Lecture 9:
Entropy Methods for Financial Derivatives

Marco Avellaneda

G63.2936.001

Spring Semester 2009

Table of Contents

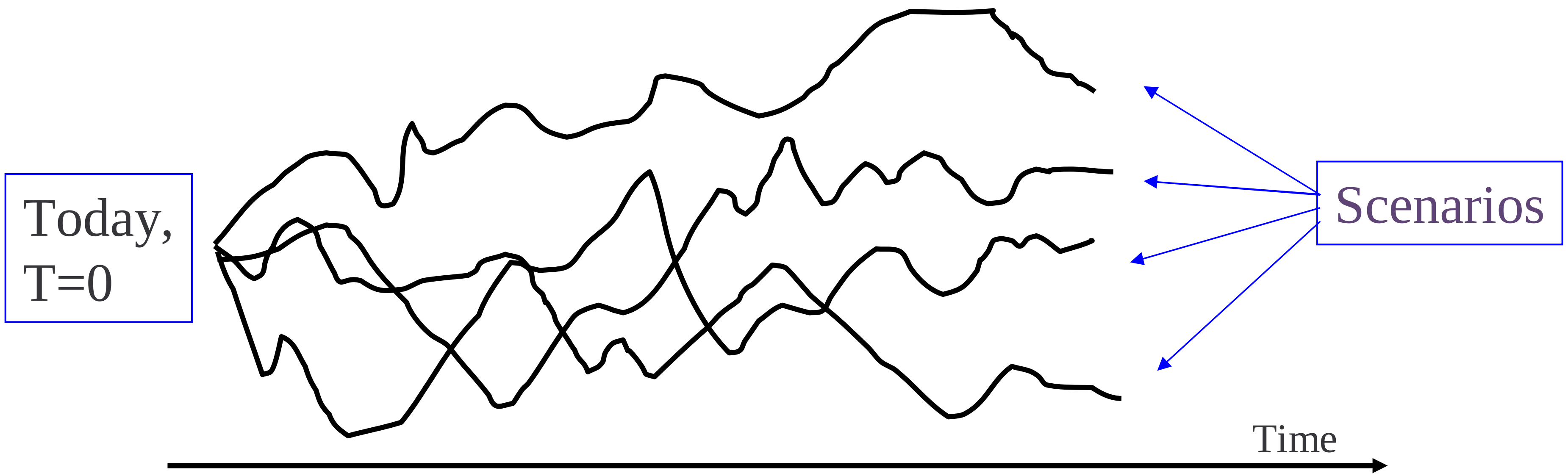
1. Review of risk-neutral valuation and model selection
2. One-dimensional models, yield curves
3. Fitting volatility surfaces
4. The principle of Maximum Entropy
5. Weighted Monte Carlo

1. Risk-Neutral Valuation and Model Selection

Risk-neutral valuation

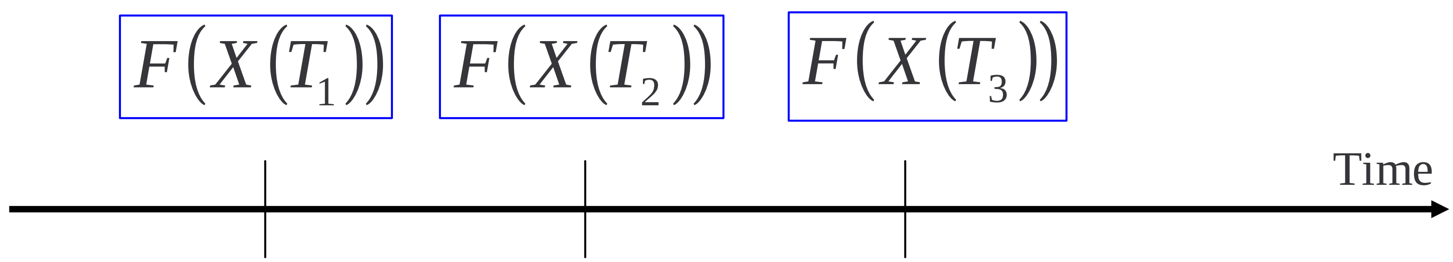
Future states of the economy or market are represented by scenarios described with state variables (prices, yields, credit spreads)

$$X(t) = [X_1(t), X_2(t), \ldots, X_n(t)] \qquad t \geq 0$$



Derivative securities & Cash-flows

Securities produce a stream of state-contingent cash-flows…



Present value of future cash-flows along each scenario:

$$G(X) = \sum_i \delta(t_i, X) F_i(X(t_i))$$

Discount factor

Arbitrage Pricing Theory

Consider a market with M <u>reference derivative securities</u>, with discounted cash flows

$$G_1(X), G_2(X) \ldots G_M(X)$$

trading at (mid-market) prices

$$C_1, C_2, \ldots, C_M$$

If we assume no arbitrage opportunities, there exists a <u>pricing probability measure</u> on the set of future scenarios such that

$$C_j = E^P\left(G_j(X)\right), \qquad j = 1,2,\ldots,M$$

Risk-neutral valuation

Consider the <u>target derivative security</u> that we wish to price

Present value of future cash-flows along each scenario (as specified by term sheet):

$$G(X) = \sum_i \delta(t_i, X) F_i(X(t_i))$$

$$\boxed{\begin{aligned} \text{Fair Value} &= E^P\{G(X)\} \\ &= E^P\left\{\sum_i \delta(t_i, X) F_i(X(t_i))\right\} \end{aligned}}$$

Fair value= expectation cash-flows, measured in constant dollars

What goes into the selection of a pricing model?

- Known statistical facts about the market under consideration

  -- relevant risk factors
  -- model for the dynamics of the underlying stocks, rates, spreads

  Gives rise to a set of scenarios and *a-priori* probabilities for these scenarios, or a stochastic process

- Known prices of cash, forwards and reference derivative securities that trade in the same asset class

  Gives rise to calculation of current risk-premia, to take into account the current prices of derivatives in the same asset class (needed for relative-value pricing)

Example 1: The Forward Rate Curve

a system of consistent forward rates

No arbitrage => a single interest rate for each expiration date
APT = > an interest rate ``curve''

$$Z(T) = E^P\left(\delta(X,T)\right)$$

Present value of \$1 paid in T years

$$F(T) = -\frac{1}{Z(T)}\frac{dZ(T)}{dT}$$

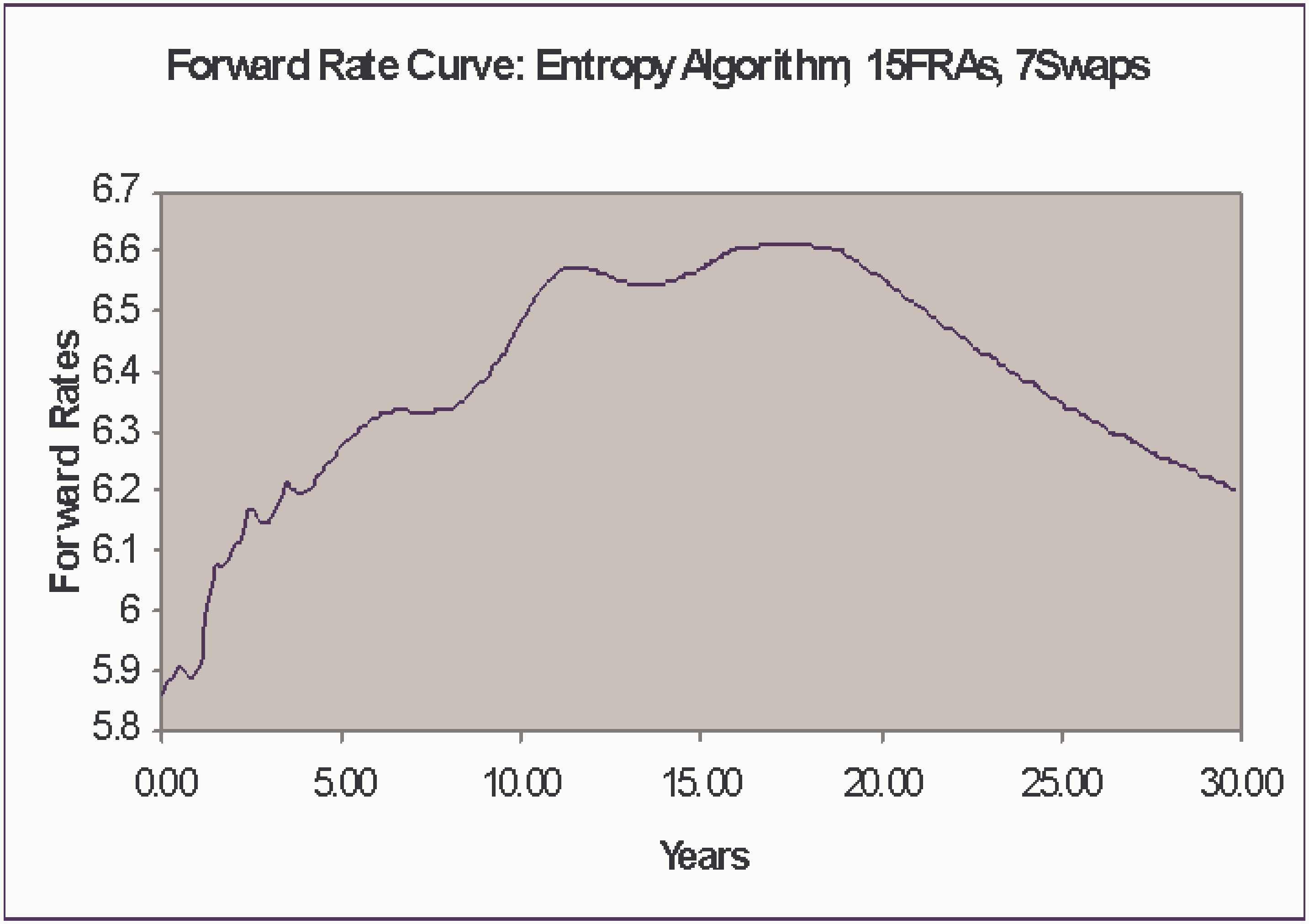
Instantaneous forward rate for loan in period (T,T+dT)

No-arbitrage implies the existence of a discount curve, or forward rate curve (interpolation, splines…)

Forward rate curve consistent with ED Futures and Swaps

Eurodollar futures/FRAs		
4m	94.2	0.0012
10m	94.14	0.0023
13m	94.08	0.003
16m	93.98	0.0044
19m	93.98	0.0092
22m	93.94	0.0131
25m	93.91	0.0176
28m	93.85	0.0234
31m	93.87	0.0232
34m	93.85	0.0371
37m	93.83	0.0447
40m	93.77	0.0522
43m	93.79	0.0637
46m	93.77	0.073
49m	93.75	0.083

Bonds/Swaps		
6y	5.815	0.3975
7y	5.8236	0.415
10y	5.847	0.4475
12y	5.8683	0.47
15y	5.9002	0.48
20y	5.9535	0.475
30y	6.05	0.375



Example #2: Equity Options

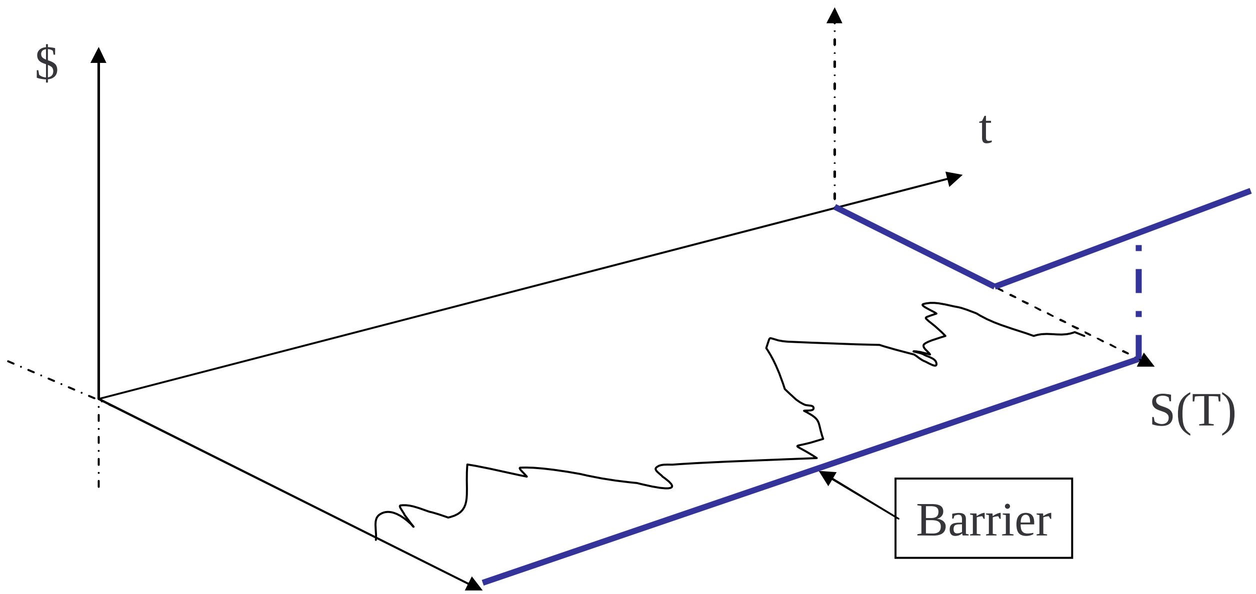
May 20, 2000 Call Series - AOL $56.500

Symbol	Issue	Intrinsic Value	bid	Ask	Volume	Open Interest
AOE EH	AOL MAY 20, 2000 $ 40.000 CALL	16.5	16.5	17	26	1159
AOE EV	AOL MAY 20, 2000 $ 42.500 CALL	14	14.125	14.625	0	0
AOE EI	AOL MAY 20, 2000 $ 45.000 CALL	11.5	12	12.5	21	79
AOE EW	AOL MAY 20, 2000 $ 47.500 CALL	9	9.75	10.125	0	1
AOO EJ	AOL MAY 20, 2000 $ 50.000 CALL	6.5	7.875	8.25	874	2009
AOO EK	AOL MAY 20, 2000 $ 55.000 CALL	1.5	4.875	5	498	13987
AOO EL	AOL MAY 20, 2000 $ 60.000 CALL	0	2.562	2.812	2429	58343
AOO EM	AOL MAY 20, 2000 $ 65.000 CALL	0	1.375	1.5	2060	48997
AOO EN	AOL MAY 20, 2000 $ 70.000 CALL	0	0.625	0.75	1470	15796
AOO EO	AOL MAY 20, 2000 $ 75.000 CALL	0	0.375	0.437	463	14290
AOO EP	AOL MAY 20, 2000 $ 80.000 CALL	0	0.125	0.25	799	8649
AOO EQ	AOL MAY 20, 2000 $ 85.000 CALL	0	0.062	0.187	16	6600
AOO ER	AOL MAY 20, 2000 $ 90.000 CALL	0	0.125	0.25	10	1493
AOO ES	AOL MAY 20, 2000 $ 95.000 CALL	0	0.062	0.125	0	1744
AOO ET	AOL MAY 20, 2000 $ 100.000 CALL	0	0.062	0.187	10	596
AOO EA	AOL MAY 20, 2000 $ 105.000 CALL	0	0.062	0.125	0	182

strikes

04/24/00 - 2:11p.m. Eastern. Current Stock Quotes are not delayed

Pricing Exotic Options



Barrier Option

Need to define a probability on stock price paths

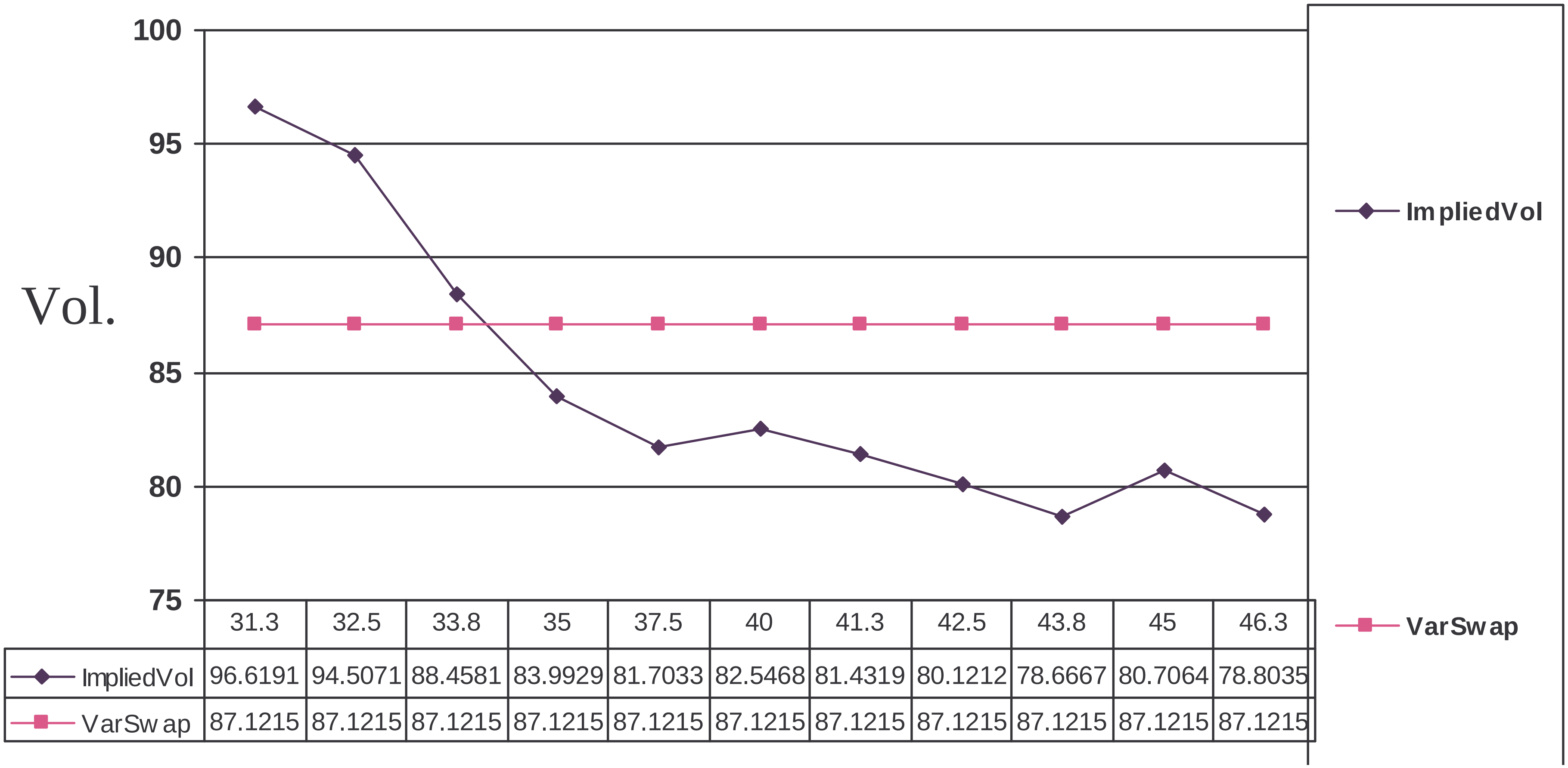
$$\text{Stock price} = S(T),\ \text{Strike price} = K$$

$$\text{Payoff } = \max(S(T) - K, 0) \text{ if } \max(S(t)) < H$$

$$G(X) = e^{-rT} \max(S(T) - K, 0) * 1_{\{\max(S(t)) < H\}}$$

AOL Jan 2001 Options:

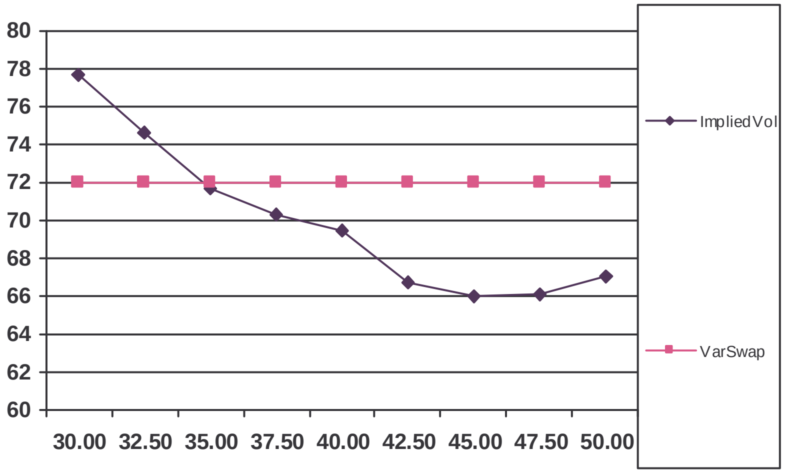
Implied volatilities on Dec 20,2000
Market close



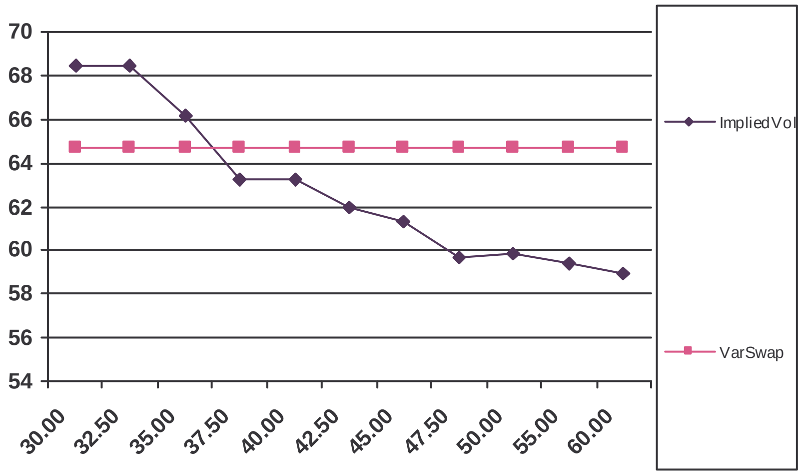
	31.3	32.5	33.8	35	37.5	40	41.3	42.5	43.8	45	46.3
ImpliedVol	96.6191	94.5071	88.4581	83.9929	81.7033	82.5468	81.4319	80.1212	78.6667	80.7064	78.8035
VarSwap	87.1215	87.1215	87.1215	87.1215	87.1215	87.1215	87.1215	87.1215	87.1215	87.1215	87.1215

Pricing probability is not lognormal

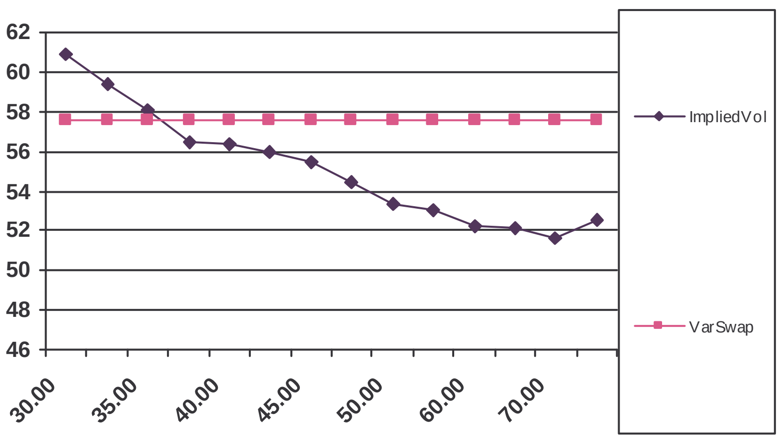
Expiration
2/17/01



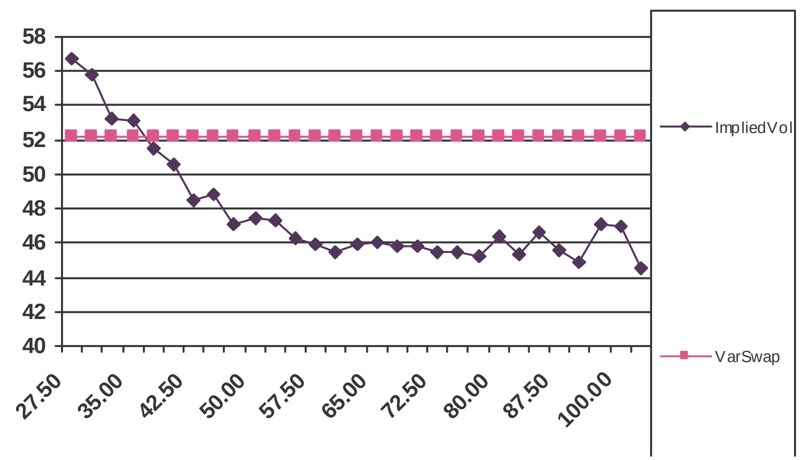
Expiration
4/21/01



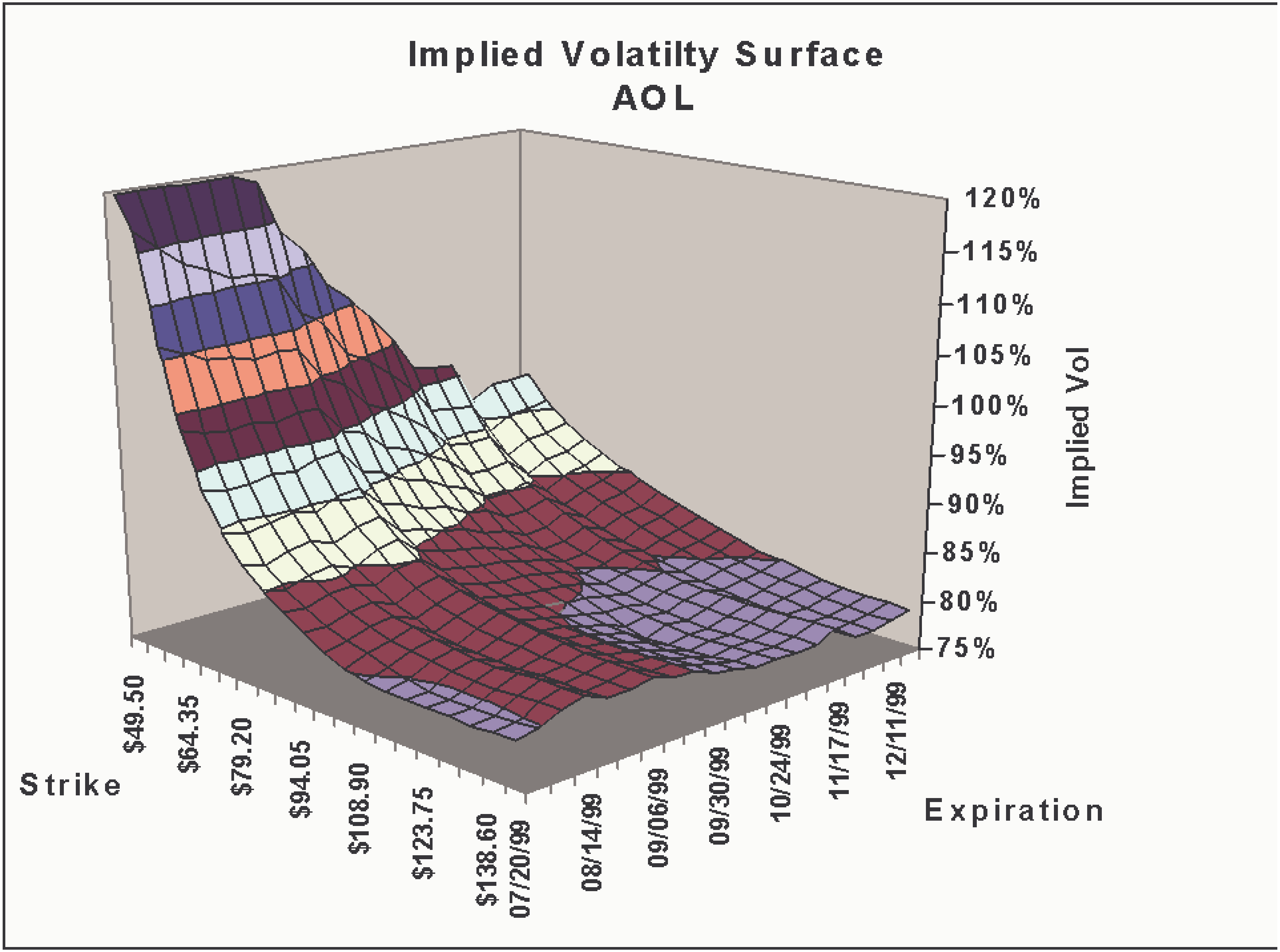
Expiration
7/21/01



Expiration
1/19/02



The AOL ``volatility skews'' for several expiration dates


Implied Volatilty Surface
AOL
120%
115%
110%
105%
100%
95%
90%
85%
80%
75%
Implied Vol
$49.50
$64.35
$79.20
$94.05
$108.90
$123.75
$138.60
Strike
07/20/99
08/14/99
09/06/99
09/30/99
10/24/99
11/17/99
12/11/99
Expiration

Dupire's Local Volatility Function

Breeden & Lizenberger 1978
Dupire 1992
Derman & Kani 1994

$C(K,T) = C^{2,1}$ differentiable function describing call prices

$$\sigma^2(S,t) = \left[\frac{\dfrac{\partial C}{\partial T}}{\dfrac{K^2}{2} \dfrac{\partial^2 C}{\partial K^2}} \right]_{T=t,\,K=S}$$

Different interpolations of $C(K_i,T_j)$ give rise to different local vols
Higher dimensions?

Model Selection Issues

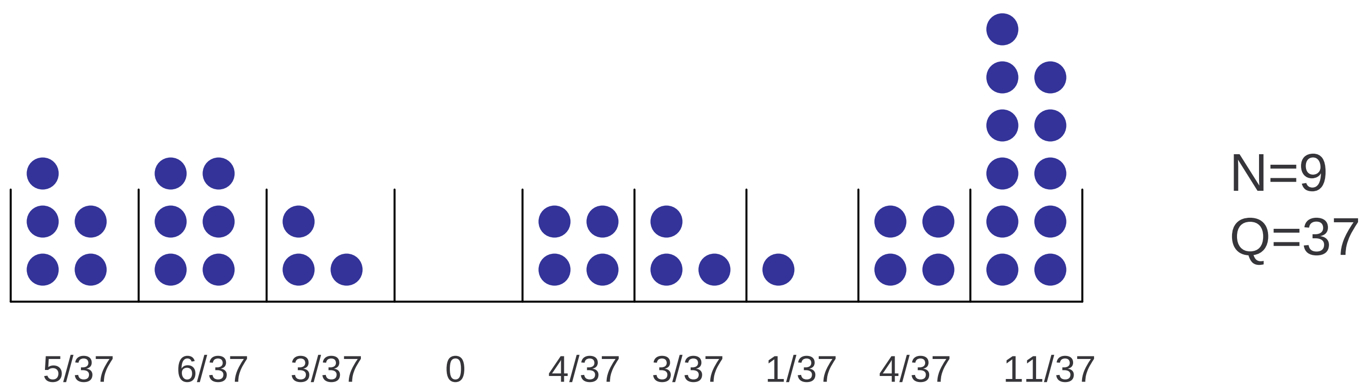
- Different interpolation mechanisms for rate curves/ volatility surfaces give rise to different valuations
- How do we take into account the historical data in conjunction with the choice of model?
- How do we generate stable and easy-to-implement model generation schemes that can be fitted to the prices of many reference derivatives?
- Few parameters (eg. Stochastic volatility) allows to calibrate to a few reference instruments; many parameters (local volatility surfaces) lead to ill-posed problems
- Curse of dimensionality: how can we write and calibrate models with many underlying assets (bespoke CDO tranches, multi-asset equity derivatives)?

2. The Principle of Maximum Entropy

Boltzmann's counting argument

N boxes, Q balls (Q>>N)

<u>Configuration</u>: an assignment or mapping of each ball to a box (or "state'')



<u>Counting probability distribution</u> associated with a configuration:

$$p_i = \frac{\text{number of balls in box } i}{\text{N}} \qquad i = 1,..,\text{N}$$

How many configurations give rise to a given probability?

$$p_i = \frac{n_i}{Q}, \qquad \sum_{i=1}^{N} n_i = Q, \qquad i = 1,\dots,N.$$

$$\nu(p_1,\dots,p_N) = \frac{Q!}{n_1!\,n_2!\dots n_N!}$$

Number of configurations consistent with p

$$m! \propto \frac{1}{\sqrt{2\pi}} n^{n+1/2} e^{-n}$$

Stirling's approximation

$$\boxed{\nu(p_1,\dots,p_N) \approx Q \cdot \sum_{i=1}^{N} p_i \ln\left(\frac{1}{p_i}\right) = -Q\left(\sum_{i=1}^{N} p_i \ln p_i\right) \qquad Q >> N}$$

Most likely probability (under constraints)

No constraints:

$$\sum_{i=1}^{N} p_i \ln\left(\frac{1}{p_i}\right) \le \ln N \quad \text{with equality iff} \quad p_i = 1/N$$

M linear moment constraints:

$$\sum_{i=1}^{N} g_{ij} p_i = c_j \qquad j = 1,\dots,M$$

$$\max\left\{ \sum_{i=1}^{N} p_i \ln\left(\frac{1}{p_i}\right) \;\middle|\; \sum_{i=1}^{N} g_{ij} p_i = c_j \qquad j = 1,\dots,M \right\}$$

Dual Method

Solve

$$\min_{\mathbf{p}}\left\{-\sum_{i} p_i \ln p_i + \sum_{j=1}^{M} \lambda_j \left(\sum_{i=1}^{N} p_i g_{ij} - c_j \right) + \lambda_0 \left(\sum_{i=1}^{N} p_i - 1 \right) \right\}$$

$$-\ln p_i - 1 + \sum_{j=1}^{M} \lambda_j g_{ij} + \lambda_0 = 0 \qquad \therefore$$

$$p_i = \frac{1}{Z(\lambda)} \exp\left(\sum_{j=1}^{M} \lambda_j g_{ij} \right), \qquad Z(\lambda) = \sum_{i=1}^{N} \exp\left(\sum_{j=1}^{M} \lambda_j g_{ij} \right)$$

Calibration Problem for Equity Derivatives

Given a group, or collection of stocks, build a stochastic model for the joint evolution of the stocks with the following properties:

- The associated probability measure on market scenarios is risk-neutral: all traded securities are correctly priced by discounting cash-flows
- The associated probability measure is such that stock prices, adjusted for interest and dividends, are martingales (local risk-neutrality)
- The model simulates the joint evolution of ~ 100 stocks
- All options (with reasonable OI), forward prices, on all stocks, must be fitted to the model. Number of constraints ~50 to ~1000 or more
- Efficient calibration, pricing and sensitivity analysis

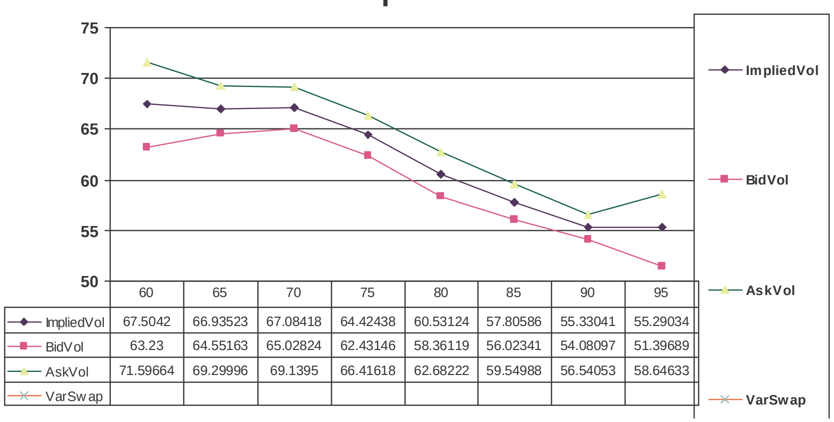
Example: Basket of 20 Biotechnology Stocks (Components of BBH)

Ticker	**Price**	**ATM ImVol**	**Ticker**	**Price**	**ATM ImVol**
ABI	17.85	55	GILD	30.05	46
AFFX	17.19	64	HGSI	16.99	84
ALKS	5.79	106	ICOS	23.62	64
AMGN	44.1	40	IDPH	43.31	72
BGEN	35.36	41	MEDI	27.75	82
CHIR	32.03	37	MLNM	11.8	92
CRA	10.2	55	QLTI	9.36	64
DNA	33.27	53.5	SEPR	6.51	84
ENZN	22.09	81	SHPGY	25.2	47
GENZ	21.66	56	BBH	81.5	32

Implied Volatility Skews

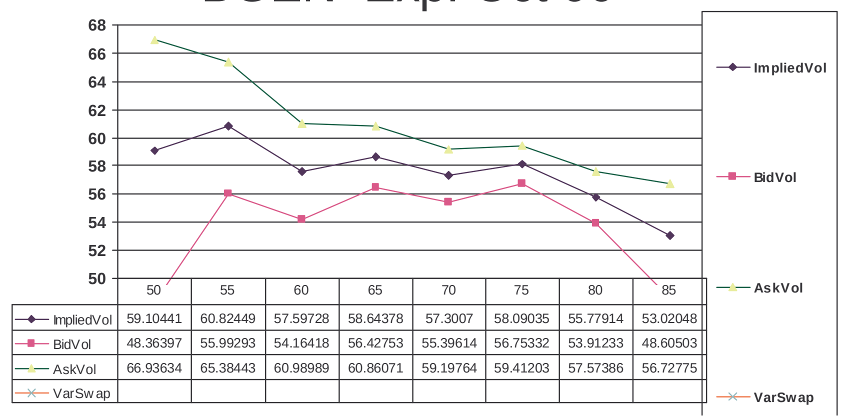
Multiple Names, Multiple Expirations

AMGN Exp: Oct 00



	60	65	70	75	80	85	90	95
ImpliedVol	67.5042	66.93523	67.08418	64.42438	60.53124	57.80586	55.33041	55.29034
BidVol	63.23	64.55163	65.02824	62.43146	58.36119	56.02341	54.08097	51.39689
AskVol	71.59664	69.29996	69.1395	66.41618	62.68222	59.54988	56.54053	58.64633
VarSwap								

BGEN Exp: Oct 00



	50	55	60	65	70	75	80	85
ImpliedVol	59.10441	60.82449	57.59728	58.64378	57.3007	58.09035	55.77914	53.02048
BidVol	48.36397	55.99293	54.16418	56.42753	55.39614	56.75332	53.91233	48.60503
AskVol	66.93634	65.38443	60.98989	60.86071	59.19764	59.41203	57.57386	56.72775
VarSwap								

MEDI Exp: Dec 00

	53.4	56.6	58.4	60	65	70	75	80	85	90
ImpliedVol	74.2145	73.3906	71.4854	68.4688	68.7068	64.2811	65.1807	64.3257	62.4619	63.1047
BidVol	57.9276	60.658	59.7874	57.4033	62.3039	58.7537	62.3079	59.6994	59.4478	60.8186
AskVol	0	0	81.818	78.353	74.8939	69.7241	68.0496	68.9602	65.4771	65.3854
VarSwap										

Needed:

- 20-dimensional stochastic process
- fits option data (multiple expirations)
- martingale property

Multi-Dimensional Diffusion Model

$$\frac{dS_i}{S_i} = \sigma_i dZ_i + \mu_i dt \qquad \mu_i = r - d_i$$

ensures martingale property

$$dZ_i = \text{Brownian motion increment}$$

$$E\left(dZ_i dZ_j\right) = \rho_{ij} dt$$

1-Dimensional Problems

Dupire: local volatility as a function of stock price $\sigma = \sigma(S,t)$

Hull-White, Heston: more factors to model stochastic volatility

Rubinstein, Derman-Kani: implied ``trees''

These methods do not generalize to higher dimensions.

They are ``rigid'' in terms of the modeling assumptions that can be made.

Main Challenges in Multi-Asset Models

- Modeling correlation, or co-movement of many assets
- Correlation may have to match market prices if index options are used as price inputs (time-dependence)
- Fitting single-asset implied volatilities which are time- and strike-dependent
- Large body of literature on 1-D models, but much less is known on intertemporal multi-asset pricing models

Beware of ``magic fixes'', e.g. Copulas

Weighted Monte Carlo

Avellaneda, Buff, Friedman, Grandchamp, Kruk: IJTAF 1999

- Build a discrete-time, multidimensional process for the asset price
- Generate many scenarios for the process by Monte Carlo Simulation
- Fit all price constraints using a Maximum-Entropy algorithm

MC with Non-Uniform Probabilities

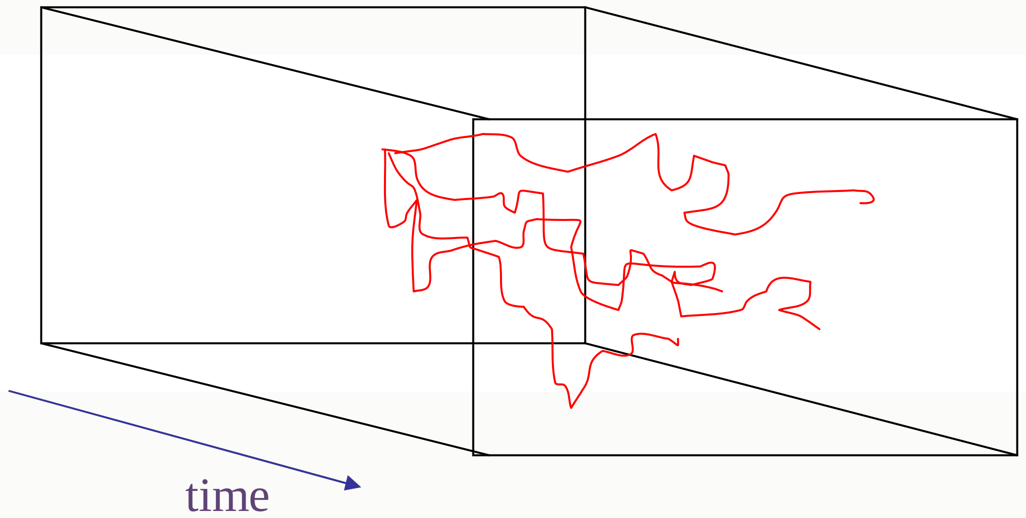
Avellaneda, Buff, Friedman, Kruk, Grandchamp: IJTAF, 1999

- SDE is used to sample the path space

$$dX = \Sigma \cdot dW + B \cdot dt$$

- SDE represents Bayesian prior, *e.g.* subjective probability

- Reweighted probabilities reflect prices of traded securities - Arrow-Debreu probabilities



MC with Non-Uniform Probabilities

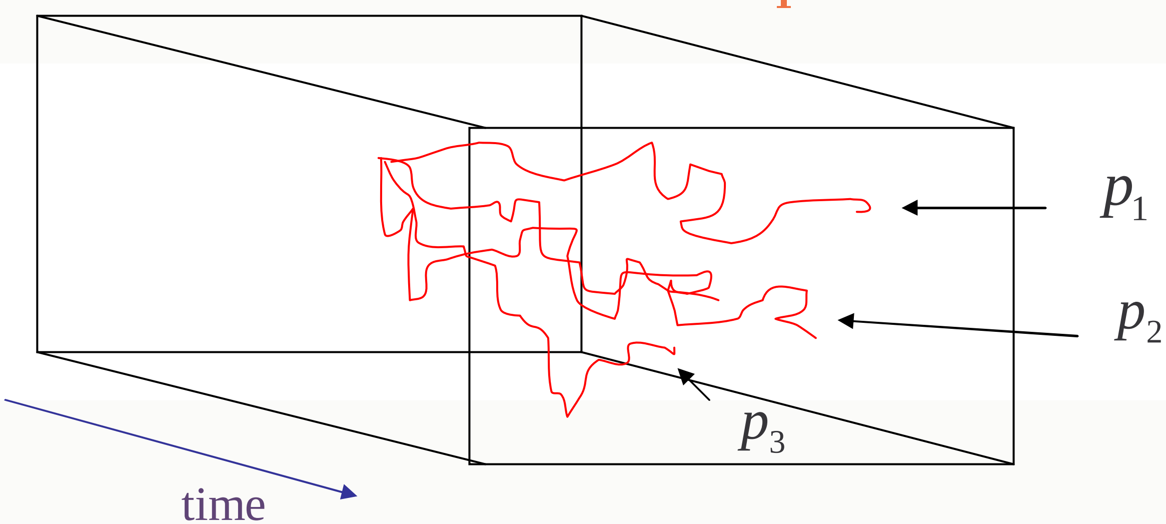
Avellaneda, Buff, Friedman, Kruk, Grandchamp: IJTAF, 1999

- SDE is used to sample the path space

$$dX = \Sigma \cdot dW + B \cdot dt$$

- SDE represents Bayesian prior, *e.g.* subjective probability

- Reweighted probabilities reflect prices of traded securities - Arrow-Debreu probabilities



Example 1: Discrete-Time Multidimensional Markov Process

Modeled after a diffusion

$$S_{n+1}^{(i)} = S_n^{(i)} \cdot \left[1 + \sigma_n^{(i)} \left(\sum_{j=1}^{N} \alpha_{ij} \xi_{n,j} \right) \sqrt{\Delta t} + \mu_n^{(i)} \Delta t \right]$$

$$\xi_{n,j} = \text{i.i.d. normals}$$

- Correlations estimated from econometric analysis
- Vols are ATM implied or estimated from data
- Time-dependence, seasonality effects, can be incorporated

Example 2: Multidimensional Resampling

$$S_{ni} = \text{historical data matrix} \qquad n \le \nu \ \left(\text{sample size}\right)$$

$$X_{ni} = \frac{S_{ni} - S_{(n-1)i}}{S_{(n-1)i}} \qquad Y_{ni} = \frac{X_{ni}}{\sqrt{\sum_{m=1}^{\nu}\left(X_{mi} - \overline{X}_i\right)^2}}$$

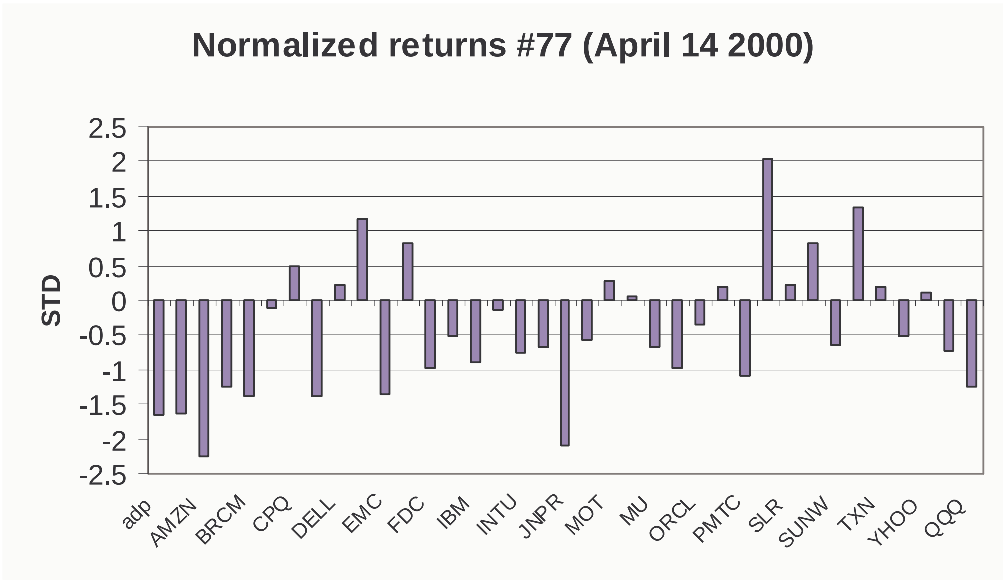
Use resampled standardized moves to generate scenarios

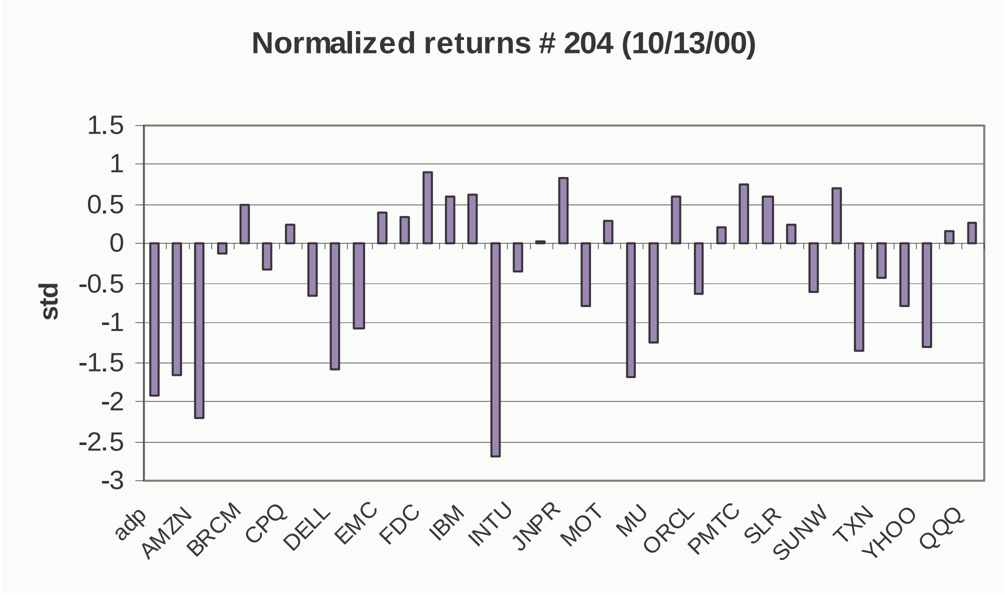
$$S_{n+1}^{(i)} = S_n^{(i)} \cdot \left[1 + \sigma_n^{(i)} Y_{R(n),i} \sqrt{\Delta t} + \mu_n^{(i)} \Delta t\right]$$

$$R(n) = \text{ random number between 1 and } \nu$$

R(n) can be uniform or have temporal correlation

Two draws from the empirical distribution (12/99-12/00)





Simulation consists of sequence of random draws from standardized empirical distribution

Calibration to Option and Forward Prices

- Evaluate Discounted Payoffs of reference instruments along different paths

$$g_{ij} = e^{-rT_j} \max\left(S_{i,T_j}^{a_j} - K_j, 0 \right)$$

$$i = 1, \ldots, N \quad \text{(number of simulated paths)}$$

$$j = 1, \ldots, M \quad \text{(number of reference instruments)}$$

$$C_j = \text{midmarket price of } j^{th} \text{ reference instrument}$$

- Solve

$$\begin{pmatrix} C_1 \\ \ldots \\ C_M \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} & \ldots & \ldots & g_{1N} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ g_{M1} & \ldots & \ldots & \ldots & g_{MN} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \ldots \\ p_N \end{pmatrix}$$

- Repricing condition

$$\boxed{C_j = E^P\left(g_j(S) \right), \qquad j = 1, 2, \ldots, M}$$

Maximum-Entropy Algorithm

$$H(p) = -\sum_{i=1}^{N} p_i \log p_i = -D(p \,\|\, u) \qquad u = \left(\frac{1}{N}, \ldots, \frac{1}{N} \right)$$

Algorithm: solve

$$\max_p H(p) \quad \text{subject to price constraints}$$

$$\min_p D(p \,\|\, u) \qquad \qquad \text{"}$$

Stutzer, 1996; Buchen and Kelly, 1997;
Avellaneda, Friedman, Holmes, Samperi, 1997; Avellaneda 1998
Cont and Tankov, 2002, Laurent and Leisen, 2002,
Follmer and Schweitzer, 1991; Marco Frittelli MEM

Calibrated Probabilities are Gibbs Measures

Lagrange multiplier approach for solving constrained optimization gives rise to M-parameter family of Gibbs-type probabilities

$$p_i = p_i = \frac{1}{Z(\lambda)} \exp\left[\sum_{j=1}^{M} \lambda_j g_{ij}\right], \qquad i = 1,2,\ldots, N$$

Unknown parameters

$$Z(\lambda) = \sum_{i=1}^{N} \exp\left[\sum_{j=1}^{M} \lambda_j g_{ij}\right]$$

Boltzmann-Gibbs partition function

Calibration Algorithm
How do we find the lambdas?

- Minimize in lambda

$$W(\lambda) = \log Z(\lambda) - \sum_{j=1}^{M} \lambda_j C_j$$

- W is a convex function
- The minimum is unique, if it exists
- W is differentiable in C, lambda with <u>explicit gradient</u>
- Use L-BFGS Quasi-Newton gradient-based optimization routine

Boltzmann-Gibbs formalism

$$\frac{\partial W(\lambda)}{\partial \lambda_j} = E^{P_\lambda}\left(G_j(X)\right) - C_j$$

Gradient=difference between
market px and model px

$$\frac{\partial^2 W(\lambda)}{\partial \lambda_j \partial \lambda_k} = E^{P_\lambda}\left(G_j(X) G_k(X)\right) - C_j C_k = Cov^{P_\lambda}\left(G_j(X), G_k(X)\right)$$

Hessian=covariance of
cash-flows under pricing measure

Numerical optimization with known gradient & Hessian also possible

Least-Squares Version

$$\chi^2 = \sum_{j=1}^{M}\left(\sum_{i=1}^{N} g_{ij} p_i - C_j\right)^2 = \sum_{j=1}^{M}\left(E^P\left(g_j(S)\right) - C_j\right)^2$$

$$\min_p \left\{ -H(p) + \frac{\chi^2}{2\varepsilon^2} \right\}$$

Max entropy with least-squares constraint

$$\min_\lambda \left\{ \ln Z(\lambda) + \sum_{j=1}^{M} \lambda_j C_j + \frac{\varepsilon^2}{2} \sum_{j=1}^{M} \lambda_j^2 \right\}$$

Equivalent to adding quadratic term to objective function

Sensitivity Analysis

$$h(X) = \text{payoff function of ``target security''}$$

$$E^{P_\lambda}(h(X)) = \text{model value of} \qquad \text{''}$$

$$\frac{\partial E^{P_\lambda}(h(X))}{\partial C_j} = \frac{\partial E^{P_\lambda}(h(X))}{\partial \lambda_k} \frac{\partial \lambda_k}{\partial C_j}$$

$$= Cov^{P_\lambda}(h(X), g_k(X)) \cdot \left(\frac{\partial C}{\partial \lambda_k}\right)^{-1}_{kj}$$

$$= Cov^{P_\lambda}(h(X), g_k(X)) \cdot \left(Cov^{P_\lambda}(g_\bullet(X), g_\bullet(X))\right)^{-1}_{jk}$$

Price-Sensitivities= Regression Coefficients

Solve LS problem:

$$\min_{\beta,\alpha} \sum_{i=1}^{\nu} p_i \left(h(X_i) - \alpha - \sum_{j=1}^{M} \beta_j G_j(X_i) \right)^2$$

Uncorrelated to gj(X)

$$h(X) = \alpha + \sum_{j=1}^{M} \beta_j g_j(X) + \varepsilon(X)$$

Minimal Martingale Measure?

Michael Fischer, Ph D Thesis, 2003

- Boltzmann-Gibbs posterior measure with price constraints is not a local martingale

- Remedy: include additional constraints:

$$g(S) = \psi\left(S_{t_1}, \ldots, S_{t_N}\right)\left(S_{t_{N+1}} - S_{t_N}\right) \qquad \psi\left(S_{t_1}, \ldots, S_{t_N}\right) = \text{polynomial function}$$

$$\text{Martingale constraint:} \qquad E^P\left(g(S)\right) = 0 \quad \text{for all } \psi$$

- Constrained Max-Entropy problem with martingale constraints: Follmer-Schweitzer MEM under constraints

- In practice, use only low-degree polynomials (deg=0 or deg=1)