measures

  - Squared Returns

  - Range

  - Realized Volatility

# THE CAUSES OF VOLATILITY

It is possible to distinguish three main causes of financial market volatility:

- ▶ Informational considerations - the arrival of new information in the market results in price adjustments as market agents evaluate the implications of the new information [French (1980)].
- ▶ Hedging or speculative pressure - an increase in activity by one class of agents, say speculators, forces the price to adjust in such a way that the other group, in this case hedgers, can accommodate the increased speculative position. [Volatility is caused by trading.] [French and Roll (1986)]
- ▶ (Only for commodities) The physical availability of the commodity - when a commodity is in short supply, a change in demand will have a larger price impact than when supplies are plentiful [Gilbert (1989)].

# MODELLING THE MEAN (I): AR(P) & MA(Q) MODELS

## ▶ AutoRegressive models - AR

▶ AR(1)  $y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t$

▶ AR(p)  $y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$

## ▶ Moving Average models - MA

▶ MA(1)  $y_t = \mu + \theta_1 \varepsilon_{t-1} + \varepsilon_t$

▶ MA(q)  $y_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$

# MODELLING THE MEAN (2): ARMA(P,Q)

- ▶ ARMA(p,q) Model - Box & Jenkins (1976)

$$\phi(L)(y_t - \mu) = \theta(L)\varepsilon_t$$

where  $\mu$  is the mean of  $y_t$ ,  $y_t$  is  $I(0)$ ,  $\varepsilon_t$  is white noise and

$$\phi(L) = 1 - \sum_{j=1}^p \phi_j L^j, \quad \theta(L) = 1 + \sum_{j=1}^q \theta_j L^j$$

- ▶ ARMA(1,1)

$$y_t = \mu + \phi y_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t$$

# MODELLING THE MEAN (2): ARIMA(P,D,Q)

- ▶ If  $y_t$  is integrated of order  $d \implies$  ARIMA( $p,d,q$ ) Model

$$\phi(L)(1-L)^d(y_t - \mu) = \theta(L)\varepsilon_t$$

- ▶ ARIMA(1,1,1)

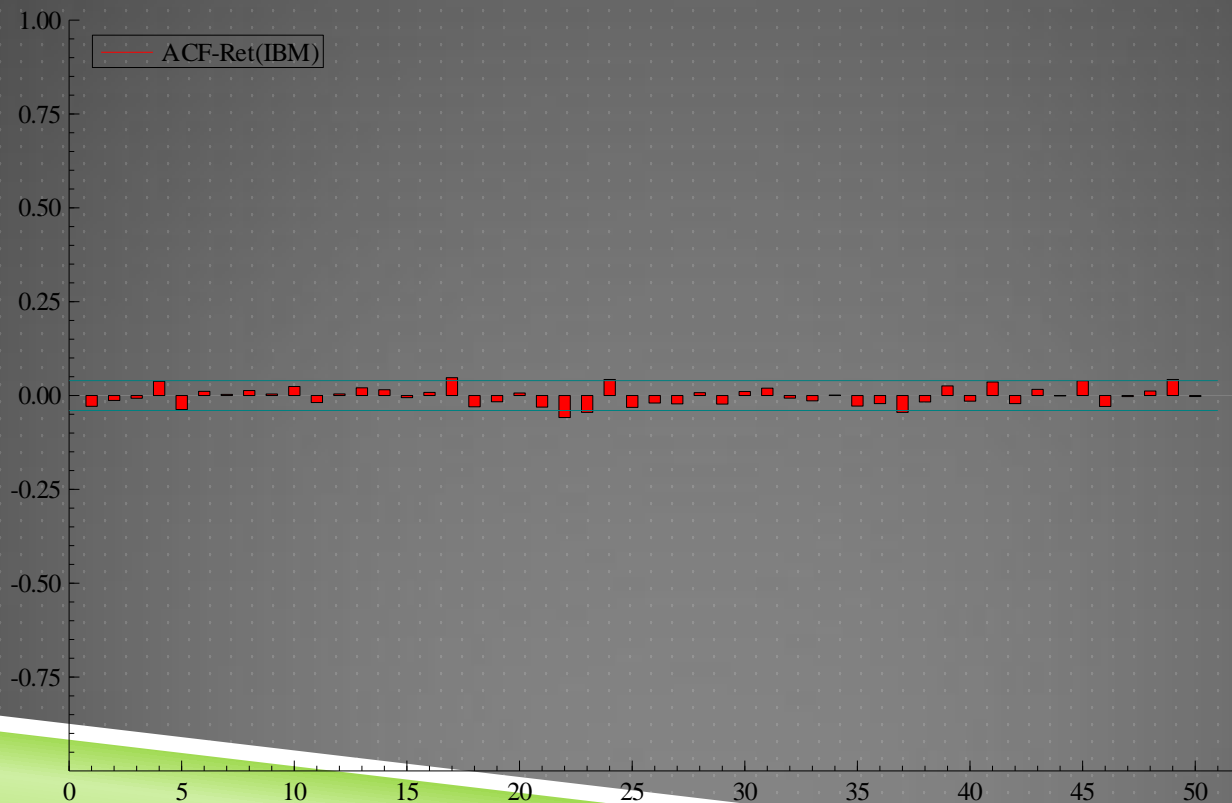
$$\Delta y_t = y_t - y_{t-1} = \mu + \phi y_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t$$

- ▶ This is what we do with returns!
  1. Take natural logarithm of price:  $P_t = \ln(p_t)$
  2. Compute first difference:  $\Delta P_t = P_t - P_{t-1} = \ln(p_t / p_{t-1})$
  3. Model the return process

# MODELLING THE MEAN (3): EMPIRICAL APPLICATION TO IBM RETURNS

$$y_t = 0.0548 - 0.029y_{t-1} - 0.030\varepsilon_{t-1} + \varepsilon_t$$

(.0415)   (.0199)   (.0200)



# MODELLING THE VARIANCE: ARCH-GARCH MODELS (I)

- ▶ AutoRegressive Conditional Heteroskedasticity:ARCH(q) - Engle (1982)

$$r_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = h_t^{1/2} e_t$$

$$e_t \sim N(0,1)$$

$$h_t | \Omega_{t-1} = \alpha_0 + \alpha(L)\varepsilon_t^2$$

where  $\alpha(L)$  is a lag polynomial of order  $q$ ,  $\Omega_{t-1}$  is the information set available at time  $t-1$  and  $r_t$  could be any regression function for the mean process

- ▶ The ARCH model enables us to distinguish between the conditional and the unconditional second order moment (variance). While the unconditional variance may be time invariant (homoskedasticity) or may not even exist, the conditional variance varies over time and depends on the past states of the world

## ARCH-GARCH MODELS (2)

- ▶ Bollerslev (1986) (see also Taylor, 1986) proposed the Generalised ARCH or GARCH(p,q) model

$$r_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = h_t^{1/2} e_t$$

$$e_t \sim N(0,1)$$

$$h_t | \Omega_{t-1} = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t$$

- ▶ The GARCH model allows to estimate both the conditional mean and the conditional variance of a process
- ▶ In the ARCH-GARCH class of models, the variance of the current error term is an increasing function of the magnitude of the past error term ==> large (small) errors are followed by large (small) errors. Hence these models are able to capture the volatility clustering phenomenon observed in financial data
- ▶ Engle (1982) shows that the kurtosis of an ARCH(1) process is greater than 3 (the kurtosis of the normal distribution) ==> ARCH-GARCH models are able to capture the leptokurtosis behaviour of financial time series

# GARCH(1,1)

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}$$

$$E[\varepsilon_{t-1}^2] = h_t, \quad \text{Unconditional Variance :} \quad h = \frac{\alpha_0}{1 - \alpha_1 - \beta}$$

$$\text{Stationarity :} \quad \alpha_1 + \beta < 1, \quad \text{Persistency :} \quad \alpha_1 + \beta$$

*Forecasting*

$$E[h_{t+1}] = \alpha_0 + \alpha_1 E[\varepsilon_t^2] + \beta E[h_t] = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta h_t$$

$$E[h_{t+2}] = \alpha_0 + \alpha_1 E[\varepsilon_{t+1}^2] + \beta E[h_{t+1}] = \alpha_0 + (\alpha_1 + \beta) E[h_{t+1}]$$

...

$$E[h_{t+n}] = \alpha_0 + \alpha_1 E[\varepsilon_{t+n-1}^2] + \beta E[h_{t+n-1}] = \alpha_0 + (\alpha_1 + \beta) E[h_{t+n-1}]$$

# ARCH-GARCH MODELS (4)

- ▶ Defining the innovations in the conditional variance process as

$$v_t = \varepsilon_t^2 - h_t$$

- ▶ The GARCH(p,q) model might be expressed as an **ARMA** process in  $v_t$ ,

$$(1 - \alpha_1 L - \beta L) \varepsilon_t^2 = \alpha_0 + (1 - \beta L) v_t$$

$$(1 - \phi L) \varepsilon_t^2 = \alpha_0 + [1 - \beta L] v_t$$

- ▶ Engle and Bollerslev (1986) considered the case of a unit root in the lag polynomial  $(L)$  and introduced the Integrated GARCH or IGARCH model

$$(1 - \phi L)(1 - L) \varepsilon_t^2 = \alpha_0 + [1 - \beta L] v_t$$

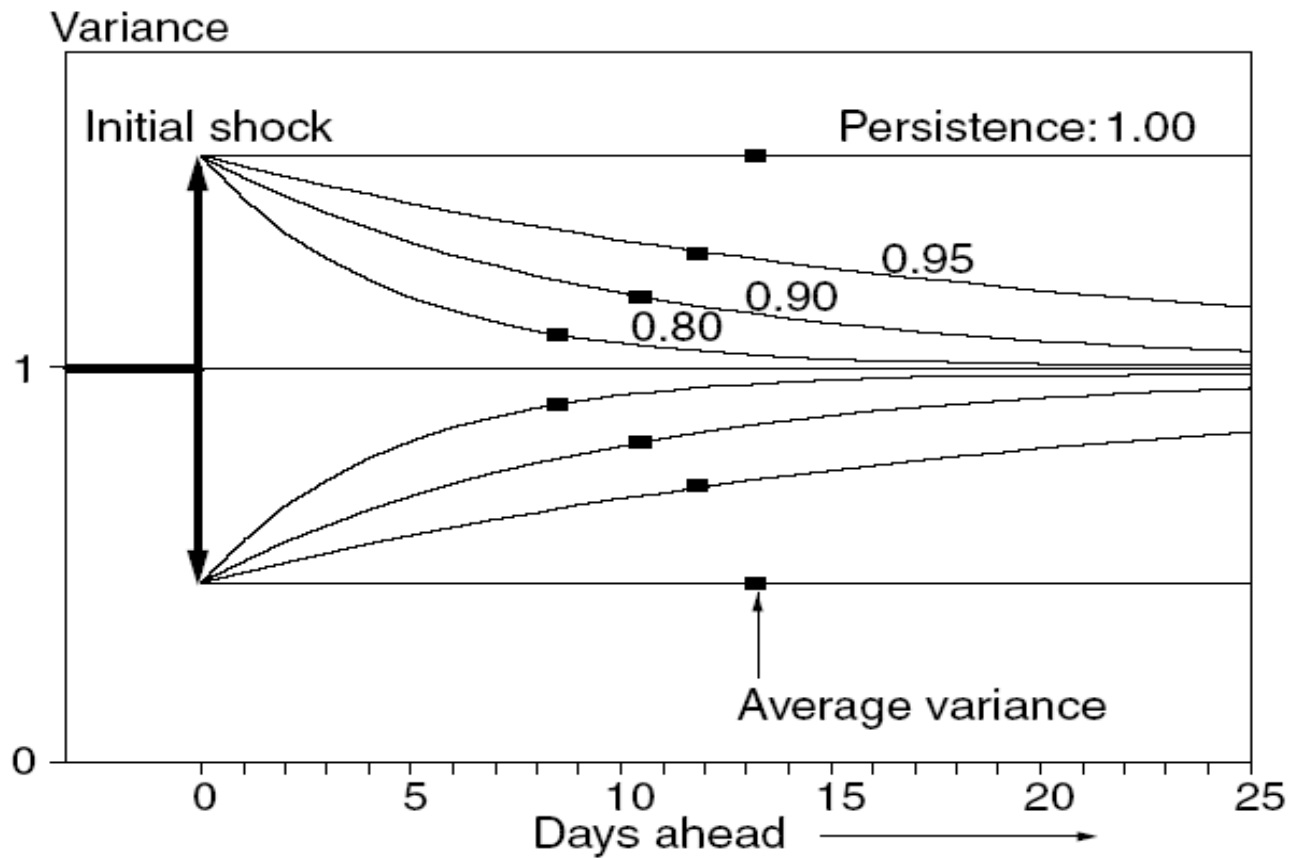
- ▶ This equation may be interpreted as the **ARIMA** representation of the IGARCH process

# GARCH(1,1): IBM EXAMPLE

	Parameter	Standard Error	t-value
$\mu$ (constant in the mean)	0.079345	0.038437	2.064
$\alpha_0$	0.037263	0.032199	1.157
$\alpha_1$	0.038015	0.012648	3.005
$\beta$	0.903020	0.040123	22.51

# PERSISTENCY

**FIGURE 14-4 Shocks to a GARCH Process**



# SQUARE-ROOT-OF-TIME RULE DOES NOT WORK

The square-root-of-time rule for extrapolating returns does not apply when risk is time-varying. If the initial value of the variance is greater than the long-run average, simply extrapolating the one-day variance to a longer horizon will overstate the average variance. Conversely, starting from a lower value and applying the square-root-of-time rule will understate risk

## **Key concept:**

Use of the square-root-of-time rule to scale one-day returns to longer horizons is generally inappropriate when risk is time-varying

# RISK METRICS

## Exponentially weighted moving-average (EWMA)

EWMA model places geometrically declining weights on past observations, assigning greater importance to recent observations

$$h_t = \lambda h_{t-1} + (1 - \lambda) \varepsilon_{t-1}^2$$

$$h_t = (1 - \lambda) [\varepsilon_{t-1}^2 + \lambda \varepsilon_{t-2}^2 + \lambda^2 \varepsilon_{t-3}^2 + \dots]$$

The weights thus decrease at a geometric rate. The lower  $\lambda$ , the more quickly older observations are forgotten

RiskMetrics has chosen  $\lambda = 0.94$  for daily data and  $\lambda = 0.97$  for monthly data

# RISK METRICS

This model is a special case of the GARCH process, where  $\alpha_0 = 0$ , and  $\alpha_0 + \beta = 1$

The model therefore has permanent persistence → Shocks to the volatility do not decay

Thus longer-term extrapolation from the GARCH and EWMA models may give quite different forecasts

Over a one-day horizon, however, the two models are quite similar and are often indistinguishable from each other

# RISK METRICS

Table 14-4 shows how to build the EWMA forecast using a parameter of  $\lambda = 0.95$ , which is consistent with the previous GARCH example. At time 0, we start with variance  $h_0 = 1.1$ , as before. The next day we have a return of 3%. The new variance forecast is then  $h_1 = 0.05 \times 3^2 + 0.95 \times 1.1 = 1.50$ . The next day this moves to  $h_2 = 0.05 \times 0^2 + 0.95 \times 1.50 = 1.42$ , and so on.

**TABLE 14-4 Building an EWMA Forecast**

Time	Return	Conditional Variance	Conditional Risk	Conditional 95% Limit
$t - 1$	$r_{t-1}$	$h_t$	$\sqrt{h_t}$	$2\sqrt{h_t}$
0	0.0	1.10	1.05	$\pm 2.1$
1	3.0	1.50	1.22	$\pm 2.4$
2	0.0	1.42	1.19	$\pm 2.4$
3	0.0	1.35	1.16	$\pm 2.3$

# STOCHASTIC VOLATILITY (I)

- ▶ Another response to non-normality of returns is to assume that there is a random variable (stochastic volatility) *conditional upon which* returns are normal. This variable (stochastic volatility) is not directly observed
- ▶ This kind of assumption is often made in continuous-time theoretical models where volatility is driven by a process separate from returns: there are two different processes, one for asset returns and one for volatility
- ▶ Example of stochastic volatility model

$$y_t = \varepsilon_t e^{\sigma_t/2}$$
$$\sigma_t = \varphi\sigma_{t-1} + \xi_t$$

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$
$$\xi_t \sim N(0, \sigma_\xi^2)$$

# STOCHASTIC VOLATILITY (2)

- ▶ Usually it is assumed that the two processes,  $\varepsilon_t$  and  $\xi_t$ , are serially uncorrelated and independent of each other
- ▶ If we square the return equation,  $y_t$ , and take logs, we get

$$\begin{aligned}\log(y_t^2) &= \sigma_t + \log(\varepsilon_t^2) \\ \sigma_t &= \varphi\sigma_{t-1} + \xi_t\end{aligned}$$

- ▶  $\sigma_t$  measures the difference between the conditional log standard deviation of returns and its mean;  $\sigma_t$  follows a zero-mean AR(1) process

# DIRECT MEASURES OF VOLATILITY

- ▶ Squared returns
- ▶ Absolute returns:
  - Luce (1980) proved that absolute returns belong to a class of risk measure
- ▶ Taylor (1987) first found that absolute returns of speculative assets have significant serial correlation over long lags. He also found that autocorrelations of absolute returns are greater than that of squared returns
- ▶ Granger and Ding (1995) proved that absolute return and any power transformation of this return, may be interpret as a measure of risk
- ▶ Empirical evidence shows that the power transformation of the absolute return  $|r_i|^a$ , has high autocorrelation for long lags and that this property is strongest when  $a = 1$  [Ding, Granger and Engle (1993), Granger and Ding (1995), Ding and Granger (1996)]

# DIRECT MEASURES OF VOLATILITY: RANGE

Underlying assumptions:  $dP = \sigma dW$

$$P = \ln(p)$$

Range:  $I = \text{Max}(P) - \text{Low}(P) \rightarrow I = \ln(\text{high}) - \ln(\text{low})$

Comparing two estimators:

$$\sigma_{t,Range}^2 = \delta I_t^2 \qquad \sigma_{t,Rt}^2 = \varphi r_t^2$$

$$\delta = \frac{1}{4 \ln(2)}, \quad \varphi = 1, \quad r_t = \ln(\text{Close}) - \ln(\text{Open})$$

$$\sigma_{t,Range} = \delta I_t \qquad \sigma_{t,Rt} = \alpha |r_t|$$

$$\delta = \sqrt{\frac{\pi}{8}}, \quad \varphi = \sqrt{\frac{\pi}{2}}$$

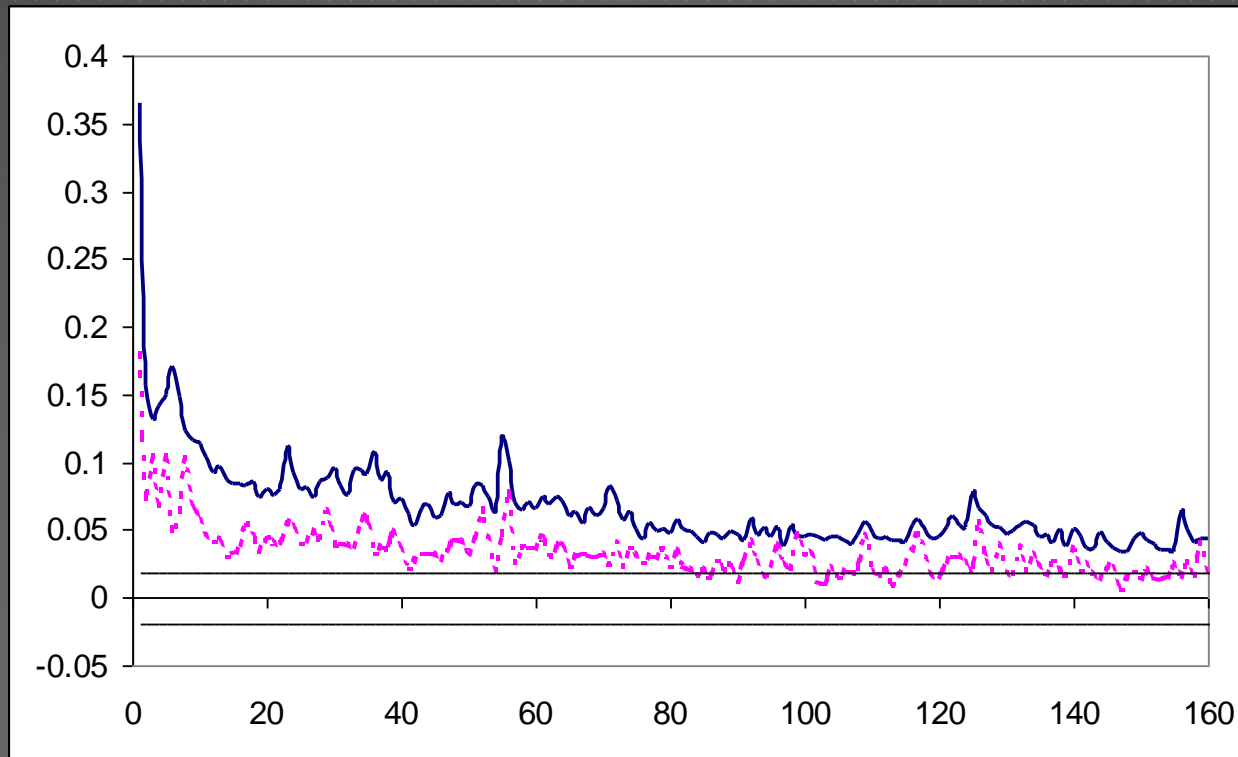
Parkinson(1980): range-based variance estimates of daily volatility are 5 times more efficient than return-based variance estimate

Brunetti Lildholdt(2005): range-based standard deviation estimates of daily volatility are 6 times more efficient than return-based estimators

# AUTOCORRELATION FUNCTION ALCOA INC (AA)

SOLID LINE: RANGE-BASED VARIANCE ESTIMATES

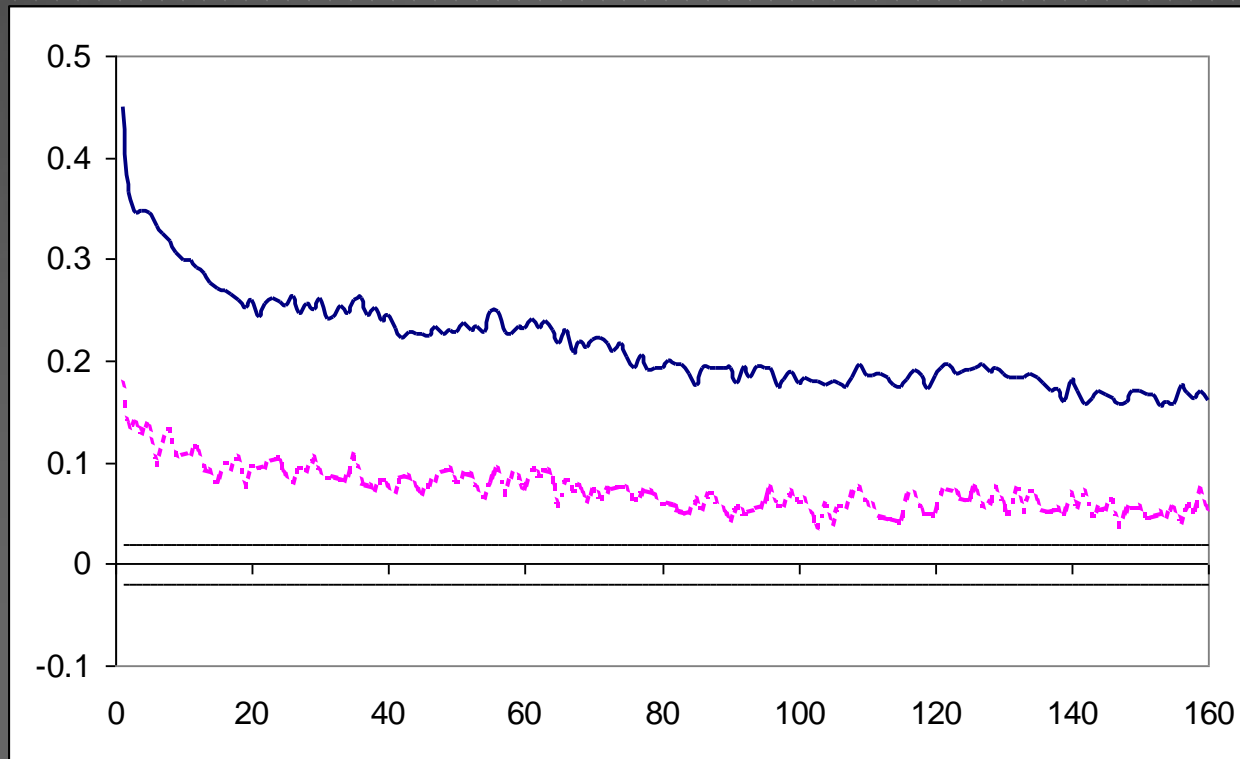
BROKEN LINE: RETURN-BASED VARIANCE ESTIMATES



# AUTOCORRELATION FUNCTION ALCOA INC (AA)

SOLID LINE: RANGE-BASED STANDARD DEVIATION ESTIMATES

BROKEN LINE: RETURN-BASED STANDARD DEVIATION ESTIMATES



# REALIZED VOLATILITY

## ► Integrated Volatility

$$\sigma_{t,t+\delta t}^2 = \int_t^{t+\delta t} \sigma^2(s) ds$$

$$\sigma_{RV,t}^2 = \sum_{i=1}^N r_i^2$$

$$r_i = \ln(P_t / P_{t-1})$$

# REALIZED VOLATILITY

To compute realized volatility we need to use high-frequency data:  
information about transaction within a day

TAQ (Trades and Quotes) from NYSE

Basic idea (very simple):

Construct 5 minutes time intervals

Compute returns

Square returns

Sum the 5-minute squared returns over the day

# REALIZED VOLATILITY: PROBLEMS

Realized volatility → market microstructure noise

Realized volatility does not take into account market microstructure noise

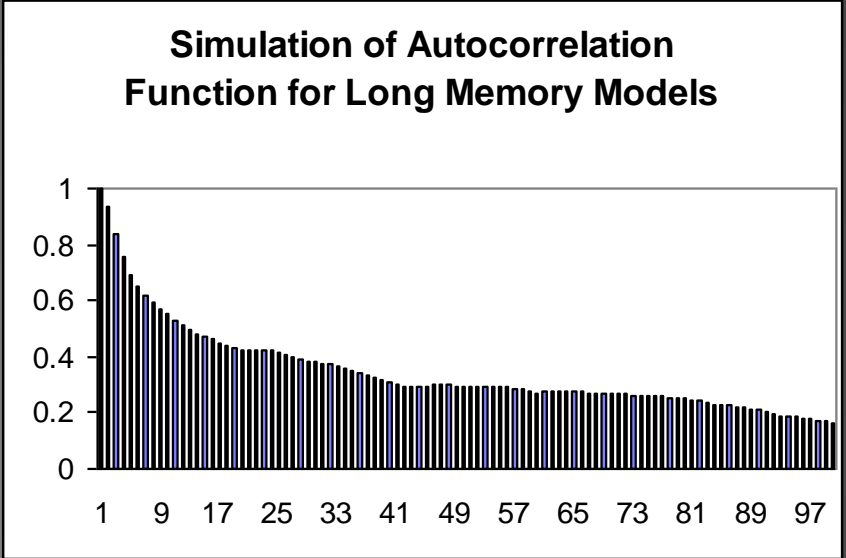
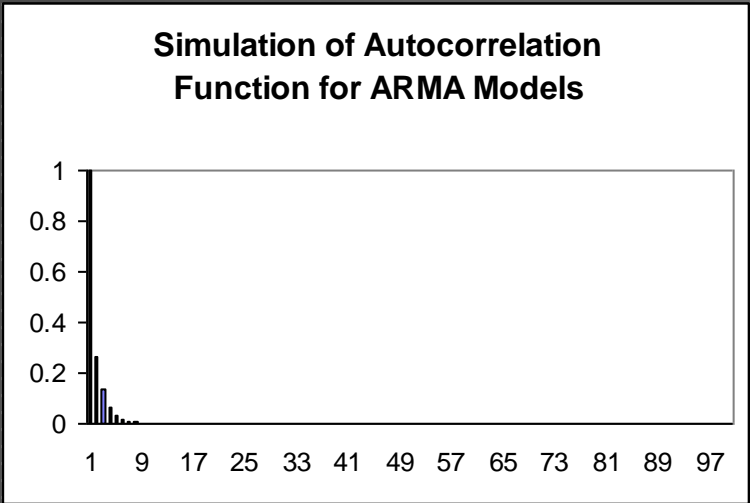
Solutions:

- ▶ Bandi and Russell (2005): Optimal Sampling
- ▶ Zhang, Mykland and Ait-Sahalia (2005): Use all information
- ▶ Hansen and Lunde (2005): Kernel estimator

# LONG MEMORY (I)

- ▶ The phenomenon of long memory has been known since the time that ancient Egyptian hydrologists studied the flows and inflows of the river Nile. The idea is very simple and states that the effects of an event (shock) persist over a long period of time
- ▶ From an empirical point of view, the presence of long memory may be defined in terms of the observed autocorrelations, which show high dependence between very distant observations
- ▶ The autocorrelation function of a long memory process decays at a hyperbolic rate, while the autocorrelation function of an ARMA process decays exponentially

# LONG MEMORY (2)



# BLACK-SCHOLES

$$c = Se^{-r^*\tau} N(d_1) - Ke^{-r\tau} N(d_2)$$

$$p = Se^{-r^*\tau} [N(d_1) - 1] - Ke^{-r\tau} [N(d_2) - 1]$$

$$\text{where } d_1 = \frac{\ln(Se^{-r^*\tau} / Ke^{-r\tau})}{\sigma\sqrt{\tau}} + \frac{\sigma\sqrt{\tau}}{2}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}$$

$$N(d) = \int_{-\infty}^d \Phi(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^d \exp\left(-\frac{1}{2} x^2\right) dx$$

# IMPLIED VOLATILITY

- ▶ The only parameter in the B-S pricing formula that cannot be observed directly in the market is the volatility of the underlying asset (stock in our example)  $\implies$  we solve B-S to find volatility  $\implies$  *Implied volatility*
- ▶ IV: It is the volatility implied by an option price observed in the market. Many analysts believe that IV represents the expectation of market participants on future volatility

# VOLATILITY

- ▶ Historical estimates: Squared returns, absolute returns, range, realized volatility → backward looking
- ▶ Implied volatility → forward looking
- ▶ Options on implied volatility → important for the market
- ▶ We can combine historical measures with implied volatility  
Historical estimates → model

# IMPLIED VOLATILITY

- ▶ Black-Scholes: volatility is constant → implied volatility from options on the same underlying asset and same maturity but with different strike price are different – i.e. they are not constant



Volatility smile (or smirk)

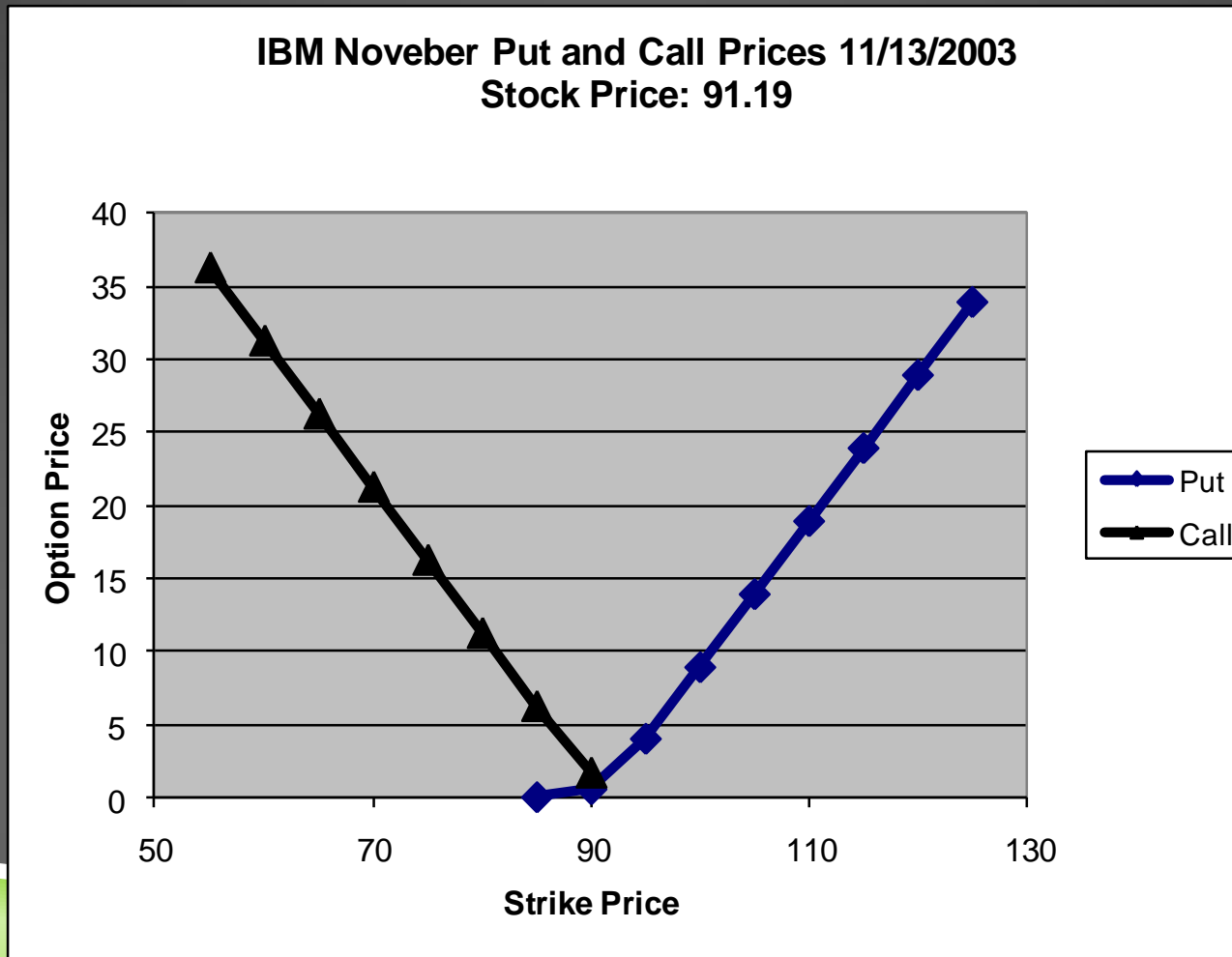
Out-of-the-money puts appear overpriced → different implied volatilities are clearly inconsistent with the joint assumption of a lognormal distribution for prices and efficient markets

This effect was pronounced after the stock market crash of 1987, raising the possibility that the market expected another crash, although with low probability

# IMPLIED VOLATILITY

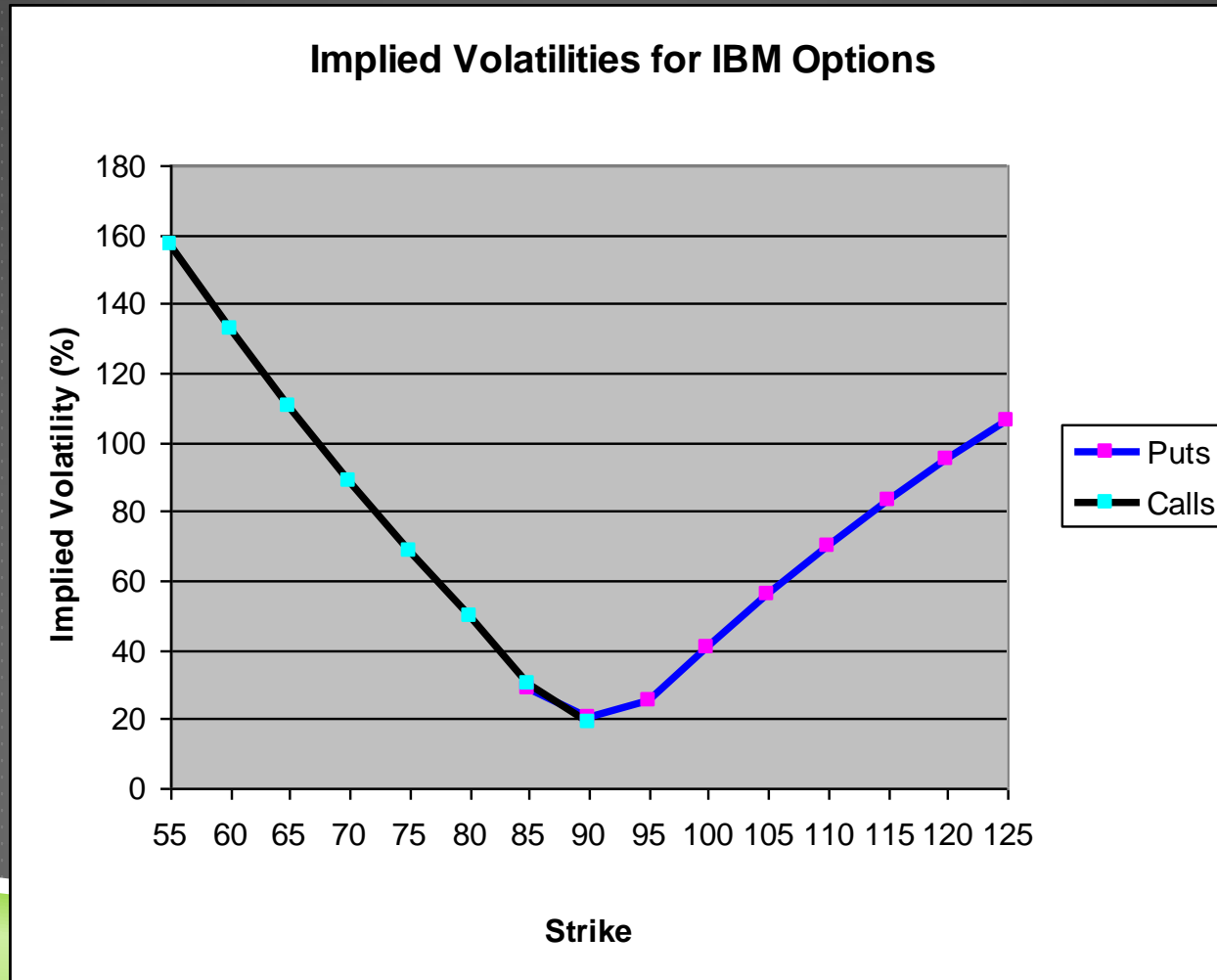
- ▶ The Black-Scholes equation fundamentally assumes lognormally distributed assets.
  - ▶ One would expect that two options written on the same stock and the same maturity date with different strikes would have prices based upon the same volatility.
  - ▶ This is not generally the case.
  - ▶ Consider implied volatility data on IBM and Bank of America options from 11/13/2003.
  - ▶ IBM was selling for 91.15 per share when these quotes were taken. November options expire on 11/22/2003.
  - ▶ Bank of America was selling for 75.48 per share when these quotes were taken. December options expire in 37 days on 12/20/2003.

# IBM OPTION PRICES



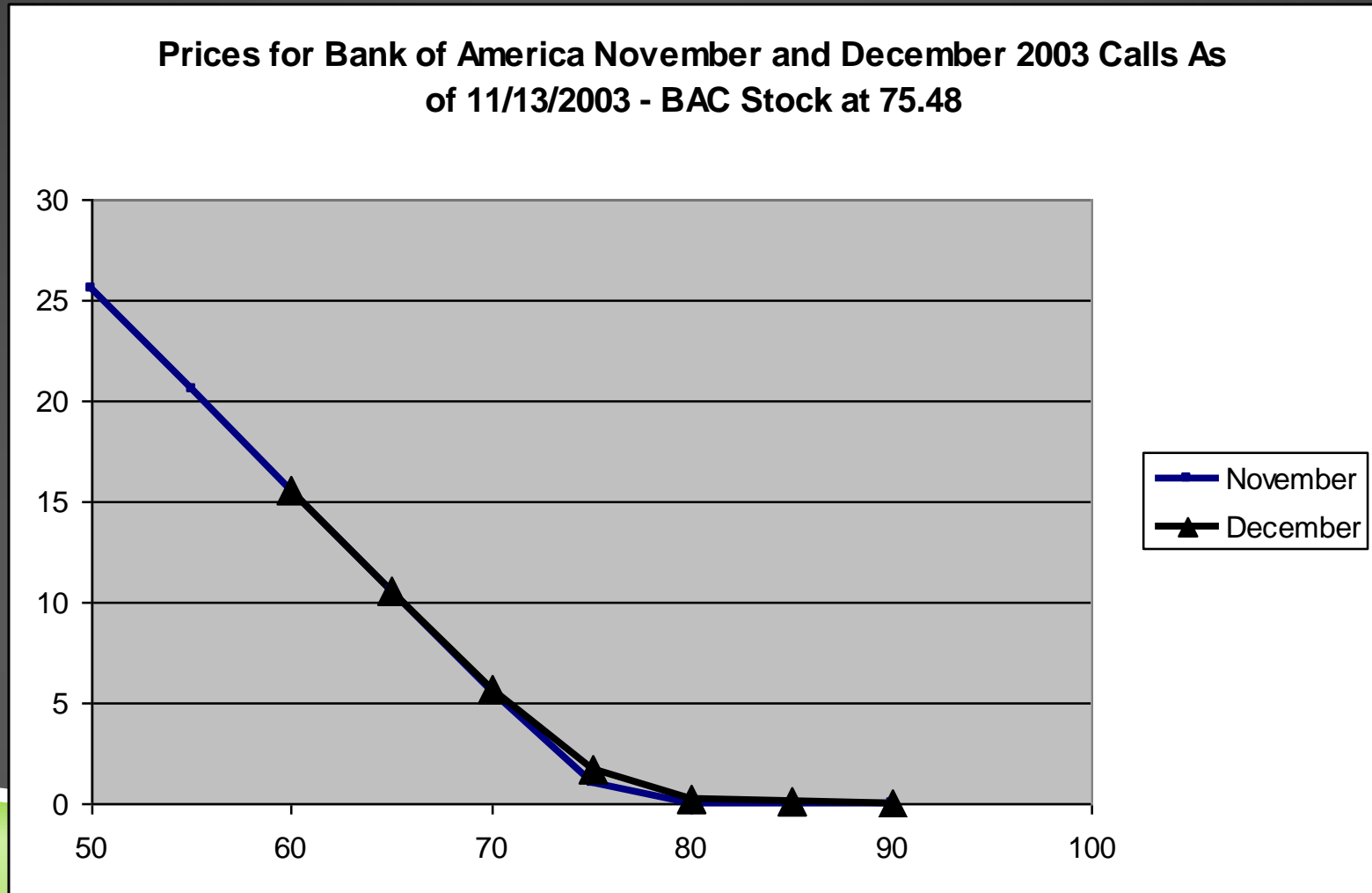
Source: Bloomberg

# IBM OPTIONS IMPLIED VOLATILITIES



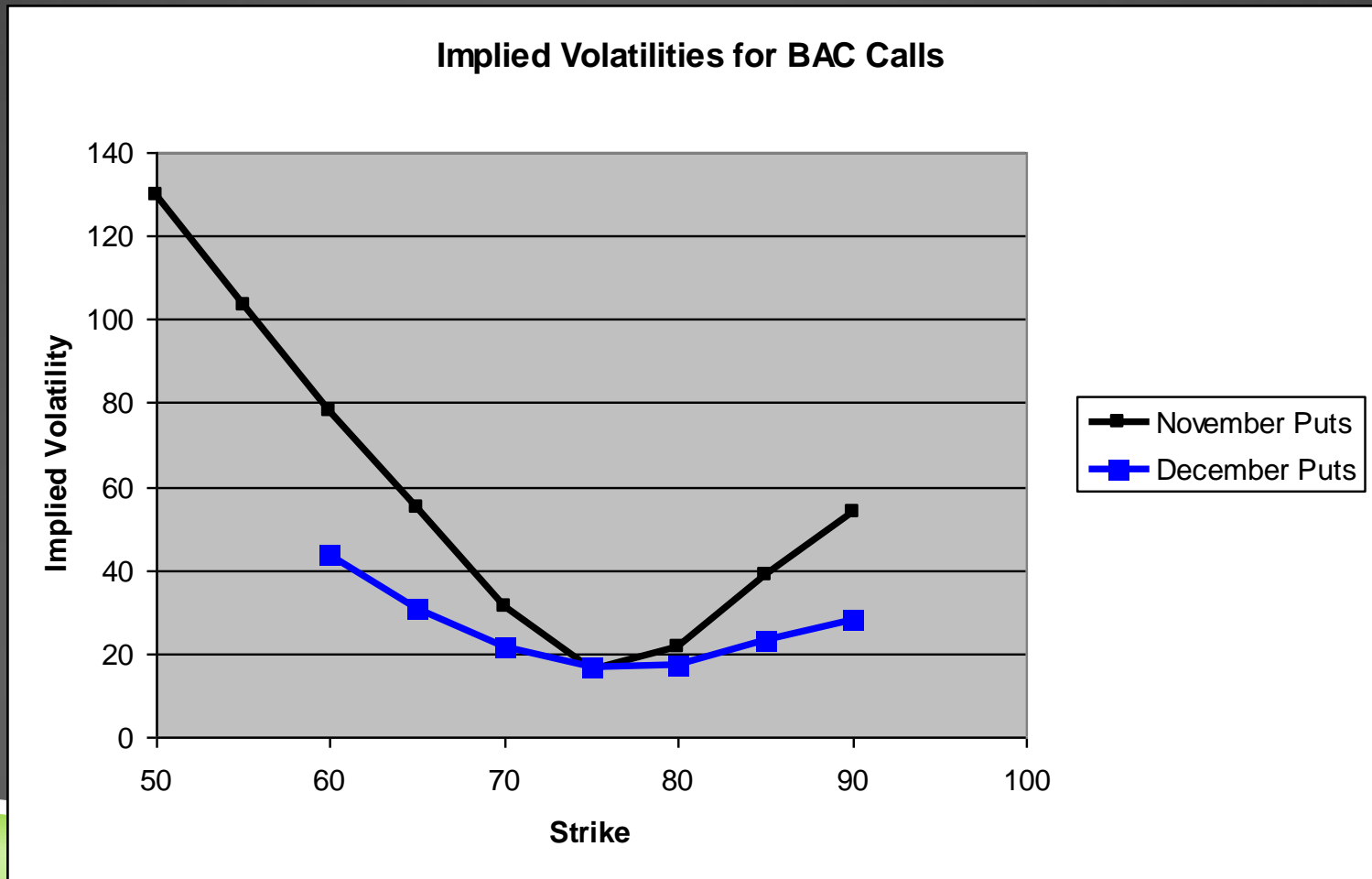
Source: Bloomberg

# BANK OF AMERICA OPTION PRICES



Source: Bloomberg

# BAC PUT IMPLIED VOLATILITIES



Source: Bloomberg

# VOLATILITY SMILES

- ▶ Notice how in both cases the lower strike price volatilities are much higher than the higher strike price volatilities.
  - ▶ This is referred to as the “volatility skew”.

# VOLATILITY SMILES

- ▶ Now, recall that in our discussion of Black-Scholes (or the binomial models, for that matter), we treated volatility as a parameter.
  - ▶ We never treated it as a function of anything, and never as a function of the strike price.
  - ▶ Although we never talked about volatility as being a function of time to maturity, it is somewhat easier to accept that it could be.
    - ▶ One could think of the volatility as an expected volatility until maturity, and if volatility was expected to change after the first maturity date, then one could get different expected volatilities over two time periods.
  - ▶ One must also rule out simple issues, such as asynchronous trading, bad quotes, and minimum prices (such as \$.05 for BAC options.)

# VOLATILITY SMILES

- ▶ The simple fact is, however, that we consistently observe volatility smiles across almost all types of options, underlying, etc.
  - ▶ Admittedly the actual shapes may be “frowns” or “smirks”, but the basic notion of a volatility that varies with strike price is what matters.
- ▶ The most important idea you must remember is that Black-Scholes does not determine option prices, the market does that. Black-Scholes is simply a sophisticated model that **approximates** (very well) market prices.
  - ▶ In one sense volatility smiles are simply the limits of the model showing!

# VOLATILITY SMILES

- ▶ Put-Call parity defines a general relationship that must exist to prevent arbitrage:

$$c_{mkt} + Ke^{-rT} = p_{mkt} + Se^{-qT}$$

(assuming that we have a dividend yield in this case).

- ▶ Since Black-Scholes obeys put-call parity, we can state

$$c_{BS} + Ke^{-rT} = p_{BS} + Se^{-qT}$$

- ▶ And since  $K, S, q, r,$  and  $T$  are independent of the pricing model, we know that:

$$c_{Mkt} - p_{Mkt} = Se^{-qT} - Ke^{-rT}$$

$$c_{BS} - p_{BS} = Se^{-qT} - Ke^{-rT}$$

So that:

$$c_{BS} - p_{BS} = c_{Mkt} - p_{Mkt}$$

or:

$$c_{BS} - c_{Mkt} = p_{BS} - p_{Mkt}$$

# VOLATILITY SMILES

- ▶ This implies that the difference between Black-Scholes option values and market values should be the same regardless as to whether we are pricing puts or calls.
- ▶ Another way of saying this is that for a given strike, maturity and volatility, the Black-Scholes call and put prices will differ from the market price by the same amount.
  - ▶ Since, by definition, the implied volatility is the value for  $\sigma$  that makes that difference 0, then the implied volatilities of calls and puts with the same maturity and strike price should be the same.
  - ▶ It is only when you have different strikes (or maturities) that you observe different implied volatilities, not when you have different types of options.
  - ▶ Thus, if you are trying to price a put option with a given strike, the appropriate volatility to use is the implied volatility from the equivalent call.

# IMPLIED DISTRIBUTIONS

- ▶ It is possible to extract from options prices the implied distribution of the underlying asset price.
  - ▶ In the appendix, Hull demonstrates the following, assuming that a European call has a risk-neutral probability density function  $g()$  and a constant interest rate  $r$ :

$$c = e^{-rT} \int_{S_T=K}^{\infty} (S_T - K) g(S_T) dS_T$$

- ▶ Of course, what we are interested in is extracting  $g()$ . First if we differentiate with respect to  $K$ , we get:

$$\frac{\partial c}{\partial K} = -e^{-rT} \int_{S_T=K}^{\infty} g(S_T) dS_T$$

- ▶ Differentiating again yields:

$$\frac{\partial^2 c}{\partial K^2} = e^{-rT} g(K)$$

# IMPLIED DISTRIBUTIONS

- ▶ From this we can solve for  $g(K)$ :

$$g(K) = e^{rT} \frac{\partial^2 c}{\partial K^2}$$

- ▶ Hull points out that for a given strike  $K$  and for “sufficiently small values of  $\delta$ , you can estimate  $g(K)$  by:

$$g(K) = e^{rT} \left( \frac{c_1 + c_3 - 2c_2}{\delta^2} \right)$$

- ▶ Where  $c_2$  has strike of  $K$ ,  $c_1$  has strike of  $K-\delta$ , and  $c_3$  has strike of  $K+\delta$ .
- ▶ The difficulty is that in practice one rarely has enough traded options to build a “good” distribution.

# IMPLIED DISTRIBUTIONS

- ▶ To see how this can be done, however, and to illustrate the potential problems with it, let's look at January options on GM as of 11/19/2003.
  - ▶ GM closed at 42.21
  - ▶ Options Expire on January 17 or in 59 days ( $t=59/365=.16164$ ).
  - ▶ Here are the January options prices.

Strike	Price
25	17.1
27.5	15.2
30	12.1
32.5	10.4
35	7.1
37.5	5.2
40	2.8
42.5	1.5
45	0.5
47.5	0.25
50	0.1

- ▶ Since we have \$2.5 strike increments, that is what we will use for  $\delta$ .
- ▶ Since we need strikes both before and after  $K$ , we cannot get values for  $K=25$  and  $K=60$ . So we will start with  $K=40$ , such that:  $c_1=5.2$ ,  $c_2=2.8$ , and  $c_3=1.5$ , such that our estimate of  $G(k)$  is

$$g(K) = e^{.096(.1616)} \left( \frac{5.2 + 1.5 - 2 * 2.8}{2.5^2} \right) = 0.176$$

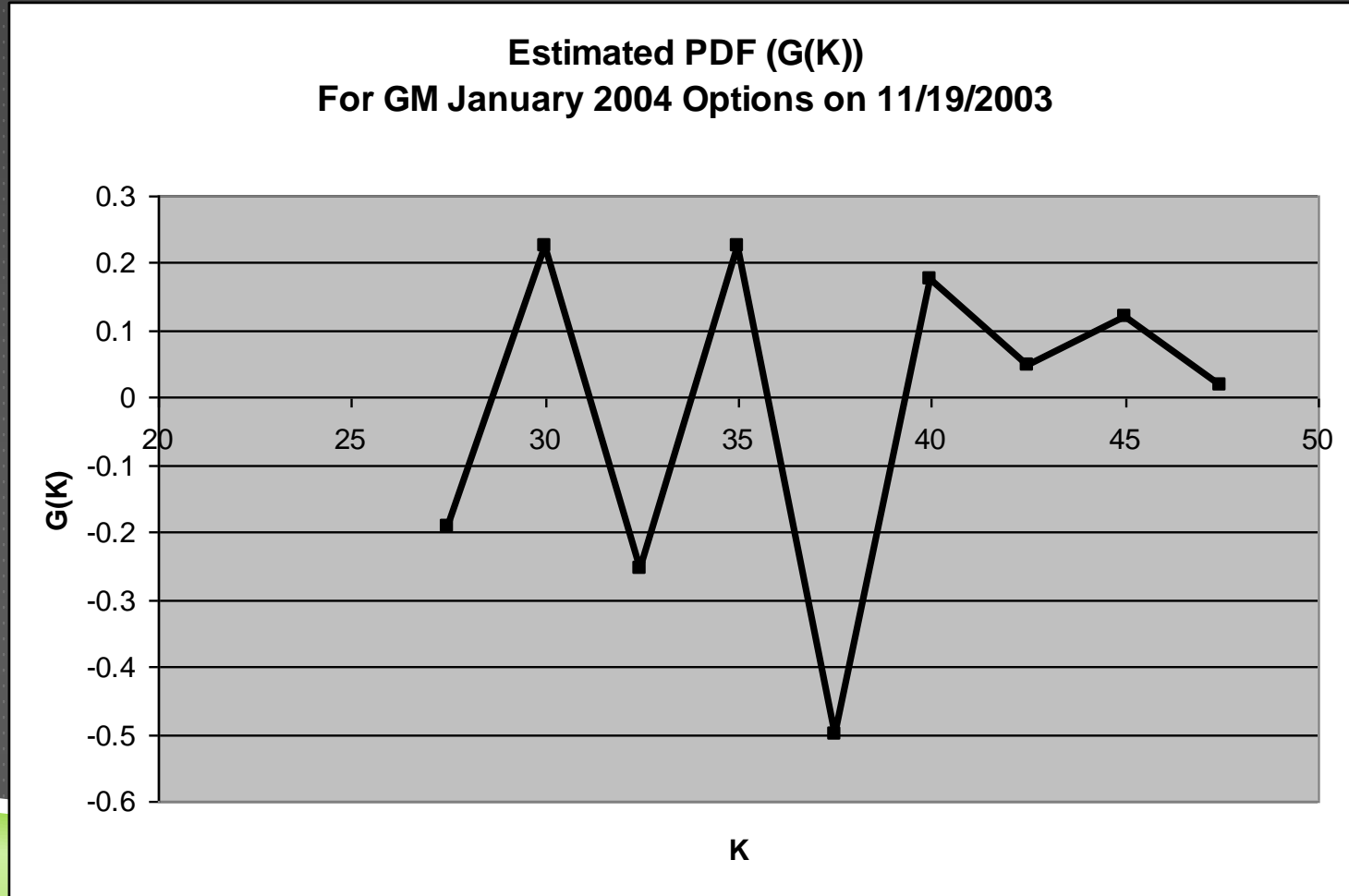
# IMPLIED DISTRIBUTIONS

- ▶ The difficulty is that we are not guaranteed that for a given set of observed prices we will even get positive values for  $g(K)$ .
  - ▶ For example, consider  $g(37.5)$ ,  $c_1=7.1$ ,  $c_2=5.2$ ,  $c_3=2.8$ . From this we can calculate our  $g()$  estimate:

$$g(K) = e^{0.0096(.1616)} \left( \frac{7.1 + 2.8 - 2 * 5.2}{2.5^2} \right) = -0.080124$$

- ▶ There are, of course, several reasons for this, including:
  - ▶ Asynchronous prices
  - ▶ Differences in volumes “behind” the prices
  - ▶ Our  $\delta$  is quite large.
- ▶ But it does illustrate the difficulties of this estimation method, the graph on the next page demonstrates the distribution.

# IMPLIED DISTRIBUTIONS

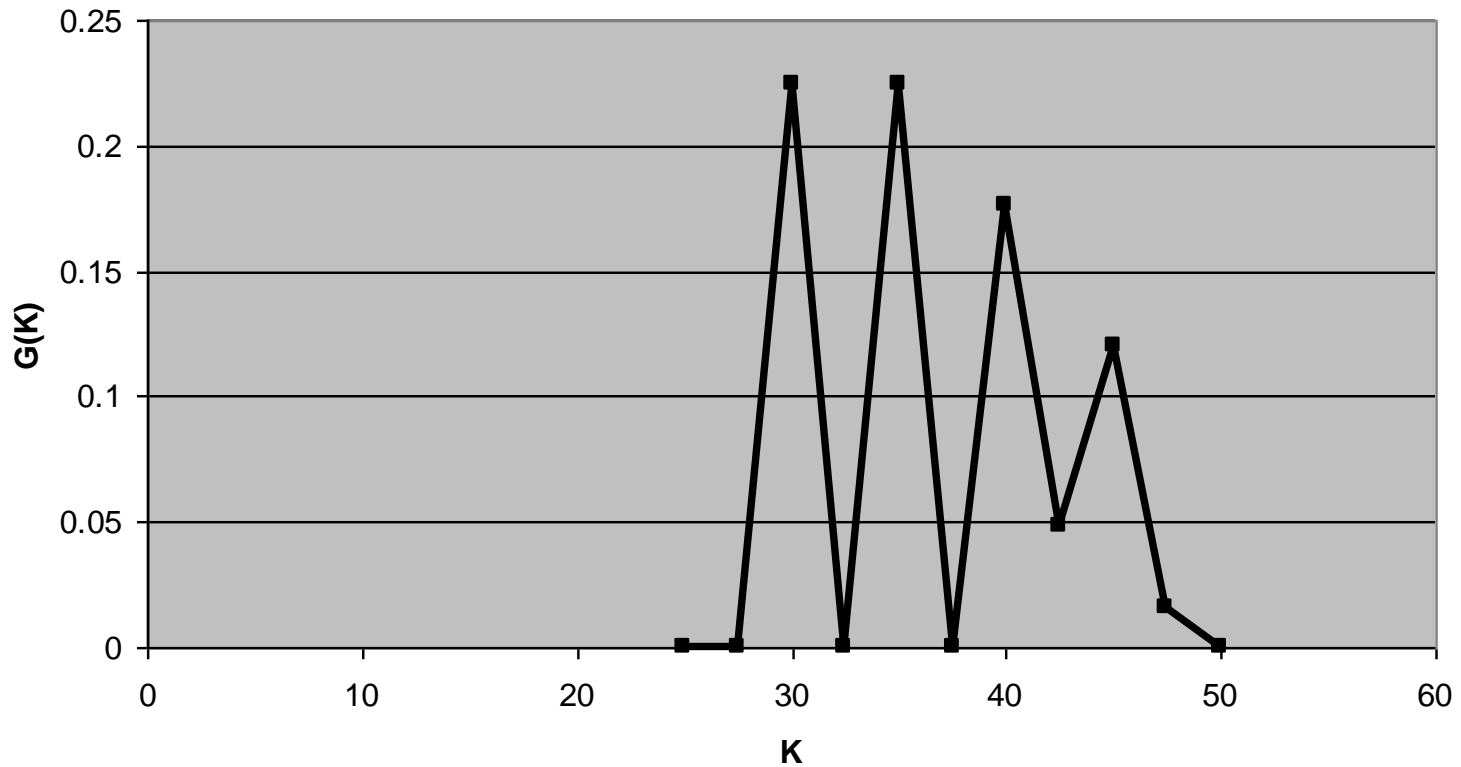


# IMPLIED DISTRIBUTIONS

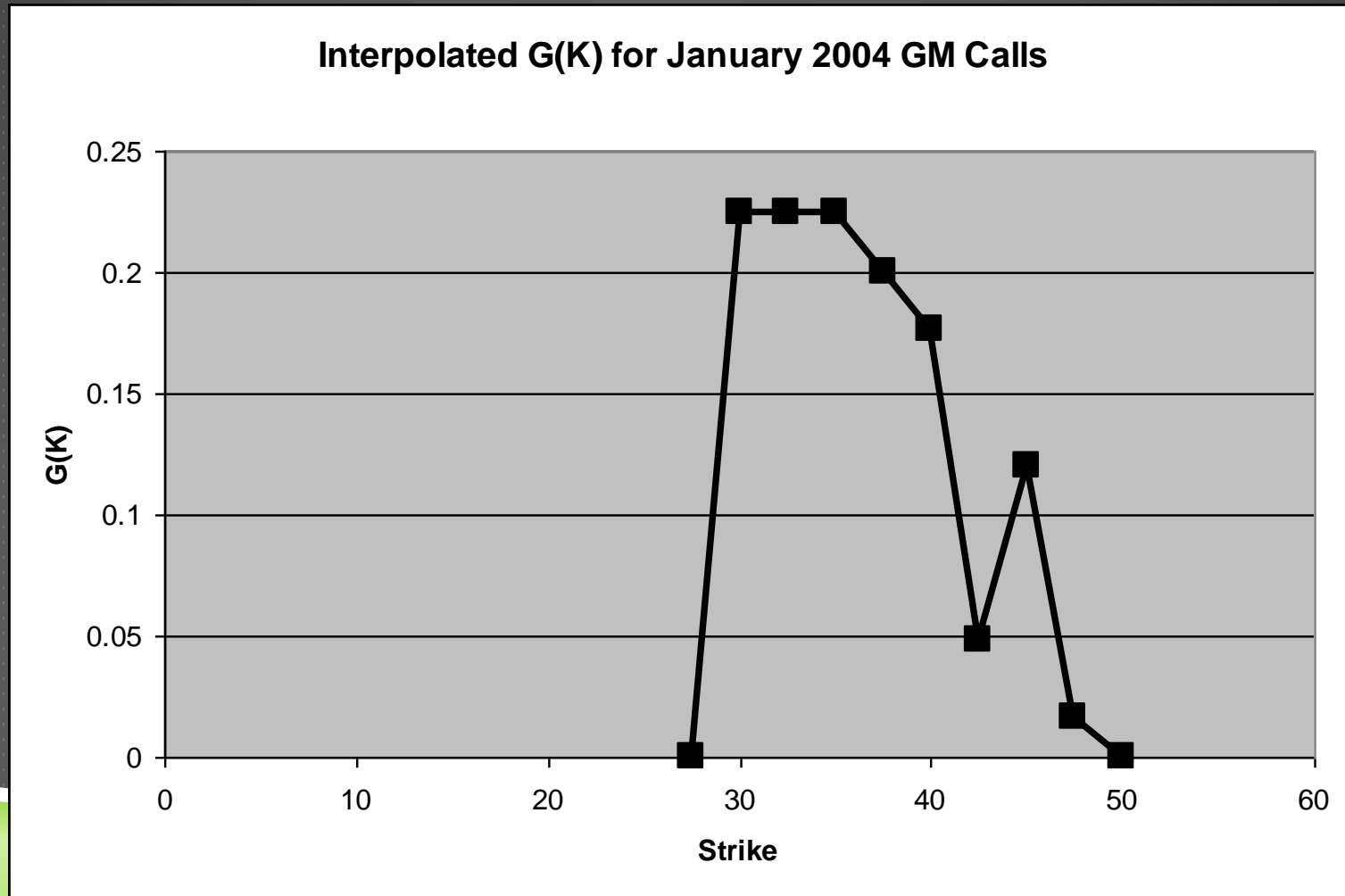
- ▶ Obviously this distribution does not look terribly lognormal.
  - ▶ You can get some better results simply by assuming that any negative value is really a 0.
  - ▶ It looks even more interesting if you interpolate for any negative values...

# IMPLIED DISTRIBUTIONS

**Estimated  $G(K)$ , Treating Negative Values as Zero  
GM January 2004 Options as of 11/19/2003**



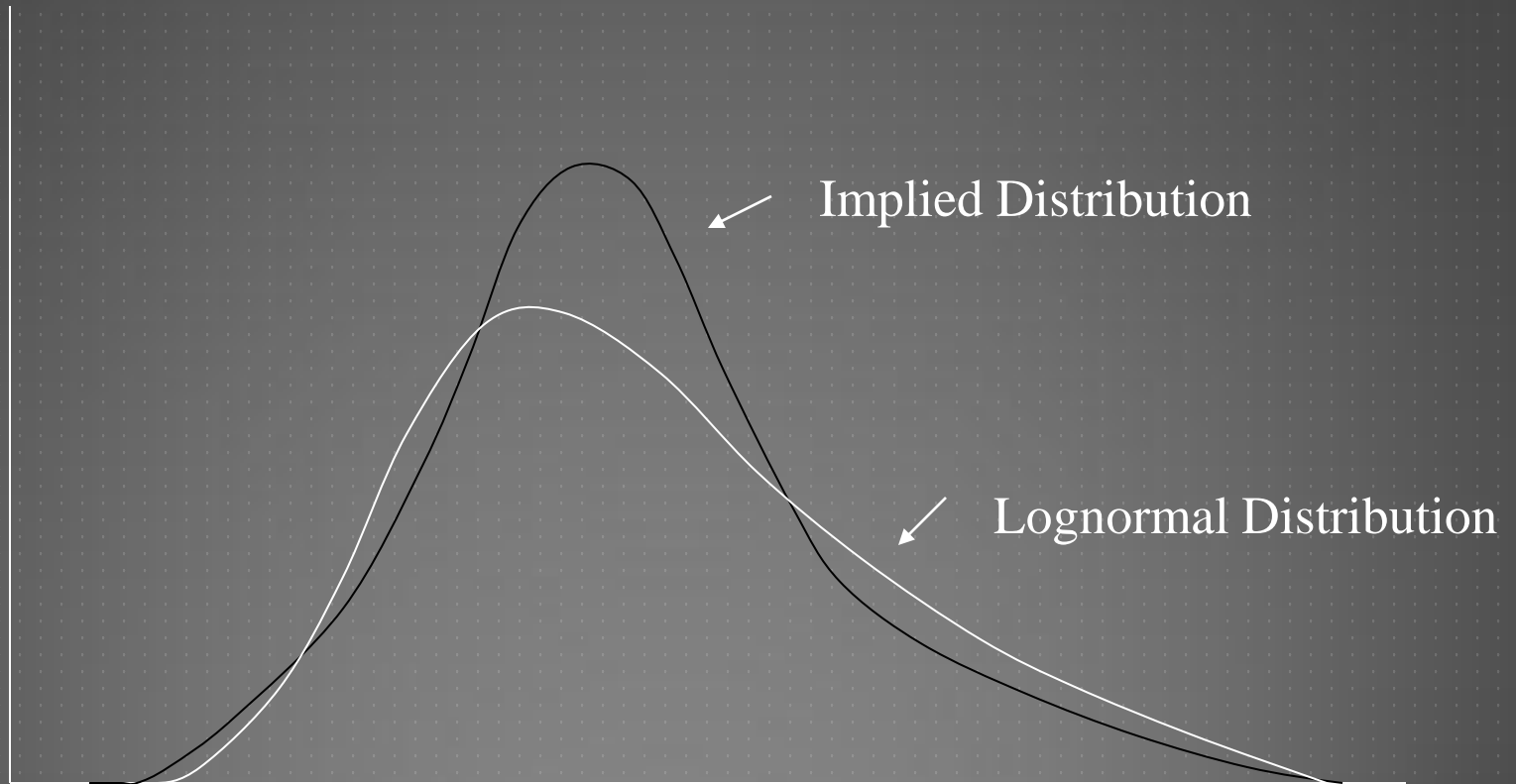
# IMPLIED DISTRIBUTIONS



# IMPLIED DISTRIBUTIONS

- ▶ There have, however, been a number of studies, notably by Jackworth and Rubinstein, that examine (with very good econometrics) the implied distributions for stock prices implied by option price.
- ▶ What they generally find is that the implied distribution is not particularly lognormal. Indeed, the graphs generally look something like this:

# IMPLIED DISTRIBUTIONS



# IMPLIED DISTRIBUTIONS

- ▶ The main issue from these studies is that the implied distributions are not really that close to a lognormal.
  - ▶ Other studies have examined actual stock returns and note that stock returns more closely resemble a logistic as opposed to a normal distribution.
  - ▶ This does seem to imply that Black-Scholes is not a perfect approximation to market prices.
    - ▶ But it is still probably good enough for most trading, since transactions costs are still larger than “errors” in Black-Scholes.

# IMPLIED DISTRIBUTIONS

- ▶ One interpretation of the volatility smile in general (or volatility skew), is that it represents some form of a conditional expectation.
  - ▶ That is, that if the option were to get “in the money” the firm would have to have volatility in this range.
  - ▶ Hull notes this by appealing to the leverage argument.
  - ▶ The basic argument is that as a firm’s equity declines in value, then its leverage, by definition, increases, unless the company has no debt.
    - ▶ That is, the firm’s debt becomes a higher percentage of its equity value.
    - ▶ *Ceteris Paribus*, increased leverage tends to increase the volatility of a stock (since it is riskier). As a result, when equity declines, the volatility of the stock should increase.
    - ▶ Rubinstein notes that the volatility skew really became pronounced after the October, 1997 crash.

# VOLATILITY SURFACES

- ▶ We have demonstrated that volatility is related to strike, and our earlier graph of BAC presented some evidence that it is related to maturity as well.
- ▶ In practice, traders realize that volatility is at least related to both strike and maturity. They normally refer to the **volatility surface**.
  - ▶ Normally what a trader will do is collect a series of implied volatilities from trades and fill out a grid that is the volatility surface.
  - ▶ When they are then asked to price an option they then look at their grid. If the strike/maturity of that option is one for which they have market prices, they use that price (or volatility, which is the same thing!) If they do not have that strike/maturity, they interpolate the volatility from their “good” data and then plug that volatility into Black-Scholes to get the price at which they will trade.
  - ▶ In this sense, you can see that Black-Scholes is really used primarily as a way of **interpolating** market prices.

# A RETURN TO GREECE

- ▶ Obviously the fact that volatility is non-constant has some impact on our analysis if the Greek letters.
- ▶ Because there is not a good formal theory of why implied volatility will change, traders have developed a number of rules of thumb that they observe and use.
- ▶ These rules of thumb are:
  - ▶ Sticky strike rule
  - ▶ Sticky delta rule
  - ▶ Implied volatility function/Implied Tree method

# STICKY RULES

## ▶ Sticky Strike Rule:

- ▶ Simplest of the rules. Simply assume that the implied volatility of an option remains the same from one day to the next, even though options on the same underlying asset, but with different strikes or maturities, may not be the same.
- ▶ Under this rule, all of the standard Greeks from Black-Scholes are correct, provided that you use the current implied volatility in calculating those Greeks.

## ▶ Sticky Delta Rule:

- ▶ This is more complex, but essentially says that we will adjust our volatility as the moneyness of the option changes.
- ▶ In reality this rule is simply an advanced heuristic so that we say how we adjust our volatility as the option price changes.

# STICKY RULES

## ▶ Sticky Delta Rule

- ▶ So how does this rule work?
- ▶ Assume that the relationship between  $S/K$  is constant, and as  $S/K$  changes adjust the volatility proportionately.
- ▶ Example:
  - ▶ Say that stock was priced at 100, and that an at-the-money option had volatility of 25%. The relationship between volatility and  $S/K$  would be:
$$.25 = 100/100 = 1$$
  - ▶ If tomorrow the stock price were to decline to \$98, this would mean that  $S/K$  would decline to  $98/100 = 0.98$ , so you would adjust your implied volatility to be  $.25(.98) = .245$

# STICKY RULES

- ▶ Under this rule, our Greeks are not the same as those that would be reported by Black-Scholes.
- ▶ To understand this just consider this logic:
  - ▶ We know that the Black-Scholes Delta is the partial derivative of the Black-Scholes equation with respect to the underlying, and that this is simply given by  $N(d_1)$  and  $d_1$  is given by:

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

- ▶  $\sigma$  is treated as a constant in  $d_1$ , but we know it would be more appropriate to treat it as  $\sigma(S)$ . Which means there should be a chain rule effect in our derivative.
- ▶ We also know that a Call option will increase in value as volatility increases (since Vega > 0).
- ▶ Under the Sticky Delta rule, we know that as you increase  $S$ ,  $S/K$  increases, and so the implied volatility will increase, i.e. that  $\frac{\partial \sigma_{\text{Implied}}}{\partial S} > 0$
- ▶ Thus we know that our Black-Scholes Delta must be **understating** the options sensitivity.

# STICKY RULES

- ▶ Under this rule, the delta of the call option would actually be given by:

$$Delta = \frac{\partial c_{BS}}{\partial S} + \frac{\partial c_{BS}}{\partial \sigma_{Implied}} \frac{\partial \sigma_{Implied}}{\partial S}$$

- ▶ That is, you must adjust it for the cross-partials.
- ▶ The problem is that these simple stick-rules are not self-consistent. That is, you can find arbitrage between them, so the market cannot strictly follow them.
  - ▶ The error in their approximation to market prices exceeds the transactions costs barrier in the market.
  - ▶ Thus they are simply “quick and dirty” rules of thumb and not really pricing models as we think of them.

# STICKY RULES

- ▶ Jackworth and Rubinstein (and others) have demonstrated that it is possible to build internally self-consistent lattices that allow for volatility to vary with strike prices and maturities.
- ▶ These lattices basically work by taking the implied distributions that we calculated earlier, and then building a lattice that is exactly consistent with that implied distribution.

# LARGE-JUMP MODELS

- ▶ Hull notes that when there is a single large jump anticipated in a stock, that it is possible for implied volatilities to be significantly higher than will occur after the jump, regardless of the direction of the jump.
  - ▶ He presents the example of an option on a stock, where some announcement on the stock is expected in the very near future that will send the stock into one of two regimes. The stock is currently priced at \$50.
    - ▶ If the news is bad, the stock will fall to \$48, but will have constant volatility of 20%.
    - ▶ If the news is good, the stock will rise to \$58, but will have constant volatility of 20%.
    - ▶ Assume that the probability of the news being good is 50%.

# LARGE-JUMP MODELS

- ▶ Assume that this news will come out in the next short period of time.
- ▶ How much would an option on this stock, with .25 years to maturity and with a risk-free rate of 5% be worth?
- ▶ Since the announcement is imminent, we ignore any discounting effects, and simply calculate the value of the option in both of the two potential new states of the world, and then multiply them by their probabilities of occurring.
- ▶ In this case, Black-Scholes tells us that the price of the two options would be:
  - ▶  $BS(S=42, K=50, r=.05, \sigma=.20, T=.25)=0.101$
  - ▶  $BS(S=58, K=50, r=.05, \sigma=.20, T=.25)=8.743$
- ▶ Thus, the value of the option at time 0 would be:
  - ▶  $Value = .5(0.101) + .5(8.743) = 4.422.$
- ▶ If we then use this price and ask Black-Scholes to solve for the implied volatility, we get an implied volatility of 41.48%

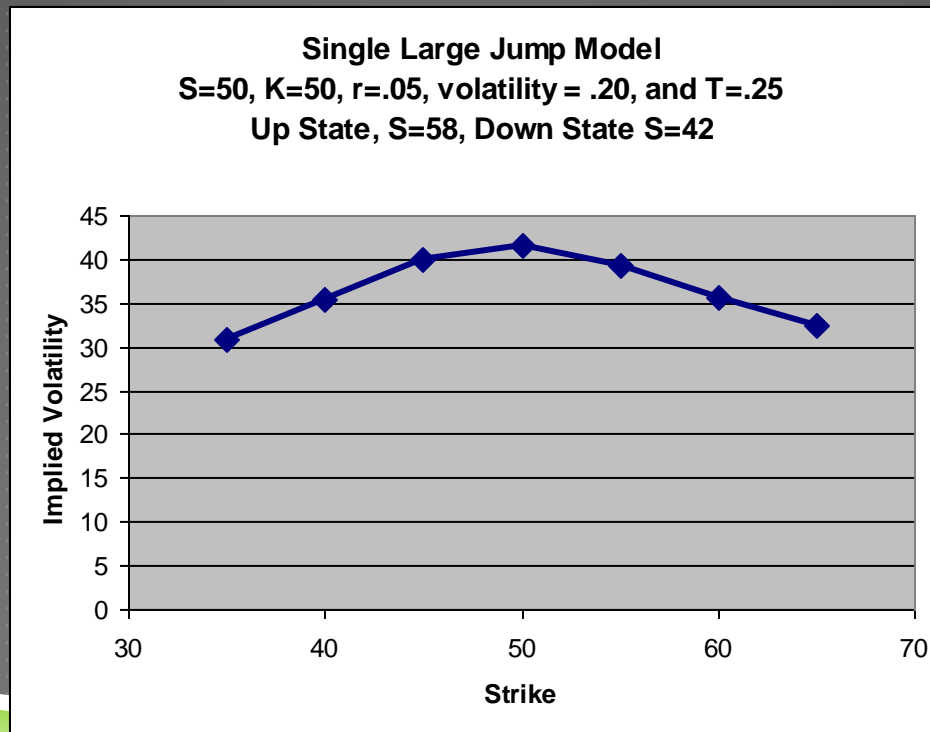
# LARGE-JUMP MODELS

- Indeed, the Black-Scholes implied values and volatilities can be calculated for strike-prices from 35 to 65 dollars:

Strike	Good News	Bad News	Price Today	Implied Volatility
35	23.435	7.471	15.453	30.95
40	18.497	3.169	10.833	35.46
45	13.565	0.771	7.168	39.94
50	8.743	0.101	4.422	41.48
55	4.546	0.008	2.277	39.27
60	1.764	0.000	0.882	35.66
65	0.494	0.000	0.247	32.55

# LARGE-JUMP MODELS

- ▶ This leads us to a “volatility frown”!



# OPTIONS PRICES: 3 EPOCHS

Before Black-Scholes: random (almost)

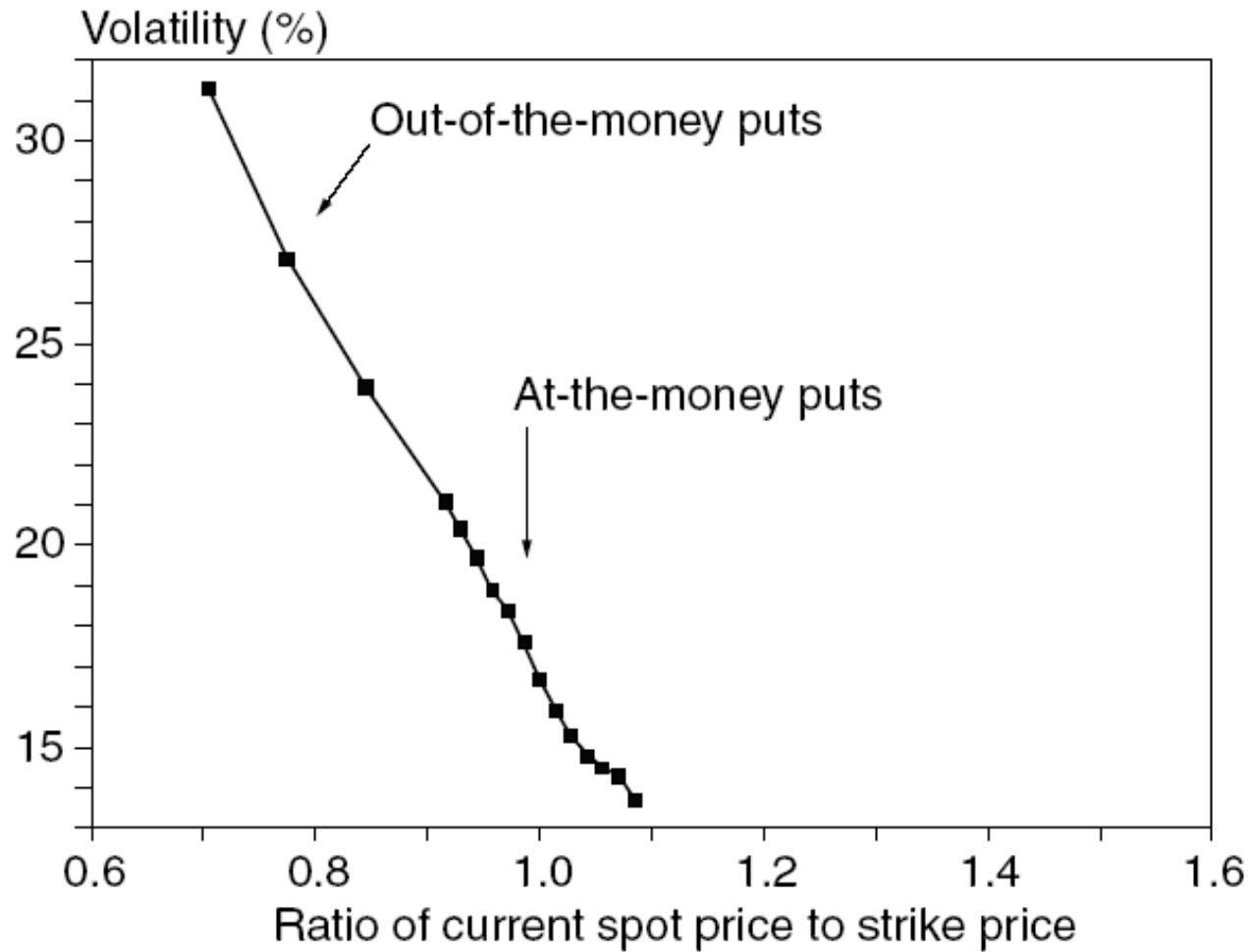
From Black-Scholes till October 1987: Black-Scholes prices

After October 1987: At-the money options → Black-Scholes prices

Out-of-the money and in-the money options →  
no Black-Scholes

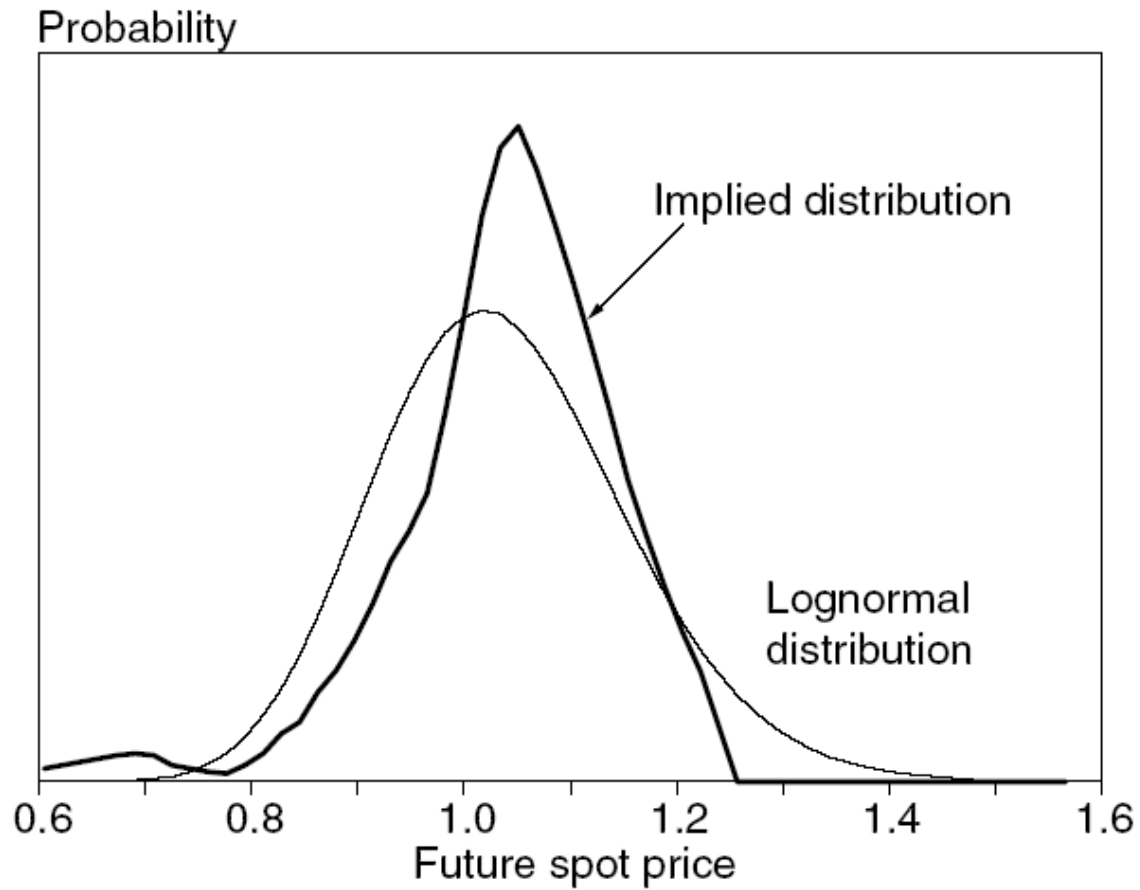
# VOLATILITY SMILE/SMIRK

**FIGURE 14-6 Smile Effect**



# IMPLIED DISTRIBUTION - RUBINSTEIN (1994)

**FIGURE 14-7 Implied Distribution**



# IMPLIED DISTRIBUTION DIFFERENT FROM LOG-NORMAL

This puzzling result can be given two interpretations

1) The market predicts a small probability of a future crash

2) This distribution derived from option prices assumes risk-neutrality, since the Black-Scholes approach values options assuming investors are risk-neutral → this distribution may differ from the true, objective distribution due to a **risk premium**

Intuitively, investors may be very averse to a situation where they have to suffer a large fall in the value of their stock portfolios. As a result, they will bid up the price of put options, which is reflected in a higher-than-otherwise implied volatility