HEC Montreal
Affiliated to the University of Montreal

Term-Structure Models and The Libor Market Model: Calibration to the Canadian market

By

Patrick Drouot

Management Sciences

This thesis has been submitted for the Masters degree in Finance (M. Sc.)

March 8, 2006

HEC Montreal
Affiliated to the University of Montreal

Term-Structure Models and The Libor Market Model: Calibration to the Canadian market

By

Patrick Drouot

Management Sciences

This thesis has been submitted for the Masters degree in Finance (M. Sc.)

September 2, 2005

DÉCLARATION DE L'ÉTUDIANTE, DE L'ÉTUDIANT

ÉTHIQUE EN RECHERCHE AUPRÈS DES ÊTRES HUMAINS

Recherche sans collecte directe d'informations

Cette recherche n'impliquait pas une collecte directe d'informations auprès de personnes (exemples : entrevues, questionnaires, appels téléphoniques, groupes de discussion, tests, observations participantes, communications écrites ou électroniques, etc.).

Cette recherche n'impliquait pas une consultation de documents, de dossiers ou de banques de données existants qui ne font pas partie du domaine public et qui contiennent des informations sur des personnes.

Titre de la recherche : Term-Structure models and the Libor Market Model: Calibration to the Canadian market.

Nom de l'étudiante, de l'étudiant : DROUOT PATRICK

Signature :

Date : 8-03-2006

Abstract

The purpose of this study is to provide an overview of term structure models and to calibrate the Brace, Gatarek and Musiela (BGM - 1997) model to the Canadian cap and swaption market. BGM is also known as the Libor Market Model (LMM). In order to calibrate the model we use the current term structure of interest rates, cap and swaption volatilities as quoted by the market using Black's (1976) formula. We make modeling choices on the instantaneous volatilities of forward rates, the correlation structure among forward rates and an approximation formula for the calibration of swaption volatilities. We look at the cap and swaption structures as well as the correlation structures between forward rates. We also identify the factors which influenced the Canadian market in the first half of 2005. Finally, we look at the fit of the LMM to the swaption volatilities.

Résumé

Le but de cette étude est de fournir un survol des différents modèles de structure à terme des taux d'intérêts pour ensuite appliquer le modèle de Brace, Gatarek et Musiela (BGM - 1997) au marché Canadien des caps et des swaptions. BGM est aussi connu comme étant le Libor Market Model (LMM). Afin de calibrer le modèle, nous utilisons la structure des taux d'intérêts, les volatilités cap et swaptions quottées sur le marché à l'aide de la formule de Black (1976). Nous faisons des choix de modélisations en ce qui concerne la structure de volatilité instantanées des taux forward, la structure de corrélation entre les taux forward et une formule d'approximation pour calibrer les volatilités swaption. Nous regardons les structures cap et swaptions ainsi que les structures de corrélation entre les taux forward. Nous identifions aussi les facteurs qui influencent le marché Canadien au premier semestre 2005. Finalement nous observons l'écart d'approximation entre le LMM et les volatilités swaption.

Table of contents

Tables

Figures

Acknowledgments

I would like to thank my family for their encouragements, Nicolas Papageorgiou my thesis director for his ongoing support on a theoretical and financial basis, and Simon Lalancette for his teachings on term structure models.

I would also like to thank Pierre Gagnon from Desjardins Securities in Montreal who has provided me with some good insights on the practical aspects of term structure models.

Abbreviations and Notation

- AMs = Arbitrage Models;
- BDT = Black Derman Toy;
- BGM = Brace, Gatarek, Musiela model;
- CIR = Cox-Ingersoll-Ross model;
- EMM = Equivalent Martingale Measure (Q);
- HJM = Heath-Jarrow-Morton model;
- K = Strike;
- LEH = Local Expectations Hypothesis;
- LMM = Libor Market Model (also known as the BGM model);
- O-U: Ornstein-Uhlenbeck process;
- PCA = Principal Component Analysis;
- SDE = Stochastic differential equation;
- P(t,T) : Zero coupon bond price at time t for the maturity T;
- r(t), r_t: Instantaneous spot interest rate at time t;
- f(t,T): the continuously compounded forward rate observed at time t for an instantaneous transaction starting at time T;
- F(t;T): Simply compounded forward (Libor) rate at time t for expiry-maturity pair T, T_{+1};
- k : Mean reversion coefficient;
- μ : drift term for the short rate, forward rate or any quantity that we choose to model;
- μ_t : Time varying drift;
- λ : Market price of risk;
- ϕ : Parameter in CIR and Longstaff and Schwartz zero coupon bond price formula;
- Φ, ψ, θ : Parameters affecting the LMM calibration through the instantaneous volatility specification found in formulation 7 of Brigo Mercurio (2001).
- $\eta(t,T)$: Deterministic function in the nearly proportional volatility in the HJM setting;
- y : Arbitrary but deterministic function in Markovian HJM;
- m(t): Index for the next reset at time t. It is the smallest integer such that $t \leq Tm(t)$;
- $T_1, T_2, \ldots, T_{i-1}, T_i, \ldots$: An increasing set of maturities;
- τ_i : The year fraction between T_{i-1} and T_i (tenor);
- S: fixed rate in a fixed leg of a swap;
- Q_0 : Physical/Objective/Real-World measure;
- Q : Risk neutral measure, equivalent martingale measure, risk-adjusted measure;
- Q^i : T_i-forward adjusted measure;
- W_t : Brownian motions under the Physical/Objective/Real-World Measure;
- ${W^Q}_t$: Brownian motions under the Risk Neutral Measure;

- W_t^i: Brownian motions under the T_i-forward adjusted measure;
- ': Transposition;
- $\rho^i(X,Y)$: correlation between X and Y under the Ti forward adjusted measure Q^i; i can be omitted if clear from the context or under the risk-adjusted measure;
- $\sigma_i(t)$: Volatility of the ith forward rate $F(t,T_i)$.
- $\sigma_i^X(t,T_i)$: Volatility of X under the T_i-forward adjusted measure Q^i: I can be ommited if clear from the context or under the risk neutral measure;
- ~: distributed as;
- **Cpl**(t,T,S,τ_0,N,σ) : Price at time t of a caplet resetting at time T and paying at time S at a fixed strike-rate σ; As usual τ_0 is the year fraction between T and S and can be omitted, and N is the nominal amount and can be omitted;
- **Fll**(t,T,S,τ_0,N,y) : Price at time t of a floorlet resetting at time T and paying at time S at a fixed rate y; As usual τ_0 is the year fraction between T and S and can be omitted, and N is the nominal amount and can be omitted;
- **Cap**(t, τ , τ, N, y): Price at time t of a cap first resetting at time T_1 and paying at times $T_2,\ldots,T_n$ at a fixed rate y; As usual τ_i is the year fraction between T_{i-1} and T_i *and* can be omitted, and N is the nominal amount and can be omitted;
- **Flr**(t, τ , τ, N, y): Price at time t of a floor first resetting at time T_1 and paying at times $T_2,\ldots,T_n$ at a fixed rate y; As usual τ_i is the year fraction between T_{i-1} and T_i *and* can be omitted, and N is the nominal amount and can be omitted;
- **FPS**(t, τ , τ, N,R): Price at time t of a payer forward-start interest rate swap with first reset date T_1 and payment dates $T_2,\ldots,T_n$ at the fixed rate R; As usual τ_i is the year fraction between T_{i-1} and T_i *and* can be omitted, and N is the nominal amount and can be omitted;
- **RFS**(t, τ , τ, N,R): Same as above but for a receiver swap;
- **PS**(t, τ , τ, N,R): Price of a payer swaption maturing at time T, which gives its holder the right to enter at time T an interest rate swap with first reset date T_1 and payment dates $T_2,\ldots,T_n$ (with $T_1 \geq T$) at the fixed strike-rate R; As usual τ_i is the year fraction between T_{i-1} and T_i *and* can be omitted, and N is the nominal amount and can be omitted;
- **RS**(t, τ , τ, N,R): Same as above but for a receiver swaption;

Introduction

Le marché des produits dérivés sur titres à revenus fixes est de loin le plus important au monde et la demande est croissante pour les nouvelles émissions de dettes et dérivés en tout genre. Si les titres à revenus fixes peuvent varier sensiblement d'un produit à l'autre, leur prix est toujours fixé en partant d'une bonne estimation de la structure à terme de taux d'intérêts sous-jacent. Il est donc important de disposer d'un modèle qui capture la dynamique de la structure à terme. De plus, le modèle doit répliquer la structure à terme actuelle. Le développement d'un modèle robuste et flexible qui peut être directement calibré aux données de marché, est un problème auquel les académiciens et les praticiens se sont confrontés durant de nombreuses années.

Les premiers modèles étaient difficilement calibrables aux données de marchés. Les modèles d'équilibres qui tentent de modéliser le processus de taux court et qui utilisent la prime de risque telle que cotée par le marché, ne prennent pas la structure à terme présente des taux comme intrants et donc ne peuvent être utilisés pour établir le prix des produits dérivés sur titres à revenus fixes. La génération suivante de modèles appelée « modèles de marchés », ne se consacre plus à capturer la dynamique du taux court mais à répliquer la structure à terme des taux actuels. Les modèles proposés par Black Derman Toy (1990), Hull & White (1999) et Black Karasinsky (1991) prennent la structure actuelle des taux comme intrants mais les paramètres sont difficilement calibrables aux caps et aux swaptions. Une contribution importante est faite par Heath, Jarrow et Morton (HJM, 1992), qui proposent un modèle basé sur le taux forward instantané. Même si ce modèle est théoriquement robuste, il est difficile à implémenter car les taux instantanés sont inobservables. Toutefois, il prépare les bases pour le Libor Market Model qui est la génération de modèle le plus récente.

Brace, Gatarek et Musiela (BGM, 1997) proposent un environnement compatible avec HJM connut sous le nom de Libor Market Model (LMM) qui prend comme intrant les taux forward discrets ou swap. Ceci est pratique dés lors que l'on peut travailler en utilisant les volatilités caps et swaptions de façon cohérente avec le modèle de Black (1976). Ceci est important car les traders utilisent le modèle de Black (1976) en pratique mais il n'ya jamais eu de justification au niveau théorique. Le LMM établit un lien formel entre les modèles de structure à terme et leur implémentation au niveau pratique.

Dans cette étude nous montrons comment calibrer le LMM aux volatilités caps et swaptions au marché canadien. Nous montrons la justesse du modèle et identifions les facteurs qui influencent le plus le marché des titres à revenus fixes canadiens au cours des six premiers mois de 2005.

Nous procédons an analysant les modèles de taux courts en distinguant les modèles d'équilibre des modèles d'arbitrages. Nous présentons ensuite le modèle de HJM (1992). Troisièmement, nous regardons le modèle de Black (1976) pour établir le prix des caps et des swaptions. En quatrième partie, nous présentons les hypothèses derrière le LMM et la façon dont elles lient la théorie et la pratique. Ensuite, nous précisons les données utilisées pour notre étude. Sixièmement, nous calibrons le LMM aux caps et swaptions en faisant certaines hypothèses sur les structures de volatilités et de corrélations des taux forwards. Nous présentons ensuite des conclusions et certains thèmes restant à développer.

Introduction

The market for fixed income securities is by far the largest of the capital markets, and the demand for debt issues, collateralized bonds, and fixed-income derivatives is still growing rapidly. Although fixed-income securities vary greatly in structure and risk, the pricing of all these securities relies on the proper estimation of the underlying term structure of interest rates. Therefore, it is imperative that we have at our disposal a model that can capture the dynamics of this term structure. Furthermore, for the pricing of interest rate derivatives it is important that the model be capable of replicating the present term-structure. The development of a robust and flexible model that can be readily calibrated to market data is a problem that has confronted academics and practitioners for many years.

Early interest rate models were difficult to calibrate to market data. Equilibrium models, which attempt to model the short rate process and make use of the market price of risk, do not take the current yield curve as an input, and therefore cannot be used to price interest rate derivatives. Subsequent models, often referred to as "arbitrage models", no longer focused on capturing the dynamics of the short rate, but rather on replicating the present term-structure. Models proposed by Black Derman Toy (1990), Hull and White (1999) and Black Karasinsky (1991) take the current yield curve as an input, however the parameters are difficult to calibrate to the price of caps and swaptions. An important contribution was made by Heath Jarrow Morton (HJM, 1992), who propose a model based on the instantaneous forward rate. Although theoretically robust, the model is difficult to implement as instantaneous forward rates are unobservable. It does provide the basis however, for the most recent development in interest rate modelling, the Libor Market Model.

Brace, Gatarek and Musiela (BGM, 1997) propose a compatible framework with HJM also known as the Libor Market Model (LMM) which take market observable data such as discrete libor forward rates or swap rates as inputs. This is an interesting class of models because it can be built around cap and swaption volatilities and is consistent with the Black (1976) model. This is an important feature because in practice, traders price derivatives using the Black (1976) model but there has never been a theoretically sound justification for doing so. The Libor market model provides a formal link between theoretical term-structure modelling and practical implementation.

In this study we will show how to calibrate the LMM to cap and swaption volatilities for the Canadian market. We will show the goodness of fit of the model and outline the driving factors of the Canadian fixed income market in the first half of 2005.

We proceed by reviewing short rate based models which are separated in two categories, the equilibrium and the arbitrage models. We then discuss the HJM (1992) framework. Thirdly we show the Black (1976) model to price cap and swaptions. In a fourth part, we show the hypothesis behind the LMM and how they bridge the gap between theory and practice. Than, we show the data we use in our study. In a sixth part, we calibrate the LMM to caps and swaptions by making a number of assumptions on the volatility and correlation structures of forward rates. We will then offer concluding remarks and address certain remaining issues.

1. Short rate based models: Equilibrium and arbitrage models

There are two classes of short rate based models, equilibrium and arbitrage models (AMs). They both offer insight into the dynamics of the term structure but differ in that equilibrium models do not take the current term structure as an input and arbitrage models do.

1.1. One factor equilibrium models

One factor equilibrium models are easy to implement which explains their popularity. Between 80% and 90% of the variance of the dynamics of the interest rate term structure can be explained by the first factor, which is considered to be the level of the interest rate (Rebonato 1998) (see appendix 1 "*notes on the short rate*"). Unfortunately, these models don't fit the current yield curve exactly which limits their effectiveness when we want to price fixed income derivatives.

Lets now review two of the most popular one-factor short rate models: the Vasicek (1977) and the CIR (1985) models.

1.1.1. Vasicek (1977)

The Vasicek (1977) model assumes that the instantaneous spot rate under the real world measure (Q_0) evolves according to an Ornstein-Uhlenbeck (O-U) process (the coefficients are constant in time).

$$dr(t)=\kappa\left[\mu - r(t)\right]dt + \sigma dW_t, \qquad \text{with } r(0)=r_0 \qquad (1)$$

r_0 is the instantaneous short rate, κ is the coefficient that measures the speed of mean reversion, μ is the long run mean to which r(t) is reverting and σ is the instantaneous volatility of the short rate. Note that all these parameters are positive constants. W_t is a Brownian motion under the physical measure (real world).

If the market price of risk (λ) is a constant, the risk neutral process for the short rate is

$$dr(t) = \kappa\left[\bar{\mu} - r(t)\right]dt + \sigma dW_t^Q \qquad (2)$$

Where $\bar{\mu} = \mu - \frac{(\lambda\sigma)}{\kappa}$ and W_t^Q is a Brownian motion under the risk neutral measure.

Hence we can state that the short rate r(t) follows an O-U process with constant coefficients under the risk neutral measure, Q, as well. Therefore, the process for the short rates is the same under Q_0 and Q except for a change in the long run mean.

The price of a pure discount bond can be derived and be shown to be[1]:

$$P(t,T) = A(\tau)\exp^{-B(\tau)r(t)} \qquad (3)$$

With $A(t,T) = \exp\left\{\left(\mu - \frac{\sigma^2}{2\kappa^2}\right)\left[B(t,T) - T + t\right] - \frac{\sigma^2}{4\kappa}B(t,T)^2\right\}$

$$B(t,T) = \frac{1}{\kappa}\left[1 - e^{-\kappa(T-t)}\right]$$

Jamshidian (1989) proposes a formula for evaluation of European options on zero coupon bonds, coherent with Vasicek (1977).

Let $t<T_{i-1}<T_i$ and the underlying security $P(t, T_i)$, then:

[1] For ease of notation $\tau = T - t$.

$$\text{Call}(t, T_{i-1}) = P(t, T_i)N(d_1) - KP(t, T_{i-1})N(d_2) \tag{4}$$

$$\text{Put}(t, T_{i-1}) = KP(t, T_{i-1})N(-d_2) - P(t, T_i)N(-d_1) \tag{5}$$

With $d_1 = \dfrac{\ln\left[P(t,T_{i+1})/KP(t,T_i)\right]}{\sigma_p} + \dfrac{\sigma_p}{2}$

$d_2 = d_1 - \sigma_p$

$$\sigma_p = \frac{\sqrt{\dfrac{\sigma^2(1-e^{-2\kappa(T_{i-1}-t)})}{2\kappa}}}{\kappa}\left(1-e^{-\kappa(T_i-T_{i-1})}\right)$$

The Vasicek model is the first to incorporate mean reversion in the interest rate dynamic so it is quite popular but can generate negative interest rates (due to the Normal distribution assumption of the O-U process). Additionally, the variance of interest rates is independent of the level.

Lets now review another popular short rate model, the Cox-Ingersoll-Ross (1985) model.

1.1.2. Cox Ingersoll Ross (1985)

The Cox-Ingersoll-Ross (CIR, 1985) model incorporates a square root term in the diffusion coefficient of the short rate.

Under Q_0:

$$dr(t) = \kappa\left[\mu - r(t)\right]dt + \sigma\sqrt{r(t)}dW_t, \text{ with } r(0) = r_0 \tag{6}$$

This short rate dynamic has the same mean reverting features as in the Vasicek (1977) model but additionally displays an instantaneous variance that depends of the interest rate level.

For r(t)=0, $2\kappa\mu > \sigma^2$ is the boundary that permits the short rate to always be positive.

Unlike previously, the market price of risk depends of the short rate but, under Q, it displays the same square root process structure.

The market price of risk depends of the short rate in the following sense:

$$\lambda(r) = \lambda_0 \sqrt{\frac{r}{\sigma}}$$

And under the measure Q, the short rate process becomes:

$$dr(t)= \bar{\kappa}\left[\bar{\mu} - r(t)\right]dt + \sigma\sqrt{r(t)}dW_t^Q \quad (7)$$

Moving from the real probability measure to the risk neutral one makes the drift be modified in the same way as in the Vasicek case (2) and for the same reasons. The goal is to preserve the same structure under both measures. In the Vasicek case, the change of measure is to maintain a linear dynamics and in the CIR case it is to maintain a square root-process structure.

Much like for the Vasicek model, the zero-coupon bond price is exponentially linear in the short rate and takes the following form:

$$P(t,T)= e^{A(t)-B(t)r_t} \quad (8)$$

With $B(\tau) = \frac{2(e^{\phi_1(\tau)} - 1)}{(\phi_1 + \kappa + \lambda)(e^{\phi_1(\tau)} - 1) + 2\phi_1}$ $\quad A(\tau) = \ln\left(\frac{2\phi_1 e^{\frac{(\phi_1+\kappa+\lambda)(\tau)}{2}}}{(\phi_1 + \kappa + \lambda)(e^{\phi_1(\tau)} - 1) + 2\phi_1}\right)^{\phi_2}$

$$\phi_1 = \sqrt{(\kappa + \lambda)^2 + 2\mu^2}$$

$$\phi_2 = \frac{2\kappa\mu}{\mu^2}$$

Closed form formulas exist to price derivatives in the CIR framework but they are rarely used because of their complexity. Traders favour the Jamshidian (1989) formulas to price interest derivatives (in the Vasicek framework).

We now present multifactor equilibrium models.

1.2. Multifactor equilibrium models

Single factor models describe the evolution of the interest rate term structure in a simple way. Adding subsequent factors permit to explain more sophisticated yield curve movements. These factors can be represented by macroeconomic shocks or linked to the level, slope and curvature of the yield curve. This is why we now provide an overview of multifactor equilibrium models.

The Brennan-Schwartz model (1979) is based on the dynamics of two yields on the curve: the short rate and the console yield. The difference between the short rate and the console yield provides a proxy for the slope of the curve. Therefore, this model allows the account for level and slope effects of the term structure.

The Fong-Vasicek (1991) model is a Vasicek model with two factors where the O-U process for the short rate includes a stochastic variance that follows a square root process.

The Longstaff and Schwartz (1992) model is a two factor model that describes the dynamic of the short rate and its variance within the CIR general equilibrium framework. Just like the Vasicek and CIR models this is an affine class model.
The first factor is the short rate r(t) and the second is the stochastic variance of the short rate (c.f. appendix 2). The calibration involves six parameters therefore it increases the level of fitting of the model to the yield curve. The problem is that the parameters that are obtained through calibration are unstable in a dynamic setting. Additionally, this model is very hard to calibrate to caps and swaptions

(which are the most liquid interest rate derivatives on the market today) and the dynamics of the yield curve are dubious.
This model is part of the stochastic variance model class that is not yet used by practitioners but should be in the years to come. The stochastic volatility models allow to model long term rate shocks.

We now provide remarks on equilibrium models.

1.3. Remarks on equilibrium models

What is common to all these models is that to specify the "real world" dynamics of bond prices, an econometric analysis of the statistical properties of the short rate is needed to determine the drift and volatility parameters. Additionally, we have to estimate the utility function of the bond investors in order to find the market price of risk (λ). This tells us everything about the real-world dynamics of the bond process and their price today. In practice though, the estimation of the market price of risk is tedious. Another way to tackle this issue is to use market data to estimate the market price of risk.
To overcome the lack of applicability of the theory, practitioners use the market data and imply these quantities. Unfortunately this is irrelevant because the model is incorrectly specified. Brown and Dybvig (1986) and Brown and Schaefer (1994) found the estimates to be non stationary over long period of time as well as on a day-to-day basis. This is a rather serious problem as the model implies that parameters are constant. An additional problem is that because of their stationary nature, neither the Vasicek nor the CIR models can recover for an arbitrary observed yield curve. Therefore, these models force the user to make a trade off. Relative value bond trading is possible but it isn't practical to price interest rate derivatives because we can't recover the prices of the underlying bonds.

We now review arbitrage models (AMs).

1.4. Arbitrage models

The models previously discussed are derived from equilibrium frameworks and are relevant for explaining the observed historical patterns in the dynamics of the term structure which help understand which factors move the economy.
On the other hand, this approach is inconsistent to price derivatives because empirically fitted models using past data won't guarantee that the model term structure matches the current term structure.

In practice, a derivatives trader needs to use the prevailing term structure and not the term structure derived from a model in order to adequately hedge his derivatives positions with underlying financial products.

In AMs, the goal is to match the current yield curve and one way to do this is to make the coefficients in a factor model vary deterministically with time.

AMs take the market prices of bonds (i.e. the current yield term structure) as inputs in order to price interest rate derivatives. It follows that these models won't find mispricing in the bonds (like the equilibrium models could) but will permit to price derivatives in the same fashion as the Black & Scholes (1973) framework for stock options.
To illustrate the discussion we will present the Ho-Lee (1985), the Hull-White (1990) and Black Derman Toy (1990) models.

1.4.1. Ho-Lee (1985)

In the Ho-Lee (1985) model, the short rate follows a random walk which allows the drift to be time varying. The short rate process still depends on the drift and volatility parameter. The volatility is constant and the yield curve is matched using the drift parameter. Ho-Lee (1985) give a specification of the drift which depends of the instantaneous forward rate in function of time and the volatility. This specification permits to find closed form formula for the price of European options on discount bonds in a manner very close to Black and Scholes (1973) for stock options. American style options can be evaluated through a binomial tree implementation.

The problem with this model is that it doesn't take into consideration the mean reversion process in the short rate and allows for negative outputs. Additionally, the volatility is flat for all rates, which is not realistic. This is what led to the Hull-White (1990) model.

1.4.2. Hull-White (1990)

To counter the problems of the Ho-Lee (1985) model, Hull and White (HW, 1990) extend the Vasicek (1977) model to match the initial term structure (c.f. appendix 3).

The Hull-White (1990) model is a short rate based model with a Gaussian distribution and the process is mean reversing. One version of HW's extended Vasicek model gives the short-rate dynamic as:

$$dr = [\mu_t - kr]dt + \sigma dW \tag{9}$$

Where, bond prices at time t are given by

$$P(t,T)=e^{A(t,T)-B(t,T)r_t} \tag{10}$$

and A(t,T) is related to the time varying drift, $\mu(t)$. Then, the calibration to the yield curve is performed through:

$$\mu(t) = \frac{\partial f(0,t)}{\partial t} + kf(0,t) + \frac{\sigma^2}{2k}(1-e^{-2kt}). \tag{11}$$

In this case, the volatility structure is richer than in Ho-Lee (1985) and for P(t,T) is:

$$\frac{\sigma}{k}\left[1-e^{-k(T-t)}\right]$$

The volatility of f (t,T) is $\sigma e^{-k(T-t)}$.

In this context, prices of European options on bonds are found and the implementation of the model is done through trinomial trees.

1.4.3. Black Derman Toy (1990)

The Black Derman Toy (1990, BDT) model is similar to Ho-Lee (1985) but precludes rates from being negative and allows for a mean reversion.
If we consider that r_t follows a Gaussian distribution:

$r_t = \mu_t \exp^{\sigma_t W_t}$

so that $\ln r_t = \ln \mu_t + \ln \sigma_t W_t$

Since $r_t = f(t, W_t)$, Ito's lemma is applied an we obtain:

$$d \ln r_t = \left[\frac{\partial \ln \mu_t}{\partial t} + \frac{\partial \ln \sigma_t}{\partial t} [\ln \mu - \ln r_t] \right] dt + \sigma_t dW_t \qquad (12)$$

Consider the volatility is time dependent in the following manner:

$\sigma_t = \sigma e^{-v(T-t)}$ where $\upsilon > 0$

In that case (11) becomes:

$$d \ln r_t = \left[\frac{\partial \ln \mu_t}{\partial t} + v [\ln \mu - \ln r_t] \right] dt + \sigma_t dW_t \qquad (13)$$

If $\upsilon = 0$ so that the volatility of the short rate is constant, the Ho-Lee specification is approximately obtained.

The shortcoming of the BDT model is that once the market prices of caplets are recovered, the resulting short rate volatility is usually time-decaying. This implies that future yield curves are less and less volatile.
Nevertheless, this has been a very popular model because it permits direct calibration to the term structure via the drift term and to caplets via the volatility parameter υ.

1.5. Remarks on equilibrium and arbitrage models

Vasicek (1977) and CIR (1985) do not take bond prices as given. They use the current term structure to deduce the risk premium in expected returns. On the other hand AMs take bond prices as given and the assumption is made that we don't need a risk premium.

Equilibrium models have the market price of risk (which is hard to obtain) as input and require statistical examination of past data. Because equilibrium models don't take the term structure of interest rates as given, they let us know which bonds are mispriced (according to the model) and therefore can be used to implement bond trading strategies.

On the other hand when it comes to trading interest rate derivatives we would need to check if the underlying bonds are correctly priced meaning these models are often useless to price derivatives. The advantage of this class of model is that it can be used overtime without having to reestimate parameters.

For AMs, we need the term structure of spot rates as inputs, which is not hard to obtain. Attention must be brought on any misquotes of spot rates (due to errors or liquidity) because otherwise errors will be built into the model. Bond trading is not the purpose of AMs because they assume that all bonds are correctly priced. The purpose of these models and Heath Jarrow Morton (HJM, 1992) in particular is to trade and hedge derivatives by telling us which price is too high or too low in comparison to the market observed prices. The problem with AMs versus equilibrium models is that AMs' use is inconsistent over time, i.e. we need daily recalibration.

We now present the HJM framework.

2. Heath Jarrow Morton (1992)

We will now review the Heath Jarrow Morton (HJM, 1992) framework. It is interesting because it is also based on the no arbitrage condition and can accommodate nearly all existing interest rate models distributed normally.

Because of this feature the martingale measure is implied. HJM is built in the same spirit as Ho-Lee[2] and HW because it takes the initial term structure as given and prices other interest rate derivatives according to the no-arbitrage condition.

HJM differs from previous methodologies because its stochastic structure is based on forward rates[3]. Secondly, unlike certain early short rate models, it doesn't require the inversion of the term structure to eliminate the market price of risk from contingent claim values[4]. Finally, it proposes a stochastic forward rate process with multiple stochastic factors influencing the term structure.

We now look at the setting of the HJM model.

---

[2] The difference is that Ho-Lee is a single factor model whereas HJM can admit many factors to drive interest rates. Additionally, unlike HJM impose the exogenous stochastic structure to forward rates (not on zero coupon bonds).

[3] That is because the volatility of zero coupon bond prices changes over time since prices are a fixed amount at maturity whereas constant forward rate volatilities are consistent with a fixed value for a zero coupon bond at maturity.

[4] Inversion is required due to the two step process utilized to price options. First, zero coupon bonds are priced which introduces dependency on market price of risk. Then options are valued. To remove the dependency on λ, the bond price formula is inverted after step one. This is difficult computationally because bond pricing formulas are non linear. Spot rate and bond price processes parameters are dependant of λ hence models can admit arbitrage opportunities.

2.1. The setting

The model is set in a continuous trading economy and uncertainty is characterized by the probability space (Ω, F, Q) (see appendix 5 for a discussion on the probabilistic setting).

There is absence of arbitrage opportunities and the markets are complete[5].

There are N-basis bonds and all bonds are priced as linear combinations of the basis bonds. The risk premium, λ, need not to be specified because the information is already contained in the market price of the N-basis bonds. In fact, the no arbitrage and complete markets conditions imply market prices of risk are uniquely determined by and contained in the market prices of bonds.

We now discuss term structure movement in the HJM setting.

2.2. Term structure movements

Discounting can be made by using forward rates, hence P(t,T) = $e^{-\int_t^T f(t,T)ds}$. We can also extract a forward rate by differentiating the above with respect to T:

$$f\,(\mathrm{t,T}) = -\frac{\partial \log P(t,T)}{\partial T} = \frac{\log P(t,T) - \log P(T,T_{+\tau})}{T_{+\tau} - T} \tag{14}$$

Let f(t,T) be the continuously compounded forward rate observed at time t for an instantaneous transaction starting at time T, and let r(t)= f (t,t) be the only spot rate we pay attention to.

HJM propose that f (0,T) changes as follows between 0 and T:

$$f(t,T) = f(0,T) + \int_0^t \mu(u,T)du + \sum_{i=1}^{n} \int_0^t \sigma_i(u,t)dW_i(u) \tag{15}$$

[5] All payoffs in the market can be replicated by a combination of traded securities. The market is complete if there is a unique Equivalent Martingale Measure (EMM). The risk neutral (martingale measure, Q) measure can be constructed from a set of N different discount bonds (see Rebonato 1998 chapter 14).

With the assumption that f (0,t) is a fixed, non random initial forward rate curve, that $\int_0^T \mu(t,T,w)dt < +\infty$ and that volatilities are measurable (i.e. $\int_0^T \sigma_i^2(t,T,w)dt < +\infty$).

This is an "n" factor model where the factors are captured by $\sigma_i(u,t)$ and $dW_i(u)$. $\sigma_i(u,t)$ is the volatility of factor "i" observed at time "u" for the forward rate at time T[6]. $dW_i(u)$ is a Weiner process representing the source of uncertainty of factor "i" at time "u".

The expression means that forward rate started off with a value of $f(0,T)$ and evolves over time to a value of $f(t,T)$. These changes in the forward rate reflect the accumulation of the infinitesimal changes that consist of the drift and volatility that occurred over the period 0 to T.

Then, HJM go on to derive their most important result, namely the "Forward Rate Drift Condition". The condition of no arbitrage implies that a martingale probability measure exists and implies a restriction on the drift coefficients of the forward rates.

$$\mu(t,T) = \sum_{i=1}^{n} \sigma_i(t,T) \left[\int_t^T \sigma_i(t,s)ds - \lambda_i(t) \right] \tag{16}$$

Where $\mu(t,T)$ and $\sigma_i(t,T)$, i=1,…,n, are the coefficient functions in the forward rate process under the objective measure Q_0. When the "Forward Rate Drift Condition" holds, the price of the risk function is unique, $\lambda_s(t) \equiv \lambda(t)$. To transform to the Equivalent Martingale Measure (EMM) Q, $\lambda_s(t)$=0.

Therefore in HJM, we don't have to separately model a price of risk (unless we are calibrating to time series data). The preferred calibration method is to current cross-sectional market prices, so the price of risk is marginalised.

[6] σ_i can depend on the entire past of the Brownian motions.

HJM is truly preference free because we price interest rate derivatives relative to the current yield curve (forward curve), and the yield curve reflects all relevant investor preferences.

As we can tell, the drift[7] of the forward rate is entirely determined by its volatility structure under both the physical measure and the risk neutral measure.

Proof: *Under the risk neutral probability measure Q,*

$$df(t,T) = \alpha^Q(t,T)dt + \sum_{i=1}^{N}\sigma_i(t,T)dW_i^Q(t), \textit{where}$$

$$\mu^Q(t,T) = \mu(t,T) + \sum_{i=1}^{N}\lambda_i(t)\sigma_i(t,T). \textit{Then}$$

$$\mu^Q(t,T) = \sum_{i=1}^{N}\sigma_i(t,T)\int_t^T \sigma_i(t,s)ds.$$

At time t, the drift of the forward rate to start at time T is obtained by integrating (adding) all the volatilities over the time periods from t to T and multiplying by the sum of the volatilities across all the factors observed at time t for the rate to start at time T.

In continuous time and under the measure Q, HJM describe the process for the instantaneous forward rate as:

$$df_t(T) = \mu(t,T,\omega)dt + \sigma(t,T,\omega)dW_t^Q, \qquad t \leq T, \qquad (17)$$

where ω is a sample point in the sample space Ω. The drift is restricted to the above.

Since the drift depends entirely on the volatility of the instantaneous forward rate, it is critical to choose the right volatility specification. We now show the different volatility functions available in HJM.

---

[7] The drift can't be zero under EMM. HJM (1994) and Ritchken (1996) use a binomial tree example to show that if a drift of 0 is assumed then there is an arbitrage opportunity.

2.3. Volatility functions in HJM

HJM depends on the specification of the volatility structure of the forward rates. The volatility structure is given according to two criteria. The first is the cross sectional volatility of factor *i* at a given time point for a set of different forward rates, $\sigma_i(0,1), \sigma_i(0,2), ..., \sigma_i(0,T)$. The second is the time series volatility i.e. the volatility of factor *i* associated with a given forward rate over time, $\sigma_i(0,T), \sigma_i(1,T), \sigma_i(2,T)$. If it is a multifactor version of HJM there is a factor volatility. If we have "n" factors in the model, there is a different set of time series and cross sectional volatilities for each factor *i*.

Note that there need not be any formal mathematical structure between these volatilities. In other words, they don't need to be related to each other.

We now present possible volatility specifications: namely the Ho-Lee, exponentially dampened and the nearly proportional volatilities.

2.3.1. Ho-Lee type volatility

We now show the Ho-Lee type volatility.

Lets first consider the case of a one factor (n=1) specification with constant volatility $\sigma_1(t,T) = \sigma$. The drift term can then be expressed as:

$$\mu(t,T) = \sigma_1(t,T)\left[\int_t^T \sigma_1(t,T)ds - \lambda_1(t)\right]$$

$$= \sigma_1^2(T-t) - \sigma\lambda_1(t) \quad (18)$$

Under the risk neutral measure, the instantaneous forward rate evolves as:

$$f(t,T) = f(0,T) + \sigma_t^2(T - \frac{t}{2}) + \sigma W^Q(t) \quad (19)$$

Where the bond price process becomes:

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt - \sigma(T-t)dW^{Q}(t)\,. \qquad (20)$$

r(t)=f(t,t)= the instantaneous short rate. Note that this is the Ho & Lee specification. It is tractable but unrealistic[8].

2.3.2. Exponentially dampened volatility

Here, we consider an exponentially dampened volatility, which is consistent with the Vasicek (1977) model. It is also called exponentially dampened volatility because it results in the volatility declining at an exponential rate.

The volatility is expressed as:

$$\sigma(t,T) = \sigma \exp^{-\varphi(T-t)}, \qquad \text{with } \sigma \text{ and } \varphi \text{ constant.} \qquad (21)$$

This structure is convenient because it permits many closed forms for options and other derivatives (see Jarrow Turnbull 2000). Although is performs better than the constant Ho-Lee type volatility, it is not capable of replicating the observed volatility term structure.
This leads to the nearly proportional volatility.

2.3.3. Nearly proportional volatility

The nearly proportional volatility is called that way because it sets the σ proportional to the current forward rate and bounds it on the upper end so that it won't get unreasonably high.

The volatility is expressed as:

$$\sigma(t,T) = \eta(t,T)\min(f(t,T);M) \qquad (22)$$

η (t,T) is a deterministic function and M a large constant.
We just reviewed a few common types of volatilities but there are other ones[9].

---

[8] Rejected by Flesaker (1993) on Eurodollar Futures and futures option data.

Beyond the simplistic approach of the Ho-Lee type volatility structure, the HJM model does not produce closed-form formulas for interest rate derivatives prices. We now address the limits of the HJM framework.

2.4. HJM Limits

With a general form for the volatility function $\sigma_i(t,T)$, the evolution of the bond price depends on the whole history of interest rates. This path dependency makes implementation difficult. The tree structure in numerical techniques is non-recombining (bushy) and Monte Carlo simulation techniques are generally inapt in dealing with American style options.

Hence, many common versions of HJM are non Markovian (path dependent) which increases computational complexity.

The solution to these issues is the Markovian HJM models, because they impose more structure on the volatility function (see Ritchken & Sankarasubramanian 1995 and Inui & Kijima 1998).

If we restrict the volatility function $\sigma_1(t,T)$ to solve:

$$\frac{\partial \sigma_i(s,t)}{\partial t} = -y_i(t)\sigma_i(s,t) \qquad (23)$$

y_i(t)=arbitrary but deterministic function

$$\gamma_i(t) = \int_0^t \sigma_i^2(s,t)dt \qquad (24)$$

Then the process for r(t) and γ(t) becomes jointly Markovian thus the tree structure is recombining.

---

[9] For more on volatility structures under HJM see Ritchken and Sankarasubramanian (1995).

All previous models are special cases of this specification. For example if we assume y_i(t)= y_i and $\sigma_1(s,s) = \sigma_i$, then this class of Markovian HJM models reduces to the HW model. When k=0, we rediscover the Ho-Lee model.
For practical purposes we now discuss HJM in a discreet setting.

2.5. Discretizing HJM

With the exception of the exponentially dampened volatility and a few other (see Brenner and Jarrow 1993), the HJM model does not produce closed-form solutions for the prices and risk measures of interest rate derivatives. Hence numerical methods are normally required.

We now discretize HJM in order to use it in a binomial tree setting. It will permit us to gain better understanding of the drift restriction. We choose to focus on the one factor version.

$$df(t,T) = \mu(t,T)dt + \sigma(t,T)dW_t^Q \qquad (25)$$

Where the arbitrage free drift restriction is:

$$\mu(t,T) = \sigma(t,T)\sum_t^T \sigma(t,T)du \qquad (26)$$

To generate binomial versions of the model we need prices of a number of bonds maturing at discrete time points 1, 2, 3… (We then have T forward rates from f(0,0) to f(0,T-1). Additionally we need volatilities for maturities 1,2,…,T-1. That is enough to build a binomial tree of T-1 time steps.

The stochastic process for forward rates in discrete form is:

$$\Delta f(t,T) = \mu(t,T)\Delta t + \sigma(t,T)\Delta W_t^Q \qquad (27)$$

Assume:

- Δ=1 as a discrete time step;

- We convert the Wiener process to a random variable with value of +1, -1 at each time step;
- Martingale probabilities of 0,5;
- Each time step has a defined length of one unit.

Hence:

$$\Delta f(t,T) = \mu(t,T) \pm \sigma(t,T) \tag{28}$$

So, at each given time, for a given forward rate, we move on one step ahead to the next time in the following way:

$$\begin{aligned} f(t+1,T)^{+} &= f(t,T) + \mu(t,T) + \sigma(t,T) \\ f(t+1,T)^{-} &= f(t,T) + \mu(t,T) - \sigma(t,T) \end{aligned} \tag{29}$$

We also remind that to prevent arbitrage, the local expectations hypothesis must hold[10].

From now on probabilities and expectations are always under the EMM, Q.

$$\frac{P(t,T)}{P(t,t+i)} = E^{Q}\left[P(t+i,T)\right] \tag{30}$$

Using P(t,T)=e $^{-\int_t^T f(t,u)du}$

$$\exp\left[-\int_t^T f(t,u)du\right]\exp\left[\int_t^{t+h} f(t,u)du\right] = \exp\left[-\int_{t+h}^T f(t,u)du\right] \tag{31}$$

This must be equal to $E^{Q}\left[P(t+i,T)\right]$ which can be found by evaluating:

[10] Local expectations hypothesis says that the expected return on any instrument over the shortest period of time is the Risk Free Rate. The expectation that is taken is using the martingale probability measure that is:

$$P(t,T) = P(t,t+i)E^{Q}\left[P(t+i,T)\right]$$

$$\frac{1}{2}\exp\left[-\int_{t+h}^{T}(f(t,u)+\mu(t,u)+\sigma(t,u))du\right]+\frac{1}{2}\exp\left[-\int_{t+h}^{T}(f(t,u)+\mu(t,u)-\sigma(t,u))du\right] \quad (32)$$

These two terms are the next possible bond prices which are obtained by discounting at the sequence of forward rates over the remaining lives of the bonds (taking into account the binomial probabilities).

After some additional derivations (see HJM 1992) we obtain,

$$\mu(t,T)=\sigma(t,T)\int_{t}^{T}\sigma(t,u)du\,, \quad (33)$$

which is the result obtained before.

In order to work with HJM in discrete time, we need to obtain a discretized version of the drift restriction. HJM provide one but it is incorrect as shown by Grand and Vora (1996,1999).

Grant and Vora (1999) go on to derive the correct formula and start at:

$$\text{P(t,T)}=E^{Q}\left[P(t+i,T)\right]P(t,t+i) \quad (34)$$

They then make use of the fact that a Weiner process follows a normal distribution and that the correlations between all forward rates in a one factor model is perfect. After some algebra[11] they find:

$$\mu(t,T)=\frac{1}{2}\left[\sigma^{2}(t,T)+2\sigma(t,T)\sum_{j=t+1}^{T-1}\sigma(t,j)\right] \quad (35)$$

This formula is a bit confusing because when t=0 and T=1 we are summing from j=1 to 0. This whole term drops out in fact. To avoid confusion there is an alternative equivalent formula:

$$\mu(t,T)=\sigma(t,T)\sum_{j=t+1}^{T}\sigma(t,j)-\frac{\sigma^{2}(t,T)}{2} \quad (36)$$

[11] For proof see Grant & Vora in Journal of Fixed Income (March 1999).

Grant and Vora (1999) call it the Drift Adjustment Term but in fact it is simply the drift.

2.6. Additional Problems with HJM

In the HJM framework, interest rates can be negative with positive probability because Gaussian models are not excluded. Additionally, calibrating parameters to fit the current term structure can result in negative rates (In CIR $2k\mu_t > \sigma^2$ permits r(t)>0. This restriction is not guaranteed when μ_t (the drift) has to match current bond prices).

The negative interest rate positive probability issue as well as the non recombining tree one can also be addressed in certain versions of the model[12]
We now conclude our discussion on HJM.

2.7. Conclusion

We have presented the HJM model in a one factor environment, which captures changes in the level of interest rates but not in the slope of the curvature of the term structure. Multifactor models are required for that effect. They are complex, for example a tree version of a two factor HJM model requires a trinomial tree and the number of paths for T time steps is 3^T.

Tradeoffs are therefore required before deciding to go to models of more than one factor.

The solution to these problems was found in the Brace Gatarek Musiela model (BGM, 1997) which will be discussed subsequently. But before, lets look at the different market instruments for which term structure models are typically used.

---

[12] See Munnik 1994 for recombining tree, negative rates can be adressed by volatility specifications that dampen the volatility sufficiently to prevent it from moving further downward when at or near zero.

3. Market instruments

In the previous section, we saw the HJM model and in the next section we will review the Brace, Gatarek and Musiela (1997) model, also known as the Libor Market Model (LMM), before implementing it to Canadian market data. However, given the fact that the market prices caps and swaptions using Black (1976) and that the BGM model attempts to reconcile theory and practice, we first review how these products are valued.

3.1. Caps and caplets

Caps are market instruments which protect from an increase in rates over a specific time period. It is like a series of calls on interest rates.
A cap payoff is equal to:

$$\text{PayoffT(0)}= N\sum_{i=1}^{T-1} P(0,T_{i+1})\tau_{(i)}(F(t,T_i)-K)_+ \qquad (37)$$

Where N is the notional and T is the time where the cap ends. $F(t,Ti)$ is the forward rate seen from time t starting in T_i and ending in T_{i+1}, and linear over τ_i. It is therefore function of the present term structure.

A cap is the net present value of a portfolio of caplets. Each caplet is like a call on a forward rate starting at different periods in the future. If we hypothesise that all forward rates follow a lognormal distribution then we can use the Black model which permits to find a price given the volatility and vice versa. This is precisely what the market does.

The market uses Black (1976) to price caps in the following manner:

$$\mathbf{Cap}\left(0,\tau,N,K,\sigma_{\alpha,\beta}\right)=N\sum_{i=\alpha+1}^{\beta}P(0,T_i)\tau_i Bl(K,F(t,T_i),\upsilon_i,1) \quad (38)$$

Where

$$Bl(K,F,T,\upsilon)=FN(d_1)-KN(d_2)$$

and

$$d_1(K,F,\upsilon)=\frac{\ln(F/K)+\upsilon^2/2}{\upsilon}$$

$$d_2(K,F,\upsilon)=\frac{\ln(F/K)-\upsilon^2/2}{\upsilon}$$

$$\upsilon_i=\sigma_{\alpha,\beta}\sqrt{T_{i-1}}$$

Note that all market participants agree that caplets and caps should be quoted using volatilities.

Now, because within the BGM framework we use caplet volatilities we show how to price caplets.

To price caplets A caplet is a call on a forward rate at a specific period in the future, for a specific time period. It protects against unwanted rate variations. It is an option on the i-th forward rate. It is therefore necessary to take the $P(t,T_{i+1})$ numeraire. The T_{i+1}-forward neutral measure is then used to price the caplet because all the assets in the market using the numeraire $P(t,T_{i+1})$, are martingales under this measure:

$$\frac{f(0,T_i)}{P(0,T_{i+1})}=E_{i+1}\left(\frac{F(t,T_i)}{P(t,T_{i+1})}\right) \quad (39)$$

So for the caplet:

$$\mathbf{Cpl}(0,T_i)=P(0,T_{i+1})E_{i+1}(\mathbf{Cpl}(t,T_i))$$
$$=P(0,T_{i+1})N\tau_i\,E_{i+1}((F(t,T_i)-K)+)$$

$$= P(0,T_{i+1}) N \tau_i \left(E_{i+1}(F(t,T_i)\, n(d_1) - K n(d_2)\right) \quad (40)$$

$$d1 = \frac{\ln\left(\frac{E_{i+1}F(t,T_i)}{k}\right) + \frac{\sigma^2_{caplet,i}T_i}{2}}{\sigma_i\sqrt{T_i}}$$

$$d2 = d1 - \sigma_{caplet,i}\sqrt{T_i}$$

Where $\sigma^2_{caplet,i}T_i = \int_0^{T_i} \sigma_i^2\, dt$ and $E_{i+1}(F(t,T_i)) = F(0,T_i)$

The market quotes cap volatilities but not caplet volatilities. We now show how to recover caplet volatilities from cap volatilities.

3.2. Stripping caps into caplets[13]

One of the inputs of the model is the caplet volatilities which can be stripped from the cap volatilities. As mentioned, caps are portfolios of caplets. For example a one year cap is comprised of three three-month caplets or one six month caplet depending on the tenor structure. On the other hand a two-year cap has eight three-month caplets or four six-month caplets.
The convention we will use is a six-month tenor structure.

Using the fact that the cap volatility is the equivalent of a vega weighted sum of each caplet volatility:

$$\sigma_{i,cap} = \frac{\sum_{i=\alpha+1}^{\beta} \upsilon_i \sigma_{i,caplet}}{\sum_{i=\alpha+1}^{\beta} \upsilon_i} \quad (41)$$

[13] Source: Thomas Weber (2005).

We will use the following algorithm in order to strip the caplets volatilities:

1. Set $\sigma_{1,caplet} = \sigma_{1,cap}$

2. Solve for $\sigma_{2,caplet}$ from $\dfrac{\upsilon_1 \sigma_{1,caplet} + \upsilon_2 \sigma_{2,caplet}}{\upsilon_1 + \upsilon_2} = \sigma_{2,cap}$

3. And so forth.

After discussing caps and caplets, we now present the Black (1976) formulas which permit to price swaptions.

3.3. Swaptions

Swaptions are derived indirectly from swaps. We discuss the swap rate before presenting swaptions.

Market participants use swaps to prevent from adverse interest rate variations. A company entering a swap agreement will pay a fixed rate a specified time periods for a certain notional amount. In exchange for that it receives the libor rate at those periods. This is a payer swap. Should the company wish to pay libor and receive a fixed rate it is then a receiver swap.
We don't need a model to price swaps because the price of an at-the-money swap is simply zero. So for the fixed leg:

$$\text{Pv(fixed leg)} = \sum_{i=\alpha+1}^{\beta} C_i P(t, T_i) \tag{42}$$

Where C_i is the coupon.

The floating leg is equal to:

$$\text{Pv(floating leg)} = N \sum_{i=\alpha+1}^{\beta} \left[P(t, T_{i-1}) - P(t, T_i) \right] \tag{43}$$

For simplicity it is assumed that tenors are the same for the floating and fixed legs.

The swap rate is the rate which makes both legs equal to one another.

$$S_{\alpha,\beta}(t) = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau_i P(t,T_i)} \tag{44}$$

A swaption is an option on the swap rate. So a swap in two years for five years gives the right but not the obligation to enter into a five year swap in t+2 years, where t is today`s date. This swap would be noted a 2x5. In our notation this would be noted a $\text{Swaption}_{4,10}(t)$ if we consider the tenor structure to be six month intervals.

Black (1976) provided with the following formula to price swaptions:

$$\mathbf{PS}(0,\tau,\mathrm{N},\mathrm{K},\sigma_{\alpha,\beta}) = \mathrm{N}\, Bl\left(K, S_{\alpha,\beta}(0), \sigma_{\alpha,\beta}\sqrt{T_\alpha}, 1\right)\sum_{i=\alpha,\beta}^{\beta} \tau_i P(0,T_i) \tag{45}$$

$$Bl\left(K, S_{\alpha,\beta}(0), \sigma_{\alpha,\beta}\sqrt{T_\alpha}, 1\right)\sum_{i=\alpha,\beta}^{\beta} \tau_i P(0,T_i) = FPS(0,T_\alpha,T_\beta)N(d_1) - KN(d_2)$$

where

$$d_1 = \frac{\ln\left(\frac{FPS(0,T_\alpha,T_\beta)N(d_1) - KN(d_2)}{K} + \frac{\sigma_{\alpha,\beta}{}^2 T}{2}\right)}{\sigma_{\alpha,\beta}\sqrt{T}}$$

$$d_2 = d_1 - \sigma_{\alpha,\beta}\sqrt{T}$$

Note that $\sigma_{\alpha,\beta}$ is a volatility parameter quoted in the market and FPS is a forward start swap underlying the swaption. The formula for a FPS is as follows:

$$FPS(\alpha,\beta) = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau_i P(t,T_i)} \tag{46}$$

A similar formula exists for the receiver swaption, which gives the holder the right to enter at time T_α a receiver swap, with payment dates between α and β. The formula is the same but

$$Bl\left(K, S_{\alpha,\beta}(0), \sigma_{\alpha,\beta}\sqrt{T_\alpha}, 1\right)\sum_{i=\alpha,\beta}^{\beta}\tau_i P(0,T_i) = KN(-d_2) - FPS(0,T_\alpha,T_\beta)N(-d_1) \qquad (47)$$

The market lists prices of swaptions using the volatility found in Black's formula $\sigma_{\alpha,\beta}$.

The problem with the Black environment is that we can't price caps and swaptions simultaneously while preserving a lognormal distribution for both the swap and the forward rates. But the fact that the market uses the same formula both for the forward and the swap rate does not lead to arbitrage opportunities because the discrepancy is too small and it is simply due to a difference in the conventions used to quote forward and swap rates.

Now that we have provided a review of the Black (1976) cap and swaption pricing formulas, we can present the BGM model which bridges the gap between market practice and financial theory.

4. The Brace, Gatarek, Musiela Model (1997)

The BGM model, also called the Libor Market Model (LMM), is an extension and a generalization of HJM. In HJM we model the instantaneous forward rates whereas in BGM we model the discrete, observable, forward and swap rates.

We now review the characteristics of the BGM model.

4.1. Forward rates are lognormal

The first hypothesis of the model is that forward rates are lognormal. The second assumption is the absence of arbitrage opportunities. Under a single measure, when there is absence of arbitrage opportunities, all traded assets are martingales (i.e. driftless motions). Take the i-th forward and the numeraire $P(t,T_{i+1})$.

$$F(t,T_i)=\frac{1}{\tau_i}\left(\frac{P(t,T_i)-P(t,T_{i+1})}{P(t,T_{i+1})}\right) \tag{48}$$

F(t,T) is the simply compounded forward (libor) rate observed at time t for an for the maturity pair T_i-T_{i+1}.
If we consider the tenor, τ, and the top part of the right hand side as a traded asset, then there exists a probability measure for which this asset is a martingale. It is the T_i-forward adjusted measure Q^{i+1}, and its associated Brownian motion is W_t^{i+1}.

We can then write:

$$\frac{dF(t,T_i)}{F(t,T_i)}=\sigma(t,T_i)dW_t^{i+1} \tag{49}$$

This is the fundamental BGM equation. It is then understandable that the whole model lies around the specification of the instantaneous volatility function $\sigma(t,T_i)$. We now discuss the probability measure under which we perform the calibration: the forward equivalent measure.

4.2. The forward equivalent measure

The hypothesis that forward rates are log-normally distributed combined with the expectation market hypothesis[14] permit to obtain Black's model. Additionally, under the probability measure Q^{i+1}, $F(t,T_i)$ [15] is a martingale:

$$F(t,T_i) = \frac{1}{\tau}\left[\frac{P(t,T_i)}{P(t,T_{i+1}} - 1\right] \tag{50}$$

$$= \frac{1}{\tau}\left[\frac{P(t,T_i) - P(t,T_{i+1})}{P(t,T_{i+1})}\right] \tag{51}$$

$$\begin{aligned} &\Rightarrow F(t,T_i)P(t,T_{i+1}) = \frac{1}{\tau}\left[P(t,T_i) - P(t,T_{i+1})\right] \\ &\Rightarrow F(u,T_i)P(u,T_{i+1}) = \frac{1}{\tau}\left[P(u,T_i) - P(u,T_{i+1})\right] \end{aligned} \tag{52}$$

One can consider the last two equations to be traded assets in which case the relative prices are martingales:

$$\frac{1}{\tau}\left[\frac{P(t,T_i) - P(t,T_{i+1})}{P(t,T_{i+1})}\right] = E_t^{i+1}\left[\frac{1}{\tau}\frac{P(u,T_i) - P(u,T_{i+1})}{P(u,T_{i+1})}\right] \tag{53}$$

$$\Rightarrow F(t,T_i) = E_t^{i+1}\left[F(u,T_{i+1})\right] \tag{54}$$

---

[14] $E_t(L(T_i)) = F(t,T_i)\ \forall t$ and $T_i \geq t$.

[15] The forward expiring at time T_i and maturing at time T_{i+1}.

$F(t,T_i)$ is then a martingale under the forward neutral measure Q^{i+1} (i.e. when the numeraire is $P(t,T_{i+1})$).

We now know that each forward rate is a martingale under its own probability measure. Note that a single measure is used for Monte Carlo simulations. It is then important to identify the diffusion processes of the forward rate over a certain time interval $[t,T_{i+1}]$ under the forward neutral measure Q^{i+1}.

For example we have $i=3$ and T^1, T^2, T^3, T^4 is our time scale. $F(t,T_i)$ is the forward rate expiring in T_i and maturing in T_{i+1}. We therefore seek the diffusion processes for:

$$\frac{dF(t,T_3)}{F(t,T_3)} = \mu^{i+1}(t,T_3)dt + \sigma(T_3)dW^{i+1}{}_{3t} \tag{55}$$

$$\frac{dF(t,T_2)}{F(t,T_2)} = \mu^{i+1}(t,T_2)dt + \sigma(T_2)dW^{i+1}{}_{2t} \tag{56}$$

$$\frac{dF(t,T_1)}{F(t,T_1)} = \mu^{i+1}(t,T_1)dt + \sigma(T_1)dW^{i+1}{}_{1t} \tag{57}$$

All the σ s above are caplet volatilities extracted from market cap volatilities. According the preceding result, we know that $\mu^{i+1}(t,T_3)=0$. So, under Q^{i+1}:

$$\frac{dF(t,T_3)}{F(t,T_3)} = \sigma(t,T_3)dW^{i+1}{}_{3t} \tag{58}$$

For $\mu^{i+1}(t,T_2)$, if we use Vaillant's brackets (Rebonato 1999) and the Ito Lemma, the diffusion in Q^{i+1} corresponds to the following equation[16]:

$$\frac{dF(t,T_2)}{F(t,T_2)} = \left[\frac{\sigma_{F(t,T_2)}\sigma_{F(t,T_3)}\tau F(t,T_3)}{1+\tau F(t,T_3)}\rho_{F(t,T_2)F(t,T_3)}\right]dt + \sigma_{F(t,T_2)}dW^{i+1}{}_{2t} \quad (59)$$

and for $\mu^{i+1}(t,T_1)$:

$$\frac{dF(t,T_1)}{F(t,T_1)} = \sigma_{F(t,T_1)}\sum_{i=1}^{2}\left[\frac{\sigma_{F(t,T_{i+1})}\tau F(t,T_{i+1})}{1+\tau F(t,T_{i+1})}\rho_{F(t,T_1)F(t,T_{i+1})}\right]dt + \sigma_{F(t,T_1)}dW_{1t}^{i+1} \quad (60)$$

$\rho_{i,j}$ is the instantaneous correlation between two given forward rates. Since this quantity appears we will be required to specify a functional specification for it. The three equations we have just shown are the diffusion processes for our three forward rates under the forward neutral measure, Q^{i+1}.

The market model is interesting because it takes into consideration the instantaneous volatility of the forward rates as well as the instantaneous correlations. It is therefore richer than other models discussed before and it is critical to make the right modeling choices for these quantities as the whole forward rate process depends on it.

We now move on to the empirical part of our study and present the data set we use to calibrate the LMM to Canadian market quoted cap and swaption volatilities.

---

[16] Following Rebonato (1999), if $Z = \frac{f}{b}$ where Z is a martingale and where f and b are lognormaly distributed then: $\mu_f = -[f,b] = -(-\sigma_f - \sigma_b\rho_{f,b})$ and $\left[f,\frac{b}{c}\right] = [f,b]-[f,c]$

5. The data

All the data we use is for the Canadian market. We conduct our study using three types of data as of July 21, 2005: Swap data defining the term structure of interest rates, market implied volatilities for European swaptions, and market implied volatilities for Libor interest-rate caps. These implied volatilities define the market price of swaptions and caps. The source of all the data is the Bloomberg system where we collected market quotations from brokers and dealers in the OTC swap and fixed-income derivative market.

The term structure data consists of the mid-market one to ten-year par swap rates, with one year intervals. We use a standard cubic spline algorithm to interpolate the swap curve at semiannual intervals. From these swap rates we derive semi-annual spot rates using a bootstrapping method (i.e. we use the one-year T-bill as the first spot rate to solve for the one-year and a half spot rate using swap rates and so on). Finally, we solve for six-month forward rates by bootstrapping the spot curve.

We also invert the spot rates to find the zero coupon bond prices giving us the discount factor.

Table 1 reports this data. The term structure of forward rates is also graphed in Figure 1. Figure 1' shows the forward rates for the sample period from January 4, 2005 to June 30, 2005 which will be used to extract the historical correlation matrix.

Table 1. Canadian data on July 21, 2005.

Source: Bloomberg LP.

Year	Swap rates	Spot	Discount Factor	Forward
1	3.05	2.972	0.9713	3.21
1.5	3.14	3.09	0.9557	3.21
2	3.23	3.17	0.9404	3.27
2.5	3.30	3.25	0.9249	3.34
3	3.37	3.32	0.9095	3.40
3.5	3.44	3.39	0.8940	3.48
4	3.52	3.46	0.8783	3.55
4.5	3.60	3.54	0.8625	3.63
5	3.68	3.62	0.8466	3.72
5.5	3.75	3.71	0.8307	3.80
6	3.83	3.79	0.8148	3.88
6.5	3.89	3.87	0.7991	3.95
7	3.96	3.94	0.7837	4.02
7.5	4.02	4.01	0.7688	4.08
8	4.07	4.07	0.7542	4.14
8.5	4.12	4.13	0.7400	4.19
9	4.16	4.19	0.7262	4.24
9.5	4.21	4.24	0.7129	4.29

Figure 1. Term structure of Libor forward rates on July 21, 2005 with maturities ranking from one year to ten years.

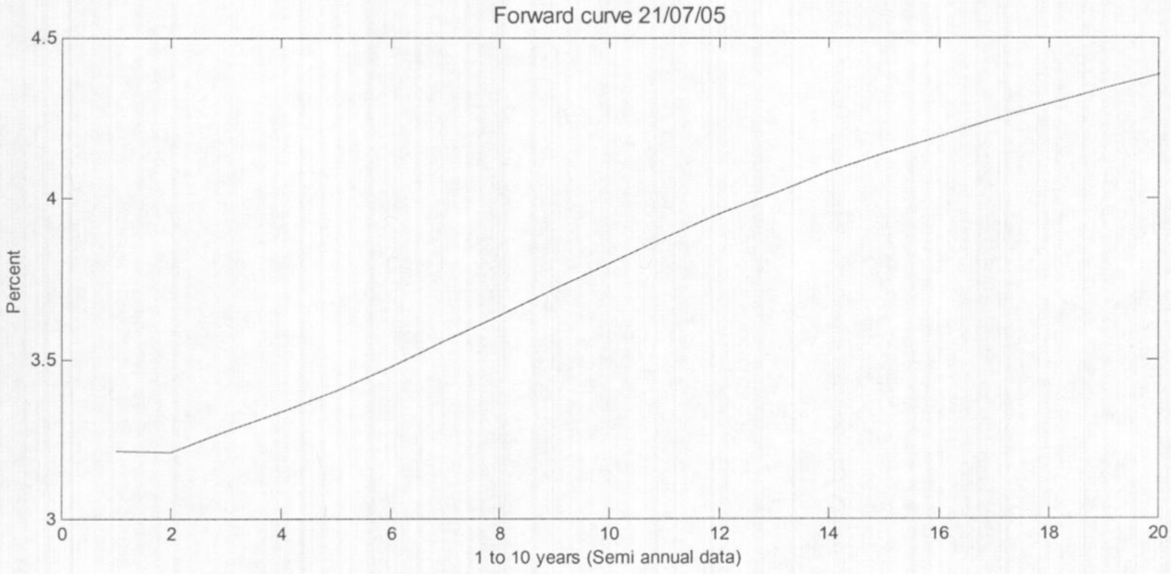
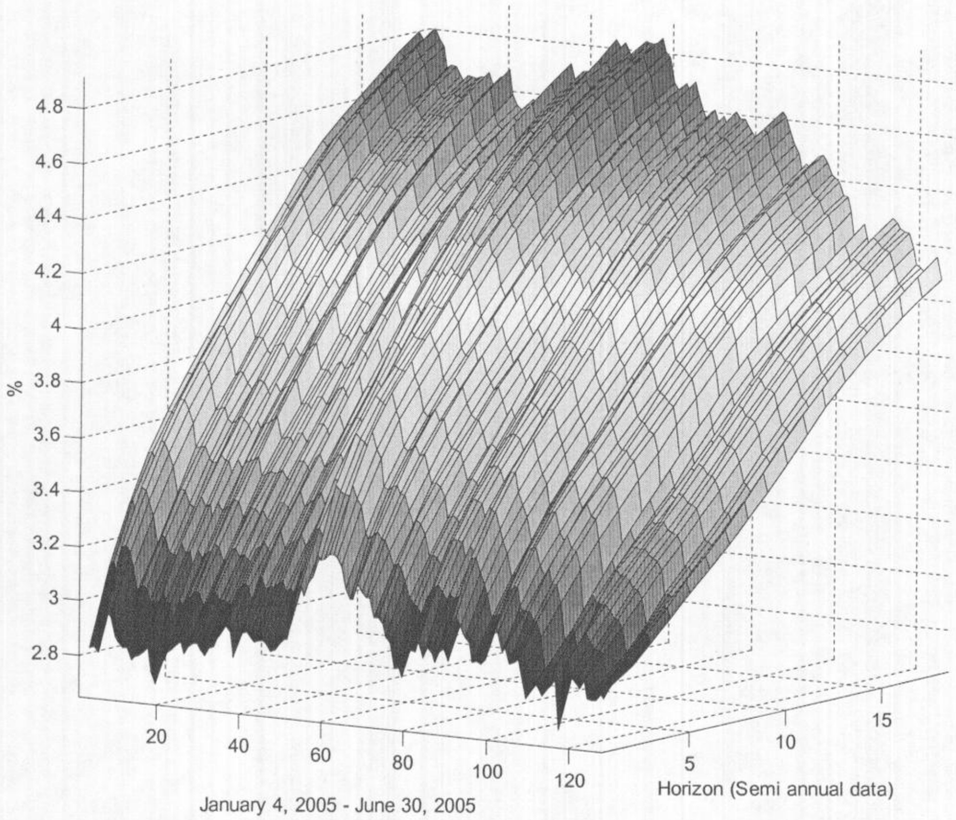


Figure 1'. Time Series of Six-Month Libor Forward Rates.

The data set consists of daily observations of six-month Libor forward rates starting at 1 to 9.5 years, for the period from January 4, 2005 to June 30, 2005. The forward rates are computed from the one year cdor as well as the two-year, three-year, five-year, seven-year and ten-year mid-market swap rates using a cubic spline to interpolate the curve and then bootstrapping the forward curve. All data comes from the Bloomberg system. The daily data for interest rates represents the closing rates. The total number of observations in the sample is 128.



We also use the daily term structure data from January 4, 2005 to June 30, 2005 to estimate the historical correlation matrix from which the eigenvectors are determined[17].

This correlation matrix is shown in Table 2.

[17] We choose to use a 6 month time frame to minimize as much as possible the impact of older market conditions which would not apply today.

Table 2. Correlation Matrix of Log Changes in Six-Month Libor Forward Rates.

The correlation matrix is based on daily changes in the logarithm of individual six-month Libor forward rates for the period from January 4, 2005 to June 30, 2005. The forward rates are computed from the one year cdor as well as the two-year, three-year, five-year, seven-year and ten-year mid-market swap rates using a cubic spline to interpolate the curve and then bootstrapping the forward curve. All data comes from the Bloomberg system. The daily data for interest rates represents the closing rates. The horizons of the six-month forward rates used to compute the correlation matrix range from 1 year to 9.5 years forward, giving a total of 18 time series of forward rates. The total number of observations is 128.

	1,00	1,50	2,00	2,50	3,00	3,50	4,00	4,50	5,00	5,50	6,00	6,50	7,00	7,50	8,00	8,50	9,00	9,50
1,00	1,000																	
1,50	0,964	1,000																
2,00	0,622	0,802	1,000															
2,50	0,324	0,549	0,934	1,000														
3,00	0,336	0,533	0,868	0,954	1,000													
3,50	0,344	0,503	0,768	0,851	0,962	1,000												
4,00	0,259	0,411	0,686	0,777	0,865	0,945	1,000											
4,50	0,157	0,310	0,627	0,743	0,808	0,883	0,975	1,000										
5,00	0,037	0,184	0,546	0,703	0,763	0,785	0,808	0,889	1,000									
5,50	-0,060	0,064	0,414	0,587	0,644	0,638	0,617	0,716	0,942	1,000								
6,00	-0,110	-0,014	0,293	0,460	0,518	0,536	0,548	0,622	0,765	0,884	1,000							
6,50	-0,197	-0,106	0,210	0,390	0,437	0,464	0,509	0,584	0,686	0,799	0,967	1,000						
7,00	-0,313	-0,206	0,172	0,383	0,405	0,404	0,450	0,566	0,739	0,781	0,718	0,795	1,000					
7,50	-0,347	-0,242	0,143	0,362	0,376	0,358	0,389	0,516	0,729	0,761	0,618	0,671	0,960	1,000				
8,00	-0,326	-0,238	0,123	0,338	0,366	0,354	0,369	0,478	0,686	0,781	0,755	0,740	0,756	0,835	1,000			
8,50	-0,318	-0,253	0,057	0,259	0,303	0,306	0,316	0,403	0,580	0,703	0,761	0,723	0,586	0,656	0,956	1,000		
9,00	-0,436	-0,391	-0,096	0,133	0,208	0,243	0,268	0,350	0,509	0,633	0,716	0,719	0,615	0,628	0,816	0,884	1,000	
9,50	-0,442	-0,433	-0,246	-0,056	0,030	0,091	0,128	0,182	0,269	0,343	0,410	0,472	0,483	0,420	0,363	0,418	0,792	1,000

The swaption data consists of daily midmarket implied volatility for 42 at-the-money European swaptions for July 21, 2005. These 42 swaptions represent all of the standard quoted N by M European swaption structures where the final maturity date of the underlying swap is less than or equal to ten years, $T \leq 10$. As mentioned earlier, the market convention is to quote swaptions in terms of there implied volatility according to the Black (1976) model for at-the-money European swaptions. Note that the market price of swaptions is given by substituting the implied volatilities into the Black model. Table 3 will shows the swaption data we use for the calibration. Figure 2 shows the shape of the swaption implied volatility surface.

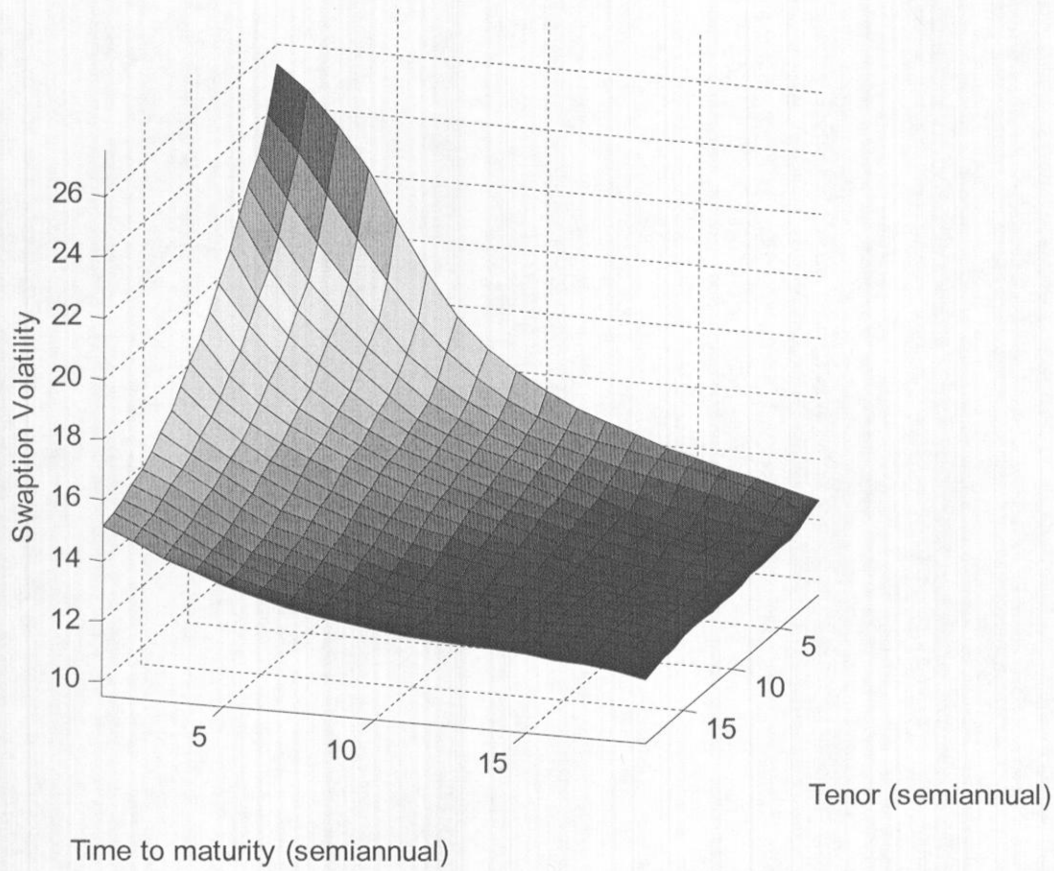
Table 3. At-the-money European Cap/Floor and Swaption Volatilities. The data consists of 42 mid-market implied Black-model swaption volatilities and 7 mid-market implied Black cap/floor volatilities for July 21, 2005.

	1y	2y	3y	4y	5y	7y	10y
cap/floor	19.25	24.88	26	25	23.75	21.5	19.25

Swaptions		Swap							
		1y	2y	3y	4y	5y	7y	10y	30y
Option	1m	18.63	21	20.63	20.75	20	17.5	15.88	9.75
	3m	22.5	24.75	23.75	22.63	21.63	18.5	16.38	10.38
	6m	22.38	23.63	22.5	21.13	20.13	17.63	15.75	11.13
	1y	25.38	23	21.5	19.88	18.63	16.63	15.13	11.38
	2y	23.5	20.25	18.63	17.5	16.63	15.38	14.25	11.38
	3y	20.63	17.88	16.88	16	15.5	14.38	13.5	11.38
	4y	17.63	16.13	15.38	15	14.63	13.75	13	11.5
	5y	16	15	14.63	14.25	14	13.13	12.63	11.5
	7y	14.25	13.75	13.25	13.13	12.88	12.5	12.25	11.25
	10y	12.63	12.25	12.13	12.13	12	12	11.63	11.25

Figure 2. Example of Swaption Volatility Surface.

This figure plots the quoted volatilities of swaptions on July 21, 2005. The maturities range from 1 year to 10 years on underlying swaps with horizons at the maturity of the options between 1 and 10 years.



The interest-rate cap data consists of daily midmarket implied volatilities for one-year, two-year, three-year, four-year, five-year, seven-year and ten-year caps on the same date as for the swaption data. By market convention, the strike price of a T-year cap is the T-year swap rate. In order to simplify the analysis of the data, we assume that caps are on the six-month Libor rate rather than the three-month rate. We therefore use a six-month tenor structure on both swaptions and caps[18]. The market prices of caps are then given by substituting the quoted implied volatilities

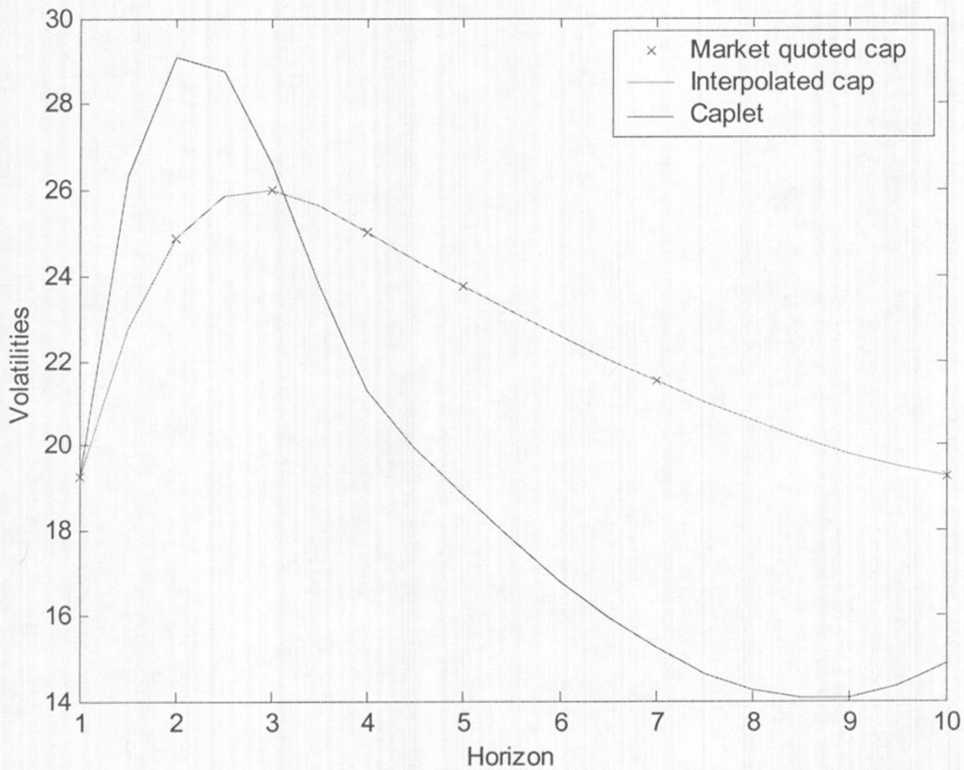
[18] This assumption has no effect on the empirical results according to Longstaff, Sana-Clara and Schwartz (2000).

into the Black (1976) formula. Table 3 presents the cap data for our sample period.

We will be working with caplet volatilities to implement the LMM. To extract them from cap volatilities we use the algorithm presented in section 3.3. Figure 3 shows the cap and the extracted caplet volatilities on July 21, 2005.

Figure 3. Cap and caplet volatilities.

Midmarket implied Black-model cap and caplet volatilities on July 21, 2005.



Now that we are comfortable with the data set, we can start with the calibration of the LMM to Canadian cap and swaption volatilities.

6. BGM calibration

In this section we address the heart of the matter which is to recover market data, according to the Black setting, using the LMM. It is quite straightforward to calibrate the model to the caplet volatilities but we show how to simultaneously calibrate to swaption volatilities as well. We do this although it poses problems on both a theoretical and practical level because rates can't follow a lognormal distribution for caps and swaptions simultaneously (as outlined in 3.3). We review all the different assumptions that we need to make and the modeling choices for the calibration.

In our "step-by-step" implementation we review the different instantaneous volatility specifications available in the literature for forward rates. The second step is to take a look at the correlation structure between forward rates. Thirdly, we perform a Principal Component Analysis (PCA) and identify the contribution of each factor to the total volatility.
Then we provide a closed form formula to approximate for the price of swaptions. This allows us to offer calibration results and discus possible remaining issues.

6.1. Step-by-step implementation

In this section, we outline each required step to perform a proper calibration of the LMM to cap and swaption volatilities.

6.1.1. Specification of the instantaneous volatility function

Within the modeling framework we need to specify a functional form for the instantaneous volatility of forward rates.

The simplest way is to choose a total parameterization of the volatility by considering that every $\sigma_{i,j}$ is independent and to fit the cap and swaption market. As outlined in Brigo and Mercurio (2001) this causes overparamerization problems. Therefore it is preferable to use a semi-parametric form.

Brigo and Mercurio (2001) outline different functional forms for the instantaneous volatility function.

The most interesting one is the following:

$$\sigma(t,T_i) = \Phi(T_i)\psi(T_i - t; a,b,c,d) := \Phi(T_i)\left(\left[a(T_i - t) + d\right]e^{-b(T_i - t)} + c\right) \qquad (61)$$

It is the richest form encountered which allows for a humped shape of the instantaneous volatilities of the forward rate ($F(t,T_i)$) as a function of the time to maturity. It has a parametric core ψ that is altered by the Φ's for each maturity T_i.

The formula we use to calibrate to the market volatilities is:

$$\sigma^2{}_{caplet,i} = \Phi_i^2 \sum_{j=1}^{i} \tau_{j-2,j-1}\psi_{i-j+1}^2 \qquad (62)$$

Given that the squares of the caplet volatilities (multiplied by time) are read from the market, we can identify the parameters Φ using the ψ's through:

$$\Phi_i^2 = \frac{(\sigma_{caplet,i})^2}{\sum_{j=1}^{i} \tau_{j-2,j-1}\psi_{i-j+1}^2} \qquad (63)$$

The caplet volatilities are therefore incorporated in the model by determining the thetas in terms of the psis. The psis together with the instantaneous correlations of the forward rates are then used to calibrate to swaption volatilities.

The target function for the volatility extraction is:

$$\min_{a,b,c,d} \sum_i \eta_i \left(\sigma_{caplet,i}\sqrt{t_i} - \Phi_i \sqrt{\int_0^{T_i} \left(\left[a(T_{i-1} - 1) + d\right]e^{-b(T_{i-1} - t)} + c\right) dt} \right)^2 \qquad (64)$$

The weights η_i account for the different quality of option prices (and volatilities).

Overall, this form allows for more flexibility and improves the joint calibration of the model to caps and swaptions.

We now discuss the other important aspect of the calibration procedure: the specification of the correlation structure.

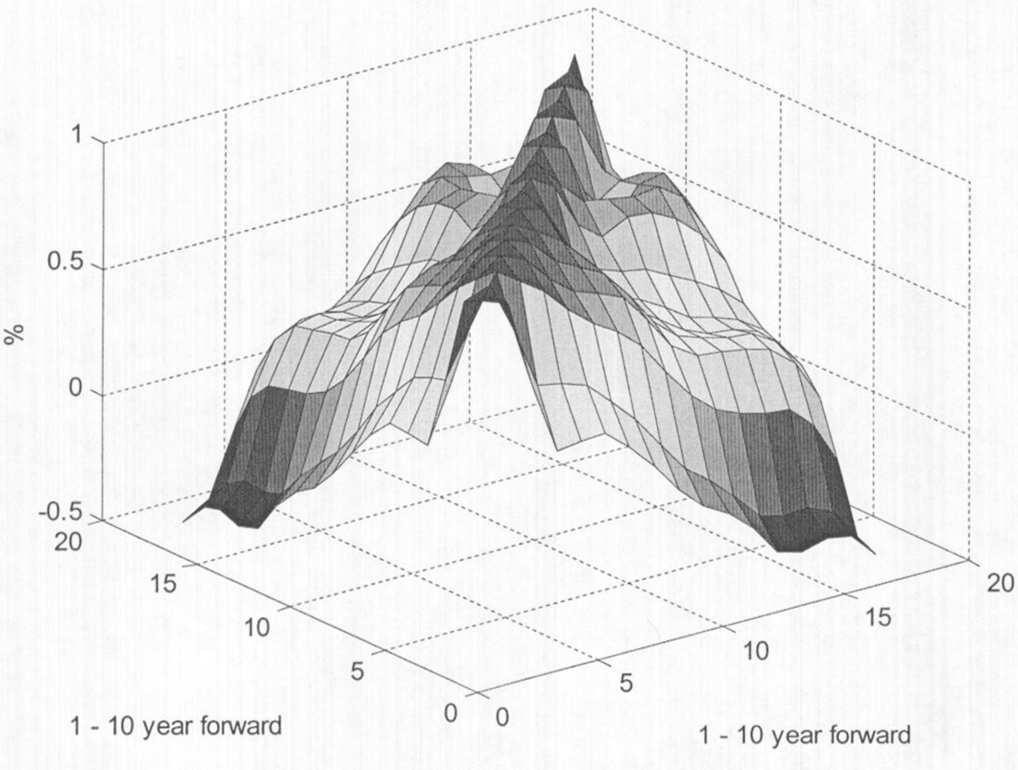
6.1.2. Specification of the correlation structure

Many articles in the literature present theoretically sound instantaneous correlations structure specifications. They often contain two, three or more factors. The correlation structure is one of the key to the multifactor BGM model. The reason is that forward rates have a lognormal distribution. In practice, they are not completely independent from one another. They are linked by arbitrage like we saw earlier, but we also consider these processes to be correlated.

We have a choice between two types of correlation structures: one based on historical data and the other on parametric forms.

- **Historical data**: This is the most straightforward way to estimate correlations. Unfortunately, in practice this procedure proves to contain noise in the data and is difficult to use. This is why many authors propose parametric forms which permit to smooth the data. Figure 4 shows the historical correlation surface of log changes in six-month Libor forward rates. Table 2 shows the correlation matrix of log changes in the six-month Libor forward rates for our sample. It is interesting to note that the correlation between short and long term rates is negative. This shows us that as the Bank of Canada increased short term rates, the bond market diminished the risk premium for inflation. Therefore bond yields where decreasing. This is in opposition to past market conditions where correlation where positive.

Figure 4. Historical Correlation Surface of Log Changes in Six-Month Libor Forward Rates. The correlation matrix is based on daily changes in the logarithm of individual six-month Libor forward rates for the period from January 4, 2005 to June 30, 2005. The forward rates are computed from the one year cdor as well as the two-year, three-year, five-year, seven-year and ten-year mid-market swap rates using a cubic spline to interpolate the curve and then bootstrapping the forward curve. All data comes from the Bloomberg system. The daily data for interest rates represents the closing rates. The horizons of the six-month forward rates used to compute the correlation matrix range from 1 year to 9.5 years forward, giving a total of 18 time series of forward rates. The total number of observations is 128.



- **Parametric forms**: The simplest correlation structure is exponentially decreasing and is function of the distance to the diagonal (Brigo and Mercurio 2001), (Rebonato and Joshi 2001).

$$\rho_{i,j} = \exp^{-\rho|i-j|} \tag{65}$$

Unfortunately, in practice, it doesn't permit a very precise calibration because it doesn't contain any information on the speed of decorrelation.

Coffey and Schoenmakers (2002) propose a semi-parametric form which allows the correlation of forward rates to increase with maturity:

$$\rho_{i,j}(\eta_1,\eta_2,\rho_\infty)=\exp\left[-\frac{|i-j|}{m-1}\left(\ln\rho_\infty+\eta_1\frac{i^2+j^2+ij-3mi-3mj+3i+3j+2m^2-m-4}{(m-2)(m-3)}\right.\right.$$
$$\left.\left.+\eta_2\frac{i^2+j^2+ij-mi-mj-3i-3j+3m+2}{(m-2)(m-3)}\right)\right] \quad (66)$$

$$(i,j)\in[1,m]^2, 3\eta_1\geq\eta_2\geq 0,\ 0\leq\eta_1+\eta_2\leq-\ln\rho_\infty$$

This is more complicated but more robust. According to Mony Lim (2001), it accommodates matrices which are qualitatively acceptable.

For instantaneous correlation there is also the two factor Rebonato (1999) angle formulation.

$$\rho_{i,j}=\cos(\theta_i-\theta_j)\,. \quad (67)$$

Rebonato (1999) also proposes the following three parameters form:

$$\rho_{i,j}=\rho_\infty+(1-\rho_\infty)\exp\left(|T_i-T_j|\,\beta(T_i,T_j)\right)$$
$$\beta(T_i,T_j)=d1-d2\max(T_i,T_j) \quad (68)$$

This allows to include some information in the infinite correlation and to generate a large number of matrices. It can provide surprising results like values greater than one which is why it isn't ideal.

For our study, we use Rebonato's two factor angle formulation as it has been used by many authors successfully.

Now that we have discussed the instantaneous volatility and the instantaneous correlation of forward rates, we address the question of how many factors should be used to perform the calibration. Alternatively, we also identify how and which factors affected the Canadian market in the first half of 2005.

6.1.3. How many implied factors? The Principal Components Analysis

Researchers find that two or three factors are sufficient to capture the historical variation in the term structure. We are going to look at a few cases and identify how much of the variation of the Libor forward rates the first few factors explain. Principal Component Analysis (PCA) consists in taking the diagonal of the correlation matrix and to retain the largest eigenvectors.

We look at the eigenvectors of the instantaneous correlation matrix in Table 4 and see that the first two represent 89.15% of all eigenvectors, that the first three factors represent 95.85% of the sum of eigenvectors and that the first four eigenvectors represent 99.33%.

Table 4. Eigenvector Weights.

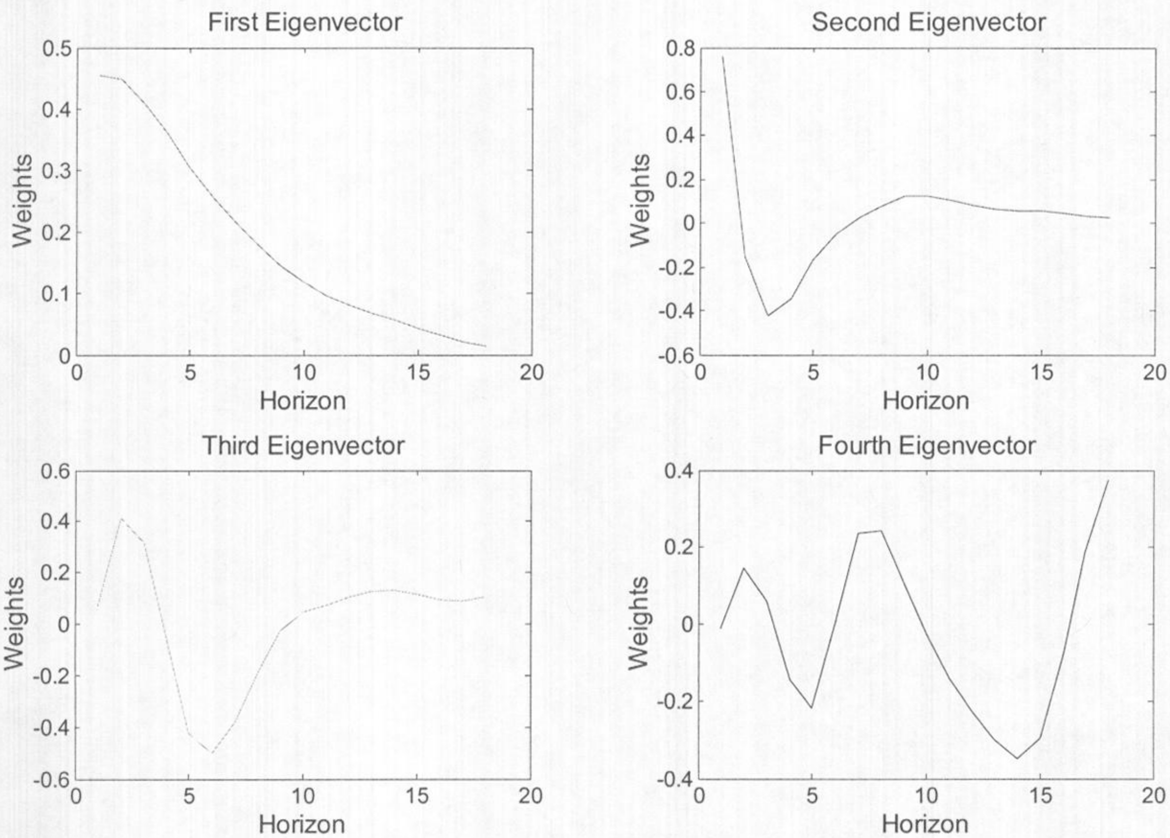
The table below shows the weights of the first four eigenvectors of the historical correlation matrix. The correlation matrix is based on the daily changes in the logarithm of the individual six-month Libor forward rates for the period from January 4, 2005 to June 30, 2005. The forward rates are computed from the one year cdor as well as the two-year, three-year, five-year, seven-year and ten-year mid-market swap rates using a cubic spline to interpolate the curve and then bootstrapping the forward curve. All data comes from the Bloomberg system. The daily data for interest rates represents the closing rates. The horizons of the six-month forward rates used to compute the correlation matrix range from 1 year to 9.5 years forward, giving a total of 18 time series of forward rates. The total number of observations is 128.

Eigenvectors weights	1st	2nd	3rd	4^{th}	Sum of the other 15
Value	2.79	3.30	0.47	0.25	0.27
%	40.80%	48.35%	6.70%	3.48%	3.77%

To provide some insight into the four implied factors that traders view as driving the term structure we refer to Figure 5 which graphs the first four eigenvectors, which define the weights of the first four factors, from the historical matrix in Table 2.

Figure 5. Eigenvector Weights.

The four graphs below show the weights of the first four eigenvectors of the historical correlation matrix. The correlation matrix is based on daily changes in the logarithm of individual six-month Libor forward rates for the period from January 4, 2005 to June 30, 2005. The forward rates are computed from the one year cdor as well as the two-year, three-year, five-year, seven-year and ten-year mid-market swap rates using a cubic spline to interpolate the curve and then bootstrapping the forward curve. All data comes from the Bloomberg system. The daily data for interest rates represents the closing rates. The horizons of the six-month forward rates used to compute the correlation matrix range from 1 year to 9.5 years forward, giving a total of 18 time series of forward rates. The total number of observations is 128.



As we can see, these factors are comparable to those found in previous works, done for the US markets (Longstaff & Schwartz (2000). According to Rebonato (1998) the first factor generates the level of the curve which is also an approximate for the short rate. The second factor generates shifts in the yield curve. The third factor is a "curvature factor which generates movements in the term structure where short term and long term rates move in opposite directions from the mid-term rates". The fourth factor represents the shape of the very short end of the curve, which might be affected by central bank influence or other monetary authorities according to Longstaff and Schwartz (2000).

Our results imply that that 95.85% of the variance can be explained by the first three factors. Therefore the 4.15% that are left can be considered as noise. To refine the analysis we can say that 99.33% of the variance can be explained by the first four factors where the remaining 0.67% is noise.

What is interesting to note is that the weight of the second factor is superior to the weight of the first factor (48.35% and 40.88% respectively). In Canada, during the first semester of 2005, the yield curve has flattened out and the short rate hasn't moved much. Typically, the short rate influences the yield curve and tends to be the most significant factor. What we understand here, is that the eigenvector represent market conditions where there has been a steepening of the curve and therefore, traders paid more attention to the shift in the yield curve than they would have under other circumstances.

We see that three factors is satisfactory to explain the term structure movements of the Canadian market in the first half of 2005. We now discuss the simultaneous calibration of the LMM to cap and swaption volatilities.

6.1.4. LMM calibration to cap and swaption volatilities

In this section we show the methods that can be used to calibrate the LMM to cap and swaptions volatilities. We show the closed form formula identified by Rebonato (2000).

The general diffusion process of the forward rate used to do Monte Carlo simulations is the one outlined by Rebonato (2000). He says that the realization at time T of the k-th log-normal forward rate, of value $F_k(0,T)$ today, in terms of its time-dependent instantaneous volatility, $\sigma_k^{inst}(u,T)$, is given by:

$$F_k(T,T) = F_k(t_0,T)\exp\left[\int_0^T \mu_k(u,T) - \frac{\sigma_k^{inst}(u,T)^2}{2} du\right]\exp\left[\int_0^T \sigma_k^{inst}(u,T)dW_u\right] \quad (69)$$

Where $\mu_k(u,T)$ is the value at time u of the time-dependent drift that ensures the model is arbitrage free. The drift, (source: Lutz Molgedey (2002)) under the T_i-forward equivalent measure, Q^{i+1}, has the following form:

$$dL(t,T_i) = L(t,T_i)\left(\mu_i^{Q^{i+1}}(t)dt + dW_t^{i+1}\right) \quad (70)$$

where given the libor rate $L(t,T_i) = \frac{1}{\delta_i}\left(\frac{P(t,T_i)}{P(t,T_{i+1})} - 1\right)$ and P(t,T) is the zero coupon bond prices with maturity T and with

$$\mu_i^{Q^{i+1}} = -\sum_{j=i+1}^{N} \frac{\tau_j L(t,T_j)\operatorname{cov}(T_i,T_j)}{1+\sigma_j L(t,T_j)} \quad \text{with} \quad \langle dW_i(t)dW_j(t)\rangle = \operatorname{cov}(T_i,T_j)d_t \quad (71)$$

More generally, in the case of r orthogonal driving factors[19]

$$F_k(T,T) = F_k(t_0,T)\exp\left[\int_0^T \mu_k(u,T) - \frac{\sigma_k^{inst}(u,T)^2}{2} du\right]\exp\left[\int_0^T \sum_{m=1,r}\sigma_{km}^{inst}(u,Tm)dW_m(u)\right] \quad (72)$$

under the constraint that

$$\sum_{m=1,r}\sigma_{km}^{inst}(u)^2 = \sigma_k^{inst}(u)^2 \quad (73)$$

Where m is indicating the factor m=1…r. As long as the last two equations are satisfied, the variance of the forward rates, and therefore the caplet prices, will be recovered irrespective of the number of driving factors. According to Rebonato (2000), as far as the pricing of either caplets or European swaptions is concerned, any number of factors can be used to obtain their exact prices. He says that one factor is enough as the quantities to be matched are

[19] The assumption is made that $E\lfloor dW_i, dW_j \rfloor = \rho(T_i,T_j)dt$.

$$\sigma_{Black}{}^{2}(T-t_0)=\int_0^T \sigma^{inst}(u)^2\,du \tag{74}$$

Here, σ_{Black} is the market implied black volatility for the forward rate or swap rate expiring at time T and $\sigma^{inst}(u)$ is the instantaneous volatility of the same rate from today to the expiry.

As we have seen earlier, the instantaneous volatility of forward rates and swap rates has a functional form. We need to determine what these parameters must be before using Monte Carlo simulations to price caps and swaptions.

We choose to simultaneously calibrate the caps and the swaptions using the following formula for swaption volatilities:

$$v^2(T_a,T_b)=\sum_{i,j=a+1}^{b}\frac{w_i(0)w_j(0)F(t,T_{i-1},T_i)F(t,T_{j-1},T_j)\rho_{i,j}}{T_a FS(0,T_a,T_b)^2}\sigma_i{}^{inst}\sigma_j{}^{inst} \tag{75}$$

Where $v^2(T_a,T_b)$ is the variance of the swaption starting with maturity a and expiry b. $FPS(0,T_a,T_b)$ is a forward start swap rate observed in time 0.

$$w_i(0)=\frac{\tau P(0,T_i)}{\sum_{m=a+1}^{b}\tau P(0,T_m)} \tag{76}$$

σ_i^{inst} is the instantaneous volatility of the forward rate $F(t,T_{i-1}, T_i)$ with the specification outlined earlier (61) and $\rho_{i,j}$ is the instantaneous correlation $\frac{dF(T,T_{i-1},T_i).dF(T,T_{j-1},T_i)}{\sigma_{dF(T,T_{i-1},T_i)}\sigma_{dF(T,T_{j-1},T_i)}}$ with the specification we outlined earlier as well (68).

As noted previously, we should find the instantaneous volatility and correlation parameters to perform Monte-Carlo simulations on the forward rates in order to obtain swaption prices.

Performing Monte-Carlo simulations at each time step necessitates a long time when running computer programs. It is why we prefer a closed form formula which permits to calibrate the model to market data.

Hull and White (2000) propose a different approximation formula but Brigo and Mercurio (2001) outline that the difference between Rebonato's formula and Hull and White's is negligible. We therefore opt for Rebonato's closed form formula (72).

We have now identified the closed form formula (72) for the joint calibration of caps and swaptions. We now present the calibration procedure in details and some results.

6.2. Calibration Results[20]

In this section, we present some numerical results regarding the simultaneous goodness of fit of the LMM to the cap and swaption market quoted volatilities. We study the impact of the choice of initial parameters on the goodness of fit as well as on the instantaneous correlation matrix.

In order to obtain a satisfactory calibration we calibrate the caps and swaptions using the volatility structure outlined in (61) with a local algorithm minimization for finding the best fitting parameters $\psi_1,\ldots,\psi_{19}$ and $\theta_1,\ldots,\theta_{19}$.

We tested many calibration cases involving different initial guesses for the $\Phi's, \psi's$ and $\theta's$ as well as for the initial parameters a,b,c,d in the instantaneous volatility (61) and have decided to present the best two performing cases here. We believe that they summarize adequately how the calibration of the LMM to Canadian swaption and cap data can be performed.

[20] The methodology is inspired by Brigo and Mercurio (2001).

First case

The initial guesses for the ψ 's, the Φ 's and the θ 's are outlined in table 5[21].

Table 5. The initial guesses for the ψ 's, the Φ 's and the θ 's.

Index	Psi	Phi	Theta
1	2,583	0,149	0,015
2	1,633	0,159	0,064
3	1,108	0,153	0,103
4	1,811	0,145	0,15
5	0,232	0,136	0,197
6	2,387	0,127	0,224
7	0,401	0,121	0,277
8	1,513	0,118	0,295
9	1,269	0,114	0,363
10	0	0,111	0,381
11	3,234	0,108	0,422
12	0	0,105	0,484
13	0,509	0,102	0,52
14	1,177	0,099	0,542
15	0	0,098	0,579
16	0	0,097	0,65
17	0	0,097	0,668
18	0	0,097	0,713

By using formula (61) for the instantaneous volatilities, we obtain the Φ 's as functions of the ψ 's by using the (annualized) caplet volatilities and formula (63).

For the instantaneous volatility parameters *a, b, c, d,* we choose a=0.0285, b=0.20004, c=0.11, d=0.057 as initial values.

Using the parameters outlined for the instantaneous volatilities, we compute swaption prices using Rebonato's formula (72). We also impose the constraints

$$-\frac{\pi}{2} < \theta_i - \theta_{i-1} < \frac{\pi}{2}$$

21 Source: table 8.1 of Brigo and Mercurio (2001).

to the correlation angles, which imply that correlations $\rho_{i,i-1}>0$ (adjacent rates are positively correlated). Table 6 shows the obtained parameters.

Table 6. Calibration results: first case: parameter values.

Index	Phi	Psi	Theta
1	0.090	2.583	0.015
2	0.107	1.633	-0.069
3	0.119	1.108	-0.025
4	0.125	1.811	-0.446
5	0.129	0.232	0.449
6	0.130	2.387	-0.676
7	0.130	0.401	0.599
8	0.130	1.513	-0.803
9	0.131	1.269	0.563
10	0.131	0.000	-1.137
11	0.131	3.234	0.403
12	0.132	0.000	0.719
13	0.132	0.509	0.917
14	0.132	1.177	0.983
15	0.133	0.000	1.187
16	0.133	0.000	1.614
17	0.134	0.000	1.233
18	0.135	0.000	1.009

The fitting is exact for caplets whereas for swaptions there are some discrepancies. We attribute the discrepancies to the fact that the Canadian swaption market is not liquid for all maturities as it is in the US and in Europe.

Table 7 and Figure 6 show the matrix and plot of percentage errors between market swaption volatilities and the LMM's calibrated volatilities. Figure 7 shows the difference in volatility surface between the market swaption volatilities and the fit of the calibration. The calibration was performed on swaptions with maturities ranging from one year to five years in the option and the underlying swaps because this is where the Canadian market is the most liquid and it saves time during the optimization.

Table 7: Percentage Difference between Market Swaption Volatilities and LMM Volatilities (first case).

The swaptions which are the most liquid on the Canadian market are highlighted in Blue.

Option Maturity	Swap maturity								
	1Y	1.5Y	2Y	2.5Y	3Y	3.5Y	4Y	4.5Y	5Y
1Y	5.74%	1.69%	1.61%	2.24%	2.24%	3.97%	2.97%	5.31%	3.26%
1.5Y	-1.14%	-1.68%	-2.65%	-1.47%	-1.47%	-0.48%	2.47%	0.63%	0.58%
2Y	0.87%	-2.31%	-2.49%	-2.47%	-2.47%	1.39%	-0.34%	-0.10%	0.90%
2.5Y	2.09%	-2.45%	-1.63%	-3.58%	-3.58%	-0.09%	-0.53%	0.14%	0.52%
3Y	2.15%	-2.66%	-3.18%	0.48%	0.48%	0.81%	2.17%	3.12%	4.99%
3.5Y	1.79%	-5.14%	3.89%	-0.68%	-0.68%	-0.55%	-0.73%	0.18%	4.05%
4Y	-4.50%	-2.61%	-2.61%	-0.47%	-0.47%	2.65%	4.27%	9.22%	7.81%
4.5Y	8.73%	-1.75%	-4.92%	-5.88%	-5.88%	-6.32%	-2.95%	-4.11%	-6.22%
5Y	-1.62%	1.53%	0.39%	-1.58%	-1.58%	1.50%	-0.93%	-3.67%	-6.07%

Figure 6. Plot of the percentage error in the swaptions calibration (first case).

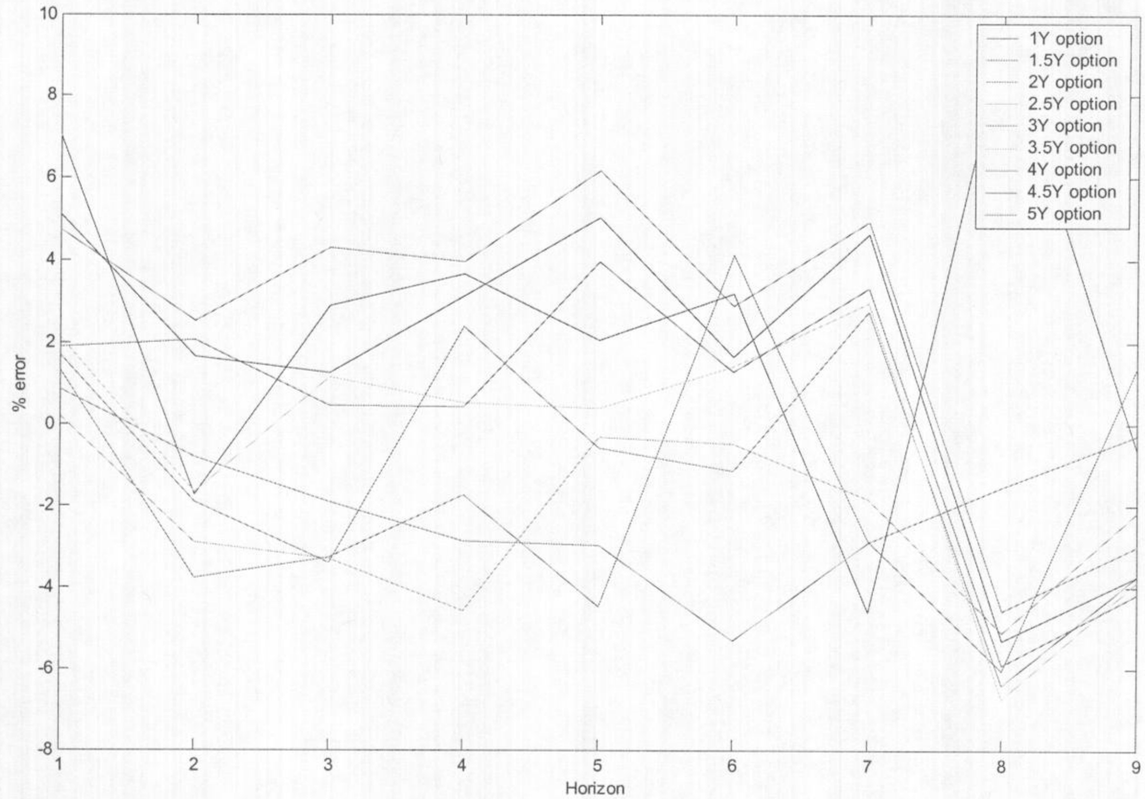
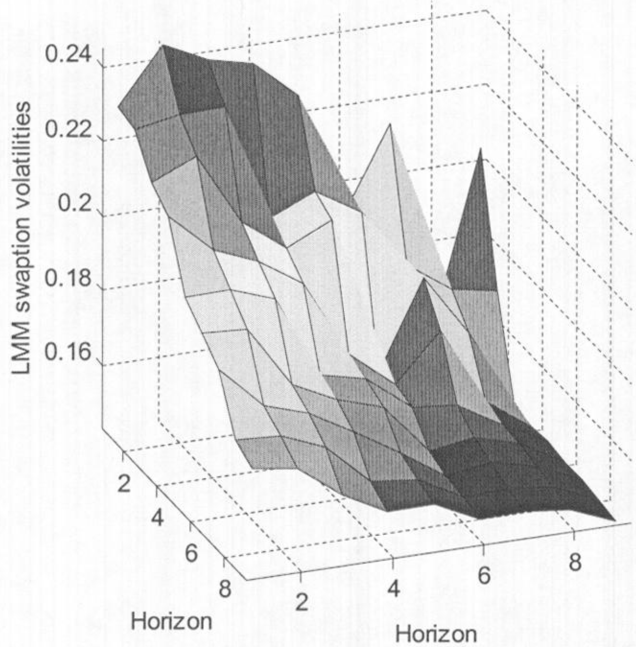
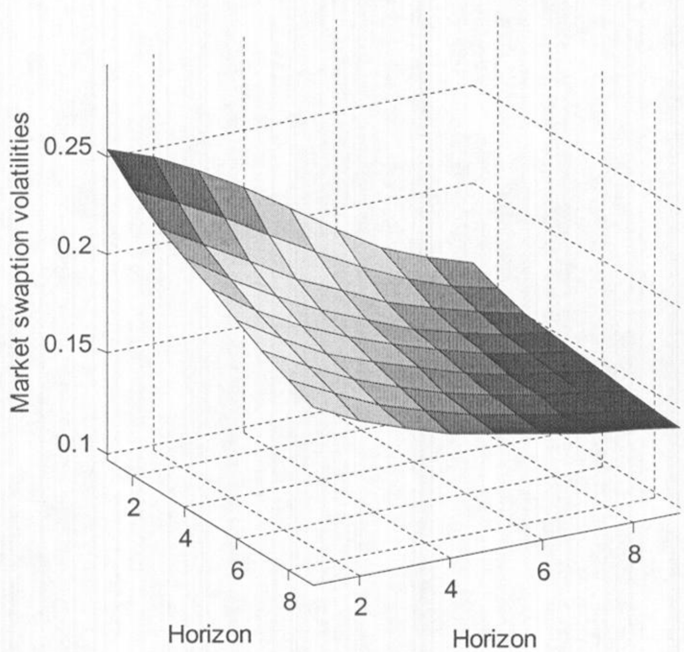


Figure 7. Comparison of the market swaption volatilities and the swaption calibration (first case).





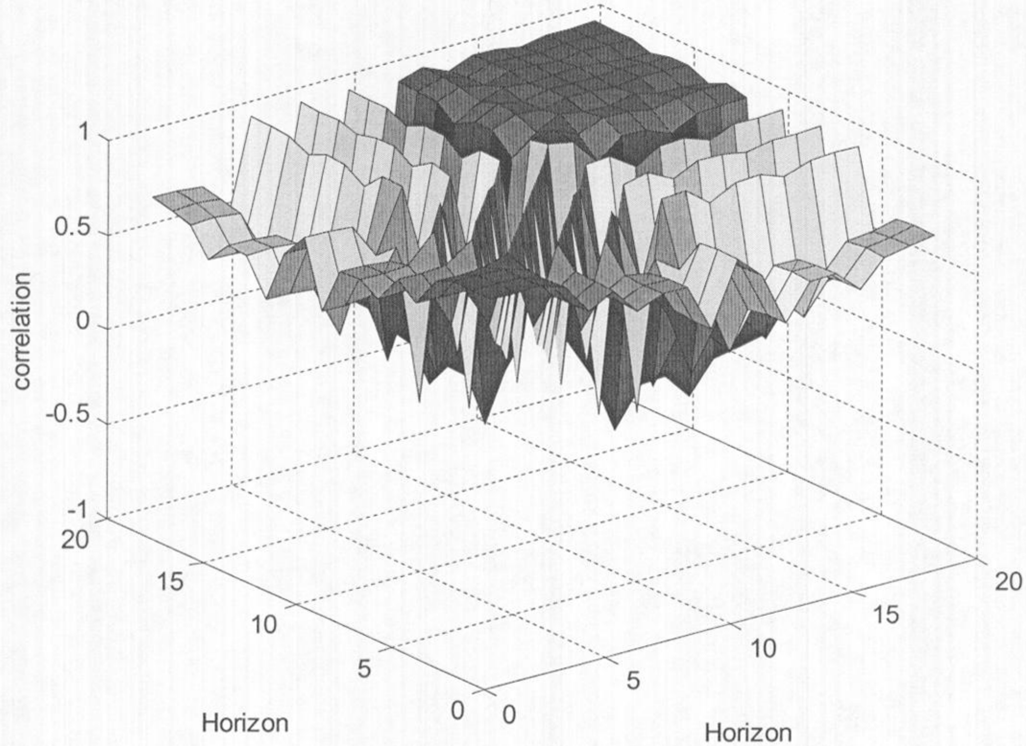
We note errors are acceptable and actually small for most maturities considering we are trying to fit 19 caplets and 81 swaption volatilities.

However, the calibrated θ's imply strange instantaneous correlations. The instantaneous correlations matrix, in Table 8 and Figure 8, displays correlations which vary from positive to negative values. This pattern is not satisfactory and shows us that the fitting quality is not the only criterion by which the calibration should be judged.

Table 8. Calibration results: Instantaneous correlation matrix (first case).

	1,00	1,50	2,00	2,50	3,00	3,50	4,00	4,50	5,00	5,50	6,00	6,50	7,00	7,50	8,00	8,50	9,00	9,50
1,00	1,000																	
1,50	1,000	1,000																
2,00	1,000	1,000	1,000															
2,50	0,952	0,952	0,957	1,000														
3,00	0,794	0,795	0,783	0,570	1,000													
3,50	0,870	0,869	0,879	0,979	0,391	1,000												
4,00	0,728	0,729	0,716	0,484	0,995	0,296	1,000											
4,50	0,780	0,778	0,790	0,934	0,238	0,987	0,138	1,000										
5,00	0,776	0,777	0,765	0,546	1,000	0,364	0,997	0,209	1,000									
5,50	0,519	0,518	0,534	0,756	- 0,108	0,873	- 0,208	0,940	- 0,137	1,000								
6,00	0,863	0,864	0,854	0,668	0,992	0,502	0,975	0,357	0,988	0,017	1,000							
6,50	0,735	0,736	0,723	0,493	0,996	0,306	1,000	0,148	0,998	- 0,198	0,977	1,000						
7,00	0,448	0,450	0,432	0,153	0,899	- 0,051	0,939	- 0,211	0,912	- 0,532	0,838	0,936	1,000					
7,50	0,641	0,642	0,627	0,376	0,976	0,180	0,993	0,019	0,982	- 0,323	0,941	0,992	0,973	1,000				
8,00	0,642	0,644	0,629	0,377	0,976	0,181	0,993	0,020	0,982	- 0,322	0,941	0,992	0,973	1,000	1,000			
8,50	0,521	0,522	0,506	0,235	0,933	0,032	0,964	- 0,129	0,943	- 0,459	0,881	0,962	0,997	0,989	0,989	1,000		
9,00	0,662	0,663	0,649	0,401	0,981	0,207	0,996	0,046	0,987	- 0,297	0,950	0,995	0,967	1,000	1,000	0,985	1,000	
9,50	0,686	0,688	0,673	0,431	0,987	0,239	0,998	0,079	0,991	- 0,265	0,960	0,998	0,958	0,998	0,998	0,978	1,000	1,000

Figure 8. Instantaneous Correlation Surface of Six-Month Libor Forward Rates Implied by the Calibration of the LMM (first case). The correlation matrix is based on Rebonato's angle formulation. The horizons of the six-month forward rates used to compute the correlation matrix range from 1 year to 9.5 years forward.



We now present a different scenario.

Second case

Because the instantaneous correlation wasn't satisfactory in the previous example, we choose to constrain the θ's to range in smaller intervals. We choose the following constraint:

$$-\frac{\pi}{8} < \theta_i - \theta_{i-1} < \frac{\pi}{8}$$

This implies that 'sub-adjacent' rates have positive correlations up to the 'fourth level'. All other inputs are maintained as before. Once the optimization process is performed we obtain the ψ's and θ's shown in table 9.

Table 9. Calibration results: second case: parameter values.

Index	Phi	Psi	Theta
1	0,1116	2,5834	0,1948
2	0,1454	1,6333	0,5038
3	0,1724	1,1076	0,4762
4	0,1924	1,8114	0,4897
5	0,2058	0,2315	0,7473
6	0,2141	2,3874	0,0468
7	0,2191	0,4012	0,4395
8	0,2226	1,5133	0,8322
9	0,2252	1,2685	1,2249
10	0,227	0	-0,3797
11	0,228	3,2337	0,013
12	0,2285	0	0,4057
13	0,2285	0,5089	0,7984
14	0,2283	1,1768	1,1911
15	0,228	0	1,1194
16	0,2276	0	0,9545
17	0,2274	0	0,8048
18	0,2274	0	0,7809

The overall fitting quality deteriorates from the previous case with errors usually higher and going up to 24.7%. Table 10 displays the matrix of percentage errors for swaption volatilities. Figures 9, 10 and 11 show the plot of the percentage errors and the difference in volatility surface between the market swaption volatilities and the fit of the calibration. Note that the longer maturities show the largest errors, on the downside.

Table 10. Percentage Difference between Market Swaption Volatilities and LMM Volatilities (second case).

The swaptions which are the most liquid on the Canadian market are highlighted in Blue.

Option Maturity	Swap maturity								
	1Y	1.5Y	2Y	2.5Y	3Y	3.5Y	4Y	4.5Y	5Y
1Y	0,13%	10,45%	9,43%	10,94%	10,54%	10,31%	11,46%	14,15%	13,47%
1.5Y	2,58%	1,17%	1,77%	1,02%	1,86%	4,20%	7,70%	7,49%	6,28%
2Y	-3,20%	-2,26%	-4,78%	-4,54%	-1,58%	3,03%	3,24%	2,37%	2,11%
2.5Y	-3,27%	-6,89%	-7,97%	-5,26%	0,69%	1,13%	0,19%	0,25%	1,88%
3Y	-12,12%	-10,42%	-7,12%	-1,60%	-1,00%	-1,68%	-1,46%	0,90%	2,24%
3.5Y	-16,84%	-14,17%	-4,41%	-3,66%	-4,00%	-3,52%	-1,40%	-0,22%	-0,04%
4Y	-20,91%	-1,60%	-1,97%	-4,05%	-3,93%	-1,48%	-0,23%	-0,07%	-0,44%
4.5Y	7,77%	0,24%	-4,14%	-4,08%	-0,80%	0,63%	0,76%	0,49%	0,03%
5Y	-24,70%	-20,38%	-14,83%	-7,14%	-3,09%	-2,02%	-2,05%	-1,98%	-2,05%

Figure 9. Plot of the percentage error in the swaptions calibration (second case).

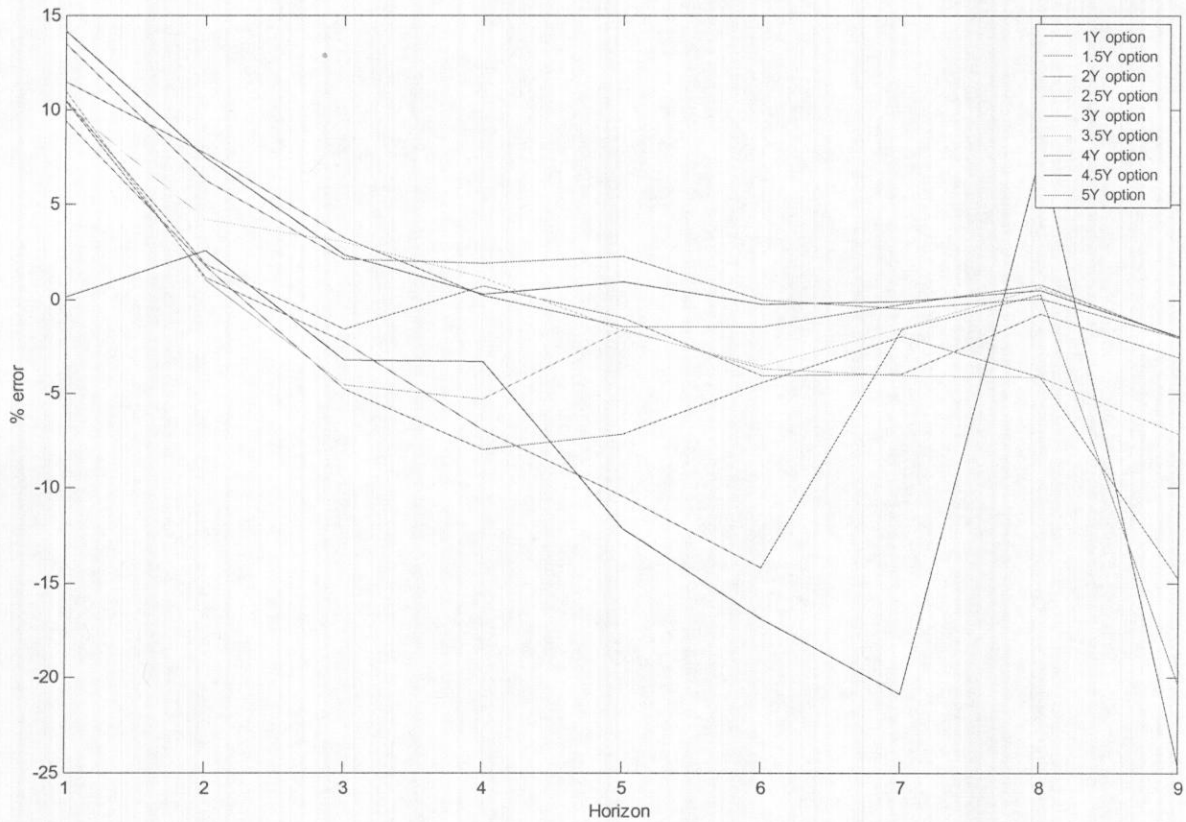
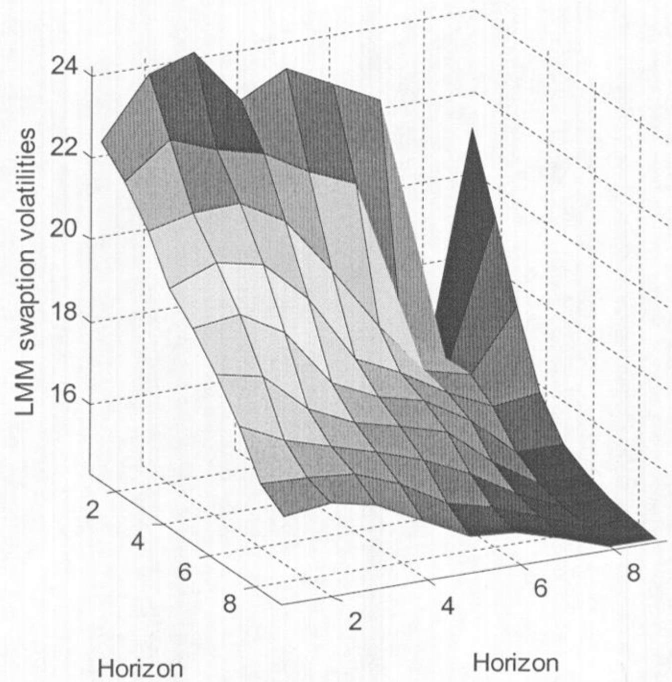


Figure 10. Comparison of the market swaption volatilities and the swaption calibration (second case).



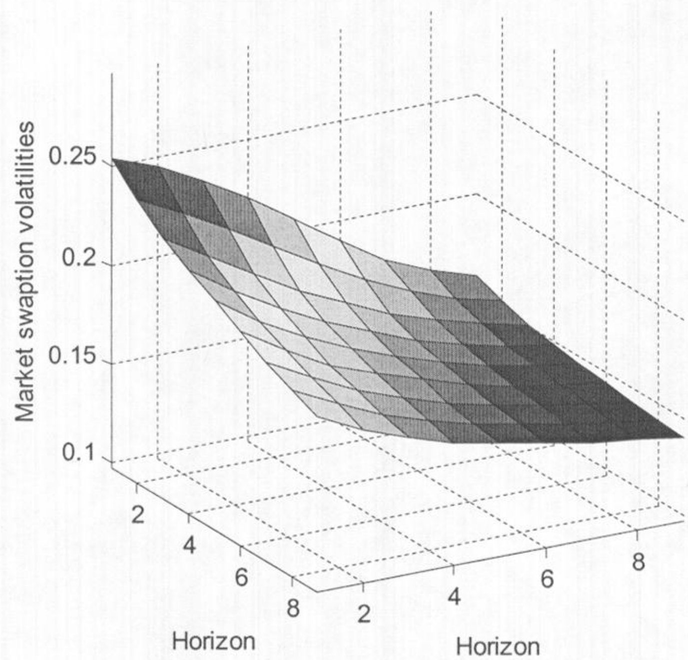
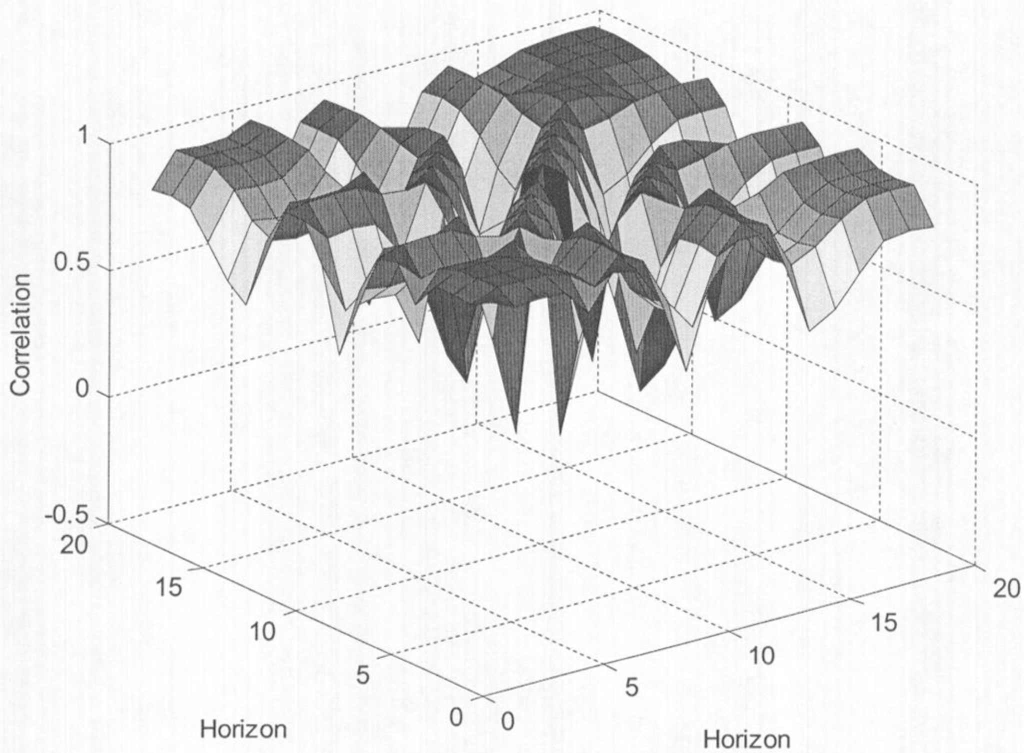


Figure 11. Instantaneous Correlation Surface of Six-Month Libor Forward Rates Implied by the Calibration of the LMM (second case). The correlation matrix is based on Rebonato's angle formulation. The horizons of the six-month forward rates used to compute the correlation matrix range from 1 year to 9.5 years forward.



The instantaneous correlations displayed in Table 11 and Figure 11 look better than in the previous case as almost all correlations are positive.

Table 11. Calibration results: Instantaneous correlation matrix (second case).

	1,00	1,50	2,00	2,50	3,00	3,50	4,00	4,50	5,00	5,50	6,00	6,50	7,00	7,50	8,00	8,50	9,00	9,50
1,00	1.0000																	
1,50	0.9526	1.0000																
2,00	0.9607	0.9996	1.0000															
2,50	0.9568	0.9999	0.9999	1.0000														
3,00	0.8512	0.9705	0.9635	0.9670	1.0000													
3,50	0.9891	0.8974	0.9092	0.9035	0.7645	1.0000												
4,00	0.9702	0.9979	0.9993	0.9987	0.9530	0.9239	1.0000											
4,50	0.8037	0.9466	0.9373	0.9419	0.9964	0.7071	0.9239	1.0000										
5,00	0.5148	0.7511	0.7326	0.7417	0.8881	0.3827	0.7071	0.9239	1.0000									
5,50	0.8395	0.6345	0.6555	0.6453	0.4294	0.9104	0.6828	0.3513	-0.0337	1.0000								
6,00	0.9835	0.8820	0.8946	0.8885	0.7423	0.9994	0.9104	0.6828	0.3513	0.9239	1.0000							
6,50	0.9778	0.9952	0.9975	0.9965	0.9422	0.9363	0.9994	0.9104	0.6828	0.7071	0.9239	1.0000						
7,00	0.8233	0.9569	0.9486	0.9527	0.9987	0.7306	0.9363	0.9994	0.9104	0.3827	0.7071	0.9239	1.0000					
7,50	0.5434	0.7729	0.7552	0.7639	0.9031	0.4136	0.7306	0.9363	0.9994	-0.0000	0.3827	0.7071	0.9239	1.0000				
8,00	0.6021	0.8164	0.8002	0.8082	0.9315	0.4778	0.7776	0.9590	0.9944	0.0716	0.4479	0.7559	0.9489	0.9974	1.0000			
8,50	0.7251	0.9001	0.8878	0.8939	0.9786	0.6155	0.8703	0.9925	0.9637	0.2344	0.5886	0.8532	0.9878	0.9721	0.9864	1.0000		
9,00	0.8197	0.9550	0.9465	0.9508	0.9983	0.7262	0.9340	0.9996	0.9131	0.3768	0.7026	0.9214	1.0000	0.9263	0.9509	0.9888	1.0000	
9,50	0.8331	0.9618	0.9539	0.9579	0.9994	0.7424	0.9423	0.9987	0.9031	0.3988	0.7193	0.9304	0.9998	0.9171	0.9433	0.9850	0.9997	1.0000

What this shows us is that the modeller has to make a decision between the goodness of fit and the behaviour of instantaneous correlations. Rebonato (1998) argues that because the markets are incomplete, it is better to seek positive correlations and have errors in the root mean square error.

Therefore, it is preferable to have a good fit to the instantaneous correlation matrix than to the swaption matrix.

6.3. Calibration issues and conclusion

In our implementation of the LMM to Canadian cap and swaption data, we used the rate curve, caplet volatilities and market quoted swaption prices (volatilities) as inputs. Then we hypothesized that forward rates are lognormal and stripped caplet volatilities from cap volatilities using (41).
Finally, we made modeling choices on the correlation structure (67), on the instantaneous volatility structure (60) and on the swaption approximation formula choice (72).

We showed that the LMM fits well to the Canadian data. The results we obtain are satisfactory in terms of the fit offered by the approximation formula (72) to swaption volatilities but it is still insufficient to price swaptions on the whole matrix. It is not possible to obtain a perfect fit with all the parameters we used (θ, Φ ,a,b,c and d). This is perhaps because the model is still incomplete in that it does not offer a perfect vision of the 'real world'. Another reason might be that the quotations given by brokers on the Canadian market are not all up to date because some swaptions are more liquid than other and therefore some entries in the matrix do not reflect current market conditions. In order to trust the model one needs to price only the swaptions that corresponds to the time horizon of the product to price. For example we should price only coterminal swaptions for a given maturity instead of trying to fit the whole matrix. This would allow the model to provide a better fit while preserving an instantaneous correlation structure that makes sense.

On the choice of the correlation matrix to be used in the closed form formula (88), we saw that historical correlations are not the first choice because of noise in the matrix and that historical values contain less information on current market conditions than do implied correlations, this why we chose to perform the calibration using implied correlations. On the other hand, the use of the implied correlation showed to be unstable. For certain values, we saw that the minimisation algorithm converges towards certain local solutions (like the short term options in the second case). It is therefore important to make a compromise between the instantaneous correlations and the goodness of fit to the swaption matrix.

One of the key theoretical issues concerning the simultaneous calibration of the LMM is that the caps and swaptions are lognormal under the LMM exclusively but not mutually.
Because of the nature of swaptions, if a forward rate is lognormal under the forward neutral measure in a cap, this same rate can't be in a swaption. Luckily for us, Rebonato (2000) shows that the discrepancy in practice is negligible unless we want to price spread options or rigger swaps.

The reason why we couldn't obtain simultaneous fit on the swaption matrix and the instantaneous correlation matrix is possibly due to the lack of liquidity for certain swaptions on the Canadian market (see table 10) as compared to the European or US markets.

Conclusion

This study offers a review of the term structure models literature to arrive at the LMM model which permits to work using the black (1976) framework in a theoretically and practically consistent way. We show the calibration procedure of the LMM to the Canadian market in a step-by-step approach. Our conclusion is that the fit of the LMM to the cap and swaption volatilities in Canada is satisfactory but not perfect. This is perhaps due to the lack of liquidity for certain maturities.

In that respect, our analysis is interesting because the Canadian swaption and cap market is not as developed as in Europe or the US. We believe that no study on the Canadian market has been performed as far as the LMM is concerned. Additionally our study provides an update on the factors that drove the yield curve in the first half of 2005.

We look at examples of calibrating the LMM to market data in Canada by making certain choices. We can extend the study further by changing the functional forms for the instantaneous forward volatilities and correlations or make other assumptions on the constraints to the thetas. We can also employ statistical testing or econometric analysis in our analysis but the attempt is more to provide an overview of the applicability of the LMM to market data.

The smile effect can also be addressed but it seems that in most interest rate models it tends to be of a much smaller magnitude than in the FX or equity markets (Rebonato 2000). What we would like to do in the next study is to perform delta hedging and compare the performance of the LMM to other well known models.

Conclusion

Cette étude offre une revue de littérature des modèles de structure à terme pour arriver au LMM qui est cohérant avec Black (1976) au niveau théorique et pratique. Nous montrons la procédure de calibration du LMM au marché canadien en détaillant chaque étape. Notre conclusion est que la justesse du modèle aux données canadiennes est satisfaisante mais imparfaite. Cela peut être due au manque de liquidité pour certaines maturités.

Notre analyse est intéressante car le marché canadien des caps et des swaptions n'est pas aussi développé qu'en Europe ou aux États-Unis. Nous croyons qu'aucune étude sur le marché canadien n'a été réalisée avec le LMM. De plus, nous montrons les facteurs qui influencent le marché canadien aux cours des six-premiers mois de 2005.

Nous présentons certains exemples de calibration du LM aux données de marché en faisant certains choix. Nous pouvons étendre notre étude en modifiant les formes fonctionnelles pour les volatilités et corrélations instantanées des taux forwards et faire d'autres hypothèses pour les thêtas.

Le « smile effect » peut aussi être adressé mais il semble que pour les modèles sur titres à revenus fixes, son effet soit moins important que dans le marché des changes ou des actions (Rebonato 2000). Dans une prochaine étude, nous aimerions pratiquer la gestion des risques en utilisant le « delta hedging ».

REFERENCES

[1] Alexander, C. (2002). Common Correlation Structures for Calibrating the Libor model. ISMA discussion papers in finance 2002-18.

[2] Alexander, C., Lvov, D. (2003). Statistical properties of Forward Libor Rates. ISMA discussion papers in finance 2003-03.

[3] Black, F. (1976). The Pricing of Commodity Contracts. Journal of Financial Economics.

[4] Black, F., Scholes M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, vol. 81, pp. 637-654.

[5] Brace, A., Gatarek, D., Musiela, M. (1997). The Market Model of Interest Rate Dynamics. Mathematical Finance, vol. 7, no.2, pp. 127-155.

[6] Brenner, R.J., Jarrow, R.A. (1993). A Simple Formula for Options on Discount Bonds. Advances in Futures and Options Research, 6, pp.45-51.

[7] Brigo, D., Mercurio, F. (2001). Interest Rate Models Theory and Practice. Springer Finance.

[8] Brown, S., J., Dybvig, P., H. (1994). Term Structure of Real Interest Rates and the Cox, Ingersoll and Ross model. Journal of Financial Economics. Vol. 35 pp. 3-42.

[9] Brown, S., J., Dybvig, P., H. (1986). The Empirical Implications of the Cox, Ingersoll, Ross Theory of Term Structure of Interest Rates.

[10] Cox, J., Ingersoll, J. and Ross, S. (1985). A Theory of the Term Structure of Interest Rates. Econometrica.

[11] Grant, D., Vora, G. (1999). Implementing No-Arbitrage Term Structure of Interest Rate Models in Discrete Time when Interest Rates are Normally Distributed. Journal of Fixed Income, 8 4, pp. 85-98.

[12] Heath, D., Jarrow, R., Morton, A. (1992). Bond Pricing and the Term Structure of Interest Rates: a new methodology for contingent claims valuation. Econometrica, vol. 60, no.1, pp. 77-105.

[13] Hull, J., White, A. (1999). Forward Rate Volatilities, Swap Rate Volatilities, and the implementation of the Libor Market Model. Joseph L. Rotman school of management University of Toronto.

[14] Inui, K., Kijima, M. (1998). A Markovian framework in multi-factor Heath-Jarrow-Morton models. Journal of Financial and Quantitative Analysis, 33, pp. 423-440.

[15] Jamshidian, F., (1997). Libor and Swap Market Models and Measures. Finance and Stochastics.

[16] Jarrow, R., A., Turnbull, S., M. (2000). Derivative Securities. South Western College Publications.

[17] Kagraoka, Y. (2000). Comparaison of the HJM and the BGM models: application to the Japanese market. MTEC Jounrnal.

[18] Longstaff, F., Shwartz, E. (1992). Interest Rate Volatility and the Term Structure : A two-factor General Equitlibrium Model. Journal of Finance.

[19] Lim, M. (2004). Calibration d'un modèle BGM à trois facteurs. Ecole polytechnique.

[20] Mercurio, F., Moraleda, J.M. (1996). An analytically tractable interest rate model with humped volatility. European journal of operational research 120, pp. 205-214. Elsevier.

[21] Molgedey, L. (2002) Calibration of the deterministic and stochastic volatility Libor Market Model. Dfine, Frankfurt MathFinance workshop.

[22] Rebonato, R. (2000). "On the simultaneous calibration of multi-factor log-normal interest-rate models to Black volatilities and to the correlation matrix". Working paper.

[23] Rebonato, R. (1999). On the simultaneous calibration of multi-factor log-normal interest-rate models to Black volatilities and to the correlation matrix. Working paper.

[24] Rebonato, R. (1998). Interest Rate Models. Second Edition, Whiley.

[25] Ritchken, P., Sankarasubramanian, L. (1995). Volatility Structures of Forward Rates and the Dynamics of the Term Structure. Mathematical Finance, 5, no. 1, pp. 55-72.

[26] Schoenmakers, J.G.M. (2002). Calibration of libor models to caps and swaptions: a way around intrinsic instabilities via parsimonious structures and a collateral market criterion. Preprint no. 740, Weierstrass Institute Berlin.

[27] Schoenmakers, J.G.M., Coffey, B. (2000). Stable implied calibration of a multi-factor Libor model via a semi-parametric correlation structure.

[28] Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. Journal of Financial Economics.

[29] Weber, T. (2005). Efficient calibration for libor market models: alternative strategies and implementation issues. Frankfurt MathFinance workshop Frankfurt am Main, 14./15.4.2005.

APPENDIX

1. Notes on the short rate

Let's outline a few specifications about the short rate. The short rate is the instantaneous interest rate and is believed to be the main factor governing the interest rate term structure dynamics.

The short rate dynamic is explained through the following SDE:

$$dr(t) = \mu(r)dt + \sigma(r)dW_t \tag{1}$$

This equation shows that the short rate dynamic is composed of two elements. The first is the drift over the time period (t, t + dt) which is $\mu(r)$dt and the second is a random shock represented by an increment of a Weiner process (otherwise known as a Brownian motion) dW(t) multiplied by the instantaneous volatility $\sigma(r)$.

Once the short rate dynamic is known, one can use it to determine the price of a traded asset which is the discount bond. The return on the bond is expressed through:

$$\frac{dP(t,T)}{P(t,T)} = \mu p(t,T)dt + \sigma p(t,T)\, dW_t \tag{2}$$

Where, by Ito's Lemma :

$$\mu p(t,T)\, P(t,T) = [\, (\frac{\partial P}{\partial t}) + (\frac{\partial P}{\partial r})\, \mu(r)\,] + (\frac{1}{2})(\frac{\partial^2 P}{\partial r^2})\, \sigma^2(r) \tag{3}$$

and

$$\sigma p(t,T)P(t,T) = (\frac{\partial P}{\partial r})\sigma(r) \tag{4}$$

Which shows how $\mu p(t,T)$ is directly related to the drift and volatility of the short rate and $\sigma p(t,T)$ to the volatility of the short rate.

Under the risk neutral measure Q, the short rate is expressed as:

$$dr = [\mu(r) - \lambda(r)\sigma(r)]dt + \sigma(r) d\,W_t^Q \qquad (5)$$

Where λ(r) is the market price of risk (i.e. expected excess return over the risk free-rate for bearing one unit of risk as measured by the volatility of returns $\sigma p(t,T)$).

It follows that the risk neutral process for the bond price is:

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt + \sigma p(t,T)\, d\,W_t^Q \qquad (6)$$

Under Q, all traded securities have an instantaneous expected return equal to the r(t) which is considered to be the risk free rate.

Under the risk neutral measure Q, the price of a discount bond is:

$$P(t,T) = E_t^Q\, [\, e^{-\int_{(t,T)} r(s)ds} \, . \, 1] \qquad (7)$$

This formula is quite simple given that, for the case of the discount bond, the payoff at maturity will be a dollar. The expected return is taken under the risk neutral probability measure. A more general formula is:

$$C(t,T) = E_t^Q\, [\, e^{-\int_t^T r(s)ds} \, . \, C(T,T)] \qquad (8)$$

Where C(t,T) is a random derivative product and C(T,T), it's payoff at maturity.

2. The Longstaff and Schwartz model (1992)

Two state variables, X and Y, represent the state of the economy (they both follow a square root process à la CIR).

$$dX = (a - bX) + c\sqrt{X}\, dW_{1t}$$
(9)

$$\text{dY}=(d-eY)+f\sqrt{Y}dW_{2t} \tag{10}$$

a, b, c, d, e and f are parameters of the model and dW_1 and dW_2 are Brownian motions (Weiner processes).

r(t) and the instantaneous variance have the following processes:

$$\text{dr}_t = \mu_r(r_t, v_t, t)dt + v_t(\mu_t, t)dW_{3t} \tag{11}$$

$$\text{d}v_t = \mu_v(r_t, v_t, t)dt + \sigma_v dW_{4t} \tag{12}$$

Longstaff and Schwartz make the following supposition

$$\text{r}_t = \alpha x_t + \beta y_t \text{ and } v_t = \alpha^2 x_t + \beta^2 y_t \tag{13}$$

After applying the Ito Lemma and manipulating a partial derivatives equation, they obtained a closed from formula for the price of zero-coupon bonds:

$$\text{P(t,T)} = A^{2\gamma}(\tau)B^{2\eta}(\tau)e^{(\kappa(\tau)+c(\tau)r_t+d(\tau)\upsilon_t)} \tag{14}$$

Where

$$\text{A}(\tau) = \frac{2q}{(\delta+q)\left[e^{q(\tau)}-1\right]+2q}$$

$$\text{B}(\tau) = \frac{2\psi}{(\gamma+\psi)\left[e^{\psi(\tau)}-1\right]+2\psi}$$

$$\text{C}(\tau) = \frac{2q\left[e^{\psi(\tau)}-1\right]B(\tau) - B\psi\left[e^{\psi(\tau)}-1\right]A(\tau)}{q\psi(B-2)}$$

$$\text{D}(\tau) = \frac{-q\left[e^{\psi(\tau)}-1\right]B(\tau) + \psi\left[e^{q(\tau)}-1\right]A(\tau)}{q\psi(B-2)}$$

$$\upsilon = \lambda + \xi$$

$$q = \sqrt{2\alpha + \delta^2}$$

$$\psi = \sqrt{2\beta + \gamma^2}$$

$$\kappa = \gamma(\delta+q) + \eta(\gamma+q)$$

Using (11), (12) and (13), it is possible to obtain a closed form formula for var(dP(t,T)) and for car(dR(0,T)).

3. The Extended Vasicek model

3.1. The Vasicek Model (1977) with a time-dependent drift

Under the risk neutral measure, the short rate evolves according to the following SDE:

$$drt = \kappa(\theta - rt)dt + \sigma d\,W_t^Q \tag{15}$$

Theta is the risk neutral mean and prices of fixed income derivatives depend only on the distribution of r_t under the risk neutral probability measure Q. It follows that the process under the original probability measure and the true drift are not essential here.

The issue with equation (15) is that it will not fit the time 0 yield curve perfectly. Therefore, it is necessary to make the SDE a risk neutral process with a time dependent mean:

$$drt = \kappa(\theta(t) - rt)dt + \sigma d\,W_t^Q \tag{16}$$

To which the solution is:

$$r_t = e^{-\kappa t} r_0 + \int_0^t e^{-\kappa(t-s)} \kappa\theta(s)ds + \sigma \int_0^t e^{-\kappa(t-s)} dW_s^Q \tag{17}$$

To ease things we denote m(t) and xt by,

$$m_t = e^{-\kappa t} r_0 + \int_0^t e^{-\kappa(t-s)} \kappa\theta(s)ds \tag{18}$$

and

$$dx_t = -\kappa x_t dt + \sigma dW^Q(t) \qquad \text{with } x_0=0 \tag{19}$$

Now we can write r_t like

$$r_t = m_t + x_t \tag{20}$$

(17) and (20) being the same equation, the latter will be used from now on for facilitating the calibration process purposes. From (20) it follows that with the absence of arbitrage, the price of zero-coupon bonds is given by the risk neutral expectation formula (7) and:

$$P(t,T) = e^{-\int_t^T m(s)ds} . E_t^Q [e^{-\int_t^T x(s)ds} .1] \tag{21}$$

$$= \exp\left(- \int_t^T m(s)\, ds \right) . \exp\left[A(T-t) + B(T-t) x_T \right]$$

The last equation is the same as in the original Vasicek (1977) with the exception that the unconditional mean = 0.

And (with (T-t) $= \tau$):

$$B(\tau) = \frac{e^{-\kappa\tau} - 1}{\kappa} \tag{22}$$

$$A(\tau) = \frac{1}{2}\sigma^2 \int_0^\tau B^2(s) ds \tag{23}$$

$$= \frac{1}{2}\left(\frac{\sigma}{\kappa}\right)^2 \left[\frac{1 - e^{-2\kappa\tau} - 4(1 - e^{-\kappa\tau})}{2\kappa} + \tau \right]$$

P (t,T) in (4.7) is written as the product of a deterministic factor and the bond price in a classic Vasicek setting with zero mean (under Q).

3.2. Calibration of the time dependent parameters

We have the discount function:

$$\text{d(T)} = \exp\left[-\int_0^T m(s)ds + A(T)\right] = \text{P(0, T)} \tag{24}$$

That represents the initial yield curve and we want to fit this initial yield curve. Since we stated x_0=0, if we take the log of (24):

$$\int_0^T m(s)ds = -\log d(T) + A(T) \tag{25}$$

Now if we differentiate with respect to T on both sides, we get the drift:

$$\text{m(T)} = \frac{-d\log d(T)}{dT} + \frac{dA(T)}{dT} = f(0,T) + \frac{1}{2}\sigma^2 B^2(T) \tag{26}$$

Now, the drift is directly related to the initial forward curve, f (0,T). It remains that the constants κ and σ have to be chosen ad-hoc prior to this calculation. One usually can obtain these two elements from market data (through interest rate caps for example or swaptions).

To obtain Theta (the risk neutral mean) lets look at the derivative of m_t:

$$m'_t = -\kappa e^{-\kappa t} r_0 + \kappa\theta_t - \kappa\int_0^t e^{-\kappa(t-s)}\kappa\theta_s ds \tag{27}$$

$$= \kappa\theta_t - \kappa m_t \tag{28}$$

And

$$\kappa\theta_t = \kappa m_t + m'_t \tag{29}$$

Using the definition of m_t in (26),

$$\kappa\theta_t = \kappa f(0,t) + \frac{1}{2}\kappa\sigma^2 B^2{}_t + \frac{\partial f(0,t)}{\partial t} + \sigma^2 B_t B'_t$$

$$= \kappa f(0,t) + \frac{\partial f(0,t)}{\partial t} + \phi_t$$

where

$$\phi_t = \frac{1}{2}\kappa\sigma^2 B^2{}_t + \sigma^2 B_t B'_t \tag{30}$$

$$= \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})$$

Now that we have $\kappa\theta_t$, from (16) we can write the SDE for r_t in the following manner:

$$dr_t = \left(\kappa(f(0,t) - r_t + \frac{\partial f(0,t)}{\partial t} + \phi_t\right)dt + \sigma dW_t^Q \tag{31}$$

This is an important result since it shows how the time-dependent parameters of the SDE are obtained from the forward curve $f(0,t)$.

3.3. Distribution of future bond prices

The purpose of calibrating the drift to the initial yield curve is to price fixed income derivatives at time t=0.

Lets consider a call option, maturing at time t, on a zero coupon bond maturing at time T, with strike K. The uncertain payoff is given by:

$$C(r_t) = \max[P(r_t, t, T) - K, 0] \tag{32}$$

The current price of this claim is given by (just like in 8, because of the no arbitrage condition).

$$V_0 = E_0^Q\left[e^{-\int_0^t r_s ds} C(r_t)\right]$$

From (21), the bond price at time t is given by:

$$P(t,T) = \exp\left(- \int_t^T m(s)\, ds \right) . \exp\left[A(\tau) + B(\tau)(r_t - m_t)\right] \tag{33}$$

m(s), $t \leq s \leq T$, is obtained from the calibration to the initial term structure.

"We are looking at the distribution of the future bond price given the current (t=0) information. Once we observe P(t,T), we can re-calibrate the function m(s) for $s \geq t$, and the new function will generally differ from the one obtained from f (0,t). This is the inherent inconsistency of the calibration approach. However we are only interested in prices of contingent claims at time t=0, which leads us to ignore the problem" (L. Cathcart, 1999).

After many lengthy calculations, we get:

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left[-\frac{1}{2} B^2(\tau)\phi_t + B(\tau)(r_t - f(0,t)) \right] \tag{34}$$

As one can tell, this equation only involves the current forward curve.

3.4. Calibration in other cases: Hull and White (1990)

Only the drift was time dependent in the above model, Sigma and Kappa where constant parameters. The natural extension to the model is therefore:

$$dr_t = \kappa_t(\theta_t - r_t)dt + \sigma_t dW_t^Q \tag{35}$$

And the equivalent generalization for the Cox, Ingersoll, Ross model

$$dr_t = \kappa_t(\theta_t - r_t)dt + \sigma_t \sqrt{r_t}\, dW_t^Q \tag{36}$$

The advantages of letting all the parameters time dependant is that the model can match the current volatility structure additionally to the yield curve.

Unfortunately, this feature makes the calibration of the time parameters complex. The relationship between θ_t and the initial forward curve in (29) is no longer that simple. Hull and White (1990) analyze these two models and discuss different approaches for calibration of time dependent parameters.

4. No-arbitrage pricing and numeraire change

The absence of arbitrage opportunities is the fundamental assumption in the Black and Scholes (1973) model to prices stock options. Absence of arbitrage means that it is impossible to invest zero today and to lock in a profit with a positive probability. It therefore means that two portfolios of securities displaying the same future payoff patterns must have the same price today.

4.1. Market, Portfolio and Arbitrage

We consider an economy in continuous time just like the one analyzed by Harrison and Kreps (1979) and Harrison and Pliska (1981, 1983). It is done so that the terminology set forth in this paper is introduced.

The time horizon is set from t=0 to T>0, there is a probability space and right continuous filtration. There are n+1 non dividend paying securities in this economy and they are traded continuously throughout the whole time span. The prices of those securities are modeled by a D+1 dimensional adapted semi-martingale whose components are positive.

There is an asset which is indexed by 0, it is a bank account. The price of this asset evolves according to $dB_t = r_t B_t \, dt$ with $B_0 = 1$ and where r_t is the instantaneous short rate.

A trading strategy is self financing if its value changes only due to the evolution of the asset prices. Another way of seeing it is that no additional cash inflows or outflows occur after the starting date.

An arbitrage opportunity is defined as a self financing strategy such that it's value at time 0=0 but at time T is >0. The existence of an equivalent risk neutral measure (or martingale measure or risk adjusted risk measure) implies the absence of arbitrage opportunities (HaPl, 81).

Now we assume there exists an equivalent martingale measure Q and let C be an attainable contingent claim, then for each time t, 0<t<T, there exists a unique price P associated with C:

$$P_t = E(\frac{1}{B_T} C), \text{ with the information set at time t.} \quad (37)$$

"When the set of all equivalent martingale measure is non empty, it is then possible to derive a unique no-arbitrage price associated to any attainable contingent claim. Such a price is given by the expectation of the discounted claim payoff under the measure Q equivalent to Q_0" (Brigo, Mercurio 2001). Note that for the price to be uniquely given, the market is arbitrage free but not necessarily complete.

The market is complete if and only if there exists a sole martingale measure. Therefore, the existence of a unique martingale measure makes the economy free of arbitrage opportunities and allows for the derivation of a unique price for any contingent claim.

4.2. The change of numeraire technique

Since the interest rates are stochastic, the presence, for example, of the stochastic discount factor $\frac{1}{B_T}$ renders quite complex the calculation of the expectation in (37). is is where a change of numeraire can be helpful as shown by Jamshidian (1989) in calculating bond-option prices under the Vasicek (1977) model.

A numeraire is a reference asset, with a self financing strategy, so as to normalize all other asset prices with respect to it. By choosing a numeraire Z, what we consider are the relative prices S^k/Z (with k=0,1,…,K), instead of the security S itself. An attainable claim is attainable under any numeraire (Geman et al. 1995). So far, we have only considered the bank account as possible numeraire but there can be more convenient ones for the calculation of claim prices.

Geman and al. (1995) proved that if we assume a numeraire N, a probability measure Q^N, equivalent to the initial Q_0, such that the price of any asset X relative to N is a martingale under Q^N,

$$\frac{X_t}{N_t} = E^N\left(\frac{X_T}{N_T}\right), 0 \le t \le T$$

(38)

then there exists another probability measure, Q^U for example, such that the price any attainable claim Y normalized to U is a martingale under Q^U.

$$\frac{Y_t}{U_t} = E^U\left(\frac{Y_T}{U_T}\right), 0 \le t \le T$$

4.3. Change of numeraire toolkit

The change of numeraire technique permits to derive asset-price dynamics under different numeraires.

Suppose that two numeraire S and U evolve under Q^U according to:

$$dS_t = (\ldots)dt + \sigma_t^S C dW_t^U, \quad Q^U$$

$$dU_t = (\ldots)dt + \sigma_t^U C dW_t^U, \quad Q^U$$

Where σ_t^S and σ_t^U are 1*n vectors, W^U is an n-dimensional driftless Brownian motion and CC'=instantaneous correlation matrix ρ.

Now consider an n-dimensional Itô process X whose dynamics is given respectively, under Q^S and Q^U, by:

$$dX_t = \mu_t^S(X_t)dt + \sigma_t(X_t)CdW_t^S, \quad Q^S$$

$$dX_t = \mu_t^U(X_t)dt + \sigma_t(X_t)CdW_t^U, \quad Q^U$$

Where μ_t^S and μ_t^U are an n*1 vector and σ_t is an n*n diagonal matrix.

Then the drift of the process X under the numeraire U is

$$\mu_t^U(X_t) = \mu_t^S(X_t) - \sigma_t(X_t)\rho\left(\frac{\sigma_t^S}{S_t} - \frac{\sigma_t^U}{U_t}\right)'$$

(39)

4.4. Note on the forward measure

In many situations (as we will see later), zero coupon bonds are a useful numeraire whose maturity T coincide with the one of the derivative to price. The forward measure is the measure associated with bonds that mature at time T are referred to as the T-forward risk-adjusted measure (Q^T). The related expectation is denoted by E^T.

Equivalent martingale measure: An EMM, Q, is a probability measure on the space (Ω, F) such that:

i. Q_0 (the objective measure) and Q are equivalent measures, that is Q_0(A)=0 if and only if Q(A)=0 for every A belonging to *F*.

ii. The Radon Nikodyn derivative[22] $\frac{dQ}{dQ_0}$ is square integrable with respect to Q_0.

iii. The discounted asset price process $\frac{1}{P(t,T)}$ is a right continuous filtration and EMM martingale.

Harrison and Pliska (1981) proved that the existence of an equivalent martingale measure implies the absence of arbitrage opportunities (which, in turn, implies a complete market if there is a unique EMM).

5. Probabilistic definitions and representing the flow of information

Define Ω as a **sample set**. f is a **partition** which is comprised of **events** $\omega 1, \omega 2, ..., \omega n$ (with n=1,N), where f and its events belongs to Ω. f is defined by (i). the intersection of any **subset** of f equals the **null set** and (ii). the union of all elements of f giving Ω). Define x, a **random variable** which associates with each event in Ω a number in $\Re$. x is said to be **measurable** if it associates the same values to elements of the same subsets of a partition. The **coarsest partition** on which a function x is measurable is said to be the partition generated by x (i.e. f(x)). An ordered sequence of partitions (the first being coarser than the second) is called an **information structure** or a **filtration**.

A sequence of random variables, $x_{(0)},...,x_{(n)}$, is said to be an **adapted process** to the information structure if each random variable $x_{(k)}$ is measurable on f(k). A sequence of random variables, $x_{(0)},...,x_{(n)}$, is a **predictable process** (on f(I)) if each $x_{(j)}$ is measurable on $f_{(j-1)}$. The price random variable is an adapted process to the information structure g_i (i=0,1,2) (where g_i is a sequence of partitions).

---

[22] $E^q\left[X|F_t\right] = \frac{E\left[X\frac{dQ}{dQ_0}\middle|F_t\right]}{\rho_t}$, where $\rho_t = \left.\frac{dQ}{dQ_0}\right|_{F_t}$

Two partitions f and g and **independent** if $P(f_i \cap g_i)=P(f_i)P(g_i)\ \forall i,j$. Two random variables are independent if the partitions generated by them (i.e. the coarsest partitions over which x and y are measurable) are independent.

For each element of a partition, we can define the **conditional expectation** as:

$$E[X|f_i]=\frac{\sum_{\omega\in f_i(k)} P(\omega)x(\omega)}{\sum P(\omega)}$$

Filtration is made up of algebras $\Im_0 \supset \Im_1 \supset \Im_2 \supset \Im_3$ meaning $\Im_3$ is included in $\Im_2$ etc…

$E[E[X|\Im_j]\ |\Im_i] = E[\ X\ |\ \Im_i\ ]$, i.e. the conditional expectation gives an unbiased estimator of X.

$E[X-E[X|\ \Im_i\ ] < \mathrm{var}\ [X-2]$: for any $\Im_i$ measurable random variable Z, the conditional expectation is the minimum variance estimator of X.

5.1. Brownian motions and random walks

For a unitary symmetric random walk:

$E[X_{(t+1)}\ |\ \Im_t\ ] = 0$

$\mathrm{Var}\ [X_{(t+1)}\ |\ \Im_t\ ] = 1$

$E\ [\ RW_{(s)}\ |\ \Im_i\ ] = RW_{(n)}$

$\mathrm{Var}\ [\ RW_n - RW_m\ |\ \Im_i\ ] = n-m$

In order to "scale" the process in such a way that the expectation and variance properties are retained, define:

$$B_m(t) = \frac{1}{\sqrt{m}} RW_k \qquad if\ m_t = k$$

$$B_m(t) = \frac{1}{\sqrt{m}} RW_k + 1 \qquad if\ mt=k+1$$

$$B_m(t) = RW_{(mt)}$$

by construction, so:

$$E\ [\ BM_{(tk+1)} - BM_{(tk)}\ ] = 0$$

$$\text{Var}\,[BM_{(tk+1)} - BM_{(tk)}] = t_{k+1} - t_k$$

We have shown that the Symmetric Random Walk is defined on any t and that it shares on "t_i", the properties of the unitary Symmetric Random Walk.

The resulting process is the Brownian motion $B_{(t)}$ which:

$B(0) = 0$

$E[\,B_{(s)} - B_{(t)} \mid \Im_{(t)}] = 0$

$\text{Var}\,[B_{(s)} - B_{(t)} \mid \Im_{(t)}] = s - t \qquad s \geq t$

$\Delta B_{(s,t)} = B_{(s)} - B_{(t)} \in N(0, s-t)$ i.e. the increments of a Brownian Motion are independent normal variables with zero expectation and variances = (s-t).

The last property holds since a binomial distribution "tends", for a finer and finer spacing, to the normal distribution.

$B_{(t)}$ is a continuous function of t! (to be differentiable, a function has to be continuous). But Brownian Motions (BM) aren't differentiable because:

$$\lim_{\zeta \to 0} SV_2(B)_{[0T]} = V_2(B)_{[0T]} = E[SV_2(B)] = T$$

$$\zeta = \max\{\Delta tk\}$$

$V_n = the\ n^{th}$ variation of B on [0,T]

An n-dimensional Brownian motion, $B_n(t) = \begin{bmatrix} B_0(t) \\ . \\ B_i(t) \end{bmatrix}$, where each element is the one dimensional BM we have studied so far. If E $[dB_j(t)dB_k(t)] = 0$, then the n-dimensional vector process $B_m(t)$ is an orthogonal n-dimensional BM.

5.2. Martingales and Ito Integrals

We have a probability space $(\Omega, \Im, P)$, with $\Im_n (0 \leq n \leq N)$. $\forall \omega \in \Omega$, $P(\omega) > 0$ and $\{Z_n\}$is a predictable process.

$\{Z_n\}$ is a martingale if $E[Z_{n+1} \mid \Im_n] = Z_n$, $\forall n \leq N-1$.

Consider an adapted (predictable) sequence of random variables $\{hn\}$, with respect to $\Im_n$. So

$$\Pi_0 = h_0 Z_0$$
$$\Pi_n = \Pi_0 + \sum_i \Delta\Pi_i$$

The process $\{\Pi_n\}$is called a martingale transform. Z_n could be prices and h_n, holdings in a portfolio. We want to express $\{\Pi_n\}$in terms of a Brownian Motion to use properties of the latter.

Z_n could be expressed as the martingale transform obtainable from a predictable sequence $\{kn\}$and a BM whish is a martingale.

$$Z_0 = k_0 B_0$$
$$Z_i = Z_0 + \sum_{j=1,i} k_j \Delta B_j$$
$$\Delta Z_i = k_i \Delta B_i$$

$$\Pi_n = \Pi_0 + \sum_i h_i k_i \Delta B_i = \Pi_0 + \sum_i \sigma_i \Delta B_i$$

$\Delta\Pi_n = \sigma_n \Delta B_n$ is an increment of the martingale transform.

We want to identify σ_n with the not yet defined volatility of the gain process Π_n.

$$I_{(t)} = \sum_{j=0,k-1} s(t_j)\left[B(t_{j+1}) - B(t_j)\right] + s(t_k)\left[B(t) - B(tk)\right]$$

This is the Ito integral of elementary process S(t) that is a function.

So:

$$I_{(t)} = \int_0^t s(u)dB(u)$$

We now want to relax the assumption that the integrand function should be an elementary process. So put σ that is adapted to $\Im_t$, $0 < t < T$, but which now depends continuously on time.

With the usual assumptions (c.f. Rebonato 98), we can define:

$$I(t) = \int_0^t \sigma(u)dB(u) = \lim_{n\to\infty} \int_0^t s_n(u)dB(u) \ , \ 0 \le t \le T \, .$$

It is the Ito Integral of the adapted process $\sigma(s)$.

We can then prove that an Ito integral is a martingale and that the Ito integral $I_{(t)} = \int_0^t \sigma(s)dB(s)$ has a quadratic variation given by :

$$V_2(I)_{[0,T]} = \int_0^t \sigma(s)^2 ds$$

The Ito integral has a variance given by:

$$VAR[I_t] = E[I_t^2] - E[I_t]^2 = E[I_t^2] = E\left[\int_0^t \sigma(s)^2 ds \right]$$

5.3. Ito's lemma and the rule of stochastic differentiation

Consider $C(t) = \int_0^t \mu(s)ds$ and $A_{(t)} = A_{(0)} + C_{(t)} + I_{(t)}$.

μ is a function of time so $A_{(t)} = A_{(0)} + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dB(s)$

The Stochastic Differential Equation (SDE) notation is $dA_{(t)} = \mu_{(t)}dt + \sigma_{(t)}dB_{(t)}$

In the limit, as $\Delta t \to 0$:

$$\Delta t^2 \Rightarrow 0$$
$$\Delta B \Delta t \Rightarrow 0$$
$$\Delta B \Delta B \Rightarrow \Delta t$$

When $\Delta t \to 0$, the SDE obtained by F is the Ito Lemma:

$$F_{(T)} = F_{((t)} + \int_t^T \left[\frac{\partial F}{\partial S} + \frac{\partial F}{\partial A}\mu(s) + \frac{1}{2}\frac{\partial^2 F}{\partial A^2}\sigma(s)^2 \right] ds + \int_0^T \frac{\partial F}{\partial A}\sigma(s)dB(s)$$

In the 2-dimensionnal case, $B_1(t)$ and $B_2(t)$ are independent components of an orthogonal two dimensional Brownian vector.