Interest Rate Models: Paradigm shifts in recent years

Damiano Brigo
Q-SCI, Managing Director and Global Head
DerivativeFitch, 101 Finsbury Pavement, London

Columbia University Seminar, New York, November 5, 2007

This presentation is based on the book
"Interest Rate Models: Theory and Practice - with Smile, Inflation and Credit"
by D. Brigo and F. Mercurio, Springer-Verlag, 2001 (2nd ed. 2006)
`http://www.damianobrigo.it/book.html`

Overview

- No arbitrage and derivatives pricing.

- Modeling suggested by no-arbitrage discounting.
  1977: Endogenous **short-rate** term structure models

- Reproducing the initial market interest-rate curve exactly.
  1990: Exogenous **short rate** models

- A general framework for no-arbitrage rates dynamics.
  1990: HJM - modeling **instantaneous forward rates**

- Moving closer to the market and consistency with market formulas
  1997: **Fwd market-rates models** calibration and diagnostics power

- 2002: Volatility smile extensions of **Forward market-rates models**

No arbitrage and Risk neutral valuation

Recall shortly the risk-neutral valuation paradigm of Harrison and Pliska's (1983), characterizing no-arbitrage theory:

A future stochastic payoff, built on an underlying fundamental financial asset, paid at a future time T and satisfying some technical conditions, has as unique price at current time t the *risk neutral world* expectation

$$E_t^Q\left[\underbrace{\exp\left(-\int_t^T r_s\ ds\right)}_{\text{Stochastic discount factor}}\text{Payoff}(\text{Asset})_T\right]$$

where r is the risk-free instantaneous discount rate

Risk neutral valuation

$$E_t^Q \left[\exp\left(-\int_t^T r_s \, ds \right) \text{Payoff}(\text{Asset})_T \right]$$

"Risk neutral world" means that all fundamental underlying assets must have as locally deterministic drift rate (expected return) the risk-free interest rate r:

$$\frac{d \, \text{Asset}_t}{\text{Asset}_t} = \boxed{r_t} \, dt + \text{Asset-\%-Volatility}_t \, (0\text{−mean } dt\text{−variance Normal shock under Q})_t$$

Nothing strange at first sight. To value **future unknown** quantities now, we discount at the relevant interest rate and then take **expectation**, and the mean is a reasonable estimate of unknown quantities.

Risk neutral valuation

But what is surprising is that we do not take the mean in the **real world** (statistics, econometrics) but rather in the risk neutral world, since the actual **growth rate of our asset** (e.g. a stock) **in the real world does not enter the price and is replaced by the risk free rate** r.

$$\frac{d\ \text{Asset}_t}{\text{Asset}_t} = \underbrace{\boxed{r_t}}_{\text{risk free rate}} dt + \text{Asset-\%-Volatility}_t\ (0-\text{mean } dt-\text{variance Normal shock under Q})_t$$

Even if two investors do not agree on the expected return of a fundamental asset in the real world, they still agree on the price of derivatives (e.g. options) built on this asset.

Risk neutral valuation

This is one of the reasons for the enormous success of Option pricing theory, and partly for the Nobel award to Black, Scholes and Merton who started it.

According to Stephen Ross (MIT) in the Palgrave Dictionary of Economics:
"... options pricing theory is the most successful theory not only in finance, but in all of economics".

From the risk neutral valuation formula we see that one fundamental quantity is r_t, the instantaneous interest rate. In particular, if we take $\text{Payoff}_T = 1$, we obtain the Zero-Coupon Bond

Zero-coupon Bond, LIBOR rate

A T–**maturity zero–coupon bond** guarantees the payment of one unit of currency at time T. The contract value at time $t < T$ is denoted by $P(t,T)$:

$$P(T,T) = 1, \quad P(t,T) = E_t^Q \left[\exp\left(-\int_t^T r_s \, ds \right) 1 \right] \quad \boxed{\begin{array}{ccc} t & \longleftarrow & T \\ \downarrow & & \downarrow \\ P(t,T) & & 1 \end{array}}$$

All kind of rates can be expressed in terms of zero–coupon bonds and vice-versa. ZCB's can be used as fundamental quantities or building blocks of the interest rate curve.

Zero-coupon Bond, LIBOR and swap rates

- Zero Bond at t for maturity T: $\boxed{P(t,T)} = E_t\left[\exp\left(-\int_t^T \boxed{r_s}\, ds\right)\right]$

- Spot LIBOR rate at t for maturity T: $L(t,T) = \frac{1-P(t,T)}{(T-t)\ P(t,T)}$;

- Fwd Libor at t, expiry T_{i-1} maturity T_i: $F_i(t) := \frac{1}{T_i-T_{i-1}}\left(\frac{P(t,T_{i-1})}{P(t,T_i)} - 1\right)$. This is a market rate, it underlies the Fwd Rate Agreement contracts.

- Swap rate at t with tenor $T_\alpha, T_{\alpha+1}, \ldots, T_\beta$. This is a market rate- it underlies the Interest Rate Swap contracts:

$$S_{\alpha,\beta}(t) := \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{i=\alpha+1}^{\beta}(T_i - T_{i-1})P(t,T_i)} \ .$$

Zero-coupon Bond, LIBOR and swap rates

$L(t, T_j), F_i(t), S_{\alpha,\beta}(t), ...$ these rates at time t, for different maturities $T = T_j, T_{i-1}, T_i, T_\alpha, T_\beta$, are completely known from bonds $T \mapsto P(t, T)$.

Bonds in turn are defined by expectations of functionals of future paths of r_t. So if we know the probabilistic behaviour of r from time t on, we also know the bonds and rates for all maturities at time t:

term structure at time t : $\quad T \mapsto L(t, T)$, initial point $r_t \approx L(t, t+\text{small } \epsilon)$

dr_t future random dynamics $\Rightarrow$ Knowledge of $T \mapsto L(t, T)$ at t;

$T \mapsto L(t, T)$ at $t \not\Rightarrow$ Knowledge of dr_t random future dynamics;

Zero-coupon curve

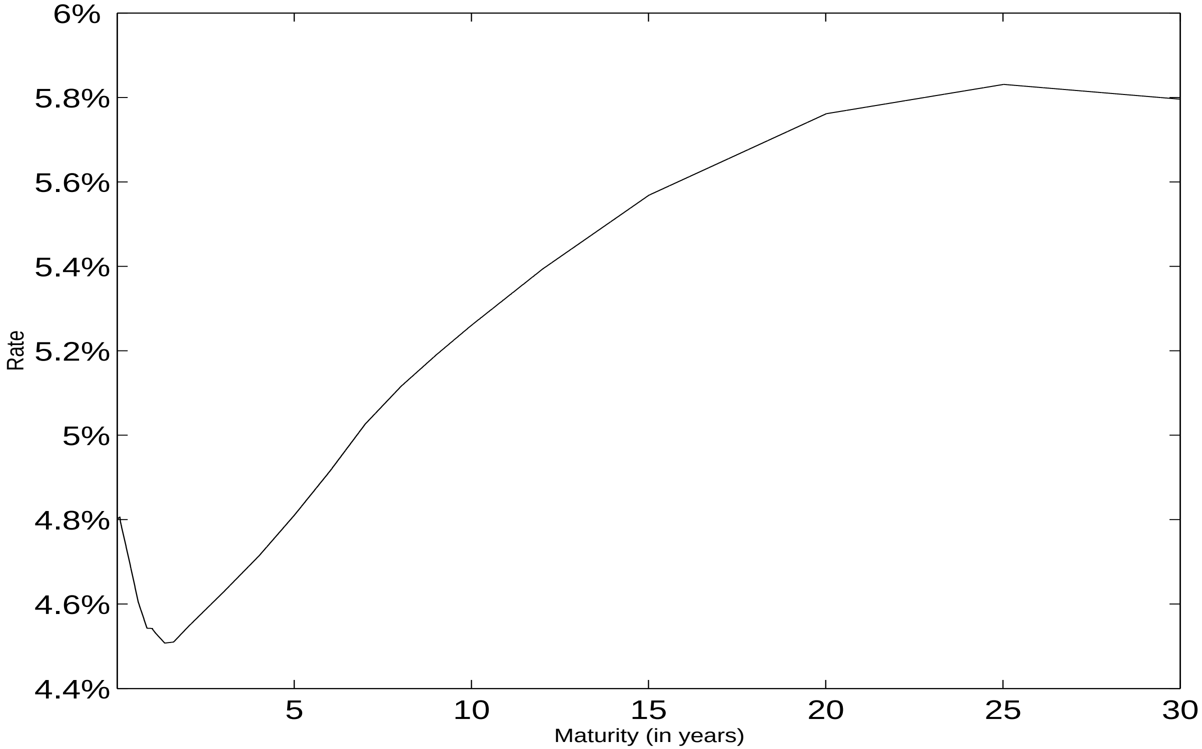

Figure 1: Zero-coupon curve $T \mapsto L(t, t+T)$ stripped from market EURO rates on February 13, 2001, $T = 0, \ldots, 30y$

Which variables do we model?

The curve today does not completely specify how rates will move in the future. For derivatives pricing, we need specifying a stochastic dynamics for interest rates, i.e. choosing an **interest-rate model**.

- Which quantities dynamics do we model? Short rate r_t? LIBOR $L(t,T)$? Forward LIBOR $F_i(t)$? Forward Swap $S_{\alpha,\beta}(t)$? Bond $P(t,T)$?
- How is randomness modeled? i.e: What kind of stochastic process or stochastic differential equation do we select for our model?
- Consequences for valuation of specific products, implementation, goodness of calibration, diagnostics, stability, robustness, etc?

First Choice: short rate r (Vasicek, 1977)

This approach is based on the fact that the zero coupon curve at any instant, or the (informationally equivalent) zero bond curve

$$T \mapsto P(t,T) = E_t^Q \exp\left(- \int_t^T \boxed{\;r_s\;}\, ds \right)$$

is completely characterized by the probabilistic/dynamical properties of r. So we write a model for r, typically a stochastic differential equation

$$d\, r_t = \text{local\_mean}(t, r_t)dt + \text{local\_standard\_deviation}(t, r_t) \times \boxed{\;0 - \text{mean } dt - \text{variance\_normal\_shock}_t\;}$$

which we write $dr_t = \underbrace{b(t, r_t)}_{\text{drift}}\, dt + \underbrace{\sigma(t, r_t)}_{\text{diffusion coeff. or absolute volatility}} \boxed{\; dW_t \;}$.

First Choice: short rate r

Dynamics of r_t under the risk–neutral-world probability measure

1. **Vasicek (1977):** $dr_t = k(\theta - r_t)dt + \sigma dW_t, \quad \alpha = (k, \theta, \sigma)$.

2. **Cox-Ingersoll-Ross (CIR, 1985):**

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t, \quad \alpha = (k, \theta, \sigma), \quad 2k\theta > \sigma^2 \ .$$

3. **Dothan / Rendleman and Bartter:**

$$dr_t = ar_t dt + \sigma r_t dW_t, \ (r_t = r_0 \ e^{(a-\frac{1}{2}\sigma^2)t+\sigma W_t}, \ \alpha = (a, \sigma)).$$

4. **Exponential Vasicek:** $r_t = \exp(z_t), \quad dz_t = k(\theta - z_t)dt + \sigma dW_t$.

First Choice: short rate r. Example: Vasicek

$$dr_t = k(\theta - r_t)dt + \sigma dW_t.$$

The equation is linear and can be solved explicitly: Good.

Joint distributions of many important quantities are Gaussian. Many formula for prices (i.e. expectations): Good.

The model is mean reverting: The expected value of the short rate tends to a constant value θ with velocity depending on k as time grows, while its variance does not explode. Good also for risk management, rating.

Rates can assume negative values with positive probability. Bad.

Gaussian distributions for the rates are not compatible with the market implied distributions.

First Choice: short rate r. Example: CIR

$$dr_t = k[\theta - r_t]dt + \sigma \boxed{\sqrt{r_t}} dW(t)$$

For the parameters k, θ and σ ranging in a reasonable region, this model implies **positive** interest rates, but the instantaneous rate is characterized by a **noncentral chi-squared distribution**. The model is mean reverting as Vasicek's.

This model maintains a certain degree of analytical tractability, but is **less tractable** than Vasicek

CIR is closer to market implied distributions of rates (fatter tails).

Therefore, the CIR dynamics has both some advantages and disadvantages with respect to the Vasicek model.

Similar comparisons affect lognormal models, that however lose all tractability.

First Choice: Modeling r. Endogenous models.

Model	**Distrib**	**Analytic** $P(t,T)$	**Analytic Options**	**Multifactor** Extensions	Mean Reversion	$r > 0$?
Vasicek	Gaussian	Yes	Yes	Yes	Yes	No
CIR	n.c. χ^2, Gaussian^2	Yes	Yes	Yes but..	Yes	Yes
Dothan	e^{Gaussian}	"Yes"	No	Difficult	"Yes"	Yes
Exp. Vasicek	e^{Gaussian}	No	No	Difficult	Yes	Yes

These models are **endogenous**. $P(t,T) = E_t(e^{-\int_t^T r(s)ds})$ can be computed as an expression (or numerically in the last two) depending on the model parameters.

E.g. in Vasicek, at $t = 0$, the interest rate curve is an **output** of the model, rather than an input, depending on k, θ, σ, r_0 in the dynamics.

First Choice: Modeling r. Endogenous models.

Given the observed curve $T \mapsto P^{\text{Market}}(0,T)$, we wish our model to incorporate this curve. Then we need forcing the model parameters to produce a curve as close as possible to the market curve. This is the **calibration of the model to market data**.

k, θ, σ, r_0?: $T \mapsto P^{\text{Vasicek}}(0,T;k,\theta,\sigma,r_0)$ is closest to $T \mapsto P^{\text{Market}}(0,T)$

Too few parameters. Some shapes of $T \mapsto L^{\text{Market}}(0,T)$ (like an inverted shape in the picture above) can never be obtained.

To improve this situation and calibrate also option data, **exogenous** term structure models are usually considered.

First Choice: Modeling r. Exogenous models.

The basic strategy that is used to transform an endogenous model into an exogenous model is the inclusion of “time-varying” parameters. Typically, in the Vasicek case, one does the following:

$$dr(t) = k[\theta - r(t)]dt + \sigma dW(t) \longrightarrow dr(t) = k[\boxed{\vartheta(t)} - r(t)]dt + \sigma dW(t) \; .$$

$\vartheta(t)$ can be defined in terms of $T \mapsto L^{\text{Market}}(0,T)$ in such a way that the model reproduces exactly the curve itself at time 0.

The remaining parameters may be used to calibrate option data.

1. **Ho-Lee:** $dr_t = \theta(t)\ dt + \sigma\ dW_t$.

2. **Hull-White (Extended Vasicek):** $dr_t = k(\theta(t) - \ r_t)dt + \sigma dW_t$.

3. **Black-Derman-Toy (Extended Dothan):** $r_t = r_0\ e^{u(t)+\sigma(t)W_t}$

4. **Black-Karasinski (Extended exponential Vasicek):**

$$r_t = \exp(z_t), \quad dz_t = k\left[\theta(t) - z_t\right]dt + \sigma dW_t.$$

5. **CIR++ (Shifted CIR model, Brigo & Mercurio):**

$$r_t = x_t + \phi(t;\alpha), \quad dx_t = k(\theta - x_t)dt + \sigma\sqrt{x_t}dW_t$$

In general other parameters can be chosen to be time–varying so as to improve fitting of the volatility term–structure (but...)

Extended Model	**Distrib**	**Analytic bond form.**	**Analytic Analytic**	**Multifactor extension**	**Mean Reversion**	$r > 0?$
Ho-Lee	Gaussian	Yes	Yes	Yes	No	No
Hull-White	Gaussian	Yes	Yes	Yes	Yes	No
BDT	e^{Gaussian}	No	No	difficult	Yes	Yes
BK	e^{Gaussian}	No	No	difficult	Yes	Yes
CIR++	s.n.c. χ^2 $\approx$ Gaussian^2	Yes	Yes	Yes	Yes	Yes but..

Tractable models are more suited to risk management thanks to computational ease and are still used by many firms

Pricing models need to be more precise in the distribution properties so lognormal models were usually preferred (BDT, BK). For pricing these have been supplanted in large extent by our next choices below.

First choice: Modeling r. Multidimensional models

In these models, typically (e.g. shifted two-factor Vasicek)

$$\begin{aligned} dx_t &= k_x(\theta_x - x_t)dt + \sigma_x dW_1(t), \\ dy_t &= k_y(\theta_y - y_t)dt + \sigma_y dW_2(t), \quad dW_1\, dW_2 = \rho\, dt, \\ r_t &= x_t + y_t + \phi(t,\alpha), \quad \alpha = (k_x, \theta_x, \sigma_x, x_0, k_y, \theta_y, \sigma_y, y_0) \end{aligned}$$

More parameters, can capture more flexible options structures and especially give **less correlated** rates at future times: 1-dimens models have corr$(L(1y,2y), L(1y,30y)) = 1$, due to the single random shock dW. Here we may play with ρ in the **two** sources of randomness W_1 and W_2.

We may retain analytical tractability with this Gaussian version, whereas Lognormal or CIR can give troubles.

First choice: Modeling r. Numerical methods

Monte Carlo simulation is the method for payoffs that are path dependent (range accrual, trigger swaps...): it goes forward in time, so at any time it knows the whole past history but not the future.

Finite differences or recombining lattices/trees: this is the method for early exercise products like american or bermudan options. It goes back in time and at each point in time knows what will happen in the future but not in the past.

Monte carlo with approximations of the future behaviour regressed on the present (Least Squared Monte Carlo) is the method for products both path dependent and early exercise (e.g. multi callable range accrual).

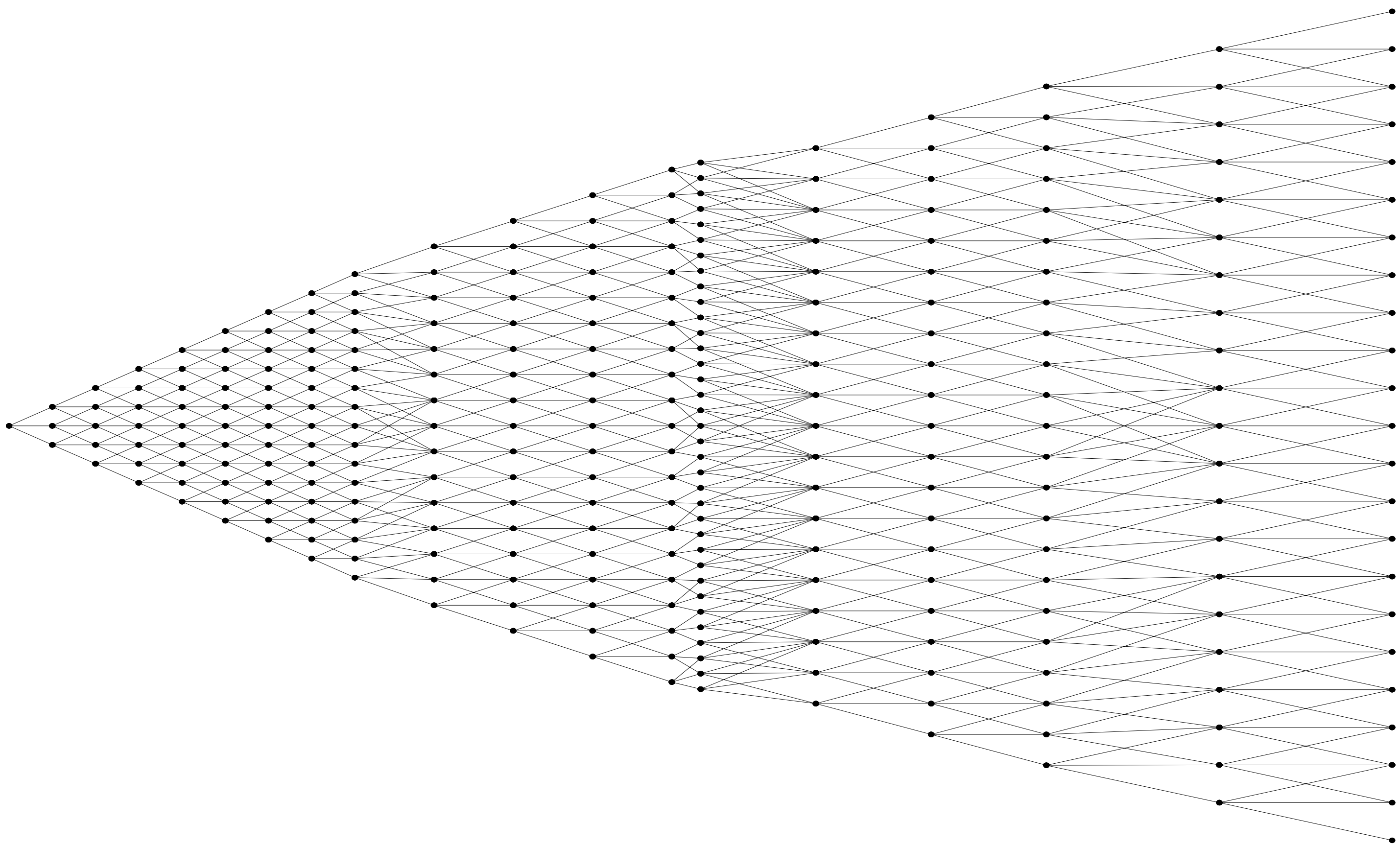

Figure 2: A possible geometry for the discrete-space discrete-time tree approximating r.

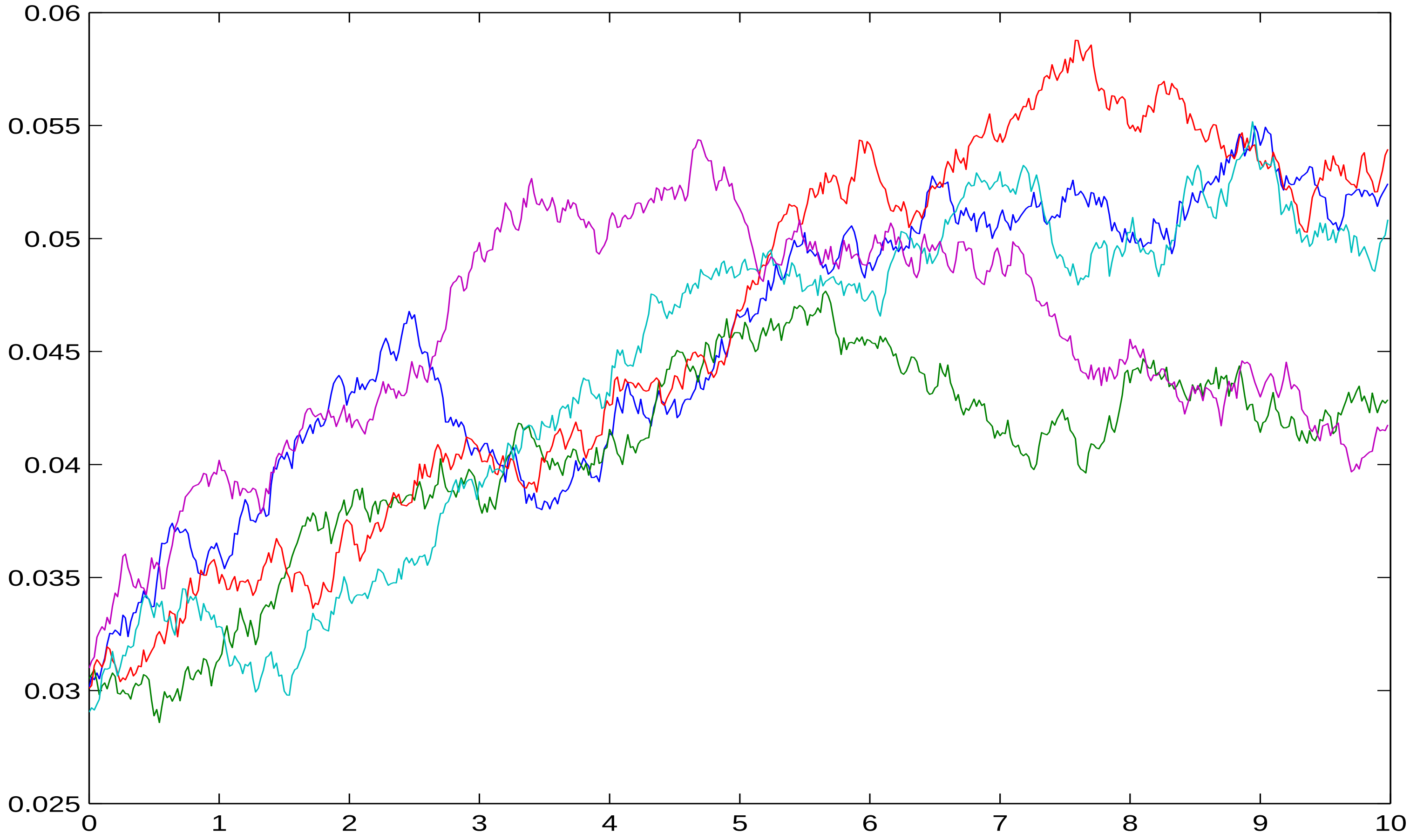

Figure 3: Five paths for the monte carlo simulation of r

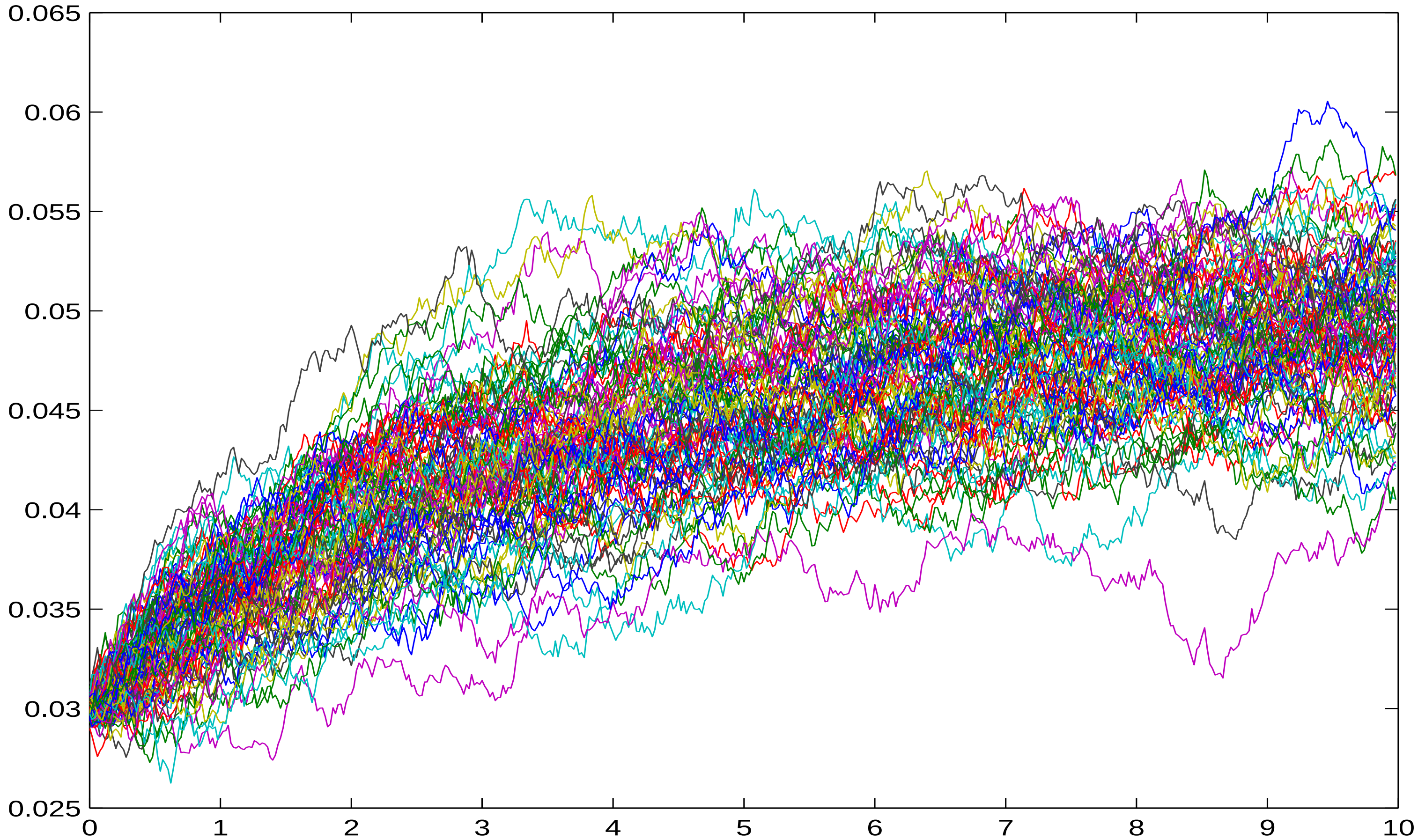

Figure 4: 100 paths for the monte carlo simulation of r

2nd Choice: Modeling inst forward rates $f(t,T)$ (Heath Jarrow Morton, 1990)

Recall the market-based forward LIBOR at time t between T and S, $F(t;T,S) = (P(t,T)/P(t,S) - 1)/(S-T)$. When S collapses to T we obtain *instantaneous* forward rates:

$$f(t,T) = \lim_{S \to T^+} F(t;T,S) = -\frac{\partial \ln P(t,T)}{\partial T}, \quad \lim_{T \to t} f(t,T) = r_t.$$

Why should one be willing to model the f's? The f's are not observed in the market, so that there is no improvement with respect to modeling r in this respect. Moreover notice that f's are more structured quantities:

$$f(t,T) = -\frac{\partial \ln E_t \left[\exp\left(-\int_t^T \boxed{\; r(s) \;} ds \right) \right]}{\partial T}$$

Second Choice: Modeling inst forward rates $f(t, T)$

Given the rich structure in r, we expect restrictions on the dynamics that are allowed for f. Indeed, there is a fundamental theoretical result due to Heath, Jarrow and Morton (HJM, 1990): Set $f(0, T) = f^{\text{Market}}(0, T)$. We have

$$df(t, T) = \boxed{\sigma(t, T)\left(\int_t^T \sigma(t, s)ds\right)} dt + \sigma(t, T)dW(t),$$

under the risk neutral world measure, if no arbitrage has to hold. Thus we find that the no-arbitrage property of interest rates dynamics is here clearly expressed as a **link between the local standard deviation (volatility or diffusion coefficient) and the local mean (drift)** in the dynamics.

Second Choice: Modeling inst forward rates $f(t,T)$

$$df(t,T) = \boxed{\sigma(t,T)\left(\int_t^T \sigma(t,s)ds\right)} dt + \sigma(t,T)dW(t),$$

Given the volatility, there is no freedom in selecting the drift, contrary to the more fundamental models based on dr_t, where the whole risk neutral dynamics was free:

$$dr_t = b(t,r_t)dt + \sigma(t,r_t)dW_t$$

b and σ had no link due to no-arbitrage.

Second Choice, modeling f (HJM): retrospective

Useful to study arbitrage free properties, but when in need of writing a concrete model, most useful models coming out of HJM are the already known short rate models seen earlier (these are particular HJM models, especially exogenous Gaussian models) or the market models we see later.

r model	$\rightarrow$ HJM $\rightarrow$	Market models (LIBOR and SWAP)
Risk Management,	???	Accurate Pricing, Hedging
Rating, easy pricing	???	Accurate Calibration, vol Smile

HJM may serve as a unifying framework in which all categories of no-arbitrage interest-rate models can be expressed.

Third choice: Modeling Market rates. The LIBOR amd SWAP MARKET MODELS (1997)

Recall the forward LIBOR rate at time t between T_{i-1} and T_i,

$$F_i(t) = (P(t, T_{i-1})/P(t, T_i) - 1)/(T_i - T_{i-1}),$$

associated to the relevant Forward Rate Agreement. **A family of F_i with spanning i is modeled in the LIBOR market model instead of r or f.**

To further motivate market models, let us consider the time-0 price of a T_2-maturity libor rate call option - caplet - resetting at time T_1 $(0 < T_1 < T_2)$ with strike X. Let τ denote the year fraction between T_1 and T_2. Such a contract pays out at time T_2 the amount

$$\tau(L(T_1, T_2) - X)^+ = \tau(F_2(T_1) - X)^+.$$

Third Choice: Market models.

$$E\left[\underbrace{\exp\left(-\int_0^{T_2} r_s ds\right)}\tau\underbrace{(L(T_1,T_2)-X)^+}\right] = P(0,T_2)E^{Q^2}[\tau(F_2(T_1)-X)^+],$$

Discount from 2 years Call option on Libor(1year,2years)

(change to measure Q^2 associated to numeraire $P(t,T_2)$, leading to a driftless F_2) with a **lognormal distribution** for F:

$$dF_2(t) = \boxed{v}\, F_2(t)dW_2(t), \quad \text{mkt } F_2(0)$$

where v is the instantaneous volatility, and W_2 is a standard Brownian motion under the measure Q^2.

Third Choice: Market models.

Then the expectation $E^{Q^2}[(F_2(T_1) - X)^+]$ can be viewed as a T_1-maturity call-option price with strike X and underlying volatility v.

The zero-curve $T \mapsto L(0,T)$ is calibrated through the market $F_i(0)$'s. We obtain from **lognormality of F**:

$$\mathrm{Cpl}(0, T_1, T_2, X) := \; P(0, T_2)\; \tau\; E(F_2(T_1) - X)^+$$

$$= P(0,T_2)\tau\ [\ F_2(0)\ N(d_1(X, F_2(0), v\sqrt{T_1}))\ -\ X\ N(d_2(X, F_2(0), v\sqrt{T_1}))],$$

$$d_{1,2}(X, F, u) = \frac{\ln(F/X) \pm u^2/2}{u},$$

where N is the standard Gaussian cumulative distribution function. This is the Black formula that has been used in the market for years to convert prices Cpl in volatilities v and vice-versa.

Third Choice: Market models.

The LIBOR market model is thus compatible with Black's market formula and indeed **prior to this model there was no rigorous arbitrage free justification of the formula**, a building block for all the interest rate options market.

A similar justification for the market swaption formula is obtained through the swap market model, where forward swap rates $S_{\alpha,\beta}(t)$ are modeled as lognormal processes each under a convenient measure.

LIBOR and SWAP market models are inconsistent in theory but consistent in practice.

Third Choice: Market models.

The quantities in the LIBOR market models are forward rates coming from expectations of objects involving r, and thus are structured.

As for HJM, we may expect restrictions when we write their dynamics.

We have already seen that each F_i has no drift (local mean) under its forward measure. Under other measures, like for example the (discrete tenor analogous of the) risk neutral measure, the F_i have local means coming from the volatilities and correlations of the family of forward rates F being modeled.

Third Choice: Market models.

This is similar to HJM

$$dF_k(t) = \underbrace{\boxed{\sigma_k(t)}\, F_k(t) \sum_{j:T_j>t}^{k} \frac{\tau_j \boxed{\rho_{j,k}\,\sigma_j(t)}\, F_j(t)}{1+\tau_j F_j(t)}}_{\text{local mean or drift}} dt + \boxed{\sigma_k(t)}\, F_k(t)\, dZ_k(t)$$

$$''volatility''(dF_j(t)) = \boxed{\sigma_j(t)}, \quad ''correlation''(dF_i, dF_j) = \boxed{\rho_{i,j}}.$$

Third Choice: Market models.

$$''volatility''(dF_j(t)) = \boxed{\sigma_j(t)}, \quad ''correlation''(dF_i, dF_j) = \boxed{\rho_{i,j}}.$$

This direct modeling of vol and correlation of the movement of real market rates rather than of abstract rates like r or f is one of the reasons of the success of market models.

Vols and correlations refer to objects the trader is familiar with.

On the contrary the trader may have difficulties in translating a dynamics of r in facts referring directly to the market. Questions like "what is the volatility of the 2y3y rate" or "what is the correlation between the 2y3y and the 9y10y rates" are more difficult with r models.

Third Choice: Market models.

Furthermore r models are not consistent with the market standard formulas for options.

Also, market models allow easy diagnostics and extrapolation of the volatility future term structure and of terminal correlations that are quite difficult to obtain with r models.

Not to mention that the abundance of clearly interpretable parameters makes market models able to calibrate a large amount of option data, a task impossible with reasonable short rate r models.

The model can easily account for 130 swaptions volatilities consistently.

The most recent paradigm shift: Smile Modeling

In recent years, after influencing the FX and equity markets, the volatility smile effect has entered also the interest rate market.

The volatility smile affects directly volatilities associated with option contracts such as caplets and swaptions, **and is expressed in terms of market rates volatilities**. For this reason, models addressing it are more naturally market models than r models or HJM.

Smile Modeling: Introduction

The option market is largely built around Black's formula. Recall the T_2-maturity caplet resetting at T_1 with strike K. Under the lognormal LIBOR model, its underlying forward follows

$$dF_2(t) = \sigma_2(t) F_2(t) dW_2(t)$$

with deterministic time dependent instantaneous volatility σ_2 not depending on the level F_2. Then we have Black's formula

$$\mathsf{Cpl}^{\mathsf{Black}}(0, T_1, T_2, K) = P(0, T_2)\tau \mathsf{Black}(K, F_2(0), v_2(T_1)), \quad v_2(T_1)^2 = \frac{1}{T_1}\int_0^{T_1} \sigma_2^2(t)dt\,.$$

Volatility is a characteristic of the **underlying** and not of the contract. Therefore the averaged volatility $v_2(T_1)$ should not depend on K.

Smile Modeling: The problem

Now take two different strikes K_1 and K_2 for market quoted options. Does there exist a *single* volatility $v_2(T_1)$ such that both

$$\mathrm{Cpl}^{\mathrm{MKT}}(0, T_1, T_2, K_1) = P(0, T_2)\tau \mathrm{Black}(K_1, F_2(0), v_2(T_1))$$
$$\mathrm{Cpl}^{\mathrm{MKT}}(0, T_1, T_2, K_2) = P(0, T_2)\tau \mathrm{Black}(K_2, F_2(0), v_2(T_1))$$

hold? The answer is a resounding "no". Two *different* volatilities $v_2(T_1, K_1)$ and $v_2(T_1, K_2)$ are required to match the observed market prices if one is to use Black's formula

Smile Modeling: The problem

$$\mathsf{Cpl}^{\mathsf{MKT}}(0, T_1, T_2, K_1) = P(0, T_2)\tau\mathsf{Black}(K_1, F_2(0), v_2^{\mathsf{MKT}}(T_1, K_1)),$$
$$\mathsf{Cpl}^{\mathsf{MKT}}(0, T_1, T_2, K_2) = P(0, T_2)\tau\mathsf{Black}(K_2, F_2(0), v_2^{\mathsf{MKT}}(T_1, K_2)).$$

Each caplet market price requires its own Black volatility $v_2^{\mathsf{MKT}}(T_1, K)$ depending on the caplet strike K.

The market therefore uses Black's formula simply as a metric to express caplet prices as volatilities, but the dynamic and probabilistic assumptions behind the formula, i.e. $dF_2(t) = \sigma_2(t)F_2(t)dW_2(t)$, do not hold.

The typically smiley shape curve $K \mapsto v_2^{\mathsf{MKT}}(T_1, K)$ is called the volatility smile.

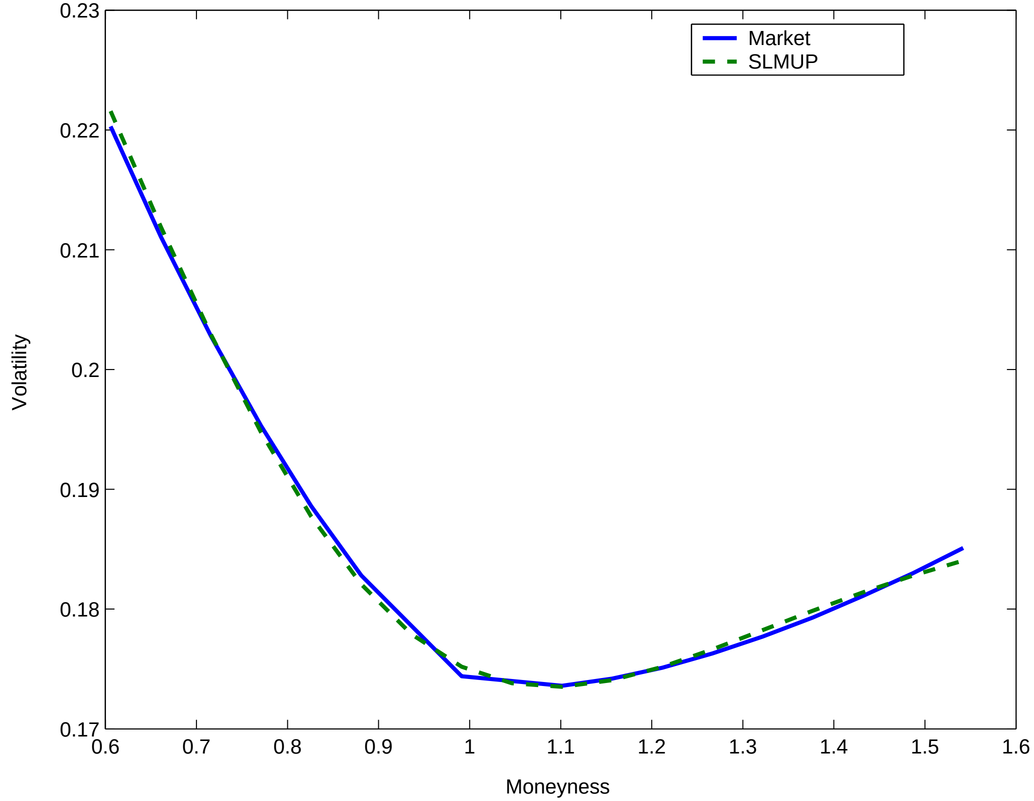

Figure 5: Example of smile with $T_1 = 5y,\ T_2 = 5y6m$

Smile Modeling: The "solutions"

We need to postulate a new dynamics beyond the lognormal one. Mostly there are two solutions.

Local volatility models. Make σ_2 a function of the underlying, such as $\sigma_2(t, F_2(t)) = a\sqrt{F_2(t)}$ (CEV), $\sigma_2(t, F_2(t)) = a(t)(F_2(t) - \alpha)$ (Displaced diffusion) or other more complex and flexible solutions like the mixture diffusion.

Local volatility models are believed to imply a volatility smile that flattens in time: no new randomness is added into the system as time moves on, all randomness in the volatility coming from the underlying F_2.

Smile Modeling: The ”solutions”

Stochastic volatility models. Make σ_2 a new stochastic process, adding new randomness to the volatility, so that in a way volatility becomes a variable with a new random life of its own, possibly correlated with the underlying. Heston's model (1993), adapted to interest rates by Wu and Zhang (2002):)

$$dF_2(t) = \sqrt{v(t)}F_2(t)dW_2(t),$$
$$dv(t) = k(\theta - v(t))dt + \eta\sqrt{v(t)}dZ_2(t), \quad dZdW = \rho dt.$$

Smile Modeling: The "solutions"

A different, more simplistic and popular stochastic volatility model is the so-called SABR (2002):

$$dF_2(t) = b(t)(F_2(t))^{\beta} dW_2(t),$$
$$db(t) = \nu b(t) dZ_2(t), \quad b_0 = \alpha, \quad dZ dW = \rho\, dt$$

with $0 < \beta \leq 1$, ν and α positive. Analytical approximations provide formulas in closed form. When applied to swaptions, this model has been used to quote and interpolate implied volatilities in the swaption market across strikes.

Finally, uncertain volatility models where σ_2 takes at random one among some given values are possible, and lead to models similar to the mixture diffusions in the local volatility case.

Conclusions

1977: short rate models dr_t

1990: HJM $df(t,T)$

1997: Market models $dF_i(t)$, $dS_{\alpha,\beta}(t)$

2002: Volatility smile inclusive Market models $dF_i(t)$, $dS_{\alpha,\beta}(t)$

All these formulations are still operating on different levels, ranging from risk management to rating practice to advanced pricing.

The models have different increasing levels of complexity but in some respects, while having richer parameterization, dF_i market models are more transparent than simpler dr models.

No family of models wins across the whole spectrum of applications and all these models are still needed for different applications.