Interest Rate Modelling and Derivative Pricing

Sebastian Schlenkrich

HU Berlin, Department of Mathematics / FRAME Consulting GmbH, Berlin

Summer term, 2023

Part I

Introduction and Preliminaries

Outline

Outline

What is this lecture about?

Interbank swap deal example

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(annually, 30/360 day count, modified following, Target calendar)

Pays 6-months Euribor floating rate on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(semi-annually, act/360 day count, modified following, Target calendar)

Suppose, Bank A may decide to early terminate deal in 10, 11, 12,.. years

How does early termination option affect the present value and risk of the deal?

Organisational details first

- Lecture: Fri, 13:15 - 14:45 s.t., RUD26, R. 1.304 (plus some additional times)
- Exercises: Fri, 15:00 - 16:30, RUD26, R. 1.304 (every second week, some exceptions)
- Office times: Fridays on request before or after the lecture

Exercises:

- Discuss and analyse practical examples and theory details
- Main tool: QuantLib (open source financial library)
- Implementation: Python, some Excel

Requirements:

- Present at least once during exercises
- exam planned for July 28, 2023

Literature and resources you will need

- Literature
  - L. Andersen and V. Piterbarg. *Interest rate modelling, volume I to III.*
    Atlantic Financial Press, 2010
  - D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice.*
    Springer-Verlag, 2007
  - S. Shreve. *Stochastic Calculus for Finance II - Continuous-Time Models.*
    Springer-Verlag, 2004
- QuantLib web site `www.quantlib.org`
- Official source repository `www.github.com/lballabio`
- Some extensions which we might use `www.github.com/sschlenkrich`
- `https://www.applied-financial-mathematics.de/interest-rate-modelling-summer-term-2023`

Let's revisit the introductory example

Interbank swap deal example

Fixed interest rate

Pays 3% on 100mm EUR

Notional

Start date: Oct 30, 2020

Dates

End date: Oct 30, 2040

Market conventions

(annually, 30/360 day count, modified following, Target calendar)

Stochastic interest rates

Pays 6-months Euribor floating rate on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(semi-annually, act/360 day count, modified following, Target calendar)

Optionalities

Bank A may decide to early terminate deal in 10, 11, 12,.. years

Agenda covers static yield curve modelling, Vanilla rates models and term structure models

Interest Rate Modelling

- Stochastic calculus basics
- Static yield curve modelling and linear products
- Vanilla interest rate models
- HJM term structure modelling framework
- Classical Hull-White interest rate model
- Pricing methods for Bermudan swaptions

Model Calibration

- Multi-curve yield curve calibration
- Hull-White model calibration
- Numerical methods for model calibration

Sensitivity Calculation

- Delta and Vega specification
- Numerical methods for sensitivity calculation

Outline

We will work along three streams

Probability space & filtration	Brownian Motion	Self-financing trading strategy & arbitrage
Change of measure	Ito integral	Equivalent martingale measure & FTAP
Martingale	Martingale representation	Change of equiv. martingale meas.
Density process	Ito's lemma	Permissible trading strategy

Risk-neutral derivative pricing formula

Outline

Measure theory is independent of financial application

Probability space & filtration

Brownian Motion

Self-financing trading strategy & arbitrage

Change of measure

Ito integral

Equivalent martingale measure & FTAP

Martingale

Martingale representation

Change of equiv. martingale meas.

Density process

Ito's lemma

Permissible trading strategy

Risk-neutral derivative pricing formula

We start with stochastic processes and probability space

Stochastic process (for assets or interest rate components)

$$X(t) = [X_1(t), \ldots, X_p(t)]^\top .$$

Probability space that drives stochastic process $(\Omega, \mathcal{F}, \mathbb{P})$

- Ω sample space with outcomes ω (typically increments of Brownian motions),
- $\mathcal{F}$ σ-algebra on Ω,
- $\mathbb{P}$ probability measure on $\mathcal{F}$.

Information flow is realised via filtration $\{\mathcal{F}_t,\ t \in [0, T]\}$

- $\mathcal{F}_t$ sub-algebra of $\mathcal{F}$ with $\mathcal{F}_t \subseteq \mathcal{F}_s$ for $t \leq s$,
- Assume $X(t)$ is adapted to filtration $\mathcal{F}_t$, i.e. $X(t)$ is fully observable at time t.

Measures can be linked by Radon–Nikodym derivative

Theorem (Radon–Nikodym derivative)

Let $\mathbb{P}$ and $\hat{\mathbb{P}}$ be equivalent probability measures on $(\Omega, \mathcal{F})$. Then there exists a unique (a.s.) non-negative random variable $R(\omega)$ with $\mathbb{E}^{\mathbb{P}}[R] = 1$, such that for all $A \in \mathcal{F}$

$$\hat{\mathbb{P}}(A) = \mathbb{E}^{\mathbb{P}}\left[R\,\mathbb{1}_{\{A\}}\right].$$

R is denoted Radon–Nikodym derivative.

It follows

$$\hat{\mathbb{P}}(A) = \int_A d\hat{\mathbb{P}} = \int_A R\,d\mathbb{P} = \mathbb{E}^{\mathbb{P}}\left[R\,\mathbb{1}_{\{A\}}\right].$$

and also for all measurable functions X (via algebraic induction)

$$\mathbb{E}^{\hat{\mathbb{P}}}[X] = \mathbb{E}^{\mathbb{P}}[R\,X].$$

Thus we may write

$$R = d\hat{\mathbb{P}}/d\mathbb{P}.$$

We will frequently need the change of measure for conditional expectations

Definition (Conditional expectation)

Let X be a random variable. The conditional expectation $\mathbb{E}^{\mathbb{P}}[X \mid \mathcal{F}_t]$ is defined as the stochastic variable that satisfies:

- $\mathbb{E}^{\mathbb{P}}[X \mid \mathcal{F}_t]$ is $\mathcal{F}_t$-measurable and
- for all $A \in \mathcal{F}_t$ we have

$$\int_A \mathbb{E}^{\mathbb{P}}[X \mid \mathcal{F}_t]\, d\mathbb{P} = \int_A X d\mathbb{P}.$$

Theorem (Baye's rule for conditional expectation)

Let $R = d\hat{\mathbb{P}}/d\mathbb{P}$ be the Radon–Nikodym derivative associated with $(\Omega, \mathcal{F}, \mathbb{P})$ and $\left(\Omega, \mathcal{F}, \hat{\mathbb{P}}\right)$ and X a random variable. Then

$$\mathbb{E}^{\hat{\mathbb{P}}}[X \mid \mathcal{F}_t] = \frac{\mathbb{E}^{\mathbb{P}}[R\, X \mid \mathcal{F}_t]}{\mathbb{E}^{\mathbb{P}}[R \mid \mathcal{F}_t]}.$$

We sketch the proof for change of measure (1/2)

We use the definition of conditional expectation and show that (for all $A \in \mathcal{F}_t$)

$$\int_A \mathbb{E}^{\mathbb{P}}\left[R\, X \,|\, \mathcal{F}_t\right] d\mathbb{P} = \int_A \mathbb{E}^{\mathbb{P}}\left[R \,|\, \mathcal{F}_t\right] \mathbb{E}^{\hat{\mathbb{P}}}\left[X \,|\, \mathcal{F}_t\right] d\mathbb{P}.$$

We have for the left side using conditional expectation and Radon–Nikodym derivative

$$\int_A \mathbb{E}^{\mathbb{P}}\left[R\, X \,|\, \mathcal{F}_t\right] d\mathbb{P} = \int_A X\, R\, d\mathbb{P} = \int_A X d\hat{\mathbb{P}}.$$

For the right side we get using conditional expectation

$$\begin{aligned}\int_A \mathbb{E}^{\mathbb{P}}\left[R \,|\, \mathcal{F}_t\right] \mathbb{E}^{\hat{\mathbb{P}}}\left[X \,|\, \mathcal{F}_t\right] d\mathbb{P} &= \int_A \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\hat{\mathbb{P}}}\left[X \,|\, \mathcal{F}_t\right] R \,|\, \mathcal{F}_t\right] d\mathbb{P} \\ &= \int_A \mathbb{E}^{\hat{\mathbb{P}}}\left[X \,|\, \mathcal{F}_t\right] R\, d\mathbb{P}.\end{aligned}$$

We sketch the proof for change of measure (2/2)

Applying Radon–Nikodym derivative and again conditional expectation yields

$$\int_A \mathbb{E}^{\hat{\mathbb{P}}}\left[X \mid \mathcal{F}_t\right] R \, d\mathbb{P} = \int_A \mathbb{E}^{\hat{\mathbb{P}}}\left[X \mid \mathcal{F}_t\right] d\hat{\mathbb{P}} = \int_A X d\hat{\mathbb{P}}.$$

We will use Frobenius norm in martingale definition

Sum of squares notation (Frobenius norm, L^2 norm for vectors)

For a matrix or vector $A \in \mathbb{R}^{m \times n}$ with elements $\{a_{i,j}\}_{i,j}$ we denote

$$|A| = \sqrt{\operatorname{tr}(AA^\top)} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j}^2}.$$

Martingales allow derivation of expected future values

Definition (Martingale)

Let $X(t)$ be an adapted vector-valued process with finite absolute expectation $\mathbb{E}^{\mathbb{P}}[|X(t)|] < \infty$ (under the probability measure $\mathbb{P}$) for all $t \in [0, T]$.
$X(t)$ is a martingale under $\mathbb{P}$ if for all $t, s \in [0, T]$ with $t \leq s$

$$X(t) = \mathbb{E}^{\mathbb{P}}[X(s) \mid \mathcal{F}_t] \quad a.s.$$

- Typically, martingale property is derived (by other results) for a process.
- Then we can use martingale property to obtain expectation of future values $X(T)$.

Density process describes change of measure for processes

Definition (Density process)

Denote $\zeta(t) = \mathbb{E}^{\mathbb{P}}\left[d\hat{\mathbb{P}}/d\mathbb{P} \mid \mathcal{F}_t\right]$ the density process of $\hat{\mathbb{P}}$ (relative to $\mathbb{P}$).

- Then $\zeta(t)$ is a $\mathbb{P}$-martingale with $\zeta(0) = \mathbb{E}^{\mathbb{P}}\left[\zeta(t)\right] = 1$.

Lemma (Change of measure for processes)

Let $X(t)$ be a $\mathcal{F}_t$ measurable random variable. Then

$$\mathbb{E}^{\hat{\mathbb{P}}}\left[X(T) \mid \mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\frac{\zeta(T)}{\zeta(t)} X(T) \mid \mathcal{F}_t\right].$$

Proof.

Recall that $R = d\hat{\mathbb{P}}/d\mathbb{P}$. We have $\mathbb{E}^{\hat{\mathbb{P}}}\left[X(T) \mid \mathcal{F}_t\right] = \frac{\mathbb{E}^{\mathbb{P}}[R\, X(T) \mid \mathcal{F}_t]}{\mathbb{E}^{\mathbb{P}}[R \mid \mathcal{F}_t]}$. Then

$$\mathbb{E}^{\mathbb{P}}\left[R\, X(T) \,|\, \mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[R\, X(T) \,|\, \mathcal{F}_T\right] \,|\, \mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[R \,|\, \mathcal{F}_T\right] X(T) \,|\, \mathcal{F}_t\right].$$

The result follows from the definition of $\zeta(t)$ via $\zeta(t) = \mathbb{E}^{\mathbb{P}}\left[R \,|\, \mathcal{F}_t\right]$. □

Density process may be used to define a new measure

Let $\zeta(t)$ be a $\mathbb{P}$-martingale with $\zeta(0) = 1$. We choose a final horizon time T and define the Radon–Nikodym derivative as $R(\omega) = \zeta(T, \omega)$.

The corresponding measure $\hat{\mathbb{P}}$ on $(\Omega, \mathcal{F}_T)$ is

$$\hat{\mathbb{P}}(A) = \mathbb{E}^{\mathbb{P}}\left[R\, 1_{\{A\}}\right] = \mathbb{E}^{\mathbb{P}}\left[\zeta(T, \omega)\, \mathbb{1}_{\{A\}}\right].$$

We show that the density of $\hat{\mathbb{P}}$ indeed equals $\zeta(t)$.

Denote $\bar{\zeta}(t) = \mathbb{E}^{\mathbb{P}}[R \,|\, \mathcal{F}_t]$ the density of $\hat{\mathbb{P}}$. Then we get with the martingale property of $\zeta(t)$

$$\bar{\zeta}(t) = \mathbb{E}^{\mathbb{P}}[\zeta(T, \omega) \,|\, \mathcal{F}_t] = \zeta(t).$$

Outline

Diffusion processes are the basis for our models

- Probability space & filtration
- Brownian Motion
- Self-financing trading strategy & arbitrage
- Change of measure
- Ito integral
- Equivalent martingale measure & FTAP
- Martingale
- Martingale representation
- Change of equiv. martingale meas.
- Density process
- Ito's lemma
- Permissible trading strategy
- Risk-neutral derivative pricing formula

Stochastic process is driven by Brownian motion

Information is generated by Brownian motion

- $W(t) = [W_1(t), \ldots, W_d(t)]^\top$ d-dimensional Brownian motion.
- $W_i(\cdot)$ independent of $W_j(\cdot)$ for $i \neq j$.
- Independent Gaussian increments $W_i(s) - W_i(t) \sim \mathcal{N}(0, s-t)$ for $s \geq t$.
- Typically, filtration $\mathcal{F}_t$ is generated by Brownian motion $W(\cdot)$, i.e. $\mathcal{F}_t = \sigma\{W(u),\, 0 \leq u \leq t\}$.

Definition (H^2 for volatility processes σ)

Let $\sigma : \mathbb{R} \times \Omega \to \mathbb{R}^{p \times d}$ be a volatility process adapted to the filtration generated by $\mathcal{F}_t$. We say that σ is in H^2 if for all $t \in [0, T]$ we have

$$\mathbb{E}^{\mathbb{P}}\left[\int_0^t |\sigma(s,\omega)|^2\, ds\right] < \infty.$$

Stochastic process is described as Ito process with Ito integral

$$X(t) = X(0) + \int_0^t \mu\left(s,\omega\right) ds + \int_0^t \sigma\left(s,\omega\right) dW(s)$$

or in differential notation

$$dX(t) = \mu\left(t,\omega\right) dt + \sigma\left(t,\omega\right) dW(t),$$

- vector-valued drift $\mu : \mathbb{R} \times \Omega \to \mathbb{R}^p$,
- matrix of volatilities $\sigma : \mathbb{R} \times \Omega \to \mathbb{R}^{p\times d}$,
- assume drift μ and volatility σ are adapted to $\mathcal{F}_t$ and σ is in H^2.

We consider the Ito integral as

$$\int_0^t \sigma\left(s,\omega\right) dW(s) = \lim_{n\to\infty} \sum_{i=1}^{n} \sigma\left(s_{i-1},\omega\right)\left[W(s_i) - W(s_{i-1})\right], \quad s_i = \frac{i}{n}t.$$

Ito integrals are important martingales for modelling

Theorem (Ito Integral properties)

Define the Ito integral $X(t) = \int_0^t \sigma(u, \omega)\, dW(u)$ *with* σ *is in* H^2*. Then*

1. $X(t)$ *is* $\mathcal{F}_t$*-measurable (i.e. we can calculate the distribution of* $X(t)$ *using* $(\Omega, \mathcal{F}, \mathbb{P})$*)*

2. $X(t)$ *is a continuous martingale*

3. $\mathbb{E}^{\mathbb{P}}\left[|X(t)|^2\right] = \mathbb{E}^{\mathbb{P}}\left[\int_0^t |\sigma(u, \omega)|^2\, du\right] < \infty$ *(Ito isometry)*

4. $\mathbb{E}^{\mathbb{P}}\left[X(t)X(s)^\top\right] = \mathbb{E}^{\mathbb{P}}\left[\int_0^{\min\{t,s\}} \sigma(u, \omega)\, \sigma(u, \omega)^\top\, dt\right]$ *(auto-covariance)*

Stochastic processes can be represented as Ito integrals

Theorem (Martingale representation theorem)

If $X(\cdot)$ is a (local) martingale adapted to the filtration $\mathcal{F}_t$ which is generated by Brownian motion $W(\cdot)$ then there exists a volatility process $\sigma(t, \omega)$ such that

$$dX(t) = \sigma(t, \omega)\, dW(t).$$

Moreover, if $X(\cdot)$ is a square-integrable martingale then σ is in H^2.

Ito's Lemma is one of the most relevant tools

Theorem (Ito's Lemma)

Let $X(t)$ be an Ito process and $f(\cdot)$ a twice continuous differentiable scalar function. Then

$$df\left(X(t)\right) = \nabla_X f(X)^\top dX(t) + \frac{1}{2} dX(t)^\top H_X f(x) dX(t)$$

with $\nabla_X f$ being the gradient of f and $H_X f(x)$ being the Hessian of f.

Here we use *calculus* $dW_i(t)dW_i(t) = dt$ and $dW_i(t)dW_j(t) = 0$ for $i \neq j$.

Corollary (Ito product rule)

Let $X_1(t)$ and $X_2(t)$ be scalar Ito processes. Then

$$d\left[X_1(t)X_2(t)\right] = X_1(t)dX_2(t) + X_2(t)dX_1(t) + dX_1(t)dX_2(t).$$

Outline

Pricing builds on measure theory and stochastic processes

Probability space & filtration

Brownian Motion

Self-financing trading strategy & arbitrage

Change of measure

Ito integral

Equivalent martingale measure & FTAP

Martingale

Martingale representation

Change of equiv. martingale meas.

Density process

Ito's lemma

Permissible trading strategy

Risk-neutral derivative pricing formula

We specify our market based on assets and trading strategies

Financial Market

We assume p (dividend-free[1]) assets $X(t) = [X_1(t), \ldots, X_p(t)]^\top$ which are driven by Ito processes

$$dX(t) = \mu\,(t,\omega)\,dt + \sigma\,(t,\omega)\,dW(t).$$

Trading Strategy

A trading strategy represents a predictable adapted process (of asset weights)

$$\phi(t,\omega) = [\phi_1(t,\omega), \ldots, \phi_p(t,\omega)]^\top.$$

The value of the trading strategy (or corresponding portfolio) is

$$\pi(t) = \phi(t)^\top X(t).$$

[1] I.e. no intermediate payments

Self-financing strategies and arbitrage

Trading Gains and Self-financing Strategy

Trading gains (over a short period of time) are $\phi(t)^\top [X(t+dt) - X(t)]$.
This leads to the general specification $\int_t^T \phi(t)^\top dX(t)$.
A trading strategy is self-financing if portfolio changes are only induced by asset returns (no money inflow or outflow). That is

$$\pi(T) - \pi(t) = \int_t^T \phi(s)^\top dX(s).$$

Definition (Arbitrage)

An arbitrage opportunity is a self-financing strategy $\phi(\cdot)$ with $\pi(0) = 0$ and, for some $t \in [0, T]$,

$$\pi(t) \geq 0 \text{ a.s., and } \mathbb{P}\left(\pi(t) > 0\right) > 0.$$

Arbitrage needs to be precluded in a financial model.

Absence of arbitrage is closely related to equivalent martingale measures

Definition (Numeraire and equivalent martingale measure)

A numeraire is a positive asset $N(t)$ of our market. An equivalent martingale measure (corresponding to the numeraire $N(t)$) is a measure $\mathbb{Q}$ such that the normalised asset prices $[X_1(t)/N(t), \ldots, X_p(t)/N(t)]^\top$ are $\mathbb{Q}$-martingales.

Fundamental theorem of asset pricing

Assuming some restrictions on permissible trading strategies one can show that absence of arbitrage is "nearly equivalent" to the existence of an equivalent martingale measure.

Our models are all based on the assumption of no-arbitrage and the existence of an equivalent martingale measure.

Equivalent martingale measures exists for any numeraire (1/2)

Suppose we have a numeraire $N(t)$ and an equivalent martingale measure $\mathbb{Q}^N$. Suppose we also have another numeraire $M(t)$. Define

$$\zeta(t) = \frac{M(t)}{N(t)} \frac{N(0)}{M(0)}.$$

Then

- $\mathbb{E}^N [\zeta(T) \,|\, \mathcal{F}_t] = \mathbb{E}^N \left[\frac{M(T)}{N(T)} \,|\, \mathcal{F}_t \right] \frac{N(0)}{M(0)} = \frac{M(t)}{N(t)} \frac{N(0)}{M(0)} = \zeta(t)$, thus $\zeta(t)$ is a $\mathbb{Q}^N$-martingale
- $\zeta(0) = \frac{M(0)}{N(0)} \frac{N(0)}{M(0)} = 1$

Equivalent martingale measures exists for any numeraire (2/2)

Define the new measure $\mathbb{Q}^M$ via the density $\zeta(t)$. Then for an asset $X_i(t)$

$$\mathbb{E}^M\left[\frac{X_i(T)}{M(T)}\,|\,\mathcal{F}_t\right] = \mathbb{E}^N\left[\frac{\zeta(T)}{\zeta(t)}\,\frac{X_i(T)}{M(T)}\,|\,\mathcal{F}_t\right] = \mathbb{E}^N\left[\frac{M(T)}{N(T)}\,\frac{N(\mathrm{t})}{M(t)}\,\frac{X_i(T)}{M(T)}\,|\,\mathcal{F}_t\right].$$

Taking out what is known and using the martingale property of measure $\mathbb{Q}^N$ yields

$$\mathbb{E}^M\left[\frac{X_i(T)}{M(T)}\,|\,\mathcal{F}_t\right] = \frac{N(\mathrm{t})}{M(t)}\,\mathbb{E}^N\left[\frac{X_i(T)}{N(T)}\,|\,\mathcal{F}_t\right] = \frac{N(\mathrm{t})}{M(t)}\,\frac{X_i(t)}{N(t)} = \frac{X_i(t)}{M(t)}.$$

$X_i(t)/M(t)$ is a $\mathbb{Q}^M$-martingale. Thus $\mathbb{Q}^M$ is an equivalent martingale measure for $M(t)$.

Trading strategies need to be permissible

Definition (Permissible trading strategy)

Let $X(t)$ be an Ito process and $\mathbb{Q}$ an equivalent martingale measure with numeraire $N(t)$. A self-financing trading strategy $\phi(t)$ is called permissible if

$$\int_0^t \phi(s)^\top d\left(\frac{X(s)}{N(s)}\right)$$

is a $\mathbb{Q}$-martingale.

Recall that $X(t)/N(t)$ is a $\mathbb{Q}$-martingale by construction. If $\phi(t)$ is sufficiently bounded then it is also permissible.

Theorem (Martingale property for trading strategies)

For any self-financing and permissible trading strategy $\phi(t)$ and an equivalent martingale measure $\mathbb{Q}$ with numeraire $N(t)$ the discounted portfolio price process $\pi(t)/N(t)$ is a martingale.

On average you can not beat the market when trading in the assets.

We proof the martingale property for trading strategies

Proof.

Recall that $\pi(t) = \phi(t)^\top X(t)$. The self-financing condition may be written as $d\pi(t) = \phi(t)^\top dX(t)$. Applying Ito's product rule yields

$$
\begin{aligned}
d\left[\frac{\pi(t)}{N(t)}\right] &= d\left[\pi(t)\frac{1}{N(t)}\right] = \frac{d\pi(t)}{N(t)} + \pi(t)d\left[\frac{1}{N(t)}\right] + d\pi(t)d\left[\frac{1}{N(t)}\right] \\
&= \frac{\phi(t)^\top dX(t)}{N(t)} + \phi(t)^\top X(t)d\left[\frac{1}{N(t)}\right] + \phi(t)^\top dX(t)d\left[\frac{1}{N(t)}\right] \\
&= \phi(t)^\top \left[\frac{dX(t)}{N(t)} + X(t)d\left[\frac{1}{N(t)}\right] + dX(t)d\left[\frac{1}{N(t)}\right]\right] \\
&= \phi(t)^\top d\left[\frac{X(t)}{N(t)}\right].
\end{aligned}
$$

Now the assertion follows directly from the condition that $\phi(t)$ is permissible. $\square$

Derivative pricing is closely related to trading strategies

Definition (Contingent claim)

A derivative security (or contingent claim) pays at time T the random variable $V(T)$ (no intermediate payments). We assume $V(T)$ has finite variance and is attainable. That is there exists a permissible trading strategy $\phi(\cdot)$ such that

$$V(T) = \phi(T)^\top X(T) \, a.s.$$

Then absence of arbitrage yields that the fair price $V(t)$ of the derivative security becomes

$$V(t) = \phi(t)^\top X(t) \text{ for all } t \in [0, T].$$

Consequently,

$$\frac{V(t)}{N(t)} = \frac{\phi(t)^\top X(t)}{N(t)} = \mathbb{E}^{\mathbb{Q}} \left[\frac{\phi(T)^\top X(T)}{N(T)} \,|\, \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}} \left[\frac{V(T)}{N(T)} \,|\, \mathcal{F}_t \right].$$

Above arbitrage pricing formula is the foundation of derivative pricing.

Outline

We summarize the key results

Probability space & filtration	Brownian Motion	Self-financing trading strategy & arbitrage
Change of measure	Ito integral	Equivalent martingale measure & FTAP
Martingale	Martingale representation	Change of equiv. martingale meas.
Density process	Ito's lemma	Permissible trading strategy

Risk-neutral derivative pricing formula

We summarize the key results (cheat sheet)

$$(\Omega, \mathcal{F}, \mathbb{P}),\ \mathcal{F}_t,\ t \in [0, T]$$

$$W(t) = [W_1(t), \ldots, W_d(t)]^\top$$

$$d\pi(T) = \phi(t)^\top dX(t)$$

$$\mathbb{E}^{\hat{\mathbb{P}}}[X \mid \mathcal{F}_t] = \frac{\mathbb{E}^{\mathbb{P}}[R\,X \mid \mathcal{F}_t]}{\mathbb{E}^{\mathbb{P}}[R \mid \mathcal{F}_t]}$$

$$X(t) = \int_0^t \sigma(u, \omega)\, dW(u)$$

$$\frac{X(t)}{N(t)} = \mathbb{E}^{\mathbb{Q}}\left[\frac{X(T)}{N(T)} \mid \mathcal{F}_t\right]$$

$$X(t) = \mathbb{E}^{\mathbb{P}}[X(s) \mid \mathcal{F}_t]$$

$$dX(t) = \sigma(u, \omega)\, dW(u)$$

$$\mathbb{E}^{M}\left[\frac{X_i(T)}{M(T)} \mid \mathcal{F}_t\right] = \mathbb{E}^{N}\left[\frac{N(t)}{M(t)}\, \frac{X_i(T)}{N(T)} \mid \mathcal{F}_t\right]$$

$$\zeta(t) = \mathbb{E}^{\mathbb{P}}\left[d\hat{\mathbb{P}}/d\mathbb{P} \mid \mathcal{F}_t\right]$$

$$df = f'\, dX + \frac{f''}{2}\, dX^2$$

$$\phi(t)^\top d\left[\frac{X(t)}{N(t)}\right] = \bar{\sigma}\, dW(t)$$

$$V(t)/N(t) = \mathbb{E}^{\mathbb{Q}}\left[V(T)/N(T) \mid \mathcal{F}_t\right]$$

Outline

Outline

First we need to specify the assets in the market (1/2)

Example (Overnight bank account)

- Suppose bank A deposits 1 EUR at ECB at time $T_0 = 0$ (today) with the right to withdraw money at T_1, say the next day.
- Bank A may leave deposit with ECB as long as they want
- Time T_i is measured in years (or year fraction) for simplicity
- ECB pays annualized interest rate r_i from T_i to T_{i+1}

Example also holds for deposits between two banks, e.g. bank A and bank B.

What is the value of the deposit at a future time T_N?

First we need to specify the assets in the market (2/2)

Denote B_i the value of the deposit at time T_i. Then

$$B_0 = 1$$

and

$$B_i = B_{i-1} + r_{i-1} \cdot (T_i - T_{i-1}) \cdot B_{i-1} = [1 + r_{i-1} (T_i - T_{i-1})] \cdot B_{i-1}.$$

The most basic asset is the money market bank account

Definition (Short rate and (abstract) bank account)

Assume a process $r(t)$ (adapted to the filtration $\mathcal{F}_t$) for the instantaneous interest rate. The rate $r(t)$ is denoted the short rate. The continuous compounded bank account (or money market account) is an asset with price $B(t)$ given by $B(0) = 1$ and

$$dB(t) = r(t) \cdot B(t) \cdot dt.$$

It follows that the future price of the bank account becomes

$$B(t) = \exp\left\{\int_0^t r(s)ds\right\}.$$

Short rate $r(t)$ is considered the *risk-free rate* at which market participants can lend and borrow money.

The most relevant assets are zero coupon bonds (ZCBs) (1/2)

ZCBs are fixed future cash flows of unit notional, e.g. 1 EUR in 10y.

Definition (Zero Coupon Bond)

A zero coupon bond for maturity T is an asset with time-t asset price $P(t, T)$ for $t \leq T$ and $P(T, T) = 1$.

What is the time-t asset price of a zero coupon bond?

The most relevant assets are zero coupon bonds (ZCBs) (2/2)

Use risk-neutral pricing formula!

Select money market account $B(t)$ as numeraire and denote $\mathbb{Q}$ the equivalent martingale measure.

Then (with $\mathbb{E}_t^{\mathbb{Q}}[\cdot] = \mathbb{E}^{\mathbb{Q}}[\cdot|\mathcal{F}_t]$)

$$\frac{P(t,T)}{B(t)} = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{P(T,T)}{B(T)}\right] = \mathbb{E}_t^{\mathbb{Q}}\left[B(T)^{-1}\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{-\int_0^T r(s)ds\right\}\right].$$

Multiplying with $B(t) = \exp\left\{\int_0^t r(s)ds\right\}$ yields

$$P(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{-\int_t^T r(s)ds\right\}\right].$$

And what is the ZCB price in terms of money ...?

- Formula $P(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{-\int_t^T r(s)ds\right\}\right]$ is a model-independent result
- To calculate it more concrete we need to specify a model/dynamics for short rate $r(t)$
- Suppose short rate is known deterministic function, then

$$P(t,T) = \exp\left\{-\int_t^T r(s)ds\right\}.$$

- Suppose short rate is fixed, i.e. $r(t) = r_0$, then (even simpler)

$$P(t,T) = e^{-r_0(T-t)}.$$

For our market we assume that today's prices $P(0,T)$ of all ZCBs (with maturity $T \geq 0$) are known.

Interest rate market consists of money market bank account and zero coupon bonds

Interest rate market

We consider a market consisting of the money market account $B(t)$ and zero coupon bonds $P(t, T)$ for $t \leq T$ as financial assets.

Interest rate derivatives

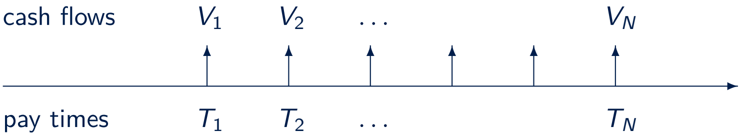
Interest rate derivatives are contingent claims (or baskets of contingent claims) depending on realisations of future zero coupon bonds.

- We may restrict modelling to discrete set of ZCBs $\{P(t, T_i)\}_i$ (vanilla models).
- Full continuum of ZCBs $\{P(t, T) \,|\, t \leq T\}$ is modelled via term structure models.

Outline

Discounted cash flow (DCF) pricing methodology ...

cash flow stream (or leg)



$$\frac{V(t)}{B(t)} = \sum_{i=1}^{N} \mathbb{E}^{\mathbb{Q}} \left[\frac{V_i}{B(T_i)} \,|\, \mathcal{F}_t \right]$$

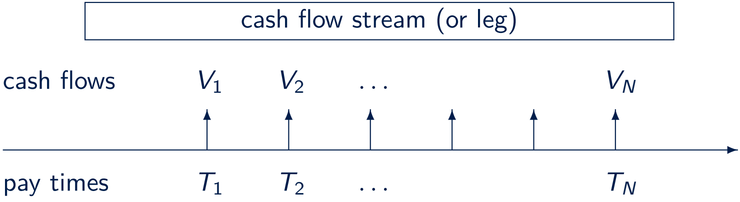
Denote $\mathbb{E}^{T_i}[\cdot]$ expectations in T_i-forward measures with zero coupon bond numeraire $P(t, T_i)$ $(i = 1, \ldots, N)$. Then (change of measure)

$$\frac{V(t)}{B(t)} = \sum_{i=1}^{N} \mathbb{E}^{T_i} \left[\frac{P(t, T_i)}{B(t)} \cdot \frac{V_i}{P(T_i, T_i)} \,|\, \mathcal{F}_t \right].$$

With $P(T_i, T_i) = 1$ follows

$$V(t) = \sum_{i=1}^{N} P(t, T_i) \cdot \mathbb{E}^{T_i} [V_i \,|\, \mathcal{F}_t].$$

(DCF) ... is a model-independent concept



$$V(t) = \sum_{i=1}^{N} P(t, T_i) \cdot \mathbb{E}^{T_i} \left[V_i \,|\, \mathcal{F}_t \right]$$

- Present value is sum of discounted expected future cash flows.
- If future cash flows are known (i.e. deterministic), then

$$V(t) = \sum_{i=1}^{N} P(t, T_i) \cdot V_i$$

.

- In general, challenge lies in calculating $\mathbb{E}^{T_i} \left[V_i \,|\, \mathcal{F}_t \right]$ using a model.

Part II

Yield Curves and Linear Products

Outline

Static Yield Curve Modelling and Market Conventions

Multi-Curve Discounted Cash Flow Pricing

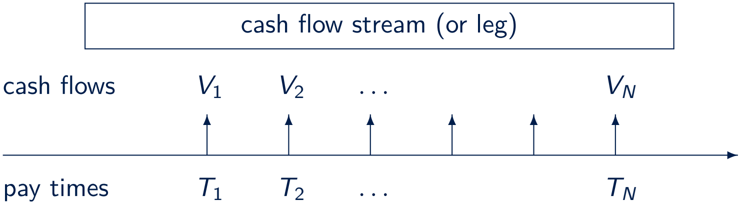
Linear Market Instruments

Credit-risky and Collateralized Discounting

Outline

Outline

DCF method requires knowledge of today's ZCB prices



- Assume $t = 0$ and deterministic cash flows, then

$$V(0) = \sum_{i=1}^{N} P(0, T_i) \cdot V_i.$$

How do we get today's ZCB prices $P(0, T_i)$?

Yield curve is fundamental object for interest rate modelling

- A yield curve (YC) at an observation time t is the function of zero coupon bonds $P(t,\cdot) : [t,\infty) \to \mathbb{R}^+$ for maturities $T \geq t$.

- YCs are typically represented in terms of interest rates (instead of zero coupon bond prices).

- Discretely compounded zero rate curve $z_p(t,T)$ with frequency p, such that
$$P(t,T) = \left(1 + \frac{z_p(t,T)}{p}\right)^{-p\cdot(T-t)}.$$

- Simple compounded zero rate curve $z_0(t,T)$ (i.e. $p = 1/(T-t)$), such that
$$P(t,T) = \frac{1}{1 + z_0(t,T)\cdot(T-t)}.$$

- Continuous compounded zero rate curve $z(t,T)$ (i.e. $p = \infty$), such that
$$P(t,T) = \exp\left\{-z(t,T)\cdot(T-t)\right\}.$$

For interest rate modelling we also need continuous compounded forward rates

Definition (Continuous Forward Rate)

Suppose a given observation time t and zero bond curve $P(t,\cdot):[t,\infty)\to\mathbb{R}^+$ for maturities $T \geq t$. The continuous compounded forward rate curve is given by

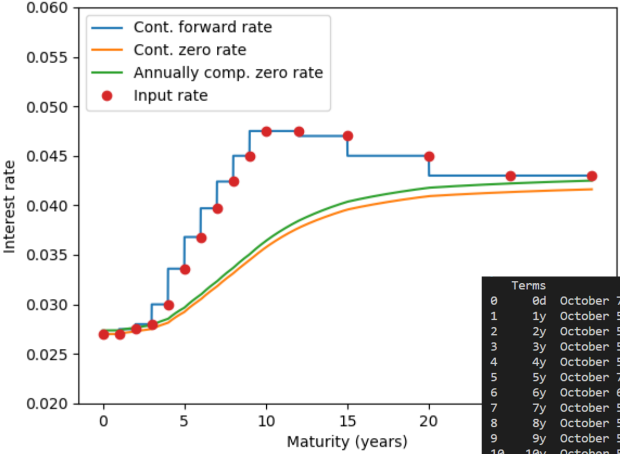
$$f(t,T) = -\frac{\partial \ln\left(P(t,T)\right)}{\partial T}.$$

From the definition follows

$$P(t,T) = \exp\left\{-\int_t^T f(t,s)ds\right\}.$$

- For static yield curve modelling and (simple) linear instrument pricing we are interested particularly in curves at $t=0$.
- For (more complex) option pricing we are interested in modelling curves at $t>0$.

We show a typical yield curve example



	Terms	Dates	Times	Rates
0	0d	October 7th, 2019	0	0.027
1	1y	October 5th, 2020	0.99726	0.027
2	2y	October 5th, 2021	1.99726	0.0275
3	3y	October 5th, 2022	2.99726	0.028
4	4y	October 5th, 2023	3.99726	0.03
5	5y	October 7th, 2024	5.00548	0.0336
6	6y	October 6th, 2025	6.00274	0.0368
7	7y	October 5th, 2026	7	0.0397
8	8y	October 5th, 2027	8	0.0424
9	9y	October 5th, 2028	9.00274	0.045
10	10y	October 5th, 2029	10.0027	0.0475
11	12y	October 6th, 2031	12.0055	0.0475
12	15y	October 5th, 2034	15.0055	0.047
13	20y	October 5th, 2039	20.0082	0.045
14	25y	October 5th, 2044	25.0137	0.043
15	30y	October 5th, 2049	30.0164	0.043

The market data for curve calibration is quoted by market data providers

Euribor vs 6 mth

1 Yr	-0.226/-0.266	16Yrs	1.295/1.255
2 Yrs	-0.128/-0.168	17Yrs	1.334/1.294
3 Yrs	0.010/-0.030	18Yrs	1.367/1.327
4 Yrs	0.154/0.114	19Yrs	1.393/1.353
5 Yrs	0.293/0.253	20Yrs	1.415/1.375
6 Yrs	0.429/0.389		
7 Yrs	0.558/0.518	21Yrs	1.432/1.392
8 Yrs	0.678/0.638	22Yrs	1.446/1.406
9 Yrs	0.790/0.750	23Yrs	1.457/1.417
10Yrs	0.892/0.852	24Yrs	1.465/1.425
		25Yrs	1.471/1.431
11Yrs	0.983/0.943		
12Yrs	1.064/1.024	26Yrs	1.476/1.436
13Yrs	1.135/1.095	27Yrs	1.480/1.440
14Yrs	1.197/1.157	28Yrs	1.484/1.444
15Yrs	1.250/1.210	29Yrs	1.486/1.446
		30Yrs	1.488/1.448
		35Yrs	1.491/1.451
		40Yrs	1.486/1.446
		45Yrs	1.476/1.436
		50Yrs	1.466/1.426
Disclaimer <IDIS>		60Yrs	1.456/1.416

3/6 basis
Spot Starting Date

1 Yr	4.30
2 Yrs	4.80
3 Yrs	5.35
4 Yrs	5.90
5 Yrs	6.40
6 Yrs	6.70
7 Yrs	6.85
8 Yrs	6.90
9 Yrs	6.90
10Yrs	6.85

Swap Spreads
(Gadget)

5y	59.3
10y	66.0

Page live in
London hours ONLY
(between 0700 - 1800)

This page will close 30th April
6.00pm and re open 7.00am 2nd May

10X12	0.192/0.152
10X15	0.378/0.338
10X20	0.543/0.503
10X25	0.599/0.559
10X30	0.616/0.576
10X35	0.619/0.579
10X40	0.614/0.574
10X45	0.604/0.564
10X50	0.594/0.554
10X60	0.584/0.544

Outline

Recall the introductory swap example

Interbank swap deal example

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

Dates

End date: Oct 30, 2040

Market conventions

(annually, 30/360 day count, modified following, Target calendar)

Pays 6-months Euribor floating rate on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(semi-annually, act/360 day count, modified following, Target calendar)

How do we get from description to cash flow stream?

There are a couple of market conventions that need to be taken into account in practice

- Holiday calendars define at which dates payments can be made.
- Business day conventions specify how dates are adjusted if they fall on a non-business day.
- Schedule generation rules specify how regular dates are calculated.
- Day count conventions define how time is meassured between dates.

Outline

Dates are represented as triples day/month/year or as serial numbers

	Date	Serial	EUR Payment System (TARGET)	London Bank Holiday
	Friday, July 27, 2018	43308	FALSE	FALSE
	Monday, August 27, 2018	43339	FALSE	TRUE
	Thursday, September 27, 2018	43370	FALSE	FALSE
	Saturday, October 27, 2018	43400	TRUE	TRUE
	Tuesday, November 27, 2018	43431	FALSE	FALSE
	Thursday, December 27, 2018	43461	FALSE	FALSE
	Sunday, January 27, 2019	43492	TRUE	TRUE
	Wednesday, February 27, 2019	43523	FALSE	FALSE
	Wednesday, March 27, 2019	43551	FALSE	FALSE
	Saturday, April 27, 2019	43582	TRUE	TRUE
	Monday, May 27, 2019	43612	FALSE	TRUE
	Sunday, January 1, 1900	1		

A calender specifies business days and non-business days

Holiday Calendar

A holiday calendar $\mathcal{C}$ is a set of dates which are defined as holidays or non-business days.

- A particular date d is a non-business day if $d \in \mathcal{C}$.
- Holiday calendars are specific to a region, country or market segment.
- Need to be specified in the context of financial product.
- Typically contain weekends and special days of the year.
- May be joined (e.g. for multi-currency products), $\bar{\mathcal{C}} = \mathcal{C}_1 \cup \mathcal{C}_2$.
- Typical examples are TARGET calendar and LONDON calendar.

Outline

A business day convention maps non-business days to adjacent business days

Business Day Convention (BDC)

- A business day convention is a function $\omega_{\mathcal{C}} : \mathcal{D} \to \mathcal{D}$ which maps a date $d \in \mathcal{D}$ to another date $\bar{d}$.
- It is applied in conjunction with a calendar $\mathcal{C}$.
- Good business days are unchanged, i.e. $\omega_{\mathcal{C}}(d) = d$ if $d \notin \mathcal{C}$.

Following

$$\omega_{\mathcal{C}}(d) = \min\left\{\bar{d} \in \mathcal{D} \backslash \mathcal{C} \mid \bar{d} \geq d\right\}$$

Preceding

$$\omega_{\mathcal{C}}(d) = \max\left\{\bar{d} \in \mathcal{D} \backslash \mathcal{C} \mid \bar{d} \leq d\right\}$$

Modified Following

$$\omega_{\mathcal{C}}(d) = \begin{cases} \omega_{\mathcal{C}}^{\text{Following}}(d), & \text{if Month}\,[d] = \text{Month}\left[\omega_{\mathcal{C}}^{\text{Following}}(d)\right] \\ \omega_{\mathcal{C}}^{\text{Preceeding}}(d), & \text{else} \end{cases}$$

Outline

Schedules represent sets of regular reference dates

	Annual Frequency	TARGET Calendar	Modified Following
Start	Fri, 30 Oct 2020	FALSE	Fri, 30 Oct 2020
	Sat, 30 Oct 2021	TRUE	Fri, 29 Oct 2021
	Sun, 30 Oct 2022	TRUE	Mon, 31 Oct 2022
	Mon, 30 Oct 2023	FALSE	Mon, 30 Oct 2023
	Wed, 30 Oct 2024	FALSE	Wed, 30 Oct 2024
	Thu, 30 Oct 2025	FALSE	Thu, 30 Oct 2025
	Fri, 30 Oct 2026	FALSE	Fri, 30 Oct 2026
	Sat, 30 Oct 2027	TRUE	Fri, 29 Oct 2027
	Mon, 30 Oct 2028	FALSE	Mon, 30 Oct 2028
	Tue, 30 Oct 2029	FALSE	Tue, 30 Oct 2029
	Wed, 30 Oct 2030	FALSE	Wed, 30 Oct 2030
	Thu, 30 Oct 2031	FALSE	Thu, 30 Oct 2031
	Sat, 30 Oct 2032	TRUE	Fri, 29 Oct 2032
	Sun, 30 Oct 2033	TRUE	Mon, 31 Oct 2033
	Mon, 30 Oct 2034	FALSE	Mon, 30 Oct 2034
	Tue, 30 Oct 2035	FALSE	Tue, 30 Oct 2035
	Thu, 30 Oct 2036	FALSE	Thu, 30 Oct 2036
	Fri, 30 Oct 2037	FALSE	Fri, 30 Oct 2037
	Sat, 30 Oct 2038	TRUE	Fri, 29 Oct 2038
	Sun, 30 Oct 2039	TRUE	Mon, 31 Oct 2039
End	Tue, 30 Oct 2040	FALSE	Tue, 30 Oct 2040

Schedule generation follows some rules/conventions as well

1. Consider direction of roll-out: forward or backward (relevant for front/back stubs).
   1.1 Forward, roll-out from start (or effective) date to end (or maturity) date
   1.2 Backward, roll-out from end (or maturity) date to start (or effective) date
2. Roll out unadjusted dates according to frequency or tenor, e.g. annual frequency or 3 month tenor
3. If first/last period is broken consider short stub or long stub.
   3.1 Short stub is an unregular last period smaller then tenor.
   3.2 Long stub is an unregular last period larger then tenor
4. Adjust unadjusted dates according to calendar and BDC.

Outline

Day count conventions map dates to times or year fractions

Day Count Convention

A day count convention is a function $\tau : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ which measures a time period between dates in terms of years.

We give some examples:

Act/365 Fixed Convention

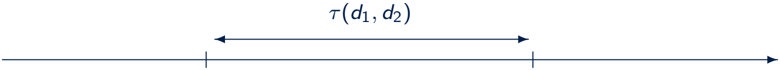
$\tau(d_1, d_2) = (d_2 - d_1)/365$

- Typically used to describe time in financial models.

Act/360 Convention

$\tau(d_1, d_2) = (d_2 - d_1)/360$

- Often used for Libor floating rate payments.



30/360 methods are slightly more involved

General 30/360 Method

- Consider two dates d_1 and d_2 represented as triples of day/month/year, i.e. $d_1 = [D_1, M_1, Y_1]$ and $d_2 = [D_2, M_2, Y_2]$ with $D_{1/2} \in \{1, \dots, 31\}$, $M_{1/2} \in \{1, \dots, 12\}$ and $Y_{1/2} \in \{1, 2, \dots\}$.

- Obviously, only valid dates are allowed (no Feb. 30 or similar).

- Adjust $D_1 \mapsto \bar{D}_1$ and $D_2 \mapsto \bar{D}_2$ according to specific rules.

- Calculate

$$\tau(d_1, d_2) = \frac{360 \cdot (Y_2 - Y_1) + 30 \cdot (M_2 - M_1) + \left(\bar{D}_2 - \bar{D}_1\right)}{360}.$$

Some specific 30/360 rules are given below

30/360 Convention (or 30U/360, Bond Basis)

1. $\bar{D}_1 = \min\{D_1, 30\}$.
2. If $\bar{D}_1 = 30$ then $\bar{D}_2 = \min\{D_2, 30\}$ else if $\bar{D}_2 = D_2$.

30E/360 Convention (or Eurobond)

1. $\bar{D}_1 = \min\{D_1, 30\}$.
2. $\bar{D}_2 = \min\{D_2, 30\}$.

Outline

Now we have all pieces to price a deterministic coupon leg

Coupon is calculated as

$$\begin{aligned} \text{Coupon} &= \text{Notional} \times \text{Rate} \times \text{YearFraction} \\ &= 100,000,000\text{EUR} \times 3\% \times \tau \end{aligned}$$

ValDate	Thu, 01 Oct 2020

Sum	41,787,559

	Annual Frequency	TARGET Calendar	Modified Following	_D1	_D2	tau	Rate	Coupon	P(0,T)	P(0,T)*Cpn
Start	Fri, 30 Oct 2020	FALSE	Fri, 30 Oct 2020							
	Sat, 30 Oct 2021	TRUE	Fri, 29 Oct 2021	30	29	0.997	3.00%	2,991,667	0.9713	2,905,943
	Sun, 30 Oct 2022	TRUE	Mon, 31 Oct 2022	29	31	1.006	3.00%	3,016,667	0.9451	2,850,916
	Mon, 30 Oct 2023	FALSE	Mon, 30 Oct 2023	30	30	1.000	3.00%	3,000,000	0.9192	2,757,657
	Wed, 30 Oct 2024	FALSE	Wed, 30 Oct 2024	30	30	1.000	3.00%	3,000,000	0.8927	2,678,166
	Thu, 30 Oct 2025	FALSE	Thu, 30 Oct 2025	30	30	1.000	3.00%	3,000,000	0.8646	2,593,664
	Fri, 30 Oct 2026	FALSE	Fri, 30 Oct 2026	30	30	1.000	3.00%	3,000,000	0.8345	2,503,445
	Sat, 30 Oct 2027	TRUE	Fri, 29 Oct 2027	30	29	0.997	3.00%	2,991,667	0.8031	2,402,572
	Mon, 30 Oct 2028	FALSE	Mon, 30 Oct 2028	29	30	1.003	3.00%	3,008,333	0.7704	2,317,730
	Tue, 30 Oct 2029	FALSE	Tue, 30 Oct 2029	30	30	1.000	3.00%	3,000,000	0.7373	2,211,969
	Wed, 30 Oct 2030	FALSE	Wed, 30 Oct 2030	30	30	1.000	3.00%	3,000,000	0.7039	2,111,644
	Thu, 30 Oct 2031	FALSE	Thu, 30 Oct 2031	30	30	1.000	3.00%	3,000,000	0.6713	2,013,762
	Sat, 30 Oct 2032	TRUE	Fri, 29 Oct 2032	30	29	0.997	3.00%	2,991,667	0.6401	1,915,033
	Sun, 30 Oct 2033	TRUE	Mon, 31 Oct 2033	29	31	1.006	3.00%	3,016,667	0.6103	1,841,155
	Mon, 30 Oct 2034	FALSE	Mon, 30 Oct 2034	30	30	1.000	3.00%	3,000,000	0.5822	1,746,731
	Tue, 30 Oct 2035	FALSE	Tue, 30 Oct 2035	30	30	1.000	3.00%	3,000,000	0.5555	1,666,418
	Thu, 30 Oct 2036	FALSE	Thu, 30 Oct 2036	30	30	1.000	3.00%	3,000,000	0.5300	1,590,074
	Fri, 30 Oct 2037	FALSE	Fri, 30 Oct 2037	30	30	1.000	3.00%	3,000,000	0.5060	1,518,029
	Sat, 30 Oct 2038	TRUE	Fri, 29 Oct 2038	30	29	0.997	3.00%	2,991,667	0.4833	1,445,981
	Sun, 30 Oct 2039	TRUE	Mon, 31 Oct 2039	29	31	1.006	3.00%	3,016,667	0.4617	1,392,766
End	Tue, 30 Oct 2040	FALSE	Tue, 30 Oct 2040	30	30	1.000	3.00%	3,000,000	0.4413	1,323,902

Outline

Static Yield Curve Modelling and Market Conventions

Multi-Curve Discounted Cash Flow Pricing

Linear Market Instruments

Credit-risky and Collateralized Discounting

Outline

Recall the introductory swap example

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(annually, 30/360 day count, modified following, Target calendar)

Stochastic interest rates

Pays 6-months Euribor floating rate on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(semi-annually, act/360 day count, modified following, Target calendar)

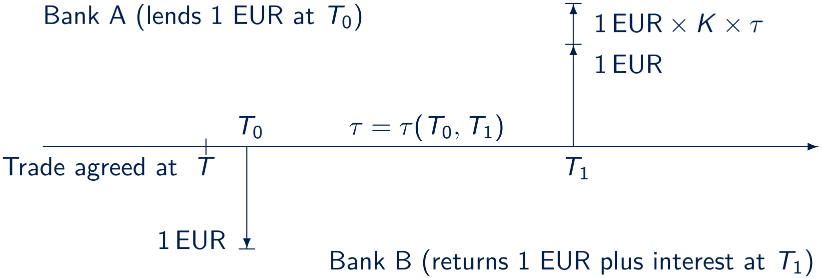
How do we model floating rates?

We start with some introductory remarks

- London Interbank Offered Rates (Libor) used to be the key building blocks of interest rate derivatives (for USD, GBP, JPY, CHF).
- EUR equivalent rate is Euribor rate - we will use Libor synonymously for Euribor.
- Libor rate modelling has undergone significant changes since financial crisis in 2008.
- This is typically reflected by the term Multi-Curve Interest Rate Modelling.
- Recent developments in the market lead to a shift away from Libor rates to alternative reference rates (Ibor Transition or Benchmark Reform).
- Alternative rates specifications lead to overnight index swaps.

Let's start with the classical Libor rate model

What is the fair interest rate K bank A and Bank B can agree on?



We get (via DCF methodology)

$$
\begin{aligned}
0 = V(T) &= P(T, T_0) \cdot \mathbb{E}^{T_0}\left[-1 \mid \mathcal{F}_T\right] + P(T, T_1) \cdot \mathbb{E}^{T_1}\left[1 + \tau K \mid \mathcal{F}_T\right], \\
0 &= -P(T, T_0) + P(T, T_1) \cdot (1 + \tau K).
\end{aligned}
$$

Spot Libor rates are fixed daily and quoted in the market

$$0 = -P(T, T_0) + P(T, T_1) \cdot (1 + \tau K)$$

Spot Libor rate

The fair rate for an interbank lending deal with trade date T, spot starting date T_0 (typically 0d or 2d after T) and maturity date T_1 is

$$L(T; T_0, T_1) = \left[\frac{P(T, T_0)}{P(T, T_1)} - 1 \right] \frac{1}{\tau}.$$

- Panel banks submit daily estimates for interbank lending rates to calculation agent.
- Relevant periods (i.e. $[T_0, T_1]$) considered are 1m, 3m, 6m and 12m.
- Trimmed average of submissions is calculated and published.

Libor rate fixings used to be the most important reference rates for interest rate derivatives. Nowadays, overnight rates become the key reference rates.

Example publication at Intercontinental Exchange (ICE) and EMMI

theice.com/marketdata/reports/170

ICE LIBOR Historical Rates

TENOR	PUBLICATION TIME*	USD ICE LIBOR 06-SEP-2018
Overnight	11:55:04 AM	1.91838
1 Week	11:55:04 AM	1.96100
1 Month	11:55:04 AM	2.13256
2 Month	11:55:04 AM	2.20950
3 Month	11:55:04 AM	2.32706
6 Month	11:55:04 AM	2.54419
1 Year	11:55:04 AM	2.84906

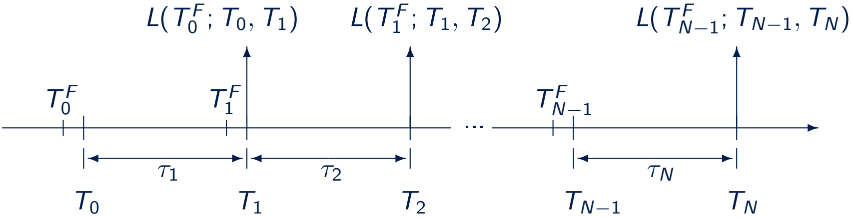
https://www.emmi-benchmarks.eu/benchmarks/

SS

Euribor

Date	1 Week	1 Month	3 Months	6 Months	12 Months
19 Apr 2022	-0.572	-0.560	-0.468	-0.333	-0.010

A plain vanilla Libor leg pays periodic Libor rate coupons



We get (via DCF methodology)

$$
\begin{aligned}
V(t) &= \sum_{i=1}^{N} P(t, T_i) \cdot \mathbb{E}^{T_i}\left[L(T_{i-1}^F; T_{i-1}, T_i) \cdot \tau_i \mid \mathcal{F}_t \right] \\
&= \sum_{i=1}^{N} P(t, T_i) \cdot \mathbb{E}^{T_i}\left[L(T_{i-1}^F; T_{i-1}, T_i) \mid \mathcal{F}_t \right] \cdot \tau_i .
\end{aligned}
$$

Thus all we need is

$$
\mathbb{E}^{T_i}\left[L(T_{i-1}^F; T_{i-1}, T_i) \mid \mathcal{F}_t \right] = ?
$$

Libor rate is a martingale in the terminal measure (1/2)

Theorem (Martingale property of Libor rate)

The Libor rate $L(T; T_0, T_1)$ *with observation/fixing date* T*, accrual start date* T_0 *and accrual end date* T_1 *is a martingale in the* T_1*-forward measure and*

$$\mathbb{E}^{T_1}\left[L(T; T_0, T_1) \mid \mathcal{F}_t\right] = \left[\frac{P(t, T_0)}{P(t, T_1)} - 1\right]\frac{1}{\tau} = L(t; T_0, T_1).$$

Libor rate is a martingale in the terminal measure (2/2)

Proof.

Fair Libor rate at fixing time T is $L(T; T_0, T_1) = [P(T, T_0)/P(T, T_1) - 1]/\tau$. The zero coupon bond $P(T, T_0)$ is an asset and $P(T, T_1)$ is the numeraire in the T_1-forward meassure. Thus FTAP yields that the discounted asset price is a martingale, i.e.

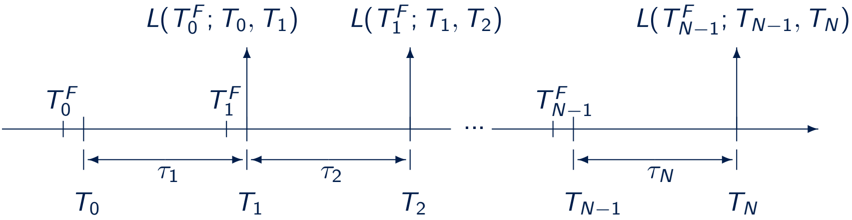
$$\mathbb{E}^{T_1}\left[\frac{P(T, T_0)}{P(T, T_1)} \mid \mathcal{F}_t\right] = \frac{P(t, T_0)}{P(t, T_1)}.$$

Linearity of expectation operator yields

$$\begin{aligned}
\mathbb{E}^{T_1}\left[L(T; T_0, T_1) \mid \mathcal{F}_t\right] &= \left[\mathbb{E}^{T_1}\left[\frac{P(T, T_0)}{P(T, T_1)} | \mathcal{F}_t\right] - 1\right]\frac{1}{\tau} \\
&= \left[\frac{P(t, T_0)}{P(t, T_1)} - 1\right]\frac{1}{\tau} \\
&= L(t; T_0, T_1).
\end{aligned}$$

$\square$

This allows pricing the Libor leg based on today's knowledge of the yield curve only



Libor leg becomes

$$
\begin{aligned}
V(t) &= \sum_{i=1}^{N} P(t, T_i) \cdot \mathbb{E}^{T_i} \left[L(T_{i-1}^F; T_{i-1}, T_i) \cdot \tau_i \mid \mathcal{F}_t \right] \\
&= \sum_{i=1}^{N} P(t, T_i) \cdot L(t; T_{i-1}, T_i) \cdot \tau_i
\end{aligned}
$$

Libor leg may be simplified in the current single-curve setting

We have

$$V(t) = \sum_{i=1}^{N} P(t, T_i) \cdot L(t; T_{i-1}, T_i) \cdot \tau_i$$

with

$$L(t; T_{i-1}, T_i) = \left[\frac{P(t, T_{i-1})}{P(t, T_i)} - 1 \right] \frac{1}{\tau_i}.$$

This yields

$$\begin{aligned} V(t) &= \sum_{i=1}^{N} P(t, T_i) \cdot \left[\frac{P(t, T_{i-1})}{P(t, T_i)} - 1 \right] \frac{1}{\tau_i} \cdot \tau_i \\ &= \sum_{i=1}^{N} P(t, T_{i-1}) - P(t, T_i) \\ &= P(t, T_0) - P(t, T_N). \end{aligned}$$

We only need discount fators $P(t, T_0)$ and $P(t, T_N)$ at first date T_0 and last date T_N.

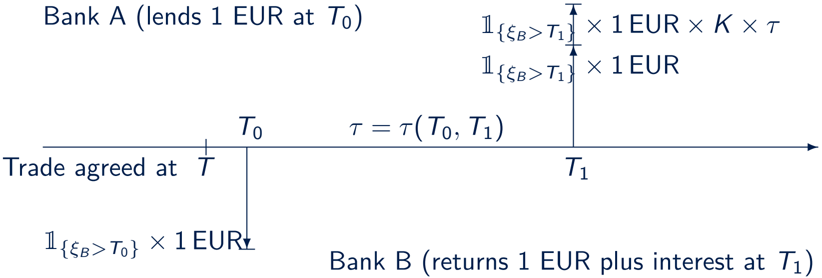
Outline

The classical Libor rate model misses an important detail

What if a counterparty defaults?

What if Bank B defaults prior to T_0 or T_1?

What is the fair rate K bank A and Bank B can agree on given the risk of default?



- Cash flows are paid only if no default occurs.
- We apply a simple credit model.
- Denote $\mathbb{1}_D$ the indicator function for an event D and random variable ξ_B the first time bank B defaults.

Credit-risky trade value can be derived using derivative pricing formula

$$\frac{V(T)}{B(T)} = \mathbb{E}^{\mathbb{Q}}\left[-\mathbb{1}_{\{\xi_B > T_0\}} \cdot \frac{1}{B(T_0)} + \mathbb{1}_{\{\xi_B > T_1\}} \cdot \frac{1 + K \cdot \tau}{B(T_1)}\right].$$

(all expectations conditional on $\mathcal{F}_T$)

Assume independence of credit event $\left\{\xi_B > T_{0/1}\right\}$ and interest rate market, then

$$\frac{V(T)}{B(T)} = -\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\xi_B > T_0\}}\right] \cdot \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{B(T_0)}\right] + \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\xi_B > T_1\}}\right] \cdot \mathbb{E}^{\mathbb{Q}}\left[\frac{1 + K \cdot \tau}{B(T_1)}\right].$$

Abbreviate survival probability $Q(T, T_{0,1}) = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\xi_B > T_{0,1}\}} \,|\, \mathcal{F}_T\right]$ and apply change of measure

$$V(T) = -P(T, T_0)Q(T, T_0)\mathbb{E}^{T_0}[1] + P(T, T_1)Q(T, T_1)\mathbb{E}^{T_1}[1 + K \cdot \tau].$$

This yields the fair spot rate in the presence of credit risk

$$V(T) = -P(T, T_0)Q(T, T_0)\mathbb{E}^{T_0}[1] + P(T, T_1)Q(T, T_1)\mathbb{E}^{T_1}[1 + K \cdot \tau].$$

If we solve $V(T) = 0$ and set $K = L(T; T_0, T_1)$ we get

$$L(T; T_0, T_1) = \left[\frac{P(T, T_0)}{P(T, T_1)} \cdot \frac{Q(T, T_0)}{Q(T, T_1)} - 1\right]\frac{1}{\tau}.$$

We need a model for the survival probability $Q(T, T_{1,2})$.

Consider, e.g., hazard rate model $Q(T, T_{1,2}) = \exp\left\{-\int_T^{T_{1,2}} \lambda(s)ds\right\}$ with deterministic hazard rate $\lambda(s)$. Then basis factor $D(T_0, T_1)$ with

$$D(T_0, T_1) = \frac{Q(T, T_0)}{Q(T, T_1)} = \exp\left\{-\int_{T_0}^{T_1} \lambda(s)ds\right\}$$

is independent of observation time T.

Deterministic hazard rate assumption preserves the martingale property of forward Libor rate

Theorem (Martingale property of credit-risky Libor rate)

Consider the credit-risky Libor rate $L(T; T_0, T_1)$ with observation/fixing date T, accrual start date T_0 and accrual end date T_1. If the basis factor $D(T_0, T_1)$ is deterministic such that

$$L(T; T_0, T_1) = \left[\frac{P(T, T_0)}{P(T, T_1)} \cdot D(T_0, T_1) - 1 \right] \frac{1}{\tau},$$

then $L(t; T_0, T_1)$ is a martingale in the T_1-forward measure and

$$\mathbb{E}^{T_1} \left[L(T; T_0, T_1) \mid \mathcal{F}_t \right] = L(t; T_0, T_1) = \left[\frac{P(t, T_0)}{P(t, T_1)} \cdot D(T_0, T_1) - 1 \right] \frac{1}{\tau}.$$

Proof.

Follows analogously to classical Libor rate martingale property. $\square$

Outline

Multi-Curve Discounted Cash Flow Pricing

- Classical Interbank Floating Rates
- Tenor-basis Modelling
- Projection Curves and Multi-Curve Pricing

Forward Libor rates are typically parametrised via projection curve

- Hazard rate $\lambda(u)$ in $Q(T, T_{1,2}) = \exp\left\{-\int_T^{T_{1,2}} \lambda(u)du\right\}$ is often considered as a tenor basis spread $s(u)$.

- Survival probability $Q(T, T_{1,2})$ can be interpreted as discount factor.

- Suppose we know time-t survival probabilities $Q(t, \cdot)$ for a forward Libor rate $L(t, T_0, T_0 + \delta)$ with tenor δ (typically 1m, 3m, 6m or 12m). Then we define the projection curve

$$P^\delta(t, T) = P(t, T) \cdot Q(t, T).$$

- With projection curve $P^\delta(t, T)$ the forward Libor rate formula is analogous to the classical Libor rate formula, i.e.

$$L^\delta(t, T_0) = L(t; T_0, T_0 + \delta) = \left[\frac{P^\delta(t, T_0)}{P^\delta(t, T_1)} - 1\right]\frac{1}{\tau}.$$

This yields the multi-curve modelling framework consisting of discount curve $P(t, T)$ and tenor-dependent projection curves $P^\delta(t, T)$.

There is an alternative approach to introduce multi-curve modelling

Define forward Libor rate $L^{\delta}(t, T_0)$ for a tenor δ as

$$L^{\delta}(t, T_0) = \mathbb{E}^{T_1}\left[L(T; T_0, T_0 + \delta) \mid \mathcal{F}_t\right].$$

(Without any assumptions on default, survival probabilities etc.)

Postulate a projection curve parametrisation

$$L^{\delta}(t, T_0) = \left[\frac{P^{\delta}(t, T_0)}{P^{\delta}(t, T_1)} - 1\right]\frac{1}{\tau}.$$

- We will discuss calibration of projection curve $P^{\delta}(t, T)$ later.
- This approach alone suffices for linear products (e.g. Libor legs) and simple options.
- It does not specify any relation between projection curve $P^{\delta}(t, T)$ and discount curve $P(t, T)$.

Projection curves can also be written in terms of zero rates and continuous forward rates

Consider a projection curve given by (pseudo) discount factors $P^\delta(t, T)$ (observed today).

- Corresponding continuous compounded zero rates are

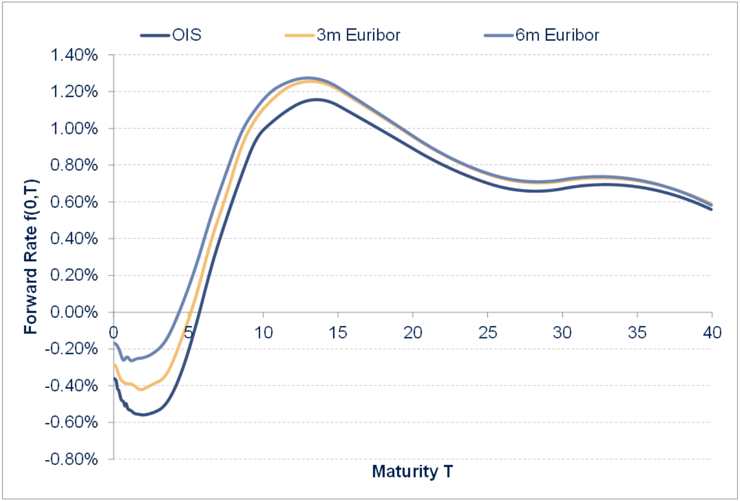
$$z^\delta(t, T) = -\frac{\ln\left[P^\delta(t, T)\right]}{T - t}.$$

- Corresponding continuous compounded forward rates are

$$f^\delta(t, T) = -\frac{\partial \ln\left[P^\delta(t, T)\right]}{\partial T}.$$

We illustrate an example of a multi-curve set-up for EUR

Market data as of July 2016



Libor leg pricing needs to be adapted slightly for multi-curve pricing

Classical single-curve Libor leg price is

$$\begin{aligned} V(t) &= \sum_{i=1}^{N} P(t, T_i) \cdot L(t; T_{i-1}, T_i) \cdot \tau_i \\ &= P(t, T_0) - P(t, T_N). \end{aligned}$$

Multi-curve Libor leg pricing becomes

$$V(t) = \sum_{i=1}^{N} P(t, T_i) \cdot L^{\delta}(t, T_{i-1}) \cdot \tau_i$$

with

$$L^{\delta}(t, T_{i-1}) = \left[\frac{P^{\delta}(t, T_{i-1})}{P^{\delta}(t, T_i)} - 1 \right] \frac{1}{\tau_i}.$$

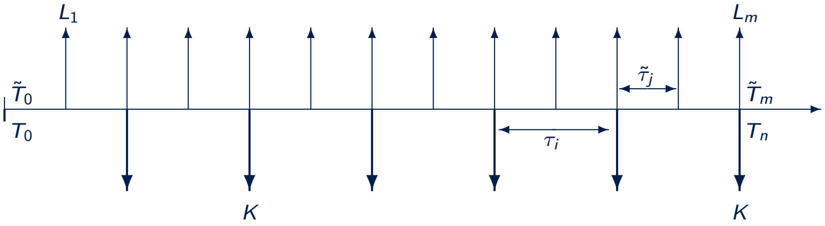
- Note that we need different yield curves for Libor rate projection and cash flow discounting.
- Single-curve pricing formula simplification does not work for multi-curve pricing.

Outline

Outline

With the fixed leg and Libor leg pricing available we can directly price a Vanilla interest rate swap

float leg (EUR conventions: 6m Euribor, Act/360)



fixed leg (EUR conventions: annual, 30/360)

Present value of (fixed rate) payer swap with notional N becomes

$$V(t) = \sum_{j=1}^{m} N \cdot L^{6m}(t, \tilde{T}_{j-1}) \cdot \tilde{\tau}_j \cdot P(t, \tilde{T}_j) - \sum_{i=1}^{n} N \cdot K \cdot \tau_i \cdot P(t, T_i).$$

Vanilla swap pricing formula allows us to price the underlying swap of our introductory example

Interbank swap deal example

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(annually, 30/360 day count, modified following, Target calendar)

Pays 6-months Euribor floating rate on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

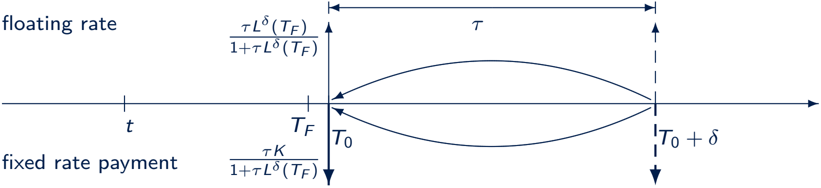
(semi-annually, act/360 day count, modified following, Target calendar)

We illustrate swap pricing with QuantLib/Excel...

- see YieldCurvesAndLegs.xlsx

Outline

Forward Rate Agreement yields exposure to single forward Libor rates



- Fixed rate K agreed at trade inception (prior to t).
- Libor rate $L^\delta(T_F, T_0)$ fixed at T_F, valid for the period T_0 to $T_0 + \delta$.
- Payoff paid at T_0 is difference $\tau \cdot \left[L^\delta(T_F, T_0) - K\right]$ discounted from T_1 to T_0 with discount factor $\left[1 + \tau \cdot L^\delta(T_F, T_0)\right]^{-1}$, i.e.

$$V(T_0) = \frac{\tau \cdot \left[L^\delta(T_F, T_0) - K\right]}{1 + \tau \cdot L^\delta(T_F, T_0)}.$$

Time-T_F FRA price can be obtained via deterministic basis spread model

Note that payoff $V(T_0) = \frac{\tau \cdot \left[L^\delta(T_F, T_0) - K\right]}{1 + \tau \cdot L^\delta(T_F, T_0)}$ is already determined at T_F. Thus (via DCF)

$$V(T_F) = P(T_F, T_0) \cdot V(T_0) = P(T_F, T_0) \cdot \frac{\tau \cdot \left[L^\delta(T_F, T_0) - K\right]}{1 + \tau \cdot L^\delta(T_F, T_0)}.$$

Recall that (with $T_1 = T_0 + \delta$)

$$1 + \tau \cdot L^\delta(T_F, T_0) = \frac{P^\delta(T_F, T_0)}{P^\delta(T_F, T_1)} = \frac{P(T_F, T_0)}{P(T_F, T_1)} \cdot D(T_0, T_1).$$

Then

$$\begin{aligned} V(T_F) &= P(T_F, T_0) \cdot \tau \cdot \left[L^\delta(T_F, T_0) - K\right] \cdot \frac{1}{D(T_0, T_1)} \cdot \frac{P(T_F, T_1)}{P(T_F, T_0)} \\ &= P(T_F, T_1) \cdot \tau \cdot \left[L^\delta(T_F, T_0) - K\right] \cdot \frac{1}{D(T_0, T_1)}. \end{aligned}$$

Present value of FRA can be obtained via martingale property

Derivative pricing formula in T_1-terminal measure yields

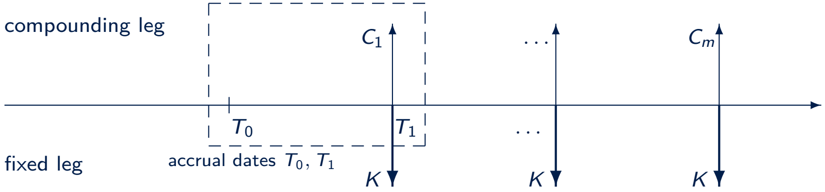
$$\begin{aligned}\frac{V(t)}{P(t,T_1)} &= \mathbb{E}^{T_1}\left[\frac{P(T_F,T_1)}{P(T_F,T_1)}\cdot\tau\cdot\left[L^\delta(T_F,T_0)-K\right]\cdot\frac{1}{D(T_0,T_1)}\right]\\ &= \tau\cdot\left[\mathbb{E}^{T_1}\left[L^\delta(T_F,T_0)\right]-K\right]\cdot\frac{1}{D(T_0,T_1)}\\ &= \tau\cdot\left[L^\delta(t,T_0)-K\right]\cdot\frac{1}{D(T_0,T_1)}.\end{aligned}$$

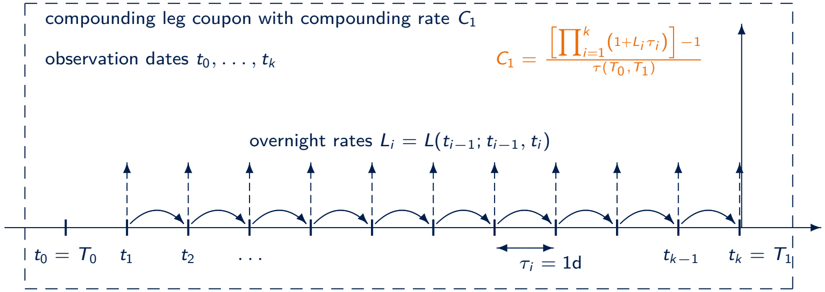
Using $1+\tau\cdot L^\delta(t,T_0)=\frac{P(t,T_0)}{P(t,T_1)}\cdot D(T_0,T_1)$ (deterministic spread assumption) yields

$$\begin{aligned}V(t) &= P(t,T_0)\cdot\tau\cdot\left[L^\delta(t,T_0)-K\right]\cdot\left[\frac{P(t,T_0)}{P(t,T_1)}\cdot D(T_0,T_1)\right]^{-1}\\ &= P(t,T_0)\cdot\frac{\left[L^\delta(t,T_0)-K\right]\cdot\tau}{1+\tau\cdot L^\delta(t,T_0)}.\end{aligned}$$

Outline

Overnight index swap (OIS) instruments are further relevant instruments in the market





We need to calculate the compounding leg coupon rate

- Assume overnight rate $L_i = L(t_{i-1}; t_{i-1}, t_i)$ is a credit-risk free Libor rate. In practice often simply called risk-free rate (RFR)
- Compounded rate (for a period $[T_0, T_1]$) is specified as

$$C_1 = \left\{ \left[\prod_{i=1}^{k} (1 + L_i \tau_i) \right] - 1 \right\} \frac{1}{\tau(T_0, T_1)}.$$

- Coupon payment is at T_1.
- For pricing we need to calculate

$$\begin{aligned} \mathbb{E}^{T_1}\left[C_1 \,|\, \mathcal{F}_t\right] &= \mathbb{E}^{T_1}\left[\left\{ \left[\prod_{i=1}^{k} (1 + L_i \tau_i) \right] - 1 \right\} \frac{1}{\tau(T_0, T_1)} \,|\, \mathcal{F}_t \right] \\ &= \left\{ \mathbb{E}^{T_1}\left[\prod_{i=1}^{k} (1 + L_i \tau_i) \,|\, \mathcal{F}_t \right] - 1 \right\} \frac{1}{\tau(T_0, T_1)}. \end{aligned}$$

How do we handle the compounding term?

Overall compounding term is

$$\prod_{i=1}^{k}(1+L_i\tau_i)=\prod_{i=1}^{k}\left[1+L(t_{i-1};t_{i-1},t_i)\tau_i\right].$$

Individual compounding term is

$$1+L(t_{i-1};t_{i-1},t_i)\tau_i=1+\left[\frac{P(t_{i-1},t_{i-1})}{P(t_{i-1},t_i)}-1\right]\frac{1}{\tau_i}\tau_i=\frac{P(t_{i-1},t_{i-1})}{P(t_{i-1},t_i)}.$$

We get

$$\prod_{i=1}^{k}(1+L_i\tau_i)=\prod_{i=1}^{k}\frac{P(t_{i-1},t_{i-1})}{P(t_{i-1},t_i)}=\prod_{i=1}^{k}\frac{1}{P(t_{i-1},t_i)}.$$

We need to calculate the expectation of $\prod_{i=1}^{k}\frac{1}{P(t_{i-1},t_i)}$.

Expected compounding factor can easily be calculated

Lemma (Compounding rate)

Consider a compounding coupon period $[T_0, T_1]$ with overnight observation and maturity dates $\{t_0, t_1, \ldots, t_k\}$, $t_0 = T_0$ and $t_k = T_1$. Then

$$\mathbb{E}^{T_1}\left[\prod_{i=1}^{k} \frac{1}{P(t_{i-1}, t_i)} \,|\, \mathcal{F}_{T_0}\right] = \frac{1}{P(T_0, T_1)}.$$

For the proof we use the notation $\mathbb{E}^{T_1}[\cdot \,|\, \mathcal{F}_t] = \mathbb{E}_t^{T_1}[\cdot]$.

We proof the result via Tower Law of conditional expectation

$$
\begin{aligned}
\mathbb{E}^{T_1}_{T_0}\left[\prod_{i=1}^{k}\frac{1}{P(t_{i-1},t_i)}\right] &= \mathbb{E}^{T_1}_{T_0}\left[\mathbb{E}^{T_1}_{t_{k-2}}\left[\prod_{i=1}^{k}\frac{1}{P(t_{i-1},t_i)}\right]\right] \\
&= \mathbb{E}^{T_1}_{T_0}\left[\prod_{i=1}^{k-1}\frac{1}{P(t_{i-1},t_i)}\,\mathbb{E}^{T_1}_{t_{k-2}}\left[\frac{P(t_{k-1},t_{k-1})}{P(t_{k-1},t_k)}\right]\right] \\
&= \mathbb{E}^{T_1}_{T_0}\left[\prod_{i=1}^{k-1}\frac{1}{P(t_{i-1},t_i)}\,\frac{P(t_{k-2},t_{k-1})}{P(t_{k-2},t_k)}\right] \\
&= \mathbb{E}^{T_1}_{T_0}\left[\prod_{i=1}^{k-2}\frac{1}{P(t_{i-1},t_i)}\,\frac{1}{P(t_{k-2},t_k)}\right] \\
\dots &= \mathbb{E}^{T_1}_{T_0}\left[\frac{1}{P(t_0,t_k)}\right] \\
&= \frac{1}{P(T_0,T_1)}.
\end{aligned}
$$

Expected compounding rate equals Libor rate

- Expected compounding rate as seen at start date T_0 becomes

$$\mathbb{E}^{T_1}\left[C_1 \,|\, \mathcal{F}_{T_0}\right] = \left[\frac{1}{P(T_0, T_1)} - 1\right] \frac{1}{\tau(T_0, T_1)} = L(T_0; T_0, T_1).$$

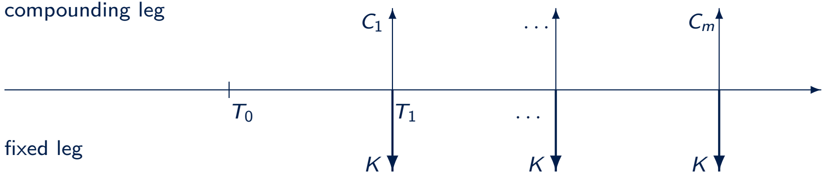
- Consequently, expected compounding rate equals Libor rate for full period.
- Moreover, expectations as seen of time-t are

$$\mathbb{E}^{T_1}\left[\prod_{i=1}^{k} \frac{1}{P(t_{i-1}, t_i)} \,|\, \mathcal{F}_t\right] = \frac{P(t, T_0)}{P(t, T_1)}$$

and

$$\mathbb{E}^{T_1}\left[C_1 \,|\, \mathcal{F}_t\right] = \left[\frac{P(t, T_0)}{P(t, T_1)} - 1\right] \frac{1}{\tau(T_0, T_1)} = L(t; T_0, T_1).$$

Compounding swap pricing is analogous to Vanilla swap pricing



$$V(t) = \sum_{j=1}^{m} N \cdot \mathbb{E}^{T_j}\left[C_j \mid \mathcal{F}_t\right] \cdot \tau_j \cdot P(t, T_j) - \sum_{j=1}^{m} N \cdot K \cdot \tau_j \cdot P(t, T_j)$$

$$= \sum_{j=1}^{m} N \cdot L(t; T_{j-1}, T_j) \cdot \tau_j \cdot P(t, T_j) - \sum_{j=1}^{m} N \cdot K \cdot \tau_j \cdot P(t, T_j).$$

Outline

As a summary we give an overview of linear products pricing

Vanilla (Payer) Swap

$$\mathrm{Swap}(t) = \underbrace{\sum_{j=1}^{m} N \cdot L^{\delta}(t, \tilde{T}_{j-1}) \cdot \tilde{\tau}_j \cdot P(t, \tilde{T}_j)}_{\text{float leg}} - \underbrace{\sum_{i=1}^{n} N \cdot K \cdot \tau_i \cdot P(t, T_i)}_{\text{fixed Leg}}$$

Market Forward Rate Agreement (FRA)

$$\mathrm{FRA}(t) = \underbrace{P(t, T_0)}_{\text{discounting to } T_0} \cdot \underbrace{\left[L^{\delta}(t, T_0) - K\right] \cdot \tau}_{\text{payoff}} \cdot \underbrace{\frac{1}{1 + \tau \cdot L^{\delta}(t, T_0)}}_{\text{discounting from } T_0 \text{ to } T_0+\delta}$$

Compounding Swap / OIS Swap

$$\mathrm{CompSwap}(t) = \underbrace{\sum_{j=1}^{m} N \cdot L(t; T_{j-1}, T_j) \cdot \tau_j \cdot P(t, T_j)}_{\text{compounding leg}} - \underbrace{\sum_{j=1}^{m} N \cdot K \cdot \tau_j \cdot P(t, T_j)}_{\text{fixed leg}}$$

Further reading on yield curves, conventions and linear products

- F. Ametrano and M. Bianchetti. Everything you always wanted to know about Multiple Interest Rate Curve Bootstrapping but were afraid to ask (April 2, 2013).
  Available at SSRN: http://ssrn.com/abstract=2219548 or http://dx.doi.org/10.2139/ssrn.2219548, 2013

- M. Henrard. Interest rate instruments and market conventions guide 2.0.
  Open Gamma Quantitative Research, 2013

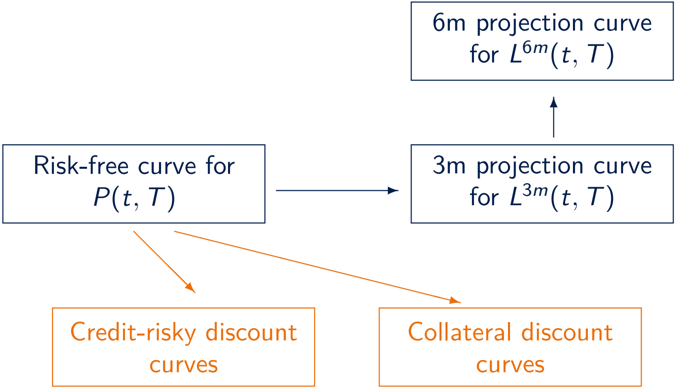
- P. Hagan and G. West. Interpolation methods for curve construction.
  Applied Mathematical Finance, 13(2):89–128, 2006

On current discussion of Libor alternatives, e.g.

- M. Henrard. A quant perspective on ibor fallback proposals.
  `https://ssrn.com/abstract=3226183`, 2018

Outline

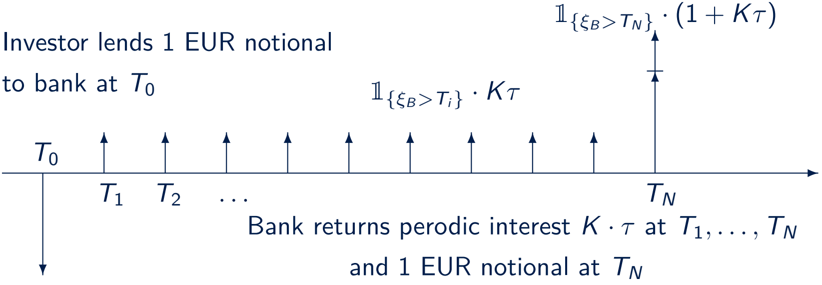
So far we discussed risk-free discount curves and tenor forward curves - now it is getting a bit more complex



Specifying appropriate discount and projection curves for a financial instrument is an important task in practice.

Outline

Discounting of bond or loan cash flows is subject to credit risk



- Cash flows are paid only if no default occurs.
- Denote $\mathbb{1}_D$ the indicator function for an event D and random variable ξ_B the first time bank defaults.
- Assume independence of credit event $\{\xi_B > T\}$ and interest rate market

We repeat credit-risky valuation from multi-curve pricing

Consider an observation time t with $T_0 < t \leq T_N$ then present value of bond cash flows becomes

$$\frac{V(t)}{B(t)} = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\xi_B > T_N\}}\frac{1}{B(T_N)} + \sum_{T_i \geq t}\mathbb{1}_{\{\xi_B > T_i\}}\frac{K\tau}{B(T_i)} \,\middle|\, \mathcal{F}_t\right].$$

Independence of credit event $\{\xi_B > T\}$ and interest rate market yields (all expectations conditional on $\mathcal{F}_t$)

$$\frac{V(t)}{B(t)} = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\xi_B > T_N\}}\right]\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{B(T_N)}\right] + \sum_{T_i \geq t}\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\xi_B > T_i\}}\right]\mathbb{E}^{\mathbb{Q}}\left[\frac{K\tau}{B(T_i)}\right].$$

Denote survival probability $Q(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\xi_B > T\}} \,\middle|\, \mathcal{F}_t\right]$ and change to forward measure, then

$$V(t) = Q(t,T_N)P(t,T_N) + \sum_{T_i \geq t} Q(t,T_i)P(t,T_i)K\tau.$$

Survival probabilities are parameterized in terms of spread curves - this leads to credit-risky discount curves

Assume survival probability $Q(t, T)$ is given in terms of a credit spread curve $s(t)$ and

$$Q(t, T) = \exp\left\{-\int_t^T s(u)du\right\}.$$

Also recall that discount factors may be represented in terms of forward rates $f(t, T)$ and

$$P(t, T) = \exp\left\{-\int_t^T f(t, u)du\right\}.$$

We may define a credit-risky discount curve $P^B(t, T)$ for a bond or loan as

$$P^B(t, T) = Q(t, T)P(t, T) = \exp\left\{-\int_t^T \left[f(t, u) + s(u)\right] du\right\}.$$

We can adapt the discounted cash flow pricing method to cash flows subject to credit risk

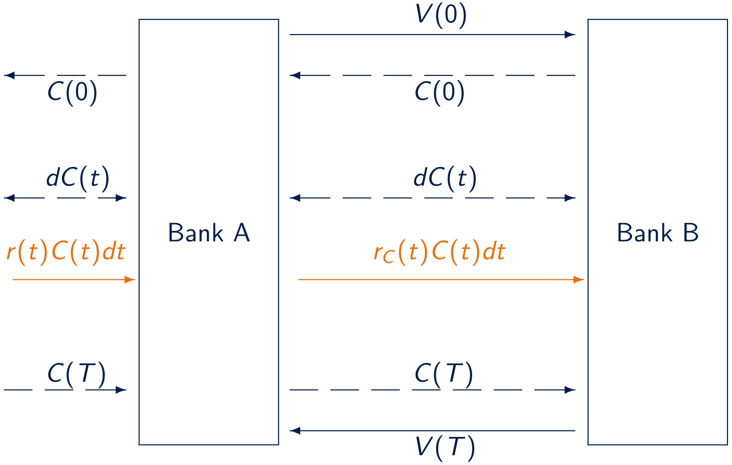
Present value of bond or loan cash flows become

$$V(t) = P^B(t, T_N) + \sum_{T_i \geq t} P^B(t, T_i) K \tau.$$

- Bonds are issued by many market participants (banks, corporates, governments, ...)
- Credit spread curves and credit-risky discount curves are specific to an issuer, e.g. Deutsche Bank has a different credit spread than Bundesrepublik Deutschland
- Many bonds are actively traded in the market. Then we may use market prices and infer credit spreads $s(t)$ and credit-risky discount curves $P^B(t, T)$

Outline

For derivative transactions credit risk is typically mitigated by posting collateral



Pricing needs to take into account interest payments on collateral.[2]

[2] Collateral amounts $C(t)$ and collateral rates are agreed in *Credit Support Annexes* (CSAs) between counterparties.

Collateralized derivative pricing takes into account collateral cash flows

Collateralized derivative price is given by (expectation of) sum of discounted payoff

$$e^{-\int_t^T r(u)du} V(T)$$

plus sum of discounted collateral interest payments

$$\int_t^T e^{-\int_t^s r(u)du} \left[r(s) - r_C(s) \right] C(s) ds.$$

That gives

$$V(t) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r(u)du} V(T) + \int_t^T e^{-\int_t^s r(u)du} \left[r(s) - r_C(s) \right] C(s) ds \,|\, \mathcal{F}_t \right].$$

Pricing is reformulated to focus on collateral rate (1/2)

From

$$V(t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(u)du} V(T) + \int_t^T e^{-\int_t^s r(u)du}\left[r(s) - r_C(s)\right] C(s)ds \,|\, \mathcal{F}_t\right]$$

we can derive:

Theorem (Collateralized Discounting)

Consider the price of an option $V(t)$ at time t which pays an amount $V(T)$ at time $T \geq t$ (and no intermediate cash flows).
The option is assumed collateralized with cash amounts $C(s)$ (for $t \leq s \leq T$). For the cash collateral a collateral rate $r_C(s)$ (for $t \leq s \leq T$) is applied.
Then the option price $V(t)$ becomes

$$\begin{aligned} V(t) &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_C(u)du} V(T) \,|\, \mathcal{F}_t\right] \\ &\quad - \mathbb{E}^{\mathbb{Q}}\left[\int_t^T e^{-\int_t^s r_C(u)du}\left[r(s) - r_C(s)\right]\left[V(s) - C(s)\right] ds \,|\, \mathcal{F}_t\right] \end{aligned}$$

Pricing is reformulated to focus on collateral rate (2/2)

For further details on collateralized discounting see, e.g.

- V. Piterbarg. Funding beyond discounting: collateral agreements and derivatives pricing.
  Asia Risk, pages 97–102, February 2010
- M. Fujii, Y. Shimada, and A. Takahashi. Collateral posting and choice of collateral currency - implications for derivative pricing and risk management (may 8, 2010).
  Available at SSRN: https://ssrn.com/abstract=1601866, May 2010

Collateralized discounting result is proved in three steps

1. Define the discounted collateralized price process

$$X(t) = e^{-\int_0^t r(u)du} V(t) + \int_0^t e^{-\int_0^s r(u)du} \left[r(s) - r_C(s) \right] C(s) ds$$

   and show that it is a martingale

2. Analyse the dynamics $dX(t)$ and deduce the dynamics for $dV(t)$

3. Solve the SDE for $dV(t)$ and calculate price via conditional expectation

Step 1 - discounted collateralized price process (1/2)

Consider $T \geq t$, then

$$\begin{aligned}
X(T) &= e^{-\int_0^T r(u)du} V(T) + \int_0^T e^{-\int_0^s r(u)du} \left[r(s) - r_C(s) \right] C(s) ds \\
&= e^{-\int_0^T r(u)du} V(T) + \int_0^t e^{-\int_0^s r(u)du} \left[r(s) - r_C(s) \right] C(s) ds + \\
&\quad \int_t^T e^{-\int_0^s r(u)du} \left[r(s) - r_C(s) \right] C(s) ds \\
&= e^{-\int_0^t r(u)du} \underbrace{\left[e^{-\int_t^T r(u)du} V(T) + \int_t^T e^{-\int_t^s r(u)du} \left[r(s) - r_C(s) \right] C(s) ds \right]}_{K(t,T)} + \\
&\quad \int_0^t e^{-\int_0^s r(u)du} \left[r(s) - r_C(s) \right] C(s) ds.
\end{aligned}$$

Step 1 - discounted collateralized price process (2/2)

We have from collateralized derivative pricing that

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[K(t,T)\,|\,\mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(u)du}V(T) + \int_t^T e^{-\int_t^s r(u)du}\left[r(s)-r_C(s)\right]C(s)ds\,|\,\mathcal{F}_t\right]\\ &= V(t).\end{aligned}$$

This yields

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[X(T)\,|\,\mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^t r(u)du}K(t,T) + \int_0^t e^{-\int_0^s r(u)du}\left[r(s)-r_C(s)\right]C(s)ds\,|\,\mathcal{F}_t\right]\\ &= e^{-\int_0^t r(u)du}\mathbb{E}^{\mathbb{Q}}\left[K(t,T)\,|\,\mathcal{F}_t\right] + \int_0^t e^{-\int_0^s r(u)du}\left[r(s)-r_C(s)\right]C(s)ds\\ &= e^{-\int_0^t r(u)du}V(t) + \int_0^t e^{-\int_0^s r(u)du}\left[r(s)-r_C(s)\right]C(s)ds\\ &= X(t).\end{aligned}$$

Thus, $X(t)$ is indeed a martingale.

Step 2 - dynamics $dX(t)$ and $dV(t)$

From $X(t) = e^{-\int_0^t r(u)du} V(t) + \int_0^t e^{-\int_0^s r(u)du} \left[r(s) - r_C(s)\right] C(s)ds$ follows

$$
\begin{aligned}
dX(t) = &-r(t)e^{-\int_0^t r(u)du} V(t)dt + e^{-\int_0^t r(u)du} dV(t) + \\
&e^{-\int_0^t r(u)du} \left[r(t) - r_C(t)\right] C(t)dt \\
&= e^{-\int_0^t r(u)du} \left[dV(t) - r(t)V(t)dt + \left[r(t) - r_C(t)\right] C(t)dt\right] \\
&= e^{-\int_0^t r(u)du} \underbrace{\left[dV(t) - r_C(t)V(t)dt + \left[r(t) - r_C(t)\right]\left[C(t) - V(t)\right] dt\right]}_{dM(t)}.
\end{aligned}
$$

Since $X(t)$ is a martingale we must have that $dM(t)$ are increments of a martingale.
We get

$$dV(t) = r_C(t)V(t)dt - \left[r(t) - r_C(t)\right]\left[C(t) - V(t)\right] dt + dM(t).$$

Step 3 - solution for $V(t)$ (1/2)

For the SDE $dV(t) = r_C(t)V(t)dt - [r(t) - r_C(t)]\,[C(t) - V(t)]\,dt + dM(t)$ we may guess a solution as

$$V(t) = e^{\int_{t_0}^{t} r_C(s)ds} V(t_0) - \int_{t_0}^{t} e^{\int_s^t r_C(u)du} \{[r(s) - r_C(s)]\,[C(s) - V(s)]\,ds - dM(s)\}$$

Differentiating confirms that

$$\begin{aligned}
dV(t) = {} & r_C(t) e^{\int_{t_0}^{t} r_C(s)ds} V(t_0) \\
& - r_C(t) \int_{t_0}^{t} e^{\int_s^t r_C(u)du} \{[r(s) - r_C(s)]\,[C(s) - V(s)]\,ds - dM(s)\} \\
& - e^{\int_t^t r_C(u)du} \{[r(t) - r_C(t)]\,[C(t) - V(t)]\,dt - dM(t)\} \\
= {} & r_C(t) \left[e^{\int_{t_0}^{t} r_C(s)ds} V(t_0) - \right. \\
& \left. \int_{t_0}^{t} e^{\int_s^t r_C(u)du} \{[r(s) - r_C(s)]\,[C(s) - V(s)]\,ds - dM(s)\} \right] \\
& - [r(t) - r_C(t)]\,[C(t) - V(t)]\,dt + dM(t) \\
= {} & r_C(t)V(t) - [r(t) - r_C(t)]\,[C(t) - V(t)]\,dt + dM(t).
\end{aligned}$$

Step 3 - solution for $V(t)$ (2/2)

Substituting $t \mapsto T$ and $t_0 \mapsto t$ yields the representation

$$V(T) = e^{\int_t^T r_C(s)ds} V(t) - \int_t^T e^{\int_s^T r_C(u)du} \left\{ [r(s) - r_C(s)] \, [C(s) - V(s)] \, ds - dM(s) \right\}$$

Solving for $V(t)$ gives

$$\begin{aligned} V(t) = {} & e^{-\int_t^T r_C(s)ds} V(T) - \\ & \int_t^T e^{-\int_t^s r_C(u)du} \left\{ [r(s) - r_C(s)] \, [V(s) - C(s)] \, ds - dM(s) \right\} \end{aligned}$$

The result follows now from taking conditional expectation

$$\begin{aligned} V(t) = {} & \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_C(s)ds} V(T) - \int_t^T e^{-\int_t^s r_C(u)du} [r(s) - r_C(s)] \, [V(s) - C(s)] \, ds \,\middle|\, \mathcal{F}_t \right] \\ & + \underbrace{\mathbb{E}^{\mathbb{Q}} \left[\int_t^T e^{-\int_t^s r_C(u)du} dM(s) \,\middle|\, \mathcal{F}_t \right]}_{0} \end{aligned}$$

A very important special case arises for full collateralization

Corollary (Full collateralization)

If the collateral amount $C(s)$ equals the full option price $V(s)$ for $t \leq s \leq T$ then the derivative price becomes

$$V(t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_C(s)ds} V(T) \,|\, \mathcal{F}_t\right].$$

- Fully collateralized price is calculated analogous to uncollateralized price.
- Discount rate must be equal to the collateral rate $r_C(s)$.
- Pricing is independent of the risk-free rate $r(t)$.
- Collateral bank account $B^C(t) = \exp\left\{\int_0^t r_C(s)ds\right\}$ can be considered as numeraire in this setting

The collateralized zero coupon bond can be used to adapt DCF method to collateralized derivative pricing

Consider a fully collateralized instrument that pays $V(T) = 1$ at some time horizon T. The price $V(t)$ for $t \leq T$ is given by
$V(t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_C(s)ds} 1 \,|\, \mathcal{F}_t\right].$

Definition (Collateralized zero coupon bond)

The collateralized zero coupon bond price (or collateralized discount factor) for an observation time t and maturity $T \geq t$ is given by

$$P^C(t,T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_C(s)ds} \,|\, \mathcal{F}_t\right].$$

Consider a time horizon T and the time-t price process of a collateralized zero coupon bond $P^C(t,T)$:

- Collateralized zero coupon bond is an asset in our economy,
- price process $P^C(t,T) > 0$.

Thus collateralized zero coupon bond is a numeraire.

The collateralized zero coupon bond can be used as numeraire for pricing

Define the collateralized forward measure $\mathbb{Q}^{T,C}$ as the equivalent martingale measure with $P^C(t,T)$ as numeraire and expectation $\mathbb{E}^{T,C}[\cdot]$. The density process of $\mathbb{Q}^{T,C}$ (relative to risk-neutral measure $\mathbb{Q}$) is

$$\zeta(t) = \frac{P^C(t,T)}{B^C(t)} \cdot \frac{B^C(0)}{P^C(0,T)}.$$

This yields

$$\begin{aligned}
\mathbb{E}^{T,C}\left[V(T) \,|\, \mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{Q}}\left[\frac{\zeta(T)}{\zeta(t)} V(T) \,|\, \mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\frac{P^C(T,T)}{B^C(T)} \cdot \frac{B^C(t)}{P^C(t,T)} V(T) \,|\, \mathcal{F}_t\right] \\
&= \frac{1}{P^C(t,T)} \mathbb{E}^{\mathbb{Q}}\left[\frac{B^C(t)}{B^C(T)} \cdot V(T) \,|\, \mathcal{F}_t\right] \\
&= \frac{1}{P^C(t,T)} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_C(s)ds} V(T) \,|\, \mathcal{F}_t\right] = \frac{V(t)}{P^C(t,T)}.
\end{aligned}$$

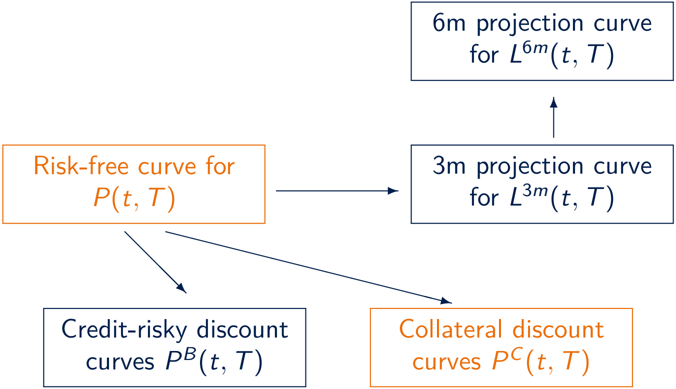
Discounted cash flow method pricing requires to use the appropriate discount curve representing collateral rates

We have

$$V(t) = P^C(t,T) \cdot \mathbb{E}^{T,C}\left[V(T) \,|\, \mathcal{F}_t\right].$$

- Requires discounting curve $P^C(t,T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_C(s)ds} \,|\, \mathcal{F}_t\right]$ capturing collateral costs and
- calculation of expected future payoffs $\mathbb{E}^{T,C}\left[V(T) \,|\, \mathcal{F}_t\right]$ in the collateralized forward measure.

We summarise the multi-curve framework widely adopted in the market



- Standard collateral curve is also considered as risk-free curve.
- In 2020 standard collateral curves move to €STR collateral rate (EUR) and SOFR collateral rate (USD).
- Projection curves are potentially not required anymore in the future if Libor (and Euribor) indices are decommissioned.

Part III

Vanilla Option Models

Outline

Outline

Vanilla Interest Rate Options

SABR Model for Vanilla Options

Summary Swaption Pricing

Outline

Now we have a first look at the cancellation option

Interbank swap deal example

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(annually, 30/360 day count, modified following, Target calendar)

Pays 6-months Euribor floating rate on 100mm EUR

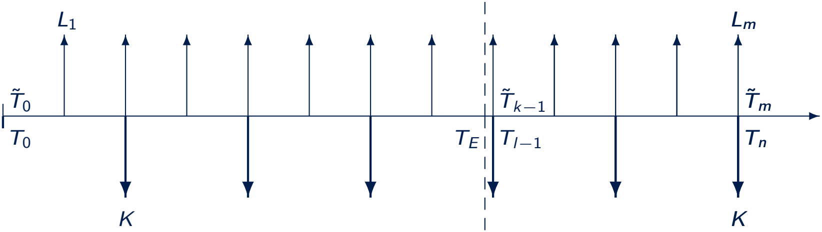
Start date: Oct 30, 2020

End date: Oct 30, 2040

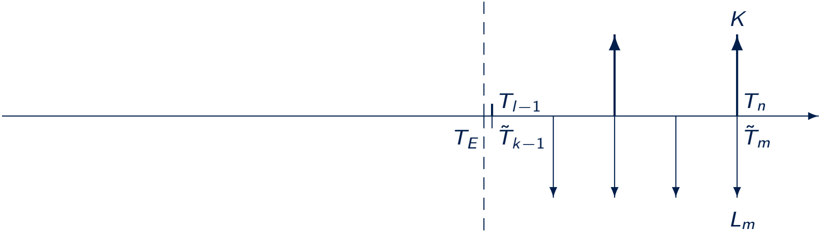
(semi-annually, act/360 day count, modified following, Target calendar)

Bank A may decide to early terminate deal in 10, 11, 12,.. years.

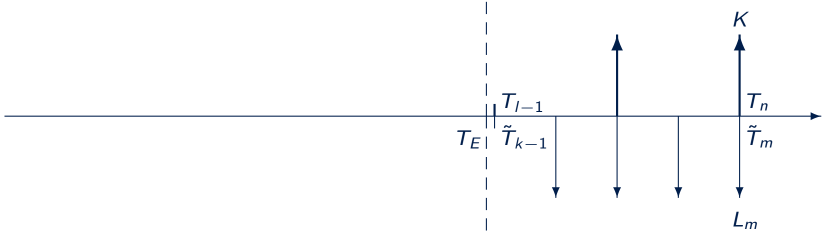
We represent cancellation as entering an opposite deal



[cancelled swap] = [full swap] + [opposite forward starting swap]



Option to cancel is equivalent to option to enter opposite forward starting swap (1/2)

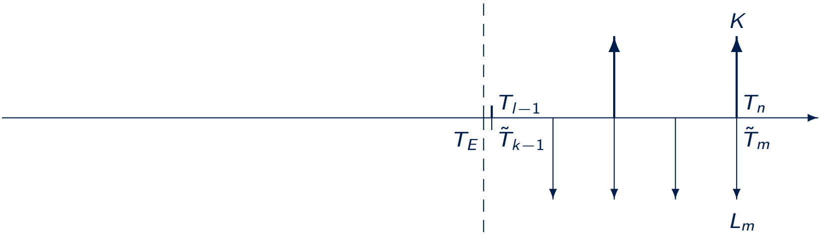


- At option exercise time T_E present value of remaining (opposite) swap is

$$V^{\text{Swap}}(T_E) = \underbrace{K \cdot \sum_{i=l}^{n} \tau_i \cdot P(T_E, T_i)}_{\text{future fixed leg}}$$

$$- \underbrace{\sum_{j=k}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta) \cdot \tilde{\tau}_j \cdot P(T_E, \tilde{T}_j)}_{\text{future float leg}}.$$

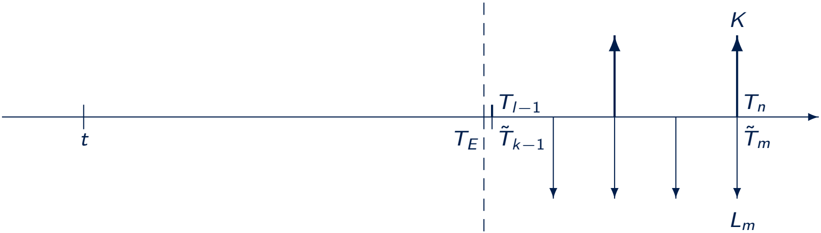
Option to cancel is equivalent to option to enter opposite forward starting swap (2/2)



- Option to enter represents the right but not the obligation to enter swap.
- Rational market participant will exercise if swap present value is positive, i.e.

$$V^{\text{Option}}(T_E) = \max\left\{V^{\text{Swap}}(T_E), 0\right\}.$$

Option can be priced via derivative pricing formula



- Using risk-neutral measure, today's present value of option is

$$V^{\text{Option}}(t) = B(t) \cdot \mathbb{E}^{\mathbb{Q}} \left[\frac{V^{\text{Option}}(T_E)}{B(T_E)} | \mathcal{F}_t \right]$$
$$= B(t) \cdot \mathbb{E}^{\mathbb{Q}} \left[\frac{\max\left\{ V^{\text{Swap}}(T_E), 0 \right\}}{B(T_E)} | \mathcal{F}_t \right].$$

- Calculation requires dynamics of future zero bonds $P(T_E, T)$ and numeraire $B(T_E)$.

Option pricing requires specific model for interest rate dynamics.

Outline

A European Swaption is an option to enter into a swap (1/2)

Physically Settled European Swaption

A physically settled European Swaption is an option with exercise time T_E. It gives the option holder the right (but not the obligation) to enter into a

- fixed rate payer (or receiver) Vanilla swap with specified
- start time T_0 and end time T_n ($T_E \leq T_0 < T_n$),
- floating rate Libor index payments $L^{\delta}(T^F_{j-1}, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)$ paid at $\tilde{T}_j$, and
- fixed rate K paid at T_i.

All properties are specified at inception of the deal.

A European Swaption is an option to enter into a swap (2/2)

At exercise time T_E swaption value or swaption payoff is

$$V^{\mathrm{Swpt}}(T_E) = \left[\phi\left(\sum_{j=0}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(T_E, \tilde{T}_j) - K\sum_{i=0}^{n}\tau_i P(T_E, T_i)\right)\right]^{+}.$$

Here $\phi = \pm 1$ is payer/receiver swaption, $[\cdot]^{+} = \max\{\cdot, 0\}$.

A European Swaption is also an option on a swap rate (1/2)

$$
\begin{aligned}
&V^{\text{Swpt}}(T_E)\\
&= \left[\phi\left(\sum_{j=0}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(T_E, \tilde{T}_j) - K\sum_{i=0}^{n}\tau_i P(T_E, T_i)\right)\right]^+\\
&= \sum_{i=0}^{n}\tau_i P(T_E, T_i)\cdot\\
&\qquad \left[\phi\left(\frac{\sum_{j=0}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(T_E, \tilde{T}_j)}{\sum_{i=0}^{n}\tau_i P(T_E, T_i)} - K\right)\right]^+.
\end{aligned}
$$

A European Swaption is also an option on a swap rate (2/2)

Float leg, annuity and swap rate

$$
\begin{aligned}
\text{float leg} \qquad Fl(T_E) &= \sum_{j=0}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(T_E, \tilde{T}_j) \\
\text{annuity} \qquad An(T_E) &= \sum_{i=0}^{n} \tau_i P(T_E, T_i) \\
\text{swap rate} \qquad S(T_E) &= \frac{\sum_{j=0}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(T_E, \tilde{T}_j)}{\sum_{i=0}^{n} \tau_i P(T_E, T_i)} \\
&= \frac{Fl(T_E)}{An(T_E)}
\end{aligned}
$$

$$V^{\mathrm{Swpt}}(T_E) = An(T_E) \cdot [\phi(S(T_E) - K)]^{+}$$

Swap rate is the key quantity for Vanilla option pricing

- Swap rate $S(T_E)$ always needs to be interpreted in the context of its underlying swap with float schedule $\{\tilde{T}_j\}_j$, Libor index rates $L^\delta(\cdot)$ and fixed schedule $\{T_i\}_i$.
- We omit swap details if underlying swap context is clear.
- Fixed rate K is the strike rate of the option.
- At-the-money strike $K = S(T_E)$ is the fair fixed rate as seen at T_E which prices underlying swap at par (i.e. zero present value).
- Float leg can be considered an asset with time-t value $(t \leq T_E)$

$$Fl(t) = \sum_{j=0}^{m} L^\delta(t, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)\tilde{\tau}_j P(t, \tilde{T}_j).$$

- Annuity can be considered a positive asset with time-t value $(t \leq T_E)$

$$An(t) = \sum_{i=0}^{n} \tau_i P(t, T_i).$$

Libor rates can be seen as one-period swap rates

- Consider single period swap rate $S(T_E)$ with $m = n = 1$ and $\tau = \tilde{\tau}_1 = \tau_1$, then

$$S(T_E) = \frac{L^\delta(T_E, \tilde{T}_0, \tilde{T}_0 + \delta)\tilde{\tau}_1 P(t, \tilde{T}_1)}{\tau_1 P(t, T_1)} = L^\delta(T_E, \tilde{T}_0, \tilde{T}_0 + \delta).$$

- Option on Libor rate $L^\delta(T_E)$ is called Caplet ($\phi = +1$) or Floorlet ($\phi = -1$) with strike K, pay time T_1 and payoff

$$\tau \cdot \left[\phi \left(L^\delta(T_E, \tilde{T}_0, \tilde{T}_0 + \delta) - K\right)\right]^+.$$

- Time-T_E price of caplet/floorlet (i.e. optionlet) is

$$V^{\mathrm{Opl}}(T_E) = \tau \cdot P(T_E, T_1) \cdot \left[\phi \left(L^\delta(T_E, \tilde{T}_0, \tilde{T}_0 + \delta) - K\right)\right]^+.$$

- Optionlet payoff is analogous to swaption payoff.

Pricing caplets and floorlets is analogous to pricing swaptions. We focus on swaption pricing.

Swap rate is a martingale in the annuity measure

Definition (Annuity measure)

Consider a swap rate $S(\cdot)$ with corresponding underlying swap details. The annuity $An(t)$ ($t \leq T_E$) is a numeraire. The annuity measure is the equivalent martingale measure corresponding to $An(t)$. Expectation under the annuity measure is denoted as $\mathbb{E}^A[\cdot]$.

Theorem (Swap rate martingale property)

The swap rate $S(t)$ is a martingale in the annuity measure and for $t \leq T \leq T_E$

$$S(t) = \mathbb{E}^A\left[S(T) \,|\, \mathcal{F}_t\right] = \frac{\sum_{j=0}^{m} L^\delta(t, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)\tilde{\tau}_j P(t, \tilde{T}_j)}{\sum_{i=0}^{n} \tau_i P(t, T_i)} = \frac{Fl(t)}{An(t)}.$$

Swap rate $S(t)$ is denoted forward swap rate.

Proof.

Annuity measure is well defined via FTAP. The swap rate $S(T) = Fl(T)/An(T)$ is a discounted asset. Thus martingale property follows directly from definition of equivalent martingale measure. □

Swaption becomes call/put option in annuity measure

$$V^{\text{Swpt}}(T_E) = An(T_E) \cdot [\phi (S(T_E) - K)]^+ .$$

Derivative pricing formula yields

$$\frac{V^{\text{Swpt}}(t)}{An(t)} = \mathbb{E}^A \left[\frac{V^{\text{Swpt}}(T_E)}{An(T_E)} \,|\, \mathcal{F}_t \right] = \mathbb{E}^A \left[[\phi (S(T_E) - K)]^+ \,|\, \mathcal{F}_t \right].$$

- $[\phi (S(T_E) - K)]^+$ is call $(\phi = +1)$ or put $(\phi = -1)$ option payoff.
- Requires modelling of terminal distribution of $S(T_E)$.
- Must comply with martingale property, i.e. $S(t) = \mathbb{E}^A [S(T_E) \,|\, \mathcal{F}_t]$.

Put-call-parity for options is an important property

We can decompose a forward payoff into a long call and a short put option

$$S(T_E) - K = [S(T_E) - K]^+ - [K - S(T_E)]^+ \, ,$$

$$\mathbb{E}^A \left[S(T_E) - K \,|\, \mathcal{F}_t \right] = \mathbb{E}^A \left[[S(T_E) - K]^+ \,|\, \mathcal{F}_t \right] - \mathbb{E}^A \left[[K - S(T_E)]^+ \,|\, \mathcal{F}_t \right] ,$$

$$\underbrace{S(t) - K}_{\text{forward minus strike}} = \underbrace{\mathbb{E}^A \left[[S(T_E) - K]^+ \,|\, \mathcal{F}_t \right]}_{\text{undiscounted call}} - \underbrace{\mathbb{E}^A \left[[K - S(T_E)]^+ \,|\, \mathcal{F}_t \right]}_{\text{undiscounted put}} .$$

Put-call-parity is a general property and not restricted to Swaptions.

General swap rate dynamics are specified by martingale representation theorem

Theorem (Swap rate dynamics)

Consider the swap rate $S(t)$ and a Brownian motion $W(t)$ in the annuity measure. There exists a volatility process $\sigma(t, \omega)$ adapted to the filtration $\mathcal{F}_t$ generated by $W(t)$ such that

$$dS(t) = \sigma(t, \omega) dW(t).$$

Proof.

$S(t)$ is a martingale in annuity measure. Thus, statement follows from martingale representation theorem. □

- Theorem provides a general framework for all swap rate models.
- Swap rate models (in annuity measure) only differ in specification of volatility function $\sigma(t, \omega)$.

We will discuss several models and their volatility specification.

Outline

Normal model is the most basic swap rate model

Assume a fixed absolute volatility parameter σ and $W(t)$ a scalar Brownian motion in annuity measure, then

$$dS(t) = \sigma \cdot dW(t).$$

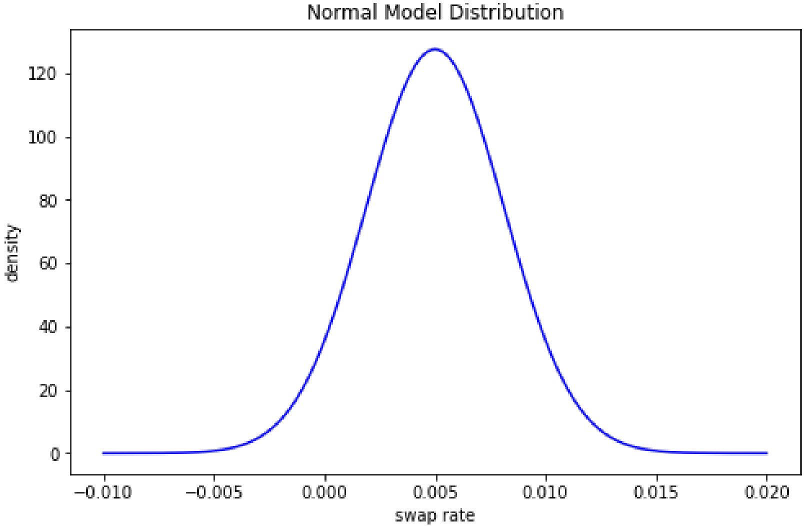
Swap rate $S(T)$ for $t \leq T$ becomes

$$S(T) = S(t) + \sigma \cdot [W(T) - W(t)] .$$

Swap rate is normally distributed with

$$S(T) \sim N\left(S(t), \sigma^2 (T - t)\right) .$$

Normal model terminal distribution of $S(T)$ for $S(0) = 0.50\%,\ T = 1,\ \sigma = 0.31\%$



Option price in normal model is calculated via Bachelier formula

Theorem (Bachelier formula)

Suppose $S(t)$ follows the normal model dynamics

$$dS(t) = \sigma \cdot dW(t).$$

Then the forward Vanilla option price becomes

$$\mathbb{E}^A \left[[\phi \left(S(T_E) - K \right)]^+ \mid \mathcal{F}_t \right] = Bachelier\left(S(t), K, \sigma\sqrt{T-t}, \phi \right)$$

with

$$Bachelier(F, K, \nu, \phi) = \nu \cdot \left[\Phi\left(\frac{\phi\,[F-K]}{\nu} \right) \cdot \frac{\phi\,[F-K]}{\nu} + \Phi'\left(\frac{\phi\,[F-K]}{\nu} \right) \right]$$

and $\Phi(\cdot)$ being the cumulated standard normal distribution function.

We derive the Bachelier formula... (1/2)

$$\mathbb{E}^A\left[[S(T_E)-K]^+ \mid \mathcal{F}_t\right] = \int_K^\infty \underbrace{[s-K]}_{\text{payoff}}$$

$$\cdot \underbrace{\frac{1}{\sqrt{2\pi\sigma^2(T-t)}}\exp\left\{-\frac{[s-S(t)]^2}{2\sigma^2(T-t)}\right\}}_{\text{density}} ds.$$

Substitute $x = [s-S(t)]/\left(\sigma\sqrt{T-t}\right)$, then

$$\mathbb{E}^A[.] = \int_{[K-S(t)]/\left(\sigma\sqrt{T-t}\right)}^\infty \left[\sigma\sqrt{T-t}x + S(t) - K\right] \underbrace{\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}}_{\Phi'(x)} dx$$

$$= \sigma\sqrt{T-t}\int_{[K-S(t)]/\left(\sigma\sqrt{T-t}\right)}^\infty \left[x + \frac{S(t)-K}{\sigma\sqrt{T-t}}\right]\Phi'(x)dx.$$

Use

$$\int x\Phi'(x)dx = -\Phi'(x).$$

We derive the Bachelier formula... (2/2)

$$\mathbb{E}^A[.] = \sigma\sqrt{T-t}\int_{[K-S(t)]/\left(\sigma\sqrt{T-t}\right)}^{\infty}\left[x + \frac{S(t)-K}{\sigma\sqrt{T-t}}\right]\Phi'(x)dx$$

$$= \sigma\sqrt{T-t}\left[-\Phi'(x) + \frac{S(t)-K}{\sigma\sqrt{T-t}}\Phi(x)\right]_{[K-S(t)]/\left(\sigma\sqrt{T-t}\right)}^{+\infty}$$

$$= \sigma\sqrt{T-t}\left[0 + \Phi'\left(\frac{K-S(t)}{\sigma\sqrt{T-t}}\right) + \frac{S(t)-K}{\sigma\sqrt{T-t}}\left[1 - \Phi\left(\frac{K-S(t)}{\sigma\sqrt{T-t}}\right)\right]\right]$$

$$= \sigma\sqrt{T-t}\left[\Phi'\left(\frac{S(t)-K}{\sigma\sqrt{T-t}}\right) + \frac{S(t)-K}{\sigma\sqrt{T-t}}\Phi\left(\frac{S(t)-K}{\sigma\sqrt{T-t}}\right)\right].$$

Log-normal model is the classical swap rate model

Assume a fixed relative volatility parameter σ and $W(t)$ a scalar Brownian motion in annuity measure, then

$$dS(t) = \sigma \cdot S(t) \cdot dW(t).$$

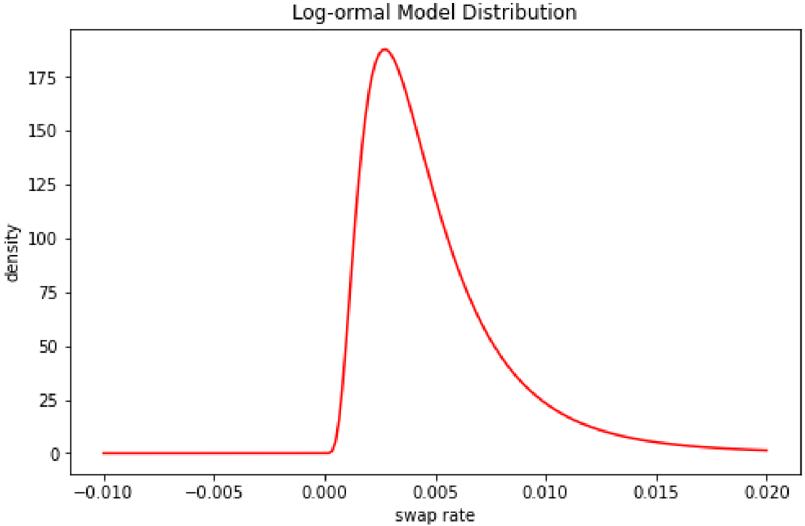
We can substitute $X(t) = \ln(S(t))$, and get with Ito formula

$$dX(t) = -\frac{1}{2}\sigma^2 \cdot dt + \sigma \cdot dW(t).$$

Log-swap rate $\ln(S(T))$ is normally distributed with

$$\ln(S(T)) \sim N\left(\ln(S(t)) - \frac{1}{2}\sigma^2 \cdot (T-t), \sigma^2(T-t)\right).$$

Log-normal model terminal distribution of $S(T)$ for $S(0) = 0.50\%,\ T = 1,\ \sigma = 63.7\%$



Option price in log-normal model is calculated via Black formula

Theorem (Black formula)

Suppose $S(t)$ follows the log-normal model dynamice

$$dS(t) = \sigma \cdot S(t) \cdot dW(t).$$

Then the forward Vanilla option price becomes

$$\mathbb{E}^A \left[[\phi (S(T_E) - K)]^+ \mid \mathcal{F}_t \right] = Black\left(S(t), K, \sigma\sqrt{T - t}, \phi \right)$$

with

$$Black(F, K, \nu, \phi) = \phi \cdot [F \cdot \Phi(\phi \cdot d_1) - K \cdot \Phi(\phi \cdot d_2)],$$
$$d_{1,2} = \frac{\ln(F/K)}{\nu} \pm \frac{\nu}{2}$$

and $\Phi(\cdot)$ being the cumulated standard normal distribution function.

Proof see exercises.

Shifted log-normal model allows *interpolating* between log normal and normal model

Assume a fixed relative volatility parameter σ, a positive shift parameter λ and a scalar Brownian motion $W(t)$ in annuity measure, then

$$dS(t) = \sigma \cdot [S(t) + \lambda] \cdot dW(t).$$

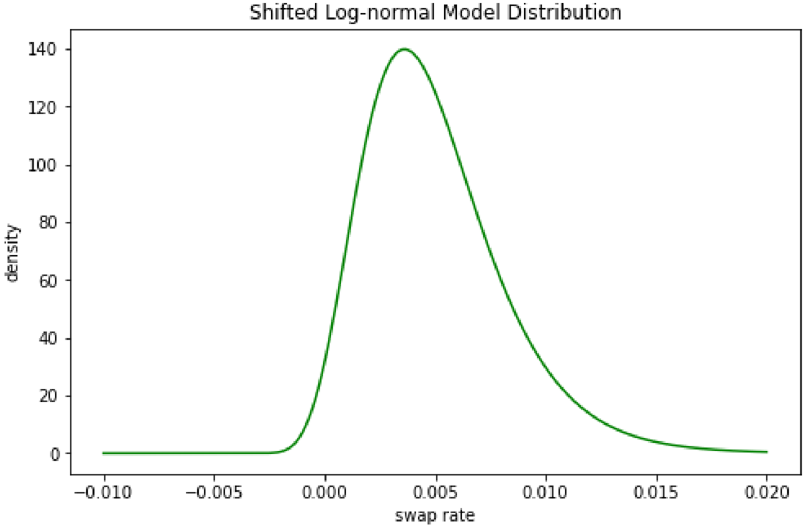
We can substitute $X(t) = \ln(S(t) + \lambda)$, and get with Ito formula

$$dX(t) = -\frac{1}{2}\sigma^2 \cdot dt + \sigma \cdot dW(t).$$

Log of shifted swap rate $\ln(S(T) + \lambda)$ is normally distributed with

$$\ln(S(T) + \lambda) \sim N\left(\ln(S(t) + \lambda) - \frac{1}{2}\sigma^2 \cdot (T - t), \sigma^2 (T - t)\right).$$

Shifted log-normal model terminal distribution of $S(T)$ for $S(0) = 0.50\%,\ T = 1,\ \lambda = 0.5\%\ \sigma = 31.5\%$



In general option pricing formula in shifted model can be obtain via un-shifted pricing formula

Theorem (Shifted model pricing formula)

Suppose an underlying process $S(t)$ *with a Vanilla call option pricing formula* $\mathbb{E}\left[(S(T)-K)^+ \mid \mathcal{F}_t\right] = V\left(S(t), K\right)$. *For a shift parameter* λ *and a shifted underlying process* $\tilde{S}(t)$ *with*

$$\tilde{S}(t) = S(t) - \lambda$$

we get the Vanilla call option pricing formula

$$\mathbb{E}\left[\left(\tilde{S}(T)-K\right)^+ \mid \mathcal{F}_t\right] = V\left(\tilde{S}(t)+\lambda, K+\lambda\right).$$

The same result holds for put option.

We prove shifted model pricing formula

Proof.

With $\tilde{S}(t) = S(t) - \lambda$ we get

$$\begin{aligned}
\mathbb{E}\left[\left(\tilde{S}(T) - K\right)^{+} \mid \mathcal{F}_t\right] &= \mathbb{E}\left[\left(S(T) - [K+\lambda]\right)^{+} \mid \mathcal{F}_t\right] \\
&= V\left(S(t), K+\lambda\right) \\
&= V\left(\tilde{S}(t)+\lambda, K+\lambda\right)
\end{aligned}$$

$\square$

- Shifted pricing formula result is model-independent.
- We will apply it to several model.

Now we can apply the previous result to shifted log-normal model

Corollary (Shifted Black formula)

Suppose $\widetilde{S}(t)$ follows the shifted log-normal model dynamics

$$d\tilde{S}(t) = \sigma \cdot \left(\tilde{S}(t) + \lambda\right) \cdot dW(t).$$

Then the forward Vanilla option price becomes

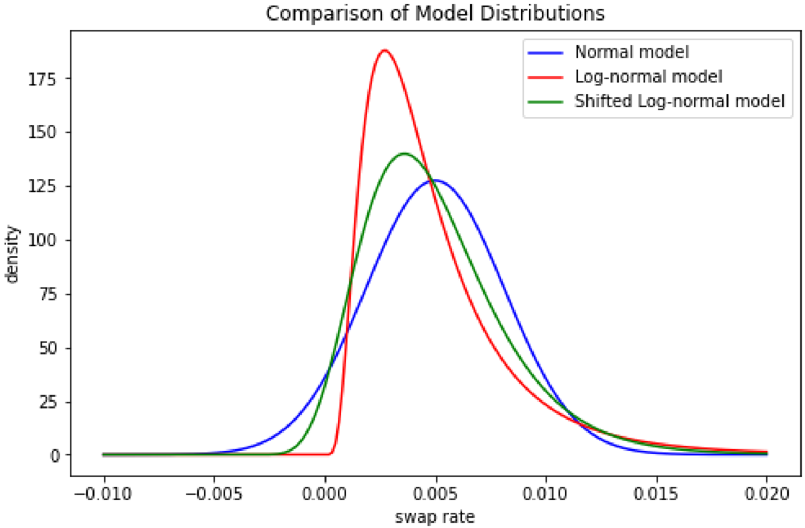
$$\mathbb{E}^A \left[\left[\phi\left(\tilde{S}(T_E) - K\right)\right]^+ \mid \mathcal{F}_t \right] = Black\left(\tilde{S}(t) + \lambda, K + \lambda, \sigma\sqrt{T-t}, \phi\right).$$

Proof.

Set $S(t) = \tilde{S}(t) + \lambda$. Then $S(T)$ is log-normally distributed and Vanilla options are priced via Black formula. Pricing formula for shifted log-normal model follows from previous theorem. $\square$

We compare the distribution examples for models calibrated to same forward ATM price

$$\mathbb{E}^A\left[[S(T) - S(t)]^+\right] = 0.125\%,\ S(0) = 0.50\%,\ T = 1,\ \lambda = 0.5\%$$



Outline

Implied Volatilities are a convenient way of representing option prices

Definition (Implied volatility)

Consider a Vanilla call ($\phi = 1$) or put option ($\phi = -1$) on an underlying $S(T)$ with strike K, and time to option expiry $T - t$. Assume that $S(t)$ is a martingale with $S(t) = \mathbb{E}\left[S(T) \,|\, \mathcal{F}_t\right]$. For a given forward Vanilla option price $V(K, T-t) = \mathbb{E}\left[(\phi\left[S(T) - K\right])^+ \,|\, \mathcal{F}_t\right]$ we define the

1. implied normal volatility σ_N such that

$$V(K, T-t) = \text{Bachelier}\left(S(t), K, \sigma_N \cdot \sqrt{T-t}, \phi\right),$$

2. implied log-normal volatility σ_{LN} such that

$$V(K, T-t) = \text{Black}\left(S(t), K, \sigma_{LN} \cdot \sqrt{T-t}, \phi\right),$$

3. implied shifted log-normal volatility σ_{SLN} for a shift parameter λ such that

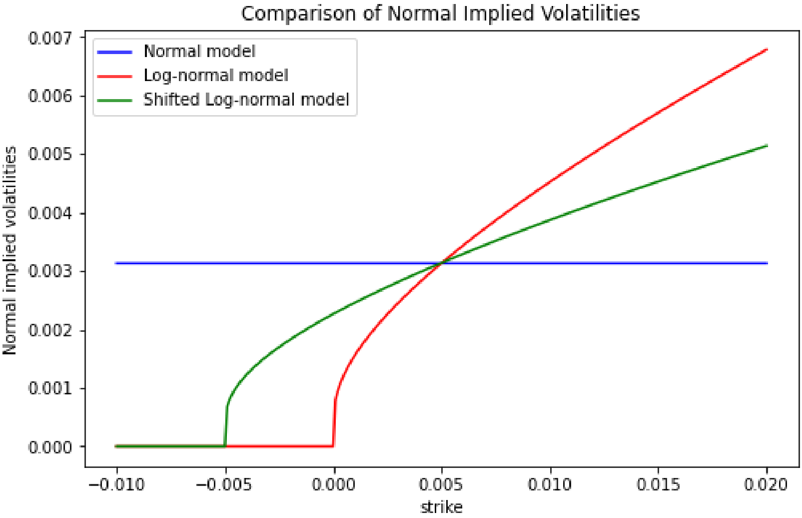
$$V(K, T-t) = \text{Black}\left(S(t) + \lambda, K + \lambda, \sigma_{SLN} \cdot \sqrt{T-t}, \phi\right).$$

We give some remarks on implied volatilities

- Implied (normal/log-normal/shifted-log-normal) volatility is only defined for attainable forward prices $V(\cdot, \cdot)$.
- Implied volatility (for swaptions) is independent from notional and annuity.
- For a given (arbitrage-free) model, implied volatilities are equal for respective call and put options.
- Typically model prices or market prices are expressed in terms of implied volatilities for comparison. This yields model-implied or market-implied volatlities.

For rates models, prices are often expressed in terms of model-implied normal volatilities

$$\mathbb{E}^A\left[[S(T)-S(t)]^+\right] = 0.125\%,\ S(0) = 0.50\%,\ T = 1,\ \lambda = 0.5\%$$



Market participants quote ATM swaptions and skew

EUR ATM Swaption Straddles - BP Volatilities (Calendar day vols)
Please call +44 (0)20 7532 3080 for further details

	1Y	2Y	3Y	4Y	5Y	7Y	10Y	15Y	20Y	25Y	30Y
1M Opt	45.3	45.0	48.0	53.8	55.3	61.6	70.1	78.6	85.2	88.7	90.0
2M Opt	38.8	40.9	44.8	48.3	51.4	58.6	67.0	76.3	82.5	84.5	85.5
3M Opt	35.6	37.3	41.7	46.8	50.9	58.3	66.7	75.0	80.5	82.5	84.1
6M Opt	34.9	37.7	42.1	46.9	51.0	59.3	66.3	74.1	78.7	80.1	81.3
9M Opt	35.4	38.0	43.1	47.3	51.5	59.1	66.9	73.8	77.5	78.7	79.0
1Y Opt	37.0	40.3	44.3	48.1	52.4	59.8	67.0	73.2	76.0	77.2	77.4
18M Opt	41.3	44.7	48.0	50.6	55.0	61.6	68.3	72.6	74.8	75.7	76.1
2Y Opt	46.5	49.4	52.6	55.0	58.2	63.9	69.8	73.0	74.2	75.1	75.5
3Y Opt	56.9	58.8	60.6	62.5	64.4	68.3	72.6	73.2	72.9	73.4	73.7
4Y Opt	64.1	65.5	66.0	67.4	68.6	71.1	73.9	72.4	71.5	71.1	71.0
5Y Opt	68.7	69.2	70.0	70.8	71.5	73.0	74.7	71.8	70.2	69.3	69.0
7Y Opt	73.0	73.3	73.6	73.8	74.1	74.5	74.8	70.1	67.6	66.4	66.0
10Y Opt	73.2	73.8	74.1	74.1	73.8	73.8	72.9	67.7	64.9	64.1	63.3
15Y Opt	70.8	71.2	71.1	71.0	70.7	69.9	68.4	62.9	59.1	57.8	56.8
20Y Opt	67.7	68.4	68.0	67.3	66.6	65.5	63.6	58.5	54.3	53.0	51.8
25Y Opt	64.6	64.8	64								
30Y Opt	60.4	60.9	59								

EUR market data as of Feb2016

EUR Vega - Normal Vol Skews

	Receivers						Payers				
	-200	-150	-100	-50	-25	ATM	+25	+50	+100	+150	+200
1y2y		22.29	14.02	5.40	1.84	40.72	0.91	4.20	13.83	24.45	
1y5y		0.20	-2.25	-2.44	-1.59	52.79	2.29	5.14	11.97	19.60	
1y10y		0.24	-1.69	-2.09	-1.40	67.86	2.10	4.80	11.53	19.32	
1y20y		13.45	7.57	2.33	0.63	76.97	0.67	2.64	9.62	18.89	
1y30y		7.75	4.34	1.43	0.46	79.14	0.16	1.00	4.59	10.15	
2y2y		11.95	6.40	1.54	0.14	49.98	1.32	3.90	11.32	19.98	
2y5y		-3.21	-3.26	-2.23	-1.28	58.62	1.61	3.52	8.09	13.38	
2y10y		-3.50	-2.97	-1.83	-1.01	70.41	1.21	2.63	6.04	10.10	
2y20y		1.10	0.20	-0.30	-0.28	75.04	0.57	1.44	4.09	7.81	
2y30y		4.86	2.51	0.58	0.07	76.50	0.46	1.48	5.08	10.29	
5y2y		-1.06	-1.41	-1.15	-0.70	69.84	0.95	2.14	5.18	8.91	
5y5y		-3.97	-2.93	-1.73	-0.94	72.02	1.11	2.39	5.42	9.00	
5y10y		-3.73	-2.55	-1.40	-0.74	75.23	0.84	1.80	4.04	6.69	
5y20y		-1.66	-1.12	-0.68	-0.38	70.67	0.49	1.10	2.72	4.85	
5y30y		-1.51	-0.99	-0.61	-0.35	69.54	0.47	1.07	2.69	4.86	
10y2y		-3.45	-2.56	-1.43	-0.75	74.34	0.83	1.74	3.79	6.11	
10y5y		-4.90	-3.28	-1.70	-0.87	74.37	0.92	1.89	4.00	6.33	
10y10y		-3.04	-1.95	-1.03	-0.54	73.36	0.60	1.29	2.91	4.89	
10y20y		-2.31	-1.32	-0.64	-0.33	65.38	0.36	0.78	1.79	3.08	
10y30y		-1.95	-1.17	-0.65	-0.36	63.77	0.46	1.02	2.53	4.54	

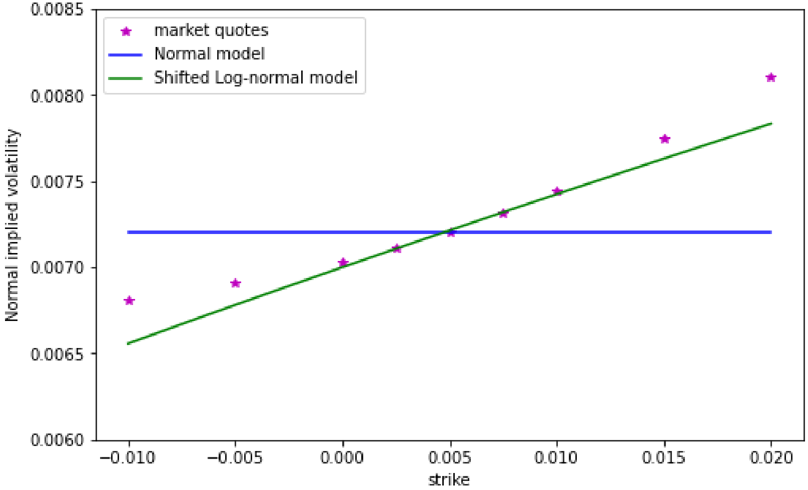
How do the market data compare to our basic swaption pricing models?

- We pick the skew data for 5y (expiry) into 5y (swap term) swaption.
- Quoted data: relative strikes and normal volatility spreads in bp:

	Receiver					Payer			
	-150	-100	-50	-25	ATM	+25	+50	+100	+150
5y5y	-3.97	-2.93	-1.73	-0.94	72.02	1.11	2.39	5.42	9.00
Vols	68.05	69.09	70.29	71.08	72.02	73.13	74.41	77.44	81.02

- Assume 5y into 5y forward swap rate $S(t)$ at 50bp (roughly corresponds to Feb'16 EUR market data).

We can fit ATM and volatility skew (i.e. slope at ATM) with a shifted log-normal model and 8% shift



However, there is no chance to fit the smile (i.e. curvature at ATM) with a basic model.

In practice Vanilla option pricing is about interpolation

Suppose we want to price a swaption with $7.6y$ expiry, on an $8y$ swap with strike 3.15%

1. Interpolate ATM volatilities in expiry dimension.
   - Typically use linear interpolation in variance $\sigma_N^2 (T - t)$.
2. Interpolate ATM volatilities in swap term dimension.
   - Typically use linear interpolation.

This yields interpolated ATM volatility σ_N^{ATM}. Then

3. Calibrate models for available skew market data.
   - We will discuss models with sufficient flexibility.
4. Interpolate smile models and combine with ATM volatility.
   - This yields a Vanilla model for the smile section $7.6y$ expiry, on an $8y$ swap term.
5. Use interpolated model to price swaption with strike 3.15%.

Outline

Vanilla Interest Rate Options

SABR Model for Vanilla Options

Summary Swaption Pricing

The SABR model was the de-facto market standard for Vanilla interest rate options until the financial crisis 2008

- Stochastic Alpha Beta Rho model is named after (some of) the parameters involved.

- Original reference is: P. Hagan, D. Kumar, A. S. Lesniewski and D. E. Woodward: *Managing Smile Risk*. Wilmott Magazine, July 2002, 86-108.

- Motivation for SABR model was less smile fit but primarily modelling smile dynamics.
  - Smile fit could (in principle) also be realised via local volatility model
  $$dS = \sigma(S) \cdot dW(t)$$
  with sufficiently complex local volatility function $\sigma(S)$.
  - We will address smile dynamics later.

- We discuss the model based on the original reference.

Outline

The SABR model extends log-normal model by local volatility term and stochastic volatility term

Swap rate dynamics in annuity meassure in SABR model are

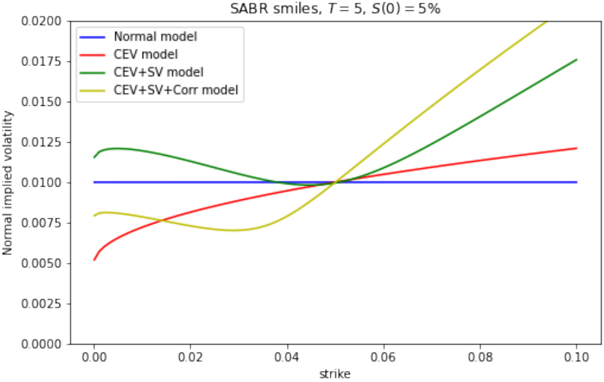
$$
\begin{aligned}
dS(t) &= \hat{\alpha}(t) \cdot S(t)^{\beta} \cdot dW(t), \\
d\hat{\alpha}(t) &= \nu \cdot \hat{\alpha}(t) \cdot dZ(t), \\
\hat{\alpha}(0) &= \alpha, \\
dW(t) \cdot dZ(t) &= \rho \cdot dt.
\end{aligned}
$$

Initial condition for $S(0)$ is given by today's yield curve.

- Elasticity parameter $\beta \in (0, 1)$ (extends local volatility).
- Stochastic volatility $\hat{\alpha}(t)$ with volatility-of-volatility $\nu > 0$ and initial condition $\alpha > 0$.
- $W(t)$ and $Z(t)$ Brownian motions, correlated via $\rho \in (-1, 1)$.

There is no analytic formula for Vanilla options. We analyse classical approximations.

First we give some intuition of the impact of the model parameters on implied volatility smile



	SABR	Normal	CEV	CEV+SV	CEV+SV+Corr
$S(t) = 5\%$	α	1.00%	4.50%	4.05%	4.20%
	β	0	0.5	0.5	0.5
$T = 5y$	ν	0	0	50%	50%
	ρ	0	0	0	70%

Outline

Approximation result is formulated for auxilliary model

Consider *a small* $\varepsilon > 0$ and a model with general local volatility function $C(S)$. Then

$$
\begin{aligned}
dS(t) &= \varepsilon \cdot \hat{\alpha}(t) \cdot C\left(S(t)\right) \cdot dW(t), \\
d\hat{\alpha}(t) &= \varepsilon \cdot \nu \cdot \hat{\alpha}(t) \cdot dZ(t).
\end{aligned}
$$

- In the original SABR model $C(S)$ is specialised to $C(S) = S^{\beta}$.
- Approximation is accurate in the order of $\mathcal{O}\left(\varepsilon^2\right)$.

Vanilla option is approximated via Bachelier formula

$$
\mathbb{E}^A\left[\left[\phi\left(S(T_E) - K\right)\right]^+ \mid \mathcal{F}_t\right] = \text{Bachelier}\left(S(t), K, \sigma_N \cdot \sqrt{T_E - t}, \phi\right).
$$

- Black formula implied log-normal volatility approximation σ_{LN} is also derived.
  - Actually, log-normal volatility approximation was primarily used.

Key aspect for us is approximation of implied normal volatility

$$\sigma_N = \sigma_N\left(S(t), K, T_E - t\right).$$

We start with the original approximation result

The approximate implied normal volatility is[3]

$$\sigma_N(S(t), K, T) = \frac{\varepsilon\alpha\,(S(t) - K)}{\int_K^{S(t)} \frac{dx}{C(x)}} \cdot \frac{\zeta}{\hat{\chi}(\zeta)} \cdot \left[1 + I^1(S_{av}) \cdot \varepsilon^2 T\right]$$

with

$$S_{av} = \sqrt{S(t) \cdot K}, \quad \zeta = \frac{\nu}{\alpha} \cdot \frac{S(t) - K}{C(S_{av})}, \quad \hat{\chi}(\zeta) = \ln\left(\frac{\sqrt{1 - 2\rho\zeta + \zeta^2} - \rho + \zeta}{1 - \rho}\right),$$

$$I^1(S_{av}) = \frac{2\gamma_2 - {\gamma_1}^2}{24}\alpha^2 C(S_{av})^2 + \frac{\rho\nu\alpha\gamma_1}{4} C(S_{av}) + \frac{2 - 3\rho^2}{24}\nu^2,$$

$$\gamma_1 = \frac{C'(S_{av})}{C(S_{av})}, \quad \gamma_2 = \frac{C''(S_{av})}{C(S_{av})}$$

There are some difficulties with above formula which we discuss subsequently.

[3] Eg. A.59 in Hagen et.al, 2002.

We adapt the original approximation result

Geometric average $S_{av} = \sqrt{S(t) \cdot K}$

- Inspiried by assumption that rates are more log-normal than normal.
- Not applicable if forward rate $S(t)$ or strike K is negative, we use arithmetic average
$$S_{av} = [S(t) + K]/2.$$
- Arithmetic average is also suggested as viable alternative in Hagan et al., 2002.

Approximation for $\zeta = \nu/\alpha \cdot [S(t) - K]/C(S_{av})$

- Eq. (A.57c) in Hagan et.al., 2002 specifies
$$\zeta = \frac{\nu}{\alpha} \int_K^{S(t)} \frac{dx}{C(x)} \approx \frac{\nu}{\alpha} \cdot \frac{S(t) - K}{C(S_{av})}.$$
- We use integral representation; consistent with an improved SABR approximation[4] .

[4] See J. Obloj, *Fine-tune your smile*. Imperial College working paper. 2008

Adapting the ζ term allows simplifying the volatility formula

With

$$\zeta = \frac{\nu}{\alpha} \int_K^{S(t)} \frac{dx}{C(x)}$$

we get

$$\begin{aligned}
\sigma_N\left(S(t), K, T\right) &= \frac{\varepsilon\alpha\left(S(t) - K\right)}{\int_K^{S(t)} \frac{dx}{C(x)}} \cdot \frac{\zeta}{\hat{\chi}(\zeta)} \cdot \left[1 + I^1\left(S_{av}\right) \cdot \varepsilon^2 T\right] \\
&= \frac{\varepsilon\alpha\left(S(t) - K\right)}{\int_K^{S(t)} \frac{dx}{C(x)}} \cdot \frac{\frac{\nu}{\alpha} \int_K^{S(t)} \frac{dx}{C(x)}}{\hat{\chi}(\zeta)} \cdot \left[1 + I^1\left(S_{av}\right) \cdot \varepsilon^2 T\right] \\
&= \nu \cdot \frac{\varepsilon\left(S(t) - K\right)}{\hat{\chi}(\zeta)} \cdot \left[1 + I^1\left(S_{av}\right) \cdot \varepsilon^2 T\right].
\end{aligned}$$

Further, we set $\varepsilon = 1$, i.e. omit small time expansion.

This yields normal volatility SABR approximation

SABR model normal volatility $\sigma_N(S, K, T)$

The approximated implied normal volatility $\sigma_N(K, T)$ in the SABR model with general local volatility function $C(S)$ is given by

$$\sigma_N\left(S(t), K, T\right) = \nu \cdot \frac{S(t) - K}{\hat{\chi}(\zeta)} \cdot \left[1 + I^1\left(S_{av}\right) \cdot T\right]$$

with

$$S_{av} = \frac{S(t) + K}{2}, \quad \zeta = \frac{\nu}{\alpha} \cdot \int_K^{S(t)} \frac{dx}{C(x)}, \quad \hat{\chi}(\zeta) = \ln\left(\frac{\sqrt{1 - 2\rho\zeta + \zeta^2} - \rho + \zeta}{1 - \rho}\right),$$

$$I^1(S_{av}) = \frac{2\gamma_2 - \gamma_1^2}{24}\alpha^2 C(S_{av})^2 + \frac{\rho\nu\alpha\gamma_1}{4} C(S_{av}) + \frac{2 - 3\rho^2}{24}\nu^2,$$

$$\gamma_1 = \frac{C'(S_{av})}{C(S_{av})}, \quad \gamma_2 = \frac{C''(S_{av})}{C(S_{av})}.$$

More concrete, we get with $C(S) = S^\beta$ and $\beta \in (0, 1)$

$$\zeta = \frac{\nu}{\alpha} \cdot \frac{S(t)^{1-\beta} - K^{1-\beta}}{1 - \beta}, \quad \gamma_1 = \frac{\beta}{S_{av}}, \quad \gamma_2 = \frac{\beta\left(\beta - 1\right)}{S_{av}^2}.$$

SABR model ATM volatility needs special treatment

- Implementing $\sigma_N(S(t), K, T) = \nu \cdot \frac{S(t)-K}{\hat{\chi}(\zeta)} \cdot \left[1 + I^1(S_{av}) \cdot T\right]$ yields division by zero for $K = S(t)$, i.e. $\zeta = 0$.

- Use L'Hôpital's rule for $\lim_{S(t)\to K}(\sigma_N(S(t), K, T))$,

$$\lim_{S(t)\to K}\left(\frac{S(t)-K}{\hat{\chi}(\zeta)}\right) = \frac{1}{\left[\hat{\chi}'(\zeta)\cdot\frac{d\zeta}{dS}\right]_{S(t)=K}},$$

$$\hat{\chi}'(\zeta) = \frac{1}{\sqrt{\zeta^2 - 2\rho\zeta + 1}}, \quad \hat{\chi}'(0) = 1,$$

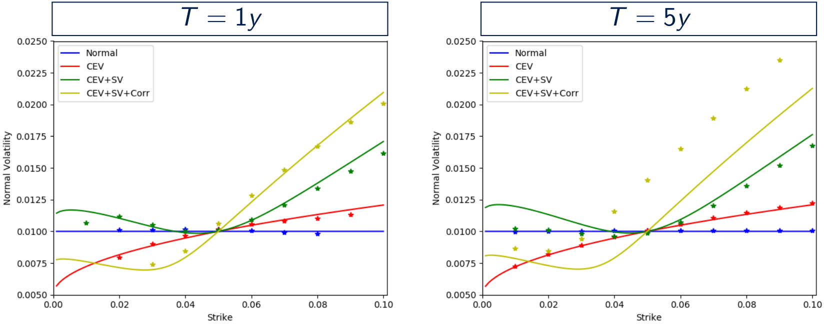
$$\left.\frac{d\zeta}{dS}\right|_{S(t)=K} = \frac{\nu}{\alpha}\cdot\frac{d}{dS}\left[\int_K^{S(t)}\frac{dx}{C(x)}\right]_{S(t)=K} = \frac{\nu}{\alpha C(S(t))}.$$

- With $\lim_{S(t)\to K} S_{av} = S(t)$ this yields ATM volatility approximation

$$\sigma_N(S(t), T) = \alpha \cdot C(S(t)) \cdot \left[1 + I^1(S(t)) \cdot T\right].$$

Outline

We compare analytic approximation (coloured lines) with Monte Carlo simulation (coloured stars)



- $S(0) = 5\%$, $\sigma_N^{ATM} = 100bp$, $\beta = 0.5$ (CEV), $\nu = 0.5$ (SV), $\rho = 0.7$ (Corr).
- 10^3 Monte Carlo paths, 100 time steps per year (stars in graphs).
- Approximation less accurate for longer maturities, low strikes, non-zero ν and ρ.

Poor approximation accuracy is less problematic in practice since SABR model is primarily used as parametric interpolation of implied volatilities.

Terminal distribution of swap rate $S(T)$ can be derived from put prices

Consider the forward put price

$$V^{\text{put}}(K) = \mathbb{E}^A\left[(K - S(T))^+\right] = \int_{-\infty}^{K} (K - s) \cdot p_{S(T)}(s) \cdot ds.$$

Here $p_{S(T)}(s)$ is the density of the terminal distribution of $S(T)$.

We get (via Leibniz integral rule)

$$\begin{aligned}
\frac{\partial}{\partial K} V^{\text{put}}(K) &= (K - K) \cdot p_{S(T)}(K) \cdot 1 - \lim_{a \downarrow -\infty} \left[(K - a) \cdot p_{S(T)}(a) \cdot 0\right] \\
&\quad + \int_{-\infty}^{K} \frac{\partial}{\partial K}\left[(K - s) \cdot p_{S(T)}(s)\right] \cdot ds \\
&= \int_{-\infty}^{K} p_{S(T)}(s) \cdot ds \; = \; \mathbb{P}^A\left\{S(T) \leq K\right\}
\end{aligned}$$

and

$$\frac{\partial^2}{\partial K^2} V^{\text{put}}(K) = p_{S(T)}(K).$$

We may also use call prices for density calculation

Recall put-call parity

$$V^{\text{call}}(K) - V^{\text{put}}(K) = \mathbb{E}^A\left[(S(T)-K)^+\right] - \mathbb{E}^A\left[(K-S(T))^+\right] = S(t) - K.$$

Differentiation yields

$$\frac{\partial}{\partial K}\left[V^{\text{call}}(K) - V^{\text{put}}(K)\right] = -1$$

and

$$\frac{\partial^2}{\partial K^2}\left[V^{\text{call}}(K) - V^{\text{put}}(K)\right] = 0.$$

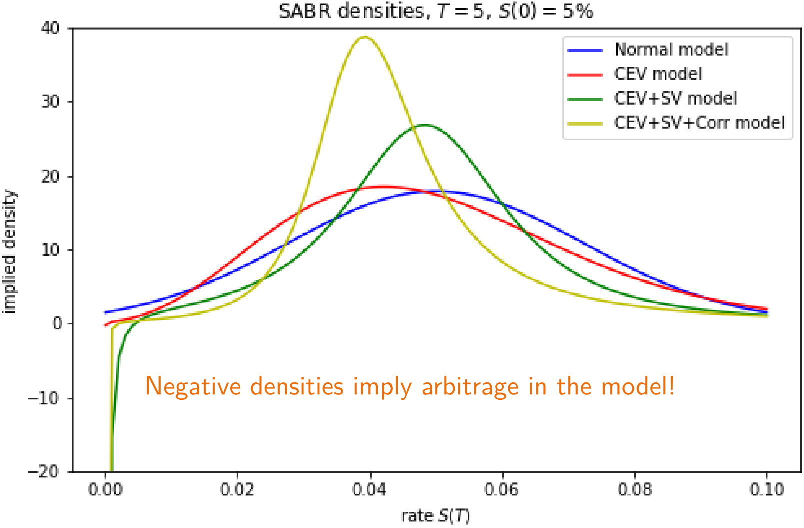
Consequently

$$\frac{\partial}{\partial K} V^{\text{call}}(K) = \frac{\partial}{\partial K} V^{\text{put}}(K) - 1 = \mathbb{P}^A\left\{S(T) \le K\right\} - 1$$

and

$$\frac{\partial^2}{\partial K^2} V^{\text{call}}(K) = \frac{\partial^2}{\partial K^2} V^{\text{put}}(K) = p_{S(T)}(K).$$

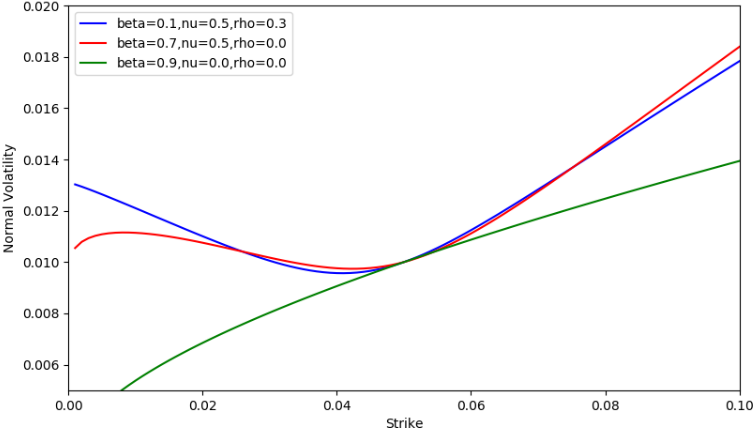
Implied Densities for example models illustrate difficulties of SABR formula for longer expiries and small strikes



Negative densities imply arbitrage in the model!

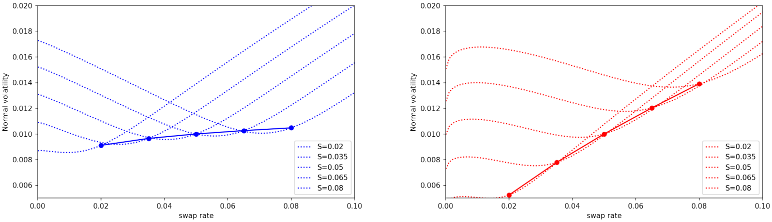
Outline

Static skew can be controlled via β and ρ



- Pure local volatility (i.e. CEV) model does not exhibit curvature.
- We can model similar skew/smile with low and high β and adjusted correlation ρ.
- What is the difference between both stochastic volatility models?

How does ATM volatility and skew/smile change if forward moves?

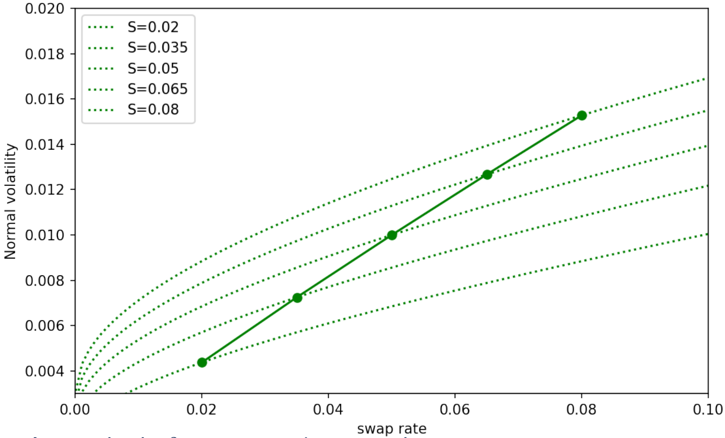


- Low $\beta = 0.1$ (left) yields horizontal shift, high $\beta = 0.7$ (right) moves smile upwards.
- Observation is consistent with expectation about *backbone function* $\sigma_N^{ATM}\left(S(t)\right)$ (solid lines in graphs),

$$\sigma_N^{ATM}\left(S(t)\right) \approx \alpha \cdot C\left(S(t)\right) = \alpha S(t)^{\beta}.$$

- β also impacts smile on the wings (i.e. low and high strikes).

What is the picture in the pure local volatility model?



- Again, high β moves smile upwards.
- Vol shape yields appearance the smile moves left if forward moves right.
- Observation is sometimes considered contradictory to market observations.

Backbone also impacts sensitivities of the option

Recall e.g. option price

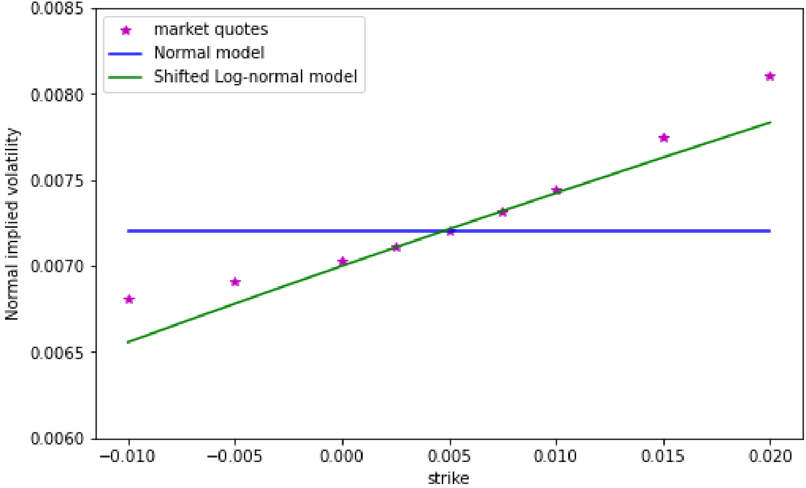
$$V(t) = \text{Bachelier}\left(S(t), K, \sigma_N\left(S(t), K, T_E\right) \cdot \sqrt{T_E}, \phi\right).$$

We get for the Delta sensitivity

$$\begin{aligned}\Delta =& \frac{dV(t)}{dS(t)} \\ =& \underbrace{\frac{\partial}{\partial S}\text{Bachelier}\left(S(t), K, \sigma_N\left(S(t), K, T_E\right) \cdot \sqrt{T_E}, \phi\right)}_{\text{Bachelier-Delta}} + \\ & \underbrace{\frac{\partial}{\partial \sigma}\text{Bachelier}\left(S(t), K, \sigma_N\left(S(t), K, T_E\right) \cdot \sqrt{T_E}, \phi\right)}_{\text{Bachelier-Vega}} \cdot \underbrace{\frac{d\sigma_N\left(S(t), K, T_E\right)}{dS}}_{\text{related to backbone slope}}.\end{aligned}$$

Outline

Recall market data example from basic Swaption pricing models



Model needs to allow negative interest rates. SABR model with $C(S) = S^{\beta}$ does not allow negative rates (unless $\beta = 0$).

Shifted SABR model allows extending the model domain to negative rates

Define $\tilde{S}(t) = S(t) - \lambda$ where $S(t)$ follows standard SABR model. Then

$$
\begin{aligned}
d\tilde{S}(t) = dS(t) &= \hat{\alpha}(t) \cdot \left[\tilde{S}(t) + \lambda\right]^{\beta} \cdot dW(t), \\
d\hat{\alpha}(t) &= \nu \cdot \hat{\alpha}(t) \cdot dZ(t), \\
\hat{\alpha}(0) &= \alpha, \\
dW(t) \cdot dZ(t) &= \rho \cdot dt.
\end{aligned}
$$

- Initial condition for $\tilde{S}(0)$ is given by today's yield curve.
- Shift parameter $\lambda \geq 0$ extends model domain to $[-\lambda, +\infty)$.
- Elasticity parameter $\beta \in (0, 1)$ (extends local volatility).
- Stochastic volatility $\hat{\alpha}(t)$ with volatility-of-volatility $\nu > 0$ and initial condition $\alpha > 0$.
- $W(t)$ and $Z(t)$ Brownian motions, correlated via $\rho \in (-1, 1)$.

We can apply SABR model pricing result to shifted local volatility function $C(S) = [S + \lambda]^{\beta}$

Vanilla option is approximated via Bachelier formula

$$\mathbb{E}^A \left[\left[\phi \left(\tilde{S}(T_E) - K \right) \right]^+ \mid \mathcal{F}_t \right]$$
$$= \text{Bachelier}\left(\tilde{S}(t), K, \sigma_N(K, T_E - t) \cdot \sqrt{T_E - t}, \phi \right)$$

and

$$\sigma_N(\tilde{S}(t), K, T) = \nu \cdot \frac{\tilde{S}(t) - K}{\hat{\chi}(\zeta)} \cdot \left[1 + I^1(S_{av}) \cdot T \right].$$

Details of normal volatility formula need to be adjusted for $C(S) = [S + \lambda]^{\beta}$ compared to $C(S) = S^{\beta}$ in original SABR model.

Shifted SABR normal volatility approximation is straight forward

Recall general approximation result

$$\sigma_N(\tilde{S}(t), K, T) = \nu \cdot \frac{\tilde{S}(t) - K}{\hat{\chi}(\zeta)} \cdot \left[1 + I^1(S_{av}) \cdot T\right]$$

with

$$S_{av} = \frac{\tilde{S}(t) + K}{2}, \quad \zeta = \frac{\nu}{\alpha} \cdot \int_K^{\tilde{S}(t)} \frac{dx}{C(x)}, \quad \hat{\chi}(\zeta) = \ln\left(\frac{\sqrt{1 - 2\rho\zeta + \zeta^2} - \rho + \zeta}{1 - \rho}\right),$$

$$I^1(S_{av}) = \frac{2\gamma_2 - \gamma_1^2}{24}\alpha^2 C(S_{av})^2 + \frac{\rho\nu\alpha\gamma_1}{4} C(S_{av}) + \frac{2 - 3\rho^2}{24}\nu^2,$$

$$\gamma_1 = \frac{C'(S_{av})}{C(S_{av})}, \quad \gamma_2 = \frac{C''(S_{av})}{C(S_{av})}$$

For shifted SABR with $C(S) = [S + \lambda]^\beta$ and $\beta \in (0, 1)$ we get

$$\zeta = \frac{\nu}{\alpha} \cdot \frac{\left[\tilde{S}(t) + \lambda\right]^{1-\beta} - [K + \lambda]^{1-\beta}}{1 - \beta}, \quad \gamma_1 = \frac{\beta}{S_{av} + \lambda}, \quad \gamma_2 = \frac{\beta(\beta - 1)}{(S_{av} + \lambda)^2}.$$

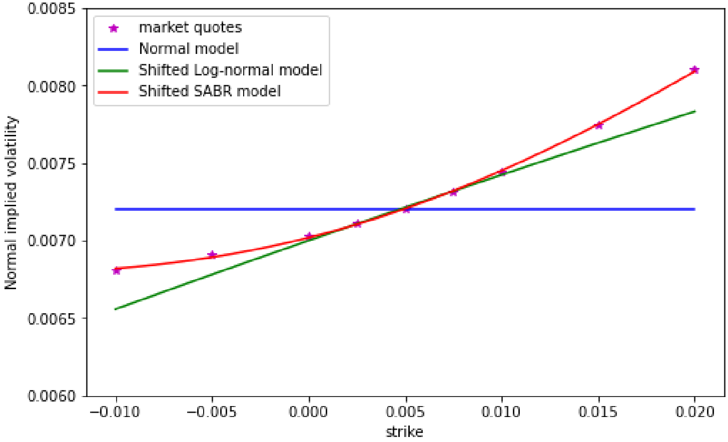
Some care is required when marking λ and β

Linearisation yields

$$\begin{aligned} C(S) &= [S+\lambda]^{\beta} \\ &\approx [S_0+\lambda]^{\beta} + \beta\,[S_0+\lambda]^{\beta-1}\,[S-S_0] \\ &= \beta\,[S_0+\lambda]^{\beta-1}\cdot\left[S+\frac{S_0+\lambda}{\beta}-S_0\right]. \end{aligned}$$

- Both λ and β impact volatility skew.
- Increasing λ is similar to decreasing β (w.r.t. skew around ATM).
- However, only λ controls domain of modelled rates.

Shifted SABR model can match example market data



- $T = 5y$, $S(t) = 0.5\%$.
- Shifted SABR: $\lambda = 5\%$, $\alpha = 5.38\%$, $\beta = 0.7$, $\nu = 23.9\%$, $\rho = -2.1\%$.

Outline

European Swaption pricing can be summarized as follows

1. Determine underlying swap start date T_0, end date T_n, schedule details and expiry date T_E.

2. Calculate annuity (as seen today), $An(t) = \sum_{i=0}^{n} \tau_i P(t, T_i)$.

3. Calculate forward swap rate (as seen today),
$S(t) = \frac{\sum_{j=0}^{m} L^{\delta}(t, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(t, \tilde{T}_j)}{\sum_{i=0}^{n} \tau_i P(t, T_i)} = \frac{Fl(t)}{An(t)}$.

4. Apply a model for the swap rate to price swaption e.g. via (shifted) SABR model, $V^{\text{Swpt}}(t) = An(t) \cdot \mathbb{E}^{A}\left[\left[\phi\left(S(T_E) - K\right)\right]^{+} \mid \mathcal{F}_t\right]$,

   4.1 determine/calibrate SABR parameters; typically depending on time to expiry $T_E - t$ and time to maturity $T_n - T_0$,

   4.2 calculate approximate normal volatility $\sigma_N\left(S(t), K, T\right)$,

   4.3 use Bachelier's formula

$$V^{\text{Swpt}}(t) = An(t) \cdot \text{Bachelier}\left(S(t), K, \sigma_N \cdot \sqrt{T_E - t}, \phi\right).$$

We illustrate Swaption pricing with QuantLib/Excel ...

Interbank swap deal example

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(annually, 30/360 day count, modified following, Target calendar)

Pays 6-months Euribor floating rate on 100mm EUR

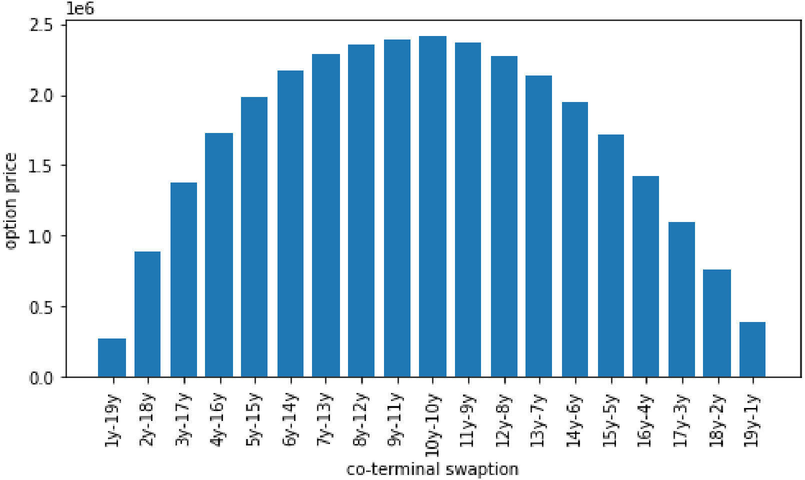
Start date: Oct 30, 2020

End date: Oct 30, 2040

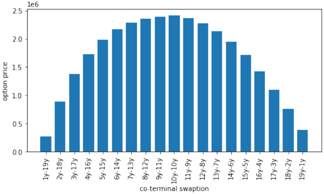
(semi-annually, act/360 day count, modified following, Target calendar)

Bank A may decide to early terminate deal in 10, 11, 12,.. years

We typically see a concave profile of European exercises



Our final swap cancellation option is related to the set of European exercise options



- Denote $V_i^{\text{Swpt}}(t)$ present value of swaption with exercise time $T_i \in \{1y, \ldots, 19y\}$.
- Denote $V^{\text{Berm}}(t)$ present value of a *Bermudan* option which allows to
  - choose any exercise time $T_i \in \{1y, \ldots, 19y\}$ and the corresponding option,
  - (as long as not exercised) postpone exercise decision on remaining options.

It follows

$$V^{\text{Berm}}(t) \geq V_i^{\text{Swpt}}(t) \quad \forall i \quad \Rightarrow \quad V^{\text{Berm}}(t) \geq \underbrace{\max_i \left\{ V_i^{\text{Swpt}}(t) \right\}}_{\text{MaxEuropean}}$$

or

$$V^{\text{Berm}}(t) = \text{MaxEuropean} + \text{SwitchOption}.$$

Further reading on Vanilla models and SABR model

- P. Hagan, D. Kumar, A. Lesniewski, and D. Woodward. Managing smile risk.
  Wilmott magazine, September 2002
- M. Beinker and H. Plank. New volatility conventions in negative interest environment.
  d-fine Whitepaper, available at www.d-fine.de, December 2012
- There are a variety of SABR extensions:
  - No-arbitrage SABR (P. Hagan et al.),
  - Free boundary SABR (A. Antonov et al.),
  - ZABR model (J. Andreasen et al.).
- Alternative local volatility-based approach:
  - D. Bang. Local-stochastic volatility for vanilla modeling.
    `https://ssrn.com/abstract=3171877`, 2018

Part IV

Term Structure Modelling

Outline

What are term structure models compared to Vanilla models?

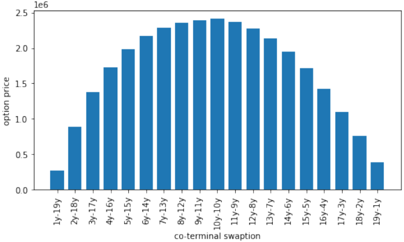
Vanilla models

- Specify dynamics for a single swap rate $S(T)$ with start/end dates T_0/T_n (and details).
- Effectively, only describes terminal distribution of $S(T)$.
- Allows pricing of European swaptions.
- Can be extended to slightly more complex options (with additional assumptions).

Term structure models

- Specify dynamics for evolution of all future zero coupon bonds $P(T, T')$ $(t \leq T \leq T')$.
- Yields (joint) distribution of *all* swap rates $S(T)$.
- Allows pricing of Bermudan swaptions and other complex derivatives.
- Typically, computationally more expensive than Vanilla model pricing.

Why do we need to model the whole term structure of interest rates?



Recall

$$V^{\text{Berm}}(t) = \text{MaxEuropean} + \text{SwitchOption}.$$

- MaxEuropean price is fully determined by Vanilla model.
- Residual SwitchOption price cannot be inferred from Vanilla model.

SwitchOption (i.e. right to postpone future exercise decisions) pricing requires modelling of full interest rate term structure.

Outline

HJM Modelling Framework

Hull-White Model

Special Topic: Options on Overnight Rates

Outline

Heath-Jarrow-Morton specify general dynamics of zero coupon bond prices

Recall our market setting with zero coupon bonds $P(t, T)$ $(t \leq T)$ and bank account $B(t) = \exp\left\{\int_0^t r(s)ds\right\}$.

Discounted bond price is martingale in risk-neutral measure.

Martingale representation theorem yields

$$d\left(\frac{P(t,T)}{B(t)}\right) = -\frac{P(t,T)}{B(t)} \cdot \sigma_P(t,T)^\top \cdot dW(t)$$

where $\sigma_P(t, T) = \sigma_P(t, T, \omega)$ is a d-dimensional process adapted to $\mathcal{F}_t$. We also impose $\sigma_P(T, T) = 0$ (pull-to-par for bond prices with $P(T, T) = 1$).

- What are dynamics of (un-discounted) zero bonds $P(t, T)$?
- What are dynamics of forward rates $f(t, T)$?
- How to specify bond price volatility?

What are dynamics of zero bonds $P(t, T)$?

Lemma (Bond price dynamics)

Under the risk-neutral measure zero bond prices evolve according to

$$\frac{dP(t, T)}{P(t, T)} = r(t) \cdot dt - \sigma_P(t, T)^\top \cdot dW(t).$$

Proof.

Apply Ito's lemma to $d\left(P(t, T)/B(t)\right)$ and compare with dynamics of discounted bond prices. $\square$

- Zero bond drift equals short rate $r(t)$.
- Zero bond volatility $\sigma_P(t, T)$ remains unchanged.
- How do we get $r(t)$?

What are dynamics of forward rates $f(t, T)$?

Theorem (Forward rate dynamics)

Consider a d-dimensional forward rate volatility process $\sigma_f(t, T) = \sigma_f(t, T, \omega)$ adapted to $\mathcal{F}_t$. Under the risk-neutral measure the dynamics of forward rates $f(t, T)$ are given by

$$df(t, T) = \sigma_f(t, T)^\top \cdot \left[\int_t^T \sigma_f(t, u) du \right] \cdot dt + \sigma_f(t, T)^\top \cdot dW(t)$$

and $f(0, T) = f^M(0, T)$. Moreover

$$\sigma_P(t, T) = \int_t^T \sigma_f(t, u) du.$$

- Once volatility $\sigma_f(t, T)$ is specified no-arbitrage pricing yields the drift.
- Model is auto-calibrated to initial yield curve via $f(0, T) = f^M(0, T)$.

We prove the forward rate dynamics (1/2)

Recall

$$f(t, T) = -\frac{\partial}{\partial T} \ln\left(P(t, T)\right).$$

Exchanging order of differentiation yields

$$df(t, T) = d\left[-\frac{\partial}{\partial T} \ln\left(P(t, T)\right)\right] = -\frac{\partial}{\partial T} d\ln\left(P(t, T)\right).$$

Applying Ito's lemma (to $d\ln\left(P(t, T)\right)$) with bond price dynamics yields

$$\begin{aligned} d\ln\left(P(t, T)\right) &= \frac{d(P(t, T))}{P(t, T)} - \frac{\sigma_P(t, T)^\top \sigma_P(t, T)}{2} \cdot dt \\ &= \left[r(t) - \frac{\sigma_P(t, T)^\top \sigma_P(t, T)}{2}\right] \cdot dt - \sigma_P(t, T)^\top \cdot dW(t). \end{aligned}$$

Differentiating $d\ln\left(P(t, T)\right)$ w.r.t. T gives

$$df(t, T) = \left[\frac{\partial}{\partial T}\sigma_P(t, T)\right]^\top \sigma_P(t, T) \cdot dt + \left[\frac{\partial}{\partial T}\sigma_P(t, T)\right]^\top \cdot dW(t).$$

We prove the forward rate dynamics (2/2)

$$df(t,T) = \left[\frac{\partial}{\partial T}\sigma_P(t,T)\right]^\top \sigma_P(t,T)\cdot dt + \left[\frac{\partial}{\partial T}\sigma_P(t,T)\right]^\top \cdot dW(t).$$

Denote

$$\sigma_f(t,T) = \frac{\partial}{\partial T}\sigma_P(t,T).$$

With terminal condition $\sigma_P(T,T) = 0$ follows integral representation

$$\sigma_P(t,T) = \int_t^T \sigma_f(t,u)du.$$

Substituting back gives the result

$$df(t,T) = \sigma_f(t,T)^\top \cdot \left[\int_t^T \sigma_f(t,u)du\right]\cdot dt + \sigma_f(t,T)^\top \cdot dW(t).$$

It will be useful to have the dynamics under the forward measure as well

Lemma (Brownian motion in T-forward measure)

Consider our HJM framework with Brownian motion $W(t)$ under the risk-neutral measure and

$$\frac{dP(t,T)}{P(t,T)} = r(t)\cdot dt - \sigma_P(t,T)^\top \cdot dW(t).$$

Under the T-forward measure the bond price dynamics are

$$\frac{dP(t,T)}{P(t,T)} = \left[r(t) + \sigma_P(t,T)^\top \sigma_P(t,T)\right]\cdot dt - \sigma_P(t,T)^\top \cdot dW^T(t)$$

with $W^T(t)$ a Brownian motion (under the T-forward measure). Moreover,

$$dW^T(t) = \sigma_P(t,T)\cdot dt + dW(t).$$

T-forward measure dynamics can be shown by Ito's lemma (1/2)

Abbrev. deflated bond prices $Y(t) = \frac{P(t,T)}{B(t)}$, then $\frac{dY(t)}{Y(t)} = -\sigma_P(t,T)^\top dW(t)$.

Now consider $1/Y(t)$ and apply Ito's lemma

$$
\begin{aligned}
d\left(\frac{1}{Y(t)}\right) &= -\frac{dY(t)}{Y(t)^2} + \frac{1}{2}\frac{2}{Y(t)^3}\left[dY(t)\right]^2 = \frac{1}{Y(t)}\left[\left(\frac{dY(t)}{Y(t)}\right)^2 - \frac{dY(t)}{Y(t)}\right] \\
&= \frac{1}{Y(t)}\left[\sigma_P(t,T)^\top \sigma_P(t,T)dt + \sigma_P(t,T)^\top dW(t)\right] \\
&= \frac{\sigma_P(t,T)^\top}{Y(t)}\left[\sigma_P(t,T)dt + dW(t)\right].
\end{aligned}
$$

T-forward measure dynamics can be shown by Ito's lemma (2/2)

However, $1/Y(t) = B(t)/P(t,T)$ is a martingale in T-forward measure and $d\left(\frac{1}{Y(t)}\right)$ must be drift-less in T-forward measure.

Define $W^T(t)$ with

$$dW^T(t) = \sigma_P(t,T)dt + dW(t).$$

Then $W^T(t)$ must be a Brownian motion in the T-forward measure.

Substituting $dW(t)$ in the risk-neutral bond price dynamics finally gives the dynamics under T-forward measure.

Outline

Short rate can be derived from forward rate dynamics

Corollary (Short rate specification)

In our HJM framework the short rate becomes

$$
\begin{aligned}
r(t) &= f(t,t) \\
&= f(0,t) + \\
&\quad \int_0^t \sigma_f(u,t)^\top \cdot \left[\int_u^t \sigma_f(u,s)ds\right] du + \int_0^t \sigma_f(u,t)^\top \cdot dW(u).
\end{aligned}
$$

Proof.

Follows directly from forward rate dynamics and integration from 0 to t. □

- Note that integrand in diffusion term $D(t) = \int_0^t \sigma_f(u,t)^\top \cdot dW(u)$ depends on t.
- In general, $D(t)$ is *not* a martingale.
- In general, $r(t)$ is *not* Markovian unless volatility $\sigma_f(t,T)$ is suitably restricted.

We analyse diffusion term in detail

$$D(t) = \int_0^t \sigma_f(u,t)^\top \cdot dW(u).$$

It follows

$$\begin{aligned}
D(T) &= \int_0^t \sigma_f(u,T)^\top \cdot dW(u) + \int_t^T \sigma_f(u,T)^\top \cdot dW(u) \\
&= D(t) + \int_t^T \sigma_f(u,T)^\top \cdot dW(u) \\
&\qquad + \int_0^t \sigma_f(u,T)^\top \cdot dW(u) - \int_0^t \sigma_f(u,t)^\top \cdot dW(u) \\
&= D(t) + \int_t^T \sigma_f(u,T)^\top \cdot dW(u) + \int_0^t \left[\sigma_f(u,T) - \sigma_f(u,t)\right]^\top \cdot dW(u).
\end{aligned}$$

- How is $\mathbb{E}^{\mathbb{Q}}\left[D(T) \mid D(t)\right]$ (knowing only last state) related to $\mathbb{E}^{\mathbb{Q}}\left[D(T) \mid \mathcal{F}_t\right]$ (knowing full history)?
- If D is Markovian then $\mathbb{E}^{\mathbb{Q}}\left[D(T) \mid D(t)\right] = \mathbb{E}^{\mathbb{Q}}\left[D(T) \mid \mathcal{F}_t\right]$ (neccessary condition).

Compare $\mathbb{E}^{\mathbb{Q}}[D(T) \mid D(t)]$ and $\mathbb{E}^{\mathbb{Q}}[D(T) \mid \mathcal{F}_t]$ (1/2)

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[D(T) \mid \mathcal{F}_t] &= \mathbb{E}^{\mathbb{Q}}\left[D(t) + \int_t^T \sigma_f(u,T)^\top dW(u) \,\Big|\, \mathcal{F}_t\right] \\
&\quad + \mathbb{E}^{\mathbb{Q}}\left[\int_0^t [\sigma_f(u,T) - \sigma_f(u,t)]^\top \, dW(u) \,\Big|\, \mathcal{F}_t\right] \\
&= D(t) + 0 + \underbrace{\int_0^t [\sigma_f(u,T) - \sigma_f(u,t)]^\top \, dW(u)}_{I(t,T)}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[D(T) \mid D(t)] &= \mathbb{E}^{\mathbb{Q}}\left[D(t) + \int_t^T \sigma_f(u,T)^\top dW(u) \,\Big|\, D(t)\right] \\
&\quad + \mathbb{E}^{\mathbb{Q}}\left[\int_0^t [\sigma_f(u,T) - \sigma_f(u,t)]^\top \, dW(u) \,\Big|\, D(t)\right] \\
&= D(t) + 0 + \mathbb{E}^{\mathbb{Q}}\left[\int_0^t [\sigma_f(u,T) - \sigma_f(u,t)]^\top \, dW(u) \,\Big|\, D(t)\right].
\end{aligned}$$

Compare $\mathbb{E}^{\mathbb{Q}}[D(T) \mid D(t)]$ and $\mathbb{E}^{\mathbb{Q}}[D(T) \mid \mathcal{F}_t]$ (2/2)

$$\mathbb{E}^{\mathbb{Q}}[D(T) \mid \mathcal{F}_t] = D(t) + \underbrace{\int_0^t [\sigma_f(u,T) - \sigma_f(u,t)]^\top \, dW(u)}_{I(t,T)}.$$

$$\mathbb{E}^{\mathbb{Q}}[D(T) \mid D(t)] = D(t) + \mathbb{E}^{\mathbb{Q}}\left[\int_0^t [\sigma_f(u,T) - \sigma_f(u,t)]^\top \, dW(u) \,\Big|\, D(t)\right].$$

- $\mathbb{E}^{\mathbb{Q}}[D(T) \mid D(t)] = \mathbb{E}^{\mathbb{Q}}[D(T) \mid \mathcal{F}_t]$ only if $I(t,T)$ is non-random or deterministic function of $D(t)$.

An important separability condition makes $D(t)$ Markovian

Assume

$$\sigma_f(t,T) = g(t) \cdot h(T)$$

with $g(t)$ (scalar) process adapted to $\mathcal{F}_t$ and $h(T)$ (scalar) deterministic and differentiable function.
Then

$$\begin{aligned} D(T) &= \int_0^t g(u) \cdot h(T) \cdot dW(u) + \int_t^T g(u) \cdot h(T) \cdot dW(u) \\ &= \frac{h(T)}{h(t)} \cdot D(t) + h(T) \cdot \int_t^T g(u) \cdot dW(u). \end{aligned}$$

Thus

$$\mathbb{E}^{\mathbb{Q}}\left[D(T) \,|\, D(t)\right] = \mathbb{E}^{\mathbb{Q}}\left[D(T) \,|\, \mathcal{F}_t\right] = \frac{h(T)}{h(t)} \cdot D(t).$$

Moreover

$$d\left(D(t)\right) = \frac{h'(t)}{h(t)} \cdot D(t) \cdot dt + g(t) \cdot h(t) \cdot dW(t).$$

Outline

We describe a very general but still tractable class of models

- We give a general description of a class of term structure models.
- Typically, these models are called Cheyette-type or quasi-Gaussian models; also associated with work by Ritchken and Sankarasubramanian (1995).
- Particular parameter choices will specialise general model to classical model (e.g. Hull-White model).
- More complex parameter choices yield powerful model instances for complex interest rate derivatives.

Quasi-Gaussian models are important models in the industry.

Separable forward rate volatility

Definition (Separable forward rate volatility)

The forward rate volatility $\sigma_f(t, T)$ of an HJM model is considered of separable form if

$$\sigma_f(t, T) = g(t)h(T)$$

for a matrix-valued process $g(t) = g(t, \omega) \in \mathbb{R}^{d \times d}$ adapted to $\mathcal{F}_t$ and a vector-valued deterministic function $h(T) \in \mathbb{R}^d$.

Corollary

For a separable forward rate volatility $\sigma_f(t, T) = g(t)h(T)$ the bond price volatility $\sigma_P(t, T)$ becomes

$$\sigma_P(t, T) = g(t) \int_t^T h(u)du.$$

Forward rate representation follows directly

Lemma

For a separable forward rate volatility $\sigma_f(t,T) = g(t)h(T)$ the forward rate becomes

$$f(t,T) = f(0,T) + h(T)^\top \int_0^t g(s)^\top g(s) \left(\int_s^T h(u)du \right) ds + h(T)^\top \int_0^t g(s)^\top dW(s)$$

and

$$r(t) = f(0,t) + h(t)^\top \left[\int_0^t g(s)^\top g(s) \left(\int_s^t h(u)du \right) ds + \int_0^t g(s)^\top dW(s) \right].$$

Proof.

Follows directly from definition. □

We need to introduce new state variables to derive Markovian representation of short rate

Re-write $h(t)^\top = \mathbf{1}^\top H(t)$ and

$$r(t) = f(0,t) + \mathbf{1}^\top H(t) \left[\int_0^t g(s)^\top g(s) \left(\int_s^t h(u) du \right) ds + \int_0^t g(s)^\top dW(s) \right]$$

with

$$\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \text{ and } H(t) = diag\left(h(t)\right) = \begin{pmatrix} h_1(t) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & h_d(t) \end{pmatrix}.$$

Introduce vector of state variables $x(t)$ with

$$x(t) = H(t) \left[\int_0^t g(s)^\top g(s) \left(\int_s^t h(u) du \right) ds + \int_0^t g(s)^\top dW(s) \right].$$

We derive the dynamics of the short rate

Theorem (Separable HJM short rate dynamics)

In an HJM model with separable volatility the short rate is given by $r(t) = f(0,t) + \mathbf{1}^\top x(t)$. *The vector of state variables* $x(t)$ *evolves according to* $x(0) = 0$ *and*

$$dx(t) = \left[y(t)\mathbf{1} - \chi(t)x(t)\right] dt + H(t)g(t)^\top dW(t)$$

with symmetric matrix of auxilliary variables $y(t)$ *as*

$$y(t) = H(t)\left(\int_0^t g(s)^\top g(s)ds\right)H(t)$$

and diagonal matrix of mean reversion parameters $\chi(t)$ *as*

$$\chi(t) = -\frac{dH(t)}{dt}H(t)^{-1}.$$

Proof follows straight forward via differentiation (1/3)

We have

$$x(t) = H(t) \underbrace{\left[\int_0^t g(s)^\top g(s) \left(\int_s^t h(u)du\right) ds + \int_0^t g(s)^\top dW(s)\right]}_{G(t)}.$$

$$\begin{aligned} dx(t) &= H'(t) \cdot G(t) \cdot dt + H(t) \cdot dG(t) \\ &= H'(t) \cdot H(t)^{-1} \cdot H(t) \cdot G(t) \cdot dt + H(t) \cdot dG(t) \\ &= -\chi(t) \cdot x(t) \cdot dt + H(t) \cdot dG(t). \end{aligned}$$

Proof follows straight forward via differentiation (2/3)

$$
\begin{aligned}
dx(t) &= -\chi(t) \cdot x(t) \cdot dt + H(t) \cdot dG(t), \\
G(t) &= \int_0^t g(s)^\top g(s) \left(\int_s^t h(u)du \right) ds + \int_0^t g(s)^\top dW(s).
\end{aligned}
$$

Leibnitz rule yields

$$
\begin{aligned}
dG(t) &= \left[g(t)^\top g(t) \left(\int_t^t h(u)du \right) + \int_0^t g(s)^\top g(s) \frac{d}{dt} \left(\int_s^t h(u)du \right) ds \right] dt \\
&\quad + g(t)^\top dW(t) \\
&= \left[0 + \int_0^t g(s)^\top g(s) \cdot H(t)\mathbf{1} \cdot ds \right] dt + g(t)^\top dW(t) \\
&= \left[\left(\int_0^t g(s)^\top g(s) ds \right) H(t)\mathbf{1} \right] dt + g(t)^\top dW(t).
\end{aligned}
$$

Proof follows straight forward via differentiation (3/3)

Combining results gives

$$\begin{aligned}
dx(t) &= -\chi(t) \cdot x(t) \cdot dt + H(t) \cdot dG(t) \\
&= \left[H(t) \left(\int_0^t g(s)^\top g(s) ds \right) H(t) \mathbf{1} - \chi(t) \cdot x(t) \right] dt \\
&\quad + H(t) \cdot g(t)^\top dW(t) \\
&= \left[y(t) \cdot \mathbf{1} - \chi(t) \cdot x(t) \right] dt + H(t) \cdot g(t)^\top dW(t).
\end{aligned}$$

- Note that $dx(t)$ depends on accumulated previous volatility via $\int_0^t g(s)^\top g(s) ds$.
- $x(t)$ is Markovian only if volatility function $g(t)$ is deterministic.
- In general, short rate dynamics can be ammended by dynamics of $y(t)$.

Short rate dynamics can be written in terms of state and auxilliary variables (1/2)

Corollary (Augmented short rate dynamics)

In an HJM model with separable volatility the short rate is given via $r(t) = f(0,t) + \mathbf{1}^\top x(t)$ *with*

$$
\begin{aligned}
dx(t) &= \left[y(t) \cdot \mathbf{1} - \chi(t) \cdot x(t)\right] dt + \sigma_r(t)^\top dW(t), \\
dy(t) &= \left[\sigma_r(t)^\top \sigma_r(t) - \chi(t) y(t) - y(t) \chi(t)\right] dt,
\end{aligned}
$$

and $x(0) = 0$, $y(0) = 0$.

Short rate dynamics can be written in terms of state and auxilliary variables (2/2)

Proof.

Set $\sigma_r(t) = g(t)H(t)$ and differentiate

$y(t) = H(t) \left(\int_0^t g(s)^\top g(s) ds \right) H(t)$. $\square$

- Model class also called Cheyette or quasi-Gaussian models.
- Typically $\sigma_r(t)$ and $\chi(t)$ are specified and $\sigma_f(t, T)$ is reconstructed via

$$H'(t) = -\chi(t)H(t), \; H(0) = 1 \quad \text{and}$$
$$g(t) = \sigma_r(t)H(t)^{-1}.$$

Forward rates and zero bonds can be written in terms of state/auxilliary variables

Theorem (Forward rate and zero bond reconstruction)

In our HJM model setting we get

$$f(t, T) = f(0, T) + \mathbf{1}^\top H(T) H(t)^{-1} \left[x(t) + y(t) G(t, T) \right]$$

and

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp \left\{ -G(t, T)^\top x(t) - \frac{1}{2} G(t, T)^\top y(t) G(t, T) \right\}$$

with

$$G(t, T) = \int_t^T H(u) H(t)^{-1} \mathbf{1} du.$$

- We prove the first part for $f(t, T)$.
- And we sketch the proof for the second part for $P(t, T)$.

We prove the first part for $f(t, T)$ (1/2)...

$$\underbrace{\mathbf{1}^\top H(T)H(t)^{-1}x(t)}_{I_1}$$
$$= h(T)^T \left[\int_0^t g(s)^\top g(s) \left(\int_s^{t} h(u)du\right) ds + \int_0^t g(s)^\top dW(s)\right].$$

$$\underbrace{\mathbf{1}^\top H(T)H(t)^{-1}y(t)G(t, T)}_{I_2}$$
$$= h(T)^\top \left(\int_0^t g(s)^\top g(s)ds\right) \int_{t}^T h(u)du.$$

We prove the first part for $f(t, T)$ (2/2)...

$$
\begin{aligned}
& I_1 + I_2 \\
& = h(T)^\top \times \\
& \quad \left[\int_0^t g(s)^\top g(s) \left(\int_s^t h(u)du \right) ds + \left(\int_0^t g(s)^\top g(s) ds \right) \int_t^T h(u)du \right] \\
& \quad + h(T)^T \int_0^t g(s)^\top dW(s) \\
& = h(T)^T \times \\
& \quad \left[\int_0^t g(s)^\top g(s) \left(\int_s^t h(u)du + \int_t^T h(u)du \right) ds + \int_0^t g(s)^\top dW(s) \right] \\
& = h(T)^T \left[\int_0^t g(s)^\top g(s) \left(\int_s^T h(u)du \right) ds + \int_0^t g(s)^\top dW(s) \right] \\
& = f(t, T) - f(0, T)
\end{aligned}
$$

... and sketch the proof for the second part for $P(t, T)$ (1/2)

$$
\begin{aligned}
P(t,T) &= \exp\left\{-\int_t^T f(t,s)ds\right\} \\
&= \exp\left\{-\int_t^T \left(f(0,s) + \mathbf{1}^\top H(s)H(t)^{-1}\left[x(t) + y(t)G(t,s)\right]\right)ds\right\} \\
&= \frac{P(0,T)}{P(0,t)} \cdot \exp\left\{-\underbrace{\left(\int_t^T \mathbf{1}^\top H(s)H(t)^{-1}ds\right)}_{G(t,T)^\top} x(t)\right\} \cdot \\
&\qquad \exp\left\{-\int_t^T \mathbf{1}^\top H(s)H(t)^{-1}y(t)G(t,s)ds\right\}
\end{aligned}
$$

... and sketch the proof for the second part for $P(t, T)$ (2/2)

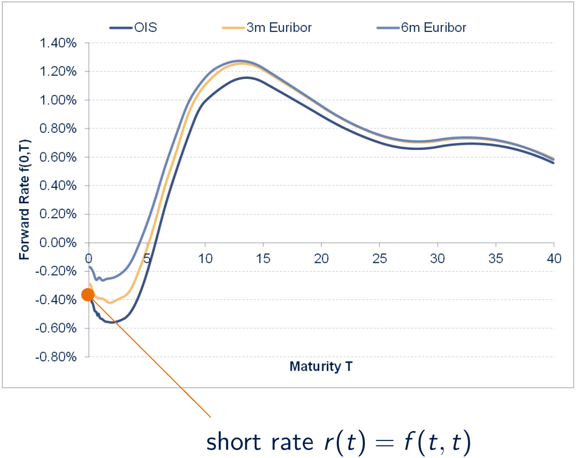
It remains to show that

$$\int_t^T \mathbf{1}^\top H(s)H(t)^{-1}y(t)G(t,s)ds = \frac{1}{2}G(t,T)^\top y(t)G(t,T).$$

We note that both sides of above equation are zero for $T = t$.
The equality for $T > t$ follows then by differentiating both sides w.r.t. T and comparing terms.

Outline

We take a complementary view to HJM framework and consider direct modelling of the short rate $r(t)$



We model short rate of the discount curve as offset point for future rates.

Short rate suffices to specify evolution of the full yield curve

Recall zero bond formula

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left\{-\int_t^T r(s)ds\right\} \mid \mathcal{F}_t\right].$$

- Once dynamics of $r(t)$ are specified all zero bonds can be derived.

Libor rates (in multi-curve setting) are

$$L(t;T_0,T_1) = \mathbb{E}^{T_1}\left[L(T;T_0,T_1) \mid \mathcal{F}_t\right] = \left[\frac{P(t,T_0)}{P(t,T_1)} \cdot D(T_0,T_1) - 1\right]\frac{1}{\tau}.$$

- With zero bonds $P(t,T)$ (and tenor basis factors $D(T_0,T_1)$) we can also derive future Libor rates.

Short rate is a natural choice of state variable for modelling evolution of interest rates.

Outline

Vasicek model and Ho-Lee model were the first models for the short rate

Vasicek (1977) assumed Ornstein-Uhlenbeck process

$$dr(t) = \kappa\,(\theta - r(t))\,dt + \sigma dW(t), \quad r(0) = r_0$$

for positive constants r_0, κ, θ, and σ.

- Model is not too different from HJM model representation.
- Constant parameters (in particular θ) limit ability to reproduce/calibrate yield curve observed today.

Ho and Lee (1986) introduce exogenous time-dependent drift parameter,

$$dr(t) = \theta(t)dt + \sigma dW(t).$$

- Drift parameter $\theta(t)$ is used to match today's zero bonds $P(0, T)$.
- Lack of mean reversion is considered main disadvantage.
- Model was historically used with binomial tree implementation.

Hull and White (1990) extended Vasicek model by $\theta(t)$

Definition (Hull-White model)

In the Hull-White model the short rate evolves according to

$$dr(t) = [\theta(t) - a(t)r(t)]\,dt + \sigma(t)dW(t)$$

with deterministic scalar functions $\theta(t)$, $a(t)$, and $\sigma(t) > 0$.

- $\theta(t)$ is mean reversion level,
- $a(t)$ is mean reversion speed, and
- $\sigma(t)$ is short rate volatility.

- Original reference is J. Hull and A. White. Pricing interest-rate-derivative securities.
  The Review of Financial Studies, 3:573–592, 1990
- To simplify analytical tractability we will assume
  - constant mean reversion speed $a(t) = a > 0$, and
  - piece-wise constant short rate volatility function on a siutable time grid $\{t_0, \ldots, t_k\}$,

$$\sigma(t) = \sum_{i=1}^{k} \mathbb{1}_{\{t_{i-1} \le t < t_i\}} \cdot \sigma_i.$$

How do we calibrate the drift $\theta(t)$?

Lemma (Hull-White drift calibration)

In the risk-neutral specification of the Hull-White model the drift term $\theta(t)$ is given by

$$\theta(t) = \frac{\partial}{\partial T} f(0,t) + a \cdot f(0,t) + \int_0^t \left[e^{-a(t-u)} \sigma(u) \right]^2 du.$$

Here $f(0,t) = f^M(0,t)$ is exogenously specified and assumed continuously differentiable w.r.t. the maturity T.

Proof follows along the following steps

- Calculate $r(s)$ via integration.
- Integrate $I(t,T) = \int_t^T r(s)ds$ and calculate distribution of $I(t,T)$.[5]
- Derive $\theta(t)$ such that $\mathbb{E}^{\mathbb{Q}}\left[e^{-I(0,t)}\right] = P(0,T)$.

[5] We will re-use distribution of integrated short rate $I(t,T)$ later for options on compounded rates.

Proof (1/4) - calculate $r(s)$

We show that for $s \geq t$

$$r(s) = e^{-a(s-t)} \left[r(t) + \int_t^s e^{a(u-t)} \left[\theta(u) du + \sigma(u) dW(u) \right] \right].$$

$$\begin{aligned} dr(s) &= -ar(s)ds + e^{-a(s-t)} \left[e^{a(s-t)} \left[\theta(s) ds + \sigma(s) dW(s) \right] \right] \\ &= \left[\theta(s) - ar(s) \right] ds + \sigma(s) dW(s). \end{aligned}$$

Use notation $[\cdot]'(t, T) = \frac{\partial}{\partial T} [\cdot]$. Set $I(t, T) = \int_t^T r(s) ds$, then $I'(t, T) = \frac{\partial I(t,T)}{\partial T} = r(T)$. We show

$$I(t, T) = G(t, T) r(t) + \int_t^T G(u, T) \left[\theta(u) du + \sigma(u) dW(u) \right]$$

with

$$G(t, T) = \int_t^T e^{-a(u-t)} du = \left[\frac{1 - e^{-a(T-t)}}{a} \right].$$

Proof (2/4) - calculate distribution $I(t, T)$

$$
\begin{aligned}
I(t,T) &= G(t,T)r(t) + \int_t^T G(u,T)\left[\theta(u)du + \sigma(u)dW(u)\right], \\
I'(t,T) &= G'(t,T)r(t) + 0 + \int_t^T G'(u,T)\left[\theta(u)du + \sigma(u)dW(u)\right] \\
&= e^{-a(T-t)}r(t) + \int_t^T e^{-a(T-u)}\left[\theta(u)du + \sigma(u)dW(u)\right] \\
&= e^{-a(T-t)}\left[r(t) + \int_t^T e^{a(u-t)}\left[\theta(u)du + \sigma(u)dW(u)\right]\right] \\
&= r(T).
\end{aligned}
$$

Conditional on $\mathcal{F}_t$, integral is normally distributed, $I(t,T)|_{\mathcal{F}_t} \sim N(\mu, \sigma^2)$ with

$$
\begin{aligned}
\mu(t,T) &= G(t,T)r(t) + \int_t^T G(u,T)\theta(u)du, \\
\sigma(t,T)^2 &= \int_t^T \left[G(u,T)\sigma(u)\right]^2 du.
\end{aligned}
$$

Proof (3/4) - calculate forward rate

$I(t,T)|_{\mathcal{F}_t} \sim N(\mu, \sigma^2)$ with

$$\mu(t,T) = G(t,T)r(t) + \int_t^T G(u,T)\theta(u)du,$$

$$\sigma^2(t,T) = \int_t^T \left[G(u,T)\sigma(u)\right]^2 du.$$

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-I(t,T)} \,|\, \mathcal{F}_t\right] = e^{-\mu(t,T)+\frac{1}{2}\sigma^2(t,T)}.$$

$$\begin{aligned}
f(t,T) &= -\frac{\partial}{\partial T}\ln\left[P(t,T)\right] = \frac{d}{dT}\left[\mu(t,T) - \frac{1}{2}\sigma^2(t,T)\right] \\
&= G'(t,T)r(t) + 0 + \int_t^T G'(u,T)\theta(u)du \\
&\quad - \frac{1}{2}\left[0 + \int_t^T 2G(u,T)G'(u,T)\sigma(u)^2 du\right] \\
&= G'(t,T)r(t) + \int_t^T G'(u,T)\theta(u)du - \int_t^T G'(u,T)G(u,T)\sigma(u)^2 du.
\end{aligned}$$

Proof (4/4) - derive drift $\theta(t)$

$$f(t,T) = G'(t,T)r(t) + \int_t^T G'(u,T)\theta(u)du - \int_t^T G'(u,T)G(u,T)\sigma(u)^2 du.$$

Use $G'(t,T) = e^{-a(T-t)}$ and $G''(t,T) = -aG'(t,T)$, then

$$\begin{aligned} f'(t,T) &= G''(t,T)r(t) + \theta(T) + \int_t^T G'(u,T)\theta(u)du - 0 \\ &\quad - \int_t^T \left[G''(u,T)G(u,T) + G'(u,T)^2\right]\sigma(u)^2 du \\ &= \theta(T) - af(t,T) - \int_t^T \left[G'(u,T)\sigma(u)\right]^2 du. \end{aligned}$$

This finally gives the result (with $t = 0$)

$$\begin{aligned} \theta(T) &= f'(t,T) + af(t,T) + \int_t^T \left[G'(u,T)\sigma(u)\right]^2 du \\ &= f'(0,T) + af(0,T) + \int_0^T \left[e^{-a(T-u)}\sigma(u)\right]^2 du. \end{aligned}$$

Do we really need the drift $\theta(t)$?

- Risk-neutral drift representation

$$\theta(t) = \frac{\partial}{\partial T} f(0,t) + a \cdot f(0,t) + \int_0^t \left[e^{-a(t-u)} \sigma(u) \right]^2 du$$

poses some obstacles.

- Derivative $\frac{\partial}{\partial T} f(0,t)$ may cause numerical difficulties.
- In some market situations you want to have jumps in $f(0,t)$.
- This is relevant in particular for the short end of OIS curve.

- Fortunately, for most applications we don't need drift term.

- HJM representation allows avoiding it alltogether.

Now we can also derive future zero bond prices I

Theorem (Zero bonds in Hull-White model)

In the Hull-White model future zero bond prices are given by

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \cdot \exp\left\{ -G(t,T)\left[r(t) - f(0,t)\right] - \frac{G(t,T)^2}{2} \int_0^t \left[e^{-a(t-u)} \sigma(u) \right]^2 du \right\}$$

with

$$G(t,T) = \int_t^T e^{-a(u-t)} du = \left[\frac{1 - e^{-a(T-t)}}{a} \right].$$

- The proof is a bit technical.
- We already derived the zero bond representation

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-I(t,T)} \mid \mathcal{F}_t \right] = e^{-\mu(t,T) + \frac{1}{2}\sigma^2(t,T)}.$$

Now we can also derive future zero bond prices II

We have from the proof of risk-neutral drift that

$$f(t,T) = G'(t,T)r(t) + \int_t^T G'(u,T)\theta(u)du - \int_t^T G'(u,T)G(u,T)\sigma^2(u)du$$

and

$$P(t,T) = e^{-G(t,T)r(t) - \int_t^T G(u,T)\theta(u)du + \frac{1}{2}\int_t^T G(u,T)^2\sigma^2(u)du}.$$

We aim at calculating the term

$$I(t,T) = -\int_t^T G(u,T)\theta(u)du + \frac{1}{2}\int_t^T G(u,T)^2\sigma^2(u)du.$$

Now we can also derive future zero bond prices III

Consider

$$
\begin{aligned}
&\log\left(\frac{P(0,t)}{P(0,T)}\right) \\
&= \left[G(0,T) - G(0,t)\right] r(0) \\
&\quad + \int_0^T G(u,T)\theta(u)du - \int_0^t G(u,t)\theta(u)du \\
&\quad - \frac{1}{2}\left[\int_0^T G(u,T)^2\sigma^2(u)du - \int_0^t G(u,t)^2\sigma^2(u)du\right] \\
&= \left[G(0,T) - G(0,t)\right] r(0) \\
&\quad + \int_t^T G(u,T)\theta(u)du + \int_0^t \left[G(u,T) - G(u,t)\right]\theta(u)du \\
&\quad - \frac{1}{2}\left[\int_t^T G(u,T)^2\sigma^2(u)du + \int_0^t \left[G(u,T)^2 - G(u,t)^2\right]\sigma^2(u)du\right].
\end{aligned}
$$

Now we can also derive future zero bond prices IV

We use $G(u,T) - G(u,t) = G(t,T)G'(u,t)$ and re-arrange terms. Then

$$
\begin{aligned}
I(t,T) = \log\left(\frac{P(0,T)}{P(0,t)}\right) &+ G(t,T)G'(0,t)r(0) \\
&+ G(t,T)\int_0^t G'(u,t)\theta(u)du \\
&- \frac{1}{2}\int_0^t \underbrace{\left[G(u,T)+G(u,t)\right]\left[G(u,T)-G(u,t)\right]}_{[G(u,T)-G(u,t)+2G(u,t)]G(t,T)G'(u,t)}\sigma^2(u)du.
\end{aligned}
$$

We use representation for forward rate $f(t,T)$ and get

$$
\begin{aligned}
I(t,T) &= \log\left(\frac{P(0,T)}{P(0,t)}\right) + G(t,T)f(0,t) \\
&\quad - \frac{1}{2}\int_0^t \left[G(u,T)-G(u,t)\right]G(t,T)G'(u,t)\sigma^2(u)du \\
&= \log\left(\frac{P(0,T)}{P(0,t)}\right) + G(t,T)f(0,t) - \frac{G(t,T)^2}{2}\int_0^t G'(u,t)^2\sigma^2(u)du.
\end{aligned}
$$

Now we can also derive future zero bond prices V

Finally, we get the result

$$
\begin{aligned}
P(t,T) &= e^{-G(t,T)r(t)+I(t,T)} \\
&= \frac{P(0,T)}{P(0,t)} e^{-G(t,T)[r(t)-f(0,t)]-\frac{G(t,T)^2}{2}\int_0^t \left[e^{-a(t-u)}\sigma(u)\right]^2 du}.
\end{aligned}
$$

- Future zero coupon bonds depend on:
  - today's yield curve $f(0,t)$,
  - mean reversion parameter a via $G(t,T)$, and
  - short rate volatility $\sigma(t)$.
- We see that drift $\theta(t)$ is not required for future zero coupon bonds.

Outline

Recall short rate dynamics in separable HJM model

We consider a one-factor model ($d = 1$)

$$
\begin{aligned}
r(t) &= f(0,t) + x(t) \\
dx(t) &= [y(t) - \chi(t) \cdot x(t)]\, dt + \sigma_r(t) \cdot dW(t) \\
dy(t) &= \left[\sigma_r(t)^2 - 2 \cdot \chi(t) \cdot y(t)\right] \cdot dt
\end{aligned}
$$

with

$$
H'(t) = -\chi(t) H(t),\; H(0) = 1 \text{ and } g(t) = H(t)^{-1} \sigma_r(t).
$$

- How does this relate to Hull-White model with

$$
dr(t) = [\theta(t) - a \cdot r(t)] \cdot dt + \sigma(t) \cdot dW(t)?
$$

Differentiate short rate in HJM model

$$
\begin{aligned}
dr(t) &= f'(0,t)dt + dx(t) \\
&= f'(0,t)dt + \left[y(t) - \chi(t)x(t)\right]dt + \sigma_r(t)dW(t) \\
&= \left[f'(0,t) + y(t) - \chi(t)\left(r(t) - f(0,t)\right)\right]dt + \sigma_r(t)dW(t) \\
&= \Bigg[\underbrace{f'(0,t) + \chi(t)f(0,t) + y(t)}_{\theta(t)} - \underbrace{\chi(t)}_{a} r(t)\Bigg] dt + \underbrace{\sigma_r(t)}_{\sigma(t)} dW(t)
\end{aligned}
$$

HJM volatility parameters become

$$
H'(t) = -aH(t), \quad H(0) = 1 \;\Rightarrow\; h(t) = H(t) = e^{-at},
$$

$$
g(t) = \sigma_r(t) \cdot H(t)^{-1} = \sigma(t)e^{at}.
$$

Deterministic volatility allows calculation of auxilliary variable $y(t)$

We have

$$y'(t) = \sigma(t)^2 - 2 \cdot a \cdot y(t), \quad y(0) = 0.$$

Solving initial value problem yields

$$y(t) = \int_0^t \sigma(u)^2 \cdot e^{-2a(t-u)} du.$$

Hull-White model in HJM notation

In the HJM framework the Hull-White model becomes

$$
\begin{aligned}
r(t) &= f(0,t) + x(t), \\
dx(t) &= \left[\int_0^t \sigma(u)^2 \cdot e^{-2a(t-u)} du - a \cdot x(t)\right] \cdot dt + \sigma(t) \cdot dW(t), \\
x(0) &= 0.
\end{aligned}
$$

We will use this representation of the Hull-White model for our implementations.

This also gives HJM representation of Hull-White model

Corollary (Forward rate dynamics in Hull-White model)

In a Hull-White model the dynamics of the forward rate $f(t, T)$ become

$$df(t,T) = \sigma(t)^2 e^{-a(T-t)} \frac{1 - e^{-a(T-t)}}{a} dt + \sigma(t) e^{-a(T-t)} dW(t).$$

Proof.

$$\begin{aligned} df(t,T) &= \sigma_f(t,T) \cdot \left[\int_t^T \sigma_f(t,u) du \right] \cdot dt + \sigma_f(t,T) \cdot dW(t) \\ &= g(t)h(T) \left[\int_t^T g(t)h(u) du \right] \cdot dt + g(t)h(T) \cdot dW(t) \\ &= \sigma(t)^2 e^{-a(T-t)} \underbrace{\left[\int_t^T e^{-a(u-t)} du \right]}_{\frac{1-e^{-a(T-t)}}{a}} \cdot dt + \sigma(t) e^{-a(T-t)} \cdot dW(t). \end{aligned}$$

Zero bond prices may also be computed in terms of $x(t)$

Corollary (Zero bonds in Hull-White model)

In the Hull-White model future zero coupon bonds are

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left\{ -G(t,T)x(t) - \frac{G(t,T)^2}{2} \int_0^t \left[e^{-a(t-u)} \sigma(u) \right]^2 du \right\}$$

with

$$G(t,T) = \int_t^T e^{-a(u-t)} du = \left[\frac{1 - e^{-a(T-t)}}{a} \right].$$

Proof.

Result follows either from Hull-White model zero bond formula with $x(t) = r(t) - f(0,T)$ or from zero bond formula for the separable HJM model with Hull-White results for $G(t,T)$ and $y(t)$. □

Outline

First we need the distribution of the state variable $x(t)$

We have
$$dx(t) = [y(t) - a \cdot x(t)] \cdot dt + \sigma(t) \cdot dW(t).$$
This yields for $t \geq s$
$$x(t) = e^{-a(t-s)} \left[x(s) + \int_s^t e^{a(u-s)} \left(y(u)du + \sigma(u)dW(u) \right) \right].$$

Lemma (State variable distribution)

In the HJM version of the Hull-White model we have that under the risk-neutral measure the state variable $x(t)$ is normally distributed with
$$\mathbb{E}^{\mathbb{Q}} [x(t) \,|\, \mathcal{F}_s] = e^{-a(t-s)} \left[x(s) + \int_s^t e^{a(u-s)} y(u) du \right] \text{ and}$$
$$Var[x(t) \,|\, \mathcal{F}_s] = \int_s^t \left[e^{-a(t-u)} \sigma(u) \right]^2 du.$$

Result follows directly from state variable representation for $x(t)$

Proof.

Result for $\mathbb{E}\left[x(t)\,|\,\mathcal{F}_s\right]$ follows from martingale property of Ito integral. Variance follows from Ito isometry

$$\begin{aligned}\mathrm{Var}\left[x(t)\,|\,\mathcal{F}_s\right] &= e^{-2a(t-s)}\int_s^t \left[e^{a(u-s)}\sigma(u)\right]^2 du \\ &= \int_s^t \left[e^{-a(t-u)}\sigma(u)\right]^2 du.\end{aligned}$$

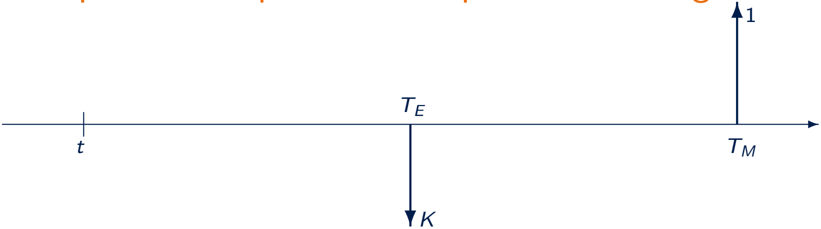
$\square$

- We will have a closer look at $\mathbb{E}^{\mathbb{Q}}\left[x(t)\,|\,\mathcal{F}_s\right] = e^{-a(t-s)}\left[x(s) + \int_s^t e^{a(u-s)}y(u)du\right]$ later on.
- Note, that we can also write

$$\mathrm{Var}\left[x(t)\,|\,\mathcal{F}_s\right] = y(t) - G'(s,t)^2 y(s).$$

Auxilliary variable $y(t)$ represents the (co-)variance process of $x(t)$.

Zero coupon bond options are important building blocks



Definition (Zero coupon bond (ZCB) option)

A zero coupon bond option is defined as an option with expiry time T_E, ZCB maturity time T_M with $T_M \geq T_E$, strike K , call/put flag $\phi \in \{1, -1\}$ and payoff

$$V^{\mathrm{ZBO}}(T_E) = \left[\phi\left(P(T_E, T_M) - K\right)\right]^+.$$

- We are interested in present value $V^{\mathrm{ZBO}}(t)$.
- We use T_E-forward measure for valuation

$$V^{\mathrm{ZBO}}(t) = P(t, T_E) \cdot \mathbb{E}^{T_E}\left[\left[\phi\left(P(T_E, T_M) - K\right)\right]^+ \mid \mathcal{F}_t\right].$$

$P(T_E, T_M)$ is log-normally distributed with known parameters

- We have for the forward bond price

$$\mathbb{E}^{T_E}\left[P(T_E, T_M)\,|\,\mathcal{F}_t\right] = P(t, T_M)/P(t, T_E).$$

- From

$$P(T_E, T_M) = \frac{P(t, T_M)}{P(t, T_E)} e^{-G(T_E, T_M)x(T_E) - \frac{G(T_E, T_M)^2}{2}\int_t^{T_E}\left[e^{-a(T_E-u)}\sigma(u)\right]^2 du}$$

  we get
  - $P(T_E, T_M)$ is log-normally distributed with log-normal variance

$$\nu^2 = \mathrm{Var}\left[G(T_E, T_M)x(T_E)\,|\,\mathcal{F}_t\right] = G(T_E, T_M)^2 \int_t^{T_E}\left[e^{-a(T_E-u)}\sigma(u)\right]^2 du,$$

  - we can apply Black's formula for option pricing.

ZCO prices are given by Black's formula

Theorem (ZCO pricing formula)

The time-t price of a zero coupon bond option with expiry time T_E, ZCB maturity time T_M with $T_M \geq T_E$, strike K , call/put flag $\phi \in \{1, -1\}$ and payoff

$$V^{ZBO}(T_E) = [\phi (P(T_E, T_M) - K)]^+$$

is given by

$$V^{ZBO}(t) = P(t, T_E) \cdot Black\left(P(t, T_M)/P(t, T_E), K, \nu, \phi\right)$$

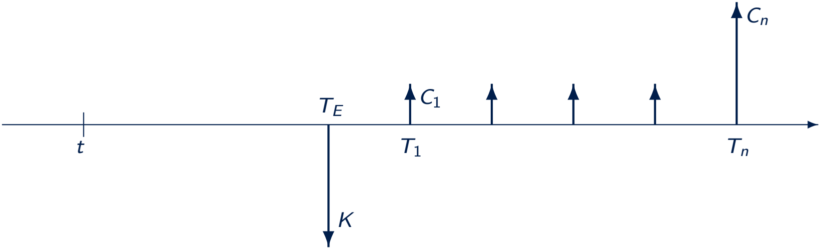
with log-normal bond price variance

$$\nu^2 = G(T_E, T_M)^2 \int_t^{T_E} \left[e^{-a(T_E - u)} \sigma(u)\right]^2 du.$$

Proof.

Result follows from log-normal distribution property. □

Coupon bond options are further building blocks



Payoff at option expiry T_E

$$V(T_E) = \left[\left(\sum_{i=1}^{n} C_i \cdot P(T_E, T_i) \right) - K \right]^+ .$$

Coupon bond options are options on a basket of future cash flows

Definition (Coupon bond option (CBO))

A coupon bond option is defined as an option with expiry time T_E, future cash flow payment times $T_1, \ldots, T_n$ (with $T_i > T_E$), corresponding cash flow values $C_1, \ldots, C_n$, a fixed strike price K, call/put flag $\phi \in \{1, -1\}$ and payoff

$$V^{\text{CBO}}(T_E) = \left[\left(\phi\left[\left(\sum_{i=1}^{n} C_i P(T_E, T_i)\right) - K\right]\right)^+\right].$$

- We cannot price CBO directly due to the basket structure.
- However, with some (not too strong) assumptions we can represent the 'option on a basket' as a 'basket of options'.
- We use monotonicity of bond prices (for $t < T$)

$$\frac{\partial}{\partial x} P(x(t); t, T) = -G(t, T) \cdot P(x(t); t, T) < 0.$$

CBO's are transformed via Jamshidian's trick I

W.l.o.g. set $\phi = 1$ (method works for $\phi = -1$ as well).
Assume underlying bond is monotone in state variable $x(T_E)$, i.e.

$$\begin{aligned}\frac{\partial}{\partial x}\sum_{i=1}^{n} C_i P(x(T_E); T_E, T_i) &= \sum_{i=1}^{n} C_i \frac{\partial}{\partial x} P(x(T_E); T_E, T_i) \\ &= -\sum_{i=1}^{n} C_i G(T_E, T_i) P(x(T_E); T_E, T_i) < 0.\end{aligned}$$

- Condition is satisfied, e.g. if $C_i \geq 0$.
- Small negative cash flows typically don't violate the assumption since last cash flow C_n is typically a large positive cash flow.

CBO's are transformed via Jamshidian's trick II

Then find $x^\star$ such that

$$\left(\sum_{i=1}^{n} C_i P(x^\star; T_E, T_i)\right) - K = 0$$

and set $K_i = P(x^\star; T_E, T_i)$.
We get (using monotonicity assumption)

$$\begin{aligned}
\left[\left(\sum_{i=1}^{n} C_i P(T_E, T_i)\right) - K\right]^+ &= \mathbb{1}_{\{x(T_E)\le x^\star\}}\left[\left(\sum_{i=1}^{n} C_i P(T_E, T_i)\right) - K\right] \\
&= \mathbb{1}_{\{x(T_E)\le x^\star\}}\left[\sum_{i=1}^{n} C_i P(T_E, T_i) - \sum_{i=1}^{n} C_i K_i\right] \\
&= \left[\sum_{i=1}^{n} C_i \left[P(T_E, T_i) - K_i\right] \mathbb{1}_{\{x(T_E)\le x^\star\}}\right] \\
&= \left[\sum_{i=1}^{n} C_i \left[P(T_E, T_i) - K_i\right]^+\right].
\end{aligned}$$

CBO's are transformed via Jamshidian's trick III

This gives

$$\mathbb{E}^{T_E}\left[\left[\left(\sum_{i=1}^{n} C_i P(T_E, T_i)\right) - K\right]^+\right] = \sum_{i=1}^{n} C_i \underbrace{\mathbb{E}^{T_E}\left[[P(T_E, T_i) - K_i]^+\right]}_{\text{Black's formula}}$$

or

$$\begin{aligned} V^{\text{CBO}}(t) &= \sum_{i=1}^{n} C_i \cdot V_i^{\text{ZBO}}(t) \\ &= \sum_{i=1}^{n} C_i \cdot P(t, T_E) \cdot \text{Black}\left(P(t, T_i)/P(t, T_E), K_i, \nu_i, \phi\right), \\ \nu_i^2 &= G(T_E, T_i)^2 \int_t^{T_E} \left[e^{-a(T_E - u)}\sigma(u)\right]^2 du. \end{aligned}$$

CBO's are prices as sum of ZBO's

Theorem (CBO pricing formula)

Consider a CBO with expiry time T_E, future cash flow payment times $T_1, \ldots, T_n$ (with $T_i > T_E$), corresponding cash flow values $C_1, \ldots, C_n$, fixed strike price K and call/put flag $\phi \in \{1, -1\}$. Assume that the underlying bond price $\sum_{i=1}^{n} C_i P(x(T_E); T_E, T_i)$ is monotonically decreasing in the state variable $x(T_E)$. Then the time-t price of the CBO is

$$V^{CBO}(t) = \sum_{i=1}^{n} C_i \cdot V_i^{ZBO}(t)$$

where $V_i^{ZBO}(t)$ is the time-t price of a corresponding ZBO with strike $K_i = P(x^\star; T_E, T_i)$ where the break-even state $x^\star$ is given by

$$\left(\sum_{i=1}^{n} C_i P(x^\star; T_E, T_i) \right) - K = 0.$$

Proof.

Follows from derivation above. □

Outline

We have another look at the expectation(s) of $x(t)$

- For general option pricing we also need expectation $\mathbb{E}^T\left[x(T) \,|\, \mathcal{F}_t\right]$.
- Then we can price

$$V(t) = P(t,T) \cdot \mathbb{E}^T\left[V(x(T);T) \,|\, \mathcal{F}_t\right] = P(t,T) \cdot \int_{-\infty}^{+\infty} V(x;T) \cdot p_{\mu,\sigma^2}(x) \cdot dx.$$

- Here $p_{\mu,\sigma^2}(x)$ is the density of a normal distribution $N\left(\mu, \sigma^2\right)$ with

$$\mu = \mathbb{E}^T\left[x(T) \,|\, \mathcal{F}_t\right] \text{ and } \sigma^2 = \mathrm{Var}\left[x(T) \,|\, \mathcal{F}_t\right].$$

- Integral $\int_{-\infty}^{+\infty} V(x;T) \cdot p_{\mu,\sigma^2}(x) \cdot dx$ is typically evaluated numerically (i.e. quadrature).
- We first calculate $\mathbb{E}^{\mathbb{Q}}\left[x(T) \,|\, \mathcal{F}_t\right]$ and then derive $\mathbb{E}^T\left[x(T) \,|\, \mathcal{F}_t\right]$.

We calculate expectation in risk-neutral measure I

Recall

$$dx(t) = [y(t) - a \cdot x(t)] \cdot dt + \sigma(t) \cdot dW(t).$$

This yields for $T \geq t$

$$x(T) = e^{-a(T-t)} \left[x(t) + \int_t^T e^{a(u-t)} \left(y(u)du + \sigma(u)dW(u) \right) \right]$$

and

$$\mathbb{E}^{\mathbb{Q}} \left[x(T) \,|\, \mathcal{F}_t \right] = e^{-a(T-t)} x(t) + \int_t^T e^{-a(T-u)} y(u) du.$$

We get

$$\begin{aligned}
\int_t^T e^{-a(T-u)} y(u) du &= \int_t^T e^{-a(T-u)} \left(\int_0^u \sigma(s)^2 e^{-2a(u-s)} ds \right) du \\
&= \int_t^T e^{-a(T-u)} \left(\int_0^t \sigma(s)^2 e^{-2a(u-s)} ds \right) du \\
&\quad + \int_t^T e^{-a(T-u)} \left(\int_t^u \sigma(s)^2 e^{-2a(u-s)} ds \right) du.
\end{aligned}$$

We calculate expectation in risk-neutral measure II

We analyse the integrals individually,

$$
\begin{aligned}
I_1(t,T) &= \int_t^T e^{-a(T-u)} \left(\int_0^t \sigma(s)^2 e^{-2a(u-s)} ds \right) du \\
&= \int_t^T \left(\int_0^t e^{-a(T-u)} \sigma(s)^2 e^{-2a(u-s)} ds \right) du \\
&= \int_0^t \left(\int_t^T e^{-a(T-u)} \sigma(s)^2 e^{-2a(u-s)} du \right) ds \\
&= \int_0^t \sigma(s)^2 \left(\int_t^T e^{-a(T-u)} e^{-2a(u-s)} du \right) ds \\
&= \int_0^t \sigma(s)^2 \left[\frac{e^{-a(T-u)} e^{-2a(u-s)}}{-a} \right]_{u=t}^T ds \\
&= \int_0^t \frac{\sigma(s)^2}{a} \left[e^{-a(T-t)} e^{-2a(t-s)} - e^{-a(T-T)} e^{-2a(T-s)} \right] ds.
\end{aligned}
$$

We calculate expectation in risk-neutral measure III

Exponential terms can be further simplified as

$$e^{-a(T-t)}e^{-2a(t-s)} - e^{-2a(T-s)} = e^{-a(T-t)}\left[1 - e^{-a(T-t)}\right]e^{-2a(t-s)}.$$

This gives

$$I_1(t,T) = e^{-a(T-t)}\frac{1 - e^{-a(T-t)}}{a}\int_0^t \sigma(s)^2 e^{-2a(t-s)}ds.$$

We calculate expectation in risk-neutral measure IV

For the second integral we get

$$
\begin{aligned}
I_2(t,T) &= \int_t^T e^{-a(T-u)} \left(\int_t^u \sigma(s)^2 e^{-2a(u-s)} ds \right) du \\
&= \int_t^T \left(\int_t^u e^{-a(T-u)} \sigma(s)^2 e^{-2a(u-s)} ds \right) du \\
&= \int_t^T \left(\int_s^T e^{-a(T-u)} \sigma(s)^2 e^{-2a(u-s)} du \right) ds \\
&= \int_t^T \sigma(s)^2 \left(\int_s^T e^{-a(T-u)} e^{-2a(u-s)} du \right) ds \\
&= \int_t^T \sigma(s)^2 \left[\frac{e^{-a(T-u)} e^{-2a(u-s)}}{-a} \right]_{u=s}^{T} ds \\
&= \int_t^T \frac{\sigma(s)^2}{a} \left[e^{-a(T-s)} e^{-2a(s-s)} - e^{-a(T-T)} e^{-2a(T-s)} \right] ds.
\end{aligned}
$$

We calculate expectation in risk-neutral measure V

Again we simplify exponential terms

$$e^{-a(T-s)} - e^{-2a(T-s)} = e^{-a(T-s)} \left[1 - e^{-a(T-s)}\right].$$

This gives

$$I_2(t,T) = \int_t^T \sigma(s)^2 e^{-a(T-s)} \frac{1 - e^{-a(T-s)}}{a} ds.$$

In summary, we get

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[x(T) \,|\, \mathcal{F}_t\right] &= e^{-a(T-t)} x(t) + I_1(t,T) + I_2(t,T) \\
&= e^{-a(T-t)} \left[x(t) + \frac{1 - e^{-a(T-t)}}{a} \int_0^t \sigma(s)^2 e^{-2a(t-s)} ds\right] \\
&\quad + \int_t^T \sigma(s)^2 e^{-a(T-s)} \frac{1 - e^{-a(T-s)}}{a} ds.
\end{aligned}$$

We calculate expectation in terminal measure I

Recall change of measure

$$dW^T(t) = dW(t) + \sigma_P(t,T)dt.$$

We have

$$\sigma_P(t,T) = \sigma(t)G(t,T) = \sigma(t) \cdot \frac{1 - e^{-a(T-t)}}{a}.$$

This gives

$$dx(t) = \left[y(t) - \sigma(t)^2 G(t,T) - a \cdot x(t)\right] \cdot dt + \sigma(t) \cdot dW^T(t)$$

and

$$x(T) = e^{-a(T-t)}.$$

$$\left[x(t) + \int_t^T e^{a(u-t)} \left(\left[y(u) - \sigma(u)^2 G(u,T)\right] du + \sigma(u) dW^T(u)\right)\right].$$

We calculate expectation in terminal measure II

We find that

$$\mathbb{E}^T\left[x(T)\,|\,\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[x(T)\,|\,\mathcal{F}_t\right] - \int_t^T \sigma(u)^2 e^{-a(T-u)} G(u,T)du.$$

It turns out that

$$\begin{aligned}\int_t^T \sigma(u)^2 e^{-a(T-u)} G(u,T)du &= \int_t^T \sigma(u)^2 e^{-a(T-u)} \frac{1-e^{-a(T-u)}}{a} du \\ &= I_2(t,T).\end{aligned}$$

As a result, we get

$$\mathbb{E}^T\left[x(T)\,|\,\mathcal{F}_t\right] = e^{-a(T-t)}\left[x(t) + \frac{1-e^{-a(T-t)}}{a}\int_0^t \sigma(s)^2 e^{-2a(t-s)} ds\right]$$

or more formally

$$\mathbb{E}^T\left[x(T)\,|\,\mathcal{F}_t\right] = G'(t,T)\left[x(t) + G(t,T)y(t)\right].$$

Outline

All the formulas serve the purpose of model calibration and derivative pricing

Model Calibration	Derivative Pricing
zero bond option (ZBO)	future zero bonds $P(x(t); t, T)$
coupon bond option (CBO)	expectation $\mathbb{E}^T[x(T) \mid \mathcal{F}_t]$ and variance $\mathrm{Var}[x(T) \mid \mathcal{F}_t]$
European swaption	payoff pricing $V(t) = P(t, T) \cdot \mathbb{E}^T[V(x(T); T) \mid \mathcal{F}_t]$

Bond option pricing is realised via ZBO's and CBO's

Zero Bond Option (ZBO)

Zero bond with expiry T_E, maturity T_M, strike K and call/put flag ϕ

$$
\begin{aligned}
V^{\text{ZBO}}(0) &= P(0, T_E) \cdot \text{Black}\left(P(0, T_M)/P(0, T_E), K, \nu, \phi\right), \\
\nu^2 &= G(T_E, T_M)^2 y(T_E).
\end{aligned}
$$

Coupon Bond Option (CBO)

Coupon bond option with strike K and underlying bond $\sum_{i=1}^{n} C_i \cdot P(T_E, T_i)$,

$$
V^{\text{CBO}}(t) = \sum_{i=1}^{n} C_i \cdot V_i^{\text{ZBO}}(t)
$$

where ZBO's $V_i^{\text{ZBO}}(t)$ with expiry T_E, maturity T_i, and strike $K_i = P(x^\star, T_E, T_i)$ and $x^\star$ s.t.

$$
\sum_{i=1}^{n} C_i \cdot P(x^\star; T_E, T_i) = K.
$$

General derivative pricing requires state variable expectation and variance

Zero Bonds (as building blocks for payoffs $V(x(T);T)$)

$$P(x(T);T,S) = \frac{P(0,S)}{P(0,T)} \exp\left\{-G(T,S)x(T) - \frac{G(T,S)^2}{2} y(T)\right\}.$$

General Derivative Pricing

$$V(t) = P(t,T) \cdot \mathbb{E}^T \left[V(x(T);T) \,|\, \mathcal{F}_t \right] = P(t,T) \cdot \int_{-\infty}^{+\infty} V(x;T) \cdot p_{\mu,\sigma^2}(x) \cdot dx$$

with $p_{\mu,\sigma^2}(x)$ density of a Normal distribution $N\left(\mu,\sigma^2\right)$ with

$$\mu = \mathbb{E}^T \left[x(T) \,|\, \mathcal{F}_t \right] = G'(t,T)\left[x(t) + G(t,T) y(t) \right]$$

and

$$\sigma^2 = \mathrm{Var}\left[x(T) \,|\, \mathcal{F}_t \right] = y(T) - G'(t,T)^2 y(t).$$

Fortunately, we only need a small set of model functions for implementation

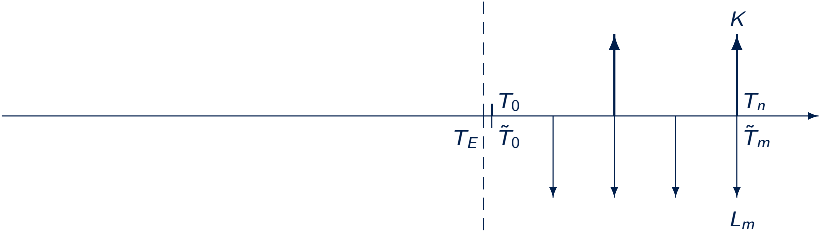
- Discount factors $P(0, T)$ from input yield curve.
- Function $G(t, T)$ with
$$G(t, T) = \frac{1 - e^{-a(T-t)}}{a}.$$
- Function $G'(t, T)$ with
$$G'(t, T) = e^{-a(T-t)}.$$
- Auxilliary variable $y(t)$ with
$$y(t) = \int_0^t \left[e^{-a(t-u)} \sigma(u) \right]^2 du = \sum_{j=1}^{k} \frac{e^{-2a(t-t_j)} - e^{-2a(t-t_{j-1})}}{2a} \sigma_j^2$$
where we assume $\sigma(t)$ piece-wise constant on a grid $0 = t_0, t_1, \ldots, t_k = t$.

Outline

It remains to show how Hull-Wite model is applied to European swaptions

Model Calibration	Derivative Pricing
zero bond option (ZBO)	future zero bonds $P(x(t); t, T)$
coupon bond option (CBO)	expectation $\mathbb{E}^T[x(T) \mid \mathcal{F}_t]$ and variance $\text{Var}[x(T) \mid \mathcal{F}_t]$
European swaption	payoff pricing $V(t) = P(t, T) \cdot \mathbb{E}^T[V(x(T); T) \mid \mathcal{F}_t]$

Recall that Swaption is option to enter into a swap at a future time



- At option exercise time T_E present value of swap is

$$V^{\text{Swap}}(T_E) = \underbrace{K \sum_{i=1}^{n} \tau_i P(T_E, T_i)}_{\text{future fixed leg}} - \underbrace{\sum_{j=1}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)\tilde{\tau}_j P(T_E, \tilde{T}_j)}_{\text{future float leg}}.$$

- Option to enter represents the right but not the obligation to enter swap.
- Rational market participant will exercise if swap present value is positive, i.e.

$$V^{\text{Swpt}}(T_E) = \max\left\{V^{\text{Swap}}(T_E), 0\right\}.$$

How do we get the swaption payoff compatible to our Hull-White model formulas?

$$V^{\text{Swap}}(T_E) = \underbrace{K \sum_{i=1}^{n} \tau_i P(T_E, T_i)}_{\text{future fixed Leg}} - \underbrace{\sum_{j=1}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)\tilde{\tau}_j P(T_E, \tilde{T}_j)}_{\text{future float leg}}$$

- Fixed leg can be expressed in terms of future state variable $x(T_E)$ via $P(x(T_E); T_E, T_i)$
- Float leg contains future forward Libor rates $L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)$ from (future) projection curve
- However, Hull-White model only provides representation of discount factors, i.e. $P(T_E, \tilde{T}_j)$

We need to model the relation between future Libor rates $L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)$ and discount factors $P(T_E, \tilde{T}_j)$.

We do have all ingredients from our deterministic multi-curve model

Recall the definition of (future) forward Libor rate

$$L^\delta(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta) = \mathbb{E}^{\tilde{T}_{j-1}+\delta}\left[L^\delta(\tilde{T}_{j-1}, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta) \mid \mathcal{F}_{T_E}\right]$$
$$= \left[\frac{P(T_E, \tilde{T}_{j-1})}{P(T_E, \tilde{T}_{j-1}+\delta)} \cdot D(\tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta) - 1\right] \frac{1}{\tau_{j-1}}$$

$(\tau_{j-1} = \tau(\tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta))$ with tenor basis factor

$$D(\tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta) = \frac{Q(T_E, \tilde{T}_{j-1})}{Q(T_E, \tilde{T}_{j-1}+\delta)}$$

and discount factors $Q(T_E, T)$ arising from credit (or funding) risk embedded in Libor rates $L^\delta(\cdot)$.

- Key assumption is that $D(\tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)$ is deterministic or independent of T_E.
- Then

$$D(\tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta) = \frac{Q(0, \tilde{T}_{j-1})}{Q(0, \tilde{T}_{j-1}+\delta)} = \frac{P^\delta(0, \tilde{T}_{j-1})}{P^\delta(0, \tilde{T}_{j-1}+\delta)} \cdot \frac{P(0, \tilde{T}_{j-1}+\delta)}{P(0, \tilde{T}_{j-1})}.$$

We use basis spread model to simplify Libor coupons

- Tenor basis factor

$$D_{j-1} = D(\tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta) = \frac{P^\delta(0, \tilde{T}_{j-1})}{P^\delta(0, \tilde{T}_{j-1} + \delta)} \cdot \frac{P(0, \tilde{T}_{j-1} + \delta)}{P(0, \tilde{T}_{j-1})}$$

  is calculated from today's projection curve $P^\delta(0, T)$ and discount curve $P(0, T)$.
- Further assume *natural* Libor payment dates and consistent year fractions

$$\tilde{T}_j = \tilde{T}_{j-1} + \delta, \quad \tau(\tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta) = \tilde{\tau}_j.$$

- Libor coupon becomes

$$\begin{aligned} L^\delta(T_E, \tilde{T}_{j-1}, \tilde{T}_j)\tilde{\tau}_j P(T_E, \tilde{T}_j) &= \left[\frac{P(T_E, \tilde{T}_{j-1})}{P(T_E, \tilde{T}_j)} D_{j-1} - 1\right] \frac{1}{\tilde{\tau}_j} \tilde{\tau}_j P(T_E, \tilde{T}_j) \\ &= P(T_E, \tilde{T}_{j-1}) D_{j-1} - P(T_E, \tilde{T}_j). \end{aligned}$$

We can write the float leg (1/2)

$$
\begin{aligned}
V^{\text{Swap}}(T_E) &= \underbrace{K\sum_{i=1}^{n}\tau_i P(T_E, T_i)}_{\text{future fixed leg}} - \underbrace{\sum_{j=1}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(T_E, \tilde{T}_j)}_{\text{future float leg}} \\
&= K\sum_{i=l}^{n}\tau_i P(T_E, T_i) - \sum_{j=1}^{m} P(T_E, \tilde{T}_{j-1})D_{j-1} - P(T_E, \tilde{T}_j) \\
&= K\sum_{i=1}^{n}\tau_i P(T_E, T_i) \\
&\quad - \left[P(T_E, \tilde{T}_0)D_0 - P(T_E, \tilde{T}_m) + \sum_{j=2}^{m} P(T_E, \tilde{T}_{j-1})\left[D_{j-1} - 1\right]\right] \\
&= K\sum_{i=1}^{n}\tau_i P(T_E, T_i) \\
&\quad - \left[P(T_E, \tilde{T}_0) - P(T_E, \tilde{T}_m) + \sum_{j=1}^{m} P(T_E, \tilde{T}_{j-1})\left[D_{j-1} - 1\right]\right].
\end{aligned}
$$

We can re-write the float leg (2/2)

Reordering terms yields

$$V^{\text{Swap}}(T_E) = -\underbrace{P(T_E, \tilde{T}_0)}_{\text{strike paid at } T_0} + \underbrace{\sum_{i=1}^{n} K \cdot \tau_i \cdot P(T_E, T_i)}_{\text{fixed rate coupons}}$$

$$\underbrace{-\sum_{j=1}^{m} P(T_E, \tilde{T}_{j-1}) \cdot [D_{j-1} - 1]}_{\text{negative spread coupons}} + \underbrace{P(T_E, \tilde{T}_m)}_{\text{notional payment}}$$

$$= \sum_{k=0}^{n+m+1} C_k \cdot P(T_E, \bar{T}_k)$$

with

$$C_0 = -1,\ C_i = K \cdot \tau_i \,(i = 1, \ldots, n),\ C_{n+j} = -[D_{j-1} - 1],\ (j = 1, \ldots, m),$$

$$\text{and } C_{n+m+1} = 1,$$

and corresponding payment times $\bar{T}_k$.

Swaptions are equivalent to coupon bond options

Corollary (Equivalence between Swaption and bond option)

Consider a European Swaption with receiver/payer flag $\phi \in \{1, -1\}$ payoff

$$V^{Swpt}(T_E) = \left[\phi\left\{K\sum_{i=1}^{n}\tau_i P(T_E, T_i) - \sum_{j=1}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(T_E, \tilde{T}_j)\right\}\right.$$

Under our deterministic basis spread assumption the swaption payoff is equal to a call/put bond option payoff

$$V^{CBO}(T_E) = \left[\phi\left\{\sum_{k=0}^{n+m+1} C_k \cdot P(T_E, \bar{T}_k)\right\}\right]^+$$

with zero strike and cash flows C_k and times $\bar{T}_k$ as elaborated above. Moreover, if the underlying bond payoff is monotonic then

$$V^{Swpt}(t) = V^{CBO}(t) = \sum_{k=0}^{n+m+1} C_k \cdot V_k^{ZBO}(t)$$

We give some comments regarding the CBO mapping

- Note that $C_0 = -1$ is a *large* negative cash flow.
- However, $\frac{\partial}{\partial x}\left[-P(T_E, \tilde{T}_0)\right] \approx -G(T_E, T_0)$ is small because $T_E - T_0$ is small.
- If $T_E = \tilde{T}_0$, i.e. no spot offset between option expiry and swap start time, then
  - set CBO strike $K = D(\tilde{T}_0, \tilde{T}_1)$,
  - remove first negative spread coupon C_{n+1} from cash flow list.
- In practice monotonicity assumption

$$\frac{\partial}{\partial x}\left[\sum_{k=0}^{n+m+1} C_k \cdot P(T_E, \bar{T}_k)\right] < 0$$

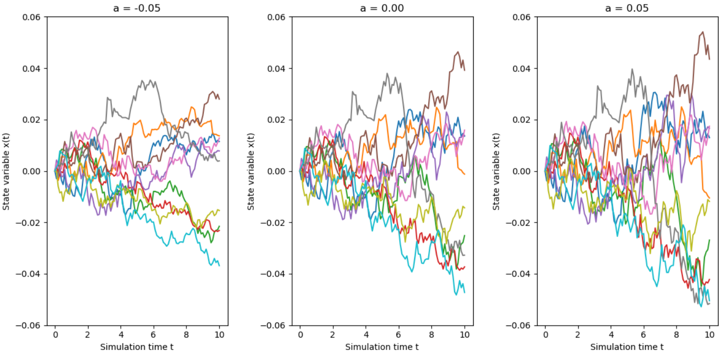
is typically no issue.

In Hull-White model calibration we will use CBO formula to match Hull-White model prices versus Vanilla model swaption prices.

Outline

How do the simulated paths *look like*?

- Model short rate volatility σ calibrated to $100bp$ flat volatility at $5y$ and $10y$, mean reversion $a \in \{-5\%, 0\%, 5\%\}$ [6]



- Higher mean reversion yields more *forward volatility.*

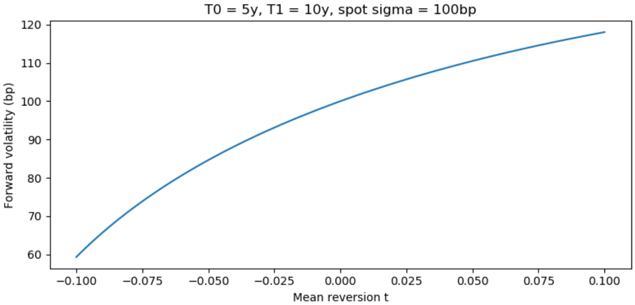
[6] Zero mean reversion is effectively approximated via $a = 1bp$. This does not change the overall behavior and avoids special treatment in formulas.

Forward volatility dependence on mean reversion can also be derived analytically

Denote forward volatility as

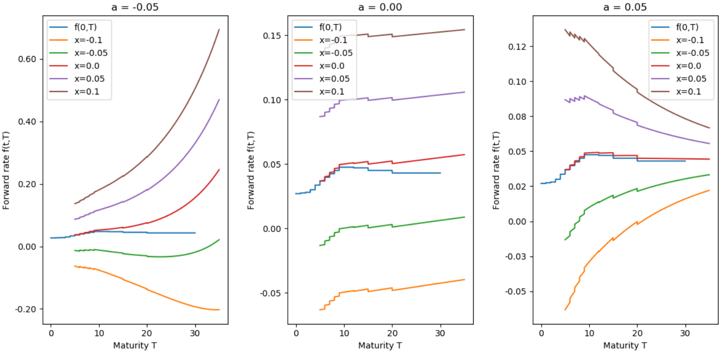
$$\sigma_{\text{Fwd}}(T_0, T_1) = \sqrt{\frac{\text{Var}\left[x(T_1) \mid \mathcal{F}_{T_0}\right]}{T_1 - T_0}} = \sqrt{\frac{y(T_1) - G'(T_0, T_1)^2 y(T_0)}{T_1 - T_0}}$$

- Suppose spot volatilities $\sigma_{\text{Fwd}}(0, T_1)$ and $\sigma_{\text{Fwd}}(0, T_0)$ (and thus $y(T_0)$ and $y(T_1)$ are fixed)
- If mean reversion a increases then $G'(T_0, T_1) = e^{-a(T_1 - T_0)}$ decreases
- Thus forward volatility $\sigma_{\text{Fwd}}(T_0, T_1)$ increases



Which kind of curves can we simulate with Hull-White model?

- Models use flat short rate volatility $\sigma = 100bp$ and mean reversion $a \in \{-5\%, 0\%, 5\%\}$ [7]



- Model *works* with negative mean reversion - however, yield curves are exploding

---

[7] Zero mean reversion is effectively approximated via $a = 1bp$. This does not change the overall behavior and avoids special treatment in formulas.

What are relevant properties of a model for option pricing?

- Vanilla models require input (ATM volatility) parameters for expiry-tenor-pairs.
  - Which shape of ATM volatilities for expiry-tenor-pairs are predicted by Hull-White model?
- SABR model allows modelling of volatility smile.
  - Which shapes of volatility smile can be modelled with Hull-White model?
  - How does the smile change if we change the model parameters?
- We aim at applying the Hull-White model to price Bermudan swaptions.
  - How do the model parameters impact prices of exotic derivatives?

For now we focus on model-implied volatilities (ATM and smile). The impact of model parameters on Bermudans is analysed later.

Model properties for option pricing are assessed by analysing model-implied volatilities

Model-implied normal volatility

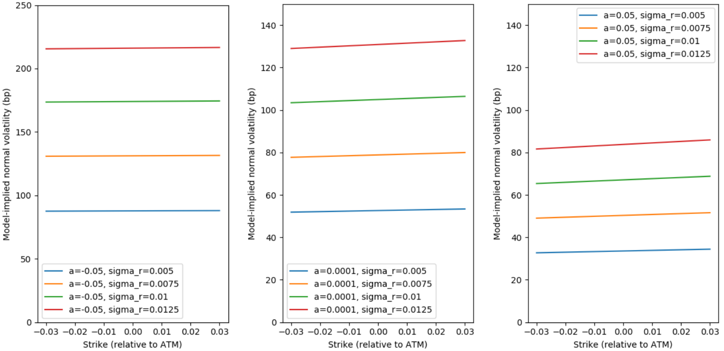
Consider a swaption with expiry/start/end-dates $T_E/T_0/T_n$ and strike rate K. For a given Hull-White model the model-implied normal volatility is calculated as

$$\sigma(T_0, T_n, K) = \text{Bachelier}^{-1}\left(S(t), K, V^{\text{CBO}}(t)/An(t), \phi\right)/\sqrt{T_E - t}.$$

Here, $S(t)$ and $An(t)$ are the forward swap rate and annuity of the underlying swap with start/end-date T_0/T_n. $V^{\text{CBO}}(t)$ is the Hull-White model price of a coupon bond option equivalent to the input swaption.

Which shapes of volatility smile can be modelled and how does the smile change if we change the model parameters?

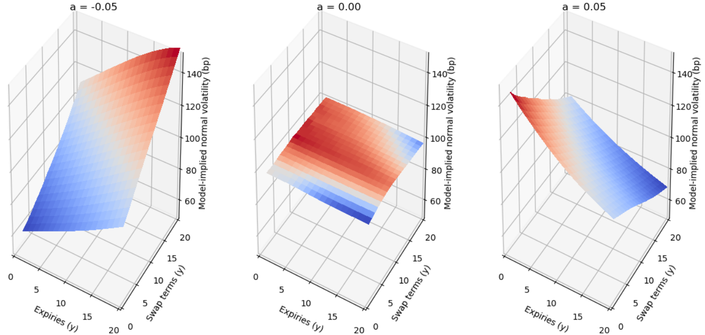
- Models use flat short rate volatility $\sigma \in \{50bp, 75bp, 100bp, 125bp\}$ and mean reversion $a \in \{-5\%, 0\%, 5\%\}$:



- We can only model flat smile - this is a major model limitation!
- Model-implied volatility decreases if mean reversion increases.

Which shape of ATM volatilities for expiry-tenor-pairs are predicted by Hull-White model?

- Models use flat short rate volatility σ - calibrated to 10y-10y swaption with $100bp$ volatility
- Mean reversion $a \in \{-5\%, 0\%, 5\%\}$:



- Mean reversion impacts slope of ATM volatilities in expiry and swap term dimension.

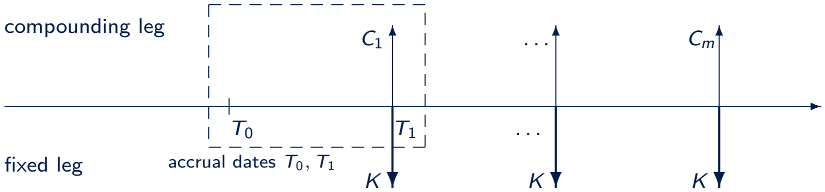
Outline

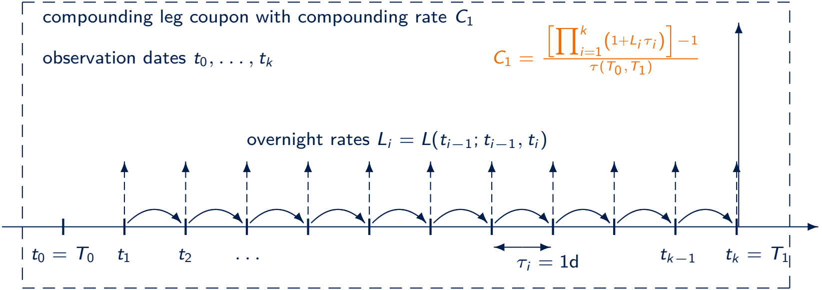
HJM Modelling Framework

Hull-White Model

Special Topic: Options on Overnight Rates

Recall overnight index swap (OIS) coupon rate calculation





The backward-looking compounded rate is composed of individual overnight rates

- Assume overnight index rate $L_i = L(t_{i-1}; t_{i-1}, t_i)$ is a credit-risk free simple compounded rate.
- Compounded rate C_1 (for a period $[T_0, T_1]$) is payed at T_1 and specified as

$$C_1 = \left\{ \left[\prod_{i=1}^{k} (1 + L_i \tau_i) \right] - 1 \right\} \frac{1}{\tau(T_0, T_1)}.$$

- Crucial part from modeling perspective is compounding factor

$$\prod_{i=1}^{k} (1 + L_i \tau_i) = \prod_{i=1}^{k} \frac{1}{P(t_{i-1}, t_i)}.$$

- Tower-law yields

$$\mathbb{E}^{T_1} \left[\prod_{i=1}^{k} \frac{1}{P(t_{i-1}, t_i)} \,|\, \mathcal{F}_{T_0} \right] = \frac{1}{P(T_0, T_1)}.$$

Outline

For pricing options on compounded rates we need the terminal distribution of the compounding factor

Use Hull-White model representation of zero bonds

$$P(t_{i-1}, t_i) = \frac{P(t, t_i)}{P(t, t_{i-1})} \exp\left\{ -G(t_{i-1}, t_i) x(t_{i-1}) - \frac{1}{2} G(t_{i-1}, t_i)^2 y(t_{i-1}) \right\},$$

$$G(t_{i-1}, t_i) = \frac{1 - \exp\{-a(t_i - t_{i-1})\}}{a},$$

$$y(t_{i-1}) = \int_t^{t_{i-1}} \sigma(u)^2 \cdot e^{-2a(t_{i-1}-u)} du.$$

Compounding factor becomes

$$\prod_{i=1}^{k} \frac{1}{P(t_{i-1}, t_i)} = \frac{P(t, T_0)}{P(t, T_1)} \exp\left\{ \sum_{i=1}^{k} G(t_{i-1}, t_i) x(t_{i-1}) + \frac{1}{2} G(t_{i-1}, t_i)^2 y(t_{i-1}) \right\}.$$

Variance of compounding factor is driven by stochastic term $\sum_{i=1}^{k} G(t_{i-1}, t_i) x(t_{i-1})$.

We write all $x(t_{i-1})$ in terms of $x(T_0)$ plus individual Ito integrals

We have in Hull-White model and risk-neutral measure

$$x(t_{i-1}) = e^{-a(t_{i-1}-T_0)} \left[x(T_0) + \int_{T_0}^{t_{i-1}} e^{a(u-T_0)} \left[y(u)du + \sigma(u)dW(u) \right] \right].$$

Abbreviate $dp(u) = y(u)du + \sigma(u)dW(u)$ (to simplify notation). Then

$$\begin{aligned}
&\sum_{i=1}^{k} G(t_{i-1}, t_i) x(t_{i-1}) \\
&= \sum_{i=1}^{k} G(t_{i-1}, t_i) \left\{ e^{-a(t_{i-1}-T_0)} \left[x(T_0) + \int_{T_0}^{t_{i-1}} e^{a(u-T_0)} dp(u) \right] \right\} \\
&= x(T_0) \sum_{i=1}^{k} G(t_{i-1}, t_i) e^{-a(t_{i-1}-T_0)} \\
&\qquad + \sum_{i=1}^{k} G(t_{i-1}, t_i) \int_{T_0}^{t_{i-1}} e^{-a(t_{i-1}-u)} dp(u).
\end{aligned}$$

We analyse above two parts individually.

First we calculate the scaling factor for $x(T_0)$

We have

$$G(t_{i-1}, t_i)e^{-a(t_{i-1}-T_0)} = \frac{1 - e^{-a(t_i - t_{i-1})}}{a} e^{-a(t_{i-1}-T_0)} = G(T_0, t_i) - G(T_0, t_{i-1}).$$

This yields the telescopic sum

$$\sum_{i=1}^{k} G(t_{i-1}, t_i)e^{-a(t_{i-1}-T_0)} = \sum_{i=1}^{k} G(T_0, t_i) - G(T_0, t_{i-1}) = G(T_0, T_1).$$

And we have

$$x(T_0) \sum_{i=1}^{k} G(t_{i-1}, t_i)e^{-a(t_{i-1}-T_0)} = G(T_0, T_1)x(T_0).$$

Second we calculate the sum of Ito integrals (1/2)

We split integration and re-order sums

$$\begin{aligned}
&\sum_{i=1}^{k} G(t_{i-1}, t_i) \int_{T_0}^{t_{i-1}} e^{-a(t_{i-1}-u)} dp(u) \\
&= \sum_{i=1}^{k} G(t_{i-1}, t_i) \sum_{j=1}^{i-1} \int_{t_{j-1}}^{t_j} e^{-a(t_{i-1}-u)} dp(u) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{i-1} \int_{t_{j-1}}^{t_j} G(t_{i-1}, t_i) e^{-a(t_{i-1}-u)} dp(u) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{i-1} \int_{t_{j-1}}^{t_j} \left[G(u, t_i) - G(u, t_{i-1}) \right] dp(u) \\
&= \sum_{j=1}^{k-1} \sum_{i=j+1}^{n} \int_{t_{j-1}}^{t_j} \left[G(u, t_i) - G(u, t_{i-1}) \right] dp(u) \\
&= \sum_{j=1}^{k-1} \int_{t_{j-1}}^{t_j} \sum_{i=j+1}^{n} \left[G(u, t_i) - G(u, t_{i-1}) \right] dp(u).
\end{aligned}$$

Second we calculate the sum of Ito integrals (2/2)

Now we can use telescopic sum property again and simplify

$$
\begin{aligned}
&\sum_{i=1}^{k} G(t_{i-1}, t_i) \int_{T_0}^{t_{i-1}} e^{-a(t_{i-1}-u)} dp(u) \\
&= \sum_{j=1}^{k-1} \int_{t_{j-1}}^{t_j} \sum_{i=j+1}^{n} \left[G(u, t_i) - G(u, t_{i-1})\right] dp(u) \\
&= \sum_{j=1}^{k-1} \int_{t_{j-1}}^{t_j} \left[G(u, t_n) - G(u, t_j)\right] dp(u) \\
&= \sum_{j=1}^{k-1} G(t_j, t_n) \int_{t_{j-1}}^{t_j} e^{-a(t_j-u)} dp(u).
\end{aligned}
$$

Putting things together yields the desired representation of the compounding factor (1/3)

$$\prod_{i=1}^{k} \frac{1}{P(t_{i-1}, t_i)} = \frac{P(t, T_0)}{P(t, T_1)} \exp\left\{ \sum_{i=1}^{k} G(t_{i-1}, t_i) x(t_{i-1}) + \frac{1}{2} G(t_{i-1}, t_i)^2 y(t_{i-1}) \right\}$$

with

$$\sum_{i=1}^{k} G(t_{i-1}, t_i) x(t_{i-1}) = G(T_0, T_1) x(T_0) + \sum_{j=1}^{k-1} G(t_j, t_n) \int_{t_{j-1}}^{t_j} e^{-a(t_j - u)} dp(u).$$

Putting things together yields the desired representation of the compounding factor (2/3)

Substituting back $dp(u) = y(u)du + \sigma(u)dW(u)$ gives

$$\sum_{i=1}^{k} G(t_{i-1}, t_i)x(t_{i-1}) = \underbrace{G(T_0, T_1)x(T_0)}_{I_0}$$
$$+ \sum_{j=1}^{k-1} \underbrace{G(t_j, t_n) \int_{t_{j-1}}^{t_j} e^{-a(t_j-u)}\sigma(u)dW(u)}_{I_j}$$
$$+ \sum_{j=1}^{k-1} G(t_j, t_n) \int_{t_{j-1}}^{t_j} e^{-a(t_j-u)}y(u)du.$$

Putting things together yields the desired representation of the compounding factor (3/3)

$$\prod_{i=1}^{k} \frac{1}{P(t_{i-1}, t_i)} = \frac{P(t, T_0)}{P(t, T_1)} \exp\left\{ \sum_{i=1}^{k} G(t_{i-1}, t_i)x(t_{i-1}) + \frac{1}{2} G(t_{i-1}, t_i)^2 y(t_{i-1}) \right\}$$

with

$$\begin{aligned}\sum_{i=1}^{k} G(t_{i-1}, t_i)x(t_{i-1}) = & \underbrace{G(T_0, T_1)x(T_0)}_{I_0} \\ & + \sum_{j=1}^{k-1} \underbrace{G(t_j, t_n) \int_{t_{j-1}}^{t_j} e^{-a(t_j-u)}\sigma(u)dW(u)}_{I_j} \\ & + \sum_{j=1}^{k-1} G(t_j, t_n) \int_{t_{j-1}}^{t_j} e^{-a(t_j-u)}y(u)du.\end{aligned}$$

Stochastic Terms I_0 and I_j are independent Ito integrals. Thus $\prod_{i=1}^{k} \frac{1}{P(t_{i-1}, t_i)}$ is log-normal with known variance.

Log-normal variance is given by sum of variances for Ito integrals I_0 and I_j

We first calculate the variance

$$\begin{aligned}\nu^2 &= \mathrm{Var}\left[\log\left(\prod_{i=1}^{k}\frac{1}{P(t_{i-1},t_i)}\right)\,|\,\mathcal{F}_t\right] = \mathrm{Var}\left[I_0+\sum_{j=1}^{k-1}I_j\,|\,\mathcal{F}_t\right]\\ &= G(T_0,T_1)^2\mathrm{Var}\left[x(T_0)\,|\,\mathcal{F}_t\right]\\ &\quad+\sum_{j=1}^{k-1}\mathbb{1}_{\{t\le t_{j-1}\}}G(t_j,t_n)^2\int_{t_{j-1}}^{t_j}\left[e^{-a(t_j-u)}\sigma(u)\right]^2du.\end{aligned}$$

Expectation is given from martingale property

Recall that expectation is also known already as

$$\begin{aligned}
\mu &= \mathbb{E}^{T_1}\left[\prod_{i=1}^{k}\frac{1}{P(t_{i-1},t_i)}\,|\,\mathcal{F}_t\right] \\
&= \frac{P(t,T_0)}{P(t,T_1)} \\
&= \prod_{i=1}^{k}\frac{P(t,t_{i-1})}{P(t,t_i)} \\
&= \prod_{i=1}^{k}\left(1+\mathbb{E}^{t_i}\left[L_i\,|\,\mathcal{F}_t\right]\tau_i\right)
\end{aligned}$$

for $t \le T_0$.

- Derivation can also be applied for partly fixed compounding periods with $T_0 < t \le T_1$.

We summarise results for compounding factor terminal distribution

Lemma (OIS compounding factor distribution)

The compounding factor $\prod_{i=1}^{k}(1+L_i\tau_i)=\prod_{i=1}^{k}\frac{1}{P(t_{i-1},t_i)}$ *of an OIS coupon in Hull-White model is log-normally distributed with expectation (in* T_1*-forward measure)*

$$\mu=\mathbb{E}^{T_1}\left[\prod_{i=1}^{k}(1+L_i\tau_i)\,|\,\mathcal{F}_t\right]=\prod_{i=1}^{k}\left(1+\mathbb{E}^{t_i}\left[L_i\,|\,\mathcal{F}_t\right]\tau_i\right)$$

and log-normal variance

$$\nu^2=G(T_0,T_1)^2 Var\left[x(T_0)\,|\,\mathcal{F}_t\right] + \sum_{j=1}^{k-1}\mathbb{1}_{\{t\le t_{j-1}\}}G(t_j,t_n)^2\int_{t_{j-1}}^{t_j}\left[e^{-a(t_j-u)}\sigma(u)\right]^2 du.$$

Note:

- If $t\ge T_0$ then $\mathrm{Var}\left[x(T_0)\,|\,\mathcal{F}_t\right]=0$.
- if $t<T_0$ then $\mathrm{Var}\left[x(T_0)\,|\,\mathcal{F}_t\right]=\int_t^{T_0}\left[e^{-a(T_0-u)}\sigma(u)\right]^2 du$.

Caplets and floorlets on OIS coupons can be calculated via Black formula

Theorem (OIS caplet and floorlet pricing)

A caplet or floorlet written on a compounded coupon rate $C_1 = \left\{ \left[\prod_{i=1}^{k} (1 + L_i \tau_i) \right] - 1 \right\} \frac{1}{\tau(T_0, T_1)}$ *with coupon period* $[T_0, T_1]$*, observation times* $T_0 = t_0, \ldots, t_k = T_1$ *and strike rate* K *pays at* T_1 *the payoff*

$$V(T_1) = \tau(T_0, T_1) \left[\phi \left(C_1 - K \right) \right]^+ .$$

In a Hull White model the option price at $t < T_1$ *is*

$$V(t) = P(t, T_1) \cdot Black\left(\mu, 1 + \tau(T_0, T_1) K, \nu, \phi\right)$$

with $\mu = \prod_{i=1}^{k} \left(1 + \mathbb{E}^{t_i} \left[L_i \,|\, \mathcal{F}_t \right] \tau_i \right)$ *and*

$$\begin{aligned} \nu^2 = {} & G(T_0, T_1)^2 Var\left[x(T_0) \,|\, \mathcal{F}_t \right] \\ & + \sum_{j=1}^{k-1} \mathbb{1}_{\{t \le t_{j-1}\}} G(t_j, t_n)^2 \int_{t_{j-1}}^{t_j} \left[e^{-a(t_j - u)} \sigma(u) \right]^2 du. \end{aligned}$$

Caplet and floorlet pricing formula follows directly from earlier derivations

Proof.

We abbreviate $\tau = \tau(T_0, T_1)$ and re-write the payoff as

$$V(T_1) = [\phi(\tau C_1 - \tau K)]^+ = \left[\phi\left(\left[\prod_{i=1}^{k}(1 + L_i\tau_i)\right] - (1 + \tau K)\right)\right]^+.$$

Consequently, we can view it as an option on the compounding factor $\prod_{i=1}^{k}(1 + L_i\tau_i)$ with strike $1 + \tau(T_0, T_1)K$. Using T_1-forward measure yields the present value

$$V(t) = P(t, T_1) \cdot \mathbb{E}^{T_1}\left\{\left[\phi\left(\left[\prod_{i=1}^{k}(1 + L_i\tau_i)\right] - (1 + \tau K)\right)\right]^+ \mid \mathcal{F}_t\right\}.$$

We established earlier that the compounding factor $\prod_{i=1}^{k}(1 + L_i\tau_i)$ is log-normally distributed with expectation μ and log-normal variance ν^2 as stated in the theorem. Thus we can apply Black's formula for call and put option pricing. $\square$

Outline

In practice, the discrete compounding factor $\prod_{i=1}^{k}(1 + L_i\tau_i)$ may be approximated to simplify valuation formulas

Typically, the compounding period t_{i-1} to t_i for an overnight rate L_i is small: one day (or two/three days for holidays/weekends).
We use the short rate $r(t)$, martingale property of bank account in t_i-forward measure and approximate

$$
\begin{aligned}
1 + L_i\tau_i = \frac{1}{P(t_{i-1}, t_i)} &= \mathbb{E}^{t_i}\left[\exp\left\{\int_{t_{i-1}}^{t_i} r(u)du\right\} \mid \mathcal{F}_{t_{i-1}}\right] \\
&\approx \exp\left\{\int_{t_{i-1}}^{t_i} r(u)du\right\}.
\end{aligned}
$$

This yields continuous compounding factor approximation

$$
\prod_{i=1}^{k}(1 + L_i\tau_i) \approx \prod_{i=1}^{k} e^{\int_{t_{i-1}}^{t_i} r(u)du} = e^{\sum_{i=1}^{k}\int_{t_{i-1}}^{t_i} r(u)du} = \exp\left\{\int_{T_0}^{T_1} r(u)du\right\}.
$$

Approximate option payoff is formulated using continuous compounding factor

(Approximate) OIS caplet payoff is

$$\left[\exp\left\{\int_{T_0}^{T_1} r(u)du\right\} - [1+\tau(T_0,T_1)K]\right]^+.$$

As before we have for $t \leq T_0$

$$\begin{aligned}
\mu &= \mathbb{E}^{T_1}\left[\exp\left\{\int_{T_0}^{T_1} r(u)du\right\} \mid \mathcal{F}_t\right] \\
&= \mathbb{E}^{T_1}\left[\mathbb{E}^{T_1}\left[\exp\left\{\int_{T_0}^{T_1} r(u)du\right\} \mid \mathcal{F}_{T_0}\right] \mid \mathcal{F}_t\right] \\
&= \mathbb{E}^{T_1}\left[\frac{1}{P(T_0,T_1)} \mid \mathcal{F}_t\right] = \frac{P(t,T_0)}{P(t,T_1)}.
\end{aligned}$$

What is the distribution of continuous compounding factor $\exp\left\{\int_{T_0}^{T_1} r(u)du\right\}$?

We already know $I(T_0, T_1) = \int_{T_0}^{T_1} r(u)du$ from drift calculation for classical Hull White model

From the proof of Lemma lem:HW-Drift-Calibration(p. 268) we have

$$\begin{aligned}
I(T_0, T_1) &= \int_{T_0}^{T_1} r(u)du \\
&= G(T_0, T_1)r(T_0) + \int_{T_0}^{T_1} G(u, T_1)\left[\theta(u) + \sigma(u)dW(u)\right]. \\
&= G(T_0, T_1)\left[f(0, T_0) + x(T_0)\right] + \int_{T_0}^{T_1} G(u, T_1)\left[\theta(u) + \sigma(u)dW(u)\right].
\end{aligned}$$

This yields

- Integrated short rate $I(T_0, T_1)$ is normally distributed, thus $\exp\left\{I(T_0, T_1)\right\}$ is log-normal.
- Variance of $I(T_0, T_1)$ can be calculated via Ito isometry

$$\bar{\nu}^2 = \mathrm{Var}\left[I(T_0, T_1) \,|\, \mathcal{F}_t\right] = G(T_0, T_1)^2 \mathrm{Var}\left[x(T_0) \,|\, \mathcal{F}_t\right] + \int_{T_0}^{T_1} \left[G(u, T)\sigma(u)\right]^2 du.$$

With continuous rate approximation compounded rate caplet can also be priced via Black formula

Corollary

With continuous rate approximation $\prod_{i=1}^{k}(1+L_i\tau_i) \approx \exp\left\{\int_{T_0}^{T_1} r(u)du\right\}$ *Theorem p.345 (thm:Ois-caplet-florlet-pricing) remains valid with the adjustment that log-variance* ν^2 *is replaced by* $\bar{\nu}^2$ *with*

$$\bar{\nu}^2 = G(T_0, T_1)^2 Var[x(T_0) \,|\, \mathcal{F}_t] + \int_{\max\{t,T_0\}}^{T_1} [G(u,T)\sigma(u)]^2 \, du.$$

How do log-variance ν^2 and $\bar{\nu}^2$ compare? (1/2)

We have (daily compounding)

$$\begin{aligned}\nu^2 &= G(T_0, T_1)^2 \mathrm{Var}\left[x(T_0) \,|\, \mathcal{F}_t\right] \\ &\quad + \sum_{j=1}^{k-1} \mathbb{1}_{\{t \le t_{j-1}\}} G(t_j, t_n)^2 \int_{t_{j-1}}^{t_j} \left[e^{-a(t_j - u)} \sigma(u)\right]^2 du \\ &\approx G(T_0, T_1)^2 \mathrm{Var}\left[x(T_0) \,|\, \mathcal{F}_t\right] + \sum_{j=1}^{k-1} \mathbb{1}_{\{t \le t_{j-1}\}} G(t_j, t_n)^2 \sigma(t_j)^2 \left(t_j - t_{j-1}\right)\end{aligned}$$

versus (continuous compounding)

$$\bar{\nu}^2 = G(T_0, T_1)^2 \mathrm{Var}\left[x(T_0) \,|\, \mathcal{F}_t\right] + \int_{\max\{t, T_0\}}^{T_1} \left[G(u, T)\sigma(u)\right]^2 du.$$

How do log-variance ν^2 and $\bar{\nu}^2$ compare? (2/2)

$$\nu^2 \approx G(T_0, T_1)^2 \mathrm{Var}\left[x(T_0) \,|\, \mathcal{F}_t\right] + \sum_{j=1}^{k-1} \mathbb{1}_{\{t \le t_{j-1}\}} G(t_j, t_n)^2 \sigma(t_j)^2 \left(t_j - t_{j-1}\right)$$

$$\bar{\nu}^2 = G(T_0, T_1)^2 \mathrm{Var}\left[x(T_0) \,|\, \mathcal{F}_t\right] + \int_{\max\{t, T_0\}}^{T_1} \left[G(u, T)\sigma(u)\right]^2 du.$$

- Variance from t to T_0, $G(T_0, T_1)^2 \mathrm{Var}\left[x(T_0) \,|\, \mathcal{F}_t\right]$, coincides in both approaches
- Variance during compounding period from T_0 to T_1 differs slightly between approaches

Log-variance ν^2 (daily compounding) can be viewed as numerical integration (or quadrature) scheme for $\bar{\nu}^2$ (continuous compounding).

Outline

Do we really need a term structure model - like Hull White model - to price caplets on compounded rates?

We establish a relation between standard (forward-looking) Libor rates and compounded (backward-looking) rates.

- Standard Libor rate with fixing time T, start time T_0 and end time T_1 (no tenor basis) is
$$L(T, T_0, T_1) = \left[\frac{P(T, T_0)}{P(T, T_1)} - 1\right] \frac{1}{\tau(T_0, T_1)}.$$
- We can define forward Libor rate $L(t, T_0, T_1)$ which *lives* for t prior to T.
- We have martingale property of forward Libor rates $L(t, T_0, T_1)$ for $t \le T$ and well understood Vanilla models
$$dL(t,) = \sigma_L(t) \cdot dW(t)$$
(e.g. Normal model, shifted SABR model, ... - depending on choice of $\sigma_L(t)$).

How can we extend Libor rate models to compounded rates $C_1 = \left\{ \left[\prod_{i=1}^{k} (1 + L_i \tau_i) \right] - 1 \right\} \frac{1}{\tau(T_0, T_1)}$?

We generalise the definition of forward Libor rates to capture backward-looking compounded rates

Use continuous rate approximation for overnight rate, $1 + L_i\tau_i \approx \exp\left\{\int_{t_{i-1}}^{t_i} r(u)du\right\}$. This yields

$$C_1 = \left\{\exp\left\{\int_{T_0}^{T_1} r(u)du\right\} - 1\right\}\frac{1}{\tau(T_0, T_1)}$$

Define generalised forward rate

$$R(t) = \frac{1}{\tau(T_0, T_1)}\begin{cases}\left[\frac{P(t,T_0)}{P(t,T_1)} - 1\right] & t \le T_0 \\ \left[\frac{\exp\left\{\int_{T_0}^{t} r(u)du\right\}}{P(t,T_1)} - 1\right] & T_0 < t \le T_1\end{cases}.$$

- $R(t)$ is a martingale in T_1-forward measure (by construction).
- $R(t)$ coincides with standard forward Libor rate $L(t, T_0, T_1)$ for all t until fixing time T.
- $R(T_1)$ is equal to compounded rate C_1.

Now we can specify a Vanilla model for the generalised forward rate

We specify a Vanilla model for the generalised forward rate as

$$dR(t) = \sigma_R(t) \cdot dW(t).$$

Here, $W(t)$ is a Brownian motion in T_1-forward measure and $\sigma_R(t)$ is an adapted volatility process.

How can we specify volatility $\sigma_R(t)$?

For $t \leq T$ $R(t) = L(t, T_0, T_1)$, thus also $dR(t) = dL(t,)$.

- We use standard Libor rate volatility $\sigma_R(t) = \sigma_L(t)$ for $t \leq T$.
- But what can we do for $T_0 < t \leq T_1$?

We need to take into account that between T_0 and T_1 more and more overnight rates get fixed

- At observation time $t \to T_1$ we get that $r(u)$, with $u \le t$ in $C_1 = \left\{ \exp\left\{ \int_{T_0}^{T_1} r(u) du \right\} - 1 \right\} \frac{1}{\tau(T_0, T_1)}$ is deterministic.
- Volatility of coupon decreases to zero as $t \to T_1$.

Assume linear decay of volatility of generalised forward rates,

$$\sigma_R(t) = \frac{T_1 - t}{T_1 - T_0} \cdot \sigma(t), \qquad T_0 < t \le T_1.$$

For backbone volatility $\sigma(t)$ we can use same type of model as for Libor volatility $\sigma_L(t)$.

Let's have a look at a simple example Vanilla model with normal dynamics and constant volatility

$$dR(t) = \min\left\{1, \frac{T_1 - t}{T_1 - T_0}\right\} \cdot \sigma \cdot dW(t).$$

- Final rate $R(T_1) = C_1$ is normally distributed. Option on C_1 can be priced with Bachelier formula
- Integrated variance of C_1 at observation (pricing) time $t < T_0$ becomes

$$\begin{aligned} \nu^2 &= \int_t^{T_1} \left[\min\left\{1, \frac{T_1 - t}{T_1 - T_0}\right\} \cdot \sigma\right]^2 dt \\ &= \sigma^2 \cdot (T_0 - t) + \frac{1}{3}\sigma^2 (T_1 - T_0) \cdot \end{aligned}$$

- Analogous derivation holds for shifted Log-normal model for $R(t)$
- Compare with integrated variance in Hull-White model for mean reversion $a \to 0$!

Outline

We can re-use Vanilla and term structure models to price caps and floors on compounded rate coupons

- Compounded overnight rate coupon rates are

$$C_1 = \left\{ \left[\prod_{i=1}^{k} (1 + L_i \tau_i) \right] - 1 \right\} \frac{1}{\tau} \approx \left\{ \exp\left\{ \int_{T_0}^{T_1} r(u) du \right\} - 1 \right\} \frac{1}{\tau}$$

- Terminal distribution of C_1 and caplets/floorlets on C_1 can be calculated using Hull-White model
- A generalisation of Libor forward rates to the compounding period T_0 to T_1 yields generalised forward rates $R(t)$ for which we can specify Vanilla models

Literature:

- A. Lyashenko and F. Mercurio. Looking forward to backward-looking rates: A modeling framework for term rates replacing libor. `https://ssrn.com/abstract=3330240`, 2019
- M. Henrard. A quant perspective on ibor fallback consultation results. `https://ssrn.com/abstract=3308766`, 2019

Part V

Bermudan Swaption Pricing

Outline

Outline

Let's have another look at the cancellation option

Interbank swap deal example

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(annually, 30/360 day count, modified following, Target calendar)

Pays 6-months Euribor floating rate on 100mm EUR

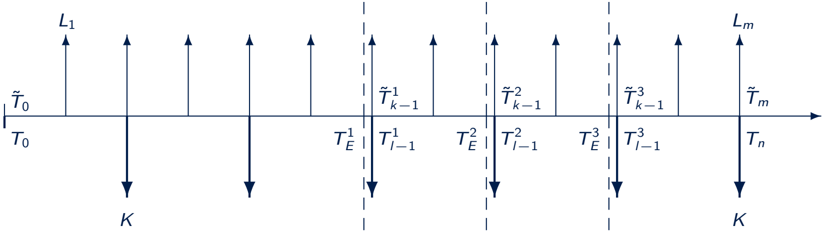
Start date: Oct 30, 2020

End date: Oct 30, 2040

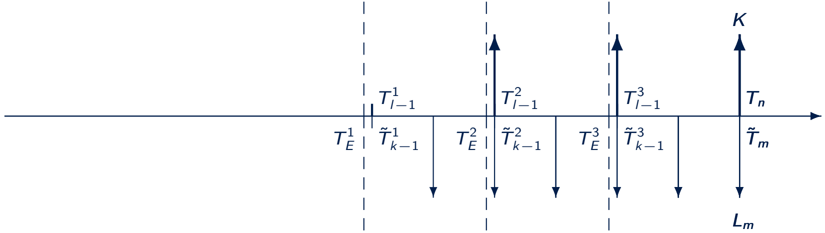
(semi-annually, act/360 day count, modified following, Target calendar)

Bank A may decide to early terminate deal in 10, 11, 12,..years.

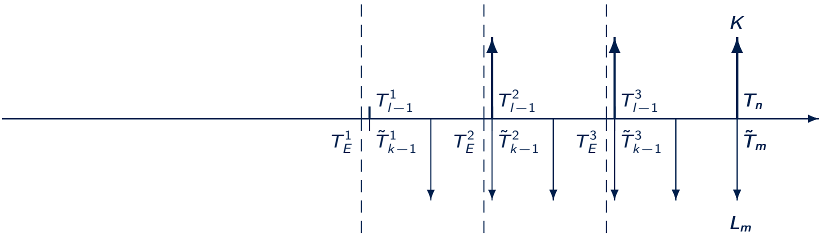
What does such a *Bermudan call right* mean?



[Bermudan cancellable swap] = [full swap] + [Bermudan option on opposite swap]



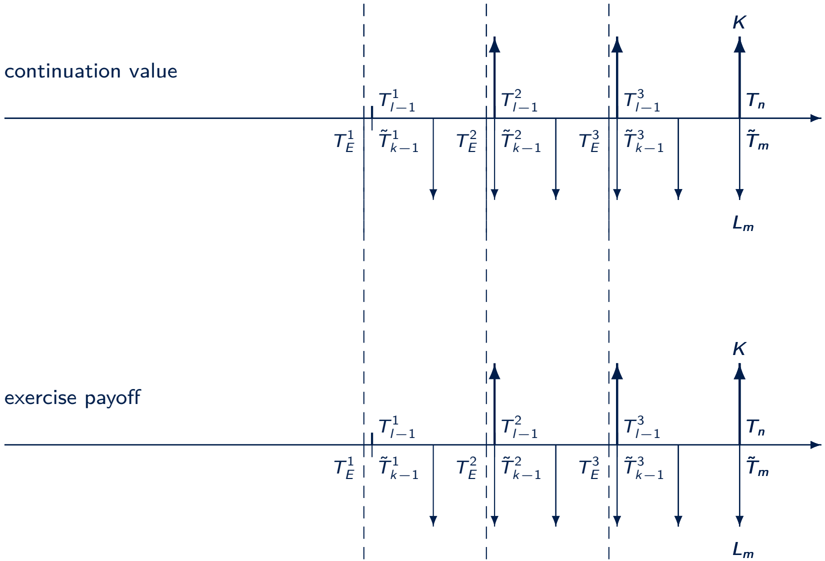
What is a Bermudan swaption?



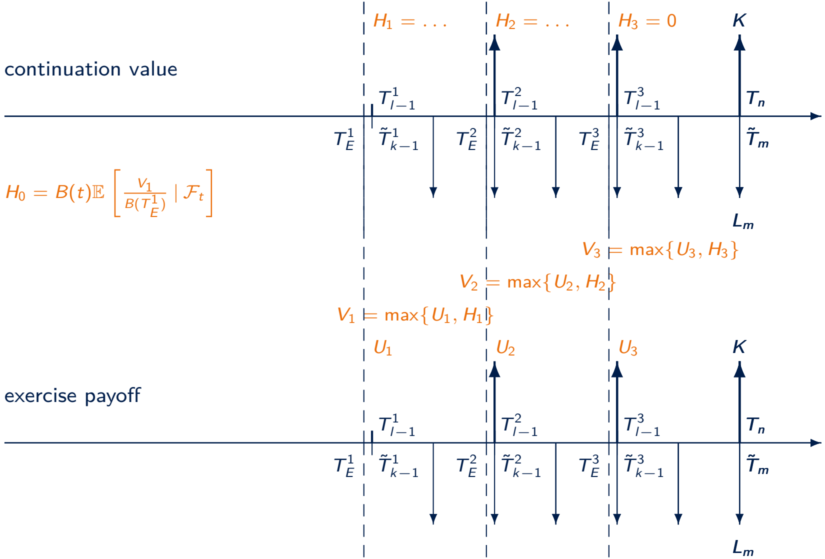
Bermudan swaption

A Bermudan swaption is an option to enter into a Vanilla swap with fixed rate K and final maturity T_n at various exercise dates $T_E^1, T_E^2, \ldots, T_E^{\bar{k}}$. If there is only one exercise date (i.e. $\bar{k} = 1$) then the Bermudan swaption equals a European swaption.

A Bermudan swaption can be priced via *backward induction*



A Bermudan swaption can be priced via *backward induction* - let's add some notation



First we specify the future payoff cash flows

- Choose a numeraire $B(t)$ and corresponding cond. expectations $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot \,|\, \mathcal{F}_t]$.
- Underlying payoff U_k if option is exercised

$$
\begin{aligned}
&U_k \\
&= B(T_E^k) \sum_{T_i \geq T_E^k} \mathbb{E}_{T_E^k}\left[\frac{X_i(T_i)}{B(T_i)}\right] \\
&= B(T_E^k) \underbrace{\left[\sum_{T_i \geq T_E^k} K\tau_i P(T_E^k, T_i) - \sum_{\tilde{T}_j \geq T_E^k} L^\delta(T_E^k, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)\tilde{\tau}_j P(T_E^k, \tilde{T}_j)\right]}_{\text{future fixed leg minus future float leg}} \\
&= B(T_E^k) \left[\sum_{T_i \geq T_E^k} K\tau_i P(T_E^k, T_i) - \left[P(T_E^k, \tilde{T}_{j_k}) - P(T_E^k, \tilde{T}_m)\right]\right. \\
&\qquad \left. - \sum_{\tilde{T}_j \geq T_E^k} P(T_E^k, \tilde{T}_{j-1}) \left[D(\tilde{T}_{j-1}, \tilde{T}_j) - 1\right]\right].
\end{aligned}
$$

Then we specify the continuation value and optimal exercise (1/2)

- Continuation value $H_k(t)$ ($T_E^k \leq t \leq T_E^{k+1}$) represents the time-$t$ value of the remaining option if not exercised.
- Option becomes worthless if not exercised at last exercise date $T_E^{\bar{k}}$. Thus last continuation value $H_{\bar{k}}(T_E^{\bar{k}}) = 0$.
- Recall that Bermudan option gives the right but not the obligation to enter into underlying at exercise.
- Rational agent will choose the maximum of payoff and continuation at exercise, i.e.
$$V_k = \max \left\{ U_k, H_k(T_E^k) \right\}.$$

Then we specify the continuation value and optimal exercise (2/2)

$$V_k = \max\left\{U_k, H_k(T_E^k)\right\}.$$

- V_k represents the Bermudan option value at exercise T_E^k. Thus we also must have for the continuation value

$$H_{k-1}(T_E^k) = V_k.$$

- Derivative pricing formula yields

$$\begin{aligned} H_{k-1}(T_E^{k-1}) &= B(T_E^{k-1}) \cdot \mathbb{E}_{T_E^{k-1}}\left[\frac{H_{k-1}(T_E^k)}{B(T_E^k)}\right] \\ &= B(T_E^{k-1}) \cdot \mathbb{E}_{T_E^{k-1}}\left[\frac{V_k}{B(T_E^k)}\right]. \end{aligned}$$

We summarize the Bermudan pricing algorithm

Backward induction for Bermudan options

Consider a Bermudan option with $\bar{k}$ exercise dates T_E^k $(k = 1, \ldots \bar{k})$ and underlying future payoffs with (time-T_E^k) prices U_k.

Denote $H_k(t)$ the option's continuation value for $T_E^k \leq t \leq T_E^{k+1}$ and set $H_{\bar{k}}\left(T_E^{\bar{k}}\right) = 0$. Also set $T_E^0 = t$ (i.e. pricing time today).

The option price can be derived via the recursion

$$
\begin{aligned}
H_k\left(T_E^k\right) &= B(T_E^k) \cdot \mathbb{E}_{T_E^k}\left[\frac{H_k(T_E^{k+1})}{B(T_E^{k+1})}\right] \\
&= B(T_E^k) \cdot \mathbb{E}_{T_E^k}\left[\frac{\max\left\{U_{k+1}, H_{k+1}(T_E^{k+1})\right\}}{B(T_E^{k+1})}\right].
\end{aligned}
$$

for $k = \bar{k} - 1, \ldots, 0$. The Bermudan option price is given by

$$
V^{\mathrm{Berm}}(t) = H_0(t) = H_0(T_E^0).
$$

Some more comments regarding Bermudan pricing ...

- Recursion for Bermudan pricing can be formally derived via theory of optimal stopping and Hamilton-Jacobi-Bellman (HJB) equation.
- For more details, see Sec. 18.2.2 in Andersen/Piterbarg (2010).
- For a single exercise date $\bar{k} = 1$ we get

$$H_0(t) = B(t) \cdot \mathbb{E}_t \left[\frac{\max\{U_1, 0)\}}{B(T_E^1)} \right].$$

  This is the general pricing formula for a European swaption (if U_1 represents a Vanilla swap).

- In principle, recursion $H_k\left(T_E^k\right) = B(T_E^k) \cdot \mathbb{E}_{T_E^k}\left[\frac{\max\left\{U_{k+1}, H_{k+1}(T_E^{k+1})\right\}}{B(T_E^{k+1})}\right]$ holds for any payoffs U_k. However, computation

$$U_k = B(T_E^k) \sum_{T_i \geq T_E^k} \mathbb{E}_{T_E^k}\left[\frac{X_i(T_i)}{B(T_i)}\right]$$

  might pose additional challenges if cash flows $X_i(T_i)$ are more complex.

How do we price a Bermudan in practice?

- In principle, recursion algorithm for $H_k()$ is straight forward.
- Computational challenge is calculating conditional expectations

$$H_k\left(T_E^k\right) = B(T_E^k) \cdot \mathbb{E}_{T_E^k}\left[\frac{\max\left\{U_{k+1}, H_{k+1}(T_E^{k+1})\right\}}{B(T_E^{k+1})}\right].$$

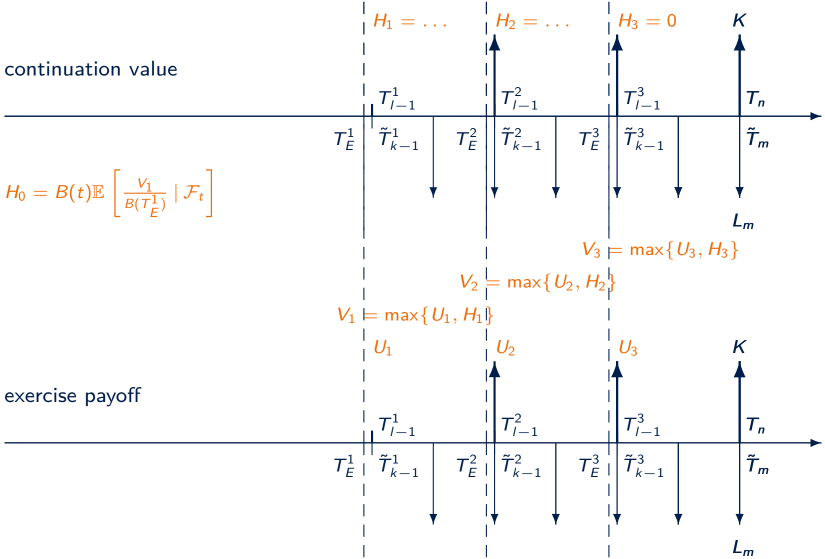
- Note, that this problem is an instance of the general option pricing problem

$$V(T_0) = B(T_0) \cdot \mathbb{E}\left[\frac{V(T_1)}{B(T_1)} \,|\, \mathcal{F}_{T_0}\right].$$

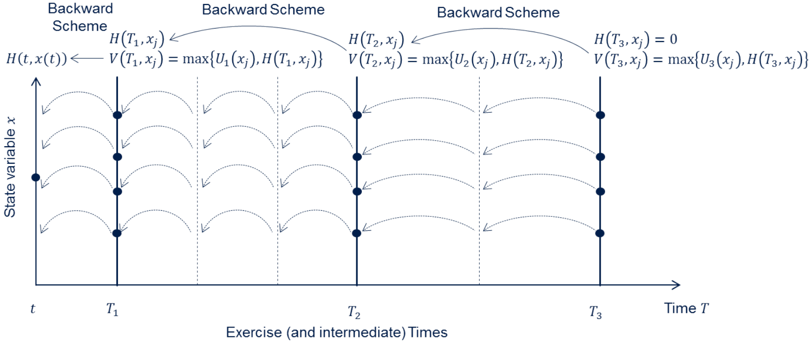
We can apply general option pricing methods to *roll-back* the Bermudan payoff.

Outline

Note that U_k, V_k and H_k depend on underlying state variable $x(T_E^k)$



Typically we need to discretise variables U_k, V_k and H_k on a grid of underlying state variables



Forthcomming, we discuss several methods to roll-back the payoffs.

Outline

Outline

Key idea is using the conditional density function in the Hull-White model

Recall that

$$V(T_0) = B(T_0) \cdot \mathbb{E}\left[\frac{V(T_1)}{B(T_1)} \,|\, \mathcal{F}_{T_0}\right].$$

We choose the T_1-maturing zero coupon bond $P(t, T_1)$ as numeraire. Then

$$\begin{aligned} V(T_0) &= P(T_0, T_1) \cdot \mathbb{E}^{T_1}\left[V(T_1) \,|\, \mathcal{F}_{T_0}\right] \\ &= P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} V(x; T_1) \cdot p_{\mu,\sigma^2}(x) \cdot dx. \end{aligned}$$

State variable $x = x(T_1)$ is normally distributed with known mean and variance.

Hull-White model results yield density parameters of the state variable $x(T_1)$

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} V(x; T_1) \cdot p_{\mu,\sigma^2}(x) \cdot dx.$$

State variable $x = x(T_1)$ is normally distributed with mean

$$\mu = \mathbb{E}^{T_1}\left[x(T_1) \,|\, \mathcal{F}_{T_0}\right] = G'(T_0, T_1)\left[x(T_0) + G(T_0, T_1) y(T_0)\right]$$

and variance

$$\sigma^2 = \mathrm{Var}\left[x(T_1) \,|\, \mathcal{F}_{T_0}\right] = y(T_1) - G'(T_0, T_1)^2 y(T_0).$$

Thus

$$p_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

and

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx.$$

Integral against normal density needs to be computed numerically by quadrature methods

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx.$$

- We can apply general purpose quadrature rules to function

$$f(x) = \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}.$$

  - Select a grid $[x_0, \ldots, x_N]$ and approximate e.g. via
  - Trapezoidal rule

$$\int_{-\infty}^{+\infty} f(x) \cdot dx \approx \sum_{i=1}^{N} \frac{1}{2} \left[f(x_{i-1}) + f(x_i)\right] (x_i - x_{i-1})$$

  - or Simpson's rule with equidistant grid $(h = x_i - x_{i-1})$ and even sub-intervalls, then

$$\int_{-\infty}^{+\infty} f(x) \cdot dx \approx \frac{h}{3} \cdot \left[f(x_0) + 2\sum_{j=1}^{N/2-1} f(x_{2j}) + 4\sum_{j=1}^{N/2} f(x_{2j-1}) + f(x_N)\right].$$

There are some details that need to be considered - Select your integration domain carefully

- Infinite integral is approximated by definite integral

$$\int_{-\infty}^{+\infty} f(x) \cdot dx \approx \int_{x_0}^{x_N} f(x) \cdot dx \approx \cdots .$$

- Finite integration boundaries need to be chosen carefully by taking into account variance of $x(t)$.
- One approach is calculating variance to option expiry T_1, $\hat{\sigma}^2 = \mathrm{Var}\left[x(T)\right] = y(T_1)$ and set

$$x_0 = -\lambda \cdot \hat{\sigma} \quad \text{and} \quad x_N = \lambda \cdot \hat{\sigma}.$$

- Note, that $\mathbb{E}^{T_1}\left[x(T_1)\right] = 0$, thus we do not apply a shift to the x-grid.

There are some details that need to be considered - Take care of the break-even point

- Note that convergence of quadrature rules depends on smoothness of integrand $f(x)$.
- Recall that

$$f(x) = V(x) \cdot p_{\mu,\sigma^2}(x) = \max\left\{U_{k+1}(x), H_{k+1}(x; T_E^{k+1})\right\} \cdot p_{\mu,\sigma^2}(x).$$

- Max-function is not smooth at $U_{k+1}(x) = H_{k+1}(x; T_E^{k+1})$.

Determine $x^\star$ (via interpolation of $H_{k+1}(\cdot)$ and numerical root search) such that

$$U_{k+1}(x^\star) = H_{k+1}(x^\star; T_E^{k+1})$$

and split integration

$$\int_{-\infty}^{+\infty} f(x) \cdot dx = \int_{-\infty}^{x^\star} f(x) \cdot dx + \int_{x^\star}^{+\infty} f(x) \cdot dx.$$

Can we exploit the structure of the integrand?

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx.$$

- Integral against normal distribution density can be solved more efficiently:

1. Use Gauss–Hermite quadrature.
2. Interpolate only $V(x; T_1)$ by cubic spline and integrate exact.

Outline

Gauss–Hermite quadrature is an efficient integration method for smooth integrands

- Gauss–Hermite quadrature is a particular form of Gaussian quadrature.
- Choose a degree parameter d, and approximate

$$\int_{-\infty}^{+\infty} f(x) \cdot e^{-x^2} dx \approx \sum_{k=1}^{d} w_k \cdot f(x_k)$$

  with x_k $(i = 1, 2, ..., d)$ being the roots of the physicists' version of the Hermite polynomial $H_d(x)$ and w_k are weights with

$$w_k = \frac{2^{d-1} d! \sqrt{\pi}}{d^2 \left[H_{d-1}(x_k)\right]^2}.$$

- Roots and weights can be obtained, e.g. via stored values and look-up tables.

Variable transformation allows application of Gauss–Hermite quadrature to Hull-White model integration

We get

$$\int_{-\infty}^{+\infty} \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} V(\sqrt{2}\sigma x + \mu; T_1) \cdot e^{-x^2} dx$$

$$\approx \frac{1}{\sqrt{\pi}} \sum_{k=1}^{d} w_k \cdot V(\sqrt{2}\sigma x_k + \mu; T_1).$$

Some constraints need to be considered:

- Payoff $V(\cdot, T_1)$ is only available on the x-grid at T_1, thus $V(\cdot, T_1)$ needs to be interpolated.
- Gauss-Hermite quadrature does not take care of non-smooth payoff at break-even state $x^\star$, thus d needs to be sufficiently large to mitigate impact.

Outline

If we apply cubic spline interpolation anyway then we can also integrate exactly

Approximate $V(\cdot, T_1)$ via cubic spline on the grid $[x_0, \dots x_N]$ as

$$V(x, T_1) \approx C(x) = \sum_{i=0}^{N-1} \mathbb{1}_{\{x_i \le x < x_{i+1}\}} \sum_{k=0}^{d} c_{i,k} \cdot (x - x_i)^k .$$

Then

$$\begin{aligned}\int_{-\infty}^{+\infty} V(x; T_1) \cdot p_{\mu,\sigma^2}(x) \cdot dx &\approx \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} \sum_{k=0}^{d} c_{i,k} \cdot (x - x_i)^k \cdot p_{\mu,\sigma^2}(x) \cdot dx \\ &= \sum_{i=0}^{N-1} \sum_{k=0}^{d} c_{i,k} \cdot \int_{x_i}^{x_{i+1}} (x - x_i)^k \cdot p_{\mu,\sigma^2}(x) \cdot dx.\end{aligned}$$

Thus, all we need is

$$I_{i,k} = \int_{x_i}^{x_{i+1}} (x - x_i)^k \cdot p_{\mu,\sigma^2}(x) \cdot dx.$$

We transform variables to make integration easier

First we apply the variable transformation $\bar{x} = (x - \mu)/\sigma$. This yields $p_{\mu,\sigma^2}(x) = p_{0,1}(\bar{x})/\sigma$ and

$$
\begin{aligned}
I_{i,k} &= \int_{\bar{x}_i}^{\bar{x}_{i+1}} (\sigma \bar{x} + \mu - x_i)^k \cdot p_{0,1}(\bar{x}) \cdot \frac{dx}{\sigma} \\
&= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^k \left(\bar{x} - \bar{x}_i\right)^k \cdot \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\bar{x}^2}{2}\right\}}_{\text{standard normal density}} d\bar{x}
\end{aligned}
$$

with the shifted grid points $\bar{x}_i = (x_i - \mu)/\sigma$.
Denote $\Phi(\cdot)$ the cumulated standard normal distribution function. Then

$$
\Phi'(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\bar{x}^2}{2}\right\} \quad \text{and} \quad \Phi''(x) = -x\Phi'(x).
$$

As a sub-step we aim at solving the integral

$$
\int_{\bar{x}_i}^{\bar{x}_{i+1}} \bar{x}^k \cdot \Phi'(\bar{x}) \cdot d\bar{x}.
$$

We use cubic splines ($d = 3$) to keep formulas reasonably simple I

It turnes out that

$$F_0(\bar{x}) = \int \Phi'(\bar{x}) d\bar{x} = \Phi(\bar{x}),$$

$$F_1(\bar{x}) = \int \bar{x}\Phi'(\bar{x}) d\bar{x} = -\Phi'(\bar{x}),$$

$$F_2(\bar{x}) = \int \bar{x}^2\Phi'(\bar{x}) d\bar{x} = \Phi(\bar{x}) - x \cdot \Phi'(\bar{x}),$$

$$F_3(\bar{x}) = \int \bar{x}^3\Phi'(\bar{x}) d\bar{x} = -\left(\bar{x}^2 + 2\right) \cdot \Phi'(\bar{x}).$$

This yields for $I_{i,0}$

$$I_{i,0} = \int_{\bar{x}_i}^{\bar{x}_{i+1}} \Phi'(\bar{x}) \cdot dx = F_0(\bar{x}_{i+1}) - F_0(\bar{x}_i)$$

We use cubic splines ($d = 3$) to keep formulas reasonably simple II

and for $I_{i,1}$

$$
\begin{aligned}
I_{i,1} &= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma \left(\bar{x} - \bar{x}_i\right) \cdot \Phi'(\bar{x}) \cdot dx \\
&= \sigma \cdot \int_{\bar{x}_i}^{\bar{x}_{i+1}} \bar{x} \cdot \Phi'(\bar{x}) \cdot dx - \sigma \cdot \bar{x}_i \cdot I_{i,0} \\
&= \sigma \cdot \left[F_1(\bar{x}_{i+1}) - F_1(\bar{x}_i)\right] - \sigma \cdot \bar{x}_i \cdot I_{i,0}.
\end{aligned}
$$

We use cubic splines ($d = 3$) to keep formulas reasonably simple III

We may proceed similarly for $I_{i,2}$

$$
\begin{aligned}
I_{i,2} &= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^2 \left(\bar{x} - \bar{x}_i\right)^2 \cdot \Phi'(\bar{x}) \cdot dx \\
&= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^2 \left[\bar{x}^2 - 2\bar{x}_i\bar{x} + \bar{x}_i^2\right] \cdot \Phi'(\bar{x}) \cdot dx \\
&= \sigma^2 \left[F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)\right] - 2\sigma^2\bar{x}_i \left[F_1(\bar{x}_{i+1}) - F_1(\bar{x}_i)\right] + \sigma^2\bar{x}_i^2 I_{i,0} \\
&= \sigma^2 \left[F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)\right] - 2\sigma\bar{x}_i \left[I_{i,1} + \sigma \cdot \bar{x}_i \cdot I_{i,0}\right] + \sigma^2\bar{x}_i^2 I_{i,0} \\
&= \sigma^2 \left[F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)\right] - 2\sigma\bar{x}_i I_{i,1} - \sigma^2\bar{x}_i^2 I_{i,0}
\end{aligned}
$$

We use cubic splines ($d = 3$) to keep formulas reasonably simple IV

and $I_{i,3}$

$$
\begin{aligned}
I_{i,3} &= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^3 \left(\bar{x} - \bar{x}_i\right)^3 \cdot \Phi'(\bar{x}) \cdot dx \\
&= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^3 \left[\bar{x}^3 - 3\bar{x}_i\bar{x}^2 + 3\bar{x}_i^2\bar{x} - \bar{x}_i^3\right] \cdot \Phi'(\bar{x}) \cdot dx \\
&= \sigma^3 \left[F_3(\bar{x}_{i+1}) - F_3(\bar{x}_i)\right] - 3\sigma^3\bar{x}_i \left[F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)\right] \\
&\quad + 3\sigma^3\bar{x}_i^2 \left[F_1(\bar{x}_{i+1}) - F_1(\bar{x}_i)\right] - \sigma^3\bar{x}_i^3 I_{i,0}.
\end{aligned}
$$

Substituting terms as before yields

$$
\begin{aligned}
I_{i,3} &= \sigma^3 \left[F_3(\bar{x}_{i+1}) - F_3(\bar{x}_i)\right] - 3\sigma\bar{x}_i \left[I_{i,2} + 2\sigma\bar{x}_i I_{i,1} + \sigma^2\bar{x}_i^2 I_{i,0}\right] \\
&\quad + 3\sigma^2\bar{x}_i^2 \left[I_{i,1} + \sigma \cdot \bar{x}_i \cdot I_{i,0}\right] - \sigma^3\bar{x}_i^3 I_{i,0} \\
&= \sigma^3 \left[F_3(\bar{x}_{i+1}) - F_3(\bar{x}_i)\right] - 3\sigma\bar{x}_i I_{i,2} - 3\sigma^2\bar{x}_i^2 I_{i,1} - \sigma^3\bar{x}_i^3 I_{i,0}.
\end{aligned}
$$

Let's summarise the formulas...

We get

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} V(x; T_1) \cdot p_{\mu,\sigma^2}(x) \cdot dx$$
$$\approx P(x(T_0); T_0, T_1) \cdot \sum_{i=0}^{N-1} \sum_{k=0}^{3} c_{i,k} \cdot I_{i,k}$$

with

$$I_{i,0} = F_0(\bar{x}_{i+1}) - F_0(\bar{x}_i)$$
$$I_{i,1} = \sigma \cdot [F_1(\bar{x}_{i+1}) - F_1(\bar{x}_i)] - \sigma \cdot \bar{x}_i \cdot I_{i,0}$$
$$I_{i,2} = \sigma^2 [F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)] - 2\sigma \bar{x}_i I_{i,1} - \sigma^2 \bar{x}_i^2 I_{i,0}$$
$$I_{i,3} = \sigma^3 [F_3(\bar{x}_{i+1}) - F_3(\bar{x}_i)] - 3\sigma \bar{x}_i I_{i,2} - 3\sigma^2 \bar{x}_i^2 I_{i,1} - \sigma^3 \bar{x}_i^3 I_{i,0}$$

and anti-derivative functions $F_k(x)$ as before.

Integrating a cubic spline versus a normal density function occurs in various contexts of pricing methods

- Method already yields good accuracy for smaller number of grid points.
- For larger number of grid points accuracy benefit compared to e.g. Simpson integration seems not too much.
- Either way, use special treatment of break-even point $x^\star$.
- Computational effort can become significant for larger number of grid points.
  - Bermudan pricing requires N^2 evaluations of $\Phi(\cdot)$ and $\Phi'(\cdot)$ per exercise.

Outline

PDE methods for finance and pricing are extensively studied in the literature

- We present the basic principles and some aspects relevant for Bermudan bond option pricing.
- Further reading:
  - L. Andersen and V. Piterbarg. *Interest rate modelling, volume I to III.* Atlantic Financial Press, 2010, Sec. 2.
  - D. Duffy. *Finite Difference Methods in Financial Engineering.* Wiley Finance, 2006

Outline

We can adapt the Black-Scholes equation to our Hull-White model setting

- Recall that state variable $x(t)$ follows the risk-neutral dynamics

$$dx(t) = \underbrace{\left[y(t) - a \cdot x(t)\right]}_{\mu(t,x(t))} dt + \sigma(t) \cdot dW(t).$$

- Consider an option with price $V = V(t, x(t))$, option expiry time T and payoff $V(T, x(T)) = g(x(T))$.

- Derivative pricing formula yields that discounted option price is a martingale, i.e.

$$d\left(\frac{V(t, x(t))}{B(t)}\right) = \sigma_V(t, x(t)) \cdot dW(t).$$

How can we use this to derive a PDE?

Apply Ito's Lemma to the discounted option price

We get

$$d\left(\frac{V\left(t,x(t)\right)}{B(t)}\right)=\frac{dV\left(t,x(t)\right)}{B(t)}+V(t)d\left(\frac{1}{B(t)}\right).$$

With $d\left(B(t)^{-1}\right)=-r(t)\cdot B(t)^{-1}\cdot dt$ follows

$$d\left(\frac{V\left(t,x(t)\right)}{B(t)}\right)=\frac{1}{B(t)}\left[dV\left(t,x(t)\right)-r(t)\cdot V(t)\cdot dt\right].$$

Applying Ito's Lemma to option price $V\left(t,x(t)\right)$ gives

$$\begin{aligned}dV\left(t,x(t)\right)&=V_t\cdot dt+V_x\cdot dx(t)+\frac{1}{2}V_{xx}\cdot\left[dx(t)\right]^2\\&=\left[V_t+V_x\cdot\mu\left(t,x(t)\right)+\frac{1}{2}V_{xx}\cdot\sigma(t)^2\right]dt+V_x\cdot\sigma(t)\cdot dW(t)\end{aligned}$$

with partial derivatives $V_t=\partial V\left(t,x(t)\right)/\partial t$, $V_x=\partial V\left(t,x(t)\right)/\partial x$ and $V_{xx}=\partial^2 V\left(t,x(t)\right)/\partial x^2$.

Combining results yields dynamics of discounted option price

$$d\left(\frac{V\left(t,x(t)\right)}{B(t)}\right) = \frac{1}{B(t)}\underbrace{\left[V_t + V_x \cdot \mu\left(t,x(t)\right) + \frac{1}{2}V_{xx}\cdot\sigma(t)^2 - r(t)\cdot V\right]}_{\mu_V(t,x(t))}dt$$

$$+ \underbrace{\frac{V_x\cdot\sigma(t)}{B(t)}}_{\sigma_V(t,x(t))}\cdot dW(t).$$

Martingale property of $\frac{V(t,x(t))}{B(t)}$ requires that drift vanishes. That is

$$\mu_V(t,x(t)) = V_t + V_x \cdot \mu\left(t,x(t)\right) + \frac{1}{2}V_{xx}\cdot\sigma(t)^2 - r(t)\cdot V = 0.$$

Substituting $\mu\left(t,x(t)\right) = y(t) - a\cdot x(t)$ and $r(t) = f(0,t) + x(t)$ yields pricing PDE.

We get the parabolic pricing PDE with terminal condition

Theorem (Derivative pricing PDE in Hull-White model)

Consider our Hull-White model setup and a derivative security with price process $V(t, x(t))$ that pays at time T the payoff $V(T, x(T)) = g(x(T))$. Further assume $V(T, x(T))$ has finite variance and is attainable.
Then for $t < T$ the option price

$$V(t, x(t)) = B(t) \cdot \mathbb{E}^{\mathbb{Q}} \left[\frac{V(T, x(T))}{B(T)} \,|\, \mathcal{F}_t \right]$$

follows the PDE

$$V_t(t,x) + [y(t) - a \cdot x] \cdot V_x(t,x) + \frac{\sigma(t)^2}{2} \cdot V_{xx}(t,x) = [x + f(0,t)] \cdot V(t,x)$$

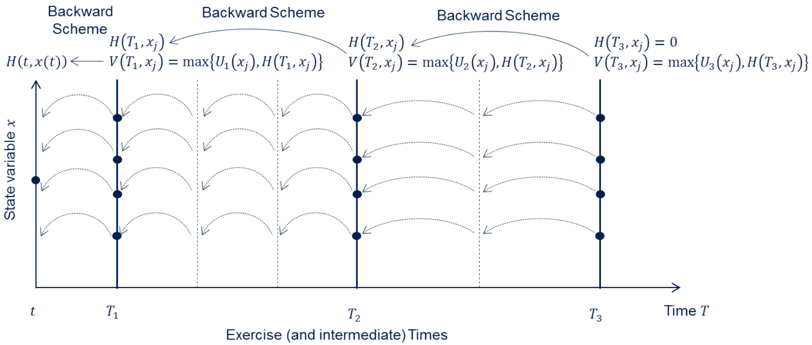
with terminal condition

$$V(T, x) = g(x).$$

Proof.

Follows from derivation above. $\square$

How does this help for our Bermudan option pricing problem?



- We need option prices on a grid of state variables $[x_0, \ldots x_N]$

Solve Hull-White option pricing PDE backwards from exercise to exercise.

Pricing PDE is typically solved via finite difference scheme and time integration

- Use *method of lines (MOL)* to solve parabolic PDE:
  - First discretise state space.
  - Then integrate resulting system of ODEs with terminal condition in time-direction.
- We discuss basic (or general purpose) approach to solve PDE; for a detailed treatment see Andersen/Piterbarg (2010) or Duffy (2006).
- Some aspects may require special attention in the context of Hull-White model:
  - more problem-specific boundary discretisation,
  - non-equidistant grids with finer grid around break-even state $x^\star$,
  - upwinding schemes to deal with potentially dominant impact of convection term $[y(t) - a \cdot x] \cdot V_x(t, x)$ at the grid boundaries of x.

Outline

How do we discretise state space?

- PDE for $V(t, x(t))$ is defined on infinite domain $(-\infty, +\infty)$.
  - We don't get explicit boundary conditions from PDE derivation.
  - Thus, we require payoff-specific approximation.
  - Finally, we are interested in $V(0, 0)$.

- We use equidistant x-grid $x_0, \ldots, x_N$ with grid size h_x centered around zero and scaled via standard deviation of $x(T)$ at final maturity T ,
$$x_0 = -\lambda \cdot \hat{\sigma} \quad \text{and} \quad x_N = \lambda \cdot \hat{\sigma}$$
with $\hat{\sigma}^2 = \mathrm{Var}\left[x(T)\right] = y(T)$ and $\lambda \approx 5$.

- Why not shift the grid by expectation $\mathbb{E}\left[x(T)\right]$ (as suggested in the literature)?
  - Pricing PDE is independent of pricing measure (used for derivation).
  - There is no *natural* measure under which $\mathbb{E}\left[x(T)\right]$ should be calculated.
  - In T-forward measure $\mathbb{E}^T\left[x(T)\right] = 0$ anyway.

Differential operators in state-dimension are discretised via central finite differences

For now leave time t continuous. We use notation $V(\cdot, x)$.
For inner grid points x_i with $i = 1, \ldots, N-1$

$$V_x(\cdot, x_i) = \frac{V(\cdot, x_{i+1}) - V(\cdot, x_{i-1})}{2h_x} + \mathcal{O}(h_x^2) \quad \text{and}$$

$$V_{xx}(\cdot, x_i) = \frac{V(\cdot, x_{i+1}) - 2V(\cdot, x_i) + V(\cdot, x_{i-1})}{h_x^2} + \mathcal{O}(h_x^2).$$

At the boundaries we impose condition

$$V_{xx}(\cdot, x_0) = \lambda_0 \cdot V_x(\cdot, x_0) \quad \text{and} \quad V_{xx}(\cdot, x_N) = \lambda_N \cdot V_x(\cdot, x_N).$$

This yields one-sided first order partial derivative approximations

$$V_x(\cdot, x_0) \approx \frac{2\left[V(\cdot, x_1) - V(\cdot, x_0)\right]}{(2 + \lambda_0 h_x)\, h_x} \quad \text{and} \quad V_x(\cdot, x_N) \approx \frac{2\left[V(\cdot, x_N) - V(\cdot, x_{N-1})\right]}{(2 - \lambda_N h_x)\, h_x}.$$

Some initial comments regarding choice of $\lambda_{0,N}$

- Often, $\lambda_{0,N} = 0$ (also suggested in the literature).

- With $\lambda_{0,N} = 0$ we have $V_{xx}(\cdot, x_0) = V_{xx}(\cdot, x_N) = 0$ and

$$V_x(\cdot, x_0) = \frac{V(\cdot, x_1) - V(\cdot, x_0)}{h_x} + \mathcal{O}(h_x^2) \quad \text{and}$$

$$V_x(\cdot, x_N) = \frac{V(\cdot, x_N) - V(\cdot, x_{N-1})}{h_x} + \mathcal{O}(h_x^2).$$

- However, for bond options the choice $V_{xx}(\cdot, x_0) = V_{xx}(\cdot, x_N) = 0$ might be a poor approximation.

- We will discuss an alternative choice for $\lambda_{0,N}$ later.

Now consider PDE for each grid point individually

Define the vector-valued function $v(t)$ via

$$v(t) = [v_0(t), \dots, v_N(t)]^\top = [V(t, x_0), \dots, V(t, x_N)]^\top \in \mathbb{R}^{N+1}.$$

Then state discretisation yields for inner points x_i with $i = 1, \dots, N-1$,

$$v_i'(t) + [y(t) - a x_i] \frac{v_{i+1}(t) - v_{i-1}(t)}{2h_x} + \frac{\sigma(t)^2}{2} \frac{v_{i+1}(t) - 2v_i(t) + v_{i-1}(t)}{h_x^2} = [x_i + f(0,t)]\, v_i(t)$$

and for the boundaries

$$v_0'(t) + \left[y(t) - a x_0 + \lambda_0 \frac{\sigma(t)^2}{2} \right] \frac{2\,[v_1(t) - v_0(t)]}{(2 + \lambda_0 h_x)\, h_x} = [x_0 + f(0,t)]\, v_0(t),$$

$$v_N'(t) + \left[y(t) - a x_N + \lambda_N \frac{\sigma(t)^2}{2} \right] \frac{2\,[v_N(t) - v_{N-1}(t)]}{(2 - \lambda_N h_x)\, h_x} = [x_N + f(0,t)]\, v_N(t).$$

As before, we have the terminal condition

$$v_i(T) = g(x_i).$$

Parabolic PDE is transformed into linear system of ODEs with terminal condition.

It is more convenient to write system of ODEs in matrix-vector notation (1/2)

We get

$$v'(t) = M(t) \cdot v(t) = \begin{bmatrix} c_0 & u_0 & & \\ l_1 & \ddots & \ddots & \\ & \ddots & \ddots & u_{N-1} \\ & & l_N & c_N \end{bmatrix} \cdot v(t)$$

with time-dependent inner components c_i, l_i, u_i $(i = 1, \dots N-1)$,

$$c_i = \frac{\sigma(t)^2}{h_x^2} + x_i + f(0,t),$$

$$l_i = -\frac{\sigma(t)^2}{2h_x^2} + \frac{y(t) - ax_i}{2h_x},$$

$$u_i = -\frac{\sigma(t)^2}{2h_x^2} - \frac{y(t) - ax_i}{2h_x}.$$

It is more convenient to write system of ODEs in matrix-vector notation (2/2)

Boundary elements of $M(t)$ become

$$
\begin{aligned}
c_0 =& \frac{2\left[y(t) - a x_0 + \lambda_0 \frac{\sigma(t)^2}{2}\right]}{\left(2 + \lambda_0 h_x\right) h_x} + x_0 + f(0,t), \\
c_N =& - \frac{2\left[y(t) - a x_N + \lambda_N \frac{\sigma(t)^2}{2}\right]}{\left(2 - \lambda_N h_x\right) h_x} + x_0 + f(0,t), \\
u_0 =& - \frac{2\left[y(t) - a x_0 + \lambda_0 \frac{\sigma(t)^2}{2}\right]}{\left(2 + \lambda_0 h_x\right) h_x}, \\
l_N =& \frac{2\left[y(t) - a x_N + \lambda_N \frac{\sigma(t)^2}{2}\right]}{\left(2 - \lambda_N h_x\right) h_x}.
\end{aligned}
$$

Outline

Linear system of ODEs can be solved by standard methods

We have
$$v'(t) = f\left(t, v(t)\right) = M(t) \cdot v(t).$$

We demonstrate time discretisation based on θ-*method.* Consider equidistant time grid $t = t_0, \ldots, t_M = T$ with step size h_t and approximation

$$\frac{v(t_{j+1}) - v(t_j)}{h_t} \approx f\left(t_{j+1} - \theta h_t, (1-\theta)v(t_{j+1}) + \theta v(t_j)\right)$$

for $\theta \in [0, 1]$.

- In general, approximation yields method of order $\mathcal{O}(h_t)$.
- For $\theta = \frac{1}{2}$, approximation yields method of order $\mathcal{O}(h_t^2)$.

For our linear ODE we set $v^j = v(t_j)$, $M_\theta = M(t_{j+1} - \theta h_t)$ and get the scheme
$$\frac{v^{j+1} - v^j}{h_t} = M_\theta \left[(1-\theta)v^{j+1} + \theta v^j\right].$$

We get a recursion for the θ-method

Rearranging terms yields

$$[I + h_t \theta M_\theta] \, v^j = [I - h_t (1 - \theta) M_\theta] \, v^{j+1}.$$

If $[I + h_t \theta M_\theta]$ is regular then we can solve for v^j via

$$v^j = [I + h_t \theta M_\theta]^{-1} [I - h_t (1 - \theta) M_\theta] \, v^{j+1}.$$

Terminal condition is

$$v^M = [g(x_0), \ldots, g(x_N)]^\top .$$

- Unless $\theta = 0$ (Explicit Euler scheme) we need to solve a linear equation system.
- Fortunately, matrix $[I + h_t \theta M_\theta]$ is tri-diagonal; solution requires $\mathcal{O}(M)$ operations.
- θ-method is A-stable for $\theta \geq \frac{1}{2}$.
- However, oscillations in solution may occur unless $\theta = 1$ (Implicit Euler scheme, L-stable).

Outline

Let's have another look at the boundary condition ...

We look at an example of a zero coupon bond option with payoff

$$V(x,T) = \left[P(x,T,T') - K\right]^+.$$

For $x \ll 0$ option is far in-the-money and $V(x,t)$ can be approximated by intrinsic value $V(x,t) \approx \tilde{V}(x,t)$ with

$$\tilde{V}(x,t) = \left[P(x,t,T') - K\right]^+ = \left[\frac{P(0,T')}{P(0,t)} e^{-G(t,T)x - \frac{1}{2}G(t,T)^2 y(t)} - K\right]^+.$$

This yields

$$\frac{\partial}{\partial x}\tilde{V}(x,t) = -G(t,T)\left[\tilde{V}(x,t) + K\right]$$

and

$$\frac{\partial^2}{\partial x^2}\tilde{V}(x,t) = \underbrace{-G(t,T)}_{\lambda}\frac{\partial}{\partial x}\tilde{V}(x,t).$$

Alternatively, for $x \gg 0$ option is far out-of-the-money and

$$\frac{\partial^2}{\partial x^2}\tilde{V}(x,t) = \frac{\partial}{\partial x}\tilde{V}(x,t) = 0.$$

We adapt approximation to our option pricing problem

- In principle, for a coupon bond underlying we could estimate $\lambda = \lambda(t)$ via option intrinsic value $\tilde{V}(x,t)$ and

$$\lambda(t) = \left[\frac{\partial^2}{\partial x^2}\tilde{V}(x,t)\right] / \frac{\partial}{\partial x}\tilde{V}(x,t) \quad \text{for} \quad \frac{\partial}{\partial x}\tilde{V}(x,t) \neq 0,$$

  otherwise $\lambda(t) = 0$.

- We take a more rough approach by approximating λ based only on previous solution

$$\begin{aligned}
\lambda_{0,N} &= \left[\frac{\partial^2}{\partial x^2}V(x,t)\right] / \frac{\partial}{\partial x}V(x,t) \\
&\approx \left[\frac{\partial^2}{\partial x^2}V(x_{1,N-1}, t+h_t)\right] / \frac{\partial}{\partial x}V(x_{1,N-1}, t+h_t) \\
&\approx \frac{v_{0,N-2}^{j+1} - 2v_{1,N-1}^{j+1} + v_{2,N}^{j+1}}{h_x^2} / \frac{v_{2,N}^{j+1} - v_{0,N-2}^{j+1}}{2h_x}
\end{aligned}$$

  for $v_{2,N}^{j+1} - v_{0,N-2}^{j+1}/(2h_x) \neq 0$, otherwise $\lambda_{0,N} = 0$.

It turns out that accuracy of one-sided first order derivative approximation is of order $\mathcal{O}(h_x^2)$ I

Lemma

Assume $V = V(x)$ is twice continuously differentiable. Moreover, consider grid points x_{-1}, x_0, x_1 with equal spacing $h_x = x_1 - x_0 = x_0 - x_{-1}$. If there is a $\lambda_0 \in \mathbb{R}$ such that

$$V''(x_0) = \lambda_0 \cdot V'(x_0)$$

then

$$V'(x_0) = \frac{2\left[V(x_1) - V(x_0)\right]}{(2 + \lambda_0 h_x)\, h_x} + \mathcal{O}(h_x^2).$$

Proof:

Denote $v_i = V(x_i)$. We have from standard Taylor approximation

$$V''(x_0) = \frac{v_{-1} - 2v_0 + v_1}{h_x^2} + \mathcal{O}(h_x^2) \quad \text{and} \quad V'(x_0) = \frac{v_1 - v_{-1}}{2h_x} + \mathcal{O}(h_x^2).$$

It turns out that accuracy of one-sided first order derivative approximation is of order $\mathcal{O}(h_x^2)$ II

From $V''(x_0) = \lambda \cdot V'(x_0)$ follows

$$\frac{v_{-1} - 2v_0 + v_1}{h_x^2} + \mathcal{O}(h_x^2) = \lambda_0 \left[\frac{v_1 - v_{-1}}{2h_x} + \mathcal{O}(h_x^2) \right].$$

Multiplying with $2h_x^2$ gives the relation

$$2\left(v_{-1} - 2v_0 + v_1\right) + \mathcal{O}(h_x^4) = \lambda_0 h_x \left(v_1 - v_{-1}\right) + \mathcal{O}(h_x^4).$$

Reordering terms yields

$$\left(2 + \lambda_0 h_x\right) v_{-1} = 4v_0 + \left(\lambda_0 h_x - 2\right) v_1 + \mathcal{O}(h_x^4).$$

And solving for v_{-1} gives
$v_{-1} = \left[4v_0 + \left(\lambda_0 h_x - 2\right) v_1\right] / \left(2 + \lambda_0 h_x\right) + \mathcal{O}(h_x^4)$.

It turns out that accuracy of one-sided first order derivative approximation is of order $\mathcal{O}(h_x^2)$ III

Now, we substitute v_{-1} in the approximation for $V'(x)$. This gives

$$
\begin{aligned}
V'(x_0) &= \frac{v_1 - \left[\left[4v_0 + (\lambda_0 h_x - 2)\, v_1 \right] / \left(2 + \lambda_0 h_x \right) + \mathcal{O}(h_x^4) \right]}{2h_x} + \mathcal{O}(h_x^2) \\
&= \frac{(2 + \lambda_0 h_x)\, v_1 - \left[4v_0 + (\lambda_0 h_x - 2)\, v_1 \right]}{2\,(2 + \lambda_0 h_x)\, h_x} + \mathcal{O}(h_x^2) + \mathcal{O}(h_x^3) \\
&= \frac{2v_1 - 4v_0 + 2v_1}{2\,(2 + \lambda_0 h_x)\, h_x} + \mathcal{O}(h_x^2) \\
&= \frac{2\,(v_1 - v_0)}{(2 + \lambda_0 h_x)\, h_x} + \mathcal{O}(h_x^2).
\end{aligned}
$$

- With constraint $V''(x_0) = \lambda \cdot V'(x_0)$ we can eliminate explicit dependence on second derivative $V''(x_0)$ and outer grid point $v_{-1} = V(x_{-1})$.

It turns out that accuracy of one-sided first order derivative approximation is of order $\mathcal{O}(h_x^2)$ IV

- Analogous result can be derived for upper boundery and down-ward approximation of first derivative.
- Resulting scheme is still second order accurate in state space direction.

Outline

We summarise the PDE pricing method

1. Discretise state space x on a grid $[x_0, \ldots, x_N]$ and specify time step size h_t and $\theta \in [0, 1]$.

2. Determine the terminal condition $v^{j+1} = \max\{U_{j+1}, H_{j+1}\}$ for the current valuation step.

3. Set up discretised linear operator M_θ of the resulting ODE system $\frac{d}{dt} v = M_\theta \cdot v$.

4. Incorporate appropriate product-specific boundary conditons.

5. Set up linear system $[I + h_t \theta M_\theta] v^j = [I - h_t (1 - \theta) M_\theta] v^{j+1}$.

6. Solve linear system for v^j by tri-diagonal matrix solver.

7. Repeat with step 3. until next exercise date or $t_j = 0$.

Outline

Monte Carlo methods are widely applied in various finance applications

- We demonstrate the basic principles for
  - path integration of Ito processes
  - exact simulation of Hull-White model paths

- There are many aspects that should also be considered, see e.g.
  - L. Andersen and V. Piterbarg. *Interest rate modelling, volume I to III.*
    Atlantic Financial Press, 2010, Sec. 3.
  - P. Glasserman. *Monte Carlo Methods in Financial Engineering.*
    Springer, 2003

Outline

Monte Carlo (MC) pricing is based on the Strong Law of Large Numbers

Theorem (Strong Law of Large Numbers)

Let $Y_1, Y_2, \ldots$ be a sequence of independent identically distributed (i.i.d.) random variables with finite expectation $\mu < \infty$. Then the sample mean $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ converges to μ a.s. That is

$$\lim_{n\to\infty} \bar{Y}_n = \mu \quad a.s.$$

- We aim at calculating $V(t) = N(t) \cdot \mathbb{E}^N\left[V(T)/N(T) \,|\, \mathcal{F}_t\right]$.
- For MC pricing simulate future discounted payoffs $\left\{\frac{V(T;\omega_i)}{N(T;\omega_i)}\right\}_{i=1,2,\ldots n}$.
- And estimate

$$V(t) = N(t) \cdot \frac{1}{n}\sum_{i=1}^{n} \frac{V(T;\omega_i)}{N(T;\omega_i)}.$$

Keep in mind that sample mean is still a random variable governed by central limit theorem (1/2)

Theorem (Central Limit Theorem)

Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables with finite expectation $\mu < \infty$ and standard deviation $\sigma < \infty$. Denote the sample mean $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Then

$$\frac{\bar{Y}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1).$$

Moreover, for the variance estimator $s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}_n\right)^2$ we also have

$$\frac{\bar{Y}_n - \mu}{s_n/\sqrt{n}} \xrightarrow{d} N(0,1).$$

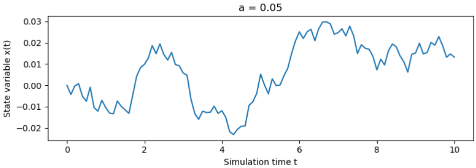
Keep in mind that sample mean is still a random variable governed by central limit theorem (2/2)

$$\frac{\bar{Y}_n - \mu}{s_n/\sqrt{n}} \xrightarrow{d} N(0,1).$$

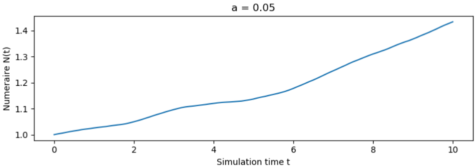
- Here, $N(0,1)$ is the standard normal distribution.
- $\xrightarrow{d}$ denotes convergence in distribution, i.e. $\lim_{n\to\infty} F_n(x) = F(x)$ for the corresponding cumulative distribution functions and all $x \in \mathbb{R}$ at which $F(x)$ is continuous.
- $s_n/\sqrt{n}$ is the standard error of the sample mean $\bar{Y}_n$.

How do we get our samples $V(T;\omega_i)/N(T;\omega_i)$?

1. Simulate state variables $x(t)$ on relevant dates t:



2. Simulate numeraire $N(t)$ on relevant dates t:



3. Calculate payoff $V(T, x(T))$ at observation/pay date T.

We need to simulate our state variables on the relevant observation dates

Consider the general dynamics for a process given as SDE

$$dX(t) = \mu(t, X(t)) \cdot dt + \sigma(t, X(t)) \cdot dW(t).$$

- Typically, we know initial value $X(t)$ ($t = 0$).
- We need $X(T)$ for some future time $T > t$.
- In Hull-White model and risk-neutral measure formulation we have

$$\mu(t, X(t)) = y(t) - a \cdot X(t), \quad \text{and}, \quad \sigma(t, X(t)) = \sigma(t).$$

There are several standard methods to solve above SDE. We will briefly discuss Euler method and Milstein method.

Euler method for SDEs is similar to Explicit Euler method for ODEs

- Specify a grid of simulation times $t = t_0, t_1, \ldots, t_M = T$.

- Calculate sequence of state variables

$$X_{k+1} = X_k + \mu(t_k, X_k)\,(t_{k+1} - t_k) + \sigma(t_k, X_k)\,[W(t_{k+1}) - W(t_k)]\,.$$

- Drift $\mu(t_k, X_k)$ and volatility $\sigma(t_k, X_k)$ are evaluated at current time t_k and state X_k.

- Increment of Brownian motion $W(t_{k+1}) - W(t_k)$ is normally distributed, i.e.

$$W(t_{k+1}) - W(t_k) = Z_k \cdot \sqrt{t_{k+1} - t_k} \quad \text{with} \quad Z_k \sim N(0,1).$$

Milstein method refines the simulation of the diffusion term (1/2)

- Again, specify a grid of simulation times $t = t_0, t_1, \ldots, t_M = T$.
- Calculate sequence of state variables

$$
\begin{aligned}
X_{k+1} = X_k + \mu(t_k, X_k)\,(t_{k+1} - t_k) + \sigma(t_k, X_k)\,[W(t_{k+1}) - W(t_k)] \\
+ \frac{1}{2}\sigma(t_k, X_k)\frac{\partial \sigma(t_k, X_k)}{\partial x}\left[(W(t_{k+1}) - W(t_k))^2 - (t_{k+1} - t_k)\right].
\end{aligned}
$$

- Drift $\mu(t_k, X_k)$ and volatility $\sigma(t_k, X_k)$ are evaluated at current time t_k and state X_k.

Milstein method refines the simulation of the diffusion term (2/2)

- Requires calculation of derivative of volatility $\frac{\partial}{\partial x}\sigma(t_k, X_k)$ w.r.t. state variable.
- Increment of Brownian motion $W(t_{k+1}) - W(t_k)$ is normally distributed, i.e.

$$W(t_{k+1}) - W(t_k) = Z_k \cdot \sqrt{t_{k+1} - t_k} \quad \text{with} \quad Z_k \sim N(0,1).$$

- With $\Delta_k = t_{k+1} - t_k$ iteration becomes

$$\begin{aligned} X_{k+1} = X_k &+ \mu(t_k, X_k)\Delta_k + \sigma(t_k, X_k) Z_k \sqrt{\Delta_k} \\ &+ \frac{1}{2}\sigma(t_k, X_k)\frac{\partial \sigma(t_k, X_k)}{\partial x}\left(Z_k^2 - 1\right)\Delta_k. \end{aligned}$$

How can we measure convergence of the methods?

- We distinguish strong order of convergence and weak order of convergence.
- Consider a discrete SDE solution $\left\{X_k^h\right\}_{k=0}^M$ with $X_k^h \approx X(t+kh)$, $h = \frac{T-t}{M}$.

Definition (Strong order of convergence)

The discrete solution X_M^h at final maturity $T = t + hM$ converges to the exact solution $X(T)$ with strong orderβ if there exists a constant C such that

$$\mathbb{E}\left[\left|X_M^h - X(T)\right|\right] \leq C \cdot h^\beta.$$

- Strong order of convergence focuses on convergence on the individual paths.
- Euler method has strong order of convergence of $\frac{1}{2}$ (given sufficient conditions on $\mu(\cdot)$ and $\sigma(\cdot)$).
- Milstein method has strong order of convergence of 1 (given sufficient conditions on $\mu(\cdot)$ and $\sigma(\cdot)$).

For derivative pricing we are typically interested in weak order of convergence

We need some context for weak order of convergence

- A function $f : \mathbb{R} \to \mathbb{R}$ is polynomially bounded if $|f(x)| \leq k(1+|x|)^q$ for constants k and q and all x.
- The set $\mathcal{C}_{\mathcal{P}}^n$ represents all functions that are n-times continuously differentiable and with 1st to nth derivative polynomially bounded.

Definition (Weak order of convergence)

The discrete solution X_M^h at final maturity $T = t + hM$ converges to the exact solution $X(T)$ with weak orderβ if there exists a constant C such that

$$\left|\mathbb{E}\left[f\left(X_M^h\right)\right] - \mathbb{E}\left[f\left(X(T)\right)\right]\right| \leq C \cdot h^{\beta} \quad \forall f \in \mathcal{C}_{\mathcal{P}}^{2\beta+2}$$

for sufficiently small h.

- Think of f as a payoff function, then weak order of convergence is related to convergence in price.
- Euler method and Milstein method can be shown to have weak order 1 convergence (given sufficient conditions on μ and σ).

Some comments regarding weak order of convergence

Error estimate

$$\left|\mathbb{E}\left[f\left(X_M^h\right)\right] - \mathbb{E}\left[f\left(X(T)\right)\right]\right| \leq C \cdot h^{\beta}$$

requires considerable assumptions regarding smoothness of $\mu(\cdot)$, $\sigma(\cdot)$ and test functions $f(\cdot)$.

- In practice payoffs are typically non-smooth at the strike.
- This limits applicability of more advanced schemes with theoretical higher order of convergence.
- A fairly simple approach of a higher order scheme is based on Richardson extrapolation:
  - this method is also applied to ODEs,
  - see Glassermann (2000), Sec. 6.2.4 for details.
- Typically, numerical testing is required to assess convergence in practice.

The choice of pricing measure is crucial for numeraire simulation

Consider risk-neutral measure, then

$$N(T) = B(T) = \exp\left\{\int_0^T r(s)ds\right\} = \exp\left\{\int_0^T \left[f(0,s) + x(s)\right] ds\right\}$$
$$= P(0,T)^{-1} \exp\left\{\int_0^T x(s)ds\right\}.$$

Requires simulation or approximation of $\int_0^T x(s)ds$.

Suppose $x(t_k)$ is simulated on a time grid $\{t_k\}_{k=0}^M$ then we approximate integral via Trapezoidal rule

$$\int_0^T x(s)ds \approx \sum_{i=1}^{M} \frac{x(t_{k-1}) + x(t_k)}{2} (t_k - t_{k-1}).$$

Numeraire simulation is done in parallel to state simulation

$$N(t_k) = \frac{P(0,t_{k-1})}{P(0,t_k)} \cdot N(t_{k-1}) \cdot \exp\left\{\frac{x(t_{k-1}) + x(t_k)}{2} (t_k - t_{k-1})\right\}.$$

Alternatively, we can simulate in T-forward measure for a fixed future time T

Select a future time $\bar{T}$ sufficiently large. Then $N(0) = P(0, \bar{T})$.
At any pay time $T \leq \bar{T}$ numeraire is directly available via zero coupon bond formula

$$N(T) = P(x(T), T, \bar{T}) = \frac{P(0, \bar{T})}{P(0, T)} e^{-G(T,T')x(T) - \frac{1}{2}G(T,T')^2 y(T)}.$$

However, $\bar{T}$-forward measure simulation needs consistent model formulation or change of measure.
In particular

$$\underbrace{dW^{\bar{T}}(t)}_{\text{B.M. in } \bar{T}\text{-forward measure}} = \underbrace{\sigma_P(t, \bar{T})}_{\text{ZCB volatility}} \cdot dt + \underbrace{dW(t)}_{\text{B.M. in risk-neutral measure}}.$$

Another commonly used numeraire for simulation is the discretely compounded bank account

- Consider a grid of simulation times $t = t_0, t_1, \ldots, t_M = T$.
- Assume we start with 1 EUR at $t = 0$, i.e. $N(0) = 1$.
- At each t_k we take numeraire $N(t_k)$ and buy zero coupon bond maturing at t_{k+1}. That is

$$N(t) = P(t, t_{k+1}) \cdot \frac{N(t_k)}{P(t_k, t_{k+1})} \quad \text{for} \quad t \in [t_k, t_{k+1}] .$$

Explicitly, define discretely compounded bank account as $\bar{B}(0) = 1$ and

$$\bar{B}(t) = P(t, t_{k+1}) \prod_{t_k < t} \frac{1}{P(t_k, t_{k+1})} .$$

We get

$$d\left(\frac{\bar{B}(t)}{P(t, t_{k+1})}\right) = \prod_{t_k < t} \frac{1}{P(t_k, t_{k+1})} \cdot d\left(\frac{P(t, t_{k+1})}{P(t, t_{k+1})}\right) = 0 \quad \text{for} \quad t \in [t_k, t_{k+1}] .$$

Simulating in $\bar{B}$-measure is equivalent to simulating in rolling t_{k+1}-forward measure.

Outline

Do we really need to solve the Hull-White SDE numerically?

Recall dynamics in T-forward measure

$$dx(t) = \left[y(t) - \sigma(t)^2 G(t,T) - a \cdot x(t)\right] \cdot dt + \sigma(t) \cdot dW^T(t).$$

That gives

$$x(T) = e^{-a(T-t)} \cdot \left[x(t) + \int_t^T e^{a(u-t)} \left(\left[y(u) - \sigma(u)^2 G(u,T) \right] du + \sigma(u) dW^T(u) \right) \right].$$

As a result $x(T) \sim N(\mu, \sigma^2)$ (conditional on t) with

$$\mu = \mathbb{E}^T \left[x(T) \,|\, \mathcal{F}_t \right] = G'(t,T) \left[x(t) + G(t,T) y(t) \right] \quad \text{and}$$

$$\sigma^2 = \mathrm{Var}\left[x(T) \,|\, \mathcal{F}_t \right] = y(T) - G'(t,T)^2 y(t).$$

We can simulate exactly

$$x(T) = \mu + \sigma \cdot Z \quad \text{with} \quad Z \sim N(0,1).$$

Expectation calculation via $\mu = \mathbb{E}^T \left[x(T) \,|\, \mathcal{F}_t \right]$ requires carefull choice of numeraire

Consider grid of simulation times $t = t_0, t_1, \ldots, t_M = T$.
We simulate

$$x(t_{k+1}) = \mu_k + \sigma_k \cdot Z_k$$

with

$$\begin{aligned}
\mu_k &= G'(t_k, t_{k+1}) \left[x(t_k) + G(t_k, t_{k+1}) y(t_k) \right], \\
\sigma_k^2 &= y(t_{k+1}) - G'(t_k, t_{k+1})^2 y(t_k), \quad \text{and} \\
Z_k &\sim N(0, 1).
\end{aligned}$$

Grid point t_{k+1} must coincide with forward measure for $\mathbb{E}^{t_{k+1}}[\cdot]$ for each individual step $k \to k+1$.
Numeraire must be discretely compounded bank account $\bar{B}(t)$ and

$$\bar{B}(t_{k+1}) = \frac{\bar{B}(t_k)}{P(x(t_k), t_k, t_{k+1})}.$$

Recursion for $x(t_{k+1})$ and $\bar{B}(t_{k+1})$ fully specifies path simulation for pricing.

Some comments regarding Hull-White MC simulation ...

- We could also simulate in risk-neutral measure or $\bar{T}$-forward measure.
  - This might be advantegous if also FX or equities are modelled/simulated.
  - Requires adjustment of conditional expectation μ_k and numeraire $N(t_k)$ calculation.
  - Variance σ_k^2 is invariant to change of meassure in Hull-White model.

- Repeat path generation for as many paths $1, \ldots, n$ as desired (or computationally feasible).

- For Bermudan pricing we need to simulate x and N (at least) at exercise dates $T_E^1, \ldots, T_E^{\bar{k}}$.

- For calculation of Z_k use
  - pseudo-random numbers or
  - Quasi-Monte Carlo sequences.

  as proxies forindependent $N(0,1)$ random variables accross time steps and paths.

We illustrate MC pricing by means of a coupon bond option example

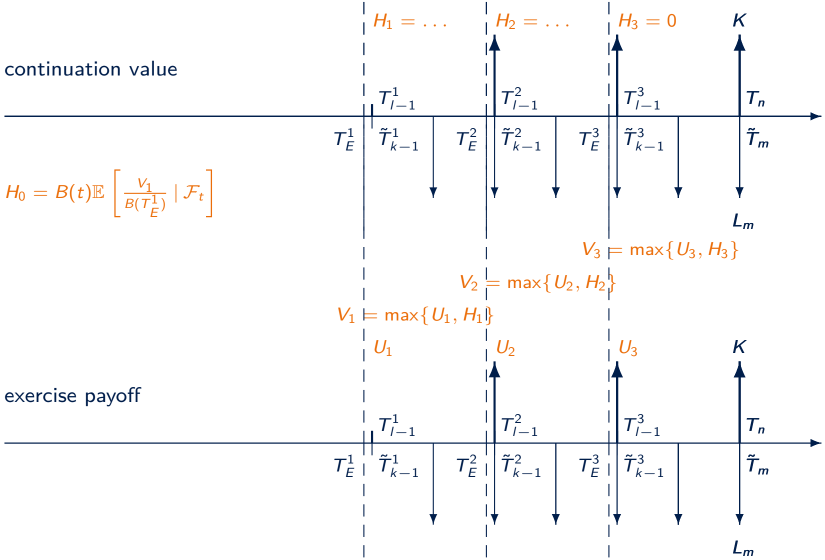
Consider coupon bond option expiring at T_E with coupons C_i paid at T_i ($i = 1, \ldots, u$, incl. strike and notional).

- Set $t_0 = 0$, $t_1 = T_E/2$ and $t_2 = T_E$ (two steps for illustrative purpose).
- Compute $2n$ independent $N(0,1)$ pseudo random numbers $Z^1, \ldots, Z^{2n}$.
- For all paths $j = 1, \ldots, n$ calculate:
  - μ_0^j, σ_0 and $\bar{B}^j(t_1)$; note μ_0^j and $\bar{B}^j(t_1)$ are equal for all paths j since $x(t_0) = 0$,
  - $x_1^j = \mu_0^j + \sigma_0 \cdot Z^j$,
  - μ_1^j, σ_1 and $\bar{B}^j(t_2)$; note now μ_1^j and $\bar{B}^j(t_2)$ depend on x_1^j,
  - $x_2^j = \mu_1^j + \sigma_1 \cdot Z^{n+j}$,
  - payoff $V^j(t_2) = \left[\sum_{i=1}^u C_i \cdot P(x_2^j, t_2, T_i)\right]^+$ at $t_2 = T_E$.
- Calculate option price (note $\bar{B}(0) = 1$)

$$V(0) = \bar{B}(0) \cdot \frac{1}{n} \sum_{j=1}^{n} \frac{V^j(t_2)}{\bar{B}^j(t_2)}.$$

Outline

Let's return to our Bermudan option pricing problem

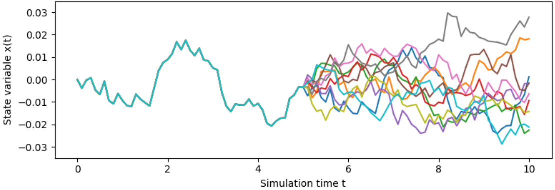


In this setting we need to calculate future conditional expectations

- Assume we already simulated paths for state variables x_k, underlyings U_k and numeraire B_k for all relevant dates t_k.
- We need continuation values H_k defined recursively via $H_{\bar{k}} = 0$ and

$$H_k = B_k \mathbb{E}_k \left[\frac{\max\left\{U_{k+1}, H_{k+1}\right\}}{B_{k+1}} \right].$$

- In principle, we could use nested Monte Carlo:



- In practice, nested Monte Carlo is typically computationally not feasible.

A key idea of American Monte Carlo is approximating conditional expectation via regression

Conditional expectation

$$H_k = \mathbb{E}_k \left[\frac{B_k}{B_{k+1}} \max \left\{ U_{k+1}, H_{k+1} \right\} \right]$$

is a function of the path $x(t)$ for $t \le t_k$.

For non-path-dependent underlyings U_k, H_k can be written as function of $x_k = x(t_k)$, i.e.

$$H_k = H_k(x_k).$$

We aim at finding a regression operator

$$\mathcal{R}_k = \mathcal{R}_k \left[Y \right]$$

which we can use as proxy for H_k.

What do we mean by regression operator?

Denote $\zeta(\omega) = [\zeta_1(\omega), \dots, \zeta_q(\omega)]^\top$ a set of basis functions (vector of random variables).

Let $Y = Y(\omega)$ be a target random variable.

Assume we have outcomes $\omega_1, \dots, \omega_{\bar{n}}$ with control variables $\zeta(\omega_1), \dots, \zeta(\omega_{\bar{n}})$ and observations $Y(\omega_1), \dots, Y(\omega_{\bar{n}})$.

A regression operator $\mathcal{R}[Y]$ is defined via

$$\mathcal{R}[Y](\omega) = \zeta(\omega)^\top \beta$$

where the regression coefficients β solve linear least squares problem

$$\left\| \begin{bmatrix} \zeta(\omega_1)^\top \beta - Y(\omega_1) \\ \vdots \\ \zeta(\omega_{\bar{n}})^\top \beta - Y(\omega_{\bar{n}}) \end{bmatrix} \right\|^2 \to \min .$$

Linear least squares system can be solved e.g. via QR factorisation or SVD.

A basic pricing scheme is obtained by replacing conditional expectation of future payoff by regression operator

Approximate $\tilde{H}_k \approx H_k$ via $\tilde{H}_{\bar{k}} = H_{\bar{k}} = 0$ and

$$\tilde{H}_k = \mathcal{R}_k \left[\frac{B_k}{B_{k+1}} \max \left\{ U_{k+1}, \tilde{H}_{k+1} \right\} \right] \quad \text{for} \quad k = \bar{k} - 1, \ldots, 1.$$

- Critical piece of this methodology is (for each step k)
  - choice of regression variables $\zeta_1, \ldots, \zeta_q$ and
  - calibration of regression operator $\mathcal{R}_k$ with coefficients β.
- Regression variables $\zeta_1, \ldots, \zeta_q$ must be calculated based on information up to t_k.
  - They must not look into the future to avoid upward bias.
- Control variables $\zeta(\omega_1), \ldots, \zeta(\omega_{\bar{n}})$ and observations $Y(\omega_1), \ldots, Y(\omega_{\bar{n}})$ for calibration should be simulated on paths independent from pricing.
  - Using same paths for calibration and payoff simulation also incorporates information on the future.

What are typical basis functions?

State variable approach

Set $\zeta_i = x(t_k)^{i-1}$ for $i = 1, \ldots, q$. Typical choice is $q \approx 4$ (i.e. polynomials of order 3). For multi-dimensional models we would set $\zeta_i = \prod_{j=1}^{d} x_j(t_k)^{p_{i,j}}$ with $\sum_{j=1}^{d} p_{i,j} \leq r$.

- Very generic and easy to incorporate.

Explanatory variable approach

Identify variables $y_1, \ldots y_{\bar{d}}$ relevant for the underlying option. Set basis functions as monomials

$$\zeta_i = \prod_{j=1}^{\bar{d}} y_j(t_k)^{p_{i,j}} \quad \text{with} \quad \sum_{j=1}^{\bar{d}} p_{i,j} \leq r.$$

- Can be chosen option-specific and incorporate information prior to t_k.
- Typical choices are co-terminal swap rates or Libor rates (observed at t_k).

Regression of the full underlying can be a bit rough - we may restrict regression to exercise decision only

For a given path consider

$$
\begin{aligned}
H_k &= \frac{B_k}{B_{k+1}} \max\left\{U_{k+1}, H_{k+1}\right\} \\
&= \frac{B_k}{B_{k+1}} \left[\mathbb{1}_{\{U_{k+1} > H_{k+1}\}} U_{k+1} + \left(1 - \mathbb{1}_{\{U_{k+1} > H_{k+1}\}}\right) H_{k+1}\right].
\end{aligned}
$$

Use regression to calculate $\mathbb{1}_{\{U_{k+1} > H_{k+1}\}}$.
Calculate $\mathcal{R}_{k+1} = \mathcal{R}_{k+1}\left[U_{k+1} - H_{k+1}\right]$, set $H_{\bar{k}} = 0$ and

$$
H_k = \frac{B_k}{B_{k+1}} \left[\mathbb{1}_{\{\mathcal{R}_{k+1} > 0\}} U_{k+1} + \left(1 - \mathbb{1}_{\{\mathcal{R}_{k+1} > 0\}}\right) H_{k+1}\right] \quad \text{for} \quad k = \bar{k}-1, \ldots, 1.
$$

- Think of $\mathbb{1}_{\{\mathcal{R}_{k+1} > 0\}}$ as an exercise strategy (which might be sub-optimal).
- This approach is sometimes considered more accurate than regression on regression.
- For further reference, see also Longstaff/Schwartz (2001).

We summarise the American Monte Carlo method

1. Simulate n paths of state variables x_k^j, underlyings U_k^j and numeraires B_k^j $(j = 1, \ldots, n)$ for all relevant times t_k $(k = 1, \ldots \bar{k})$.
2. Set $H_{\bar{k}}^j = 0$.
3. For $k = \bar{k} - 1, \ldots 1$ iterate:
   3.1 Calculate control variables $\left\{\zeta_i^j = \zeta_i(\omega_j)\right\}_{i=1,\ldots,q}^{j=1,\ldots,\hat{n}}$ and regression variables $Y^j = U_k^j - H_k^j$ for the first $\hat{n}$ paths $(\hat{n} \approx \frac{1}{4} n)$.
   3.2 Calibrate regression operator $\mathcal{R}_{k+1} = \mathcal{R}_{k+1}[Y]$ which gives coefficients β.
   3.3 Calculate control variables $\left\{\zeta_i^j = \zeta_i(\omega_j)\right\}_{i=1,\ldots,q}^{j=\hat{n}+1,\ldots n}$ for remaining paths and (for all paths)
   $$H_k^j = \frac{B_k^j}{B_{k+1}^j} \left[\mathbb{1}_{\left\{\mathcal{R}_{k+1}(\omega_j)>0\right\}} U_{k+1}^j + \left(1 - \mathbb{1}_{\left\{\mathcal{R}_{k+1}(\omega_j)>0\right\}}\right) H_{k+1}^j\right].$$
4. Calculate discounted payoffs for the paths $j = \hat{n} + 1, \ldots n$ not used for regression
$$H_0^j = \frac{B_k^j}{B_{k+1}^j} \max\left\{U_1^j, H_1^j\right\}.$$
5. Derive average $V(0) = \frac{1}{n-\hat{n}} \sum_{j=\hat{n}+1}^{n} H_0^j$.

Some comments regarding AMC for Bermudans in Hull-White model

- AMC implementations can be very bespoke and problem specific.
  - See literature for more details.
- More explanatory variables or too high polynomial degree for regression may deteriorate numerical solution.
  - This is particularly relevant for 1-factor models like Hull-White.
  - Single state variable or co-terminal swap rate should suffice.
- AMC with Hull-White for Bermudans is *not* the method of choice.
  - PDE and integration methods are directly applicable.
  - AMC is much slower and less accurate compared to PDE and integration.

AMC is the method of choice for high-dimensional models and/or path-dependent products.

Part VI

Model Calibration

Outline

Yield Curve Calibration

Calibration Methodologies for Hull-White Model

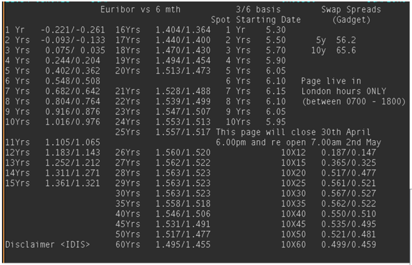
Outline

Yield Curve Calibration

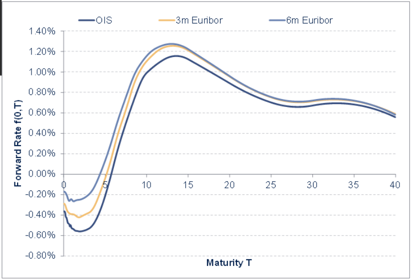
Calibration Methodologies for Hull-White Model

Outline

What is the goal of yield curve calibration?



Euribor vs 6 mth				3/6 basis Spot Starting Date		Swap Spreads (Gadget)	
1 Yr	-0.221/-0.261	16Yrs	1.404/1.364	1 Yr	5.30		
2 Yrs	-0.093/-0.133	17Yrs	1.440/1.400	2 Yrs	5.50	5y	56.2
3 Yrs	0.075/ 0.035	18Yrs	1.470/1.430	3 Yrs	5.70	10y	65.6
4 Yrs	0.244/0.204	19Yrs	1.494/1.454	4 Yrs	5.90		
5 Yrs	0.402/0.362	20Yrs	1.513/1.473	5 Yrs	6.05		
6 Yrs	0.548/0.508			6 Yrs	6.10	Page live in	
7 Yrs	0.682/0.642	21Yrs	1.528/1.488	7 Yrs	6.15	London hours ONLY	
8 Yrs	0.804/0.764	22Yrs	1.539/1.499	8 Yrs	6.10	(between 0700 - 1800)	
9 Yrs	0.916/0.876	23Yrs	1.547/1.507	9 Yrs	6.05		
10Yrs	1.016/0.976	24Yrs	1.553/1.513	10Yrs	5.95		
		25Yrs	1.557/1.517	This page will close 30th April			
11Yrs	1.105/1.065			6.00pm and re open 7.00am 2nd May			
12Yrs	1.183/1.143	26Yrs	1.560/1.520			10X12	0.187/0.147
13Yrs	1.252/1.212	27Yrs	1.562/1.522			10X15	0.365/0.325
14Yrs	1.311/1.271	28Yrs	1.563/1.523			10X20	0.517/0.477
15Yrs	1.361/1.321	29Yrs	1.563/1.523			10X25	0.561/0.521
		30Yrs	1.563/1.523			10X30	0.567/0.527
		35Yrs	1.558/1.518			10X35	0.562/0.522
		40Yrs	1.546/1.506			10X40	0.550/0.510
		45Yrs	1.531/1.491			10X45	0.535/0.495
		50Yrs	1.517/1.477			10X50	0.521/0.481
Disclaimer <IDIS>		60Yrs	1.495/1.455			10X60	0.499/0.459



We aim at finding a set of yield curves that allows re-pricing a set of market instruments.

We start with a single-curve setting example to illustrate the general principle (1/2)

Consider Vanilla swaps as market instruments with the pricing formula (single-curve setting, $t \leq T_0$)

$$\mathrm{Swap}^k(t) = \underbrace{[P(t, T_0) - P(t, T_{n_k})]}_{\text{float leg}} - \underbrace{\sum_{i=1}^{n_k} R\tau_i P(t, T_i)}_{\text{fixed Leg}}.$$

A market swap quote R_k for a T_{n_k}-maturing (and spot-starting) Vanilla swap is the fixed rate that prices the swap at par, i.e.

$$\underbrace{0}_{\mathrm{Market}(R_k)} = \underbrace{\mathrm{Swap}^k(0) = [P(0, T_0) - P(0, T_{n_k})] - \sum_{i=1}^{n_k} R_k \tau_i P(0, T_i)}_{\mathrm{Model}[P](R_k)}.$$

We start with a single-curve setting example to illustrate the general principle (2/2)

$$\underbrace{0}_{\text{Market}(R_k)} = \underbrace{\text{Swap}^k(0) = [P(0, T_0) - P(0, T_{n_k})] - \sum_{i=1}^{n_k} R_k \tau_i P(0, T_i)}_{\text{Model}[P](R_k)}.$$

We associate a calibration helper operator $\mathcal{H}_k = \mathcal{H}_k[P]$ with each market instrument which takes as input a yield curve $P(0, T)$ and calculates (for a market quote)

$$\mathcal{H}_k[P](R_k) = \text{Model}[P](R_k) - \text{Market}(R_k).$$

Yield curve calibration is formulated as minimisation problem

(Single-Curve) Yield Curve Calibration Problem

For a given set of market quotes $\{R_k\}_{k=1,\ldots q}$ with corresponding instruments and calibration helpers $\mathcal{H}_k[P]$, the yield curve calibration problem is given by

$$\min_P \left\| \left[\mathcal{H}_1[P](R_1), \ldots, \mathcal{H}_q[P](R_q)\right]^\top \right\|.$$

- Effectively, we only need a finite set of $P(0, T_i)$.
- Without further constraints there are multiple yield curves $P(0, T)$ that give optimal solution

$$\left[\mathcal{H}_1[P](R_1), \ldots, \mathcal{H}_q[P](R_q)\right]^\top = 0 \in \mathbb{R}^q.$$

- We need to add sensible regularisation to
  - make calibration problem tractable (finite dimensional domain),
  - ensure unique, accurate and sensible solution,
  - allow for efficient computation.

Regularisation is achieved by discretisation and interpolation of the yield curve

Order market quotes R_k and calibration helpers $\mathcal{H}_k[P]$ by increasing final maturity T_{n_k} $(k = 1, \ldots, q)$ of underlying instruments.
Set

$$R = [R_1, \ldots, R_q] \quad \text{and} \quad \mathcal{H}[P] = [\mathcal{H}_1[P], \ldots, \mathcal{H}_q[P]].$$

Define a vector of yield curve parameters $z = [z_1, \ldots, z_q]^\top \in \mathbb{R}^q$ which specify the yield curve via

$$P = P[z].$$

- Typically, z_k are zero, forward rates or discount factors for maturities T_{n_k}.
- Compare with interpolation traits in QuantLib.

Specify $P[z](0, T)$ via interpolation/extrapolation based on curve parameters z.

- E.g. monoton cubic spline interpolation.

We re-formulate the calibration problem in terms of model parameters

Finite Dimensional Yield Curve Calibration Problem

The yield curve calibration problem in terms of yield curve model parameters is given by

$$\min_z \left\| \mathcal{H}\left[P\left[z\right]\right]\left(R\right) \right\|$$

where

$$z = \left[z_1, \ldots, z_q\right]^\top, \; R = \left[R_1, \ldots, R_q\right]^\top \text{ and } \mathcal{H}\left[P\right] = \left[\mathcal{H}_1\left[P\right], \ldots, \mathcal{H}_q\left[P\right]\right].$$

- In general, parametrised calibration problem can be solved by general purpose optimisation methods.
- This can be computationally expensive if number of inputs and parameters q is large.
- We can also exploit the structure of the problem to reduce computational complexity.

The multi-dimensional calibration problem can be reduced to a sequence of one-dimensional calibration problems (1/3)

Lemma

Consider our parametrised calibration problem setting. Assume a yield curve parametrisation $P[z]$ such that discount factors $P[z](0,T)$ are continuously differentiable w.r.t. z for all maturities T, and parametrised locally in the sense that

$$\frac{\partial}{\partial z_k} P[z](0,T) = 0 \quad \textit{for} \quad T \le T_{n_{k-1}}.$$

Then the Jacobi matrix $\frac{d}{dz}\mathcal{H}[P[z]](R)$ is of lower triangular form.

The multi-dimensional calibration problem can be reduced to a sequence of one-dimensional calibration problems (2/3)

Proof:
Consider a component of the Jacobi matrix

$$
\begin{aligned}
\frac{d}{dz_l}\mathcal{H}_k\left[P\left[z\right]\right](R) &= \frac{d}{dz_l}\text{Model}[P\left[z\right]](R_k) \\
&= \frac{d}{dz_l}\left[P(0,T_0) - P(0,T_{n_k}) - \sum_{i=1}^{n_k} R_k \cdot \tau_i \cdot P(0,T_i)\right] \\
&= \frac{d}{dz_l}P(0,T_0) - \frac{d}{dz_l}P(0,T_{n_k}) - \sum_{i=1}^{n_k} R_k \cdot \tau_i \cdot \frac{d}{dz_l}P(0,T_i).
\end{aligned}
$$

The largest maturity is T_{n_k}. Thus, due to local parametrisaion property, for $l > k$, $\frac{d}{dz_l}P(0,T_{n_k}) = 0$. Same holds for maturities $T_i \leq T_{n_k}$.

The multi-dimensional calibration problem can be reduced to a sequence of one-dimensional calibration problems (3/3)

Consequently,

$$\frac{d}{dz_l}\mathcal{H}_k\left[P\left[z\right]\right]\left(R_k\right) = 0 \quad \text{for} \quad l > k$$

and

$$\frac{d}{dz}\mathcal{H}\left[P\left[z\right]\right]\left(R\right) = \begin{bmatrix} \star & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ \star & \dots & \dots & \star \end{bmatrix}.$$

This concludes the proof.

Sequential yield curve calibration is also called yield curve bootstrapping

- If there is an exact solution z such that $\mathcal{H}\left[P\left[z\right]\right](R) = 0$ then we can find it by solving sequence of one-dimensional equations

$$h_k(z_k) = \mathcal{H}_k\left[P\left[z_1, \ldots z_{k-1}, z_k, z_k, \ldots\right]\right](R_k) = 0 \quad \text{for} \quad k = 1, 2, \ldots, q.$$

- If there is no exact solution, we can still exploit lower triangular form of Jacobi matrix in efficiently solving

$$\min_z \left\| \mathcal{H}\left[P\left[z\right]\right](R) \right\|.$$

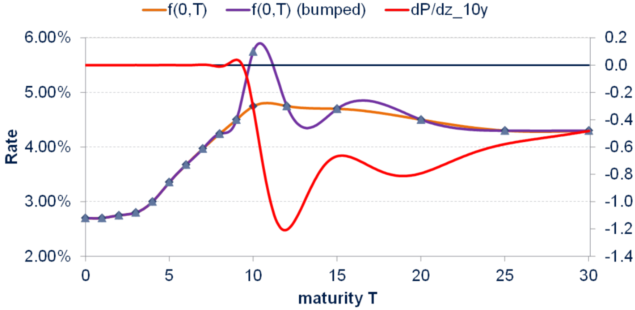
- Local parametrisation is achieved e.g. by spline interpolation methods that are fully specified by two neighboring points (e.g. linear interpolation).

- Note that local parametrisations typically yield less smooth forward rate curves than parametrisations where a change in a single parameter impacts a broader range of discount factors.

Do we really need the restriction to local parametrisation?

- In many curve parametrisations/interpolations sensitivity $\frac{\partial}{\partial z_k} P[z](0,T)$ is small for $T \le T_{n_{k-1}}$.

Example: Interpolated forward rates $f(0,T)$ with cubic C^2-splines bumped by 1% at $10y$:



- 10y rate bump does affect curve before $9y$ time point.
- However, impact is small compared to impact around $10y$ maturity.

We can extend the bootstrapping method to non-local parametrisations

Iterative Bootstrapping Method

Suppose we have a calibration problem set up via

$$\mathcal{H}\left[P\left[z\right]\right] = \left[\mathcal{H}_1\left[P\left[z\right]\right], \ldots, \mathcal{H}_q\left[P\left[z\right]\right]\right].$$

The iterative bootstrapping solves the calibration problem $\mathcal{H}\left[P\left[z\right]\right] = 0$ via the following steps:

1. Set initial solution $z^0 = \left[z_1^0, \ldots z_q^0\right]$ via standard bootstrapping.
2. If $\mathcal{H}\left[P\left[z^0\right]\right] \neq 0$ repeat the fixpoint iteration:
   2.1 For $k = 1, 2, \ldots, q$ find z_k^i such that

   $$h_k(z_k^i) = \mathcal{H}_k\left[P\left[z_1^i, \ldots z_{k-1}^i, z_k^i, z_{k+1}^{i-1}, \ldots, z_q^{i-1}\right]\right](R_k) = 0.$$

   2.2 Stop iteration if $\left\|z^i - z^{i-1}\right\| < \varepsilon$.

- Iterative bootstrapping method usually converges in a few iterations.

Outline

Single-curve calibration procedure is typically applied to discount curves from OIS swaps

Recall

$$\text{CompSwap}(t) = \underbrace{\sum_{j=1}^{m} L(t; T_{j-1}, T_j)\tau_j P(t, T_j)}_{\text{compounding leg}} - \underbrace{\sum_{j=1}^{m} R\tau_j P(t, T_j)}_{\text{fixed leg}},$$

$$L(t, T_{j-1} T_j) = \left[\frac{P(t, T_{j-1})}{P(t, T_j)} - 1\right] \frac{1}{\tau_j} \quad \text{(compounded OIS rate)}.$$

Compounding swap rate helper can be defined solely in terms of discount curve P via

$$\mathcal{H}^{\text{CS}}[P](R) = \text{CompSwap}(0) - 0.$$

Single curve calibration procedure can be applied straight away.

OIS discount curves can be derived from OIS swaps via single-curve calibration procedure.

Forward rate agreements (FRA) can be used to specify short end of projection curves (1/2)

Market quote of FRA with start date T_0 and tenor δ is the fixed rate R that prices the FRA at par as of today. Consider present value

$$\text{FRA}(t) = \underbrace{P(t, T_0)}_{\text{discounting to } T_0} \underbrace{\left[L^\delta(t; T_0, T_0 + \delta) - R\right] \tau}_{\text{payoff}} \underbrace{\frac{1}{1 + \tau L^\delta(t; T_0, T_0 + \delta)}}_{\text{discounting from } T_0 \text{ to } T_0 + \delta}.$$

Condition $\text{FRA}(t) = 0$ yields FRA calibration helper

$$\begin{aligned} \mathcal{H}^{\text{FRA}}\left[P^\delta\right](R) &= L^\delta(0; T_0, T_0 + \delta) - R \\ &= \left[\frac{P^\delta(0, T_0)}{P^\delta(0, T_0 + \delta)} - 1\right] \frac{1}{\tau} - R. \end{aligned}$$

Forward rate agreements (FRA) can be used to specify short end of projection curves (2/2)

$$\mathcal{H}^{\text{FRA}}\left[P^{\delta}\right](R) = \left[\frac{P^{\delta}(0, T_0)}{P^{\delta}(0, T_0 + \delta)} - 1\right]\frac{1}{\tau} - R.$$

- Typical tenors δ are 1m, 3m, 6m and 12m (corresponding to Libor rate indices).
- Typical expiries T_0 are up to 2y.
- Both, available tenors and expiries, depend on the market (or currency).
- Note that FRA rate helper only depends on projection curve $P^{\delta}(0, T_0)$.

Vanilla swaps are used to specify projection curves for longer maturities

Multi-curve swap price is given by

$$\text{Swap}(t) = \underbrace{\sum_{j=1}^{m} L^{\delta}(t, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)\tilde{\tau}_j P(t, \tilde{T}_j)}_{\text{float leg}} - \underbrace{\sum_{i=1}^{n} R\tau_i P(t, T_i)}_{\text{fixed Leg}}.$$

Vanilla swap rate helper becomes

$$\mathcal{H}^{\text{VS}}\left[P^{\delta}, (P)\right](R) = \sum_{j=1}^{m} L^{\delta}(0, \tilde{T}_{j-1}, \tilde{T}_{j-1}+\delta)\tilde{\tau}_j P(t, \tilde{T}_j) - \sum_{i=1}^{n} R\tau_i P(0, T_i).$$

- Rate helper depends on forward curve P^{δ} via forward Libor rates $L^{\delta}(0, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)$.
- Rate helper also depends on discount curve P via discount factors $P(t, \tilde{T}_j)$ and $P(0, T_i)$.
  - This is reflected by notation $\mathcal{H}^{\text{VS}}[\cdot, (P)]$.
  - We put dependence in parentheses (P) because usually discount curve P is calibrated earlier already from OIS swaps.

Projection curve calibration is analogous to single curve calibration (1/2)

- Specify projection curve parameters z^{δ} and projection curve $P^{\delta} = P^{\delta}\left[z^{\delta}\right]$.
  - Use methodologies/interpolations analogous to discount curves.
- Set up calibration problem in terms of z^{δ} via

$$\mathcal{H}^{\delta}\left[P^{\delta}\left[z^{\delta}\right]\right] = \begin{bmatrix} \mathcal{H}_1^{\mathrm{FRA}}\left[P^{\delta}\left[z^{\delta}\right]\right] \\ \vdots \\ \mathcal{H}_{q_{\mathrm{FRA}}}^{\mathrm{FRA}}\left[P^{\delta}\left[z^{\delta}\right]\right] \\ \mathcal{H}_1^{\mathrm{VS}}\left[P^{\delta}\left[z^{\delta}\right],(P)\right] \\ \vdots \\ \mathcal{H}_{q_{\mathrm{VS}}}^{\mathrm{VS}}\left[P^{\delta}\left[z^{\delta}\right],(P)\right] \end{bmatrix}$$

  where calibration helpers are ordered by last cash flow date.
- Obtain a set of market quotes

$$R^{\delta} = \left[R_1^{\mathrm{FRA}}, \ldots, R_{q_{\mathrm{FRA}}}^{\mathrm{FRA}}, R_1^{\mathrm{VS}}, \ldots, R_{q_{\mathrm{VS}}}^{\mathrm{VS}}\right]^{\top}.$$

Projection curve calibration is analogous to single curve calibration (2/2)

$$R^\delta = \left[R_1^{\mathrm{FRA}}, \ldots, R_{q_{\mathrm{FRA}}}^{\mathrm{FRA}}, R_1^{\mathrm{VS}}, \ldots, R_{q_{\mathrm{VS}}}^{\mathrm{VS}}\right]^\top$$

- Solve

  $$\min_{z^\delta} \left\| \mathcal{H}^\delta \left[P^\delta \left[z^\delta \right], (P) \right] \left(R^\delta \right) \right\|$$

  depending on curve parametrisation via iterative bootstrapping or multi-dimensional optimisation method.

- In principle, discount curve P and projection curve P^δ could also be solved simultanously by an augmented optimisation problem

  $$\min_{z, z^\delta} \left\| \hat{\mathcal{H}} \left[P[z], P^\delta \left[z^\delta \right] \right] \left(R, R^\delta \right) \right\|.$$

  - However, keep in mind increased computational effort and complexity.

Basis swaps are further instruments which are liquidely traded and also used for curve calibration (1/2)

Tenor Basis Swap

Floating rate payments of a longer Libor tenor are exchanged against floating rate payments of a shorter Libor tenor plus fixed spread,

$$
\begin{aligned}
\text{TenorSwap}(t) = & \sum_{j=1}^{m_1} L^{\delta_1}(t, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta_1)\tilde{\tau}_j P(t, \tilde{T}_j) \\
& - \sum_{j=1}^{m_2} \left[L^{\delta_2}(t, \hat{T}_{j-1}, \hat{T}_{j-1} + \delta) + s \right] \hat{\tau}_j P(t, \hat{T}_j).
\end{aligned}
$$

- For example, $\delta_1 = 6m$ and $\delta_2 = 3m$.
- Market quote is spread s (corresponding to maturity) which prices swap at par.

Basis swaps are further instruments which are liquidely traded and also used for curve calibration (2/2)

- Note that Libor indices are currently beeing phased out of the market.
- Consequently, tenor basis swaps will likely become less relevant.
- In EUR the following swap instruments are quoted:
  - OIS ("€STR") vs. fixed,
  - 6m Euribor vs. fixed,
  - 6m Euribor vs. 3m Euribor plus spread.
- EUR instruments allow for the following procedure:
  - First calibrate OIS (i.e. €STR) discount curve P and 6m projection curve P^{6m}.
  - Then use P and P^{6m} and calibrate P^{3m} from quoted tenor basis spreads.

Cross currency basis swaps reference overnight rates in two currencies

Cross Currency Basis Swap

In a (constant notional) cross currency basis swap floating rate payments in one currency are exchanged against floating rate payments in another currency plus fixed spread,

$$\begin{aligned}\text{XCcySwap}(t) = N_1 &\left\{ \sum_{j=1}^{m_1} \mathbb{E}_t^{\tilde{T}_j}\left[\tilde{C}_j^1\right]\tilde{\tau}_j P^1(t,\tilde{T}_j) + P^1(t,\tilde{T}_{m_1}) \right\} \\ &- Fx(t) N_2 \left\{ \sum_{j=1}^{m_2} \left[\mathbb{E}_t^{\hat{T}_j}\left[\hat{C}_j^2\right] + s\right]\hat{\tau}_j P^2(t,\hat{T}_j) + P^2(t,\hat{T}_{m_2}) \right\}.\end{aligned}$$

- $\tilde{C}_j^1$ and $\hat{C}_j^2$ are compounded overnight rates (like OIS).
- N_1 domestic currency notional, N_2 foreign currency notional.
- $Fx(t)$ spot FX rate CCY2 / CCY1.
- At trade date t_d notionals N_1 and N_2 are exchanged at time-t_d spot FX rate, i.e. $N_1 = Fx(t_d) N_2$.

We have a look at the curves involved (1/2)

projection curve from CCY-1 OIS

discount curve from CCY-1 OIS

$$
\begin{aligned}
\mathrm{XCcySwap}(t) = {} & N_1 \left\{ \sum_{j=1}^{m_1} L^1(t, \tilde{T}_{j-1}, \tilde{T}_j)\tilde{\tau}_j P^1(t, \tilde{T}_j) + P^1(t, \tilde{T}_{m_1}) \right\} \\
& - Fx(t) N_2 \left\{ \sum_{j=1}^{m_2} \left[L^2(t, \hat{T}_{j-1}, \hat{T}_j) + s \right] \hat{\tau}_j P^2(t, \hat{T}_j) + P^2(t, \hat{T}_{m_2}) \right\}
\end{aligned}
$$

projection curve from CCY-2 OIS

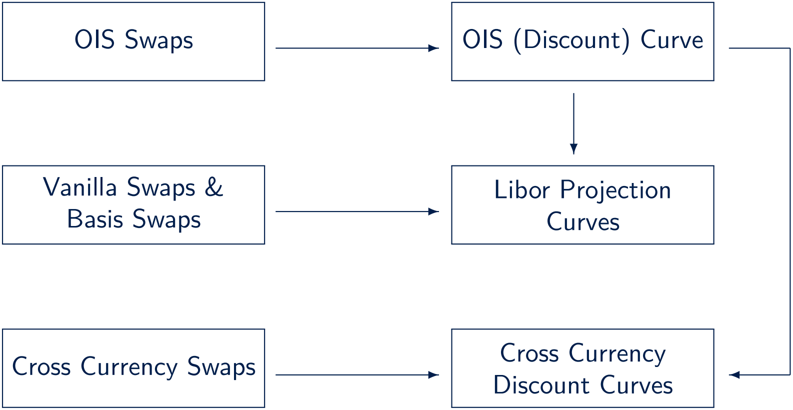
discount curve specific to XCCY discounting in CCY-2

- Cross currency swaps require particular discount curves.
- Cross currency discount curves (here P^2) are calibrated from quoted cross currency swap spreads (here s).

We have a look at the curves involved (2/2)

- Theoretical background is established via Collateralised Discounting.
- For details, see e.g. M. Fujii and Y Shimada and A. Takahashi, Collateral Posting and Choice of Collateral Currency - Implications for Derivative Pricing and Risk Management. https://ssrn.com/abstract=1601866.

In summary multi-curve calibration leads to a hierarchy of discount and projection curves



Outline

What are the parameters we need to calibrate in Hull-White model?

forward rate from initial discount curve $P(0, t)$

short rate volatility from Vanilla options

$$
\begin{aligned}
r(t) &= f(0,t) + x(t) \\
dx(t) &= \left[\int_0^t \sigma(u)^2 \cdot e^{-2a(t-u)} du - a \cdot x(t) \right] \cdot dt + \sigma(t) \cdot dW(t) \\
x(0) &= 0
\end{aligned}
$$

mean reversion e.g. from other exotic option prices

- Short rate volatility $\sigma(t)$ mainly impacts overall variance of the rates.
- Mean reversion a impacts forward volatility (and other related properties).

We first focus on volatility calibration (assuming mean reversion externally specified) and then look into mean reversion calibration.

Outline

Market instruments for Volatility calibration are European swaptions

EUR ATM Swaption Straddles - BP Volatilities (Calendar day vols)
Please call +44 (0)20 7532 3080 for further details

	1Y	2Y	3Y	4Y	5Y	7Y	10Y	15Y	20Y	25Y	30Y
1M Opt	45.3	45.0	48.0	53.8	55.3	61.6	70.1	78.6	85.2	88.7	90.0
2M Opt	38.8	40.9	44.8	48.3	51.4	58.6	67.0	76.3	82.5	84.5	85.5
3M Opt	35.6	37.3	41.7	46.8	50.9	58.3	66.7	75.0	80.5	82.5	84.1
6M Opt	34.9	37.7	42.1	46.9	51.0	59.3	66.3	74.1	78.7	80.1	81.3
9M Opt	35.4	38.0	43.1	47.3	51.5	59.1	66.9	73.8	77.5	78.7	79.0
1Y Opt	37.0	40.3	44.3	48.1	52.4	59.8	67.0	73.2	76.0	77.2	77.4
18M Opt	41.3	44.7	48.0	50.6	55.0	61.6	68.3	72.6	74.8	75.7	76.1
2Y Opt	46.5	49.4	52.6	55.0	58.2	63.9	69.8	73.0	74.2	75.1	75.5
3Y Opt	56.9	58.8	60.6	62.5	64.4	68.3	72.6	73.2	72.9	73.4	73.7
4Y Opt	64.1	65.5	66.0	67.4	68.6	71.1	73.9	72.4	71.5	71.1	71.0
5Y Opt	68.7	69.2	70.0	70.8	71.5	73.0	74.7	71.8	70.2	69.3	69.0
7Y Opt	73.0	73.3	73.6	73.8	74.1	74.5	74.8	70.1	67.6	66.4	66.0
10Y Opt	73.2	73.8	74								
15Y Opt	70.8	71.2	71								
20Y Opt	67.7	68.4	68								
25Y Opt	64.6	64.8	64								
30Y Opt	60.4	60.9	59								

EUR Vega - Normal Vol Skews

	Receivers						Payers				
	-200	-150	-100	-50	-25	ATM	+25	+50	+100	+150	+200
1y2y		22.29	14.02	5.40	1.84	40.72	0.91	4.20	13.83	24.45	
1y5y		0.20	-2.25	-2.44	-1.59	52.79	2.29	5.14	11.97	19.60	
1y10y		0.24	-1.69	-2.09	-1.40	67.86	2.10	4.80	11.53	19.32	
1y20y		13.45	7.57	2.33	0.63	76.97	0.67	2.64	9.62	18.89	
1y30y		7.75	4.34	1.43	0.46	79.14	0.16	1.00	4.59	10.15	
2y2y		11.95	6.40	1.54	0.14	49.98	1.32	3.90	11.32	19.98	
2y5y		-3.21	-3.26	-2.23	-1.28	58.62	1.61	3.52	8.09	13.38	
2y10y		-3.50	-2.97	-1.83	-1.01	70.41	1.21	2.63	6.04	10.10	
2y20y		1.10	0.20	-0.30	-0.28	75.04	0.57	1.44	4.09	7.81	
2y30y		4.86	2.51	0.58	0.07	76.50	0.46	1.48	5.08	10.29	
5y2y		-1.06	-1.41	-1.15	-0.70	69.84	0.95	2.14	5.18	8.91	
5y5y		-3.97	-2.93	-1.73	-0.94	72.02	1.11	2.39	5.42	9.00	
5y10y		-3.73	-2.55	-1.40	-0.74	75.23	0.84	1.80	4.04	6.69	
5y20y		-1.66	-1.12	-0.68	-0.38	70.67	0.49	1.10	2.72	4.85	
5y30y		-1.51	-0.99	-0.61	-0.35	69.54	0.47	1.07	2.69	4.86	
10y2y		-3.45	-2.56	-1.43	-0.75	74.34	0.83	1.74	3.79	6.11	
10y5y		-4.90	-3.28	-1.70	-0.87	74.37	0.92	1.89	4.00	6.33	
10y10y		-3.04	-1.95	-1.03	-0.54	73.36	0.60	1.29	2.91	4.89	
10y20y		-2.31	-1.32	-0.64	-0.33	65.38	0.36	0.78	1.79	3.08	
10y30y		-1.95	-1.17	-0.65	-0.36	63.77	0.46	1.02	2.53	4.54	

For Hull-White model calibration we assume that we can already price European swaptions at market level

- In practice, European swaption models depend on available market data (and business case).
- If only normal ATM volatilities are available (or should be used)[8]
  - interpolate ATM volatilities,
  - assume normal model $dS = \sigma dW$,
  - use Bachelier formula for Swaption pricing.
- If Swaption smile data is available (in addition to ATM prices/volatilities)
  - calibrate e.g. Shifted SABR models per expiry/swap term to available data,
  - interpolate models (e.g. via SABR model parameters β, ρ, ν),
  - make sure interpolated model fits (interplated) ATM swaption data (e.g. calibrate SABR α individually),
  - use interpolated model to price European swaption.

[8] Same holds for (shifted) lognormal volatilities and corresponding basic models. But keep in mind implicit smile assumption!

How can we use European swaption prices to calibrate Hull-White volatility?

$$V^{\text{Swpt}}(T_E) = \left[\phi\left\{K\sum_{i=1}^{n}\tau_i P(T_E, T_i) - \sum_{j=1}^{m} L^{\delta}(T_E, \tilde{T}_{j-1}, \tilde{T}_j)\tilde{\tau}_j P(T_E, \tilde{T}_j)\right\}\right]^+ .$$

$$V^{\text{CBO}}(T_E) = \left[\phi\left\{\sum_{k=0}^{n+m+1} C_k \cdot P(T_E, \bar{T}_k)\right\}\right]^+ .$$

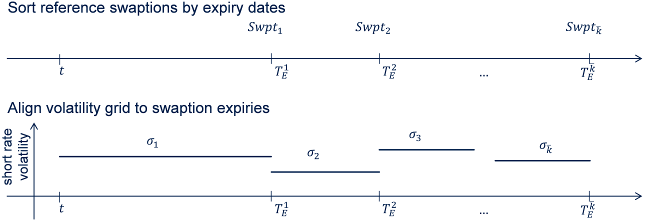
$$V^{\text{Swpt}}(t) = V^{\text{CBO}}(t) = \sum_{k=0}^{n+m+1} C_k \cdot V_k^{\text{ZBO}}(t).$$

$$V_k^{\text{ZBO}}(t) = P(t, T_E) \cdot \text{Black}\left(P(t, \bar{T}_k)/P(t, T_E), R_k, \nu_k, \phi\right),$$

$$\nu_k = G(T_E, \bar{T}_k)^2 \int_t^{T_E} \left[e^{-a(T_E-u)}\sigma(u)\right]^2 du.$$

Price of a European swaption depends on short rate volatility $\sigma(t)$ from $t = 0$ to swaption expiry T_E.

We can calibrate a piece-wise constant volatility to a strip of reference European swaptions



We set up calibration helpers

$$\mathcal{H}_k\left[\sigma\right]\left(V_k^{\text{Swpt}}\right) = \underbrace{V_k^{\text{CBO}}(t)}_{\text{Model}[\sigma]} - \underbrace{V_k^{\text{Swpt}}}_{\text{Market}\left(\sigma_N^k\right)} .$$

- $V_k^{\text{CBO}}(t)$ Hull-White model price of swaption represented as coupon bond option.
- V_k^{Swpt} (quasi-)market price of swaption obtained from Vanilla model or implied (normal) volatility.

Calibration problem is formulated in terms of short rate volatility values

Set

$$\sigma(t) = \sigma\left[\sigma_1, \ldots, \sigma_{\bar{k}}\right](t) = \sum_{k=1}^{\bar{k}} \mathbb{1}_{\{T_E^{k-1} \le t < T_E^k\}} \cdot \sigma_k.$$

- Assume distinct expiry/grid dates T_E^k for reference swaptions.
- Assume mean reversion is exogenously given.

Hull-White Volatility Calibration Problem

For a given set of market quotes (or Vanilla model prices) $\left\{V_k^{\text{Swpt}}\right\}_{k=1,\ldots,\bar{k}}$ of reference European swaptions with corresponding calibration helpers $\mathcal{H}_k\left[\sigma\left[\sigma_1, \ldots, \sigma_{\bar{k}}\right]\right]$ the Hull-White volatility calibration problem is given as

$$\min_{\sigma_1,\ldots,\sigma_{\bar{k}}} \left\| \left[\mathcal{H}_1[\sigma]\left(V_1^{\text{Swpt}}\right), \ldots, \mathcal{H}_{\bar{k}}[\sigma]\left(V_{\bar{k}}^{\text{Swpt}}\right)\right]^\top \right\|.$$

We analyse the optimisation problem in more detail.

Multi-dimensional calibration problem can be decomposed into sequence of one-dimensional problems

Note that for $l > k$

$$\frac{d}{d\sigma_l}\mathcal{H}_k\left[\sigma\left[\sigma_1,\ldots,\sigma_{\bar{k}}\right]\right] = 0$$

Thus we could write

$$\begin{aligned}
\mathcal{H}_1\left[\sigma\left[\sigma_1\right]\right] &= 0,\\
\mathcal{H}_2\left[\sigma\left[\sigma_1,\sigma_2\right]\right] &= 0,\\
&\vdots\\
\mathcal{H}_{\bar{k}}\left[\sigma\left[\sigma_1,\sigma_2,\ldots,\sigma_{\bar{k}}\right]\right] &= 0.
\end{aligned}$$

System of equations can be solved row-by-row (i.e. bootstrapping method) via one-dimensional root search method.

Sequentiel Hull-White volatility calibration is analogous to yield curve bootstrapping.

We can also formulate general optimisation problem if short rate volatilities and reference swaptions are not aligned

Suppose time grid $0 = t_0, t_1, \ldots, t_n$ and piece-wise constant volatility $\sigma(t)$ via $\bar{\sigma} = [\sigma_1, \ldots, \sigma_n]^\top$

$$\sigma(t) = \sigma[\bar{\sigma}](t) = \sum_{k=1}^{n} \mathbb{1}_{\{t_{k-1} \leq t < t_k\}} \cdot \sigma_k.$$

Denote $V^{\text{Swpt}} = \left[V_1^{\text{Swpt}}, \ldots, V_q^{\text{Swpt}} \right]$ a set of reference European swaption prices with calibration helper

$$\mathcal{H}[\sigma[\bar{\sigma}]]\left(V^{\text{Swpt}}\right) = \left[\mathcal{H}_1[\sigma[\bar{\sigma}]]\left(V_1^{\text{Swpt}}\right), \ldots, \mathcal{H}_q[\sigma[\bar{\sigma}]]\left(V_q^{\text{Swpt}}\right) \right].$$

Then calibration problem becomes

$$\min_{\bar{\sigma}} \left\| \mathcal{H}[\sigma[\bar{\sigma}]]\left(V^{\text{Swpt}}\right) \right\|.$$

The choice of reference European swaptions is critical for model calibration - What is the usage of your model?

Global calibration to available market data

General purpose calibration for yield curve simulation or pricing of a variety of products with same model.

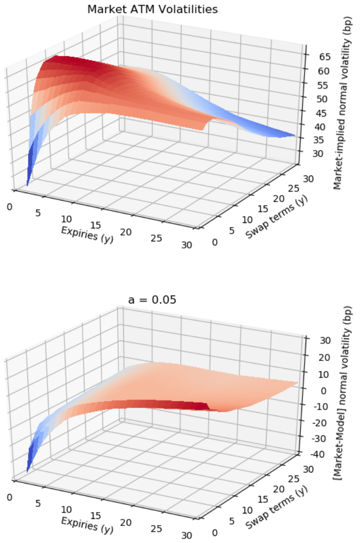
- Keep in mind model properties and limitations.
- HW model cannot model smile - use more liquidly traded ATM swaptions.
- Do not use too many reference swaptions per expiry - HW model has only one volatility parameter per expiry.

Product-specific calibration

Price a particular exotic product while focussing on consistent pricing of related simple products.

- Identify building blocks of exotic product - these are typically priced on simpler models if modelled as stand-alone product.
- Calibrate HW model to prices of building blocks obtained from simpler model.

We illustrate market volatilities and global calibration fit



Lower mean reversion appears to yield slightly better global fit.

Building blocks of Bermudan swaption are co-terminal European swaptions

Recall decomposition

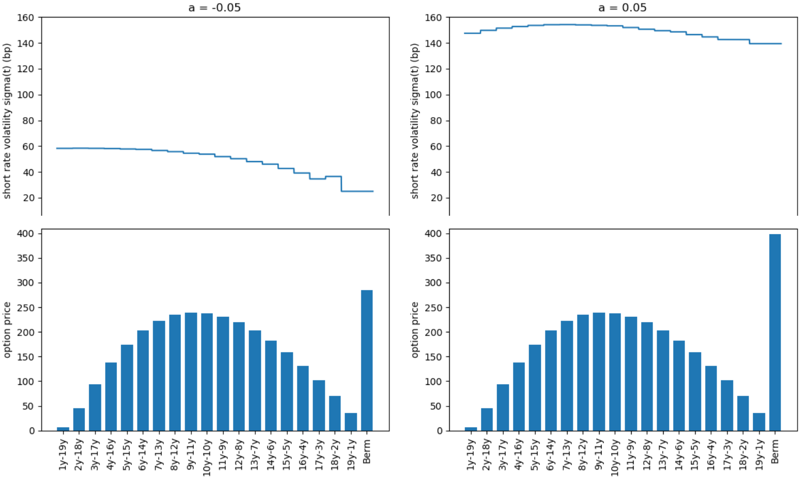
$$V^{\text{Berm}}(t) = \max_k \left\{ V_k^{\text{Swpt}}(t) \,|\, k = 1, \ldots, \bar{k} \right\} + \text{SwitchOption(t)}$$

where $V_k^{\text{Swpt}}(t)$ is price of European option to enter into swap at T_E^k (plus spot) with fixed maturity T_n.

- European swaption prices $V_k^{\text{Swpt}}(t)$ can be obtained from Vanilla model.
- Consistent Hull-White model must produce max-European price $\max_k \left\{ V_k^{\text{Swpt}}(t) \,|\, k = 1, \ldots, \bar{k} \right\}$ consistent to Vanilla model.

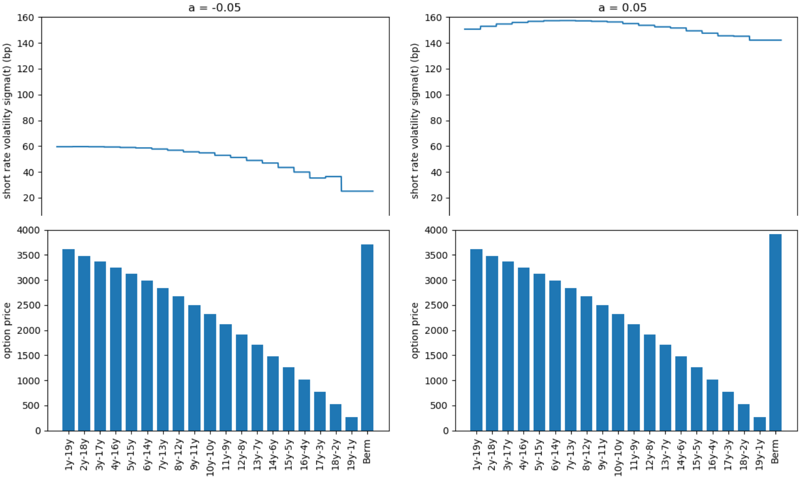
Hull-White model for Bermudan pricing is calibrated to corresponding co-terminal European swaptions.

20y-nc1y 3% Receiver Bermudan, (Fwd-)Rates at 5% (flat) and Implied Vols at 100bp (flat)



Out-of-the-money option shows concave co-terminal European swaption profile.

20y-nc1y 3% Receiver Bermudan, (Fwd-)Rates at 1% (flat) and Implied Vols at 100bp (flat)



In-the-money option shows decreasing co-terminal European swaption profile.

Outline

Mean reversion controls switch option value of Bermudan swaption

Recall decomposition of Bermudan price into max-European price plus residual switch value

$$V^{\text{Berm}}(t) = \max_k \left\{ V_k^{\text{CBO}}(t) \,|\, k = 1, \ldots, \bar{k} \right\} + \text{SwitchOption(t)}.$$

- $V_k^{\text{CBO}}(t)$ is the Hull-White price of the co-terminal European swaptions reformulated as bond option.
- SwitchOption(t) is the Hull-White price of the option to postpone exercise decision.

We get

$$\frac{\partial}{\partial a} V^{\text{Berm}}(t) = \frac{\partial}{\partial a} \max_k \left\{ V_k^{\text{CBO}}(t) \,|\, k = 1, \ldots, \bar{k} \right\} + \frac{\partial}{\partial a} \text{SwitchOption(t)}.$$

Our model calibration approach to European swaption market prices partly eliminates mean reversion dependency

We recall

$$\frac{\partial}{\partial a} V^{\text{Berm}}(t) = \frac{\partial}{\partial a} \max_k \left\{ V_k^{\text{CBO}}(t) \,|\, k = 1, \ldots, \bar{k} \right\} + \frac{\partial}{\partial a} \text{SwitchOption(t)}.$$

If model is calibrated to match co-terminal swaptions from market prices V_k^{Swpt} then

$$V_k^{\text{CBO}}(t) = V_k^{\text{Swpt}} \quad \forall a.$$

Thus

$$\frac{\partial}{\partial a} V_k^{\text{CBO}}(t) = 0 \quad (\forall k) \quad \text{and} \quad \frac{\partial}{\partial a} \max_k \left\{ V_k^{\text{CBO}}(t) \,|\, k = 1, \ldots, \bar{k} \right\} = 0.$$

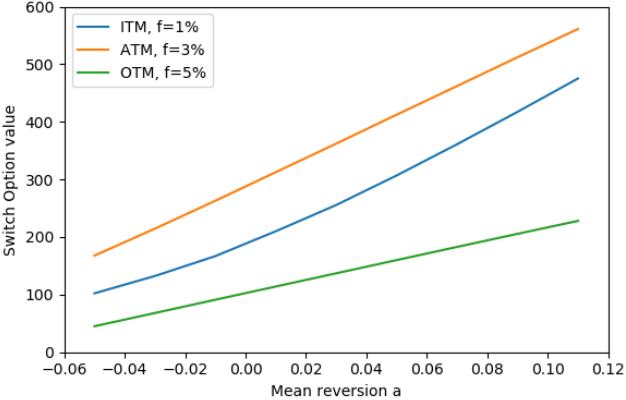
Consequently,

$$\frac{\partial}{\partial a} V^{\text{Berm}}(t) = \frac{\partial}{\partial a} \text{SwitchOption(t)}.$$

This is an important result wich shows difference between European and Bermudan Swaptions.

Switch option value (and Bermudan price) increase as mean reversion increases

- 20y-nc1y 3% Receiver Bermudan, (Fwd-)Rates $f \in \{1\%, 3\%, 5\%\}$ (flat) and Implied Vols at 100bp (flat):



If prices for reference Bermudan options are available we can use these prices to calibrate mean reversion.

If we don't have Bermudan prices available we can resort to alternative objectives to calibrate mean reversion

- Ratio of short-tenor and long-tenor option volatilities.
- Auto-correlation (or inter-temporal correlation) of historical rates.
- Payment-delay convexity adjustment.

Mean reversion impacts the slope of short-tenor volatilities versus long-tenor volatilities

- For the analysis of short- vs. long-tenor volatilities we make several approximations.
- Consider continuous forward yield

$$F(t, T_0, T_M) = \ln\left[\frac{P(t, T_0)}{P(t, T_M)}\right]\frac{1}{T_M - T_0}.$$

- We will analyse standard deviation ratio for a $T_M - T_0$ forward yield and a $T_N - T_0$ forward yield,

$$\lambda = \frac{\sqrt{\mathrm{Var}\left[F(T_0, T_0, T_M) \mid \mathcal{F}_t\right]}}{\sqrt{\mathrm{Var}\left[F(T_0, T_0, T_N) \mid \mathcal{F}_t\right]}}.$$

How are forward yields (and standard dev's) related to forward swap rates (and implied volatilities)?

We approximate swap rate by continuous forward yield I

Consider swap rate with start date T_0 and maturity T_M

$$S(t) = \frac{\sum_j L_j^\delta(t)\tilde{\tau}_j P(t, \tilde{T}_j)}{\sum_i \tau_i P(t, T_i)}.$$

First we rewrite swap rate in terms of single-curve rate plus basis spread)

$$S(t) = \frac{\sum_j L_j(t)\tilde{\tau}_j P(t, \tilde{T}_j)}{\sum_i \tau_i P(t, T_i)} + \underbrace{\frac{\sum_j \left[D_j^\delta - 1\right]\tilde{\tau}_j P(t, \tilde{T}_{j-1})}{\sum_i \tau_i P(t, T_i)}}_{b(t)}.$$

Assume $b(t)$ is deterministic (similar to assuming D_j^δ are deterministic). Simplifying single-curve swap rate yields

$$S(t) = \frac{P(t, T_0) - P(t, T_M)}{\sum_i \tau_i P(t, T_i)} + b(t).$$

We approximate swap rate by continuous forward yield II

Approximate annuity with only single long fixed-leg period T_0 to T_M with $\tau_1 = T_M - T_0$.
Then

$$S(t) \approx \frac{P(t, T_0) - P(t, T_M)}{(T_M - T_0)\, P(t, T_M)} + b(t) = \left[\frac{P(t, T_0)}{P(t, T_M)} - 1\right] \frac{1}{T_M - T_0} + b(t).$$

First-order Taylor-approximation $\ln(x) \approx x - 1$ leads to

$$S(t) \approx \ln\left[\frac{P(t, T_0)}{P(t, T_M)}\right] \frac{1}{T_M - T_0} + b(t) = F(t, T_0, T_M) + b(t).$$

Deterministic basis spread assumption for $b(t)$ yields

$$\mathrm{Var}\left[S(T_0) \mid \mathcal{F}_t\right] \approx \mathrm{Var}\left[F(T_0, T_0, T_M) \mid \mathcal{F}_t\right].$$

Also we approximate implied ATM volatility with standard deviation

Swap rate $S(t)$ is approximately normally distributed in Hull-White model. Thus

$$dS(t) \approx \sigma_S(t) dW^A(t)$$

for a deterministic volatility function $\sigma_S(t)$ depending on Hull-White model parameters.
Ito-isometry yields

$$\nu^2 = \mathrm{Var}\left[S(T_0) \mid \mathcal{F}_t\right] = \int_t^{T_0} \left[\sigma_S(t)\right]^2 dt.$$

Vanilla options depend only on terminal distribution of swap rate. Thus an alternative swap rate with

$$d\tilde{S}(t) \approx \sigma_N dW^A(t) \quad \text{with} \quad \sigma_N^2 = \nu^2 / (T_0 - t)$$

yields same Vanilla option prices.
However, by construction σ_N is also the implied normal volatility of $\tilde{S}(T_0)$ and $S(T_0)$. This yields the relation

$$\mathrm{Var}\left[S(T_0) \mid \mathcal{F}_t\right] = \sigma_N^2 (T_0 - t).$$

We get the relation of the volatility ratio I

$$\lambda = \frac{\sqrt{\mathrm{Var}\left[F(T_0, T_0, T_M) \mid \mathcal{F}_t\right]}}{\sqrt{\mathrm{Var}\left[F(T_0, T_0, T_N) \mid \mathcal{F}_t\right]}} \approx \frac{\sqrt{\left[\sigma_N^{T_0, T_M}\right]^2 (T_0 - t)}}{\sqrt{\left[\sigma_N^{T_0, T_N}\right]^2 (T_0 - t)}} = \frac{\sigma_N^{T_0, T_M}}{\sigma_N^{T_0, T_N}}.$$

It remains to calculate $\mathrm{Var}\left[F(T_0, T_0, T_M) \mid \mathcal{F}_t\right]$ with

$$F(T_0, T_0, T_M) = \ln\left[\frac{1}{P(T_0, T_M)}\right] \frac{1}{T_M - T_0} = -\frac{\ln\left[P(T_0, T_M)\right]}{T_M - T_0}.$$

From $P(T_0, T_M) = \frac{P(t, T_M)}{P(t, T_0)} e^{-G(T_0, T_M)x(T_0) - \frac{1}{2}G(T_0, T_M)^2 y(T_0)}$ we get

$$\begin{aligned} F(T_0, T_0, T_M) &= -\left\{\ln\left[\frac{P(t, T_M)}{P(t, T_0)}\right] - G(T_0, T_M)x(T_0) - \frac{1}{2}G(T_0, T_M)^2 y(T_0)\right\} / (T_M - T_0) \\ &= F(t, T_0, T_M) + \frac{G(T_0, T_M)x(T_0) - \frac{1}{2}G(T_0, T_M)^2 y(T_0)}{T_M - T_0}. \end{aligned}$$

We get the relation of the volatility ratio II

This yields

$$\mathrm{Var}\left[F(T_0, T_0, T_M) \mid \mathcal{F}_t\right] = \frac{G(T_0, T_M)^2}{(T_M - T_0)^2} \mathrm{Var}\left[x(T_0) \mid \mathcal{F}_t\right]$$

and

$$\lambda = \frac{\sqrt{\mathrm{Var}\left[F(T_0, T_0, T_M) \mid \mathcal{F}_t\right]}}{\sqrt{\mathrm{Var}\left[F(T_0, T_0, T_N) \mid \mathcal{F}_t\right]}} = \frac{G(T_0, T_M)/\,(T_M - T_0)}{G(T_0, T_N)/\,(T_N - T_0)}.$$

Substituting $G(T_0, T_1) = \left[1 - e^{-a(T_1 - T_0)}\right]/a$ yields

$$\lambda = \frac{\left[1 - e^{-a(T_M - T_0)}\right]/\,(T_M - T_0)}{\left[1 - e^{-a(T_N - T_0)}\right]/\,(T_N - T_0)}.$$

Note that

- λ is independent of short rate volatility $\sigma(t)$,
- λ only depends on mean reversion and time differences (i.e. swap terms) $T_M - T_0$ and $T_N - T_0$.

Further simplification gives a relation only depending on $T_M - T_N$

Consider second order Taylor approximation

$$e^{-a(T_M - T_0)} \approx 1 - a\left(T_M - T_0\right) + \frac{1}{2}a^2\left(T_M - T_0\right)^2.$$

This yields

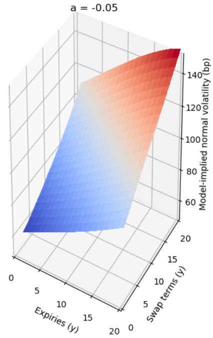
$$\begin{aligned}
\lambda &\approx \frac{\left[a\left(T_M - T_0\right) - \frac{1}{2}a^2\left(T_M - T_0\right)^2\right] / \left(T_M - T_0\right)}{\left[a\left(T_N - T_0\right) - \frac{1}{2}a^2\left(T_N - T_0\right)^2\right] / \left(T_N - T_0\right)} \\
&= \frac{1 - \frac{1}{2}a\left(T_M - T_0\right)}{1 - \frac{1}{2}a\left(T_N - T_0\right)} \approx \frac{e^{-\frac{1}{2}a(T_M - T_0)}}{e^{-\frac{1}{2}a(T_N - T_0)}} \\
&= e^{-\frac{1}{2}a(T_M - T_N)}.
\end{aligned}$$

Finally, we end up with

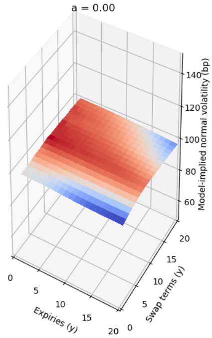
$$\frac{\sigma_N^{T_0, T_M}}{\sigma_N^{T_0, T_N}} \approx e^{-\frac{1}{2}a(T_M - T_N)}.$$

The relation $\sigma_N^{T_0,T_M}/\sigma_N^{T_0,T_N} \approx e^{-\frac{1}{2}a(T_M-T_N)}$ can be verified numerically

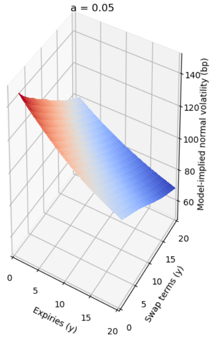
- Use flat short rate volatility σ - calibrated to 10y-10y swaption with $100bp$ volatility.
- Mean reversion $a \in \{-5\%, 0\%, 5\%\}$:



increasing



flat



decreasing

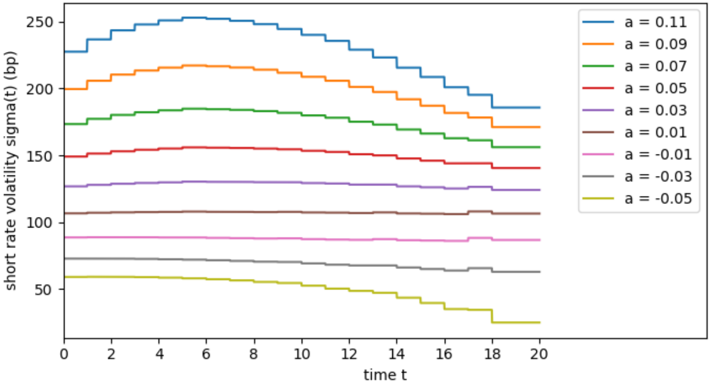
We can use volatility ratio property with co-terminal swaption volatility calibration

- Consider improvement of overall fit to ATM volatility surface as general calibration objective.
- Calibrate mean reversion to ratio of
  - first exercise and co-terminal swap rate and
  - first exercise and short-term swap rate.

	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	17y	18y	19y	20y
1y	26.2	32.5	38.2	43.1	46.8	49.5	51.8	54.0	55.2	56.0	56.5	56.9	57.3	57.8	58.2	58.3	58.4	58.4	58.5	58.5
2y	38.5	43.3	47.2	50.3	52.7	54.9	56.7	58.0	59.0	59.7	59.9	60.0	60.2	60.3	60.5	60.5	60.5	60.6	60.6	60.6
3y	50.6	52.7	54.7	56.5	58.0	59.5	60.6	61.6	62.3	62.9	62.7	62.5	62.2	62.0	61.8	61.7	61.6	61.5	61.4	61.3
4y	57.7	58.9	59.7	60.9	61.8	62.6	63.4	64.0	64.5	64.8	64.4	63.9	63.4	62.9	62.5	62.3	62.1	61.9	61.7	61.5
5y	62.1	62.6	63.1	63.7	64.3	64.8	65.3	65.8	66.1	66.2	65.6	64.9	64.3	63.6	63.0	62.6	62.3	61.9	61.6	61.2
6y	64.4	64.7	65.0	65.3	65.6	65.9	66.2	66.4	66.6	66.6	65.8	65.0	64.1	63.3	62.5	62.1	61.7	61.4	61.0	60.6
7y	66.3	66.6	66.8	66.8	66.8	66.9	67.0	67.1	67.0	66.9	66.0	65.1	64.2	63.3	62.4	61.9	61.4	61.0	60.5	60.0
8y	66.4	66.7	66.9	67.1	67.1	67.1	67.1	67.0	66.9	66.6	65.7	64.7	63.8	62.8	61.9	61.4	60.9	60.3	59.8	59.3
9y	66.5	66.8	67.1	67.2	67.3	67.3	67.2	67.0	66.7	66.3	65.3	64.3	63.4	62.4	61.4	60.9	60.3	59.8	59.2	58.6
10y	66.6	66.9	67.1	67.2	67.2	67.2	67.1	66.8	66.5	66.0	65.0	64.0	63.0	61.9	60.9	60.3	59.8	59.2	58.6	58.0
11y	65.7	66.0	66.2	66.3	66.3	66.3	66.2	65.9	65.6	65.2	64.1	63.1	62.0	61.0	60.0	59.3	58.7	58.1	57.5	56.8
12y	64.7	65.0	65.3	65.4	65.4	65.4	65.3	65.1	64.7	64.3	63.3	62.2	61.1	60.1	59.0	58.3	57.7	57.0	56.3	55.7
13y	63.8	64.1	64.4	64.5	64.5	64.5	64.4	64.2	63.9	63.5	62.4	61.3	60.2	59.1	58.1	57.3	56.6	55.9	55.2	54.5
14y	62.9	63.2	63.4	63.6	63.6	63.6	63.6	63.3	63.0	62.6	61.5	60.4	59.3	58.2	57.1	56.3	55.6	54.8	54.0	53.3
15y	61.9	62.2	62.5	62.7	62.7	62.8	62.7	62.5	62.2	61.7	60.6	59.5	58.4	57.3	56.2	55.3	54.5	53.7	52.9	52.1
16y	61.0	61.3	61.6	61.7	61.8	61.8	61.7	61.5	61.2	60.7	59.6	58.5	57.4	56.3	55.2	54.4	53.5	52.7	51.9	51.0
17y	60.1	60.4	60.6	60.8	60.9	60.9	60.8	60.5	60.2	59.7	58.6	57.5	56.4	55.3	54.2	53.4	52.5	51.7	50.8	50.0
18y	59.1	59.4	59.7	59.9	60.0	60.0	59.8	59.6	59.2	58.7	57.6	56.5	55.4	54.3	53.2	52.4	51.5	50.7	49.8	49.0
19y	58.2	58.5	58.8	59.0	59.1	59.1	58.9	58.6	58.2	57.7	56.6	55.5	54.4	53.3	52.2	51.4	50.5	49.7	48.8	47.9
20y	57.3	57.6	57.8	58.1	58.1	58.1	57.9	57.6	57.2	56.6	55.6	54.5	53.4	52.3	51.3	50.4	49.5	48.6	47.8	46.9

Another calibration objective is time-stationarity of the model

- Based on mean reversion the calibrated term-structure of short rate volatilities changes:



We can choose mean reversion such that calibrated short rate volatility is as close to constant as possible.

An alternative view on mean reversion is obtained via auto-correlation

Consider

$$F(T_0, T_0, T_M) = F(t, T_0, T_M) + \frac{G(T_0, T_M)x(T_0) - \frac{1}{2}G(T_0, T_M)^2 y(T_0)}{T_M - T_0}.$$

Then

$$\mathrm{Corr}\left[F(T_0, T_0, T_M), F(T_1, T_1, T_N)\right] = \mathrm{Corr}\left[x(T_0), x(T_1)\right].$$

We have

$$x(T) = e^{-a(T-t)}\left[x(t) + \int_t^T e^{a(u-t)}\left(y(u)du + \sigma(u)dW(u)\right)\right].$$

It follows for $T_1 > T_0$ (see exercises or literature)

$$\mathrm{Corr}\left[x(T_0), x(T_1)\right] = e^{-2a(T_1 - T_0)}\sqrt{\frac{1 - e^{-2aT_0}}{1 - e^{-2aT_1}}}.$$

Auto-correlation (or inter-temporal correlation) is independent of volatility $\sigma(t)$ and maturities T_M and T_N.

Auto-correlation property is sometimes used to calibrate mean reversion to interest rate time series

Consider limit $T_0 \to \infty$ then

$$\text{Corr}\left[x(T_0), x(T_1)\right] \approx e^{-2a(T_1 - T_0)}.$$

- Use a time-series of proxy rates $\{R(t_k)\}_{k=1,2,\ldots}$ and estimate $\rho(\Delta)=\text{Corr}\left[R(t_k), R(t_k+\Delta)\right]$.
- Find mean reversion a such that

$$\rho(\Delta) \approx e^{-2a\Delta}.$$

- However, method strongly depends on the choice of proxy rate and estimation time window.
- Also, mean reversion in risk-neutral measure needs to be distinguished from mean reversion in real-world measure, see e.g. Sec. 18 in
  - R. Rebonato. *Volatility and Correlation.* John Wiley & Sons, 2004

Outline

Summary on Hull-White model calibration

- Hull-White model calibration is distinguished between
  - short rate volatility calibration,
  - mean reversion parameter calibration.
- Short rate volatility is calibrated product-specific to match relevant Vanilla options.
  - For Bermudan swaptions these are co-terminal European swaptions.
- Mean reversion calibration involves subjective judgement regarding calibration objective.
  - Fit to reference exotic prices (e.g. Bermudans) if available.
  - Improve overall calibration fit to ATM swaption volatilities or time-stationarity of model.

Part VII

Sensitivity Calculation

Outline

Outline

Introduction to Sensitivity Calculation

Finite Difference Approximation for Sensitivities

Differentiation and Calibration

A brief Introduction to Algorithmic Differentiation

Why do we need sensitivities?

Consider a (differentiable) pricing model $V = V(p)$ based on some input parameter p. Sensitivity of V w.r.t. changes in p is

$$V'(p) = \frac{dV(p)}{dp}.$$

- Hedging and risk management.
- Market risk measurement.
- Many more applications for accounting, regulatory reporting, ...

Sensitivity calculation is a crucial function for banks and financial institutions.

Derivative pricing is based on hedging and risk replication

Recall fundamental derivative replication result

$$V(t) = V\left(t, X(t)\right) = \phi(t)^{\top} X(t) \text{ for all } t \in [0, T],$$

- $V(t)$ price of a contingent claim,
- $\phi(t)$ permissible trading strategy,
- $X(t)$ assets in our market.

How do we find the trading strategy?

Consider portfolio $\pi(t) = V(t, X(t)) - \phi(t)^{\top} X(t)$ and apply Ito's lemma

$$d\pi(t) = \mu_{\pi} \cdot dt + \left[\nabla_X \pi(t)\right]^{\top} \cdot \sigma_X^{\top} dW(t).$$

From replication property follows $d\pi(t) = 0$ for all $t \in [0, T]$. Thus, in particular

$$0 = \nabla_X \pi(t) = \nabla_X V(t, X(t)) - \phi(t).$$

This gives Delta-hedge

$$\phi(t) = \nabla_X V(t, X(t)).$$

Market risk calculation relies on accurate sensitivities (1/2)

Consider portfolio value $\pi(t)$, time horizon Δt and returns

$$\Delta\pi(t) = \pi(t) - \pi(t - \Delta t).$$

Market risk measure Value at Risk (VaR) is the lower quantile q of distribution of portfolio returns $\Delta\pi(t)$ given a confidence level $1 - \alpha$, formally

$$\mathrm{VaR}_\alpha = \inf\left\{q \quad s.t.\, \mathbb{P}\left\{\Delta\pi(t) \leq q \,|\, \pi(t)\right\} > \alpha\right\}.$$

Delta-Gamma VaR calculation method consideres $\pi(t) = \pi\left(X(t)\right)$ in terms of risk factors $X(t)$ and approximates

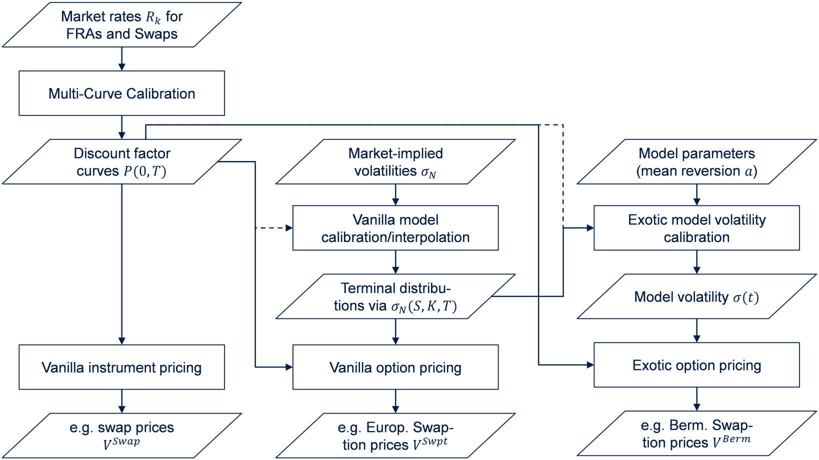
$$\Delta\pi \approx \left[\nabla_X \pi\left(X\right)\right]^\top \Delta X + \frac{1}{2}\Delta X^\top \left[H_X \pi\left(X\right)\right] \Delta X.$$

Market risk calculation relies on accurate sensitivities (2/2)

$$\Delta\pi \approx \left[\nabla_X \pi\,(X)\right]^\top \Delta X + \frac{1}{2}\Delta X^\top \left[H_X \pi\,(X)\right]\Delta X.$$

- VaR is calculated based on joint distribution of risk factor returns $\Delta X = X(t+\Delta t) - X(t)$ and sensitivities $\nabla_X \pi$ (gradient) and $H_X \pi$ (Hessian).
- Bank portfolio π may consist of linear instruments (e.g. swaps), Vanilla options (e.g. European swaptions) and exotic instruments (e.g. Bermudans).
- Common interest rate risk factors are FRA rates, par swap rates, ATM volatilities.

Sensitivity specification needs to take into account data flow and dependencies



Depending on context, risk factors can be market parameters or model parameters.

In practice, sensitivities are scaled relative to pre-defined risk factor shifts

Scaled sensitivity ΔV becomes

$$\Delta V = \frac{dV(p)}{dp} \cdot \Delta p \approx V(p + \Delta p) - V(p).$$

Typical scaling (or risk factor shift sizes) Δp are

- $1bp$ for interest rate shifts,
- $1bp$ for implied normal volatilities,
- 1% for implied lognormal or shifted lognormal volatilities.

Par rate Delta and Gamma are sensitivity w.r.t. changes in market rates (1/2)

Bucketed Delta and Gamma

Let $\bar{R} = [R_k]_{k=1,\ldots q}$ be the list of market quotes defining the inputs of a yield curve. The bucketed par rate delta of an instrument with model price $V = V(\bar{R})$ is the vector

$$\Delta_R = 1bp \cdot \left[\frac{\partial V}{\partial R_1}, \ldots, \frac{\partial V}{\partial R_q} \right].$$

Bucketed Gamma is calculated as

$$\Gamma_R = [1bp]^2 \cdot \left[\frac{\partial^2 V}{\partial R_1^2}, \ldots, \frac{\partial^2 V}{\partial R_q^2} \right].$$

- For multiple projection and discounting yield curves, sensitivities are calculated for each curve individually.

Par rate Delta and Gamma are sensitivity w.r.t. changes in market rates (2/2)

Parallel Delta and Gamma

Parallel Delta and Gamma represent sensitivities w.r.t. simultanous shifts of all market rates of a yield curve. With $\mathbf{1} = [1, \ldots 1]^\top$ we get

$$\bar{\Delta}_R = \mathbf{1}^\top \Delta_R = 1bp \cdot \sum_k \frac{\partial V}{\partial R_k} \approx \frac{V(\bar{R} + 1bp \cdot \mathbf{1}) - V(\bar{R} - 1bp \cdot \mathbf{1})}{2} \quad \text{and}$$

$$\bar{\Gamma}_R = \mathbf{1}^\top \Gamma_R = [1bp]^2 \cdot \sum_k \frac{\partial^2 V}{\partial R_k^2} \approx V(\bar{R} + 1bp \cdot \mathbf{1}) - 2V(\bar{R}) + V(\bar{R} - 1bp \cdot \mathbf{1}).$$

Vega is the sensitivity w.r.t. changes in market volatilities (1/2)

Bucketed ATM Normal Volatility Vega

Denote $\bar{\sigma} = \left[\sigma_N^{k,l}\right]$ the matrix of market-implied At-the-money normal volatilites for expiries $k = 1, \ldots, q$ and swap terms $l = 1, \ldots, r$.
Bucketed ATM Normal Volatility Vega of an instrument with model price $V = V(\bar{\sigma})$ is specified as

$$\text{Vega} = 1bp \cdot \left[\frac{\partial V}{\partial \sigma_N^{k,l}}\right]_{k=1,\ldots,q,\ l=1,\ldots,r}.$$

Vega is the sensitivity w.r.t. changes in market volatilities (2/2)

Parallel ATM Normal Volatility Vega

Parallel ATM Normal Volatility Vega represents sensitivity w.r.t. a parallel shift in the implied ATM swaption volatility surface. That is

$$
\begin{aligned}
\overline{\text{Vega}} &= 1bp \cdot \mathbf{1}^\top \left[\text{Vega}\right] \mathbf{1} \\
&= 1bp \cdot \sum_{k,l} \frac{\partial V}{\partial \sigma_N^{k,l}} \\
&\approx \frac{V\left(\bar{\sigma} + 1bp \cdot \mathbf{1}\,\mathbf{1}^\top\right) - V\left(\bar{\sigma} - 1bp \cdot \mathbf{1}\,\mathbf{1}^\top\right)}{2}.
\end{aligned}
$$

- Volatility smile sensitivities are often specified in terms of Vanilla model parameter sensitivities.
- For example, in SABR model, we can calculate sensitivities with respect to α, β, ρ and ν.

Outline

Crutial part of sensitivity calculation is evaluation or approximation of partial derivatives

Consider again general pricing function $V = V(p)$ in terms of a scalar parameter p. Assume differentiability of V w.r.t. p and sensitivity

$$\Delta V = \frac{dV(p)}{dp} \cdot \Delta p.$$

Finite Difference Approximation

Finite difference approximation with step size h is

$$\frac{dV(p)}{dp} \approx \frac{V(p+h) - V(p)}{h} \approx \frac{V(p) - V(p-h)}{h} \quad \text{(one-sided), or}$$

$$\frac{dV(p)}{dp} \approx \frac{V(p+h) - V(p-h)}{2h} \quad \text{(two-sided).}$$

- Simple to implement and calculate; only pricing function evaluation.
- Typically used for black-box pricing functions.

We do a case study for European swaption Vega I

Recall pricing function

$$V^{\text{Swpt}} = Ann(t) \cdot \text{Bachelier}\left(S(t), K, \sigma\sqrt{T-t}, \phi\right)$$

with

$$\text{Bachelier}\,(F, K, \nu, \phi) = \nu \cdot \left[\Phi\,(h) \cdot h + \Phi'\,(h)\right], \quad h = \frac{\phi\,[F-K]}{\nu}.$$

First, analyse Bachelier formula. We get

$$\begin{aligned}
\frac{d}{d\nu}\text{Bachelier}\,(\nu) &= \frac{\text{Bachelier}\,(\nu)}{\nu} + \nu\left[\left(\Phi'\,(h)\,h + \Phi\,(h)\right)\frac{dh}{d\nu} - \Phi'\,(h)\,h\frac{dh}{d\nu}\right] \\
&= \frac{\text{Bachelier}\,(\nu)}{\nu} + \nu\Phi\,(h)\,\frac{dh}{d\nu}.
\end{aligned}$$

With $\frac{dh}{d\nu} = -\frac{h}{\nu}$ follows

$$\frac{d}{d\nu}\text{Bachelier}\,(\nu) = \Phi\,(h) \cdot h + \Phi'\,(h) - \Phi\,(h) \cdot h = \Phi'\,(h)\,.$$

We do a case study for European swaption Vega II

Moreover, second derivative (Volga) becomes

$$\frac{d^2}{d\nu^2}\text{Bachelier}\left(\nu\right) = -h\Phi'\left(h\right)\frac{dh}{d\nu} = \frac{h^2}{\nu}\Phi'\left(h\right).$$

This gives for ATM options with $h = 0$ that

- Volga $\frac{d^2}{d\nu^2}\text{Bachelier}\left(\nu\right) = 0$.
- ATM option price is approximately linear in volatility ν.

Differentiating once again yields (we skip details)

$$\frac{d^3}{d\nu^3}\text{Bachelier}\left(\nu\right) = \left(h^2 - 3\right)\frac{h^2}{\nu^2}\Phi'\left(h\right).$$

It turns out that Volga has a maximum at moneyness

$$h = \pm\sqrt{3}.$$

We do a case study for European swaption Vega III

Swaption Vega becomes

$$\frac{d}{d\sigma} V^{\mathrm{Swpt}} = An(t) \cdot \frac{d}{d\nu} \mathrm{Bachelier}\,(\nu) \cdot \sqrt{T - t}.$$

Test case

- Rates flat at 5%, implied normal volatilities flat at $100bp$.
- 10y into 10y European payer swaption (call on swap rate).
- Strike at $5\% + 100bp \cdot \sqrt{10y} \cdot \sqrt{3} = 10.48\%$ (maximizing Volga).

What is the problem with finite difference approximation? I

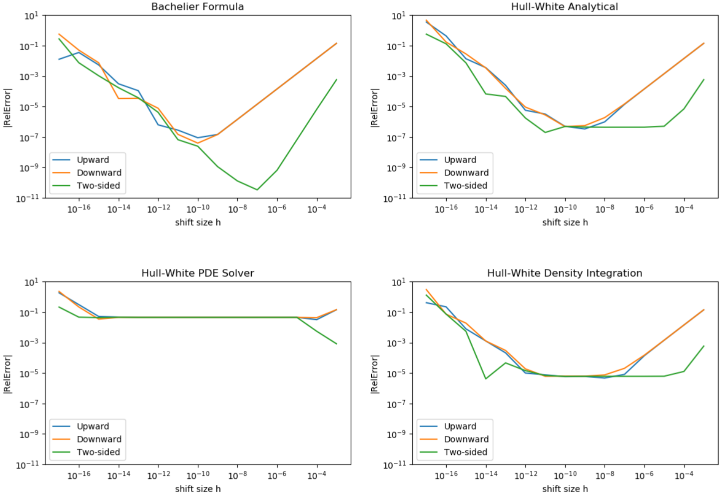
- There is a non-trivial trade-off between convergence and numerical accuracy.
- We have analytical Vega formula from Bachelier formula and implied normal volatility

$$\text{Vega} = An(t) \cdot \Phi'(h) \cdot \sqrt{T-t}.$$

- Compare one-sided (upward and downward) and two-sided finite difference approximation Vega_{FD} using
  - Bachelier formula,
  - Analytical Hull-White coupon bond option formula,
  - Hull-White model via PDE solver (Crank-Nicolson, 101 grid points, 3 stdDevs wide, 1m time stepping),
  - Hull-White model via density integration (C^2-spline exact with break-even point, 101 grid points, 5 stdDevs wide).
- Compare absolute relative error (for all finite difference approximations)

$$|\text{RelErr}| = \left| \frac{\text{Vega}_{FD}}{\text{Vega}} - 1 \right|$$

What is the problem with finite difference approximation? II



Optimal choice of FD step size h is very problem-specific and depends on discretisation of numerical method.

Outline

Derivative pricing usually involves model calibration (1/2)

Consider swap pricing function V^{Swap} as a function of yield curve model parameters z, i.e.

$$V^{\text{Swap}} = V^{\text{Swap}}(z).$$

Model parameters z are itself derived from market quotes R for par swaps and FRAs. That is

$$z = z(R).$$

This gives mapping

$$R \mapsto z \mapsto V^{\text{Swap}} = V^{\text{Swap}}\left(z(R)\right).$$

Interest rate Delta becomes

$$\Delta_R = 1bp \cdot \underbrace{\frac{dV^{\text{Swap}}}{dz}\left(z(R)\right)}_{\text{Pricing}} \cdot \underbrace{\frac{dz}{dR}\left(R\right)}_{\text{Calibration}}.$$

Derivative pricing usually involves model calibration (2/2)

$$\Delta_R = 1bp \cdot \underbrace{\frac{dV^{\text{Swap}}}{dz}(z(R))}_{\text{Pricing}} \cdot \underbrace{\frac{dz}{dR}(R)}_{\text{Calibration}}.$$

- Suppose a large portfolio of swaps:
  - Calibration Jacobian $\frac{dz(R)}{dR}$ is the same for all swaps in portfolio.
  - Save computational effort by pre-calculating and storing Jacobian.
- Brute-force finite difference approximation of Jacobian may become inaccurate due to numerical scheme for calibration/optimisation.

Can we calculate calibration Jacobian more efficiently?

Theorem (Implicit Function Theorem)

Let $\mathcal{H} : \mathbb{R}^q \times \mathbb{R}^r \to \mathbb{R}^q$ be a continuously differentiable function with $\mathcal{H}(\bar{z}, \bar{R}) = 0$ for some pair $(\bar{z}, \bar{R})$. If the Jacobian

$$J_z = \frac{d\mathcal{H}}{dz}(\bar{z}, \bar{R})$$

is invertible, then there exists an open domain $\mathcal{U} \subset \mathbb{R}^r$ with $\bar{R} \in \mathcal{U}$ and a continuously differentiable function $g : \mathcal{U} \to \mathbb{R}^q$ with

$$\mathcal{H}\left(g(R), R\right) = 0 \quad \forall R \in \mathcal{U}.$$

Moreover, we get for the Jacobian of g that

$$\frac{dg(R)}{dR} = -\left[\frac{d\mathcal{H}}{dz}(g(R), R)\right]^{-1}\left[\frac{d\mathcal{H}}{dR}(g(R), R)\right].$$

Proof.

See Analysis. □

How does Implicit Function Theorem help for sensitivity calculation? (1/4)

- Consider $\mathcal{H}(z, R)$ the q-dimensional objective function of yield curve calibration problem:
  - $z = [z_1, \ldots, z_q]^\top$ yield curve parameters (e.g. zero rates or forward rates),
  - $R = [R_1, \ldots, R_q]^\top$ market quotes (par rates) for swaps and FRAs,
  - use same number of market quotes as model parameters, i.e. $r = q$.

- Reformulate calibration helpers slightly such that

$$\mathcal{H}_k(z, R) = \text{ModelRate}_k(z) - R_k.$$

- For example, for swap rate helpers, model-implied par swap rate becomes

$$\text{ModelRate}_k(z) = \frac{\sum_{j=1}^{m_k} L^\delta(0, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta) \cdot \tilde{\tau}_j \cdot P(t, \tilde{T}_j)}{\sum_{i=1}^{n_k} \tau_i \cdot P(0, T_i)}.$$

How does Implicit Function Theorem help for sensitivity calculation? (2/4)

Suppose pair $(\bar{z}, \bar{R})$ solves calibration problem $\mathcal{H}(\bar{z}, \bar{R}) = 0$ and $\frac{d\mathcal{H}}{dz}(\bar{z}, \bar{R})$ is invertible.
Then, by Implicit Function Theorem, there exists a function

$$z = z(R)$$

in a vicinity of $\bar{R}$ and

$$\frac{dz}{dR}(R) = -\left[\frac{d\mathcal{H}}{dz}(g(R), R)\right]^{-1}\left[\frac{d\mathcal{H}}{dR}(g(R), R)\right].$$

How does Implicit Function Theorem help for sensitivity calculation? (3/4)

$$\frac{dz}{dR}(R) = -\left[\frac{d\mathcal{H}}{dz}(g(R),R)\right]^{-1}\left[\frac{d\mathcal{H}}{dR}(g(R),R)\right].$$

From reformulated calibration helpers we get

$$\frac{d\mathcal{H}}{dz}(g(R),R) = \begin{bmatrix} \frac{d}{dz}\text{ModelRate}_1(z) \\ \vdots \\ \frac{d}{dz}\text{ModelRate}_q(z) \end{bmatrix}, \quad \text{and}$$

$$\frac{d\mathcal{H}}{dR}(g(R),R) = \begin{bmatrix} -1 & & \\ & \ddots & \\ & & -1 \end{bmatrix}.$$

Consequently

$$\frac{dz}{dR}(R) = \left[\frac{d\mathcal{H}}{dz}(g(R),R)\right]^{-1} = \begin{bmatrix} \frac{d}{dz}\text{ModelRate}_1(z) \\ \vdots \\ \frac{d}{dz}\text{ModelRate}_q(z) \end{bmatrix}^{-1}.$$

How does Implicit Function Theorem help for sensitivity calculation? (4/4)

We get Jacobian method for risk calculation

$$\Delta_R = 1bp \cdot \underbrace{\frac{dV^{\text{Swap}}}{dz}(z(R))}_{\text{Pricing}} \cdot \underbrace{\begin{bmatrix} \frac{d}{dz}\text{ModelRate}_1(z) \\ \vdots \\ \frac{d}{dz}\text{ModelRate}_q(z) \end{bmatrix}^{-1}}_{\text{Calibration}}.$$

- Requires only sensitivities w.r.t. model parameters.
- Reference market intruments/rates R_k can also be chosen independent of original calibration problem.
- Calibration Jacobian and matrix inversion can be pre-computed and stored.

We can also adapt Jacobian method to Vega calculation (1/3)

Bermudan swaption is determined via mapping

$$\underbrace{\left[\sigma_N^1, \ldots \sigma_N^{\bar{k}}\right]}_{\text{market-impl. normal vols}} \quad \mapsto \quad \underbrace{\left[\sigma^1, \ldots \sigma^{\bar{k}}\right]}_{\text{HW short rate vols}} \mapsto V^{\text{Berm}}.$$

Assign volatility calibration helpers

$$\mathcal{H}_k\left(\sigma, \sigma_N\right) = \underbrace{V_k^{\text{CBO}}(\sigma)}_{\text{Model}[\sigma]} - \underbrace{V_k^{\text{Swpt}}(\sigma_N^k)}_{\text{Market}\left(\sigma_N^k\right)}.$$

- $V_k^{\text{CBO}}(\sigma)$ Hull-White model price of kth co-terminal European swaption represented as coupon bond option.
- $V_k^{\text{Swpt}}(\sigma_N^k)$ Bachelier formula to calculate market price for kth co-terminal European swaption from given normal volatility σ_N^k.

We can also adapt Jacobian method to Vega calculation (2/3)

Implicit Function Theorem yields

$$\frac{d\sigma}{d\sigma_N} = -\left[\frac{d\mathcal{H}}{d\sigma}\left(\sigma\left(\sigma_N\right),\sigma_N\right)\right]^{-1}\left[\frac{d\mathcal{H}}{d\sigma_N}\left(\sigma\left(\sigma_N\right),\sigma_N\right)\right]$$
$$= \left[\frac{d}{d\sigma}\mathrm{Model}[\sigma]\right]^{-1}\begin{bmatrix} \frac{d}{d\sigma_N}V_1^{\mathrm{Swpt}}(\sigma_N^1) & & \\ & \ddots & \\ & & \frac{d}{d\sigma_N}V_{\bar{k}}^{\mathrm{Swpt}}(\sigma_N^{\bar{k}}) \end{bmatrix}.$$

- $\frac{d}{d\sigma}\mathrm{Model}[\sigma]$ are Hull-White model Vega(s) of co-terminal European swaptions.
- $\frac{d}{d\sigma_N}V_k^{\mathrm{Swpt}}(\sigma_N^k)$ are Bachelier or market Vega(s) of co-terminal European swaptions.

We can also adapt Jacobian method to Vega calculation (3/3)

Bermudan Vega becomes

$$\frac{d}{d\sigma_N} V^{\mathrm{Berm}} = \frac{d}{d\sigma} V^{\mathrm{Berm}} \cdot \left[\frac{d}{d\sigma} \mathrm{Model}[\sigma] \right]^{-1} \cdot \frac{d}{d\sigma_N} \mathrm{Market}\left(\sigma_N^k\right).$$

Outline

What is the idea behind Algorithmic Differentiation (AD)

- AD covers principles and techniques to augment computer models or programs.
- Calculate sensitivities of output variables with respect to inputs of a model.
- Compute numerical values rather than symbolic expressions.
- Sensitivities are exact up to machine precision (no rounding/cancellation errors as in FD).
- Apply chain rule of differentiation to operations like +, $*$, and intrinsic functions like `exp(.)`.

Functions are represented as Evaluation Procedures consisting of a sequence of elementary operations

Example: Black Formula

$$\text{Black}(\cdot) = \omega \left[F\Phi(\omega d_1) - K\Phi(\omega d_2) \right]$$

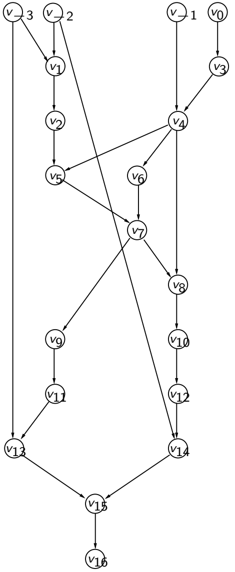
with $d_{1,2} = \frac{\log(F/K)}{\sigma\sqrt{\tau}} \pm \frac{\sigma\sqrt{\tau}}{2}$

- Inputs F, K, σ, τ
- Discrete parameter $\omega \in \{-1, 1\}$
- Output Black$(\cdot)$

$$
\begin{array}{lcll}
v_{-3} & = & x_1 = F & \\
v_{-2} & = & x_2 = K & \\
v_{-1} & = & x_3 = \sigma & \\
v_0 & = & x_4 = \tau & \\
\hline
v_1 & = & v_{-3}/v_{-2} & \equiv \; f_1(v_{-3}, v_{-2}) \\
v_2 & = & \log(v_1) & \equiv \; f_2(v_1) \\
v_3 & = & \sqrt{v_0} & \equiv \; f_3(v_0) \\
v_4 & = & v_{-1} \cdot v_3 & \equiv \; f_4(v_{-1}, v_3) \\
v_5 & = & v_2/v_4 & \equiv \; f_5(v_2, v_4) \\
v_6 & = & 0.5 \cdot v_4 & \equiv \; f_6(v_4) \\
v_7 & = & v_5 + v_6 & \equiv \; f_7(v_5, v_6) \\
v_8 & = & v_7 - v_4 & \equiv \; f_8(v_7, v_4) \\
v_9 & = & \omega \cdot v_7 & \equiv \; f_9(v_7) \\
v_{10} & = & \omega \cdot v_8 & \equiv \; f_{10}(v_8) \\
v_{11} & = & \Phi(v_9) & \equiv \; f_{11}(v_9) \\
v_{12} & = & \Phi(v_{10}) & \equiv \; f_{12}(v_{10}) \\
v_{13} & = & v_{-3} \cdot v_{11} & \equiv \; f_{13}(v_{-3}, v_{11}) \\
v_{14} & = & v_{-2} \cdot v_{12} & \equiv \; f_{14}(v_{-2}, v_{12}) \\
v_{15} & = & v_{13} - v_{14} & \equiv \; f_{15}(v_{13}, v_{14}) \\
v_{16} & = & \omega \cdot v_{15} & \equiv \; f_{16}(v_{15}) \\
\hline
y_1 & = & v_{16} &
\end{array}
$$

Alternative representation is Directed Acyclic Graph (DAG)

$$
\begin{array}{lllll}
v_{-3} & = & x_1 = F & & \\
v_{-2} & = & x_2 = K & & \\
v_{-1} & = & x_3 = \sigma & & \\
v_0 & = & x_4 = \tau & & \\
\hline
v_1 & = & v_{-3}/v_{-2} & \equiv & f_1(v_{-3}, v_{-2}) \\
v_2 & = & \log(v_1) & \equiv & f_2(v_1) \\
v_3 & = & \sqrt{v_0} & \equiv & f_3(v_0) \\
v_4 & = & v_{-1} \cdot v_3 & \equiv & f_4(v_{-1}, v_3) \\
v_5 & = & v_2/v_4 & \equiv & f_5(v_2, v_4) \\
v_6 & = & 0.5 \cdot v_4 & \equiv & f_6(v_4) \\
v_7 & = & v_5 + v_6 & \equiv & f_7(v_5, v_6) \\
v_8 & = & v_7 - v_4 & \equiv & f_8(v_7, v_4) \\
v_9 & = & \omega \cdot v_7 & \equiv & f_9(v_7) \\
v_{10} & = & \omega \cdot v_8 & \equiv & f_{10}(v_8) \\
v_{11} & = & \Phi(v_9) & \equiv & f_{11}(v_9) \\
v_{12} & = & \Phi(v_{10}) & \equiv & f_{12}(v_{10}) \\
v_{13} & = & v_{-3} \cdot v_{11} & \equiv & f_{13}(v_{-3}, v_{11}) \\
v_{14} & = & v_{-2} \cdot v_{12} & \equiv & f_{14}(v_{-2}, v_{12}) \\
v_{15} & = & v_{13} - v_{14} & \equiv & f_{15}(v_{13}, v_{14}) \\
v_{16} & = & \omega \cdot v_{15} & \equiv & f_{16}(v_{15}) \\
\hline
y_1 & = & v_{16} & &
\end{array}
$$



Evaluation Procedure can be formalized to make it more tractable

Definition (Evaluation Procedure)

Suppose $F : \mathbb{R}^n \to \mathbb{R}^m$ and $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^{m_i}$. The relation $j \prec i$ denotes that $v_i \in \mathbb{R}$ depends directly on $v_j \in \mathbb{R}$. If for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ with $y = F(x)$ holds that

$$\begin{array}{rcll} v_{i-n} & = & x_i & i = 1, \dots, n \\ v_i & = & f_i(v_j)_{j \prec i} & i = 1, \dots, l \\ y_{m-i} & = & v_{l-i} & i = m-1, \dots, 0, \end{array}$$

then we call this sequence of operations an evaluation procedure of F with elementary operations f_i. We assume differentiability of all elementary operations f_i $(i = 1, \dots, l)$. Then the resulting function F is also differentiable.

- Abbreviate $u_i = (v_j)_{j \prec i} \in \mathbb{R}^{n_i}$ the collection of arguments of the operation f_i.
- Then we may also write

$$v_i = f_i(u_i).$$

Forward mode of AD calculates tangents (1/2)

- In addition to function evaluation $v_i = f_i(u_i)$ evaluate derivative

$$\dot{v}_i \;=\; \sum_{j \prec i} \frac{\partial}{\partial v_j} f_i(u_i) \cdot \dot{v}_j .$$

Forward Mode or Tangent Mode of AD

Use abbreviations $\dot{u}_i = (\dot{v}_j)_{j \prec i}$ and $\dot{f}_i(u_i, \dot{u}_i) = f_i'(u_i) \cdot \dot{u}_i$. The Forward Mode of AD is the augmented evaluation procedure

$$\begin{array}{rcll} [v_{i-n}, \dot{v}_{i-n}] & = & [x_i, \dot{x}_i] & i = 1, \ldots, n \\ [v_i, \dot{v}_i] & = & \left[f_i(u_i), \dot{f}_i(u_i, \dot{u}_i)\right] & i = 1, \ldots, l \\ [y_{m-i}, \dot{y}_{m-i}] & = & [v_{l-i}, \dot{v}_{l-i}] & i = m-1, \ldots, 0. \end{array}$$

Here, the initializing derivative values $\dot{x}_{i-n}$ for $i = 1 \ldots n$ are given and determine the direction of the tangent.

Forward mode of AD calculates tangents (2/2)

- With $\dot{x} = (\dot{x}_i) \in \mathbb{R}^n$ and $\dot{y} = (\dot{y}_i) \in \mathbb{R}^m$, the forward mode of AD evaluates
$$\dot{y} \;=\; F'(x)\dot{x}.$$

- Computational effort is approx. 2.5 function evaluations of F.

Black formula Forward Mode evaluation procedure...

$$
\begin{array}{lcl|lcl}
v_{-3} & = & x_1 = F & \dot{v}_{-3} & = & 0 \\
v_{-2} & = & x_2 = K & \dot{v}_{-2} & = & 0 \\
v_{-1} & = & x_3 = \sigma & \dot{v}_{-1} & = & 1 \\
v_0 & = & x_4 = \tau & \dot{v}_0 & = & 0 \\
\hline
v_1 & = & v_{-3}/v_{-2} & \dot{v}_1 & = & \dot{v}_{-3}/v_{-2} - v_1 \cdot \dot{v}_{-2}/v_{-2} \\
v_2 & = & \log(v_1) & \dot{v}_2 & = & \dot{v}_1/v_1 \\
v_3 & = & \sqrt{v_0} & \dot{v}_3 & = & 0.5 \cdot \dot{v}_0/v_3 \\
v_4 & = & v_{-1} \cdot v_3 & \dot{v}_4 & = & \dot{v}_{-1} \cdot v_3 + v_{-1} \cdot \dot{v}_3 \\
v_5 & = & v_2/v_4 & \dot{v}_5 & = & \dot{v}_2/v_4 - v_5 \cdot \dot{v}_4/v_4 \\
v_6 & = & 0.5 \cdot v_4 & \dot{v}_6 & = & 0.5 \cdot \dot{v}_4 \\
v_7 & = & v_5 + v_6 & \dot{v}_7 & = & \dot{v}_5 + \dot{v}_6 \\
v_8 & = & v_7 - v_4 & \dot{v}_8 & = & \dot{v}_7 - \dot{v}_4 \\
v_9 & = & \omega \cdot v_7 & \dot{v}_9 & = & \omega \cdot \dot{v}_7 \\
v_{10} & = & \omega \cdot v_8 & \dot{v}_{10} & = & \omega \cdot \dot{v}_8 \\
v_{11} & = & \Phi(v_9) & \dot{v}_{11} & = & \phi(v_9) \cdot \dot{v}_9 \\
v_{12} & = & \Phi(v_{10}) & \dot{v}_{12} & = & \phi(v_{10}) \cdot \dot{v}_{10} \\
v_{13} & = & v_{-3} \cdot v_{11} & \dot{v}_{13} & = & \dot{v}_{-3} \cdot v_{11} + v_{-3} \cdot \dot{v}_{11} \\
v_{14} & = & v_{-2} \cdot v_{12} & \dot{v}_{14} & = & \dot{v}_{-2} \cdot v_{12} + v_{-2} \cdot \dot{v}_{12} \\
v_{15} & = & v_{13} - v_{14} & \dot{v}_{15} & = & \dot{v}_{13} - \dot{v}_{14} \\
v_{16} & = & \omega \cdot v_{15} & \dot{v}_{16} & = & \omega \cdot \dot{v}_{15} \\
\hline
y_1 & = & v_{16} & \dot{y}_1 & = & \dot{v}_{16}
\end{array}
$$

Reverse Mode of AD calculates adjoints (1/3)

- Forward Mode calculates derivatives and applies chain rule in the same order as function evaluation.
- Reverse Mode of AD applies chain rule in reverse order of function evaluation.
- Define auxiliary derivative values $\bar{v}_j$ and assume initialisation $\bar{v}_j = 0$ before reverse mode evaluation.
- For each elementary operation f_i and all intermediate variables v_j with $j \prec i$, evaluate

$$\bar{v}_j \mathrel{+}= \bar{v}_i \cdot \frac{\partial}{\partial v_j} f_i(u_i).$$

- In other words, for each arguments of f_i the partial derivative is derived.

Reverse Mode of AD calculates adjoints (2/3)

Reverse Mode or Adjoint Mode of AD

Denoting $\bar{u}_i = (\bar{v}_j)_{j \prec i} \in \mathbb{R}^{n_i}$ and $\bar{f}_i(u_i, \bar{v}_i) = \bar{v}_i \cdot f_i'(u_i)$, the *incremental reverse mode of AD* is given by the evaluation procedure

$$
\begin{array}{rcll}
v_{i-n} & = & x_i & i = 1, \ldots, n \\
v_i & = & f_i(v_j)_{j \prec i} & i = 1, \ldots, l \\
y_{m-i} & = & v_{l-i} & i = m-1, \ldots, 0 \\
\hline
\bar{v}_i & = & \bar{y}_i & i = 0, \ldots, m-1 \\
\bar{u}_i & += & \bar{f}_i(u_i, \bar{v}_i) & i = l, \ldots, 1 \\
\bar{x}_i & = & \bar{v}_i & i = n, \ldots, 1.
\end{array}
$$

Here, all intermediate variables v_i are assigned only once. The initializing values $\bar{y}_i$ are given and represent a weighting of the dependent variables y_i.

Reverse Mode of AD calculates adjoints (3/3)

- Vector $\bar{y} = (\bar{y}_i)$ can also be interpreted as normal vector of a hyperplane in the range of F.
- With $\bar{y} = (\bar{y}_i)$ and $\bar{x} = (\bar{x}_i)$, reverse mode of AD yields

$$\bar{x}^T \; = \; \nabla \left[\bar{y}^T F(x) \right] \; = \; \bar{y}^T F'(x).$$

- Computational effort is approx. 4 function evaluations of F.

Black formula Reverse Mode evaluation procedure ... I

$$
\begin{array}{rcl}
v_{-3} & = & x_1 = F \\
v_{-2} & = & x_2 = K \\
v_{-1} & = & x_3 = \sigma \\
v_0 & = & x_4 = \tau \\
\hline
v_1 & = & v_{-3}/v_{-2} \\
v_2 & = & \log(v_1) \\
v_3 & = & \sqrt{v_0} \\
v_4 & = & v_{-1} \cdot v_3 \\
v_5 & = & v_2/v_4 \\
v_6 & = & 0.5 \cdot v_4 \\
v_7 & = & v_5 + v_6 \\
v_8 & = & v_7 - v_4 \\
v_9 & = & \omega \cdot v_7 \\
v_{10} & = & \omega \cdot v_8 \\
v_{11} & = & \Phi(v_9) \\
v_{12} & = & \Phi(v_{10}) \\
v_{13} & = & v_{-3} \cdot v_{11} \\
v_{14} & = & v_{-2} \cdot v_{12} \\
v_{15} & = & v_{13} - v_{14} \\
v_{16} & = & \omega \cdot v_{15} \\
\hline
y_1 & = & v_{16} \\
\bar{v}_{16} & = & \bar{y}_1 = 1 \\
\hline
& \vdots &
\end{array}
$$

Black formula Reverse Mode evaluation procedure ... II

$$\vdots$$

$$y_1 = v_{16}$$
$$\bar{v}_{16} = \bar{y}_1 = 1$$

$$\bar{v}_{15} \mathrel{+}= \omega \cdot \bar{v}_{16}$$
$$\bar{v}_{13} \mathrel{+}= \bar{v}_{15}; \quad \bar{v}_{14} \mathrel{+}= (-1) \cdot \bar{v}_{15}$$
$$\bar{v}_{-2} \mathrel{+}= v_{12} \cdot \bar{v}_{14}; \quad \bar{v}_{12} \mathrel{+}= v_{-2} \cdot \bar{v}_{14}$$
$$\bar{v}_{-3} \mathrel{+}= v_{11} \cdot \bar{v}_{13}: \quad \bar{v}_{11} \mathrel{+}= v_{-3} \cdot \bar{v}_{13}$$
$$\bar{v}_{10} \mathrel{+}= \phi(v_{10}) \cdot \bar{v}_{12}$$
$$\bar{v}_{9} \mathrel{+}= \phi(v_{9}) \cdot \bar{v}_{11}$$
$$\bar{v}_{8} \mathrel{+}= \omega \cdot \bar{v}_{10}$$
$$\bar{v}_{7} \mathrel{+}= \omega \cdot \bar{v}_{9}$$
$$\bar{v}_{7} \mathrel{+}= \bar{v}_{8}; \quad \bar{v}_{4} \mathrel{+}= (-1) \cdot \bar{v}_{8}$$
$$\bar{v}_{5} \mathrel{+}= \bar{v}_{7}; \quad \bar{v}_{6} \mathrel{+}= \bar{v}_{7}$$
$$\bar{v}_{4} \mathrel{+}= 0.5 \cdot \bar{v}_{6}$$
$$\bar{v}_{2} \mathrel{+}= \bar{v}_{5}/v_4; \quad \bar{v}_{4} \mathrel{+}= (-1) \cdot v_5 \cdot \bar{v}_{5}/v_4$$
$$\bar{v}_{-1} \mathrel{+}= v_3 \cdot \bar{v}_{4}; \quad \bar{v}_{3} \mathrel{+}= v_{-1} \cdot \bar{v}_{4}$$
$$\bar{v}_{0} \mathrel{+}= 0.5 \cdot \bar{v}_{3}/v_3$$
$$\bar{v}_{1} \mathrel{+}= \bar{v}_{2}/v_1$$
$$\bar{v}_{-3} \mathrel{+}= \bar{v}_{1}/v_{-2}; \quad \bar{v}_{-2} \mathrel{+}= (-1) \cdot v_1 \cdot \bar{v}_{1}/v_{-2}$$

$$\bar{\tau} = \bar{x}_4 = \bar{v}_0$$
$$\bar{\sigma} = \bar{x}_3 = \bar{v}_{-1}$$
$$\bar{K} = \bar{x}_2 = \bar{v}_{-2}$$
$$\bar{F} = \bar{x}_1 = \bar{v}_{-3}$$

We summarise the properties of Forward and Reverse Mode

Forward Mode

$$\dot{y} = F'(x)\dot{x}$$

- Approx. 2.5 function evaluations.
- Computational effort independent of number of output variables (dimension of y).
- Chain rule in same order as computation.
- Memory consumption in order of function evaluation.

Reverse Mode

$$\bar{x}^T = \bar{y}^T F'(x)$$

- Approx. 4 function evaluations.
- Computational effort independent of number of input variables (dimension of x).
- Chain rule in reverse order of computation.
- Requires storage of all intermediate results (or re-computation).
- Memory consumption/management key challange for implementations.

- Computational effort can be improved by AD vector mode.
- Reverse Mode memory consumption can be managed via checkpointing techniques.

How is AD applied in practice?

- Typically, you don't want to differentiate all your source code by hand.
- Tools help augmenting existing programs for tangent and adjoint computations.

Source Code Transformation

- Applied to the model code in compiler fashion.
- Generate AD model as new source code.
- Original code may need to be adapted slightly to meet capabilities of AD tool.

Some example C++ tools:

ADIC2, dcc, TAPENADE

Operator Overloading

- provide new (active) data type.
- Overload all relevant operators/ functions with sensitivity aware arithmetic.
- AD model derived by changing intrinsic to active data type.

ADOL-C, dco/c++, ADMB/AUTODIF

- There are also tools for Python and other lamguages:

More details at `autodiff.org`

There is quite some literature on AD and its application in finance

Standard textbook on AD:

- A. Griewank and A. Walther. *Evaluating derivatives: principles and techniques of algorithmic differentiation - 2nd ed.*
  SIAM, 2008

Recent practitioner's textbook:

- U. Naumann. *The Art of Differentiating Computer Programs: An Introduction to Algorithmic Differentiation.*
  SIAM, 2012

One of the first and influencial papers for AD application in finance:

- M. Giles and P. Glasserman. Smoking adjoints: fast monte carlo greeks.
  Risk, January 2006

Part VIII

Wrap-up

Outline

What was this lecture about?

Interbank swap deal example

Trade details (fixed rate, notional, etc.)

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

Date calculations

End date: Oct 30, 2040

Market conventions

(annually, 30/360 day count, modified following, Target calendar)

Stochastic interest rates

Pays 6-months Euribor floating rate on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(semi-annually, act/360 day count, modified following, Target calendar)

Optionalities

Bank A may decide to early terminate deal in 10, 11, 12,.. years

Part IX

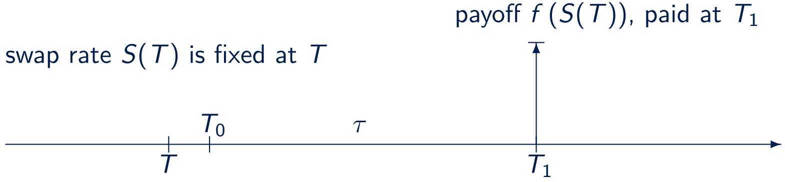
Other Topics

Outline

Outline

We analyse the pricing of more general single-rate payoffs

What is the present value of the complex payoff $f\left(S(T)\right)$?



rate is accrued from T_0 to T_1

Pricing in T_1-forward measure yields

$$V(t) = P(t, T_1) \cdot \mathbb{E}^{T_1}\left[f\left(S(T)\right) \mid \mathcal{F}_t\right].$$

- In general, $S(t)$ is not a martingale in T_1-forward measure.
- Terminal distribution of $S(T)$ in Vanilla model is specified in annuity measure.

We need to change the pricing measure to utilize Vanilla model dynamics

Pricing in annuity measure becomes

$$V(t) = P(t, T_1) \cdot \mathbb{E}^A \left[\frac{An(t)}{P(t, T_1)} \frac{P(T, T_1)}{An(T)} \cdot f\left(S(T)\right) \mid \mathcal{F}_t \right].$$

- We need to properly handle the Radon–Nikodym derivative (from T_1-forward to annuity measure)

$$\frac{An(t)}{P(t, T_1)} \frac{P(T, T_1)}{An(T)}.$$

- Take out what is known and apply tower rule of iterated expectation

$$V(t) = An(t) \cdot \mathbb{E}^A \left[\mathbb{E}^A \left[\frac{P(T, T_1)}{An(T)} \mid S(T) = s \right] \cdot f\left(S(T)\right) \mid \mathcal{F}_t \right].$$

Key challenge is modelling conditional expectation $\mathbb{E}^A \left[\frac{P(T, T_1)}{An(T)} \mid S(T) = s \right]$.

Outline

Terminal swap rate models are characterised by an annuity mappig function

Annuity Mapping Function

Consider a swap rate $S(T)$ with rate fixing at T and corresponding annuity measure. For pay times $T_p \geq T$ the annuity mapping function is defined as

$$\alpha(s, T_p) = \mathbb{E}^A \left[\frac{P(T, T_p)}{An(T)} \,|\, S(T) = s \right].$$

With annuity mapping function at hand we can calculate

$$\begin{aligned} V(t) &= An(t) \cdot \mathbb{E}^A \left[\alpha \left(S(T), T_1 \right) \cdot f \left(S(T) \right) \,|\, \mathcal{F}_t \right] \\ &= An(t) \cdot \int_{-\infty}^{\infty} \alpha \left(s, T_1 \right) \cdot f(s) \cdot d\mathbb{P}^A(s). \end{aligned}$$

Once annuity mapping function is known, we can integrate against terminal distribution $d\mathbb{P}^A(s)$ from Vanilla model.

Annuity mapping function needs to comply with model-independent properties (1/3)

No-arbitrage Condition

For all $T_p \geq T$

$$\mathbb{E}^A\left[\alpha\left(S(T), T_p\right) \mid \mathcal{F}_t\right] = \mathbb{E}^A\left[\mathbb{E}^A\left[\frac{P(T, T_p)}{An(T)} \mid S(T) = s\right] \mid \mathcal{F}_t\right] = \frac{P(t, T_p)}{An(t)}.$$

- No-arbitrage condition is closely linked to martingale property related to Radon–Nikodym derivative

$$\frac{An(t)}{P(t, T_1)} \frac{P(T, T_1)}{An(T)}.$$

- Specifies level of $\alpha(s, T_p)$ in s-direction.

Annuity mapping function needs to comply with model-independent properties (2/3)

Additivity Condition

Consider annuity of $S(T)$ given by $An(T) = \sum_{i=0}^{n} \tau_i P(T, T_i)$ then for all s

$$\sum_{i=0}^{n} \tau_i \cdot \alpha\left(s, T_i\right) = \mathbb{E}^A\left[\sum_{i=0}^{n} \tau_i \frac{P(T, T_i)}{An(T)} \mid S(T) = s\right] = 1.$$

- Additivity condition specifies overall level of $\alpha(s, T_p)$ in T_p-direction.

Annuity mapping function needs to comply with model-independent properties (3/3)

Consistency Condition

Consider swap rate representation

$$S(T) = \underbrace{\frac{\sum_j L_j(T)\tilde{\tau}_j P(T, \tilde{T}_j)}{An(T)}}_{\text{single-curve swap rate}} + \underbrace{\frac{\sum_j \left[D_j^{\delta} - 1\right] \tilde{\tau}_j P(T, \tilde{T}_{j-1})}{An(T)}}_{\text{basis spread}}$$

$$= \frac{P(T, T_0) - P(T, T_N)}{An(T)} + \frac{\sum_j \omega_j \cdot P(T, \tilde{T}_{j-1})}{An(T)}.$$

For all s we get

$$\alpha(s, T_0) - \alpha(s, T_N) + \sum_j \omega_j \cdot \alpha(s, \tilde{T}_{j-1}) = s.$$

- Note that typically $\omega_j \ll 1$, dominating term is $\alpha(s, T_0) - \alpha(s, T_N)$.
- Consistency condition specifies slope of $\alpha(s, T_p)$ in T_p-direction (relative to realisation of swap rate $S(T)$).

T-forward measure yields a very useful alternative representation of the annuity mapping function (1/3)

Theorem

In the T-forward measure the annuity mapping function becomes

$$\alpha(s, T_p) = \frac{\mathbb{E}^T\left[P(T, T_p) \,|\, S(T) = s\right]}{\mathbb{E}^T\left[An(T) \,|\, S(T) = s\right]}.$$

T-forward measure yields a very useful alternative representation of the annuity mapping function (2/3)

Proof.

Consider Radon–Nikodym derivative from annuity measure to T-forward measure $R(\omega) = \frac{P(0,T)}{An(0)} \frac{An(T)}{P(T,T)}$.

Applying Baye's rule for conditional expectation yields

$$\mathbb{E}^A \left[\frac{P(T, T_p)}{An(T)} \,|\, S(T) = s \right] = \frac{\mathbb{E}^T \left[R \frac{P(T,T_p)}{An(T)} \,|\, S(T) = s \right]}{\mathbb{E}^T \left[R \,|\, S(T) = s \right]}$$

$$= \frac{\frac{P(0,T)}{An(0)} \mathbb{E}^T \left[P(T, T_p) \,|\, S(T) = s \right]}{\frac{P(0,T)}{An(0)} \mathbb{E}^T \left[An(T) \,|\, S(T) = s \right]}$$

$$= \frac{\mathbb{E}^T \left[P(T, T_p) \,|\, S(T) = s \right]}{\mathbb{E}^T \left[An(T) \,|\, S(T) = s \right]}.$$

$\square$

T-forward measure yields a very useful alternative representation of the annuity mapping function (3/3)

Corollary

Define the conditional zero coupon bond (for $T' \geq T$) via

$$\pi(s, T') = \mathbb{E}^T \left[P(T, T') \,|\, S(T) = s \right].$$

Then the annuity mapping function becomes

$$\alpha(s, T_p) = \frac{\pi(s, T_p)}{\sum_{i=0}^{n} \tau_i \cdot \pi(s, T_i)}.$$

Proof.

Follows directly from above theorem, definition of annuity $An(T)$ and linearity of expectation. $\square$

Annuity mapping function is fully specified by conditional expectation of future zero coupon bonds.

Reformulating TSR pricing ensures consistency to initial yield curve for arbitrary annuity mapping functions (1/3)

Using tower rule we can re-write

$$
\begin{aligned}
V(t) &= An(t) \cdot \mathbb{E}^A \left[\mathbb{E}^A \left[\frac{P(T, T_1)}{An(T)} \,|\, S(T) = s \right] \cdot f\left(S(T)\right) \,|\, \mathcal{F}_t \right] \\
&= P(T, T_1) \cdot \frac{\mathbb{E}^A \left[\mathbb{E}^A \left[\frac{P(T, T_1)}{An(T)} \,|\, S(T) = s \right] \cdot f\left(S(T)\right) \,|\, \mathcal{F}_t \right]}{\mathbb{E}^A \left[\mathbb{E}^A \left[\frac{P(T, T_1)}{An(T)} \,|\, S(T) = s \right] \,|\, \mathcal{F}_t \right]} \\
&= P(T, T_1) \cdot \frac{\mathbb{E}^A \left[\alpha(s, T_1) \cdot f\left(S(T)\right) \,|\, \mathcal{F}_t \right]}{\mathbb{E}^A \left[\alpha(s, T_1) \,|\, \mathcal{F}_t \right]}.
\end{aligned}
$$

Reformulating TSR pricing ensures consistency to initial yield curve for arbitrary annuity mapping functions (2/3)

Yield curve reconstruction property

For any approximate annuity mapping function $\tilde{\alpha}(s, T_p) \approx \alpha(s, T_p)$ and any approximating expectation operator $\tilde{E} \approx \mathbb{E}^A$ (with $\tilde{E}\left[\tilde{\alpha}(s, T_p)\right] > 0$) we get that the (approximate) present value of a payoff $V(T_p) = 1$ becomes

$$V(t) = P(T, T_p) \cdot \frac{\tilde{E}\left[\tilde{\alpha}(s, T_p) \cdot V(T_p)\right]}{\tilde{E}\left[\tilde{\alpha}(s, T_p)\right]} = P(T, T_p).$$

Reformulating TSR pricing ensures consistency to initial yield curve for arbitrary annuity mapping functions (3/3)

Correcting non-arbitrage-free annuity mapping functions

We can re-write

$$
\begin{aligned}
V(t) &= P(T, T_1) \cdot \frac{\mathbb{E}^A \left[\alpha(s, T_1) \cdot f\left(S(T)\right) \,|\, \mathcal{F}_t \right]}{\mathbb{E}^A \left[\alpha(s, T_1) \,|\, \mathcal{F}_t \right]} \\
&= An(t) \cdot \mathbb{E}^A \left[\underbrace{\frac{P(t, T_1)}{An(t)} \frac{\alpha(s, T_1)}{\mathbb{E}^A \left[\alpha(s, T_1) \,|\, \mathcal{F}_t \right]}}_{\bar{\alpha}(s, T_1)} \cdot f\left(S(T)\right) \,|\, \mathcal{F}_t \right].
\end{aligned}
$$

Then, by construction, for any $\alpha(s, T_1)$

$$
\mathbb{E}^A \left[\bar{\alpha}(s, T_1) \,|\, \mathcal{F}_t \right] = \frac{P(t, T_1)}{An(t)}.
$$

For details on this aspect, see also [2], Sec. 16.6.7.

How can we actually specify annuity mapping function?

(a) Use a term structure model:

- Term structure model gives representation of future zero bonds $P(T, T')$.
- Calculate from model dynamics

$$\alpha(s, T_p) = \frac{\pi(s, T_p)}{\sum_{i=0}^{n} \tau_i \cdot \pi(s, T_i)}.$$

(b) Postulate a parametric form:

- Assume a parametric form for $\pi(s, T')$ (possibly inspired by term structure model).
- Alternatively, directly assume a parametric form of $\alpha(s, T_p)$ in terms of s and T_p.
- Calibrate parametric form(s) to model-independent properties.

Outline

We analyse Hull-White model for annuity mapping function (1/3)

Recall zero coupon bond formula

$$P(x; T, T') = \frac{P(0, T')}{P(0, T)} \exp \left\{ -G(T, T')x - \frac{G(T, T')^2}{2} y(T) \right\}.$$

Function $G(T, T')$ is specified by mean reversion

$$G(T, T') = \left[1 - e^{-a(T' - T)} \right] / a.$$

Auxilliary variable $y(T)$ represents (deterministic) variance

$$y(T) = \int_0^T \left[e^{-a(T-u)} \sigma(u) \right]^2 du.$$

We analyse Hull-White model for annuity mapping function (2/3)

For now, assume mean reversion a and volatility $\sigma(t)$ are given. Condition $S(T) = s$ is equivalent to

$$F(s,x) = \frac{P(x;T,T_0) - P(x;T,T_N)}{\sum_{i=0}^{n} \tau_i \cdot P(x;T,T_i)} + \frac{\sum_j \omega_j \cdot P(x;T,\tilde{T}_{j-1})}{\sum_{i=0}^{n} \tau_i \cdot P(x;T,T_i)} - s = 0.$$

- Obviously there is some $(\bar{x}, \bar{s})$ with $F(\bar{x}, \bar{s}) = 0$ (any x directly implies an s which solves eqation).
- Assume $\frac{\partial F}{\partial x}(s,x) > 0$ for all x.
  - Usually no restriction since $\frac{d}{dx}P(x;T,T_N) = -G(T,T_N)P(x;T,T_N) < 0$ dominates.

We analyse Hull-White model for annuity mapping function (3/3)

Implicit function theorem implies a continuous differentiable function $g(s)$ such that

$$F(s, g(s)) = 0, \quad \text{i.e.,} \quad x = g(s).$$

Thus $x(T) = g\left(S(T)\right)$ which gives

$$\begin{aligned}
\pi(s, T') &= \mathbb{E}^T\left[P(x(T); T, T') \,|\, S(T) = s\right] \\
&= \mathbb{E}^T\left[P(g\left(S(T)\right); T, T') \,|\, S(T) = s\right] \\
&= P(g(s); T, T') \\
&= \frac{P(0, T')}{P(0, T)} \exp\left\{-G(T, T')g(s) - \frac{G(T, T')^2}{2} y(T)\right\}.
\end{aligned}$$

Model requires numeric solution of $F(s, g(s)) = 0$ for a given instance of s.

How to combine Hull-White model and Vanilla model?

Hull-White TSR model is specified via

$$\pi(s, T') = \frac{P(0, T')}{P(0, T)} \exp\left\{-G(T, T')g(s) - \frac{G(T, T')^2}{2} y(T)\right\}$$

with $F(s, g(s)) = 0$.

- Mean reversion (for $G(T, T')$) is independent of Vanilla model.
  - Calibrate to market prices of related/sensitive intruments.
- We also need to specify volatility $\sigma(t)$ for calculation of $y(T)$.
- Hull-White model implies terminal distribution of $S(T)$ which, in general, is different from Vanilla model.
  - This constitutes inconsistency inherent in TSR models.
  - Calibrate Hull-White model *as close as possible* to Vanilla model.
  - Typical choice is matching ATM volatilities.

Alternative volatility choice mixes Hull-White and Vanilla model dynamics (1/2)

Hull-White model swap rate dynamics in annuity measure

$$\begin{aligned} dS\left(t, x(t)\right) &= \frac{\partial}{\partial x} S\left(t, x(t)\right) \cdot dx(t) + (\ldots)dt \\ &\approx \frac{\partial}{\partial x} S\left(0, x(0)\right) \cdot dx(t) + (\ldots)dt. \end{aligned}$$

Thus

$$\mathrm{Var}\left[S\left(T, x(T)\right)\right] \approx \left[\frac{\partial}{\partial x} S\left(0, x(0)\right)\right]^2 \cdot \mathrm{Var}\left[x(T)\right] = \left[\frac{\partial}{\partial x} S\left(0, x(0)\right)\right]^2 \cdot y(T).$$

Alternative volatility choice mixes Hull-White and Vanilla model dynamics (2/2)

$$\mathrm{Var}\left[S\left(T, x(T)\right)\right] \approx \left[\frac{\partial}{\partial x} S\left(0, x(0)\right)\right]^2 \cdot y(T).$$

This yields approximation for $y(T)$ for conditional zero coupon bond formula $\pi(s, T')$

$$y(T) = \underbrace{\left[\frac{\partial}{\partial x} S\left(0, x(0)\right)\right]^{-2}}_{\text{Hull-White model}} \cdot \underbrace{\mathrm{Var}\left[S\left(T\right)\right]}_{\text{Vanilla model}}.$$

- Sensitivity $\frac{\partial}{\partial x} S\left(0, x(0)\right)$ only depends on mean reversion.
- Variance $\mathrm{Var}\left[S\left(T\right)\right]$ is calculated solely from Vanilla model.

Outline

Linear TSR models postulate a parametric form for annuity mapping function

Linear TSR Model

In a linear TSR model the annuity mapping function is of the form

$$\alpha(s, T_p) = a(T_p)\,[s - S(t)] + \frac{P(t, T_p)}{An(t)}.$$

- Linear TSR model complies with no-arbitrage condition since

$$\begin{aligned}\mathbb{E}^A\left[\alpha(S(T), T_p)\,|\,\mathcal{F}_t\right] &= a(T_p)\cdot\underbrace{\mathbb{E}^A\left[[S(T) - S(t)]\,|\,\mathcal{F}_t\right]}_{=0} + \frac{P(t, T_p)}{An(t)} \\ &= \frac{P(t, T_p)}{An(t)}.\end{aligned}$$

- It remains to specify slope function $a(T_p)$.

Additivity and consistency condition yield constraints for lnear TSR model slope function I

Additivity condition yields

$$\sum_{i=0}^{n} \tau_i \cdot \alpha(s, T_i) = [s - S(t)] \underbrace{\sum_{i=0}^{n} \tau_i \cdot a(T_i)}_{=0} + \underbrace{\sum_{i=0}^{n} \tau_i \frac{P(t, T_i)}{An(t)}}_{=1} = 1.$$

For consistency condition we extend the index set, times and weights appropriately to

$$\alpha(s, T_0) - \alpha(s, T_N) + \sum_j \omega_j \cdot \alpha(s, \tilde{T}_{j-1}) = \sum_k \tilde{\omega}_k \cdot \alpha(s, \tilde{T}_{k-1}).$$

Then

$$\sum_k \tilde{\omega}_k \cdot \alpha(s, \tilde{T}_{k-1}) = [s - S(t)] \underbrace{\sum_k \tilde{\omega}_k \cdot a\left(\tilde{T}_{k-1}\right)}_{=1} + \underbrace{\sum_k \tilde{\omega}_k \cdot \frac{P(t, \tilde{T}_{k-1})}{An(t)}}_{S(t)} = s.$$

Additivity and consistency condition yield constraints for lnear TSR model slope function II

Additivity and consistency condition for linear TSR model

Overall slope level

$$\sum_{i=0}^{n} \tau_i \cdot a(T_i) = 0.$$

Change in slope

$$\sum_k \tilde{\omega}_k \cdot a\left(\tilde{T}_{k-1}\right) = 1$$

or equivalently

$$a\,(s,\,T_0) - a\,(s,\,T_N) + \sum_j \omega_j \cdot a(s,\,\tilde{T}_{j-1}) = 1.$$

Additivity and consistency condition fully specify a bi-linear annuity mapping function I

Bi-linear annuity mapping function

The bi-linear annuity mapping function is given by

$$\alpha(s, T_p) = \underbrace{\left[u \cdot (T_N - T_p) + v\right]}_{a(T_p)} \cdot [s - S(t)] + \frac{P(t, T_p)}{An(t)}$$

with

$$u = -\frac{\sum_i \tau_i}{\left[\sum_i \tau_i (T_N - T_i)\right] \cdot \left[\sum_k \tilde{\omega}_k\right] - \left[\sum_k \tilde{\omega}_k \left(T_N - \tilde{T}_{k-1}\right)\right] \cdot \left[\sum_i \tau_i\right]},$$

$$v = \frac{\left[\sum_i \tau_i (T_N - T_i)\right]}{\left[\sum_i \tau_i (T_N - T_i)\right] \cdot \left[\sum_k \tilde{\omega}_k\right] - \left[\sum_k \tilde{\omega}_k \left(T_N - \tilde{T}_{k-1}\right)\right] \cdot \left[\sum_i \tau_i\right]}.$$

Additivity and consistency condition fully specify a bi-linear annuity mapping function II

Result follows from

$$\sum_{i=0}^{n} \tau_i \cdot a(T_i) = u \underbrace{\sum_{i=0}^{n} \tau_i \left[T_N - T_i\right]}_{m_{11}} + v \underbrace{\sum_{i=0}^{n} \tau_i}_{m_{12}} = 0$$

$$\sum_{k} \tilde{\omega}_k \cdot a\left(\tilde{T}_{k-1}\right) = u \underbrace{\sum_{k} \tilde{\omega}_k \left[T_N - \tilde{T}_{k-1}\right]}_{m_{21}} + v \underbrace{\sum_{k} \tilde{\omega}_k}_{m_{22}} = 1$$

and Cramer's rule

$$u = \frac{0 \cdot m_{22} - 1 \cdot m_{12}}{m_{11} \cdot m_{22} - m_{12} \cdot m_{21}} \quad \text{and} \quad v = \frac{1 \cdot m_{11} - 0 \cdot m_{21}}{m_{11} \cdot m_{22} - m_{12} \cdot m_{21}}.$$

Some comments regarding bi-linear annuity mapping function...

- Method is straight forward and easy to implement.
- Appears natural due to simple linear structure and full specification via model-independent conditions.
- Linear TSR models also allow for very efficient pricing of CMS swaplets and options via power options.
- However,
  - method lacks linkage to term structure models,
  - does not allow for calibration to convexity adjustments observed in the market (e.g. via free mean reversion parameter).

Outline

Terminal Swap Rate Models

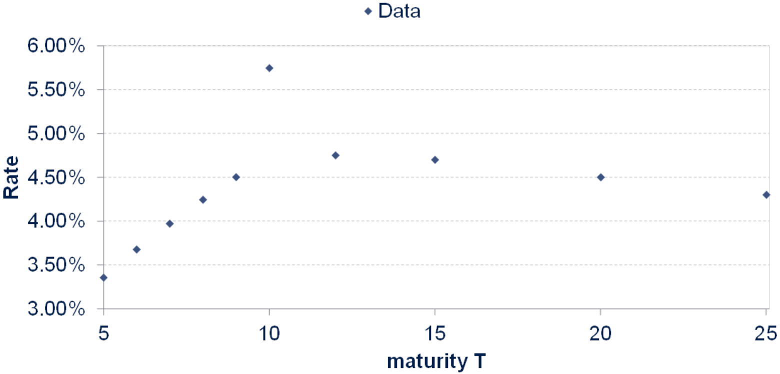
Cubic Spline Interpolation

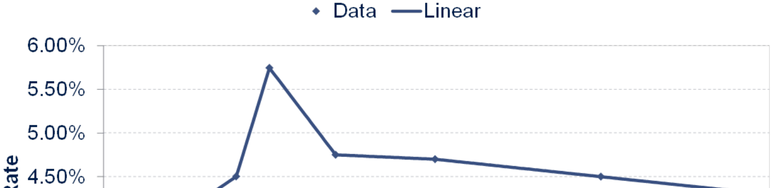
Separable HJM Revisited

Accuracy of Bermudan Pricing Methods

What is the purpose of spline interpolation?

- Suppose we want to fit a curve to a set of data points:





We analyse the example of cubic spline interpolation

First analyse a cubic function $f(t)$ on $[0,1]$ via

$$f(t) = a_3 t^3 + a_2 t^2 + a_1 t + a_0.$$

We get

$$\begin{array}{rclrcl} f(0) & = & a_0, & f'(0) & = & a_3 + a_2 + a_1 + a_0, \\ f(1) & = & a_1, & f'(1) & = & 3a_3 + 2a_2 + a_1. \end{array}$$

Solving for $a_0, \ldots, a_3$ yields

$$\begin{array}{rclrcl} a_0 & = & f(0), & a_2 & = & 3\left[f(1) - f(0)\right] - \left[f'(1) + 2f'(0)\right], \\ a_1 & = & f'(0), & a_3 & = & -2\left[f(1) - f(0)\right] + \left[f'(1) + f'(0)\right]. \end{array}$$

Cubic spline segment can be fully specified via function values and derivatives.

Cubic spline consists of segments of cubic functions

Assume we have a grid $x_0, \ldots, x_n$ with corresponding function values $y_0, \ldots, y_1$ and slopes $g_0, \ldots, g_n$ such that

$$y(x_i) = y_i \quad \text{and} \quad y'(x_i) = g_i.$$

Corresponding cubic spline is specified as

$$\begin{aligned}\bar{y}(x) = & \left[-2\left(y_i - y_{i-1}\right) + \left(g_i + g_{i-1}\right)\left(x_i - x_{i-1}\right)\right]\left(\frac{x - x_{i-1}}{x_i - x_{i-1}}\right)^3 + \\ & \left[3\left(y_i - y_{i-1}\right) - \left(g_i + 2g_{i-1}\right)\left(x_i - x_{i-1}\right)\right]\left(\frac{x - x_{i-1}}{x_i - x_{i-1}}\right)^2 + \\ & g_{i-1}\left(x_i - x_{i-1}\right)\left(\frac{x - x_{i-1}}{x_i - x_{i-1}}\right) + y_{i-1}\end{aligned}$$

for $x \in [x_{i-1}, x_i]$.
Note, spline representation follows from transformation

$$t = \frac{x - x_{i-1}}{x_i - x_{i-1}} \quad \text{and} \quad \frac{dt}{dx} = \frac{1}{x_i - x_{i-1}}.$$

Spline representation via x_i, y_i and g_i yields continuously differentiable function.

We can use slopes g_i to specify smoothness and monotonicity properties

- Usually, x_i and y_i are given; slopes g_i are a free parameter.
- Particular cubic spline methods are distinguished in how g_i are determined.

Natural Cubic (C^2) Spline Interpolation

Choose slopes such that $y(x)$ is twice continuously differentiable. Requires solving tridiagonal linear system.

Kruger Constrained Interpolation

Set slopes via harmonic mean. Abbreviate $s_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}$. Then

$$g_i = \begin{cases} 0 & s_i \cdot s_{i+1} < 0 \\ 2 s_i s_{i+1} / (s_i + s_{i+1}) & \text{else} \end{cases}$$

for $i = 1, \ldots, n-1$, $g_0 = \frac{3}{2} s_1 - \frac{1}{2} g_1$ and $g_n = \frac{3}{2} s_n - \frac{1}{2} g_{n-1}$.

There are several more cubic spline interpolation methods.[9]

[9] See e.g. Y. Iwashita. Piecewise Polynomial Interpolations. OpenGamma Quantitative Research. 2013

Outline

Outline

We have another look at the relation of $x(t)$ and $y(t)$ in the HJM model setting

We have (in risk-neutral measure)

$$x(t) = H(t)\left[\int_0^t g(s)^\top g(s)\left(\int_s^t h(u)du\right)ds + \int_0^t g(s)^\top dW(s)\right]$$

and

$$y(t) = H(t)\left(\int_0^t g(s)^\top g(s)ds\right)H(t).$$

Change of measure to T-forward measure in terms of Brownian motion becomes

$$dW^T(t) = \sigma_P(t,T)dt + dW(t)$$

with

$$\sigma_P(t,T) = g(t)\left(\int_t^T h(u)du\right).$$

In the T-forward measure the drift term of $x(t)$ may simplify I

Change of measure yields for $x(t)$

$$
\begin{aligned}
x(t) &= H(t)\left[\int_0^t g(s)^\top \left[\sigma_P(s,t) - \sigma_P(s,T)\right] ds + \int_0^t g(s)^\top dW^T(s)\right] \\
&= H(t)\left[\int_0^t g(s)^\top g(s)\left(-\int_t^T h(u)du\right) ds + \int_0^t g(s)^\top dW^T(s)\right] \\
&= H(t)\left[\int_0^t g(s)^\top g(s) ds\right] H(t)\left(-\int_t^T H(t)^{-1} h(u)du\right) \\
&\quad + H(t)\int_0^t g(s)^\top dW^T(s) \\
&= -y(t)\cdot\int_t^T H(t)^{-1} h(u)du + H(t)\int_0^t g(s)^\top dW^T(s) \\
&= -y(t)\cdot G(t,T) + H(t)\int_0^t g(s)^\top dW^T(s).
\end{aligned}
$$

In the T-forward measure the drift term of $x(t)$ may simplify II

Further

$$H(T)^{-1}x(T) - H(t)^{-1}x(t) = H(t)^{-1}y(t) \cdot G(t,T) + \int_t^T g(s)^\top dW^T(s).$$

This gives

$$x(T) = H(T)H(t)^{-1}\left[x(t) + y(t) \cdot G(t,T) + H(t)\int_t^T g(s)^\top dW^T(s)\right]$$

and

$$\mathbb{E}^T\left[x(T) \,|\, \mathcal{F}_t\right] = H(T)H(t)^{-1}\left[x(t) + y(t) \cdot G(t,T)\right],$$

$$\begin{aligned}\mathrm{Cov}^T\left[x(T) \,|\, \mathcal{F}_t\right] &= \mathbb{E}^T\left[H(T)\left(\int_t^T g(s)^\top g(s) ds\right) H(T)\right] \\ &= \mathbb{E}^T\left[y(T) - H(T)H(t)^{-1}y(t)H(t)^{-1}H(T) \,|\, \mathcal{F}_t\right].\end{aligned}$$

For implementations we need to calculate $H(T)H(t)^{-1}$and $G(t,T)$ I

We use representation in terms of short rate volatility $\sigma_r(s)^\top = H(s)g(s)^\top$and mean reversion $\chi(s)$ via $H'(s) = -\chi(s)\cdot H(s)$. It follows

$$
\begin{aligned}
H(t,T) &= H(T)H(t)^{-1} \\
&= \begin{bmatrix} \exp\left\{-\int_t^T \chi_1(s)ds\right\} & & \\ & \ddots & \\ & & \exp\left\{-\int_t^T \chi_d(s)ds\right\} \end{bmatrix}, \\
G(t,T) &= \int_t^T H(t)^{-1}h(u)du = \int_t^T H(u)H(t)^{-1}\mathbf{1}du \\
&= H(0,t)^{-1}\cdot\left[G(0,T) - G(0,t)\right].
\end{aligned}
$$

For implementations we need to calculate $H(T)H(t)^{-1}$and $G(t,T)$ II

Assume $\chi(s)$ is (piece-wise) constant on a time grid T_k. Then, for $t \in [T_{k-1}, T_k]$,

$$H(0,t) = H(0,T_{k-1}) \cdot H(T_{k-1},t)$$

with components $H_i(T_{k-1},t)$ given as

$$H_i(T_{k-1},t) = \exp\left\{-\int_{T_{k-1}}^{t} \chi_i(s)ds\right\} = e^{-\chi_i^k(t-T_{k-1})}$$

and

$$G(0,t) = G(0,T_{k-1}) + H(0,T_{k-1}) \cdot G(T_{k-1},t)$$

For implementations we need to calculate $H(T)H(t)^{-1}$and $G(t,T)$ III

with components $G_i(T_{k-1},t)$ given as

$$
\begin{aligned}
G_i(T_{k-1},t) &= \int_{T_{k-1}}^{t} \exp\left\{-\int_{T_{k-1}}^{u} \chi_i(s)ds\right\} du \\
&= \int_{T_{k-1}}^{t} \exp\left\{-\int_{T_{k-1}}^{u} \chi_i^k ds\right\} du \\
&= \int_{T_{k-1}}^{t} \exp\left\{-\chi_i^k\left(u - T_{k-1}\right)\right\} du \\
&= \left[\frac{1-\exp\left\{-\chi_i^k\left(t - T_{k-1}\right)\right\}}{\chi_i^k}\right].
\end{aligned}
$$

The quantities $H(0,T_{k-1})$ and $G(0,T_{k-1})$ can be pre-computed and cached for efficient calculation of $H(t,T)$ and $G(t,T)$.

For Gaussian models we can also calculate $y(t)$ I

We have for $t \in [T_{k-1}, T_k]$

$$y(t) = H(T_{k-1}, t)y(T_{k-1})H(T_{k-1}, t) + H(t)\left(\int_{T_{k-1}}^{t} g(s)^\top g(s) ds\right) H(t)$$

We re-write $g(s)$ in terms of short rate volatility $\sigma_r(s) = g(s)H(s)$ as

$$y(t) = H(T_{k-1}, t)y(T_{k-1})H(T_{k-1}, t) + \int_{T_{k-1}}^{t} H(s,t)\sigma_r(s)^\top \sigma_r(s)H(s,t)ds.$$

Assume $\sigma_r(s)$ is (piece-wise) constant on $[T_{k-1}, T_k]$. Then denote

$$\Sigma^2 = \left[\Sigma^2_{i,j}\right]_{i,j=1}^{d} = \sigma_r(s)^\top \sigma_r(s), \; s \in [T_{k-1}, T_k].$$

For Gaussian models we can also calculate $y(t)$ II

The matrix components $M_{i,j}$ of $M(T_{k-1}, t) = \int_{T_{k-1}}^{t} H(s,t)\Sigma^2 H(s,t)ds$ are

$$\begin{aligned} M_{i,j} &= \int_{T_{k-1}}^{t} e^{-\chi_i^k(t-s)}\Sigma_{i,j}^2 e^{-\chi_j^k(t-s)}ds = \Sigma_{i,j}^2 \int_{T_{k-1}}^{t} e^{-\left(\chi_i^k+\chi_j^k\right)(t-s)}ds \\ &= \frac{\Sigma_{i,j}^2}{\chi_i^k + \chi_j^k}\left[1 - \exp\left\{-\left(\chi_i^k + \chi_j^k\right)(t - T_{k-1})\right\}\right]. \end{aligned}$$

As a result we get

$$y(t) = H(T_{k-1}, t)y(T_{k-1})H(T_{k-1}, t) + M(T_{k-1}, t).$$

Again, $y(T_{k-1})$ can be pre-computed and cached for efficient calculation of $y(t)$.

Outline

Outline

We analyse the accuracy of numerical methods by means of a coupon bond option

Market data and model setup

Flat yield curve 3% (cont. compounding, Act/365), $100bp$ short rate volatility, mean reversion 5%.

Coupon bond option test instrument setup

- European call option, exercise in $10y$ at unit strike.
- 3% coupons at $11y, \ldots, 20y$, unit notional payment in $20y$.
- All dates and year fractions in model times.

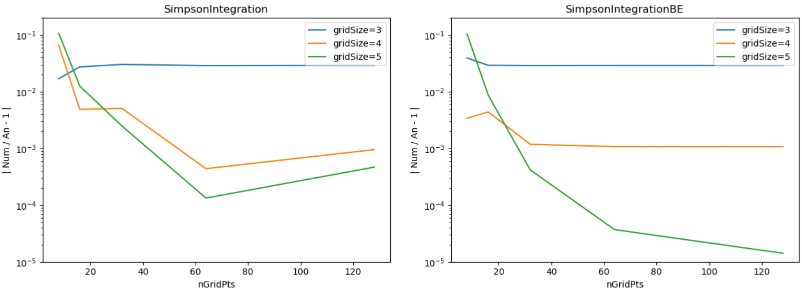
Testing approach

- Construct pseudo Bermudan option from European coupon bond option by adding zero strike exercises at $2y$ and $6y$.
- Compare numerical Bermudan option price versus analytical European option price

$$\text{RelErr} = \left| \frac{\text{BermudanPrice}}{\text{EuropeanPrice}} - 1 \right|.$$

Density integration methods are compared for scenarios of grid size, # grid points and Hermite polynomial degree I

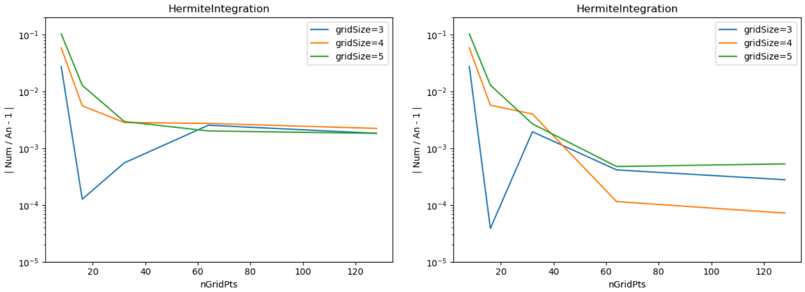
Simpson's rule - w/o (l) and w/ (r) break-even calculation



- Accuracy is mainly limited by grid size.
- Break-even calculation required to achieve higher accuracy.

Density integration methods are compared for scenarios of grid size, # grid points and Hermite polynomial degree II

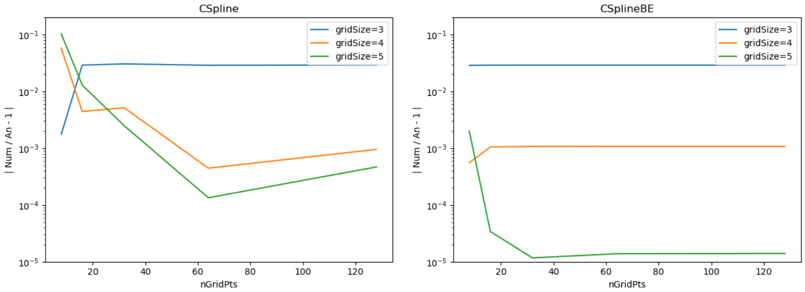
Hermite integration - degree $d = 5$ (l) and $d = 10$ (r)



- Higher polynomial degree is required to mitigate non-smooth payoff impact.
- Too large grid size seems to deteriorate accuracy.

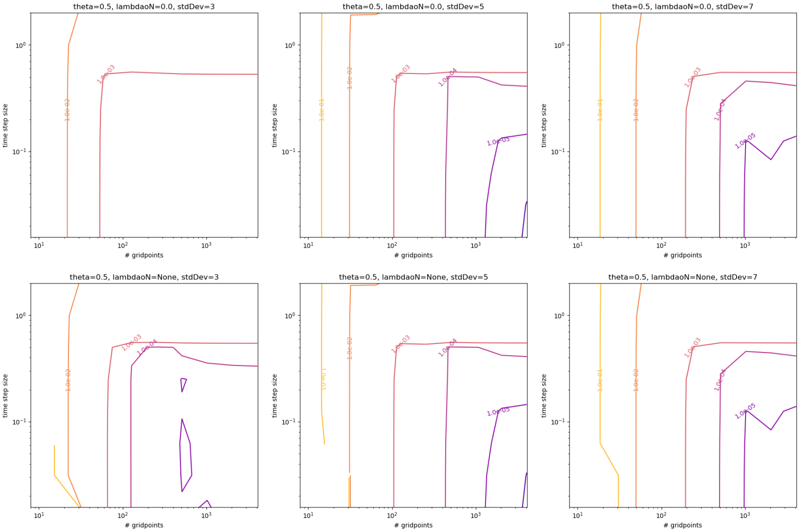
Density integration methods are compared for scenarios of grid size, # grid points and Hermite polynomial degree III

Cubic spline - w/o (l) and w/ (r) break-even calculation



- Accuracy is mainly limited by grid size and break-even calculation.
- CSpline with break-even clearly outperforms other methods for small number of grid points.

We analyse PDE methods using contour plots of error estimate for # of grid points versus time step size I

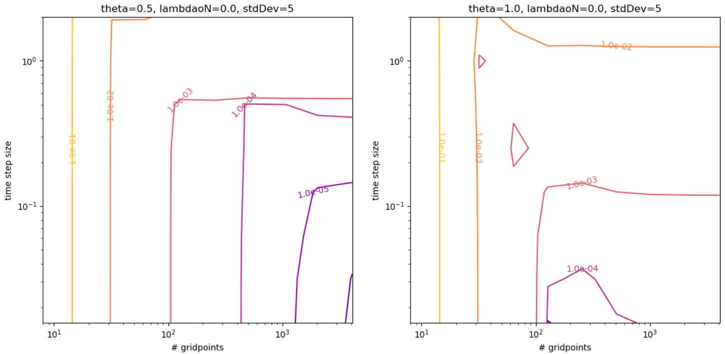


We analyse PDE methods using contour plots of error estimate for # of grid points versus time step size II

- # grid points need to be increased simultanously to reducing time step size to improve accuracy.
- Again, accuracy is limited by grid size.
- For small grid sizes approximation of boundary condition (via $\lambda_{0,N}$) improves accuracy.

We analyse PDE methods using contour plots of error estimate for # of grid points versus time step size III

Compare $\theta = \frac{1}{2}$ (l) versus $\theta = 1$, i.e. Implicit Euler (r)



- Implicit Euler requires smaller step size to achive same accuracy as for $\theta = \frac{1}{2}$ (i.e. Cranck-Nicolson).

Outline

We analyse the accuracy of numerical methods by means of a coupon bond option I

Market data and model setup

Flat yield curve 3% (cont. compounding, Act/365), $100bp$ short rate volatility, mean reversion 5%.

Coupon bond option test instrument setup

- European/Bermudan call option, exercise in $10y$ $(11y, \ldots, 19y)$ at unit strike.
- 3% coupons at $11y, \ldots, 20y$, unit notional payment in $20y$.
- All dates and year fractions in model times.

We analyse the accuracy of numerical methods by means of a coupon bond option II

Testing approach

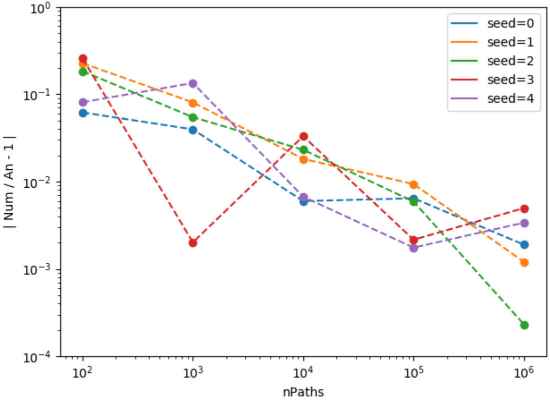
- Construct pseudo Bermudan option from European coupon bond option by adding zero strike exercises at $2y$ and $6y$.
- Compare numerical Bermudan option price versus analytical European option price.

$$\mathsf{RelErr} = \left| \frac{\mathsf{BermudanPrice}}{\mathsf{EuropeanPrice}} - 1 \right|.$$

- Compare MC Bermudan price versus density integration reference price.

MC methods are compared for scenarios of seed, # paths, as well as model and option parameters I

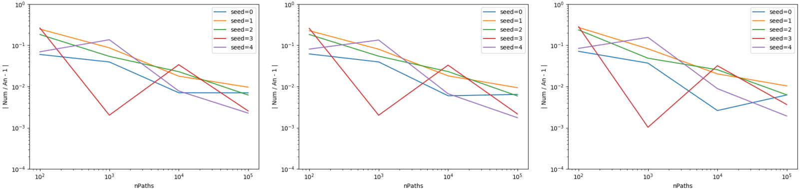
Base scenario, ATM European option



- MC estimate is a random number - dependency on seed illustrates this aspect.

MC methods are compared for scenarios of seed, # paths, as well as model and option parameters II

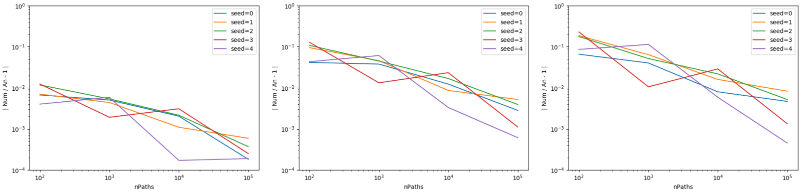
ATM European option - low volatility ($10bp$, left) and negative mean reversion (-3%, right) scenarios



- Relative (!) error more or less invariant to model parameters.
- Note that ATM option value is roughly proportional to variance (driven by volatility and mean reversion).

MC methods are compared for scenarios of seed, # paths, as well as model and option parameters III

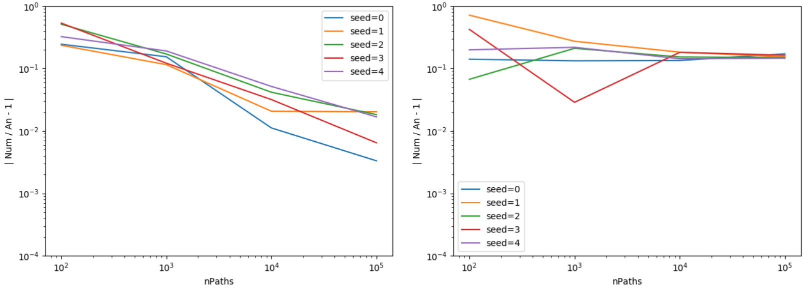
ITM European option - low volatility ($10bp$, left) and negative mean reversion (-3%, right) scenarios



- Relative error decreases for low model variance and increases for high model variance
- Note that ITM option converges to positive intrinsic value if variance decreases

AMC methods are compared for scenarios of seed, # paths, as well as AMC regression properties I

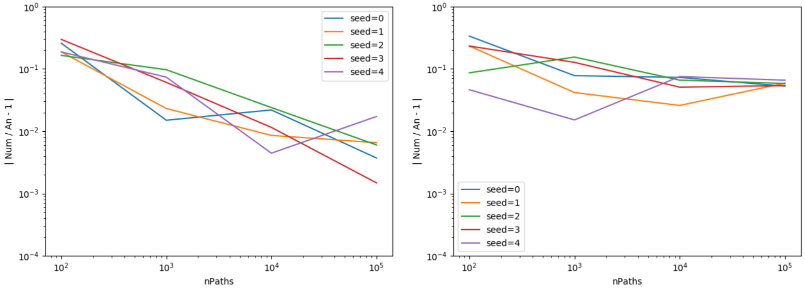
Pseudo-Bermudan option with hold value regression (left) vs. exercise decision only regression (right)



- Regression on exercise decision only does not work in this case.

AMC methods are compared for scenarios of seed, # paths, as well as AMC regression properties II

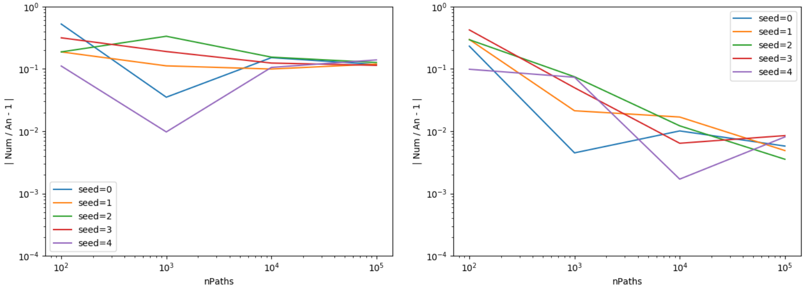
Bermudan option with hold value regression (left) vs. exercise decision only regression (right)



- Regression on exercise decision only does not work in this case.

AMC methods are compared for scenarios of seed, # paths, as well as AMC regression properties III

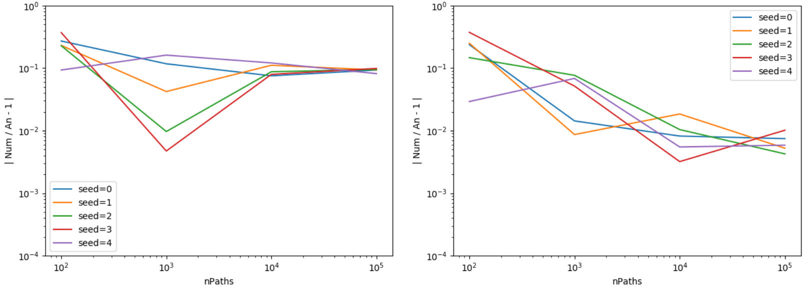
Bermudan option with max. polynomial degree 1 (left) vs. 6 (right) - default is 3



- Too small polynomial degree prevents convergence.
- Very high polynomial degree does not improve accuracy.

AMC methods are compared for scenarios of seed, # paths, as well as AMC regression properties IV

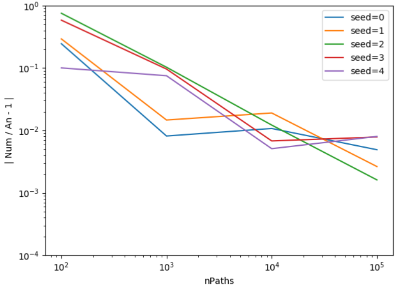
Bermudan option with co-terminal swap rate basis and max. polynomial degree 1 (left) vs. 3 (right)



- Too small polynomial degree prevents convergence.

AMC methods are compared for scenarios of seed, # paths, as well as AMC regression properties V

Bermudan option with co-terminal swap rate and Libor rate basis (max. polynomial degree 3)



- Similar result as for other basis functions.

Part X

Appendix

Outline

References

Outline

References I

F. Ametrano and M. Bianchetti.
Everything you always wanted to know about Multiple Interest Rate Curve Bootstrapping but were afraid to ask (April 2, 2013).
Available at SSRN: http://ssrn.com/abstract=2219548 or http://dx.doi.org/10.2139/ssrn.2219548, 2013.

L. Andersen and V. Piterbarg.
Interest rate modelling, volume I to III.
Atlantic Financial Press, 2010.

D. Bang.
Local-stochastic volatility for vanilla modeling.
https://ssrn.com/abstract=3171877, 2018.

M. Beinker and H. Plank.
New volatility conventions in negative interest environment.
d-fine Whitepaper, available at www.d-fine.de, December 2012.

D. Brigo and F. Mercurio.
Interest Rate Models - Theory and Practice.
Springer-Verlag, 2007.

References II

D. Duffy.
Finite Difference Methods in Financial Engineering.
Wiley Finance, 2006.

M. Fujii, Y. Shimada, and A. Takahashi.
Collateral posting and choice of collateral currency - implications for derivative pricing and risk management (may 8, 2010).
Available at SSRN: https://ssrn.com/abstract=1601866, May 2010.

M. Giles and P. Glasserman.
Smoking adjoints: fast monte carlo greeks.
Risk, January 2006.

P. Glasserman.
Monte Carlo Methods in Financial Engineering.
Springer, 2003.

A. Griewank and A. Walther.
Evaluating derivatives: principles and techniques of algorithmic differentiation - 2nd ed.
SIAM, 2008.

References III

P. Hagan, D. Kumar, A. Lesniewski, and D. Woodward.
Managing smile risk.
Wilmott magazine, September 2002.

P. Hagan and G. West.
Interpolation methods for curve construction.
Applied Mathematical Finance, 13(2):89–128, 2006.

M. Henrard.
Interest rate instruments and market conventions guide 2.0.
Open Gamma Quantitative Research, 2013.

M. Henrard.
A quant perspective on ibor fallback proposals.
https://ssrn.com/abstract=3226183, 2018.

M. Henrard.
A quant perspective on ibor fallback consultation results.
https://ssrn.com/abstract=3308766, 2019.

References IV

J. Hull and A. White.
Pricing interest-rate-derivative securities.
The Review of Financial Studies, 3:573–592, 1990.

Y. Iwashita.
Piecewise polynomial interpolations.
OpenGamma Quantitative Research, 2013.

A. Lyashenko and F. Mercurio.
Looking forward to backward-looking rates: A modeling framework for term rates replacing libor.
https://ssrn.com/abstract=3330240, 2019.

U. Naumann.
The Art of Differentiating Computer Programs: An Introduction to Algorithmic Differentiation.
SIAM, 2012.

V. Piterbarg.
Funding beyond discounting: collateral agreements and derivatives pricing.
Asia Risk, pages 97–102, February 2010.

References V

R. Rebonato.
Volatility and Correlation.
John Wiley & Sons, 2004.

S. Shreve.
Stochastic Calculus for Finance II - Continuous-Time Models.
Springer-Verlag, 2004.

Contact

Dr. Sebastian Schlenkrich

HU Berlin
Office: RUD25, R 1.201
Mail: sebastian.schlenkrich@hu-berlin.de

FRAME Consulting GmbH
www.frame-consult.de
Mail: sebastian.schlenkrich@frame-consult.de