A PRACTITIONER'S GUIDE TO PRICING AND HEDGING CALLABLE LIBOR EXOTICS IN FORWARD LIBOR MODELS

VLADIMIR V. PITERBARG

ABSTRACT. Callable Libor exotics is a class of single-currency interest-rate contracts that are Bermuda-style exercisable into underlying contracts consisting of fixed-rate, floating-rate and option legs. Bermuda swaptions, callable inverse floaters and callable range accruals are all examples of callable Libor exotics. It is commonly agreed that these instruments are best modeled using forward Libor models. There are many problems, both technical and conceptual, that arise when applying forward Libor models to callable Libor exotics. These problems span calibration, valuation and computation of risk sensitivities. This paper, to the best of our knowledge, is the first comprehensive overview of calibration, pricing and Greeks calculation techniques for callable Libor exotics in forward Libor models. Many technical results and practical methods presented in the paper are original. Others are adaptations, generalizations and extensions of known approaches. Among the technical contributions of this paper are the recommendations for basis functions for the Longstaff-Schwartz valuation algorithm, the extension of the pathwise differentiation method to callable Libor exotics and elegant Greeks formulas that result, novel smoothing techniques for Monte-Carlo, application of Markovian approximations and PDE methods to the problem of variance reduction, and practical algorithms for obtaining vegas in forward Libor models. In addition, strategies for calibrating forward Libor models for callable Libor exotics are discussed at length.

CONTENTS

1. Introduction 2
2. Literature review 4
3. Callable Libor exotics 5
3.1. The market 5
3.2. The underlying instrument 6
3.3. The callable structure 7
3.4. Examples 7
4. Volatility calibration for callable Libor exotics 8
4.1. Introduction to calibration 8
4.2. Volatility risk factors for CLE 9
4.3. Why use a forward Libor model 10
4.4. Choosing between a "fully calibrated" and a "realistic" model 11

Date: June 25, 2003.

Key words and phrases. Bermuda-style derivatives, Bermudan swaptions, callable Libor exotics, callable range accruals, callable inverse floaters, hedging, Greeks, deltas, vegas, gammas, Monte-Carlo, market model, forward Libor model, Libor market model, LMM, BGM, pathwise deltas, Markov approximation, variance reduction, control variate, smoothing.

I am indebted to Leif Andersen for a great deal of help, encouragement and support in writing this article. I also thank my current and former colleagues Paul Cloke, Phil Hinder and Jesper Andreasen for stimulating discussions, and for asking challenging questions. Participants of 2003 Risk Training course in New York provided many insightful comments. All remaining errors are mine.

4.5. Volatility smile 13
5. Forward Libor models 13
6. Valuing callable Libor exotics in a forward Libor model 15
6.1. Recursion for callable Libor exotics 15
6.2. Monte-Carlo for American-style options 16
6.3. Differences in the algorithm for CLEs vs. Bermuda swaptions 18
6.4. Choosing explanatory variables and parametric families 18
6.5. An upper bound in Monte-Carlo simulation 22
7. Risk sensitivities of callable Libor exotics: a short overview 23
8. Exercise boundary and risk sensitivities 24
9. Why CLE Greeks are hard to compute 24
10. Computing deltas in the same simulation 25
10.1. Pathwise deltas 26
10.2. Pathwise deltas of callable Libor exotics 27
11. "Sausage" Monte Carlo 28
11.1. The outline of the method 28
11.2. A model problem 29
11.3. "Sausage" Monte-Carlo for forward Libor models 30
12. Low-dimensional Markovian approximation for a forward Libor model 33
12.1. Deriving an approximation 33
12.2. Discussion of the approximation 35
12.3. Two-dimensional extension 35
13. Markovian approximation as a control variate 37
13.1. Control variate 37
13.2. Model control variate 38
13.3. How to do the Markovian approximation right 39
14. Vegas 40
14.1. Volatility calibration for forward Libor models 40
14.2. The direct method of computing vegas 41
14.3. The indirect method of computing vegas 42
14.4. Computing model vegas 45
15. Conclusions 45
References 45
Appendix A. "Sausage" Monte Carlo derivation 47
Appendix B. Proof of approximate conditional independence for "sausage" Monte-Carlo 49
Appendix C. Figures 50

1. Introduction

Callable Libor exotics are among the most challenging interest rate derivatives to price and risk-manage. These derivative contracts are loosely defined by the provision that the holder has a Bermuda-style (i.e. multiple-exercise) option to exercise into various underlying interest rate instruments. The instruments into which one can exercise can be, for instance, interest rate swaps (for Bermuda swaptions), interest rate caps (for captions, callable capped floaters, callable inverse floaters), collections of digital call and put options on Libor rates (callable range accruals), collections of options on spreads between various CMS rates (callable CMS coupon diffs), and so on.

From a modeling prospective, callable Libor exotics are difficult to handle. Simple, “first generation” interest rate models like Ho-Lee, Hull-White, Black-Karasinsky cannot be used because of their inability to calibrate to a rich enough set of market instruments. One has to use “second generation” models with richer, more flexible volatility structures. Among the latter, forward Libor models (also known as Libor Market and BGM models) are arguably the best suited for the job.

Building a pricing and risk management framework for callable Libor exotics based on forward Libor models is a formidable task. Conceptual and technical issues abound. Calibration, valuation and risk sensitivities calculations all present unique challenges.

This paper was written to address these challenges. We aim to present a comprehensive review of problems one has to deal with when developing the callable Libor exotics capability for forward Libor models. We also give solutions to these problems. The solutions fall into two categories. Some are known methods and techniques that we adapted. Others have been developed by us for problems that did not have good solutions.

To the extent that we aim for comprehensive coverage, this paper can be considered a review. Unfortunately, in the atmosphere of competing and sometimes secretive banks and financial institutions, a review of *practical* solutions to *real-world* problems is hard to do. While theoretical insights float more or less freely across different groups working on similar problems, same is not necessarily true for practical solutions and “tricks”. The latter may be perceived to confer competitive advantages, and as such are not widely publicized. For that reason, the solutions, methods and techniques presented in this paper should be taken with a grain of salt. These are not necessarily the best solutions available; these are the best solutions among those that were available to (or developed by) us at the time of writing.

Our goal in writing this paper was to make it interesting for as wide an audience as possible. Seasoned practitioners working in the area of interest rates will hopefully find the methods and techniques in this paper interesting and useful. (At the very least they can derive a certain amount of glee if they realize that their own answers to the same problems are better than ours!) Those who are just starting out in the area will find this paper a good foundation to build on. While the focus of our work is practical, we build it on a solid theoretical foundation (most of the time, anyway). That should make our work interesting to academics. They may also benefit from reading about problems that practitioners face in their day-to-day work.

While we try to write for a wide audience, we do expect a certain level of technical competence. We do not dwell on the basics of interest rate modeling, forward Libor models, volatility calibration, and interest rate markets in general.

To get an idea of the scope of the paper, one just needs to ask why using forward Libor models for callable Libor exotics (CLE for short) is so hard? Problems start with volatility calibration. Multiple types of optionality embedded in CLEs mean that they depend on volatilities of many different rates. What market instruments do we calibrate the model to? Matching today’s prices of market instruments is just part of the story. Choices affecting the dynamics of the volatility structure have a significant impact on CLE prices. How do we impose the “right” dynamics? What *are* the right dynamics? These questions, why falling in the domain of general interest rate modeling, have profound importance for callable Libor exotics, and we address them first. This part of the paper is not very technical, somewhat argumentative and can probably even be somewhat controversial.

After calibration comes pricing. Pricing must be done using Monte-Carlo, as it is the only viable numerical method available for forward Libor models. Successful pricing of Bermuda-style options in Monte-Carlo hinges on the ability to formulate good rules for choosing exercise strategies. For instruments as complex as callable Libor exotics, what are they? How do we make sure we are not significantly underpricing CLE's because our exercise boundaries are lousy? For pricing, the issues of speed and accuracy must also be addressed. A Monte-Carlo valuation is typically quite slow. What methods do we use to speed valuation up and/or make it more accurate? What variance reduction techniques work and what do not?

The Longstaff-Schwartz algorithm that we use for pricing is not new. It has been developed for Bermuda swaptions elsewhere. We extend the method to callable Libor exotics and discuss in detail how to choose good exercise strategies. As speed/accuracy issues have a much bigger impact on computing Greeks, we discuss those issues in the context of computing risk sensitivities.

As we move into the realm of Greeks calculations, problems become really hard. Obtaining good, clean and robust risk sensitivities from a Monte Carlo-based model is one of the hardest practical problems. First, why are the numerical properties of Greeks so much worse for CLEs than for other, seemingly related instruments? What methods do we use to obtain good deltas? What is a usable definition of a vega in a forward Libor model? How vegas can be computed? The methods developed here comprise the bulk of technical information presented in the paper, and are probably the most interesting overall. We explain various techniques that we know really work. We also discuss those that seem like they should work but do not.

2. Literature review

The body of literature dealing with callable Libor exotics is pretty scarce, and that was one of the motivations for writing this article. Definitions of different kinds of callable Libor exotics can be found here and there but, to the best of our knowledge, there is no comprehensive survey of the subject.

Bermuda swaptions are one kind of callable Libor exotics that have been studied extensively. Good understanding of issues around pricing and risk managing Bermuda swaptions is an important prerequisite for dealing with more general callable Libor exotics. A fair amount of subject knowledge is presented in textbooks ([BM01], [Reb02]). More extensive coverage is available from research papers ([And01], [Ped99], [And01], [AA01], [And99], [Dod02], [LSSC99] [PP03], [Sve02], to name a few).

A good grasp of general interest rate modeling is a must for anyone who is interested in modeling callable Libor exotics. Especially important are considerations that relate to volatility structure dynamics. Basics are well-covered in book such as [Reb98], [Reb99], [Reb02], [BM01], as well as other interest rate derivative textbooks. More advanced discussions are presented in papers such as [AA01], [LSSC99], [Reb03], [Sid00] and [RS95].

We discuss using forward Libor models for CLE pricing and risk management. This type of models has seen a lot of academic and practical interest since they were first introduced in [BGM96] and [Jam97]. Any modern textbook, including those cited above, will contain the necessary basics. A more theoretical development of the models is presented in [MR97].

A more practical approach, with important extensions, is discussed in considerable detail in [AA00] and [ABR01].

Volatility calibration of forward Libor models is an important topic and is directly related to the main theme of this paper. Details on volatility calibration can be found in Rebonato's and Brigo and Mercurio's books, as well as in papers [Ped98], [SC00], [LM02], [Ale02] and [Gat02]. The most recent paper on the subject with important advances is, to the best of our knowledge, [Wu03].

Valuing American-style derivatives in Monte-Carlo simulation is an established topic of research these days. Important contributions in the area include [LS98], [Rog01], [BG97]. Papers [And99], [AB01], [Ped99] deal with valuation issues in the context of forward Libor models. None are specific to callable Libor exotics, a gap we fill in this paper.

Obtaining good risk sensitivities in a Monte-Carlo based model is one of the main themes that runs through our paper. The problem has attracted a lot of attention over the years, with papers [BBG97], [GZ99], [FLLT99], [Ber99] (to name a few) proposing various ways of attacking it. None of these papers focused on American-style options, however. We present an array of techniques for obtaining good risk sensitivities that are targeted specifically at callable Libor exotics, techniques that utilize the structure of the instruments to their full advantage.

A number of researchers have studied the possibility of using lattice, or PDE, methods for forward Libor models, see e.g. [Nou99], [PPvR02]. By themselves, PDE methods as presented in these papers have severe limitations that make them useless in applications to callable Libor exotics. In this paper we demonstrate the way of making the PDE methods useful by building a variance reduction scheme around them.

Another subject we cover is measuring sensitivities of callable Libor exotics to volatilities. The problem was also considered in [PP03]. Only "diagonal" vegas were considered in that paper – an approach rather inadequate for risk managing derivatives with as complicated volatility dependence as callable Libor exotics. We present a much more general and practical approach.

3. Callable Libor exotics

In this section we discuss the market for callable Libor exotics, motivation behind these types of contracts, and define a universe of instruments we will be dealing with. We will simplify some of the definitions for brevity.

3.1. **The market.** The market in callable Libor exotics has developed in response to an increasing sophistication of corporate and institutional clients in tailoring their interest rate exposures to specific views and objectives. Also, an appetite for above-market current yield, especially in the falling interest rate environment with few attractive investment alternatives, has prompted clients to sell increasingly more complex (and more valuable) options to interest rate dealers.

Callable Libor exotics typically start their life as bonds, or notes, sold by banks to investors (institutional investors such as hedge funds and investment companies, corporate clients, and even wealthy individuals). These notes typically offer high initial coupon and, after some initial period, coupons linked to various interest rates in non-trivial ways. For example, a coupon can be equal to an observed Libor rate plus a spread, capped at a certain level. In return for these attractive features, investors give the issuing bank an option to call the bond at certain dates in the future (after a "lockout" period). For example, the bank has the right to call the bond every year until maturity, starting on the first anniversary of the bond. When the bond is called, the investor receives the principal back and stops receiving the coupons.

The investor's interest in such structures is clear – they get an above market initial coupon and/or (perceived) attractive payoffs from structured coupons down the line. For its part, the bank has an option to cancel this deal (in effect, an option *to enter a reverse transaction*), an option that can be very valuable. The bank monetizes this option by delta and vega hedging it throughout its life, hoping that its initial outlays to an investor will be recouped as hedging profits.

The higher the price of the option to cancel, the higher the coupon that the bank can pay, and the more attractive this deal becomes to investors. From these considerations it is clear that banks are interested in designing more and more esoteric structures that provide them with more and more valuable options. This drives market innovation (and keeps Quants employed).

From the investor's prospective these structures look like (callable) bonds. From the issuing bank's prospective they are Bermuda-style options to enter a swap in which the bank receives a complicated coupon and pays Libor rate (potentially with a spread). The Libor-rate leg enters the picture here as funding, i.e. a floating-rate income stream from the principal that the investor gave the bank when they bought the bond.

Getting slightly ahead of ourselves we denote the expected value of any random variable under the pricing (risk-neutral) measure by $\mathbf{E}$, and the appropriate numeraire by $\{B_t\}_{t=0}^{\infty}$. With these notations, the value at time t of a payoff paying X at time T is given by

$$B_t \mathbf{E}_t \left(B_T^{-1} X \right).$$

A callable Libor exotic is based on a tenor structure, a sequence of times spaced roughly equally apart,

$$0 = T_0 < T_1 < \cdots < T_N.$$

3.2. **The underlying instrument.** Now let us define callable Libor exotics formally. First we specify the underlying instrument for the Bermuda-style option. The underlying instrument is a stream of payments $\{X_i\}_{i=1}^{N-1}$. Each X_i is determined (fixed) at time T_i (in financial parlance, T_i is a *fixing date* for X_i). The payment is made at time T_{i+1} (so T_{i+1} is a *payment date* for X_i).

A callable Libor exotic is a Bermuda-style option to enter the underlying instrument on any of the dates $\{T_i\}_{i=1}^{N-1}$. If the option is exercised at time T_n, then the option goes away and the holder receives all payments X_i with $i \geq n$ (i.e. all payments with fixing dates on or after the exercise date). A payment at time T_i is defined as a coupon C_i minus a funding rate (which we assume to be the Libor rate F_i for simplicity),

$$X_i = \delta_i \times (C_i - F_i).$$

Here we have defined δ_i to be the day fraction for the period $[T_i, T_{i+1}]$, usually

$$\delta_i \approx T_{i+1} - T_i,$$

(we assumed for simplicity that the day fractions for the coupon leg and the Libor leg are the same).

Note that even though we say the holder receives all payments after a certain date, some of the payments can be negative, which means he has to pay those amounts to the counterparty.

We denote by $E_n(t)$ the n-th exercise value, i.e. the value of all payments one enters into if the callable Libor exotic is exercised at time T_n. Clearly

$$E_n(t) = B_t \sum_{i=n}^{N-1} \mathbf{E}_t \left(B_{T_{i+1}}^{-1} X_i \right).$$

For completeness we set

$$E_N(t) \equiv 0.$$

If a callable Libor exotic is exercised on T_n, the holder receives $E_n(T_n)$, the remaining part of the underlying.

3.3. **The callable structure.** For future considerations it is important to define a whole family of "nested" callable contracts. By $H_n(t)$ we denote the value, at time t, of a callable Libor exotic that has only the dates $\{T_{n+1}, \ldots, T_{N-1}\}$ as exercise opportunities. In particular, $H_0(0)$ is the value of the callable contract we are interested in at time zero. Necessarily

$$H_0(t) \geq H_1(t) \geq \cdots \geq H_{N-2}(t).$$

We call each H_n a "hold" value. The value of the choice of *not* exercising on date T_n ("holding") is equal to $H_n(T_n)$, hence the name.

3.4. **Examples.** Here we present a few examples of callable Libor exotics. They differ by the type of coupons C_i. As will be clear from the examples, underlying instruments for most callable Libor exotics can be described as streams of European style options on some reference rates (either Libor or CMS). For a coupon that fixes at time T_i, we denote the rate to which it is linked to by F_i (a Libor or a CMS rate that is observed, or fixed, at time T_i).

See Figure 1 for payoff diagrams.

3.4.1. *A Bermuda swaption.* A simplest example is a Bermuda swaption. The underlying instrument is a plain vanilla fixed-for-floating swap. In particular, each coupon is of the form

$$C_i = c,$$

where c is a fixed rate.

3.4.2. *A callable capped floater.* In a callable capped floater, the coupon is a floating rate with a spread, capped from above. If the cap is c and the spread is s, the i-th coupon C_i is given by

$$C_i = \min\left[F_i + s, c\right].$$

3.4.3. *A callable inverse floater.* In a callable inverse floater, the coupon is based on an inverse of a floating rate (capped and floored). If k is the strike, f is the floor and c is the cap of the inverse floating payment, the i-th coupon C_i is given by

$$C_i = \min\left[\max\left[k - F_i, f\right], c\right].$$

3.4.4. *A callable range accrual.* In a callable range accrual, a payment is based on a number of days that a reference rate is within a certain range. While the range observations are typically performed daily, for notational simplicity we assume that there is only one range observation on the fixing date. In particular,

$$C_i = c \cdot 1_{\{F_i \in [l,b]\}}.$$

Here c is the fixed rate for a range accrual payment, l is the lower range bound, b is the upper range bound.

3.4.5. *A callable capped CMS floater, a callable inverse CMS floater, a callable CMS range accrual.* These are variations of the contracts discussed above with F_i being a CMS (also known as a forward swap) rate that fixes on T_i.

3.4.6. *A callable CMS spread.* The underlying instrument for this contract consists of payments linked to a spread between two different forward swap rates. If $S_{i,1}(t)$ is one such rate (for example a 10 year CMS rate) and $S_{i,2}(t)$ is another such rate (for example a 2 year CMS rate), then the coupon X_i is given by

$$C_i = \max\left[\min\left[S_{i,1}(T_i) - S_{i,2}(T_i), c\right], f\right].$$

Here c and f are a cap and a floor on the spread $S_{i,1}(T_i) - S_{i,2}(T_i)$ between the two CMS rates.

4. Volatility calibration for callable Libor exotics

4.1. **Introduction to calibration.** The first step in using forward Libor models for callable Libor exotics is calibration. Being a Heath-Jarrow-Merton type model, a forward Libor model is defined by volatilities that it imposes on various rates. The process of choosing these volatilities is called *volatility calibration* (or just *calibration*).

We do not discuss technical aspects of forward Libor model calibration here. By now it is a standard procedure, described in great detail in many books and papers (see Literature Review). We do touch on some technical aspects later in the paper, in the context of vega calculations. This section deals with conceptual questions. It is not about *how* to calibrate a model. It is about *what* to calibrate it to.

There are no precise, mechanical rules one can follow to calibrate a model for CLEs. It is more of an art than a science. What we present here are guidelines, not recipes.

The volatilities of rates as specified by the model (the model's *volatility structure*) are chosen to match various targets. These targets can be

- Market prices of liquid interest rate derivatives;
- Modeler's beliefs about interest rate volatilities;
- Historical information about volatilities.

By far, the first type of targets, the prices of liquid derivatives, is the most important. For pricing CLEs, relevant liquid derivatives consist of options on Libor rates (Eurodollar options and caps/floors of various expiries) and options on swap rates (swaptions of various expiries and on swap rates of different maturities). Prices of these instruments are usually quoted as implied Black volatilities. Often, when talking about calibration, it is the recovery of this market-implied Black volatilities within the model that is considered. The main decision here is to choose what swaption/caplet volatilities to use in calibration.

A model, once its volatility structure is specified, *will imply certain evolution of volatilities of market instruments (caps/floors/swaptions) in time.* A volatility of a market instrument at some future time, as given by the model, is usually called a *forward volatility.* In models where state variables have deterministic volatilities, computing forward volatilities is typically straightforward. In other types of models one usually have to specify how the interest rate curve will look like on the future date for the concept of a forward volatility to be uniquely defined.

The modeler's beliefs are often expressed in terms of forward volatilities. For example, a reasonable choice may be to require that a forward volatility for a particular option be the same or close to the current (spot) volatility of the option with the same time to expiry. This corresponds to time-homogeneity of volatility structure.

One can also look at the historical behavior of volatilities and require that forward volatilities followed the same behavior. This way, statistical information on volatilities can be incorporated.

Spot volatilities of market instruments define the current snapshot of market volatility information. Forward volatilities define the evolution, or dynamics, of market volatilities (as predicted by the model). Straddling this division are correlations of different rates. Within a model, any two rates will have a certain correlation. There are different notions of correlations. The one of particular interest to us, for reasons to be revealed later, are the so called *serial correlations.* A serial correlation is a measure of dependence between two rates each observed on its fixing date. For example, for two swap rates $S_1(t)$ and $S_2(t)$, with fixing dates t_1 and t_2 correspondingly, the serial correlation is some measure of dependence between two random variables $S_1(t_1)$ and $S_2(t_2)$. The exact measure of dependence used is model-dependent. For log-normal (exactly or approximately) rates we use

$$\operatorname{corr}\left\langle \log S_1(t_1), \log S_2(t_2)\right\rangle.$$

Swap rates can be thought of as being weighted combinations of Libor rates. Therefore, certain (implied) correlation information is available in market prices of swaptions. Extracting this information is, however, very hard.

Another division of volatility parameters is into observable and unobservable ones. Observable are spot volatilities of various market rates (Libor/swap rates) as they can be implied from observed market prices of caps/swaptions. Forward volatilities and correlations are, on the other hand, unobservable. One's beliefs and statistically observed relationships are imposed on unobservable parameters; observable ones are just implied from the market.

The evolution of the volatility structure in time, under the assumption that the interest rates roll down the forward interest rate curve, is the subject of volatility structure modeling. The (much harder) problem of understanding the evolution of the volatility structure through time across different scenarios for future interest rates is the subject of volatility smile modeling. Volatility structure dynamics is understood much better than volatility smile dynamics in the context of interest rate models; we focus on it first.

Forward volatilities and correlations are extremely important for callable Libor exotics. It will probably not be an overstatement to say that the *main difficulty in modeling callable Libor exotics lies in their strong dependence on unobservable volatility parameters (time evolution of the volatility structure as expressed by forward volatilities, and correlations).*

4.2. **Volatility risk factors for CLE.** We start the discussion of the proper ways of volatility calibration for callable Libor exotics by focusing on what sorts of volatility risks a CLE is typically exposed to. For the sake of concreteness, we consider a callable inverse floater. Exercise is allowed on dates

$$T_1 < T_2 < \cdots < T_{N-1}.$$

The callable inverse floater's strike is equal to 6%, the cap is at 4% and the floor is at 0%. The coupon's payoff at time T_i is equal to

$$C_i = \min\left(\max\left(6\% - F\left(T_i\right), 0\right), 4\%\right).$$

Here $F\left(T_i\right)$ is the Libor rate observed at T_i.

The coupon can be decomposed into a portfolio of a long floorlet with strike 6% and a short floorlet with strike 2%. So, the callable inverse floater can be thought of as a Bermuda-style option on a combination of floors and a Libor leg.

Spot volatilities that this contract depends on are easy to discern. The underlying consists of two floors. Thus, spot volatilities of appropriate Libor rates are important. (Note that two different strikes are involved).

By focusing on each exercise date at a time, we see that the CLE "contains" a European-style option to enter the underlying swap. Even though the underlying swap is not vanilla (i.e. not fixed-rate-for-floating-rate), it is clearly related to one. So, the term volatility of the swap rate that fixes on T_i, and runs for the period $\left[T_i, T_N\right]$, is clearly important (it is less clear what strike should be used for this term volatility; all strikes are in fact relevant).

To summarize, this CLE is dependent on term Libor volatilities (all expiries until T_{N-1}), and term swap rate volatilities for those swaptions for which expiry+maturity is equal to T_N (so called *core*, *diagonal*, or *co-terminal* swaptions).

This is probably all as far as the spot volatility structure is concerned. Is there any dependence on the forward volatility structure? Yes. Imagine we have not exercised the contract until time T_n, $n < N-1$, and it is now exercise time T_n. At this point in time, we need to decide whether to exercise and receive $E_n\left(T_n\right)$, or hold on to the deal, a decision that is worth $H_n\left(T_n\right)$. The value of the remaining underlying depends on caplet volatilities *as observed at time* T_n, i.e. forward Libor volatilities. Likewise, the option to hold, a Bermuda-style option on the underlying, will depend on core swaption volatilities also *as observed at time* T_n. These are forward swaption volatilities. So the exercise decision at time T_n depends on the forward volatility structure at time T_n and hence, the value of the CLE today will depend on it as well.

A Bermuda-style option to enter the underlying swap can be thought of as being some kind of a "best of", or a "switch", option. A holder tries to choose the best of multiple alternatives. Clearly, the less correlation there is between the alternatives, the more value there is in a option to choose the best one. By again associating the underlying into which we can enter on exercise date T_n with a swap rate that fixes at T_n and runs for $\left[T_n, T_N\right]$, we see that the "switch" option depends on serial correlations of all these swap rates.

While we make a distinction between forward volatilities and serial correlations, they are in fact intimately related. In fact, in simple models like the Hull-White model, the two can be expressed via each other.

4.3. **Why use a forward Libor model.** It is now appropriate time to discuss why we insisted on using a forward Libor model, arguably one of the most complicated models there are, to price callable Libor exotics, and why we ruled out simpler models. It is the strong dependence of callable Libor exotics on forward volatility structure and other unobservable volatility parameters such as correlations that is the reason. A simpler, lower dimensional model might succeed in fitting spot volatility information (such as observed Libor rate and swap rate volatilities). However, it can typically achieve so only by using extremely contrived combinations of model parameters, that imply completely unrealistic evolution of the volatility structure. Having no control over correlations produced by the model, or the evolution of the volatility structure in a simpler model is too much if a price to pay for ease of valuation.

As we mentioned before, correlation information is contained in swap rate volatilities, but it is not easy to extract. A forward Libor model, however, can do just that. By calibrating the model to the whole set of swaption and cap volatilities, we find a consistent model that incorporates all available market volatility information. Correlations implied by such a model are, in a sense, the most accurate (implied) correlations one can hope to obtain. In the same spirit, a model calibrated to all swaptions and caps gives us the best available *implied* forward volatility structure, i.e. information, consistent with the market prices, of how the volatility structure will move in the future.

One should think of a forward Libor model as a tool to extract unobserved volatility information (correlations and forward volatilities) from observed volatilities of swaptions and caps.

Some people do not subscribe to the notion that one should let a model, however great, to extract forward volatilities from the market. They prefer to control it directly (more on that later). A forward Libor model is flexible enough to incorporate external beliefs about how the volatility structure should evolve (sometimes at the expense of not calibrating to some swaptions). As we mentioned before, homogeneity of the volatility structure, exact or approximate, is a popular target.

In short, the ability to extract unobserved volatility information from the observed one, and the ability to control the dynamics of volatility structure within the model (both being very important for callable Libor exotics), are the main reasons for using forward Libor models for callable Libor exotics.

4.4. **Choosing between a "fully calibrated" and a "realistic" model.** While the idea that among all models, forward Libor models are the most appropriate for callable Libor exotics is generally accepted, there are different opinions as to how it should be calibrated. (see e.g. Rebonato's book [Reb02] and survey [Reb03]) Here we present an outline of the arguments and our own opinion on the subject.

One school of thought strongly believes in calibrating a forward Libor model to the full set of available volatility instruments (all caps and swaptions). Some influence on the evolution of the volatility structure within the model can still be incorporated. The other camp advocates judiciously choosing a subset of caps/swaptions to which to calibrate, and putting significant emphasis on specifying the dynamics of the volatility structure in which one believes in (typically imposing strong time-homogeneity assumptions or observed statistical relationships).

We call the first approach "fully calibrated" and the second one "realistic". Our names should not be taken to indicate that fully-calibrated models are not realistic. The names just reflect the primary focus of the chosen approach to calibration.

The pros of the "fully calibrated" approach are as follows. All liquid volatility instruments are consistently priced within the model. This is very important for hedging, in particular *static* hedging. As one will undoubtedly use various swaptions to hedge a callable Libor exotic, mispricing them in the model would generally be a bad idea. Also, correlations between different rates are fully *implied*, meaning there is none (or little, depending on one's implementation) human judgement involved in specifying them. This is a big plus from traders' prospective who are not forced to come up with correlation estimates, as well as for risk managers naturally suspicious of any unobserved market parameter set by traders at whim.

The "realistic" approach has its strong sides as well. One should not forget that the price of an instrument in a model is equal to the model's predicted cost of re-hedging the instrument over its lifetime. Hedging profits in the future as specified by the model are directly related to the volatility structures that the model predicts for the future. For these model-predicted hedging profits to have any resemblance to the actual realized hedging profits, the dynamics of the volatility structure in the model should be a reasonable estimate of the actual dynamics of the volatility structure. Our best estimate of the future is probably what we have today (or have estimated historically), and that's why it is important to make sure that the model's evolution of volatility structure is as close to being time-homogeneous as possible. Another way to argue the same thing is to say that if the actual volatility structure in the future is very far from what was assumed by the model it would be, a trader would incur substantial volatility re-hedging costs.

The strong points of the "realistic" approach are of course the weak ones of the "fully calibrated" approach. Forward volatilities coming out of the fully calibrated model can exhibit non-stationary behavior, impairing the performance of dynamic hedging.

Likewise, mispricing certain swaptions in the "realistic" approach is troublesome. In a pragmatic view of a model as an "interpolator" that computes prices of complex instruments from prices of simple ones, if simple instruments are mispriced, how can there be any confidence that the complex ones are priced correctly? It is generally very hard to make a judgement as to what swaption volatilities are relevant for a particular callable Libor exotic (not surprisingly, given that correlation information is "spread" over all swaptions), which makes this approach so much harder to defend.

It is easy to imagine taking either approach to the extreme to generate completely silly results. The "correct" approach, as often the case, lies somewhere in the middle. We lean towards the "fully calibrated" approach. Full calibration should however be coupled with a rigorous checking of the effect of all assumptions on pricing and hedging results that come out of the model. Among the things whose impact one needs to check are

- Number of factors used;
- Relative importance of recovering all swaption prices vs. time-homogeneity of the resulting volatility structure, as specified during calibration;
- Correlation assumptions on forward Libor rates;
- Any other parameter that can "move".

The checks are usually performed by varying these parameters within "reasonable" limits and making sure that the impact on pricing/hedging is limited. If the impact of a particular parameter is large, it should raise a red flag.

One should always keep in mind that no model, even as sophisticated as a forward Libor model, is ever "correct". The best we can do is to have a model that we believe is the best possible, and (importantly) put some "confidence intervals" around it to indicate how wrong it can actually be.

4.5. **Volatility smile.** Volatility structure and its dynamics play a fundamental role in modeling callable Libor exotics. Volatility smile (dependence of volatility structure on interest rates) and its dynamics is just as important. Just as with volatilities, we divide volatility smile information into current, or spot (observed from the market) and forward (predicted by the model as time/rates move). The reasons why the volatility smile (spot and forward) is fundamentally important are basically the same as for the volatility structure. The spot volatility smile (a whole structure of them as a matter of fact, a there is one per each caplet/swaption) defines today's values of the underlying swap, and of the European options to enter portions of it in the future. This dependence is amplified by the fact that relevant strikes are usually deep in or out of the money, for which the volatility smile effect is the strongest. Forward volatility smiles define relative values of the exercise and hold pieces at the time of exercise and thus affect the exercise decision (and the today's price) for a callable Libor exotic. The interest rate levels at which it is optimal to exercise are usually significantly different from the today's levels. Therefore, the relationship between interest rates and strikes at future times on the exercise boundary is very different from today's. For a very simple example consider a situation where today's spot rate is 3%, strike on the underlying is 5%, and it is optimal to exercise when the rate is at 8%. Suppose that our model can price an instrument which is 2% out of the money today; at the time of exercise it will have to price an instrument that is 3% *in* the money. It is clear that the issue of how the smile moves with time *and* the level of rates is important for callable Libor exotics.

The state of the volatility smile modeling for interest rates is not nearly as advanced as that of the volatility structure modeling. In all fairness, it is technically a much more complicated problem to tackle. Because of that, the impact of various choices of smile dynamics on prices of callable Libor exotics (even as simple as Bermuda swaptions) is poorly understood. State of the art in smile modeling currently involves including stochastic volatility or jumps (or both) in the process for forward Libor rates, typically chosen to be the same for all Libor rates, homogeneous through time and with very simple volatility rate dependence. We present a stochastic volatility forward Libor model later. An example of a volatility smile that this model produces, versus the market smile and versus a simple displaced-diffusion type smile, is given in Figure 2.

5. Forward Libor models

Throughout the paper, Actual/Actual day counting convention is assumed for simplicity, i.e. all day counting fractions are equal to the period length as a fraction of a calendar year. A zero coupon bond paying one dollar at time T, as observed at time t, $t \leq T$, is denoted by $P(t,T)$. A forward Libor rate for the period $[T,M]$, as observed at time t, is defined by

$$F(t,T,M) \triangleq \frac{P(t,T) - P(t,M)}{(M-T)\,P(t,M)}.$$

A probability space $(\Omega, \mathcal{F}, \mathbf{P})$ is chosen, together with a sigma-algebra filtration $\{\mathcal{F}_t\}_{t=0}^{\infty}$.

Different flavors of forward Libor models are available. Not striving for completeness, we present three different choices. They are different in the ways they deal with the volatility smile. The models we consider are the standard log-normal model ([BGM96] and [Jam97]), a skew-extended one ([AA00]) and a stochastic volatility one ([ABR01]).

Let

$$\begin{aligned} 0 &= t_0 < t_1 < \cdots < t_M, \\ \tau_n &= t_{n+1} - t_n, \end{aligned}$$

be a tenor structure, i.e. a collection of approximately equally spaced (three or six months is common) maturities.

Note that the model's tenor structure $\{t_i\}_{i=0}^M$ is potentially different from contract-specific tenor structure $\{T_i\}_{i=1}^N$.

We define the n-th forward Libor rate $F_n(t)$ (the n-th "primary" Libor rate) by the expression

$$F_n(t) \triangleq F(t, t_n, t_{n+1}) = \frac{P(t, t_n) - P(t, t_{n+1})}{\tau_n P(t, t_{n+1})\,.}, \quad 0 \le n < M.$$

The log-normal model is specified by the following dynamics of each of the forward Libor rates,

$$dF_n(t)/F_n(t) = \lambda_n(t)\, dW^{T_{n+1}}(t), \quad n = 1, \ldots, M-1, \quad t \in [0, t_n], \tag{5.1}$$

Here $\lambda_n(\cdot)$ is a deterministic function of time $\mathbb{R}_+ \to \mathbb{R}^n$, and $dW^{T_{n+1}}(\cdot)$ is a one-dimensional Brownian motion under the T_{n+1}-forward measure. (We consider a one-dimensional case for brevity.)

The skew-extended forward Libor model introduces a local volatility function $\phi(x)$, independent of time, that is applied to each of the Libor rates. The dynamics (under the appropriate forward measures) is given for each F_n by

$$dF_n(t) = \lambda_n(t)\, \phi(F_n(t))\, dW^{T_{n+1}}(t), \quad n = 1, \ldots, M-1, \quad t \in [0, t_n]. \tag{5.2}$$

A popular choice for $\phi(x)$ is a linear function

$$\phi(x) = ax + b,$$

resulting in a "displaced diffusion" type model.

To obtain a stochastic volatility model, we first introduce a process for the stochastic variance

$$\begin{aligned} dz(t) &= \theta(z_0 - z(t))\, dt + \varepsilon\sqrt{z(t)}\, dZ(t), \\ z(0) &= z_0. \end{aligned}$$

Here θ is the mean reversion of variance, ε is the volatility of variance. We assume that $Z(\cdot)$ is independent of the Brownian motions that drive the rates. We add the stochastic volatility on top of the skew-expended forward Libor model to obtain

$$dF_n(t) = \sqrt{z(t)}\lambda_n(t)\, \phi(F_n(t))\, dW^{T_{n+1}}(t), \quad n = 1, \ldots, M-1, \quad t \in [0, t_n]. \tag{5.3}$$

For convenience we define

$$F_n(t) = F_n(t_n), \quad t > t_n.$$

A special numeraire is usually chosen. We define a discrete money-market numeraire B_t by

$$\begin{aligned} B_{t_0} &= 1, \\ B_{t_{n+1}} &= B_{t_n} \times (1 + \tau_n F_n(t_n)), \quad 1 \le n < M, \\ B_t &= P(t, t_{n+1}) B_{t_{n+1}}, \quad t \in [t_n, t_{n+1}]. \end{aligned}$$

The dynamics of all forward Libor rates under the same measure, the measure associated with B_t, are given by (for the stochastic volatility case)

(5.4)

$$dF_n(t) = z(t) \lambda_n(t) \phi(F_n(t)) \sum_{j=1}^{n} 1_{\{t<T_j\}} \frac{\tau_j \phi(F_j(t))}{1 + \tau_j F_j(t)} \lambda_j(t)\, dt + \sqrt{z(t)} \lambda_n(t) \phi(F_n(t))\, dW(t),$$

$$n = 1, \ldots, M-1,$$

where dW is a Brownian motion under this measure.

The measure $\mathbf{P}$ is assumed to be the probability measure associated with the numeraire B_t. The filtration $\{\mathcal{F}_t\}_{t=0}^{\infty}$ is assumed to be generated by the Brownian motion W_t (and properly augmented with the zero-probability events of $\mathbf{P}$).

For the models (5.1) and (5.2), the vector-valued process

$$\bar{F}(t) = (F_0(t), F_1(t), \ldots, F_{M-1}(t))$$

is Markov. For the stochastic volatility model (5.3), the stochastic variance process $z(t)$ needs to be added to have a Markov process.

The model defined by (5.4) is of the HJM type. In particular, in this model zero coupon discount bonds satisfy the following SDE under the risk-neutral measure,

$$dP(t, T) = r(t) P(t, T)\, dt + \Sigma(t, T) P(t, T)\, dW(t),$$

for some bond volatility process $\{\Sigma(\cdot, T), 0 \le T < \infty\}$. It is also known that we can choose the bond volatility process in such a way that

$$\Sigma(t, t_n) \equiv 0 \text{ for } t \in [t_{n-1}, t_n]$$

for any n, $1 \le n \le M$. We adopt this specification. In particular it implies

$$P(t, t_n) = P(t_{n-1}, t, t_n) \text{ for } t \in [t_{n-1}, t_n], \tag{5.5}$$

for any n, $1 \le n \le M$.

6. Valuing callable Libor exotics in a forward Libor model

6.1. **Recursion for callable Libor exotics.** If a callable contract H_0 has not been exercised up to and including time T_n, ("still alive at time T_n") then it is worth the hold value $H_n(T_n)$. If the callable contract is exercised at time T_n its value is equal to $E_n(T_n)$. Assuming optimal exercise, the value of the callable Libor exotic H_0 at time T_n is then the maximum of the two,

$$\max\{H_n(T_n), E_n(T_n)\}.$$

The value of this payoff at time T_{n-1} is then

$$B_{T_{n-1}} \mathbf{E}_{T_{n-1}} B_{T_n}^{-1} \max\{H_n(T_n), E_n(T_n)\}.$$

Clearly this is the value of the Bermuda swaption that can only be exercised at times T_n and beyond, i.e. of the Bermuda swaption H_{n-1}. These considerations define a recursion

$$\begin{aligned}(6.1)\qquad & H_{n-1}(T_{n-1}) = B_{T_{n-1}} \mathbf{E}_{T_{n-1}} B_{T_n}^{-1} \max\{H_n(T_n), E_n(T_n)\}, \quad n = N-1, \ldots, 1, \\ & H_{N-1} \equiv 0.\end{aligned}$$

The recursion starts at the final time $n = N - 1$ and progresses backward in time. For $n = 1$ we obtain the value $H_0(0)$, the value of the callable that we are after.

This is of course nothing more than a well-known algorithm for pricing American-style options in a backward induction.

Let us denote the exercise region at time T_n by R_n, $R_n \subset \Omega$,

$$R_n \triangleq \{\omega \in \Omega : H_n(T_n, \omega) \leq E_n(T_n, \omega)\}, \quad 1 \leq n \leq N-1. \tag{6.2}$$

Let $\eta = \eta(\omega)$ be the index of the first time that the exercise region is hit (or N if it is never hit),

$$\eta(\omega) = \min\{n \geq 1 : \omega \in R_n\} \wedge N.$$

The callable contract value can be re-written as

$$\begin{aligned} H_0(0) &= \mathbf{E}_0 \left(B_{T_\eta}^{-1} E_\eta(T_\eta) \right) \\ &= \mathbf{E}_0 \left(\sum_{n=\eta}^{N-1} B_{T_{n+1}}^{-1} X_n \right). \end{aligned}$$

6.2. **Monte-Carlo for American-style options.** The problem of pricing American-style options in Monte-Carlo has been considered in [LS98] and [And99]. In the latter, an algorithm for pricing Bermuda swaptions in a forward Libor model was explicitly presented. Extending both algorithms to price callable Libor exotics is a trivial exercise (in theory, not in practice). In this paper we adopt a framework of [LS98].

For more in-depth description of the algorithm for Bermuda swaptions, one can consult [Ped99] or [BM01].

Suppose an estimate of the exercise regions $\tilde{R}_n$, $n = 1, \ldots, N-1$, is available. Then the estimate of the optimal exercise time index is defined by

$$\tilde{\eta}(\omega) = \min\left\{n \geq 1 : \omega \in \tilde{R}_n\right\} \wedge N.$$

Then, a lower bound on the value of a callable contract can be computed by the standard Monte-Carlo algorithm via the formula

$$\begin{aligned}(6.3)\qquad H_0(0) &\geq \tilde{H}_0(0), \\ \tilde{H}_0(0) &= \mathbf{E}_0 \left(\sum_{n=\tilde{\eta}}^{N-1} B_{T_{n+1}}^{-1} X_n \right).\end{aligned}$$

The closer the estimated exercise regions $\tilde{R}_n$ to the actual ones, the tighter the lower bound on the value will be.

The Longstaff and Schwartz (LS) algorithm provides a way to estimate the exercise region from a collection of pre-simulated paths.

For each n, $1 \le n \le N-1$, we choose a p-dimensional $\mathcal{F}_{T_n}$-measurable "explanatory" random vector

$$\bar{V}(T_n) = \{V_m(T_n)\}_{m=1}^{p} = \{V_m(T_n,\omega)\}_{m=1}^{p}.$$

Also, for each n, $1 \le n \le N-1$, we select two parametric families of $\mathbb{R}$-valued functions $f_n(v;\alpha)$ and $g_n(v;\beta)$, $v \in \mathbb{R}^p$, $\alpha, \beta \in \mathcal{A} \subset \mathbb{R}^q$, $q \ge 1$. Without loss of generality we assume that

$$\begin{aligned} f_n(x;0) &\equiv 0, \\ g_n(x;0) &\equiv 0. \end{aligned}$$

We choose special values of the parameters α and β, denoted by $\hat{\alpha}_n$ and $\hat{\beta}_n$, such that the function f_n is a good approximation for the hold value at time T_n as a function of the explanatory vector $\bar{V}(T_n)$,

$$H_n(T_n,\omega) \approx f_n\left(\bar{V}(T_n,\omega), \hat{\alpha}_n\right),$$

and the function g_n is a good approximation for the exercise value at time T_n as a function of the explanatory vector $\bar{V}(T_n)$,

$$E_n(T_n,\omega) \approx g_n\left(\bar{V}(T_n,\omega), \hat{\beta}_n\right).$$

In particular, the Longstaff-Schwartz estimate of the R_n's will be of the form

$$\tilde{R}_n = \left\{\omega \in \Omega : f_n\left(\bar{V}(T_n,\omega), \hat{\alpha}_n\right) \le g_n\left(\bar{V}(T_n,\omega), \hat{\beta}_n\right)\right\}, \quad 1 \le n \le N-1. \tag{6.4}$$

This is similar to (6.2) except that the real hold and exercise values H_n and E_n are replaced by their proxies $f_n\left(\bar{V}(T_n)\right)$ and $g_n\left(\bar{V}(T_n)\right)$.

Let us describe the algorithm for choosing the values of parameters $\hat{\alpha}_n$, $\hat{\beta}_n$, $n = 1, \ldots, N-1$, used in (6.4). To get the best possible estimate of the exercise region $\tilde{R}_n$ for each n, we need to approximate the hold value $H_n(T_n)$ as close as possible with one of the functions from the family $f_n\left(\bar{V}(T_n,\omega), \alpha_n\right)$, and we need to approximate the exercise value $E_n(T_n)$ as close as possible with one of the functions from the family $g_n\left(\bar{V}(T_n,\omega), \beta_n\right)$. We use this as an optimality condition to find $\hat{\alpha}_n$, $\hat{\beta}_n$. We optimize the choice of α_n and β_n over a set of Monte-Carlo paths pre-simulated for that purpose.

Let ω_k, $k = 1, \ldots, K$, be a collection of Monte-Carlo simulated paths. For any random variable X, we denote its k-th simulated value by $X(\omega_k)$. We choose the optimal fit value $\hat{\beta}_n$ from the condition (a non-linear regression of the n-th exercise value on $\left\{g_n\left(\bar{V}(T_n,\omega_k);\beta\right)\right\}_{k=1}^{K}$),

$$\hat{\beta}_n = \arg\min_{\beta} \sum_{k=1}^{K} \left(B_{T_n}(\omega_k) \sum_{i=n}^{N} B_{T_i}^{-1}(\omega_k) X_i(\omega_k) - g_n\left(\bar{V}(T_n,\omega_k);\beta\right) \right)^2 \tag{6.5}$$

for $n = 1, \ldots, N-1$.

The optimal fit variables $\hat{\alpha}_n$, $n = 1, \ldots, N-1$ for the hold value are obtained in backward induction. We set

$$\alpha_{N-1} = 0$$

(so that $f_N \equiv 0$ which is consistent with the fact that $H_{N-1} \equiv 0$ by definition).Then, for each n, having computed $\hat{\alpha}_{n+1}$, $\hat{\beta}_{n+1}$ on the previous step, we obtain $\hat{\alpha}_n$ from

(6.6)

$$\hat{\alpha}_n = \arg\min_{\alpha} \sum_{k=1}^{K} \Big(B_{T_n}(\omega_k)\, B_{T_{n+1}}^{-1}(\omega_k) \max\Big\{ f_{n+1}\left(\bar{V}(T_{n+1},\omega_k), \hat{\alpha}_{n+1}\right), g_{n+1}\left(\bar{V}(T_{n+1},\omega_k), \hat{\beta}_{n+1}\right) \Big\} - f_n\left(\bar{V}(T_n,\omega_k);\alpha\right) \Big)^2 .$$

In practice the parametric families $f_n(\cdot,\alpha)$ and $g_n(\cdot,\beta)$ are most often chosen to depend linearly on parameters α and β. This makes solving for $\hat{\alpha}_n$ and $\hat{\beta}_n$ from (6.5) and (6.6) as simple as running a linear regression. The linearity assumption does not appear to be restrictive in practice, and we adopt it from now on.

The output of the Longstaff-Schwartz pre-simulation step is an estimate of a set of exercise regions $\left\{\tilde{R}_n\right\}_{n=1}^{N-1}$, and an estimate of the optimal exercise time index $\tilde{\eta}$. These are used in obtaining the value (a lower bound) of the contract in the Monte-Carlo simulation via (6.3).

6.3. **Differences in the algorithm for CLEs vs. Bermuda swaptions.** The scheme presented above is somewhat different from the one typically used for Bermuda swaptions. In particular, for a Bermuda swaption (on a vanilla fixed-for-floating swap), there is no need to run a regression on exercise values. The exercise values, being values of plain-vanilla swaps, are available at each exercise date for each path by discounting swap cashflows off an interest rate curve for the date and path ("off the curve" as we call it). For a general callable Libor exotic this is not the case. The underlying typically contains options, and valuing them on future dates is either impossible (in models without closed form expressions for European options) or too expensive. A better choice, as proposed in the section above, is to use regression to estimate the exercise values.

Another simplification that Bermuda swaptions enjoy over more complex callable Libor exotics is the relative ease with which good explanatory variables can be chosen. It is quite clear that for a Bermuda swaption, in regressing the hold value, the value of the underlying swap on each of the exercise dates is very important. It represents an overall level of rates on the exercise date. We show typical results of regressing hold values on exercise values for a Bermuda swaption in Figure 3. Going a bit further it becomes clear that a second important variable is the slope of the interest rate curve on the exercise date (a hold value is an option on rates with different tenors, and therefore their relative values, as measured by the slope, are likely to be important). For callable Libor exotics, the choice of explanatory variables is less clear-cut. We discuss it in the next section.

6.4. **Choosing explanatory variables and parametric families.** A formal scheme of valuing callable Libor exotics, as presented above, is relatively easy to implement. What is usually not so easy is to choose "good" explanatory variables and parametric families f and g for general callable Libor exotics. Recall that the result produced by the LS algorithm is a lower bound on the value. Choosing a good set of explanatory variables and parametric families is very important for getting a *tight* lower bound. The closer the estimates of the exercise boundaries to the real ones, the closer the lower estimate of the value will be to the actual value.

For the LS algorithm to work, numerical properties of the regression procedure are important. The success of the LS algorithm will depend on how robust and well-behaved the numerical problem of fitting f and g to simulated values is. Overfitting is the main danger. In particular, one should not use an overly rich set of parametric functions in the fitting. The more functions there are to choose from for the optimization procedure, the less robust it generally is. A less robust fitting procedure may find a fit that is acceptable for the values it fits but behaves completely unreasonably outside the range of fitted values. For example, high-order polynomials will generally be a bad choice as they would react very strongly to the presence of outliers in the simulation.

The danger of overfitting is somewhat increased by the fact that both explanatory variables and parametric families perform basically the same function. To avoid this problem we must separate the roles that parametric families and explanatory variables play. To that effect, we impose the following distinction on the two. We argue for using very simple parametric families (for example, polynomials of degree 2 in multiple variables work well). With simple parametric families, the dangers of overfitting are minimized.

Another reason for using low-order polynomials is given in [GY03]. The result of the paper states that the number of paths required for convergence grows exponentially with the order of polynomials used.

Choosing simple parametric families is not as restrictive as it looks. The parametric families f and g are only used to set exercise regions, see (6.4). The only place where good fit is really required is around the exercise boundary. And that can typically be achieved with simple parametric families.

Since parametric families are simple, one only needs to focus on choosing good explanatory variables. By "good" we mean financially meaningful variables that really drive the future simulated exercise and hold values.

There are no hard and fast rules for choosing explanatory variables. The analogy with Bermuda swaptions should be a good guide. Also, the following considerations should be employed.

- Use financially meaningful explanatory variables;
- Decide what are the primary financial variables affecting exercise values (often the overall level of rates on the exercise date);
- Decide what are the primary financial variables affecting hold values relative to exercise values (often the slope of the rate curve on the exercise date);
- Decide what are the secondary effects and whether they should be accounted for.

One requirement on explanatory variables mentioned in Section 6 should be explained in a bit more detail. For the LS algorithm to work, these variables, observed on any exercise time T_i, should be $\mathcal{F}_{T_i}$-measurable. What it means is that they should not be "future-looking". They should be computable by using only the state of the model i.e. the interest rate curve, as observed up to and at time T_i. This requirement, while seemingly technical, is very important to the success of the algorithm. It comes down to the fact that we one is not allowed to "see into the future" to make our exercise decisions.

We do not need to be quantitatively exact in capturing the effect of explanatory variables on exercise/hold values. As long as we got the general influence right, the fitting of parametric functions will take care of choosing the best scalings and/or linear combinations of explanatory variables. For example, one may decide that the level of rates, as measured by a certain swap rate, is important, as well as the slope of the rates as measured by the difference between the swap rate and some short-tenor Libor rate. One does not need to include the difference of the rates as an explanatory variable. The short-tenor Libor rate will work just as well. The properly weighted difference of the two will implictly be used in the fitting of parametric families.

The number of explanatory variables one typically needs is pretty small (if they are properly chosen). Two variables per exercise date are usually enough. It is hard to get a noticeable improvement in the lower bound by adding a third variable. It also appears that, almost universally, the two most useful explanatory variables are variables representing an overall level of rates and a slope of the interest rate curve. An overall level of rates is usually well captured by a core swap rate (a rate that fixes on the exercise date and matures when the whole deal matures), or something related to it. The slope is well captured by a short-tenor Libor rate (a 6 month or 1 year tenor) that fixes on the exercise date, or something related to it.

Using a core swap rate, and a short-tenor Libor rate for each exercise date as explanatory variables usually leads to acceptable results. In many cases, however, they can be improved substantially. For a typical CLE, the underlying has interest rate options in it (think of a callable capped floater as an example where the underlying is a collection of caplets) and as such has certain amount of curvature with respect to the level of rates. See the left pane of Figure 4, where simulated exercise values are plotted versus the level of rates. It turns out that using explanatory variables with approximately the same amount of curvature often gives better results. See results in the left pane of Figure 5. One can ask why is that the case, since the parametric functions have curvature, and would be able to match the underlying's features in the regression even if the explanatory variables were just rates. What comes in play here, however, is the fact that, having chosen simple functions for the parametric families for the purposes of robustness, the "resolution" we get from them is typically insufficient to match the finer features of the dependence of the underlying on the rates. For example, for a callable capped floater, the value of the underlying is capped from above. If we just fit a second-order polynomial to values of such underlying versus the swap rate (see the fitted exercise values on the left pane of Figure 4), the fitted values will not respect the cap, and will happily extend above it. It is especially troublesome given that the poorest fit is achieved where it is needed the most, i.e. around the exercise boundary. This can lead to a sub-par performance of the LS algorithm.

Good fitting of hold values, on the other hand, typically does not require any specially "curved" explanatory variables. The hold values' curvature is usually well captured by the curvature in parametric functions, especially around the area of interest (around the exercise boundary). These results are plotted in the right panes of Figures 4 and 5.

If one uses a model that allows for closed form valuation of European options on rates at any future time under any interest rate scenario, then one can just use the exact exercise value on each exercise date as an explanatory variable. This however only works for a subset of models (from the three presented in Section 5 it only works for the lognormal one, and only if the underlying consists of options on primary Libor rates). Moreover, this approach can significantly impair the speed of valuation, since one needs to compute European option prices a large number of times. Fortunately, we do not need to be all that exact in trying to match the curvature of the exercise values with the explanatory variables. Typically, a rough estimate will do. For example one can use the Black-Scholes formula with an approximately-right volatility to value the options in the underlying, even in models where it does not exactly apply. If the underlying is a strip of options, one does not need to value all options in the underlying, just one (say the one in the middle of the strip). If the speed of valuation is still an issue, one can use formulas that exhibit Black-Scholes like behavior but are faster to compute. For example, the following formula can be used as an approximation:

$$\begin{aligned}
b\left(x,K,\sigma,\tau\right) &= -G_1\left(x-K;\sigma\sqrt{\tau}\right)+\left(x-K\right)G_0\left(x-K;\sigma\sqrt{\tau}\right),\\
G_0\left(y;v\right) &= \begin{cases} \frac{1}{2}e^{\frac{\sqrt{2}}{v}y}, & y<0,\\ 1-\frac{1}{2}e^{-\frac{\sqrt{2}}{v}y}, & y\geq 0,\end{cases}\\
G_1\left(y;v\right) &= \begin{cases} \frac{e^{\frac{\sqrt{2}}{v}y}}{2\sqrt{2}}\left(\sqrt{2}y-v\right) & y<0,\\ \frac{e^{-\frac{\sqrt{2}}{v}y}}{2\sqrt{2}}\left(-\sqrt{2}y-v\right) & y\geq 0.\end{cases}
\end{aligned}$$

This is the European call option price with strike K, spot x, absolute volatility σ and time to expiry τ in a “model” where the spot process has the probability density

$$p\left(y;x,v\right)=\frac{1}{v\sqrt{2}}\exp\left(-\frac{\sqrt{2}}{v}\left|y-x\right|\right),\quad -\infty<y<\infty.$$

Volatilities to be used are the forward volatilities of the appropriate Libor rates.

To give an example, consider a callable capped floater. The coupon at time T_i is given by (we denote $L_i\left(t\right)=F\left(t,T_i,T_{i+1}\right)$)

$$C_i=\min(L_i\left(T_i\right)+s,c).$$

The coupon is received against a Libor rate payment,

$$X_i=\delta_i\times\left(C_i-L_i\left(T_i\right)\right).$$

We note that

$$\begin{aligned}
X_i &= \delta_i\times\left(\min(L_i\left(T_i\right)+s,c)-L_i\left(T_i\right)\right)\\
&= \delta_i\times s-\delta_i\times\max(L_i\left(T_i\right)-\left(c-s\right))
\end{aligned}$$

so the payment is a combination of a fixed rate payment with rate s and a call option on the Libor rate with strike $c-s$.

For the first explanatory variable we will use an approximate value of the exercise value on each exercise date

$$V_1\left(T_n\right)=\sum_{i=n}^{N-1}\delta_i\times P\left(T_n,T_{i+1}\right)\times\left(s-b\left(L_i\left(T_n\right),c-s,\sigma_{ni},T_i-T_n\right)\right).$$

Here we use the approximate option formula $b\left(\cdot\right)$ (we could have used the Black-Scholes formula if valuation speed was not a concern). The volatility σ_{ni} is the *absolute* volatility of the Libor rate $L_i\left(\cdot\right)$ over the interval $\left[T_n,T_i\right]$ as given by the model (or an approximation thereof). To reiterate, very crude approximations can be used here. In the lognormal case we may use

$$\sigma_{ni}=L_i\left(0\right)\times\bar{\lambda}$$

where $\bar{\lambda}$ is some average of $\lambda_n(t)$ over n and t in (5.1). (We need to multiply it by $L_i(0)$ because the formula asks for the absolute volatility, not a relative one.) For the skew-extended model we may use

$$\sigma_{ni} = \phi(L_i(0)) \times \bar{\lambda},$$

and for the stochastic volatility model we may use

$$\sigma_{ni} = \sqrt{z(T_n)} \times \phi(L_i(0)) \times \bar{\lambda}. \tag{6.7}$$

What should we use for our second explanatory variable? In the case of a Bermuda swaption we opted to use a front Libor rate fixing on an exercise date. We can use the same here. Alternatively, we can just use the *value* of the first coupon payment:

$$V_2(T_n) = \delta_n \times P(T_n, T_{n+1}) \times (s - b(L_n(T_n), c - s, \sigma_{nn}, 0))$$

or the last one:

$$V_2(T_n) = \delta_n \times P(T_n, T_N) \times (s - b(L_{N-1}(T_n), c - s, \sigma_{n,N-1}, T_{N-1} - T_n)).$$

In conjunction with the first explanatory variable, either one will capture the effect of the changing interest rate curve slope well.

For the parametric functions f and g, the second-degree polynomials in V_1 and V_2 are used.

The approach presented in the example is quite general and can be applied to the vast majority of callable Libor exotics.

One final note is in order. Note that in the stochastic volatility case, we proposed using an explanatory variable that is dependent on the stochastic volatility process. It turns out that this is quite important to do, even for CLEs that do not have optionality in the underlying, like a Bermuda swaption. The stochastic volatility process $z(\cdot)$ can either be incorporated into an explanatory variable like we did it in (6.7), or used as a separate explanatory variable.

6.5. **An upper bound in Monte-Carlo simulation.** The LS algorithm (as well as other algorithms based on estimating exercise boundaries or exercise decisions) only produces a lower bound on the value of a callable Libor exotic. This is somewhat unsatisfactory, as there is no way of knowing how far off the lower bound really is from the "true" model value. The problem is especially troubling for more complicated callable Libor exotics where the choice of explanatory variables in not transparent. One can experiment with different choices and try to get the value as high a possible. One is never sure, however, whether the inability to improve a value is due to the fact that it is already close to the true one, or because one was unable to find good explanatory variables.

Fortunately, recent advances in Monte-Carlo valuation of callable instruments may be able to help. The papers [Rog01], [AB01] outline the algorithm for obtaining an *upper bound* on the value of a callable instrument. In particular,[AB01] presents results for Bermuda swaptions in a forward Libor model. The results and techniques can be extended to callable Libor exotics, thus producing a reliable gauge of the quality of the lower bound approximation. The upper bound gets tighter as the approximations of the exercise boundary improve (i.e. as better and better explanatory variables and parametric families are found). This way, the tightness of the confidence interval can be used to determine whether a good set of explanatory variables/parametric families was found. The time it takes to compute the upper bound is significantly higher than that for the lower bound. Fortunately, one needs to run it only once, to determine the quality of the exercise boundary approximation. Having found a good set, the LS algorithm can be run from then on for valuation and risk management purposes.

7. Risk sensitivities of callable Libor exotics: a short overview

Valuation of callable Libor exotics was discussed in the previous sections. With this section we start our discussion of different methods for obtaining various risk sensitivities for callable Libor exotics. The risk sensitivities, also known as Greeks, that we focus on are deltas, vegas and gammas. As a group, they can be thought of as derivatives of the value of a CLE with respect to various market parameters (rates and volatilities).

Conceptually, risk sensitivities should be easy to obtain – just bump the appropriate input and revalue the security. Reality, however, is rather different. Computing Greeks in a Monte-Carlo simulation is a notoriously difficult task. It is a well-known numerical phenomena that the noise, inherently present in Monte-Carlo valuation, affects numerical derivatives (Greeks) significantly more than the underlying value. Interestingly, it is even worse for callable Libor exotics than for other interest rate derivatives. This is a consequence of the way CLEs are valued; we will discuss why this is so in more detail later.

In addition to accuracy, the speed of valuation becomes critical once computation of Greeks enters the picture. The number of various rates/volatilities that a typical CLE is exposed to is large. Computing all these sensitivities takes a lot of time and system resources. Timeliness of computations is, understandably, a big concern for traders.

Most of the rest of the paper is about solving the noise problem. As we can always trade speed of execution for accuracy, savings in valuation time are directly translatable into the quality of risk sensitivity estimates (and vice versa).

Before proceeding any further, let us recall the definitions of Greeks. Deltas are defined as sensitivities of the value of a CLE to small changes in interest rates. Since different parts of the interest rate curve can move relatively independently, interest rate deltas are usually *bucketed*. This means that the sensitivity of the value to changed in different parts of the interest rate curve are measured separately. With a forward Libor model, bucketed deltas can naturally be defined as sensitivities to primary Libor rates $F_i(0)$ as defined in Section 5.

Vegas are generally defined as sensitivities to small changes in volatilities. What we are most interested in are sensitivities to changes in market volatilities, i.e. term volatilities of caplets and swaptions. Typically, one requires sensitivities to all volatilities that the model was calibrated to. If the model is calibrated to the whole available set of swaptions, vegas are represented by a vega grid, indexed by the swaption's time to expiration and the maturity of the underlying swap. Note that these vegas are quite different from the sensitivities to model's own volatilities ($\lambda_n(\cdot)$ for the three models defined in Section 5).

Gammas are second order derivatives. They are defined as sensitivities of deltas with respect to interest rate changes. Just like the deltas, they are typically bucketed.

The risk sensitivities defined above are a minimum set required for trading CLEs. They are, of course, not sufficient. Because of a complex, highly non-linear dependence of CLEs on market inputs, risk-managing them by using only local measures of risk is extremely dangerous. Rehedging volatility exposures dynamically is expensive. It is much better to construct a semi-static hedging portfolio that tracks the exotics book well, even over big moves in rates and volatilities. To achieve this a global, scenario based, risk framework should be developed to complement the local sensitivity analysis. Conceptually, building such a framework is a significant challenge. Technically, however, it is much easier to obtain good scenario risk than good local risk, for obvious reasons. Since the focus of this paper is on technical challenges of CLE modeling, we do not pursue the topic of scenario risk further.

8. Exercise boundary and risk sensitivities

To compute risk sensitivities, a market input is shocked by a small amount and the instrument is revalued. The base value of the instrument is subtracted to get the sensitivity. For callable Libor exotics, revaluing an instrument is a two-stage procedure. First, the exercise boundary is estimated using pre-simulated paths and the LS algorithm. Then, the CLE is valued as a type of a barrier option that pays the exercise value the first time the exercise boundary is crossed (or, rather, the first time an exercise region is hit). The purpose of this section is simple. We will explain why, when computing local risk sensitivities, we do not need to perform the first step, the exercise boundary estimation. We can just reuse the exercise boundary that was obtained during the base valuation of the CLE.

Here we only explain the intuition behind this result. A formal proof is given in [Pit03].

The advantages of not having to recompute the exercise boundary for each Greek calculation are clear. First, it is faster. Typically, estimation of the exercise boundary takes as much time as valuation in post-simulation. Since the boundary only needs to be estimated once, roughly 50% of valuation time for each Greek computation is saved.

The other advantage is noise reduction. As a new boundary is not estimated with every input bumped, the noise from errors involved in estimating the boundary never enters the final result.

The basic reason why we do not have to estimate the exercise boundary is this. The exercise boundary obtained for the base valuation is *optimal*. Optimality implies that the boundary's first order sensitivity to any input that affects it is zero (first order optimality condition). So, when some input is bumped, its effect on the exercise boundary is only second-order. This means that in the limit, we can keep the same exercise boundary.

Another way of looking at this is to realize that this is a consequence of a smooth pasting condition for the optimal exercise boundary in valuation of American style options.

Here is a "proof" of our statement. Suppose we would like to compute sensitivity to parameter x, whose current value is x_0. For each x, there is an optimal exercise boundary $\zeta(x)$. We denote the value of a CLE for a given value of the parameter x and some arbitrary boundary ζ by $h(\zeta, x)$. Then, the actual CLE value, if the parameter in question is x, is equal to $h(\zeta(x), x)$. The sensitivity to x is then equal to

$$\left.\frac{\partial}{\partial x} h(\zeta(x), x)\right|_{x=x_0} = \left.\frac{\partial}{\partial \zeta} h(\zeta, x_0)\right|_{\zeta=\zeta(x_0)} \times \left.\frac{\partial}{\partial x}\zeta(x)\right|_{x=x_0} + \left.\frac{\partial}{\partial x} h(\zeta(x_0), x)\right|_{x=x_0}.$$

Note that, since $\zeta(x_0)$ is the *optimal* boundary for the CLE, we have

$$\left.\frac{\partial}{\partial \zeta} h(\zeta, x_0)\right|_{\zeta=\zeta(x_0)} = 0.$$

Hence, the full derivative of h with respect to x is equal to the partial one while keeping the exercise boundary constant.

9. Why CLE Greeks are hard to compute

When first trying to compute deltas of a callable Libor exotic, even for as "simple" as one a Bermuda swaption, one is often struck by how much more noisy the deltas look compared to, say, European swaptions' (also computed in a Monte-Carlo simulation of course). It is important to understand the reason for this. Apart from satisfying intellectual curiosity, such knowledge should help us design algorithms that produce cleaner deltas.

Armed with the estimated exercise regions and the optimal exercise index $\tilde{\eta}(\cdot)$, the value of a callable Libor exotic is given by

$$\tilde{H}_0 = J^{-1} \sum_{j=1}^{J} \sum_{i=1}^{N-1} \left[B_{T_{i+1}}^{-1}(\omega_j) X_i(\omega_j) 1_{i \geq \tilde{\eta}(\omega_j)} \right],$$

where simulated paths (for the post-simulation) are denoted by ω_j, $j = 1, \ldots, J$. The dependence of the simulated CLE value $\tilde{H}_0$ on small changes in (for example) time-zero primary Libor rates $F_m(0)$, $m = 0, \ldots, M-1$, comes from a number of different sources (we will discuss this issue in the context of deltas, but the same considerations apply to vegas as well). First, there is the dependence of $B_{T_{i+1}}^{-1}(\omega_j)$ on Libor rates. This dependence is very smooth, as clear from the formula for the numeraire. Then there is the dependence of payments $X_i(\omega_j)$. For many important classes of CLEs this is also smooth (analytic for Bermuda swaptions; continuous for callable inverse floaters and the like). Finally there is the dependence of "exercise indicators" $1_{i \geq \tilde{\eta}(\omega_j)}$ on the rates. It should be clear that these indicators do not depend smoothly on the changes in rates at all. And these are actually the culprit of our problems with deltas. Non-smooth dependence like this produces noisy deltas. We note that these indicators are not present in valuation formulas for, say, European swaptions. The indicators specify which coupons have to be added to the exercise value for each simulated path. Being binary "yes/no" indicators they cannot vary smoothly with rates. A small change in rates can add/subtract a whole payment for a particular path. Figure 6 demonstrates the problem visually. If a path of interest rates passes sufficiently close to the exercise boundary, then a small change in the initial values of primary Libor rates $\{F_m(0)\}$ can push the path outside of the exercise region for one of the exercise dates. A whole payment is lost as a result.

In effect, in a Monte-Carlo simulation, a callable Libor exotic behaves like a barrier option in a Monte-Carlo simulation. This behavior is not a result of having only a *suboptimal* exercise boundary coming from the LS algorithm. In fact, had we used the absolute optimal exercise boundary, we would still have the same problem – for some paths, a whole payment will be added/lost under a bump. Differentiating discontinuous functions is hard!

10. Computing deltas in the same simulation

As explained in the previous section, the standard "bump and revalue" method of obtaining deltas can run into problems. The valuation algorithm treats callable Libor exotics as barrier options. This introduces a lot of noise in sensitivities calculations. However, if one steps back and looks at a callable Libor exotic, it actually does not look like a barrier option at all. Because of the optimality of exercise condition, hold and exercise values are continuous across the exercise boundary. So the problem discussed in the previous section does not appear to be intrinsic to the instrument, but only to the method used to compute deltas.

Since all the payoffs of a CLE are, in a sense, smooth functions of the market inputs, one can try to take advantage of that and differentiate them *along each path*. Such approach, known as pathwise differentiation (with the deltas known as pathwise deltas) has been applied quite successfully in various Monte-Carlo applications. In the case of European options on interest rates, it was first proposed in [GZ99]. The method was extended to callable Libor exotics in [Pit03].

The method allows one to compute deltas of a CLE in the same simulation as the value. This speeds up the valuation of Greeks considerably. More importantly, the resulting deltas are significantly less noisy. The latter is the result of a numerical scheme that avoids representing CLEs as barrier-type structures. As reported in [Pit03], the standard errors for pathwise deltas can be as much as 5-6 times less than for the "bump and revalue" ones, resulting computation time savings by a factor of 25 to 35.

Before carrying on, we note that the method does not work for CLEs whose coupons are discontinuous functions of rates (such as callable range accruals). For such CLEs, a less advantageous (but still better than "bump and revalue") approach of likelihood ratio deltas can be applied. For details see [GZ99] and [Pit03].

Let us define pathwise deltas and outline the main steps in applying the pathwise deltas method to CLEs (for details and proofs see [Pit03]). The main formula for the method is (10.6) below, with pathwise deltas of forward Libor rates defined by (10.2) or (10.3).

10.1. **Pathwise deltas.** For any random payoff X, the m-th delta is defined by

$$\Delta_m X \triangleq \frac{\partial X}{\partial F_m(0)},$$

where $\{F_m(t)\}$ is the set of primary Libor rates.

Suppose $X = V\left(\bar{F}(T)\right)$ for a (reasonably smooth) deterministic function $V(\mathbf{x})$, $\mathbf{x} = (x_0, \ldots, x_{M-1})$. Then

$$\Delta_m V\left(\bar{F}(T)\right) = \sum_{i=0}^{M-1} \frac{\partial V\left(\bar{F}(T)\right)}{\partial x_i} \cdot \Delta_m F_i(T). \tag{10.1}$$

The value of the payoff X at time 0 is equal to

$$\mathbf{E}_0\left(B_T^{-1} X\right),$$

and its delta

$$\begin{aligned}\Delta_m \mathbf{E}_0\left(B_T^{-1} X\right) &= \mathbf{E}_0\left(\Delta_m\left(B_T^{-1} V\left(\bar{F}(T)\right)\right)\right)\\ &= \mathbf{E}_0\left(\Delta_m\left(B_T^{-1}\right) V\left(\bar{F}(T)\right)\right)\\ &\quad +\mathbf{E}_0\left(B_T^{-1} \Delta_m V\left(\bar{F}(T)\right)\right).\end{aligned}$$

It is clear from these formulas that, to compute pathwise deltas, all we need to be able to do is to simulate $\Delta_m F_i(T)$, the deltas of Libor rates with respect to their initial values, along with the values of the Libor rates themselves. We present relevant details for the case of a skew-extended forward Libor model (5.2). By differentiating (5.2) through with respect to $F_m(0)$, we obtain the following system of SDEs for Libor rates and their deltas. If we denote the drift for the n-th Libor rate, as a function of the vector of Libor rates $\bar{F}(t)$, by $\mu_n(t, \mathbf{x})$, $\mathbf{x} = (x_0, \ldots, x_{M-1})$, then for both the Libor rates and their derivatives ($i = 1, \ldots, M-1$,

$m = 0, \ldots, M-1$)

$$dF_i(t) = \lambda_i(t) \cdot \phi(F_i(t)) \cdot \left(\mu_i\left(t, \bar{F}(t)\right) dt + dW(t)\right), \tag{10.2}$$
$$d(\Delta_m F_i(t)) = \lambda_i(t) \cdot \phi'(F_i(t)) \cdot \Delta_m F_i(t) \cdot \left(\mu_i\left(t, \bar{F}(t)\right) dt + dW(t)\right)$$
$$+ \lambda_i(t) \phi(F_i(t)) \left(\sum \frac{\partial \mu_i\left(t, \bar{F}(t)\right)}{\partial x_j} \cdot \Delta_m F_j(t)\, dt \right),$$

with the initial conditions for the deltas given by

$$\Delta_m F_j(0) = 1_{\{m=j\}}.$$

A simplified system of SDEs for simulating values and pathwise deltas of forward Libor rates can be used,

$$dF_i(t) = \lambda_i(t) \cdot \phi(F_i(t)) \cdot \left(\mu_i\left(t, \bar{F}(t)\right) dt + dW(t)\right), \tag{10.3}$$
$$d(\Delta_m F_i(t)) = d(\Delta_m F_i(t)) \lambda_i(t) \cdot \phi'(F_i(t)) \cdot \Delta_m F_i(t) \cdot \left(\mu_i\left(t, \bar{F}(0)\right) dt + dW(t)\right)$$
$$+ \lambda_i(t) \phi(F_i(t)) \left(\sum \frac{\partial \mu_i\left(t, \bar{F}(0)\right)}{\partial x_j} \cdot \Delta_m F_j(t)\, dt \right).$$

Note that the simplification is based on evaluating drifts μ_i along the forward values $\bar{F}(0)$, instead of the actual values $\bar{F}(t)$ at time t (just for the deltas, not for the actual Libor rates). Glasserman and Zhao, who proposed this approximation, show that the loss of accuracy is very small, but speed gains are substantial. This is so because the drift in the equation for the deltas of forward Libor rates can be precomputed before the simulation.

10.2. **Pathwise deltas of callable Libor exotics.** The advantage of the pathwise deltas method is that a substantial number of differentiations can be performed analytically, before simulating a single path. Let us recall the main recursion for the value of a CLE. We have, slightly rewriting (6.1), that

$$B_{T_{n-1}}^{-1} H_{n-1}(T_{n-1}) = \mathbf{E}_{T_{n-1}} B_{T_n}^{-1} \max\{H_n(T_n), E_n(T_n)\}.$$

The order of differentiation and taking expectation can be changed to yield

$$\Delta_m \left(B_{T_{n-1}}^{-1} H_{n-1}(T_{n-1}) \right) = \mathbf{E}_{T_{n-1}} \Delta_m \left(B_{T_n}^{-1} \max\{H_n(T_n), E_n(T_n)\} \right). \tag{10.4}$$

Carrying out the differentiation under the expectation operator, we obtain that

$$\begin{aligned} \Delta_m \left(B_{T_{n-1}}^{-1} H_{n-1}(T_{n-1}) \right) &= \mathbf{E}_{T_{n-1}} \left(1_{\{E_n(T_n) > H_n(T_n)\}} \Delta_m \left(B_{T_n}^{-1} E_n(T_n) \right) \right) \\ &\quad + \mathbf{E}_{T_{n-1}} \left(1_{\{H_n(T_n) > E_n(T_n)\}} \Delta_m \left(B_{T_n}^{-1} H_n(T_n) \right) \right). \end{aligned}$$

This formula is a recursive relationship (in n, the exercise date index) between $\Delta_m \left(B_{T_{n-1}}^{-1} H_{n-1}(T_{n-1}) \right)$ and $\Delta_m \left(B_{T_n}^{-1} H_n(T_n) \right)$. Unwrapping the recursion gives us the main formula for pathwise deltas of CLEs (here η is the optimal exercise index)

$$\Delta_m H_0(0) = \mathbf{E}_0 \left(\sum_{n=\eta}^{N-1} \Delta_m \left(B_{T_{n+1}}^{-1} X_n \right) \right). \tag{10.5}$$

To estimate this delta in a Monte-Carlo simulation, η is replaced with its estimate. A pathwise delta $\Delta_m H_0(0)$ of a callable Libor exotic contract can then be approximated with $\tilde{\Delta}_m H_0(0)$ computed as follows,

$$\tilde{\Delta}_m H_0(0) = \mathbf{E}_0 \left(\sum_{n=\tilde{\eta}}^{N-1} \Delta_m \left(B_{T_{n+1}}^{-1} X_n \right) \right), \tag{10.6}$$

where $\tilde{\eta}$ is the estimate of the optimal exercise time index computed during the LS pre-simulation, as outlined in Section 6.

This elegant formula for deltas of a callable Libor exotic is easy to implement in practice. The estimate of the optimal exercise time, $\tilde{\eta}$, comes "for free" from the pre-simulation step of the valuation. Once $\tilde{\eta}$ is estimated (usually in the form of a collection of exercise regions, as explained in Section 6), deltas are computed by

1. Running a forward simulation, for each path ω determining the optimal exercise time index $\tilde{\eta}(\omega)$;
2. For each path, computing deltas of all payments X_n, $n = 1, \ldots, N-1$ (as well as the deltas of the numeraire $B_{\cdot}^{-1}$) along the path;
3. Adding up deltas $\Delta_m \left(B_{T_{n+1}}^{-1} X_n \right)$ for those coupons that occur after the exercise index $\tilde{\eta}(\omega)$;
4. Averaging the result over all paths.

The pathwise deltas of payments X_n are obtained via a chain rule, similar to how it is done in the equation (10.1). This is where the simulated values of deltas of Libor rates $\Delta_m F_i(t)$ are used. This is also the place where absolute continuity of payments X_n with respect to the Libor rates is required.

Note that some of the benefits of the pathwise deltas approach can be realized in the standard "bump-and-revalue" approach by forcing exercise in the "bumped" scenarios at exactly the same times as in the "base" valuation.

11. "Sausage" Monte Carlo

In this section we describe a different method for dealing with noisy deltas. While we encourage everybody to use the pathwise deltas approach from the previous section, system limitations may prevent its successful application. The method we describe here is easier to implement and integrate with the existing infrastructure, because it is based on the "bump-and-revalue" approach. The main idea of the method is to remove singularities that cause excessive noise in deltas.

11.1. The outline of the method.

Recall the main valuation formula for callable Libor exotics in the Monte-Carlo simulation,

$$\tilde{H}_0 \approx J^{-1} \sum_{j=1}^{J} \sum_{i=1}^{N-1} \left[B_{T_{i+1}}^{-1}(\omega_j) X_i(\omega_j) 1_{i \geq \tilde{\eta}(\omega_j)} \right], \tag{11.1}$$

where ω_j, $j = 1, \ldots, J$, are simulated paths. As analyzed in Section 9, a term $1_{i \geq \tilde{\eta}(\omega)}$ does not depend smoothly on a simulated path ω, and this is a major obstacle for computing deltas.

Arguably, the most direct way of dealing with a discontinuity in a payoff is to try and smooth it out. This entails replacing a discontinuous payoff with a continuous one that is somehow close to the original one.

In this section we develop a systematic way of deriving the appropriate smoothing procedure for (11.1). To develop the necessary machinery, we recall that (11.1) is a sampling analog to the formula

$$\tilde{H}_0 = \mathbf{E} \sum_{i=1}^{N-1} \left[B_{T_{i+1}}^{-1} (\omega_j) \, X_i (\omega_j) \, 1_{i \geq \tilde{\eta}(\omega_j)} \right]. \tag{11.2}$$

Formally, (11.1) is obtained from (11.2) by replacing the measure

$$\mathbf{P}(\cdot)$$

with an empirical approximation

$$\bar{\mathbf{P}}(\cdot) = \frac{1}{J} \sum \delta_{\omega_j} (\cdot).$$

(here $\delta_\omega(\cdot)$ is a delta-measure assigning weight 1 to the path ω). This approximation samples the payoff at each of the simulated paths. There are other approximations to the measure $\mathbf{P}$ available. Instead of just sampling the payoff along each path, the payoff can be integrated along a narrow thin tube around each sample path. This integration will result in smoothing any discontinuities in the payoff. We call this approach "sausage Monte Carlo" because we integrate the payoff along a "sausage" around each sample path.

When integrating the payoff over the sausage, various approximations are used, as described later.

11.2. **A model problem.** Before proceeding let us consider a simple model problem that introduces the main idea of our method more formally.

Consider the problem of integrating a function $f(x)$ of one variable over an interval $[0, 1]$ by using Monte-Carlo. Suppose we do not know much about the function $f(\cdot)$ except that it has a discontinuity at $\hat{x}$, $\hat{x} \in [0, 1]$, and it is smooth everywhere else. The standard Monte-Carlo method will approximate

$$\int_0^1 f \, dx \approx \frac{1}{J} \sum_{j=1}^{J} f(x_j), \tag{11.3}$$

where x_j, $j = 1, \ldots, J$, are sampled from a uniform distribution. The resulting approximation is pretty noisy because of the discontinuity in the payoff $f(\cdot)$. Also we can see that the approximation is not smooth as a function of the sample points x_j. If one moves $\{x_j\}$ around, and one of the sample points crosses over the discontinuity, there will be a large jump in the approximation (resulting in bad "deltas").

Select $\varepsilon > 0$. The "sausage" Monte-Carlo for this problem is defined by the following approximation,

$$\int_0^1 f \, dx \approx \frac{1}{\varepsilon J} \sum_{j=1}^{J} \int_{x_j - \varepsilon/2}^{x_j + \varepsilon/2} f(x) \, dx.$$

In this formula, the payoff is integrated over a ball of radius ε around each sample point x_j, and the average of these integrals is taken. Of course each sub-integral $\int_{x_j - \varepsilon/2}^{x_j + \varepsilon/2} f(x) \, dx$

cannot be computed exactly (if we could, there would be no problem computing $\int_0^1 f\,dx$ exactly!). How do we approximate each sub-integral? If the interval $[x_j - \varepsilon/2, x_j + \varepsilon/2]$ does not contain the discontinuity $\hat{x}$, then we just use the value of f at the middle of the interval, as in the original approximation (accurate to order ε^2),

$$\int_{x_j-\varepsilon/2}^{x_j+\varepsilon/2} f(x)\,dx \approx \varepsilon f(x_j).$$

If, however, the interval $[x_j - \varepsilon/2, x_j + \varepsilon/2]$ does contain the discontinuity $\hat{x}$, then the approximation above is not accurate to order ε^2. So instead, we integrate f over the interval $[x_j - \varepsilon/2, x_j + \varepsilon/2]$, assuming it is constant on the two subintervals $[x_j - \varepsilon/2, \hat{x}]$ and $[\hat{x}, x_j + \varepsilon/2]$,

$$\int_{x_j-\varepsilon/2}^{x_j+\varepsilon/2} f(x)\,dx \approx (\hat{x} - (x_j - \varepsilon/2))\, f(x_j - \varepsilon/2) + ((x_j + \varepsilon/2) - \hat{x})\, f(x_j + \varepsilon/2).$$

This approximation *is* accurate to ε^2. The resulting "sausage" Monte-Carlo approximation becomes

$$\int_0^1 f\,dx \approx \frac{1}{J}\sum_{j=1}^{J} \Big(1_{\hat{x}\notin[x_j-\varepsilon/2,x_j+\varepsilon/2]} f(x_j) \\ +1_{\hat{x}\in[x_j-\varepsilon/2,x_j+\varepsilon/2]} \left(\frac{\hat{x} - (x_j - \varepsilon/2)}{\varepsilon} f(x_j - \varepsilon/2) + \frac{(x_j + \varepsilon/2) - \hat{x}}{\varepsilon} f(x_j + \varepsilon/2)\right)\Big).$$

It is easy to see that this approximation is continuous with respect to the sample points $\{x_j\}$, unlike the naive one (11.3).

While simple, this approach will prove quite fruitful for improving the accuracy of Monte-Carlo for callable Libor exotics in forward Libor models. There, the situation is somewhat similar to the one we just discussed because we know where the exercise boundary (i.e. the discontinuity) is. That allows us to integrate around it.

11.3. **"Sausage" Monte-Carlo for forward Libor models.** All forward Libor models presented in Section 5 are Markov (although in a large number of variables). The lognormal and skew-extended models are Markov when the collection of all primary forward Libor rates at each time t, $\bar{F}(t)$, is considered as the state. The stochastic volatility model is Markov if the stochastic variance process $z(\cdot)$ is added. The same is generally true of any interest rate model – any model (certainly any model of practical interest) is Markov in a large enough (but finite) number of variables. So we define the state of a forward Libor model at time t by $x(t,\omega)$, and assume it is finite-dimensional, and the model is Markov with $x(\cdot)$ being the state variable.

Assume, as before, that sample paths ω_j, $j = 1, \ldots, J$, are drawn. We fix $\varepsilon > 0$, the width of our sausages. For each j, the ε-sausage in the state space is defined by

$$\begin{aligned} A_j^\varepsilon &= \{\omega : \|x(T_i,\omega) - x(T_i,\omega_j)\| < \varepsilon \quad \forall i = 1, \ldots, N-1\} \\ &= \bigcap_{i=1}^{N-1} A_{ji}^\varepsilon, \\ A_{ji}^\varepsilon &= \{\omega : \|x(T_i,\omega) - x(T_i,\omega_j)\| < \varepsilon\}. \end{aligned}$$

The sampling formula (11.2) is approximated with the following expression,

$$\begin{aligned}\tilde{H}_0 &\approx J^{-1}\sum_{j=1}^{J} v_j,\\ v_j &= \mathbf{E}\left(\sum_{i=1}^{N-1}\left[B_{T_{i+1}}^{-1}(\omega)\,X_i(\omega)\,1_{i\geq\tilde{\eta}(\omega)}\right]\middle|\,A_j^{\varepsilon}\right).\end{aligned}$$

Since $B_{T_{i+1}}^{-1}(\omega)$, $X_i(\omega)$ are generally smooth functions of the path ω ($X_i(\omega)$ can of course be discontinuous, but this is not our focus at the moment), we evaluate them just at the sample path (this is accurate to order ε),

$$v_j = \sum_{i=1}^{N-1} B_{T_{i+1}}^{-1}(\omega_j)\,X_i(\omega_j)\,\mathbf{E}\left(1_{i\geq\tilde{\eta}(\omega)}\middle|\,A_j^{\varepsilon}\right). \tag{11.4}$$

Recall that by definition of the optimal exercise index,

$$\begin{aligned}\{i\geq\tilde{\eta}(\omega)\} &= \bigcup_{n=1}^{i}\tilde{R}_n\\ &= \bigcup_{n=1}^{i}\left\{f_n\left(\bar{V}(T_n,\omega),\hat{\alpha}_n\right)\leq g_n\left(\bar{V}(T_n,\omega),\hat{\beta}_n\right)\right\}.\end{aligned}$$

Here, as one can (and should) recall from Section 6, $\tilde{R}_n$ are estimates of exercise regions, $f_n(\cdot,\hat{\alpha}_n)$ are parametric estimates of hold values H_n, g_n are parametric estimates of exercise values E_n, $\bar{V}$ are explanatory variables. The event equality above just says that the event that the exercise happens prior to or on the date T_i is a union of events that the exercise region is hit on the date T_n for $n\leq i$.

Since ultimately all quantities are functions of the state process $x(t)$, we define $\phi_n(x)$ by

$$\phi_n(x(T_n),\omega) = g_n\left(\bar{V}(T_n,\omega),\hat{\beta}_n\right) - f_n\left(\bar{V}(T_n,\omega),\hat{\alpha}_n\right). \tag{11.5}$$

The function $\phi_n(x)$ measures by how much the estimate of the exercise value exceeds the estimate of the hold value at time T_n if the underlying process is in state x.

We show in Appendix A that the probabilities

$$q_i(\omega_j) \triangleq \mathbf{E}\left(1_{i\geq\tilde{\eta}(\omega)}\middle|\,A_j^{\varepsilon}\right),$$

(needed in (11.4)) can approximately be computed as follows,

$$1-q_i(\omega_j) = \prod_{n=1}^{i}\left(1-p_n(\omega_j)\right),$$

$$p_n(\omega_j) \triangleq \begin{cases}1, & \phi_{nj}-\delta_{nj}>0,\\ \frac{\delta_{nj}+\phi_{nj}}{2\delta_{nj}} & \phi_{nj}-\delta_{nj}<0<\phi_{nj}+\delta_{nj},\\ 0, & \phi_{nj}+\delta_{nj}<0.\end{cases} \tag{11.6}$$

Here

$$\begin{aligned}\phi_{nj} &= \phi_n\left(x\left(T_n, \omega_j\right)\right), \\ \delta_{nj} &= \left\|\nabla \phi_{nj}\right\| \times \varepsilon, \\ \nabla \phi_{nj} &= \left.\nabla \phi_n\left(x\right)\right|_{x=x\left(T_n, \omega_j\right)}.\end{aligned}$$

The final approximation formula reads

$$v_j = \sum_{i=1}^{N-1} B_{T_{i+1}}^{-1}\left(\omega_j\right) X_i\left(\omega_j\right) q_i\left(\omega_j\right). \tag{11.7}$$

Let us try to make some sense of these formulas. The quantity ϕ_{nj} measures by how much the exercise value is exceeding the hold value at time T_n for the path ω_j. We call it the "overshoot" function. The quantity δ_{nj} is the "window" over which the overshoot function is smoothed out. It is equal to the universal constant ε (smoothing window for the state variables $x\left(\cdot\right)$) times the size of the gradient of the overshoot function. This provides consistent scaling of smoothing windows across different times/simulated paths. The weight $p_n\left(\omega_j\right)$ measures by how much the overshoot function exceeds zero (or, equivalently, by how much the exercise value exceeds the hold value). Recall that $p_n\left(\omega_j\right)$ is the conditional probability of hitting the exercise region, conditioned on the path being in the ε-sausage around ω_j. So if the exercise value is significantly higher (exceeds by δ_{nj}) than the hold value (first case in (11.6)), we declare that the exercise barrier completely breached, and set $p_n\left(\omega_j\right) = 1$. If the exercise value is significantly under the hold value (the third case in (11.6)), we say that the exercise region was not reached at all. However, in between (the second case in (11.6)), we say that the exercise barrier was "partially" breached, and assign a weight of $\frac{\delta_{nj}+\phi_{nj}}{2\delta_{nj}}$ that measures by "how much" the barrier was breached.

It is very important to note that as we vary the path ω_j (for example when computing sensitivities) the weights $p_n\left(\omega_j\right)$ change smoothly. For those n for which ω_j was in a no-exercise region, as it starts approaching the exercise region, the weight $p_n\left(\omega_j\right)$ is zero at first, then gradually starts growing as the path crosses into the exercise boundary, and then eventually becomes one once it is sufficiently deep inside the region.

As we explained, the quantities $p_n\left(\omega_j\right)$ have the interpretation of being probabilities that the exercise region was entered at time T_n (conditioned on being in the ε-neighborhood of the simulated path ω_j). Accordingly, $1 - p_n\left(\omega_j\right)$ is the probability that it was not entered at time T_n. Henceforth, the quantity $1 - q_i\left(\omega_j\right)$ is the probability that the exercise region was not hit on any of the days $T_1, \ldots, T_i$. Finally, $q_i\left(\omega_j\right)$ is the probability that the exercise region was hit on some date on or prior to T_i. And this quantity is used to weight the contribution of the i-th coupon $X_i\left(\omega_j\right)$ for that particular path.

Since p_n's are smooth with respect to ω, so are q_i's.

A convenient way of thinking of the formula (11.6) is by introducing a concept of a partial exercise. If the path ω_j is near the exercise boundary, relevant coupons get included in the exercise value only partially. The weights $q_i\left(\omega_j\right)$ then give the fractions of the coupons that get included in the exercise value. This is in contrast to the standard Monte-Carlo formula (see (11.1))

$$v_j = \sum_{i=1}^{N-1} B_{T_{i+1}}^{-1}\left(\omega_j\right) X_i\left(\omega_j\right) 1_{i \geq \tilde{\eta}\left(\omega_j\right)}.$$

In the standard formula coupons either get included in the exercise either completely or not at all, and that introduces excessive noise in the valuation.

Conveniently, the sausage Monte-Carlo formula (11.7) converges to the standard one above as ε is taken smaller and smaller.

The larger ε (the smoothing window) is, the smoother the payoff becomes, resulting in less noisy deltas. With larger ε, however, the approximations we made start breaking down, introducing bias in the final price. The way to apply the formula (11.7) in practice is to start with a small ε. Then keep increasing it for as long as the observed bias in the price is within pre-set tolerances. Once the upper acceptable bound on ε is established, use that value in risk sensitivity calculations.

12. Low-dimensional Markovian approximation for a forward Libor model

A forward Libor model is Markovian in all forward Libor rates (plus the stochastic variance variable for the Stochastic Volatility forward Libor model). Lattice and PDE (Partial Differential Equations) models are available for Markovian models, but the dimensionality of a forward Libor model is far too great for PDE methods to be useful. It is generally accepted that Markovian models with one and two state variables are ideally suited for PDE methods. The ones with three state variables are in the grey zone where it is not clear whether PDE methods have advantage over Monte-Carlo or not. An exact representation of a forward Libor model in a Markovian form with 2 or 3 state variables is impossible. However, a number of practitioners and academicians have explored ways of *approximating* a forward Libor model with a low-dimensional Markovian one, to which PDE methods can be applied. In this section we outline one possible approximation, and discuss its applicability to the problem of pricing and risk managing callable Libor exotics.

12.1. Deriving an approximation.

Deriving an approximation. The skew-extended forward Libor model has the following dynamics under the spot Libor measure, the measure for which $\{B_t\}$ is the numeraire (compare to (5.4)),

$$\begin{aligned} dF_n(t) &= \lambda_n(t)\,\phi(F_n(t))\left(\mu_n\left(t,\bar{F}(t)\right)\,dt + \,dW(t)\right), \\ \mu_n\left(t,\bar{F}(t)\right) &= \sum_{i=1}^{n} 1_{\{t<T_i\}}\phi(F_i(t))\,\frac{\tau_i}{\tau_i F_i(t)+1}\lambda_i(t), \\ n &= 1,\dots,M-1. \end{aligned} \tag{12.1}$$

(for simplicity we keep dW one-dimensional).

To get rid of the local volatility $\phi(F_n(t))$, the following transform is introduced (x_0 is an arbitrary but fixed number),

$$f(x) = \int_{x_0}^{x} \frac{d\xi}{\phi(\xi)},$$

and define new variables

$$L_n(t) \triangleq f(F_n(t)), \quad n = 1,\dots,M-1. \tag{12.2}$$

Then, in the new variables,

$$\begin{aligned} dL_n(t) &= \lambda_n(t)\left(\left(\mu_n\left(t,\bar{F}(t)\right) - \frac{1}{2}\lambda_n(t)\,\phi'(F_n(t))\right)dt + dW(t)\right), \\ n &= 1,\dots,M-1. \end{aligned}$$

The first problem in making this SDE Markovian in a small number of state variables is the drift $\mu_n\left(\cdot,\bar{F}(\cdot)\right)$. At each point in time it depends on the whole collection of forward Libor rates. The easiest way to deal with this problem is to replace μ_n with its value at today's forward Libor rates,

$$\mu_n\left(t,\bar{F}(t)\right)\approx\mu_n\left(t,\bar{F}(0)\right). \tag{12.3}$$

This can be rationalized by the fact that the drifts are pretty small, and changes in the drifts can be ignored as a first-order approximation. Also, the term $\phi'(F_n(t))$ is troublesome. For local volatility functions $\phi(x)$ that are close to linear, it is close to $\phi'(F_n(0))$ (the approximation is exact for such important cases as the lognormal model and the displaced diffusion model). Thus, the first step in building a Markovian forward Libor model is to use the following SDE instead,

$$\begin{aligned} dL_n(t) &= \lambda_n(t)\left(\left(\mu_n\left(t,\bar{F}(0)\right)-\frac{1}{2}\lambda_n(t)\,\phi'(F_n(0))\right)dt+dW(t)\right), \\ n &= 1,\ldots,M-1. \end{aligned} \tag{12.4}$$

With the drift completely deterministic, each $L_n(t)$ is an integral of $\lambda_n(t)$ against a Brownian motion. To make all the variables the functions of the same state variable, we impose the following "separability" condition of the volatility structure $\{\lambda_n(\cdot)\}$,

$$\lambda_n(t)=\sigma_n\times\nu(t),\quad n=1,\ldots,M-1,\quad t\geq 0. \tag{12.5}$$

What this condition says is that each forward Libor volatility functions is equal to a scalar (Libor-specific), multiplied by a function of time common to all Libor rates. A Markovian state variable is defined by

$$dX(t)=\nu(t)\,dW(t).$$

All variables $L_n(\cdot)$ are deterministic functions of the state variable $X(t)$,

$$\begin{aligned} L_n(t) &= L_n(0)+d_n(t)+\sigma_n X(t), \\ d_n(t) &= \int_0^t\lambda_n(s)\left(\mu_n\left(s,\bar{F}(0)\right)-\frac{1}{2}\lambda_n(s)\,\phi'(F_n(0))\right)ds, \end{aligned}$$

and all forward Libor rates are as well,

$$F_n(t)=f\left(f^{-1}(F_n(0))+d_n(t)+\sigma_n X(t)\right),\quad n=1,\ldots,M-1. \tag{12.6}$$

Now, at each point in time t, any interest-rate related payoff V can be expressed as a function of t and $X(t)$,

$$V=V(t,X(t)).$$

The function $V(t,x)$ satisfies the following PDE,

$$-\frac{\partial V(t,x)}{\partial t}=\frac{\nu^2(t)}{2}\frac{\partial^2 V(t,x)}{\partial x^2}+r(t,x)\,V(t,x).$$

Any callable Libor exotic can be valued by solving this PDE.

In the PDE above, $r(t,X(t))$ is the discounting rate applied to any instrument over an instantaneous period of time $[t,t+dt]$. It is defined as (here n is defined by the condition that $t \in (t_n, t_{n+1}])$

$$r(t, X(t)) = \frac{\partial B_t / \partial t}{B_t} = \frac{\partial P(t, t_{n+1}) / \partial t}{P(t, t_{n+1})}.$$

The bond $P(t, t_{n+1})$ is certainly a function of forward Libor rates $\bar{F}(t)$ at time t and, therefore, a function of the state $X(t)$. So is $r(t, X(t))$ then. The specific expression for the rate $r(t, X(t))$ in terms of $X(t)$ depends on the interpolation assumptions made.

12.2. **Discussion of the approximation.** A Markovian (in one variable) approximation to a forward Libor model has been derived. Can it be used to value and risk-manage callable Libor exotics? Unfortunately, the answer is no. Tests show that the drift approximation (12.3) introduces too much error in the pricing. The approximation may be acceptable for very short (under 5 years) instruments in low volatility scenarios, but it deteriorates very quickly as times get longer and volatilities higher. More sophisticated approximations to the drift can be applied (Brownian bridges, predictor-corrector schemes, ODEs for the drift, Pickand iterations) but to the best of our knowledge, none of them extend the useful range of this approach beyond 10 years.

One can argue that the errors in pricing are not so bad if one only uses this scheme to compute deltas and other risk sensitivities. When a base value of an instrument is subtracted from the value in a bumped scenario, the errors will cancel out. This is true, but to a point. If the value of a derivative is off by 20%, how can one be sure that the deltas are not off by 20% as well?

Another reason why a direct application of the Markovian approximation method fails has to do with volatility restrictions. The volatility restrictions (12.5) imposed by the need to extract a single driving factor for all Libor rates are actually quite severe. The two main advantages of a forward Libor model are its ability to calibrate to a big set of market instruments (swaptions) and its ability to generate the evolution of the volatility structure that is reasonably realistic. Both of these capabilities get severely hampered when the available volatility surfaces are restricted to "separable" ones in (12.5).

At the end, a Markovian approximation to a forward Libor model looses main advantages of a "real" forward Libor model and, moreover, produces prices too inaccurate to be useful. Why then, a reader may ask, are we talking about it? It turns out that having such an approximation in one's toolkit of models is actually quite useful, as will be explained in later sections. Before we go into that discussion however, let us quickly talk about multidimensional extensions.

12.3. **Two-dimensional extension.** Suppose we have a forward Libor model driven by a *two-dimensional* noise. Building a Markovian approximation to such a model (with two state variables) seems like a trivial extension of the method we developed, but it is not[1].

To set some notations, we assume that there are two components to the noise,

$$dW = \left(dW^1, dW^2\right),$$

[1]This important observation, and many ideas presented in this section, were shared with us by Leif Andersen.

and forward Libor volatilities are two-dimensional processes,

$$\lambda_n(t) = \left(\lambda_n^1(t), \lambda_n^2(t)\right).$$

The skew-extended forward Libor model then follows

$$\begin{aligned} dF_n(t) &= \sum_{k=1}^{2} \lambda_n^k(t)\,\phi(F_n(t))\left(\mu_n^k\left(t, \bar{F}(t)\right)\,dt + dW^k(t)\right), \\ \mu_n^k\left(t, \bar{F}(t)\right) &= \sum_{i=1}^{n} 1_{\{t<T_i\}}\phi(F_i(t))\,\frac{\tau_i}{\tau_i F_i(t) + 1}\lambda_i^k(t), \\ n &= 1, \ldots, M-1. \end{aligned}$$

Cranking through the machinery developed in Section 12.1 we come to the point where we need to postulate a condition on the volatilities similar to (12.5). A naive generalization would specify

$$\begin{aligned} \lambda_n^1(t) &= \sigma_n^1 \times \nu^1(t), \\ \lambda_n^2(t) &= \sigma_n^2 \times \nu^2(t), \\ n &= 1, \ldots, M-1, \quad t \geq 0. \end{aligned} \tag{12.7}$$

A more general extension is possible, however. It is enough to require that

$$\begin{aligned} \lambda_n^1(t) &= \sigma_n^1 \times \nu^{11}(t), \\ \lambda_n^2(t) &= \sigma_n^1 \times \nu^{21}(t) + \sigma_n^2 \times \nu^{22}(t), \\ n &= 1, \ldots, M-1, \quad t \geq 0. \end{aligned} \tag{12.8}$$

The model is still Markov in 2 variables defined by

$$\begin{aligned} dX_1(t) &= \nu^{11}(t)\,dW^1(t) + \nu^{21}(t)\,dW^2(t), \\ dX_2(t) &= \nu^{22}(t)\,dW^2(t). \end{aligned}$$

The condition (12.8) is a lot more useful than (12.7), and here is why. The first volatility component $\lambda_n^1(t)$ is typically positive for all t and n, and in general can be represented in the form (12.7) reasonably well. The second component $\lambda_n^2(t)$, however, is a different matter. For each t, it is positive for n below a certain index, and negative for others. The reason for this can be understood by thinking of different volatility components in terms of principal components (PCA). For a given $t > 0$, $\left\{\lambda_n^1(t)\right\}_{n=1}^{N-1}$ is the first principal component of the join (infinitesimal at time t) moves of the collection of forward Libor rates. It is typically a (more-or-less parallel) shift. The second volatility component $\left\{\lambda_n^2(t)\right\}_{n=1}^{N-1}$, considered for a fixed t across all n, $n = 1, \ldots, M-1$, is related to the second principal component for forward Libor rates, and is typically a twist. A "twist" component must cross zero.

Now let us look at the approximations in (12.7) and (12.8). In the former, a function of the form $\sigma_n^2\nu^2(t)$ can cross zero either "vertically" ($\sigma_n^2\nu^2(t) = 0$ for some $t = t_0$ for all n) or "horizontally" ($\sigma_n^2\nu^2(t) = 0$ for some $n = n_0$ for all t). This is not, however, how $\lambda_n^2(t)$ crosses zero. Because of imposed (or desired) time-homogeneity, $\lambda_n^2(t)$ usually crosses zero "diagonally". (Figure 7 demonstrates the point. On each of the three diagrams, we plot signs of the second volatility component for all points $(t, n) \in [0, T_{N-1}] \times \{1, \ldots, N-1\}$. The "plus" symbol indicates a positive value and the "minus" symbol represents a negative one. Diagrams (A) and (B) represent the only two possibilities for the second volatility component of the form (12.7). The diagram (C) shows how a typical second volatility component really looks like.) This seems to be a fundamental property of the second volatility component, one that is completely missed by (12.7). The representation (12.8) on the other hand does allow for this behavior.

To summarize, (12.8) allows one to capture some fundamental properties of the second volatility component whereas (12.7) does not.

13. Markovian approximation as a control variate

In this section we discuss the real reason *why* one may want to use a Markovian approximation to a forward Libor model, and *how* one should do that.

13.1. **Control variate.** First, let us recall the idea of a variance reduction technique called "control variate". Suppose we have a sequence of random draws $v_1, \ldots, v_J$ from some source (we can think of them as values of a callable Libor exotic for n random paths). We denote the sample mean of these draws by

$$\bar{v}_J = \frac{1}{J} \sum_{j=1}^{n} v_j,$$

and we use it to estimate the true expected value of the sequence,

$$\bar{v}_J \to \mathbf{E}v, \quad J \to \infty.$$

Suppose we can also find another sequence $u_1, \ldots, u_J$ such that

- The sequence $\{u_j\}$ is correlated to $\{v_j\}$; and
- The true expected value of $\{u_j\}$, $\mathbf{E}u$, is known.

Let us form a new sequence

$$w_j = v_j - \alpha \left(u_j - \mathbf{E}u\right)$$

for some α. Then the sample mean of the sequence $\{w_j\}$ can be used as an estimate of the true mean of the sequence $\{v_j\}$ because

$$\bar{w}_J \to \mathbf{E}v, \quad J \to \infty.$$

Why it is beneficial? Let us compute the standard error. We have

$$\begin{aligned} \operatorname{Var} w &= \operatorname{Var}\left(v - \alpha \left(u - \mathbf{E}u\right)\right) \\ &= \operatorname{Var} v + \alpha^2 \operatorname{Var} u - 2\alpha \operatorname{Covar}\left(vu\right). \end{aligned}$$

If $\operatorname{Covar}(vu) > 0$ then, as long as we choose

$$0 < \alpha < \frac{2\operatorname{Covar}\left(vu\right)}{\operatorname{Var} u},$$

we will have

$$\operatorname{Var} w < \operatorname{Var} v,$$

thus resulting in an estimate of $\mathbf{E}v$ with a smaller standard error (i.e. more accurate). In fact, the best value of α is given by

$$\alpha = \frac{\operatorname{Covar}\left(vu\right)}{\operatorname{Var} u}.$$

Note that this is the same as the coefficient of linear regression of v on u. Also note that the stronger the correlation between v and u, the better the error reduction achieved in this scheme. And, as always, error reduction directly translates into the decrease in the number of paths one needs to run to compute the value with suitable precision.

If we identify $\{v_j\}$ with simulated values of a callable Libor exotic, then a typical application of the control variate method would seek another instrument (or instruments) that are "similar" to the CLE but can be valued analytically in the model. Then, by setting $\{u_j\}$ to simulated values of this related instrument, one is hoping that a good correlation can be achieved. This scheme does not really work at all for callable Libor exotics. Even for Bermuda swaptions, if one tries to use underlying European swaptions as control variates, one finds that the correlations are too small to be useful. Besides, European swaptions do not really have a closed-form valuation formula in a context of a forward Libor model, only an approximate one (an approximation that typically only works well for at-the-money swaptions). And of course anything more complicated than a European swaption cannot be valued/approximated in closed form at all.

The reason why European swaptions are such bad control variates is clear. As we mentioned before, a valuation of a CLE is more similar to valuing a barrier option. This effects dominates CLE's valuation. European swaptions of course do not have barrier features, and thus do not match the payoff profile of a CLE.

13.2. **Model control variate.** In the control variate variance reduction method, nothing says that the controlling sequence $\{u_j\}$ must represent values of other instruments in the same model. In fact, in our case, a much better choice is to use values of the *same* instrument (the callable Libor exotic) in a *different* model. What is this other model? Here is where our Markovian approximation comes into play. We propose to use the value of the CLE in the Markovian approximation as the control variable.

The reason why the Markovian approximation is a good candidate for being this "control variate model" is the fact that it can value a CLE very accurately using PDE methods. Also, if a Markovian approximation is done "right" (more on that later), there will be significant correlation between simulated values of the CLE for the "base" forward Libor model and its Markovian approximation.

So how the algorithm would work? First, we create a Markovian approximation for the base forward Libor model. Then we compute the value of the CLE in the base forward Libor model using Monte-Carlo, call it V_{orig}. Then we compute the value of the same instrument in the approximate Markovian model with Monte-Carlo *using the same noise terms*, and call in V_{MC} (using the same noise terms is critical to obtaining good correlation). After that we estimate the regression coefficient between Monte-Carlo samples of the CLE values between the two models (using say first 20% of the simulated values to avoid bias) and call it α. Then we compute the CLE value in the Markovian approximation using the PDE method, V_{PDE}. Finally, we combine all these values to obtain a corrected value,

$$V_{Corrected} = V_{orig} - \alpha \left(V_{MC} - V_{PDE} \right).$$

The noise in $V_{Corrected}$ (as manifested by the standard error) is reduced, while un-biasedness is preserved:

$$\begin{aligned} \mathbf{E} V_{MC} &= V_{PDE}, \\ \mathbf{E} V_{Corrected} &= \mathbf{E} V_{orig}. \end{aligned}$$

Our tests showed that the standard error of simulation was reduced by a factor of 3 to 10 times, resulting in potential speedups of 10 to a 100 times! (of course one needs to work harder to obtain all these quantities; compared to the Monte-Carlo, the PDE valuation is instantaneous, so roughly twice the time is needed to get these accuracy improvements).

A one-dimensional forward Libor model is rarely sufficient for realistic pricing of callable Libor exotics. A two-factor model, however, is often sufficient for all but the most complex ones. For two-dimensional models, the extensions from Section 12.3 should be used.

13.3. **How to do the Markovian approximation right.** For the control variate scheme presented above, the Markovian approximation must be done "right". There are two main issues one needs to be concerned about. We discuss them in turn.

A Monte-Carlo scheme involves discretizing a continuous-time SDE (like (12.1) or (12.4)) into a discrete-time scheme. Such discretized scheme basically converts a sequence of independent standard Gaussian random variables $\xi_1, \ldots, \xi_j, \ldots$ into simulated model variables (in our case, simulated values of forward Libor models). It is very important that the discretization scheme for the Markovian approximation Monte-Carlo be the same as the one for the base forward Libor model! Both discretizations must be based on SDEs that are as similar to each other as possible. Suppose the discretization for the base forward Libor model is based on the following (same as (12.1)),

$$dF_n(t) = \lambda_n(t)\,\phi(F_n(t))\left(\mu_n\left(t, \bar{F}(t)\right) dt + dW(t)\right).$$

Then for the Markovian approximation Monte-Carlo, we must use the SDE of the following form (obtained from (12.4) and (12.2))

$$dF_n(t) = \hat{\lambda}_n(t)\,\phi(F_n(t))\left(\mu_n\left(t, \bar{F}(0)\right) - \frac{1}{2}\hat{\lambda}_n(t)\left(\phi'(F_n(0)) - \phi'(F_n(t))\right) dt + dW(t)\right).$$

Here $\hat{\lambda}_n(t)$ is of the special form

$$\hat{\lambda}_n(t) = \sigma_n \times \nu(t)$$

(or the equivalent for the two-dimensional model). The two SDEs must be discretized with the same time steps and the same integration scheme. And, of course, one has to use the same simulated realizations of Brownian paths $W(\cdot)$ in both schemes.

We can not emphasize this enough: one should not try to be clever and use SDEs that promise a faster discretization algorithm such as (12.6). Using a different scheme will destroy correlations between CLE values in the two models, rendering the scheme useless.

The second issue one needs to address in building a workable Markovian approximation is the following. Significant care must be taken in choosing the volatility structure (scalings σ_n^1 and σ_n^2 and time-dependent functions $\nu^{11}(t), \nu^{21}(t), \nu^{22}(t)$ in (12.8) for the 2D case) for the Markovian projection. The main rule here must be to match the base model's volatility structure, on *a factor-by-factor basis*, as closely as possible. Again this is motivated by the need to generate simulated CLE values in the approximate model that are as highly correlated to the base model's ones as possible. Let us explain what we mean, with a two-dimensional model in mind. The base forward Libor model comes equipped with two volatility functions, one per factor, $\left\{\lambda_n^1(\cdot)\right\}$ and $\left\{\lambda_n^2(\cdot)\right\}$. The parameters of the approximate model should be chosen from the following optimization scheme (or some variation thereof):

$$\sum_{i=1}^{M-1}\sum_{n=1}^{M-1}\left(\lambda_n^1(t_i)-\sigma_n^1\nu^{11}(t_i)\right)^2+\sum_{i=1}^{M-1}\sum_{n=1}^{M-1}\left(\lambda_n^2(t_i)-\left(\sigma_n^1\nu^{21}(t_i)+\sigma_n^2\nu^{22}(t_i)\right)\right)^2\to\min.$$

Basically a least squares fit of the approximate Markovian volatility structure to the base model's volatility structure should be performed.

Note that in this regard we are completely ambivalent about recovering market volatilities (i.e. swaption/cap volatilities) in the approximate Markovian model. We do not care how closely the model matches those. We only care about making the volatility structures look as similar as possible between the two models.

An alternative scheme, where a completely separate calibration of the Markovian approximate model to the market volatilities is performed, does not work for variance reduction purposes at all. While this approach is likely to result in a model that calibrates to the market much better, its volatility structure is going to be so different from the base model's that the control variate method will be useless.

14. Vegas

Vegas are sensitivities with respect to changes in volatilities. Most often, sensitivities to *market* volatilities (i.e. observed term volatilities of caps and swaptions) are required. Forward Libor models are defined in terms of instantaneous volatilities of forward Libor rates. We call these *model* volatilities. The discussion on obtaining good vegas cannot be carried out without a good handle of the way the market and the model volatilities are related. We start by recalling this relationship.

14.1. **Volatility calibration for forward Libor models.** Calibration of forward Libor models is a rich and complicated subject. The reader is referred to the existing body of literature (See Literature Review) for exhaustive details. Here we just consider a simple case. It should provide sufficient understanding of the issues involved without bogging us down with details.

We use the log-normal forward Libor model (5.1) as an example. We only consider a one-factor case, a (significant) simplification.

The forward Libor model calibration algorithm attempts to find model volatilities $\{\lambda_m(\cdot)\}_{m=1}^{M-1}$ to match swaption prices obtained in the model to the market's as close as possible. This is done for some selection of swaptions (and caps). Also, typically, some regularity conditions on the model volatilities are imposed. As closed form values of European swaptions are not available, an approximation formula for swaption volatilities is used. We use the following one (see [AA00]). Let $\sigma(T,S)$ be the term lognormal volatility for a swap rate $R(t,T,S)$, a rate with expiry T and the underlying swap final maturity S. Then

$$\sigma^2(T,S)\,T=\int_0^T\left(\sum_{m=1}^{M-1}w_m(T,S)\,\lambda_m(t)\right)^2dt.$$

Here the weights $w_m(T,S)$ are defined by

$$w_m(T,S)=\frac{F_m(0)}{R(0,T,S)}\frac{\partial R(0,T,S)}{\partial F_m(0)}.$$

The forward Libor volatilities $\lambda_m(\cdot)$ are usually assumed piecewise constant over time intervals $(t_i, t_{i+1}]$. Let $\{\lambda_{mi}\}$ be defined by

$$\lambda_m(t) = \sum_{i=0}^{M-1} \lambda_{mi} 1_{(t_i, t_{i+1}]}(t).$$

Then the formula for the swaption volatility can be rewritten

$$\sigma^2(T,S)\,T = \sum_{i:t_i \le T} (t_{i+1} \wedge T - t_i) \left(\sum_{m=1}^{M-1} w_m(T,S)\,\lambda_{mi} \right)^2. \tag{14.1}$$

Suppose a set of swaptions $\{(T_j, S_j)\}_{j=1}^J$ is chosen. Let their market volatilities be denoted by $\{u_j\}_{j=1}^J$. The objective function for calibration will contain a mispricing term of the form

$$O_1(\{\lambda_{mi}\}) = \sum_{j=1}^J (\sigma(T_j, S_j) - u_j)^2.$$

A second term can be introduced, a term that controls how time-homogeneous the resulting volatility structure is. We define

$$O_2(\{\lambda_{mi}\}) = \sum_{m,i} (\lambda_{mi} - \lambda_{m+1,i+1})^2.$$

Let a relative weight imposed on the time-homogeneity objective versus the market fit objective be α. Then the purpose of calibration is to find

$$\{\lambda_{mi}^0\} = \arg\min (O_1(\{\lambda_{mi}\}) + \alpha O_2(\{\lambda_{mi}\})).$$

This problem is typically solved by numerical non-linear optimization.

An important characteristic of calibration is that it is never exact. Because of the presence of the time-homogeneity term $\alpha O_2(\{\lambda_{mi}\})$, the model values of swaption volatilities for calibrated model volatilities $\{\lambda_{mi}^0\}$ will not exactly match the targets $\{u_j\}_{j=1}^J$. A typical absolute error that is deemed satisfactory is about 0.1% (with typical market swaption volatilities being in the $10\% - 50\%$ range). Even if one desired a perfect fit to market volatilities, removing the term $\alpha O_2(\{\lambda_{mi}\})$ is typically *not* a good solution. The term serves an important task of regularizing the problem. Without it, the optimization problem is usually over-specified and unstable. Other ad-hoc methods for regularizing the problem can be employed in its place (i.e. using only some of the $\{\lambda_{mi}\}$ as independent variables with some interpolation to obtain the rest), but they all result in the same outcome – the error of calibration cannot be made arbitrarily small.

For valuation purposes, this calibration error is not an issue, as the volatility fit accuracy that can be achieved is usually well within the bid-ask spreads.

14.2. **The direct method of computing vegas.** Suppose a forward Libor model is calibrated to a set of swaptions $\{(T_j, S_j)\}_{j=1}^J$ (here T_j is the expiry and S_j is the final maturity date for the swaption j) with market volatilities $\{u_j\}_{j=1}^J$. What we typically need to know is the sensitivity of a callable Libor exotic to a small change in volatility of each one of these swaptions. The direct method of computing vegas works like this. To obtain the vega to the j-th swaption, a small shock to its volatility is applied. The forward Libor model is then recalibrated, as described in the previous section. The CLE in this "bumped" model is then revalued, and subtract the base value is subtracted to obtain the appropriate sensitivity. The procedure is repeated for each j.

The direct method of computing vegas typically does not work. It does not work in the sense that the numbers it produces have usually very little relation to reality, and do not conform to one's intuition as to how vega should look like. In particular, it is not uncommon to get negative vegas for instruments that are known to be vega positive. Tell-tale "ringing" patterns are also common. These problems are apparent in Figure 8, that shows vegas for a Bermuda swaption obtained using the direct method.

The reason it does not work is the following. As explained in the previous section, calibration of a forward Libor model is an inexact procedure. Typical errors in fitting swaption volatilities are of the order 0.1%. This is the same-magnitude error as a shock applied to swaption volatilities in vega computations (significantly larger shocks cannot be applied because local sensitivities are required, and also because a single swaption volatility shocked so much that it is out of line with the rest will just break volatility calibration altogether). In a system where noise is of the same order of magnitude as the signal, distinguishing the signal from the noise is generally impossible! In effect, vegas computed this way are drowned in noise.

Finally, a full forward Libor model calibration is a slow procedure. To solve a non-linear optimization problem with (typically) dozens, if not hundreds, of variables takes a significant amount of time, comparable, if not exceeding, a CLE valuation via Monte Carlo. The extra computational burden makes this method even less appealing.

14.3. **The indirect method of computing vegas.** The mapping of market volatilities to the model ones $\{u_j\}_{j=1}^J \to \{\lambda_{mi}\}$ via the calibration process is noisy and imprecise, as explained in the previous section. The reverse mapping $\{\lambda_{mi}\} \to \{u_j\}_{j=1}^J$ is just a matter of applying the formula (14.1) and, as such, exact. The indirect method of computing vegas utilizes this observation.

In broad terms, the indirect method is based on computing sensitivities of the CLE to *model* volatilities first, and then converting them, via matrix manipulations, to sensitivities to *market* volatilities.

Not all forward Libor volatilities $\{\lambda_{mi}\}$ should necessarily be treated as independent inputs. We can choose a subset of them as individual inputs, with the others being interpolated from the "primary" ones. We choose K forward Libor rate volatilities as "primary"; to avoid multilevel notations we use a vector notation

$$\bar{e} = (e_1, \dots, e_K).$$

All other model volatilities are functions of this vector,

$$\lambda_{mi} = \lambda_{mi}(\bar{e}). \tag{14.2}$$

Similarly, we define market volatilities of swaptions $\{(T_j, S_j)\}_{j=1}^J$, as computed by the model, via

$$\bar{\sigma} = (\sigma_1, \dots, \sigma_J),$$

with

$$\begin{aligned} \bar{\sigma} &= \bar{\sigma}(\bar{e}), \\ \sigma_j &= \sigma_j(\bar{e}), \end{aligned}$$

via formula (14.1) and relationship (14.2).

Finally, let us define the value of a CLE as a function of primary model volatilities $\bar{e}$ by

$$h(\bar{e}).$$

We are ready to present the main idea. By the chain rule

$$\frac{\partial h}{\partial e_k} = \sum_j \frac{\partial h}{\partial \sigma_j} \frac{\partial \sigma_j}{\partial e_k}, \quad k = 1, \ldots, K.$$

The quantities

$$\begin{aligned} \bar{\nu} &= (\nu_1, \ldots, \nu_J), \\ \nu_j &\triangleq \frac{\partial h}{\partial \sigma_j}, \quad j = 1, \ldots, J, \end{aligned}$$

are the market vegas we are seeking. The quantities

$$\begin{aligned} \bar{\xi} &= (\xi_1, \ldots, \xi_K), \\ \xi_k &\triangleq \frac{\partial h}{\partial e_k}, \quad k = 1, \ldots, K, \end{aligned}$$

are the model vegas. These can be computed by bumping each forward Libor volatility and revaluing the CLE. The matrix

$$\begin{aligned} G &= \{G_{kj}\}, \\ G_{kj} &= \frac{\partial \sigma_j}{\partial e_k}, \end{aligned}$$

can be easily computed by numerically differentiating equations (14.1). Then, obtaining market vegas is just a matter of solving the following linear system

$$G\bar{\nu} = \bar{\xi} \tag{14.3}$$

for $\bar{\nu}$.

While in theory this is all one needs to do, in practice there are a few kinks to be sorted out before obtaining a working algorithm.

It helps to think of model volatility bumps as various volatility scenarios applied to the model. Vega hedging is about constructing a portfolio of swaptions that insulates the CLE from volatility changes as defined by these scenarios. Market vegas $\bar{\nu}$ are interpreted as the amounts of each swaption used in the hedging portfolio. (In this interpretation it is clear that a volatility scenario need not be as simple as a bump to a single forward Libor volatility e_k. It can be a more complicated shock to the whole forward Libor volatility structure. It is often beneficial to choose these shocks judiciously.)

The user typically has control over the number of volatility scenarios J that are applied, and the number of hedging instruments K to use. Let us discuss implications of different choices of J and K.

What happens if there are more scenarios than hedging instruments available ($J < K$)? In this case, the CLE cannot be insulated to all scenarios completely. The problem (14.3) should really be then interpreted as a least squares fit problem,

$$\left\| G\bar{\nu} - \bar{\xi} \right\|^2 \to \min. \tag{14.4}$$

This choice typically results in a more robust and stable set of market vegas $\bar{\nu}$.

What if we wanted to use more hedging instruments than there were volatility scenarios ($J > K$)? Now there are too many degrees of freedom. Multiple hedging portfolios $\bar{\nu}$ can be found that would provide a perfect hedge in all volatility scenarios. This is of course not good from the prospective of trying to obtain a stable set of vegas. There are different ways of dealing with this issue. One can just use some method of solving the system (14.3) that works "behind the scenes" to choose a particular solution. A primary candidate of such a method is the Singular Value Decomposition (SVD) method, see [GvL89]. The method will find a solution in the K dimensional space that is orthogonal to the null-space of G. While certainly requiring the least amount of thought, this approach is not the best. The constraints the method uses to choose one solution from many may make sense from the numerical analysis point of view, but not necessarily from the financial point of view. It is by far much better to impose some regularity conditions on the system (14.3) externally. This way the conditions can be chosen to have a *financial* meaning. For example, the condition that the market vegas were smooth across the swaption grid could be imposed. Or, a minimum-entropy hedging portfolio can be sought.

We have found that the simplest condition that works well is the minimum size condition. The cost of setting up a hedging portfolio is proportional to its size. The size is defined as the sum of absolute values of the notionals (a bid/ask spread is paid irrespective of whether one goes long or short a swaption). Replacing, for numerical purposes, absolute values of notionals with their squares, and replacing hedging notionals with vegas for simplicity of exposition, we formulate the following problem:

> Find the hedging portfolio that has the smallest sum of squared vegas and satisfies (14.3).

While constrained quadratic optimization problems can be solved relatively easily (see [Num95]), it is still easier to work with unconstrained ones. So problem we solve at the end is

> Find $\bar{\nu}$ that satisfies

$$\left\|G\bar{\nu} - \bar{\xi}\right\|^2 + \delta \left\|\bar{\nu}\right\|^2 \to \min . \tag{14.5}$$

The weight δ defines relative importance of having a small sum of vegas squared versus hedging all scenarios perfectly.

The problem (14.5) is well-posed and is trivial to solve using linear algebra.

Finally, let us look at the case $J = K$, where we choose exactly the same number of hedging swaptions as there are volatility scenarios. The matrix G is square and the equation (14.3) can be solved in the "normal" sense. We find, however, that even though it is square, solving (14.3) may not be easy at all. Unfortunately, the matrix G is typically nearly singular. While theoretically it is not, from a numerical prospective it often is. The smallest eigenvalues are usually so small that the direct solution is dominated by numerical noise. This is an unfortunate consequence of the "stiffness" of the set of equations (14.1).

The options for dealing with this problem are the same as for the case $J > K$ discussed earlier. As in that case, the method that works the best is to replace (14.3) with (14.5).

Finally, even in the case $J < K$, it is often useful to control the smoothness of vegas directly, and add a smoothness term $\delta \left\|\bar{\nu}\right\|^2$ to the problem (14.4).

Let us summarize our recommendations. In all cases, the best way to obtain market vegas from the system (14.3) is by solving a quadratic optimization problem (14.5). The weight δ must be chosen experimentally, by balancing the requirements of good hedging performance and smoothness of vegas. An example of a vega profile for a Bermuda swaption, obtained using this method, is shown in Figure 9.

14.4. **Computing model vegas.** One part of computing vegas in the indirect method is computing model vegas, i.e. sensitivities of the callable Libor exotic to small changes in the model volatility parameters (denoted by $\bar{e}$ in the previous section). The problem of computing these sensitivities is very similar to the problem of computing deltas, and the same considerations apply. The need to control noise is even more critical for model vegas. This is so because the "inversion" procedure of computing market vegas from model ones is, as we explained, a numerically not well-conditioned problem. Small noise in model vegas get amplified significantly in the inversion. While this is somewhat mitigated by the "smoothing inversion" (14.5), it is still a lot better not to have the noise in model vegas to begin with.

Since obtaining good model vegas is a lot like obtaining good deltas, the same methods work for both. We do not go into much detail here – extensions of the methods to vegas are completely straightforward.. Here are some points that one should keep in mind.

- Just like for deltas, the main source of noise for model vegas is the jumps in the exercise indicators.
- Pathwise model vegas can be computed the same way as pathwise deltas. The equation (10.5) holds for the pathwise model vegas, and the SDEs for vegas can easily be derived by differentiating the SDEs for primary Libor rates with respect to volatilities.
- The smoothing of the payoffs in the "sausage" Monte-Carlo benefits vegas as well as deltas.
- The Markovian approximation-based variance reduction does not really work for model vegas. The reason has to do with the connection between the volatilities of the original model and the Markovian approximation. Recall that the volatilities of the Markovian approximation are obtained in a least-squares fit to the volatilities of the base model (See Section 13.3). Because of this "loose" linkage, small bumps in the base model's volatilities cannot be reliably mapped to small bumps to the Markovian approximation's volatilities. The noise introduced at this stage typically renders the Markovian approximation method useless for model vegas.

15. Conclusions

Building a pricing and risk-management framework for callable Libor exotics is not an easy task. Mindless applications of generic techniques lead, here more than anywhere, to algorithms that are numerically unstable, theoretically dubious, and ultimately unworkable. One can succeed only by understanding deep, intimate details of the instruments, the fine-grained structure of the models, and the interplay between theoretical foundations and numerical tricks. Hopefully, theoretical and practical insights presented in this paper bring one closer to this goal.

References

[AA00] Leif B.G. Andersen and Jesper Andreasen. Volatility skews and extensions of the Libor Market Model. *Applied Mathematical Finance*, 7:1–32, March 2000.

[AA01] Leif B.G. Andersen and Jesper Andreasen. Factor dependence of Bermudan swaption prices: Fact or fiction? *Journal of Financial Economics*, 62:3–37, 2001.

[AB01] Leif B.G. Andersen and Mark Broadie. A primal-dual simulation algorithm for pricing of multidimensional options. Working Paper, 2001.

[ABR01] Leif B.G. Andersen and Rupert Brotherton-Ratcliffe. Extended libor market models with stochastic volatility. Working paper, 2001.

[Ale02] Carol Alexander. Common correlation structures for calibrating the Libor model. SSRN Working paper, 2002.

[And99] Leif B.G. Andersen. A simple approach to the pricing of Bermudan swaptions in the multi-factor Libor Market Model. Gen Re working paper, 1999.

[And01] Jesper Andreasen. The pricing of Bermuda swaptions. Risk conference, Paris, 2001.

[BBG97] P. Boyle, M. Broadie, and P. Glasserman. Simulation methods for security pricing. *Journal of Economic Dynamics and Control*, 21:1267–1321, 1997.

[Ber99] H.-P. Bermin. A general approach to hedging options: Applications to barrier and partial barrier options. Working paper, Lund University, Lund, Sweden, 1999.

[BG97] Mark Broadie and Paul Glasserman. Pricing American style securities using simulation. *Journal of Economic Dynamics and Control*, 21(8-9):1323–1352, 1997.

[BGM96] A. Brace, D. Gatarek, and M. Musiela. The market model of interest rate dynamics. *Mathematical Finance*, 7:127–154, 1996.

[BM01] D. Brigo and F. Mercurio. *Interest-Rate Models - Theory and Practice*. Springer Verlag, 2001.

[Dod02] Stephen Dodds. Pricing and hedging Bermudan swaptions with the BGM model. Barclays working paper, 2002.

[FLLT99] Eric Fournie, Jean-Michel Lebuchoux, Pierre-Louis Lions, and Nizar Touzi. Application of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3:391–412, 1999.

[Gat02] Dariusz Gatarek. Calibration of the Libor market model: Three prescriptions. SSRN Working paper, 2002.

[GvL89] Gene H. Golub and Charles F. van Loan. *Matrix Computations*. The John Hopkins University Press, 1989.

[GY03] Paul Glasserman and Bin Yu. Number of paths versus number of basis functions in American option pricing. Working paper, 2003.

[GZ99] Paul Glasserman and Xiaoling Zhao. Fast greeks in forward Libor models. *Journal of Computational Finance*, 3:5–39, 1999.

[Jam97] F. Jamshidian. Libor and swap market models and measures. *Finance and Stochastics*, 1(4):293–330, 1997.

[LM02] A. Lane and D. Marris. Fast calibration of Libor market models. Working paper, 2002.

[LS98] Francis Longstaff and Eduardo Schwartz. Valuing american options by simulation: A simple least-squares approach. Working paper, The Anderson School, UCLA, 1998.

[LSSC99] Francis Longstaff, Eduardo Schwartz, and E. Santa-Clara. Throwing away a billion dollars: The cost of suboptimal exercise in the swaptions market. Working paper, The Anderson School, UCLA, 1999.

[MR97] Marek Musiela and Marek Rutkowski. *Martingale Methods in Financial Modeling*. Springer, 1997.

[Nou99] Mouaoya Noubir. Prcing American and Bermudan interest rate options in the market models. Working paper, 1999.

[Num95] Visual Numerics. *IMSL C/Math/Library. C functions for mathematical applications*. 1995.

[Ped98] Morten Bjerregaard Pedersen. Calibrating Libor market models. SSRN Working paper, 1998.

[Ped99] Morten Bjerregaard Pedersen. Bermuda swaptions in the LIBOR market model. SSRN Working paper, 1999.

[Pit03] Vladimir V. Piterbarg. Computing deltas of callable LIBOR exotics in a forward LIBOR model. SSRN Working paper, 2003.

[PP03] Raoul Pietersz and Antoon Pelsser. Risk managing Bermudan swaptions in the Libor BGM model. SSRN working paper, 2003.

[PPvR02] Raoul Pietersz, Antoon Pelsser, and Marcel van Regenmortel. Fast drift approximated pricing in the BGM model. Erasmus University working paper, 2002.

[Reb98] Riccardo Rebonato. *Interest-Rate Option Models*. John Wiley and Son, 1998.

[Reb99] Riccardo Rebonato. *Volatility and Correlation*. John Wiley and Son, 1999.

[Reb02] Riccardo Rebonato. *Modern pricing of interest rate derivatives: the Libor market model and beyond*. Princeton University Press, 2002.

[Reb03] Riccardo Rebonato. Term structure models: A review. QUARC Working paper, 2003.

[Rog01] L.C.G Rogers. Monte Carlo valuation of American options. Working Paper, 2001.
[RS95] P. Ritchen and L. Sanskarasubramanyan. Volatility structures of the forward rates and the dynamics of the term structure. *Mathematical Finance*, 5:55–72, 1995.
[SC00] J. Schoenmakers and B. Coffey. Stable implied calibration of a multi-factor Libor model via a semi-parametric correlation structure. Working paper, 2000.
[Sid00] Jakob Sidenius. Libor market models in practice. *Journal of Computational Finance*, 3(Spring):75–99, 2000.
[Sve02] Mikkel Svenstrup. On the suboptimality of single-factor exercise strategies for Bermudan swaptions. SSRN Working paper, 2002.
[Wu03] Lixin Wu. Fast at-the-money calibration of the Libor market model using Lagrange multipliers. *Computational Finance*, 6(2), 2003.

Appendix A. "Sausage" Monte Carlo derivation

Recall the definition

$$q_i\left(\omega_j\right) \triangleq \mathbf{E}\left(1_{i\geq\tilde{\eta}(\omega)}\middle|\, A_j^\varepsilon\right).$$

By expressing A_j^ε in terms of the "overshoot" function ϕ (defined by (11.5)), we obtain

$$q_i\left(\omega_j\right) = 1-\mathbf{P}\left(\bigcap_{n=1}^{i}\left\{\phi_n\left(x\left(T_n,\omega\right)\right)\leq 0\right\}\middle|\, A_j^\varepsilon\right),$$

$$1-q_i\left(\omega_j\right) = \mathbf{P}\left(\bigcap_{n=1}^{i}\left\{\phi_n\left(x\left(T_n,\omega\right)\right)\leq 0\right\}\middle|\, A_j^\varepsilon\right). \tag{A.1}$$

We claim that, to order ε,

$$1-q_i\left(\omega_j\right) = \prod_{n=1}^{i}\mathbf{P}\left(\left\{\phi_n\left(x\left(T_n,\omega\right)\right)\leq 0\right\}\middle|\, A_{jn}^\varepsilon\right). \tag{A.2}$$

The proof follows by repeated applications of Lemma B.1 from Appendix B. Here is a less formal argument. Conditioning on A_j^ε $(=\bigcap_{i=1}^{N-1}A_{ji}^\varepsilon)$ basically pins down the Markov process at times T_i, $i=1,\ldots,N-1$ to known values $\left(x\left(T_i,\omega_j\right)\right)$ with ε-accuracy. If a Markov process becomes conditioned on being at a certain state on a given date, past and future events become conditionally independent. Then, the set intersection on the right-hand side of (A.1) can be "unwrapped" into a product on the right-hand side of (A.2).

Define

$$p_n\left(\omega_j\right) = 1-\mathbf{P}\left(\left\{\phi_n\left(x\left(T_n\right),\omega\right)\leq 0\right\}\middle|\, A_{jn}^\varepsilon\right),$$

so that

$$1-q_i\left(\omega_j\right) = \prod_{n=1}^{i}\left(1-p_n\left(\omega_j\right)\right).$$

We need to be able to compute

$$\begin{aligned} p_n\left(\omega_j\right) &= \mathbf{P}\left(\left\{\phi_n\left(x\left(T_n\right),\omega\right)>0\right\}\middle|\, A_{jn}^\varepsilon\right) \\ &= \mathbf{P}\left(\left\{\phi_n\left(x\left(T_n,\omega\right)\right)>0\right\}\middle|\, \left\|x\left(T_n,\omega\right)-x\left(T_n,\omega_j\right)\right\|<\varepsilon\right). \end{aligned}$$

The function ϕ_n is typically a smooth function of x. For x such that $\|x(T_n,\omega) - x(T_n,\omega_j)\| < \varepsilon$, for small $\varepsilon > 0$, we have that

$$\begin{aligned} \phi_n(x) &\approx \phi_{nj} + \nabla\phi_{nj} \times (x - x(T_n,\omega_j)), \\ \phi_{nj} &\triangleq \phi_n(x(T_n,\omega_j)), \\ \nabla\phi_{nj} &\triangleq \nabla\phi_n(x)|_{x=x(T_n,\omega_j)}, \end{aligned}$$

(here $\nabla\phi$ is the gradient of ϕ). Define $\Phi_n \subset \mathbb{R}$ to be the image of the ball $\{\|x(T_i,\omega) - x(T_i,\omega_j)\| < \varepsilon\}$ under the above mapping,

$$\begin{aligned} \Phi_n &= \phi_n(\{\|x(T_i,\omega) - x(T_i,\omega_j)\| < \varepsilon\}), \\ \Phi_n &\in \mathbb{R}. \end{aligned}$$

Approximately,

$$\Phi_n \approx \left[\phi_{nj} - \left\|\nabla\phi_{nj}\right\|\varepsilon, \phi_{nj} + \left\|\nabla\phi_{nj}\right\|\varepsilon\right],$$

(here $\left\|\nabla\phi_{nj}\right\|$ denotes the norm of the linear operator $\nabla\phi_{nj}$). Under the approximation

$$\begin{aligned} A^{\varepsilon}_{jn} &= \{\omega : \|x(T_i,\omega) - x(T_i,\omega_j)\| < \varepsilon\} \\ &\approx \{\omega : \phi_n(x(T_i,\omega)) \in \Phi_n\} \end{aligned}$$

we get

$$p_n(\omega_j) \approx \mathbf{P}\left(\phi_n(x(T_n,\omega)) > 0 \,|\phi_n(x(T_i,\omega)) \in \Phi_n\right).$$

Approximating conditional distribution of $\phi_n(x(T_n,\omega))$ by a uniform one on the set Φ_n we obtain

$$\begin{aligned} p_n(\omega_j) &\approx \frac{|\{\phi > 0\} \cap \Phi_n|}{|\Phi_n|} \\ &= \frac{\left|\{\phi > 0\} \cap \left[\phi_{nj} - \left\|\nabla\phi_{nj}\right\|\varepsilon, \phi_{nj} + \left\|\nabla\phi_{nj}\right\|\varepsilon\right]\right|}{\left|\left[\phi_{nj} - \left\|\nabla\phi_{nj}\right\|\varepsilon, \phi_{nj} + \left\|\nabla\phi_{nj}\right\|\varepsilon\right]\right|}. \end{aligned}$$

where we use $|\cdot|$ to denote the length of intervals in $\mathbb{R}$. Let us denote

$$\delta_{nj} = \left\|\nabla\phi_{nj}\right\| \times \varepsilon.$$

Then,

$$\left|\left[\phi_{nj} - \left\|\nabla\phi_{nj}\right\|\varepsilon, \phi_{nj} + \left\|\nabla\phi_{nj}\right\|\varepsilon\right]\right| = 2\delta_{nj},$$

and

$$p_n(\omega_j) = \begin{cases} 1, & \phi_{nj} - \delta_{nj} > 0, \\ \frac{\delta_{nj} + \phi_{nj}}{2\delta_{nj}} & \phi_{nj} - \delta_{nj} < 0 < \phi_{nj} + \delta_{nj}, \\ 0, & \phi_{nj} + \delta_{nj} < 0. \end{cases} \tag{A.3}$$

This completes the derivation.

Appendix B. Proof of approximate conditional independence for "sausage" Monte-Carlo

Lemma B.1. *Let $x(t,\omega)$, $t\in[0,T]$, be a Markov process with a state space $\mathbb{R}^n$ for some $n\geq 1$. Assume its transitional density*

$$\mathbf{P}\left(x(t)=z\middle|\, x(s)=y\right)$$

is differentiable in z and y for all t, $s>0$. Chose arbitrary

$$(T_1,x_1)\,,\quad (T_2,x_2)\,.$$

Also choose $\varepsilon>0$ and define

$$U_i^\varepsilon=\left\{\omega:\|x(T_i,\omega)-x_i\|<\varepsilon\right\},\quad i=1,2.$$

Let X_1 and X_2 be two subsets of the state space $\mathbb{R}^n$ and define

$$V_i=\left\{\omega:x(T_i,\omega)\in X_i\right\},\quad i=1,2.$$

Then

$$\mathbf{P}\left(V_1\cap V_2\middle|\, U_1^\varepsilon\cap U_2^\varepsilon\right)=\mathbf{P}\left(V_1\middle|\, U_1^\varepsilon\right)\mathbf{P}\left(V_2\middle|\, U_2^\varepsilon\right)\left(1+O(\varepsilon)\right)$$

as $\varepsilon\to 0$.

Proof. We have,

$$\mathbf{P}\left(V_1\cap V_2\middle|\, U_1^\varepsilon\cap U_2^\varepsilon\right)=\frac{\mathbf{P}\left(V_1\cap V_2\cap U_1^\varepsilon\cap U_2^\varepsilon\right)}{\mathbf{P}\left(U_1^\varepsilon\cap U_2^\varepsilon\right)}.$$

For the expression in the numerator we have

$$\begin{aligned}\mathbf{P}\left(V_1\cap V_2\cap U_1^\varepsilon\cap U_2^\varepsilon\right) &= \mathbf{E}\left(1_{V_1}\times 1_{V_2}\times 1_{U_1^\varepsilon}\times 1_{U_2^\varepsilon}\right)\\ &= \mathbf{E}\left(1_{V_1}\times 1_{U_1^\varepsilon}\times \mathbf{E}\left(1_{V_2}\times 1_{U_2^\varepsilon}\middle|\, x(T_1),U_1^\varepsilon\right)\right)\\ &= \mathbf{E}\left(1_{V_1}\times 1_{U_1^\varepsilon}\times \mathbf{E}\left(1_{V_2}\times 1_{U_2^\varepsilon}\middle|\, x(T_1)=x_1\right)\right)\left(1+O(\varepsilon)\right).\end{aligned}$$

We used the differentiability of the transition density in the last equality. Since the expectation $\mathbf{E}\left(1_{V_2}\times 1_{U_2^\varepsilon}\middle|\, x(T_1)=x_1\right)$ is not random it can be pulled out from under the expectation to yield

$$\mathbf{P}\left(V_1\cap V_2\cap U_1^\varepsilon\cap U_2^\varepsilon\right)=\mathbf{E}\left(1_{V_1}\times 1_{U_1^\varepsilon}\right)\times\mathbf{E}\left(1_{V_2}\times 1_{U_2^\varepsilon}\middle|\, x(T_1)=x_1\right)\times\left(1+O(\varepsilon)\right).$$

Now

$$\begin{aligned}\mathbf{E}\left(1_{V_2}\times 1_{U_2^\varepsilon}\middle|\, x(T_1)=x_1\right) &= \mathbf{E}\left(1_{V_2}\middle|\, x(T_1)=x_1,U_2^\varepsilon\right)\mathbf{E}\left(1_{U_2^\varepsilon}\middle|\, x(T_1)=x_1\right)\\ &= \left(\mathbf{E}\left(1_{V_2}\middle|\, x(T_1)=x_1,x(T_2)=x_2\right)\times\left(1+O(\varepsilon)\right)\right)\mathbf{E}\left(1_{U_2^\varepsilon}\middle|\, x(T_1)=x_1\right),\end{aligned}$$

where again we used the regularity properties of the transition density. By the Markovian property and the regularity of density,

$$\begin{aligned}\mathbf{E}\left(1_{V_2}\middle|\, x(T_1)=x_1,x(T_2)=x_2\right) &= \mathbf{E}\left(1_{V_2}\middle|\, x(T_2)=x_2\right)\\ &= \mathbf{E}\left(1_{V_2}\middle|\, U_2^\varepsilon\right)\left(1+O(\varepsilon)\right).\end{aligned}$$

Therefore, up to $O(\varepsilon)$,

$$\mathbf{P}\left(V_1\cap V_2\cap U_1^\varepsilon\cap U_2^\varepsilon\right)=\mathbf{E}\left(1_{V_1}\times 1_{U_1^\varepsilon}\right)\mathbf{E}\left(1_{V_2}\middle|\, x(T_2)=x_2\right)\mathbf{E}\left(1_{U_2^\varepsilon}\middle|\, x(T_1)=x_1\right)\left(1+O(\varepsilon)\right).$$

Hence

$$\begin{aligned}\mathbf{P}\left(V_1 \cap V_2 \middle| U_1^\varepsilon \cap U_2^\varepsilon\right) &= \frac{\mathbf{E}\left(1_{V_1} \times 1_{U_1^\varepsilon}\right)\mathbf{E}\left(1_{V_2}\middle| U_2^\varepsilon\right)\mathbf{E}\left(1_{U_2^\varepsilon}\middle| x\left(T_1\right)=x_1\right)\left(1+O\left(\varepsilon\right)\right)}{\mathbf{E}\left(1_{U_1^\varepsilon}\right)\mathbf{E}\left(1\middle| U_2^\varepsilon\right)\mathbf{E}\left(1_{U_2^\varepsilon}\middle| x\left(T_1\right)=x_1\right)\left(1+O\left(\varepsilon\right)\right)}\\ &= \frac{\mathbf{E}\left(1_{V_1} \times 1_{U_1^\varepsilon}\right)}{\mathbf{E}\left(1_{U_1^\varepsilon}\right)} \times \frac{\mathbf{E}\left(1_{V_2}\middle| U_2^\varepsilon\right)}{\mathbf{E}\left(1\middle| U_2^\varepsilon\right)} \times \left(1+O\left(\varepsilon\right)\right)\\ &= \mathbf{P}\left(V_1\middle| U_1^\varepsilon\right)\mathbf{P}\left(V_2\middle| U_2^\varepsilon\right)\left(1+O\left(\varepsilon\right)\right),\end{aligned}$$

as claimed.

Appendix C. Figures

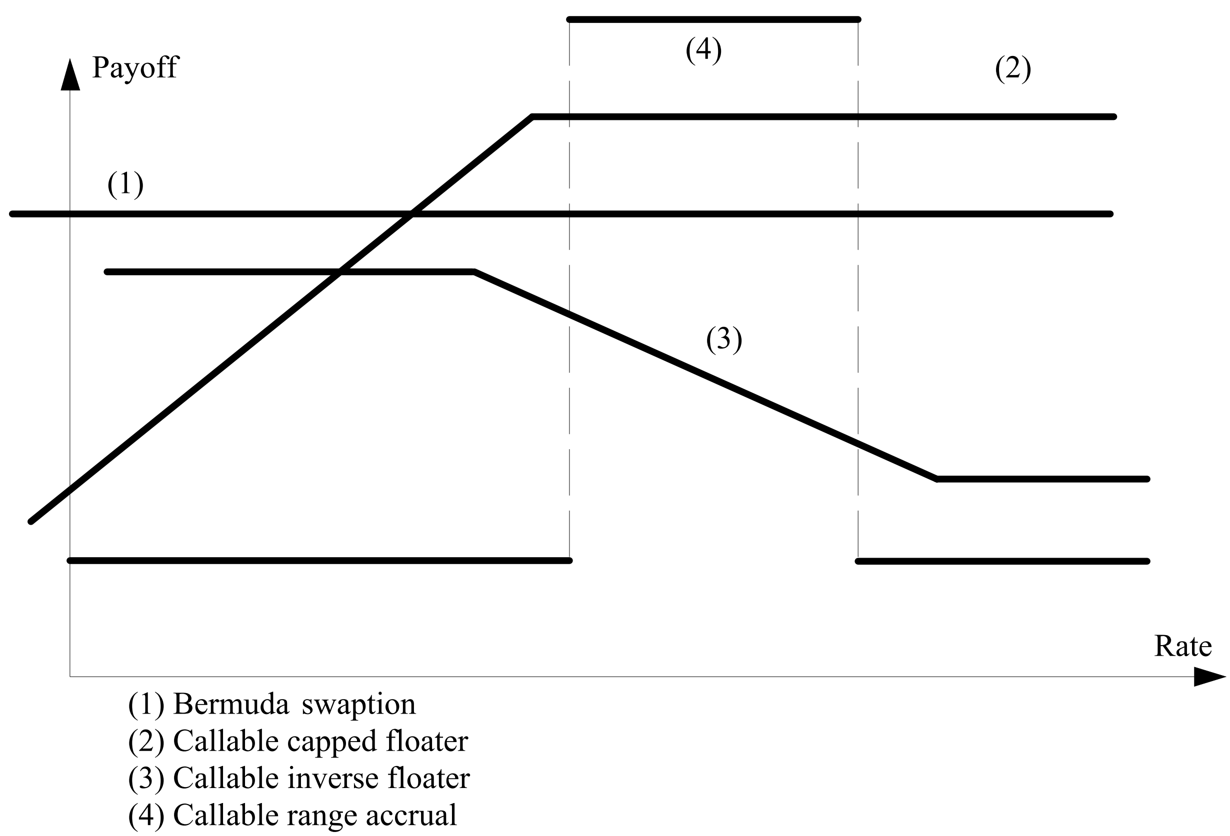


Figure 1. Coupon payoff profiles for various callable Libor exotics

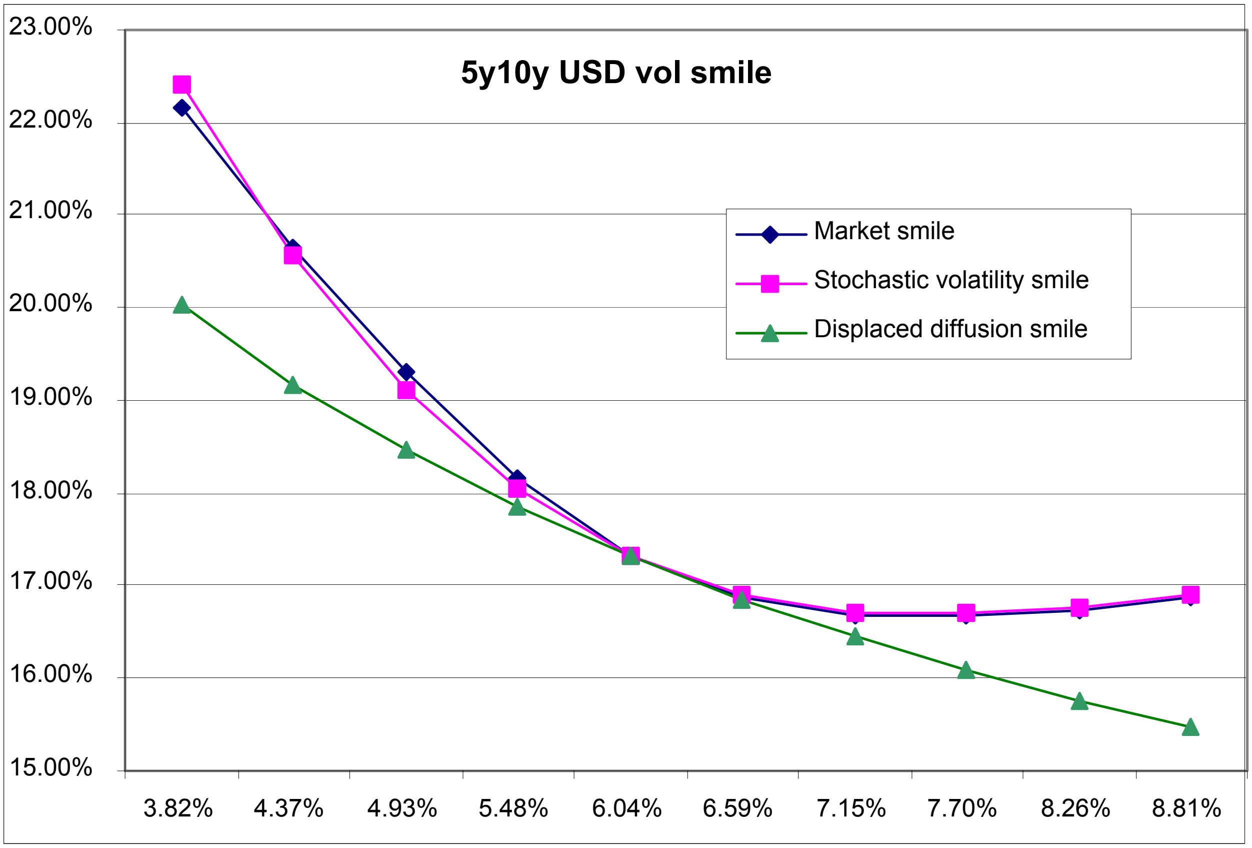


Figure 2. Volatility smile (in implied Black-Scholes volatilities) versus strike.

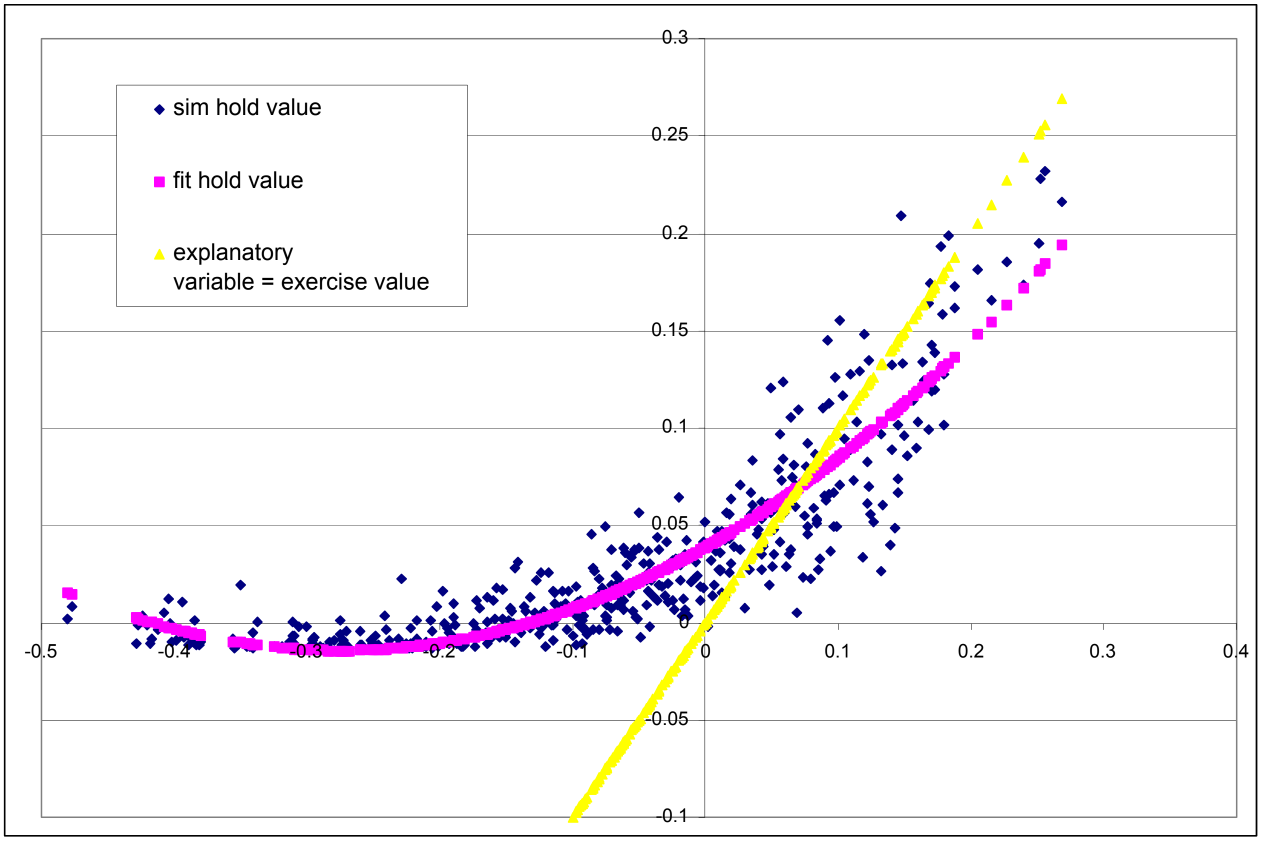


Figure 3. A scatterplot of simulated and regressed (fit) hold values as a function of the exercise value (used as an explanatory variable) for a standard Bermuda swaption, for one of the exercise dates.

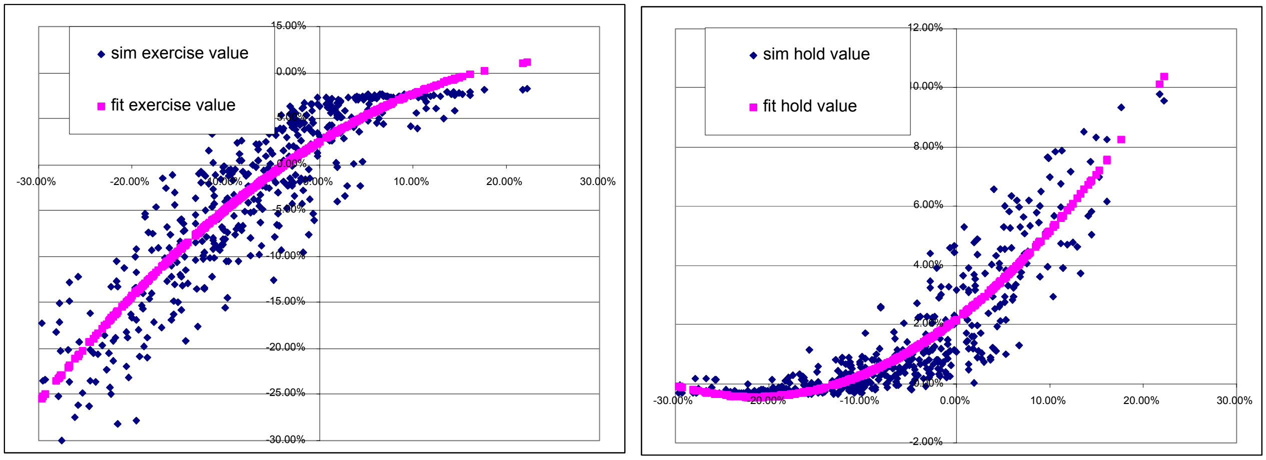


FIGURE 4. A scatterplot of simulated and regressed, or fit, exercise and hold values as functions of the explanatory variable for a callable capped floater for a particular exercise date. A case of an explanatory variable with no curvature. Exercise values on the left, hold values on the right.

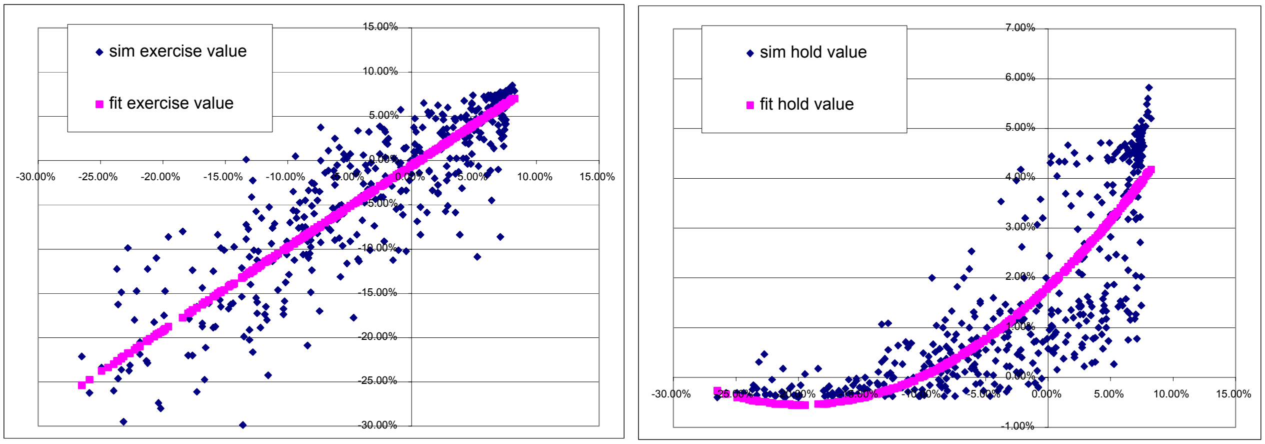


FIGURE 5. A scatterplot of simulated and regressed, or fit, exercise and hold values as functions of the explanatory variable for a callable capped floater for a particular exercise date. A case of an explanatory variable with no curvature. Exercise values on the left, hold values on the right.

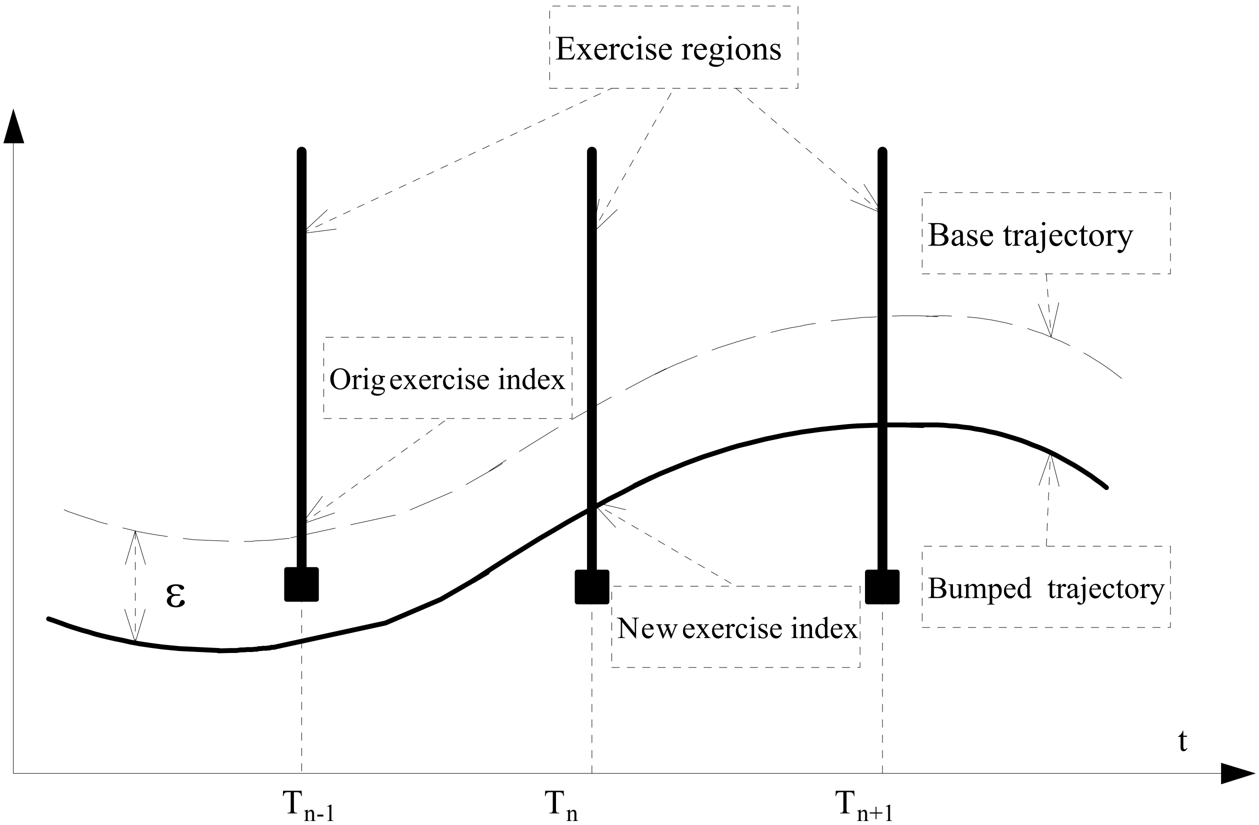
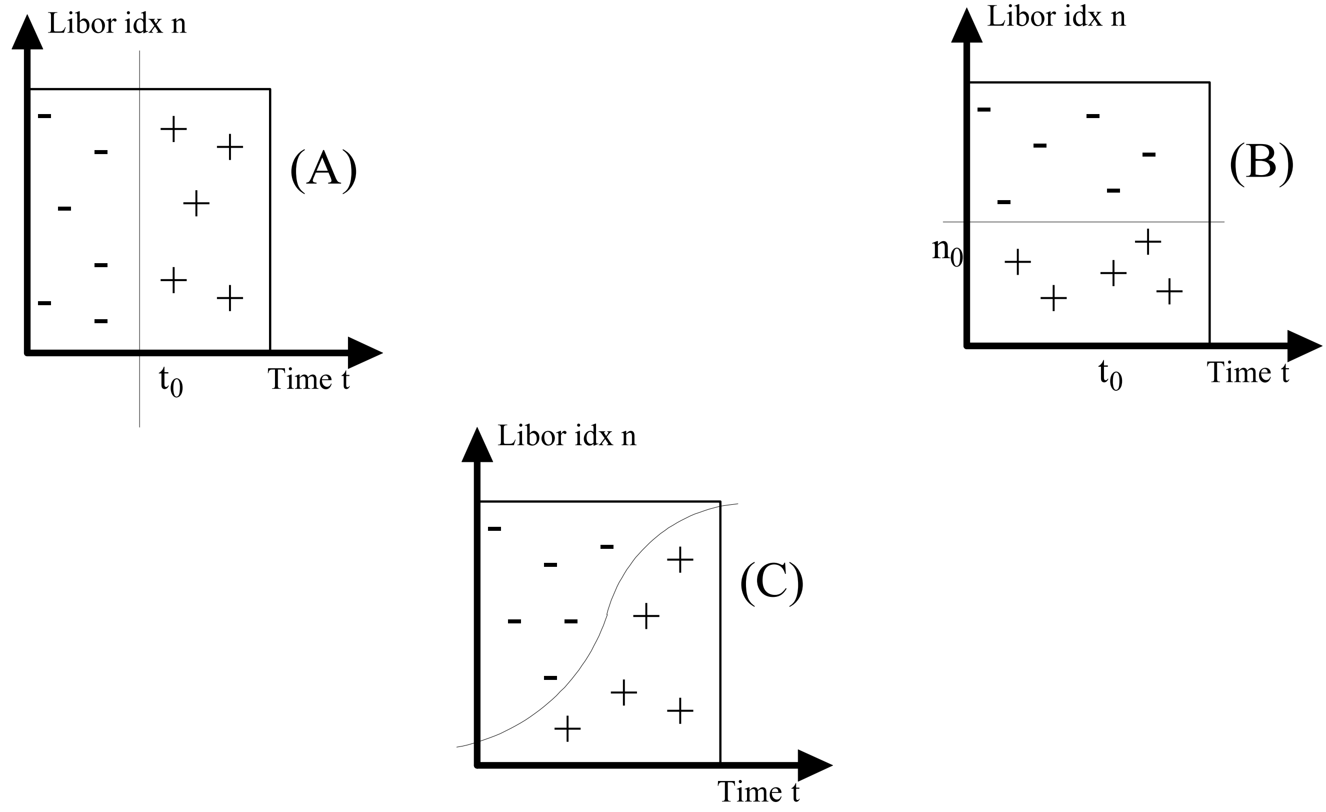


Figure 6. Non-smooth dependence of an exercise index on a simulated path



Sign of the second volatility component for the forward Libor model

FIGURE 7

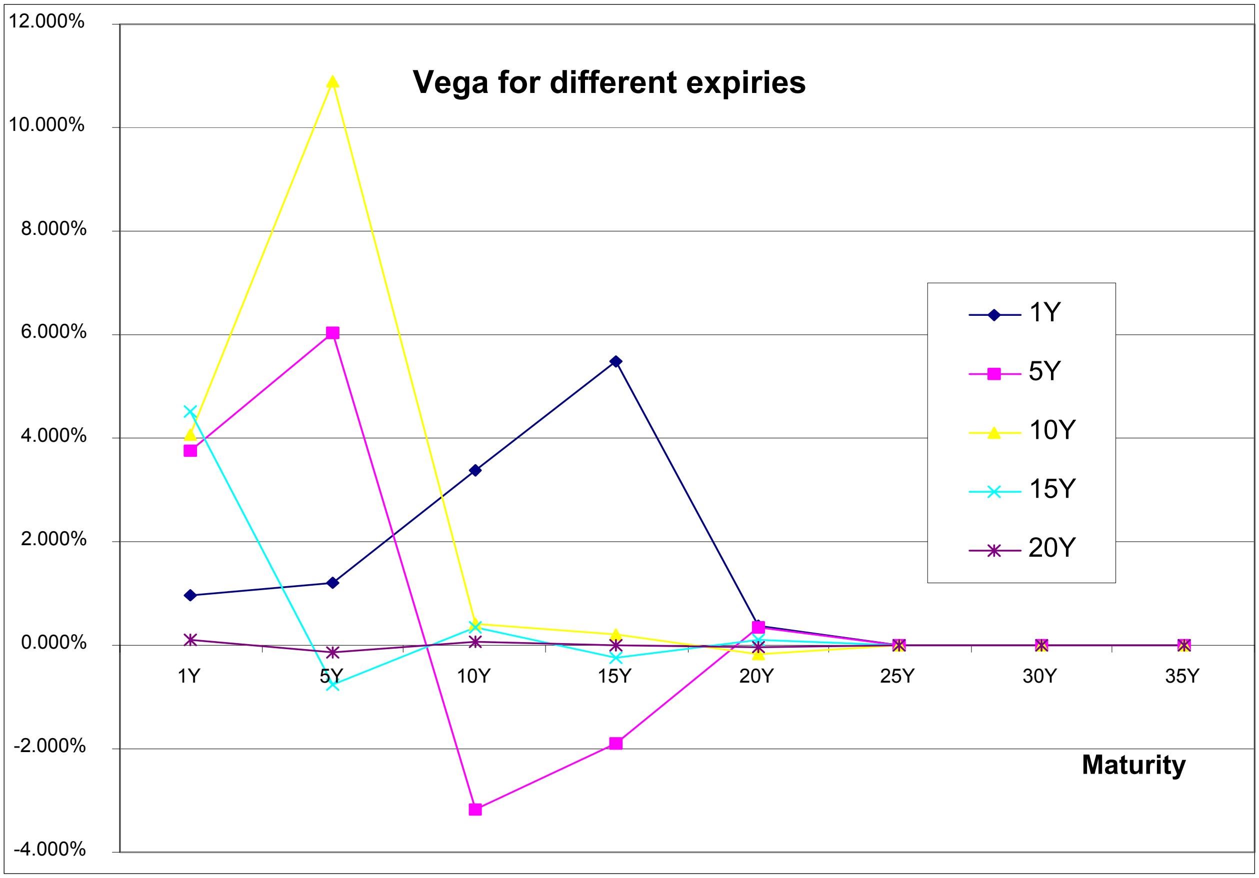


FIGURE 8. A vega grid for a callable Libor exotic, computed using the direct method. Different expiries are represented by different lines, with the x-axis representing different maturities. Unscaled vega per unit notional.

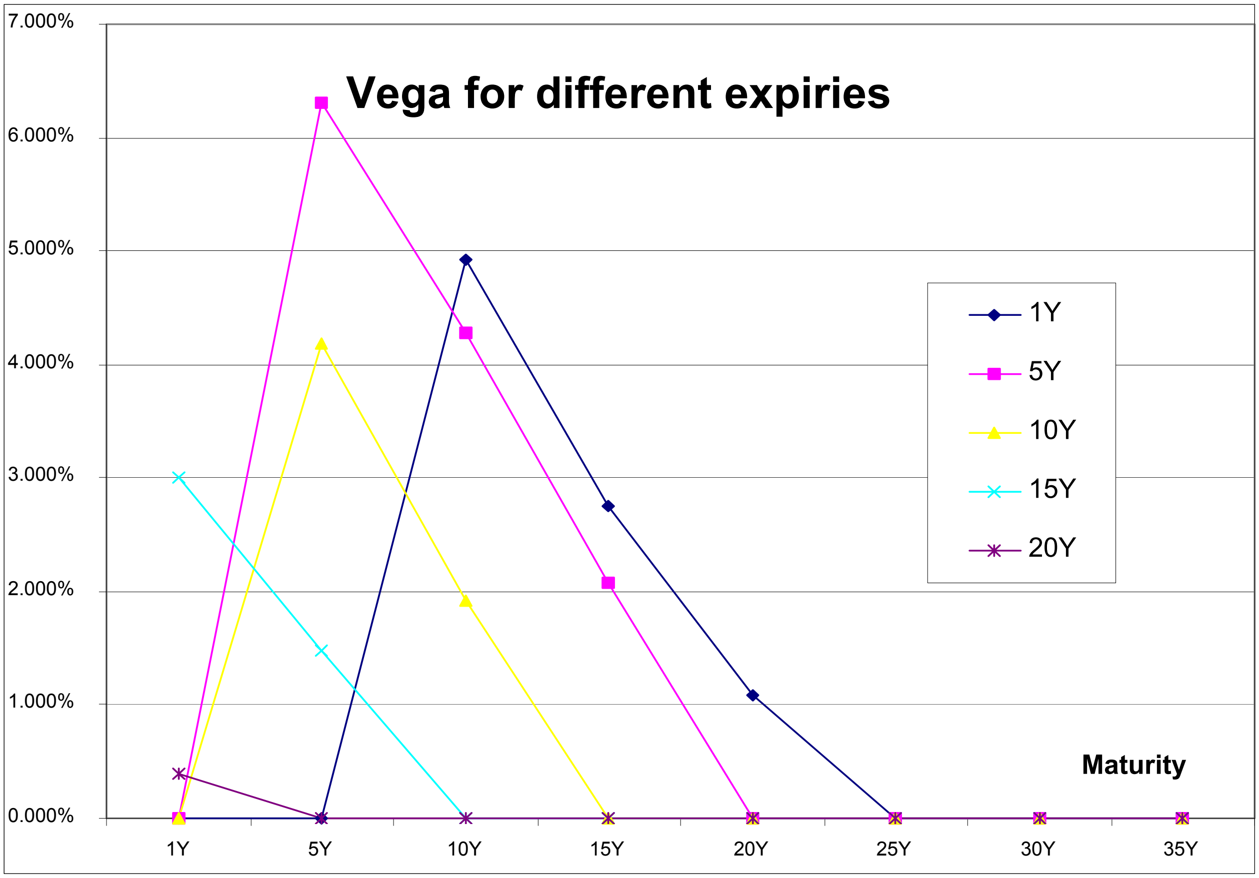


FIGURE 9. A vega grid for a callable Libor exotic, computed using the indirect method. Different expiries are represented by different lines, with the x-axis representing different maturities. Unscaled vega per unit notional.

BANK OF AMERICA, 233 SOUTH WACKER DRIVE, SUITE 2800, CHICAGO, IL 60611
E-mail address: vladimir.piterbarg@bankofamerica.com