# Modern Modeling and Pricing of Interest Rates Derivatives
# Day 2 - Session 3: Bermudan IR swaption

Vladimir Sankovich [1]
Qinghua Zhu [2]

London, June 6-7, 2018

[1] vsankovich@drwholdings.com
[2] qzhu@drwholdings.com

# Bermudan Swaptions Introduction

Swaptions by Exercise Style:

- **European Swaptions:** Single exercise into the underlying swap (or cash-settled).
- **American Swaptions:** Continuous exercise into the underlying swap (typically after a lockout period)
- **Bermudan (aka Multi-European) Swaptions:** Multiple Periodic (discrete) exercises into the underlying swap (typically after a lockout period)

Structure of Bermudan Exercise Schedule:

- A Bermudan exercise schedule is comprised of pairs $\{t_{\text{ex}}, t_{eff}\}$. Here $t_{\text{ex}}$ represent *exercise* dates on which the exercising party notifies of its intention to exercise the option. $t_{ex}$ are *effective* dates which represent the effective dates of the swap into which the parties are entering upon the option's exercise.
- Given a lockout period , $T_L$, and the final maturity of the underlying swap, $T_m$, the structure of exercises can be visualized schematically as $T_L \leq t_{ex_1} \leq t_{eff_1} < t_{ex_2} \leq t_{eff_2} < \ldots < t_{ex_n} \leq t_{eff_n} < T_m$.
- Once the exercise decision is made at $t_{ex_i}$ the exercising party will enter into a swap with the remaining maturity, $T_m - t_c^*$, where $t_c^*$ is the first coupon date of the underlying swap on or after the effective date, $t_{eff_i}$. If the effective date happens to fall inside a coupon period, the resulting accrued interest will be settled on the effective date.
- The delay between the exercise and the corresponding effective date can be material. Exercise structure of Bermudan swaptions are often constructed to mimic the call schedules of callable bonds where these delays may be as long as 45 to 90 days. Sometimes exercising the option incurs a fee, which is typically a fixed cash amount which should be factored in when making exercise decisions.

# Uses of Bermudan Swaptions - I

- Stand-alone Bermudan options are typically used for hedging optionality in investment instruments such as notes or mortgage-backed bonds.
- While MBS have built-in optionality due to their prepayment features, callable notes and bonds are issued to offer investors higher coupon compared to straightforward bullet issues. When buying a callable bond, investors effectively sell volatility (optionality) back to the issuer in exchange for higher coupons.
- To make these structures more attractive, issuers often guarantee high fixed rate payments for a period of time during which the bond cannot be called away. This is the reason for the "lockout period" we mentioned before.
- Embedding optionality in funding instruments is often referred to as "yield enhancement". However, because the note will be called away from the investor when the market moves sufficiently against the issuer, it is often difficult to claim conclusively that the overall return (yield) of a callable note is higher than that of a bullet issue.
- Issuers (e.g. corporate treasuries) seek to control their funding levels (usually against some floating benchmark index such as 3M Libor). Thus once a structured note is placed with the investors, the issuer will typically look to hedge out the embedded optionality by entering into a Bermudan-style swaption contract with a dealer bank. Some issuers such as the GSEs may keep this optionality to offset their others exposures from, e.g. mortgage bonds.
- In some circumstances an issuer may enter into a cancellable swap with the dealer which allows them to offload the Bermudan optionality and potentially switch their funding from floating to fixed or vice versa.
- A callable (cancellable) swap is just a portfolio of a bullet swap and a Bermudan option to enter into an offsetting swap. Thus pricing of callable swaps presents no additional difficulties compared to stand-alone Bermudan swaptions.

# Uses of Bermudan Swaptions - II

- Until the late 2000's the very active US mortgage market precipitated a huge flow in the Bermudan swaptions. Agencies securitized and sold an enormous amount of debt collateralized by US mortgages and correspondingly were very actively hedging their volatility exposures by buying Bermudan-style options from the dealers.
- Their hedging neeeds stimulated the Bermudan market which reached its peak liquidity just before the crisis. Essentially, despite their inherently "exotic" nature, Bermudan swaptions were traded almost on the "mark-to-market" basis.
- Another consequence of the agencies' hedging activity was the heavily one-way Bermudan market with the dealers almost always being short Bermudan optionality.
- Some of the buy-side firms developed strategies aiming to capitalize on these dislocations in the Bermudan markets. They would trade Bermudans against baskets of European swaptions to isolate the exposure to the "Bermudan volatility" or more specifically, the premium which market participants were willing to pay for the additional optionality embedded in Bermudan swaptions compared to the European counterparts (more on this below).
- Over the last several years, the flow in Bermudan swaptions has dropped significantly from its peak levels. Together with strict capital requirements imposed on banks' derivatives businesses by the regulators and the overall consolidation in the ranks of sophisticated buy-side players this caused the decline in popularity of Bermudan-based speculative strategies.

# Analysis

- From the exercise structure shown on the first slide it is clear that the option holder has the right to choose to enter into one of the swaps with maturities $M_i = T_m - t_c^*$ as defined above. These swaps are called "co-terminal" as they all share the same maturity as the overall structure.
- Assuming for the moment zero notification delay, on each exercise date the owner of the option will have to compare the value of exercise, i.e. the present value of the $i-$th co-terminal swap with the continuation value which is the present value of the remaining option.
- Mathematically this exercise decision can be expressed as:

$$V(t_{ex_i}) = \max(V_{ex}, V_{cont}) = \max(E_{t_{ex_i}}\left[V_{swap}(t_{ex_i}, T_m)\right] - \text{fee}_i, E_{t_{ex_i}}\left[V(t_{ex_{i+1}})\right]) \tag{1}$$

- If we expand the recursive expectation above, we see that the value of the Bermudan swaption at each exercise date can be seen as an option on another option with future exercise. This interpretation gives us some intuition as to what determines the price of a Bermudan.
- Another way to think about Bermudan swaptions is to consider the optionality embedded in them. Is it a basket of European swaptions? No, because only one option can be exercised.
- We can always exercise the most expensive European, so the Bermudan should be worth at least that much. However, there's value in the switch option - the option to choose which of the embedded European options to exercise.
- Thus the value of a Bermudan can be taken as the price of the most expensive European plus the value of the switch option.

# Pricing - I

- Either interpretation of Bermudan swaptions above shows that their prices depend on the set of joint distributions of the remaining embedded swaps as seen from each exercise date.
- A model which would be suitable for these products would have to be able to capture the evolution of these joint distributions through time. None of the models we've discussed so far are capable of that as they either gave only the expectations of forward rates (yield curve model) or only their *terminal* (potentially, joint, c.f. copulas) distributions.
- We need a term structure model, i.e. a model of the evolution (dynamics) of rates.
- Another requirement: the distributions implied by the dynamic model should match (in some sense) what we're observing in the market. Why? So that we could use the model to compute hedges in terms of market-traded instruments.
- The process of "forcing" a model through some points observed in the market today is known as calibration. What should we expect and require from our calibration?
- Ideally if we were able to calibrate so as to capture all model risk factors, we would be able to establish a static hedge.
- The market doesn't give us enough to do this: no liquid instruments to fix inter-temporal correlations of rates. (In the equity derivatives markets products known as cliquet options may be used to imply these parameters.) However we can fix the marginal distribution of each of the swap rates relevant to a particular Bermudan.
- The rates whose distributions we'd like to pin down are those that underlie each of the embedded European swaptions. They are known as "co-terminal" or "diagonal" swaptions.

# Pricing - II

Thus the strategy for setting up a modeling framework for Bermudans can be outlined as:

- Choose an appropriate term structure model
- Calibrate its parameters to capture the marginal distrbutions of embedded swaps as seen in the market today
- Allow the model's dynamics to evolve the rates through time between these fixed marginal distributions

With a term structure model selected and calibrated we can use it to price Bermudans:

- As always, pricing amounts to computing the risk-neutral expectation of the product's discounted payoff:

$$V(t) = E_t^Q[P(t, T_m)V(T_m)]$$

- Unlike their European-style counterparts, a Bermudan option's payoff is not uniquely specified on a single maturity date, but is defined recursively, cf. Eq. (1).
- Pricing these options amounts to finding the optimal exercise strategy on each of the exercise dates starting with the last one and moving back in time until we reach the earliest exercise date. This procedure can be visualized by applying the telescoping property of conditional expectations:

$$V(t) = E_t^Q[P(t, T)V(T_m)] =$$

$$= E_t[P(t, t_{ex_1})E^Q_{t_{ex_1}}[P(t_{ex_1}, t_{ex_2})E^Q_{t_{ex_2}}[\cdots P(t_{ex_{n-1}}, t_{ex_n})E^Q_{t_{ex_n}}[P(t_{ex_n}, T_m)V(T_m)]\cdots] =$$

$$= E_t^Q[P(t, t_{ex_1})E^Q_{t_{ex_1}}[P(t_{ex_1}, t_{ex_2})E^Q_{t_{ex_2}}[P(t_{ex_2}, t_{ex_3})V(t_{ex_3})]]] \quad (2)$$

# Pricing - III

- The equation (2) suggests that we should compute the nested expectations in it by first evaluating the final payoff at maturity, $T_m$, then moving backwards through time and applying (1) at each exercise date to compute the value of the derivative on that date, $V(t_{ex_i})$.
- This algorithm is usually referred to as "rollback" and is indeed the most common method for finding prices of Bermudan swaptions[3].
- To facilitate this rollback consistently with the chosen term structure model of rates we need to construct a discrete approximation to the model's continuous dynamics using a numerical method such as, for example, a trinomial tree.
- In the following sections we will demonstrate how this can be done and introduce the term structure models suitable for pricing Bermudan swaptions and their calibration techniques.

[3] Those who are familiar with the theory of stochastic optimization will recognize that this algorithm finds the optimal exercise strategy which is equivalent to finding a global solution to the stochastic HJB equation.

# Term Structure Models

Equivalent Martingale Measure

- Risk-neutral pricing: $V(t) = E_t^Q[e^{-\int_t^T r(s)ds} V(T)]$, where $Q$ is the risk neutral measure and money market account $\beta(t) = e^{\int_0^t r(s)ds}$ is used as the numeraire.
- Change of measure: Consider a different numeraire $N(t)$, and associated probability measure $Q^N$, then the Radon-Nikodym derivative is given by $\zeta(t) = E_t^{Q^N}\left[\frac{dQ}{dQ^N}\right] = \frac{\beta(t)/\beta(0)}{N(t)/N(0)}$, and the pricing formula becomes $V(t) = \beta(t)E_t^Q\left[\frac{V(T)}{\beta(T)}\right] = N(t)E_t^{Q^N}\left[\frac{V(T)}{N(T)}\right]$.
- Note that $N$ doesn't have to be any "identifiable" asset such as zero coupon bond or an annuity. It only has to be a positive process.

Two elements for a term structure model:

- A set of stochastic processes $X$ that drive the evolution of interest rates.
- A functional form that maps $X$ into a numeraire.

# The One Factor LGM (Linear Gaussian Markov) Model

- There is one **Gaussian** process (state variable) that drives interest rates:

  $$dX(t) = \alpha(t)dW(t); X(0) = 0.$$

  Let $x$ be the value of the state variable at time $t$, i.e., $x = X(t)$, then the entire term structure at $t$ is solely determined by $x$.

- Define the numeraire as

  $$N(t,x) = \frac{1}{P(0,t)} e^{H(t)x + A(t)}$$

  **Markovian**: The value of the numeraire depends only on $x$.
  **Linear**: $\ln N(t,x)$ is linear in $x$.

- Pricing: $V(t,x) = N(t)E_t^{Q^N}\left[\frac{V(T)}{N(T)}\middle| X(t) = x\right] = N(t)\int \frac{V(T)}{N(T)}p(t,x;T,X)dX,$
  where $p(t,x;T,X) = \frac{1}{\sqrt{2\pi v(t,T)}}e^{-\frac{1}{2}(X-x)^2/v(t,T)}$, $v(t) = \int_0^t \alpha^2(s)ds$, and $v(t,T) = v(T) - v(t)$.

# The LGM Model

- Zero Coupon Bond Price:

$$Z(t,T;x) = \frac{P(0,T)}{P(0,t)} e^{H(t)x + A(t) - A(T)} E_t^{Q^N}[e^{-H(T)X(T)} | X(t) = x]$$
$$= \frac{P(0,T)}{P(0,t)} e^{-[H(T)-H(t)]x - [A(T)-A(t)] + \frac{H^2(T)}{2} V(t,T)}$$

- Non-arbitrage condition: $Z(0,T;0) = P(0,T), \forall T$. Therefore $A(T) = \frac{1}{2}H^2(T)v(t)$.
- Rewrite numeraire: $N(t,x) = \frac{1}{P(0,t)} e^{H(t)x + \frac{1}{2}H^2(t)v(t)}$.
- Rewrite ZCB price: $Z(t,T;x) = \frac{P(0,T)}{P(0,t)} e^{-[H(T)-H(t)]x - \frac{v(t)}{2}[H^2(T)-H^2(t)]}$
- Model parameters: $H'(T)$ and $v(t)$.

# The LGM Model

Connection to the Hull-White Model

- Short rate dynamics under Hull-White model: $dr = [\theta(t) - \lambda(t)r]dt + \sigma(t)dW_t$, where $\theta(t)$ is chosen to match today's discount factor $P(0, T)$.
- Valuation under Hull-White: $V(t, r) = E_t^Q[e^{-\int_t^T r(s)ds} V(T, r(T))|r(t) = r]$.
- Hull-White parameters: $\lambda(t)$ and $\sigma(t)$.
- Connection between Hull-White and LGM:

$$\lambda(t) = -\frac{H''(t)}{H'(t)}$$

$$\sigma(t) = H'(t)\sqrt{v'(t)}$$

- LGM is exactly the Hull-White model written in a more convenient form.

# The LGM Model

Connection with HJM

- Let $f(t, T; x)$ be the instantaneous forward rate, since $Z(t, T; x) = e^{-\int_t^T f(t, T'; x) dT'}$, we have

$$f(t, T; x) = f_0(T) + H'(T)x + [H'(T)]^2 v(t)$$

  The last term $[H'(T)]^2 v(t)$ is a small convexity correction. Qualitatively, the instantaneous forward curve moves as a perturbation of the initial curve, with the perturbation weighted by the state variable $x$, and controlled by $H$. With constant mean reversion, $H'(T) = e^{-\lambda T}$, the higher the mean reversion, the lower the variation of instantaneous forwards.

- Applying Ito's lemma, it can be shown that

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

  where $\sigma_f(t, T) = H'(T)\alpha(t) = e^{-\lambda(T-t)}\sigma(t)$.
  The "forward rate drift restriction" of the HJM theory is automatically satisfied.

# European Swaption Pricing - The Jamshidian Decomposition

- Idea: rewrite the swaption payout from an option on a sum of discount bonds to a sum of options on discount bonds.
- The payout of a swaption at state $x$ is $(V_{\text{flt}} - V\text{fix})_+$, i.e.
  $V_{\text{swaption}}(T_0) = \left(1 - Z(T_0, T_N; x) - c\sum_{i=0}^{N-1} \tau_i Z(T_0, T_{i+1}; x)\right)_+$ [4].
  Each $Z(T_0, T_i; x)$ monotonically decreases in $x$, hence the swap PV increases in $x$.
- Find a "critical" value $x^*$ of the state variable, such that the swap at time $T_0$ is exactly zero:

$$1 - Z(T_0, T_N; x^*) - c\sum_{i=0}^{N-1} \tau_i Z(T_0, T_{i+1}; x^*) = 0.$$

- Define the "strikes" for zero coupon bonds:

$$K_i = Z(T_0, T_N; x^*), \qquad i = 1, ..., N.$$

  It follows that $K_N + c\sum_{i=0}^{N-1} \tau_i K_{i+1} = 1$.
- The swaption only pays out a positive amount if $x(T_0) > x^*$:

$$V_{\text{swaption}}(T_0) = \left(1 - Z(T_0, T_N; x^*) - c\sum_{i=0}^{N-1} \tau_i Z(T_0, T_N; x^*)\right) 1_{X(T_0) > x^*}$$

$$= [K_N - Z(T_0, T_N; X(T_0))]_+ + c\sum_{i=0}^{N-1} \tau_i [K_{i+1} - Z(T_0, T_{i+1}; x(T_0))]_+$$

- The swaption payout is then decomposed into $N+1$ put options on zero coupon bonds, which can be priced in closed form.

[4] This equation only holds if the swap is Libor discounted. Some adjustments need to be made with OIS discounting, assuming deterministic Libor-OIS spreads.

# The QGM (Quadratic Gaussian Markov) Model

- Same as LGM, assume one Gaussian state variable: $dX(t) = \alpha(t)dW(t); X(0) = 0$, and let $v(t) = \int_0^t \alpha^2(s)ds$ be the accumulated variance.
- Goal: Make log of zero coupon bond price $Z(t, T; x)$ quadratic in $x$.
- Note that the *reduced* price of any derivative, i.e., price in the units of the reference numeraire, $\hat{V}(t) := \frac{V(t)}{N(t)}$, is a martingale under the $Q^N$ measure, therefore

  $$\partial_t \hat{V} + \tfrac{1}{2} v'(t) \partial_{xx} \hat{V} = 0$$

  In particular, it applies to reduced prices of zero coupon bonds $\hat{Z}(t, T; x)$:

  $$\partial_t \hat{Z}(t, T; x) + \tfrac{1}{2} v'(t) \partial_{xx} \hat{Z}(t, T; x) = 0.$$

- We look for the solution of the form $\hat{Z}(t, T; x) = P(0, T) e^{-Q(t,T)x^2 - L(t,T)x - F(t,T)}$, then the above equation implies

  $$Q_t(t, T) - 2v'(t)Q^2(t, T) = 0$$
  $$L_t(t, T) - 2v'(t)Q(t, T)L(t, T) = 0$$

  The two ODEs can be solved explicitly to determine the quadratic term $Q$ and the linear term $L$.

# The QGM Model

- The constant term $F(t, T)$ can be determined from the non-arbitrage condition, i.e., the initial yield curve needs to matched: $\hat{Z}(0, T; 0) = P(0, T)$.
- The reduced zero coupon bond price is then

$$\hat{Z}(t, T; x) = \frac{P(0,T)}{\sqrt{1-C(T)v(t)}} \exp\left(-\frac{\frac{1}{2}C(T)x^2 + H(T)x + \frac{1}{2}H^2(T)v(t)}{1-C(T)v(t)}\right)$$

- It is now possible to deduce the functional form of the numeraire. Since $\hat{Z}(t, t; x) = \frac{1}{N(t,x)}$, we have

$$\hat{N}(t, x) = \frac{\sqrt{1-C(t)v(t)}}{P(0,t)} \exp\left(\frac{\frac{1}{2}C(t)x^2 + H(t)x + \frac{1}{2}H^2(t)v(t)}{1-C(t)v(t)}\right)$$

- The actual zero coupon bond price $Z(t, T; x)$ can then be easily written out by $Z(t, T; x) = \hat{Z}(t, T; x)N(t, x)$.
- Model parameters: $C(t)$, $H(t)$, $v(t)$.

# The QGM Model

- European swaption pricing: The Jamshidian trick can still be applied. Due to the quadratic nature of ZCB prices, when solving for the break-even state to make the swap PV zero, we will have roots, $x_1^*$ and $x_2^*$. Therefore, we will have to integrate the payoff over either $(\infty, x_1^*) \cup (x_2^*, \infty)$ or $(x_1^*, x_2^*)$. However, since ZCB price is log-quadratic in the state variable, integrating it against a Gaussian kernel can be explicitly calculated, therefore the European swaption price has closed form once $x_1^*$ and $x_2^*$ are found.
- Re-parameterize the model: define

$$\gamma(t) = \frac{C'(t)f_0(t)}{(H'(t))^2}$$
$$\lambda(t) = -\frac{H''(t)}{H'(t)}$$
$$\sigma(t) = H'(t)\sqrt{v'(t)}$$

then $\gamma, \lambda$ and $\sigma$ can be viewed as the CEV parameter, mean reversion rate and instantaneous vol of the short rate dynamics, respectively.

# The QGM Model

Summary

- Arbitrage free by construction.
- Closed form ZCB prices allow for efficient calibration.
- Intuitive model parameters after re-parameterization.
- Unlike LGM, the QGM model captures vol smiles via the CEV-like parameter $\gamma$.

# LGM for Berms: correlations and volatilities

Berms are about choosing the best co-terminals swap amongst the ones defined by the available exercises. Thus one one really cares about capturing these co-terminals swaps'

- correlations (inter-temporal),
- and volatilities.

When using a 1F-LGM to risk-manage a Bermudan swaption, we would use

- the mean-reversion to capture the inter-temporal correlations,
- and the instantaneous volatility to capture the volatilities of the co-terminals.

To set the level of the mean-reversion one can look at different things

- match the price of berms in Totem similar to the one at hand,
- match the inter-temporal correlations generated by a model with a more realistic correlation structure (e. g. MF-LGM, LMM),
- match the vols of swaptions with a shorter tenor, or even caplets.

The latter is a bit debatable, as these quotes are more indicative of the correlations for swaps with different tenors setting at the same time, rather than those of co-terminal swaps setting at different time (e.g. inter-temporal ones).

# LGM for Berms: effect of mean reversion

If we maintain the volatilities of the co-terminals calibrated, i.e. we fix the variance of the state variable governing the model at the various expiries,

- increasing the mean-reversion increases the price of the Berm.

Intuitively, given we're maintaining the overall variance, increasing the mean-reversion makes the process driving the state variable more "forgetful", i.e. we're lowering the inter-temporal correlation, thus gaining more value from having multiple exercise opportunities.

Equivalently, looking at the vols $\sigma(t,u) = \sigma(t)e^{-\lambda(u-t)}$ of the instantaneous forwards, if we want to maintain the overall variance for a swap when increasing $\lambda$ we'll have to increase $\sigma(t)$, which ultimately leads to an increased value for the Berm.

# LGM for Berms: effect of number of factors

It may sound a bit counter-intuitive, but increasing the number of factors would typically reduce the price of the Berm: an increased number of factors reduces the same-time correlations between (co-initial) swaps, but has the opposite effect on the inter-temporal correlation of the co-terminals.

This can be visualised in the context of an LMM model. In that case

- calibrating to caplets fixes the integral of the inst-vol of a given Libor rate,
- a swaption is pretty much a basket of the Libors covered by its underlying swap – the lower the Libors' correlations the lower the swaption's vol.

Now, let's assume we're calibrating to caplets and swaptions. Looking at swaptions with a given expiry, if we lower the Libors' correlation, to maintain the swaptions' vols, we'll have to increase the Libor instantaneous vols covered by these swaptions, but given the constraint given by the caplets that can only be done by transferring the Libors' instantaneous vols from close to the expiry towards the front. This will lower the forward vol, and thus the price of a Berm.

# LGM for Berms: what reference Europeans

In LGM swap rates are pretty much normal; using more factors doesn't change that: it only allows the freedom to have different volatilities for swaps setting at the same time but having different tenors.

Thus, given that one can calibrate to one volatility only (one per tenor in a MF setting), one has to choose judiciously the strike at which this volatility is taken from the Europeans' smile. Sticking to a 1F-LGM, where we'll be using Jamshidian's trick to price the Europeans, here are some possible choices for the strike:

- **ATM**
  Many years ago, liquidity was concentrated ATM – those swaptions would have been the most likely hedging instruments, and probably the only available quotes to calibrate to.

  One ends up with effectively the same model for different products.

  However, when applied to a non-ATM European this method doesn't capture the right price, which is clearly a major limitation.

- **At the strike calibration**
  A simple alternative is to calibrate to the co-terminal European swaptions underlying the Berm, using precisely the strike of the deal.

  Not only this guarantees that the procedure applied to a European replicates the European itself, but it also ensures that the Berm "switch" inequality is satisfied

$$V_{\mathrm{berm}} \geq \max\{V_{\mathrm{European_i}}\}\,. \tag{3}$$

# LGM for Berms: what reference Europeans – exercise-boundary

Let $x_{T_i}$ be the LGM state variable at the expiry $T_i$, one will exercise for every value of $x_{T_i}$ such that

$$\begin{aligned}&\text{vCurrentExercise}(x_{T_i}) = \\ &\quad = \max\left(\text{vCoterminalSwap}(x_{T_i}) - \text{vFutureExercises}(x_{T_i}),\, 0\right) > 0\,. \qquad (4)\end{aligned}$$

One can see that the exercise region is a half-line with boundary $x^{(b)}_{T_i}$.

Calibrating to the swaption stuck at $\text{CoterminalSwapRate}(x^{(b)}_{T_i})$, i.e. to the "exercise-boundary" is a common choice, as that is the point of maximum convexity (we match the kink of "max" with that of the calibration & hedging instrument).

Given that the exercise boundary itself depends on the level of the vol, typically one hopes for the best and runs a fixed-point iteration

1. find initial vols calibrating at the strike,
2. find the exercise boundaries,
3. calibrate vols at the exercise boundaries,
4. repeat from 2. until converged.

Depending on the product and on the shape of the smile, the above may not converge. Also, the condition in Eq. (3) may be violated.

# LGM for Berms: the variable notional case

- We've discussed already that a swap with a time-dependent notional can be seen as a basket of co-initial swaps, where the weights are stochastic, but with a slower dynamics the swaps', so people can typically "freeze" them to their initial state.
- We've also provided a way to express the volatility of the basket as a function if the vols of its constituents and their correlation.

For a Berm on such a swap, at each exercise we'll be looking at exercising into a basket of swaps. As a result, on top of the usual inter-temporal correlation between exercises, we'll be exposed to the correlation of the co-initial swaps making up each of the baskets and their vols.

This problem would require a multi-factor model. Can we try our best in a 1F-setting? We could calibrate the 1F-model to the prices of Europeans on varying notional swaps.

- The resulting calibrated instantaneous volatility would be a synthetic representation of the volatility of the underlying variable notional swap, depending implicitly on the covariance of the co-initial swaps composing the basket.
- The mean reversion would still maintain its role as the control of inter-temporal correlation, and hence the value of the switch option.

# LGM for Berms: zero-coupon notes

In a zero-coupon note, at inception the investor pays the bank a notional $N$ in exchange for a final payment equal to $N$ compounded up at a fixed rate $K$

$$c_n = N \cdot \prod_{i=1}^{n} (1 + \alpha_i K) . \tag{5}$$

Banks often hold a Bermudan option to cancel the note. When the option is exercised, the bank pays the investor the amount accrued up to that point in time, thus the exercise value to the bank will be $c_n P(T_{\mathrm{ex}}, T_n) - c_{\mathrm{ex}}$.

This exercise value can be replicated by a fixed-floating swap on the accreting notional $c_n$, paying $K$ vs the funding rate $F_i(T_{\mathrm{ex}}) = \frac{1}{\alpha_i}(\frac{P(T_{\mathrm{ex}}, T_i)}{P(T_{\mathrm{ex}}, T_{i+1})} - 1)$, because

$$\sum_{i=\mathrm{ex}}^{n-1} c_i \alpha_i (K - F_i(T_{\mathrm{ex}})) P(T_{\mathrm{ex}}, T_{i+1}) = \sum_{i=\mathrm{ex}}^{n-1} [c_{i+1} P(T_{\mathrm{ex}}, T_{i+1}) - c_i P(T_{\mathrm{ex}}, T_i)] .$$

As a result, the value of an Berm on a zc-note can be calculated as that of a Berm on an accreting notional swap, which we've covered in the previous slide.

Well... kind of. Notice how we've been careful not to call $F_i$ as $L_i$.

# LGM for Berms: more details on pricing - direct integration

As mentioned already, pricing of Berms resorts to applying

$$V(t)/N(t) = E_t^{Q^N}\left[V(T)/N(T)\right]$$

repeatedly to "roll-back" from time $T$ to time $t$ the value of the floating, fixed, and cancellation leg. We've also mentioned that for 1F-LGM, rollbacks can be expressed as the convolution

$$V(t,x) = N(t,x)\int \frac{V(T,X)}{N(T,X)}\,\frac{e^{-\frac{1}{2}(X-x)^2/v(t,T)}}{\sqrt{2\pi\, v(t,T)}}\,\mathrm{d}X\,. \qquad (6)$$

The above can be calculated analytically for the functionals representing fixed & floating flows, and even capfloorlets and digis. On the other hand, the cancellation value doesn't admit a simple and analytical expression, so that the convolution has to be performed numerically. One has to pay attention at the discontinuity generated by the max function, so no plain high order integration schemes like Gauss-Hermite can be used. Also, one has to be careful at how the direct integration method behaves for different values of $v(t,T)$.

Appealing when only rolling-back across fairly spaced exercise dates.

# LGM for Berms: more details on pricing - PDE approach

A common approach for performing roll-backs can be obtained as follows. We know $g(t, x_t) = V(t, x_t)/N(t, x_t)$ is a $Q^N$ martingale. Since Itô gives

$$\mathrm{d}g(t, x_t) = \partial_t g \,\mathrm{d}t + \partial_x g \,\mathrm{d}x_t + \frac{1}{2}\alpha^2(t)\,\partial_x^2 g \,\mathrm{d}t \tag{7}$$

for that to be the case the drift will have to be zero, which translates into the following PDE

$$\partial_t g + \frac{1}{2}\alpha^2(t)\,\partial_x^2 g = 0\,. \tag{8}$$

Now one chooses a grid $\{x_i\}$ in the spatial direction, and approximates the above on the grid-points with its finite-difference approximation. Indicating with $\mathbf{g}(t)$ the vector of values of $g(t, x)$ on the grid, one has $\partial_t \mathbf{g}(t) = \mathbf{A}(t)\mathbf{g}(t)$, where $\mathbf{A}(t)$ is a tridiagonal matrix. The solution of the space-discretized problem is then

$$\mathbf{g}(t) = e^{-\int_t^T \mathbf{A}(s)\mathrm{d}s}\mathbf{g}(T) \approx \prod_i \frac{P(\int_i \mathbf{A})}{Q(\int_i \mathbf{A})}\,\mathbf{g}(T)\,, \tag{9}$$

where we've used a rational approximation $\exp(x) = P(x)/Q(x)$ good when $x$ is small, fact we've ensured bu splitting the time-integral in smaller pieces.

- Choosing the rational approx $\longrightarrow$ implicit, explicit, Crank-Nicholson, higher order (different stability, convergence order, propagation of high freqs).
- It's just a numerical approximation of the convolution.
- Various issues with sizing of the grid and discontinuities.

# References

Hagan, Patrick S. and Woodward, Diana E., *Markov Interest Rate Models*, Applied Mathematical Finance, Feb 1999

Hagan, Patrick S, *Evaluating and Hedging Exotic Swap Instruments via LGM*

Hagan, Patrick S., *Methodology for Callable Swaps and Bermudan "Exercise Into" Swaptions*, Jan 2004