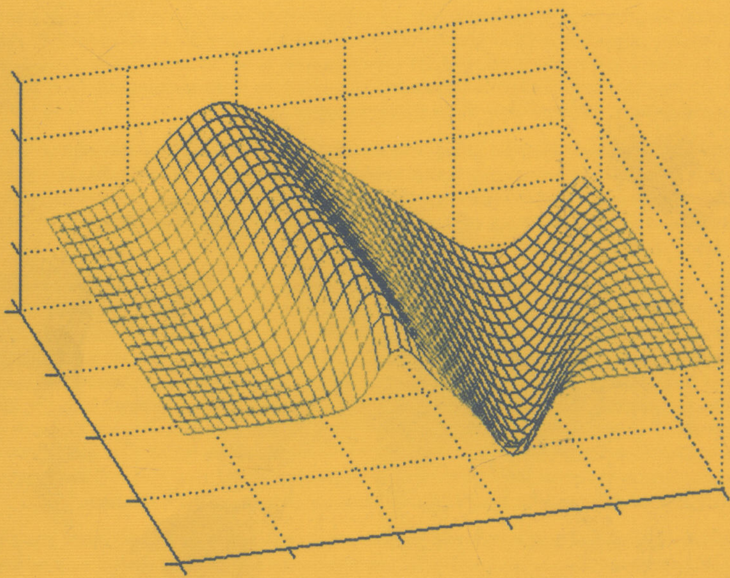
Lecture Notes in Mathematics

René A. Carmona
Ivar Ekeland
Arturo Kohatsu-Higa
Jean-Michel Lasry
Pierre-Louis Lions
Huyên Pham
Erik Taflin

Paris-Princeton Lectures on Mathematical Finance 2004

1919



Springer

René A. Carmona · Ivar Ekeland
Arturo Kohatsu-Higa · Jean-Michel Lasry
Pierre-Louis Lions · Huyên Pham
Erik Taflin

Paris-Princeton Lectures on Mathematical Finance 2004

Springer

Authors

René A. Carmona
Bendheim Center for Finance
Department of Operations Research
and Financial Engineering
Princeton University, Princeton
NJ 08544, USA
e-mail: rcarmona@princeton.edu

Ivar Ekeland
Canada Research Chair
in Mathematical Economics
University of British Columbia
Department of Mathematics
1984 Mathematics Road
V6T 1Z2 Canada
e-mail: ekeland@math.ubc.ca

Arturo Kohatsu-Higa
Graduate School of Engineering Sciences
Osaka University
Machikaneyama Cho 1–3
Osaka 560-8531, Japan
e-mail: kohatsu@sigmath.es.osaka-u.ac.jp

Jean-Michel Lasry
CALYON
9, Quai du Président Paul Doumer
92920 Paris-La Défense Cedex, France
e-mail: jeanmichel.lasry@ca-indosuez.com

Pierre-Louis Lions
Collège de France
11, place Marcelin Berthelot
75005 Paris, CX 05, France
e-mail: pierre-louis.lions@college-de-france.fr

Huyên Pham
Laboratoire de Probabilités
et Modèles Aléatoires
CNRS, UMR 7599
Université Paris 7 and Institut Universitaire
de France
Paris CX 05, France
e-mail: pham@math.jussieu.fr

Erik Taflin
Chair in Mathematical Finance
EISTI, Ecole Internationale des Sciences
du Traitement de l'Information
Avenue du Parc
95011 Cergy, France
e-mail: taflin@eisti.fr

[The addresses of the volume editors appear on page IX]

Library of Congress Control Number: 2007930225

Mathematics Subject Classification (2000): 91B28, 91B70, 60G49, 49J55, 60H07, 90C46

ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692
ISBN 978-3-540-73326-3 Springer Berlin Heidelberg New York
DOI 10.1007/978-3-540-73327-0

Springer is a part of Springer Science+Business Media
springer.com

Typesetting by the authors and SPi using a Springer LaTeX macro package

Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN: 12083206 41/SPi 5 4 3 2 1 0

Preface

This is the third volume of the Paris-Princeton Lectures in Mathematical Finance. The goal of this series is to publish cutting edge research in self-contained articles prepared by well known leaders in the field or promising young researchers invited by the editors. Particular attention is paid to the quality of the exposition, and the aim is at articles that can serve as an introductory reference for research in the field.

The series is a result of frequent exchanges between researchers in finance and financial mathematics in Paris and Princeton. Many of us felt that the field would benefit from timely exposés of topics in which there is important progress. René Carmona, Erhan Cinlar, Ivar Ekeland, Elyes Jouini, José Scheinkman and Nizar Touzi serve in the first editorial board of the Paris-Princeton Lectures in Financial Mathematics. Although many of the chapters involve lectures given in Paris or Princeton, we also invite other contributions. Given the current nature of the collaboration between the two poles, we expect to produce a volume per year. Springer Verlag kindly offered to host this enterprise under the umbrella of the Lecture Notes in Mathematics series, and we are thankful to Catriona Byrne for her encouragement and her help in the initial stage of the initiative.

This third volume contains five chapters. In the first chapter, René Carmona demonstrates how the HJM approach to the construction of dynamic models can be used for different financial markets. The original proposal of Heath, Jarrow and Morton was framed for the world of Treasury bonds, but its applicability was extended soon after its publication. However implementation of the same modeling philosophy in the case of credit and equity markets had to wait. Purely for pedagogical reasons, this chapter starts with a review of the original HJM approach to fixed income markets. Then, the recent works of Sidenius, Pitterbarg and Andersen and Schoenbucher on credit portfolio modeling are presented. Finally, the last part of the chapter explains how Carmona and Nadtochiy developed the program outlined a few years ago by Derman and Kani for equity markets.

The second chapter, by Ivar Ekeland and Erik Taflin, also develops the HJM framework. The emphasis here is on the optimal management of bond portfolios, that is, on Merton's problem of expected utility maximization. The authors introduce a special class of assets, the rollovers, which have a constant time to maturity, and describe the bond portfolios as a combination of rollovers rather than a combination of zero-coupon bonds, as is usual in the literature. The advantage of this approach is that a rollover does not mature, in contrast to a bond. By considering bond portfolios as a combination of rollovers, one brings them close to stock portfolios, which do not mature either, and one paves the way for a unified theory of money markets and equity markets. In addition, the authors derive explicit formulas for the optimal portfolios, at least in the case when the drift and volatility are deterministic processes. These advantages come at a price: a mathematical setting must be found, which will accommodate the curves describing the term structure of interest rates, and allow them to vary randomly, subject to possibly infinitely many sources of noise, while remaining sufficiently smooth. The Musiela parametrization (that is, taking time to maturity instead of date of maturity as the relevant variable), although it is natural in that context, considerably complicates matters, for it introduces an additional (and discontinuous) term in the equations of motion. Roughly speaking, that chapter complements the preceding one: taken together, they highlight the flexibility and generality of the HJM model, as well as its many unexplored consequences.

The third one is written by Arturo Kohatsu-Higa and is based on a short course given in Paris in November and December 2004. It reviews recent results on models for insider trading based on the theory of enlargement of filtrations. After reviewing the case of an insider having extra information given by the knowledge of the distribution of certain random variables, the author concentrates on the case when the insider benefits from almost sure additional information, leading naturally to the use of anticipative calculus. This review can be viewed as a natural companion to the chapter delivered by Fabrice Baudouin in the first volume of the series. The style of this chapter is purposedly pedagogical, emphasizing discrete time approximations as a way to illustrate the differences between anticipative and non-anticipative calculus, and exercises with solutions as a way to isolate proofs of technical results.

The fourth chapter is contributed to by Pierre Louis Lions and Jean Michel Lasry. It is concerned with the influence of hedging on the dynamics of the underlying asset price. The problem is completely solved in the case of a *large* investor. The results are based on a detailed analysis of liquidation and indifference prices studied via the solutions of non-standards stochastic control problems. This paper is more of a technical nature, and reads more like a research article than an introductory review. However, the importance of the issue and the originality of the mathematical approach fully justify its inclusion in the series.

The last chapter is concerned with applications of large deviations to problems in finance and insurance. It is contributed by Huyén Pham. Large deviation approximations and importance sampling methods are discussed in the context of option pricing. Credit risk losses and portfolio benchmarking are also analyzed with the tools of large deviations.

The Editors
Paris / Princeton
January 31, 2007

Editors

René A. Carmona
Paul M. Wythes '55 Professor of Engineering and Finance
ORFE and Bendheim Center for Finance
Princeton University
Princeton NJ 08540, USA
email: rcarmona@princeton.edu

Erhan Çinlar
Norman J. Sollenberger Professor of Engineering
ORFE and Bendheim Center for Finance
Princeton University
Princeton NJ 08540, USA
email: cinlar@princeton.edu

Ivar Ekeland
Canada Research Chair in Mathematical Economics
Department of Mathematics, Annex 1210
University of British Columbia
1984 Mathematics Road
Vancouver, B.C., Canada V6T 1Z2
email: ekeland@math.ubc.ca

Elyes Jouini
CEREMADE, UFR Mathématiques de la Décision
Université Paris-Dauphine
Place du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16, France
email: jouini@ceremade.dauphine.fr

José A. Scheinkman
Theodore Wells '29 Professor of Economics
Department of Economics and Bendheim Center for Finance
Princeton University
Princeton NJ 08540, USA
email: joses@princeton.edu

Nizar Touzi
Centre de Mathématiques Appliquées
Ecole Polytechnique Paris
F-91128 Palaiseau Cedex, FRANCE
email: nizar.touzi@polytechnique.edu

Contents

HJM: A Unified Approach to Dynamic Models for Fixed Income, Credit and Equity Markets
René A. Carmona 1

Optimal Bond Portfolios
Ivar Ekeland and Erik Taflin 51

Models for Insider Trading with Finite Utility
Arturo Kohatsu-Higa 103

Large Investor Trading Impacts on Volatility
Pierre-Louis Lions and Jean-Michel Lasry 173

Some Applications and Methods of Large Deviations in Finance and Insurance
Huyên Pham 191

HJM: A Unified Approach to Dynamic Models for Fixed Income, Credit and Equity Markets*

René A. Carmona

Bendheim Center for Finance
Department of Operations Research & Financial Engineering,
Princeton University, Princeton, NJ 08544, USA
email: rcarmona@princeton.edu

Summary. The purpose of this paper is to highlight some of the key elements of the HJM approach as originally introduced in the framework of fixed income market models, to explain how the very same philosophy was implemented in the case of credit portfolio derivatives and to show how it can be extended to and used in the case of equity market models. In each case we show how the HJM approach naturally yields a consistency condition and a no-arbitrage condition in the spirit of the original work of Heath, Jarrow and Morton. Even though the actual computations and the derivation of the drift condition in the case of equity models seems to be new, the paper is intended as a survey of existing results, and as such, it is mostly pedagogical in nature.

Keywords: Implied volatilty surface, Local Volatility surface, Market models, Arbitrage-free term structure dynamics, Health–Jarrow–Morton theory
Mathematics Subject Classification (2000) 91B24
JEL Classification (2000) G13

1 Introduction

The motivation for this paper can be found in the desire to understand recent attempts to implement the HJM philosophy in the valuation of options on credit portfolios. Several proposals appeared almost simultaneously in the literature on credit portfolio valuation. They were written independently by N. Bennani [3], J. Sidenius, V. Piterbarg and L. Andersen [26] and P. Shönbucher [41], the latter being most influential in the preparation of the present survey. After a sharp increase in volume and liquidity due to the coming of age of the single tranche synthetic CDOs, markets for these credit portfolios came to a stand still due to the lack of dynamic models needed to price forward starting contracts, options on options, ... So the need

* This research was partially supported by NSF DMS-0456195.

for dynamic models prompted these authors to build analogies between the original HJM approach to interest rate derivatives and derivatives on credit portfolio losses. The common starting point of these three papers is the lithany of well documented shortcomings of the market standard for the valuation of Collaterized Debt Obligations (CDOs). The standard Gaussian copula model is intrinsically a *one period* static model which cannot be used to price forward starting contracts. The valuation by expectation of these forward starting contracts require the analysis of a term structure of forward loss probabilities. The HJM modeling of the dynamics of the forward instantaneous interest rates, suggests how to choose dynamic models for these forward loss probabilities. The three papers mentioned above try to take advantage of this analogy with various degrees of generality and success.

The goal of this paper is to review the salient features of the HJM modeling philosophy as they can be applied to three different markets: the fixed income markets originally considered by Heath, Jarrow and Morton, the credit markets and the equity markets. In each of the three cases considered in this paper, the financial market model is based on a set of financial securities which are assumed to be liquidly traded. A basic assumption is that the price of each such security is observable, and any quantity of the security can be sold or bought at this observed price. These prices are used to encapsulate what the market is telling the modeler, and the thrust of the HJM modeling approach is to postulate dynamical equations for the prices of all these liquid instruments and to check that the multitude of all these equations do not introduce inconsistencies and arbitrage opportunities in the market model.

The classical HJM approach is reviewed in Section 3. Our informal presentation does not do justice to the depth of the original contribution [14] of Heath, Jarrow and Morton. It is meant as a light introduction to the modeling philosophy, our main goal being to introduce notation which are used throughout the paper, and to emphasize the crucial steps which will recur in the discussion of the other market models. Section 5 is devoted to the discussion of the recent works [26] of Sidenius, Pitterbarg and Andersen and [41] Schoenbucher on the construction of dynamic models for credit portfolios in the spirit of the HJM approach. These two papers are at the root of our renewed interest in the HJM modeling philosophy. It is while reading them that we realized the impact they could have on the classical equity models. The latter are usually calibrated to market prices by constructing an implied volatility surface, or equivalently a local volatility surface as advocated by Dupire and Derman and Kani in a series of influential works [19][16]. As we explain in Section 6, the construction of these surfaces is only the first step in the construction of a dynamic model. A dynamic version of local volatility modeling was touted by Derman and Kani in a paper [17] mostly known for its discussion of implied tree models. Motivated by the fact that the technical parts of [17] dealing with continuous models are rather informal and lacking mathematical proofs, Carmona and Nadtochiy developed in [7] the program outlined in [17]. On the top of providing a rigorous mathematical derivation of the so-called drift condition, they also provide calibration and Monte Carlo implementation recipes, and they analyze the classical Markovian spot models as well as stochastic volatility models in a generalized HJM framework. We present their results in the last section of this paper.

Acknowledgements. I would like to thank Dario Villani and Kharen Musaelian for introducing me to the intricacies of the credit markets. Their insights were invaluable: what they taught me cannot be found in textbooks !!!

2 General Mathematical Framework

This section is very abstract in nature. Its goal is to set the notation and the stage for the discussion of a common approach to three different markets.

2.1 Mathematical Notation

Throughout this paper we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\{\mathcal{F}_t\}_{t\geq 0}$ is a right continuous filtration of sub-σ-fields of $\mathcal{F}$, $\mathcal{F}_0$ containing all the null sets of $\mathbb{P}$. Most often, we assume that this filtration is a Brownian filtration in the sense that it is generated by a Wiener process $\{W_t\}_{t\geq 0}$. We allow this Wiener process to be multi-dimensional, and in fact, it can even be infinite dimensional. The facts from infinite dimensional stochastic analysis which are actually needed to prove the results discussed in this paper in the infinite dimensional setting can be found in many books and published articles. Most of them can be derived without using too much functional analysis. For the sake of my personal convenience, I chose to refer the interested reader to the book [9] for definitions and details about those infinite dimensional stochastic analysis results which we rely upon.

In order to compute cash flow *present values*, we use a discount factor which we denote by $\{\beta_t\}_{t\geq 0}$. The latter is a non-negative adapted stochastic process. Typically we use for β_t the inverse of the bank account B_t which is defined as the solution of the ordinary (possibly random) differential equation:

$$dB_t = r_t B_t \, dt, \qquad B_0 = 1, \tag{1}$$

where the stochastic process $\{r_t\}_{t\geq 0}$ has the interpretation of a short interest rate. In this case we have

$$\beta_t = e^{-\int_0^t r_s \, ds}. \tag{2}$$

Notice that $\{\beta_t\}_{t\geq 0}$ is multiplicative in the sense that

$$\beta_{s+t}(\omega) = \beta_s(\omega)\beta_t(\theta_s\omega), \qquad \omega \in \Omega,$$

where $\{\theta_t\}_{t\geq 0}$ is a semigroup of shift operators on Ω. For the sake of illustration, we should think of the ω's in Ω as functions of time, in which case $[\theta_s\omega](t) = w(s+t)$.

We shall assume that $\mathbb{P}$ is a pricing measure. This means that the market price at time $t = 0$ of any liquidly traded contingent claim which pays a random amount ξ at time T, say p_0, is given by (notice that the pay-off ξ is implicitly assumed to be a $\mathcal{F}_T$ integrable random variable):

$$p_0 = \mathbb{E}\{\beta_T \xi\}$$

where $\mathbb{E}\{\,\cdot\,\} = \mathbb{E}^{\mathbb{P}}\{\,\cdot\,\}$ denotes the expectation with respect to the probability measure $\mathbb{P}$. In other words, $\mathbb{P}$ is a pricing measure if prices of contingent claims are given by $\mathbb{P}$-expectations of present values of their future cashflows.

If we also assume that the market is free of arbitrage, then the price p_t at time $t<T$ of the same contingent claim is necessarily given by the conditional expectation

$$p_t = \frac{1}{\beta_t}\mathbb{E}\{\beta_T\xi|\mathcal{F}_t\}$$

which shows that $\{\beta_t p_t\}_{t\geq 0}$ is a $\mathbb{P}$-martingale in the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. In other words, if $\mathbb{P}$ is a pricing measure, the discounted prices are $\mathbb{P}$-martingales.

Notice that we do not assume that such a pricing measure is unique. In other words, we allow for incomplete market models in our discussion.

2.2 Liquidly Traded Instruments

We next assume that our economy is driven by a set of liquidly traded instruments whose prices at time t, we denote by P_t^α. We can think of the vector $\mathbf{P}_t = (P_t^\alpha)_{\alpha\in\mathcal{A}}$ of these observable prices as a state vector for our economy. We will not make the completeness assumption that

$$\mathcal{F}_t = \sigma\{\mathbf{P}_s;\ 0\leq s\leq t\}, \qquad t\geq 0.$$

These instruments are fundamental for the analysis of the market, and a minimal requirement on a dynamical model of the economy will be that such a model provides prices for forward starting contracts and European call and put options on these basic instruments. In particular, at each time t, we should be able to compute the quantity

$$\mathbb{E}\{\beta_T(P_T^\alpha - K)^+|\mathcal{F}_t\} \tag{3}$$

for every maturity $T>t$ and strike $K>0$. Since a measure μ on the half line $\mathbb{R}_+$ is entirely determined by the knowledge of its call transform, i.e. the values of the integrals

$$\int_{\mathbb{R}_+}(x-K)^+\ \mu(dx),$$

for $K>0$, the knowledge at time t, of the prices of all the call options completely determines the distributions under the conditional measure $\mathbb{P}_t$, of all the random variables P_T^α for all $T>t$ and all $\alpha\in\mathcal{A}$.

Here, for each $t>0$, we define the random measure $\mathbb{P}_t$ as the (regular version of the) conditional distribution given $\mathcal{F}_t$ of the discounted version of $\mathbb{P}$. In other words, $\mathbb{P}_t$ is characterized by the requirement that the equality

$$\mathbb{E}\{\beta_{t+T}\Phi\Psi\circ\theta_t\} = \mathbb{E}\{\beta_t\Phi\mathbb{E}^{\mathbb{P}_t}\{\beta_T\Psi\}\}$$

holds for all bounded random variables Φ and Ψ which are $\mathcal{F}_t$ and $\mathcal{F}_T$ measurable respectively.

Remark. Notice that if instead of simply requiring the knowledge of the prices of all the European call options we were to also require the knowledge of the prices of all the path dependent options, then for each $\alpha\in\mathcal{A}$, the entire (joint) distribution under $\mathbb{P}_t$ of $(P_T^\alpha)_{T\geq t}$ would be determined. In the situation of interest to us, only the one-dimensional marginal distributions of $\mathbb{P}_t$ are determined by the prices we can observe.

2.3 Dynamic Market Model

All the information about the market model should be contained in the specification of a pricing measure $\mathbb{P}$. However, as we explained earlier, it seems that a reasonable market model should

- *be consistent with the prices of the liquidly traded instruments quoted on the market*, in other words, the numerical values P_t^α observed on the market should be recovered as conditional expectations under the pricing measure $\mathbb{P}$ of the discounted cashflows of the corresponding instruments;
- *allow for the pricing of forward starting contracts (e.g. European call options on call options* using the identified liquidly traded instruments as underlyers. In other words, it should provide a way to compute the time evolution of the conditional (random) measures $\mathbb{P}_t$, or at least its marginal distributions.

The first bullet point involves simply reproducing the prices of the basic liquid instruments at time $t = 0$. It usually goes under the name of calibration. The restriction of the measure $\mathbb{P}$ to $\mathcal{F}_0$ is typically trivial and the computation of these prices involves only regular expectations with respect to $\mathbb{P}$ which can be computed at time $t = 0$. So this first bullet point does not seem to involve the dynamics of the stochastic evolution of the characteristics of the market model: it looks like a static requirement for a one period model.

On the other hand, the second bullet point involves information about the model (and hence the pricing measure $\mathbb{P}$) of a more dynamic nature. For this reason, if will appear to be preferable to specify this dynamic information about $\mathbb{P}$ by specifying $\{\mathbb{P}_t\}_{t\geq 0}$ as a stochastic process in the space of probability measures on the possible future time evolutions of the vectors $\{\mathbf{P}_{t+s}\}_{s\geq 0}$ of basic instruments. This is the main thrust of the HJM approach to fixed income market models as it was originally introduced by Heath, Jarrow and Morton, and this is the point of view we take to review in the remaining part of this paper, recent developments in modeling credit and equity markets.

3 The Classical HJM Approach

The goal of this section is purely of a pedagogical nature. It is not intended as a rigorous *exposé* of the original work of Heath, Jarrow and Morton: it is merely an informal discussion aimed at a very general audience. In the case of fixed income markets (also called interest rate derivatives markets), the simplest form of interest rate is the spot rate whose value at time t we denote by r_t. As we will emphasize in several instances, any market model needs to provide with the distribution of the stochastic process $\{r_t\}_{t\geq 0}$, even if its role is limited to the introduction of the bank account and the discount factor as in the previous section. Many market models have been based on the specification of the dynamics of this process. For this reason they are called *short rate models*. Despite the limitations which we are about to document, they remain very popular, mostly because of their versatility and the existence of closed form formulae for the prices of many liquidly traded instruments.

There are several sets of liquid interest rate derivatives actively traded and quoted daily. Coupon bearing bonds, caps, floors, swaptions, are some of them. But because most of them can be viewed as portfolios of zero coupon bonds, or European options on zero coupon bonds, and because this section aims at recasting classical material (which can be found in most financial mathematics textbooks) into the framework adopted in the paper, we find convenient to choose, for the set of liquidly traded securities, the ensemble of all the zero coupon non-defaultable bonds.

For the sake of definiteness, we denote by $B(t,T)$ the price at time t of such a zero coupon bond with maturity T. We shall often use the term "Treasury" (which essentially means that the bond will not default) interchangeably with "non-defaultable". The entire face value will be paid at time T by the issuer of the bond to the buyer as long as $T > t$. So at time $t = 0$, all the prices $B(0,T)$ can be observed and the entire curve

$$T \hookrightarrow B_0(T) = B(0,T) \tag{4}$$

is known. So as stated in the first bullet point of Subsection 2.3 above, a first requirement for a model given by a pricing measure $\mathbb{P}$ is to reproduce these prices exactly.

As we are about to see, this innocent looking condition cannot always be satisfied by the short interest rate models which need to be re-calibrated frequently to satisfy, at least approximatively this requirement. Indeed, short interest rate models are endogenous term structure models as the initial term structure of zero coupon bond prices (4) is an output of the model instead of being an input observed in the market place. This last point is one of the main components of the HJM approach.

Since the cash flows of a zero coupon bond reduce to paying its nominal amount (which we conveniently normalize to 1) at time T, the price has to be given by

$$B_0(T) = \mathbb{E}\{\beta_T\} = \mathbb{E}\{e^{-\int_0^T r_s ds}\}, \tag{5}$$

recall that $\beta_0 = 1$. So if the parameters of the pricing measure $\mathbb{P}$ allow for the computation of the expectation in the above right hand side, the value of this expectation will have to coincide with the observed price $B_0(T)$ if we want to satisfy the first bullet point above.

Using Instantaneous Forward Rates Instead. For reasons that will become clear later, if the zero coupon prices $B(t,T)$ are (or assumed to be) smooth in the maturity variable T, it is more convenient to work with the forward rates defined by

$$f(t,T) = -\frac{\partial}{\partial T} \log B(t,T) \tag{6}$$

rather than the bond prices directly. Since the bond prices can be recovered from the forward rates

$$B(t,T) = e^{-\int_t^T f(t,u)\,du} \tag{7}$$

the term structure of interest rates can be given equivalently by the forward curves. In particular, observing all the bond prices $B_0(T)$ at time $t = 0$ is equivalent to observing all the forward rates $f_0(T)$, and the initial forward rate curve

$$T \hookrightarrow f_0(T)$$

can be the object of the calibration efforts (in the case of short rate models) or it can serve as initial condition (in the case of HJM models).

3.1 Short Rate Models

Since the prices of the basic instruments of the market can be computed as expectations over the short interest rate, recall formula (5), the simplest prescription for a pricing measure $\mathbb{P}$ is to describe the dynamics of the short rate process. Typically, a short rate model assumes that under the pricing measure $\mathbb{P}$, the short interest rate r_t is the solution of a stochastic differential equation of the diffusion form (i.e. Markovian):

$$dr_t = \mu^{(r)}(t, r_t)\, dt + \sigma^{(r)}(t, r_t)\, dW_t \tag{8}$$

where the drift and volatility terms are given by real-valued (deterministic) functions

$$(t, r) \hookrightarrow \mu^{(r)}(t, r) \quad \text{and} \quad (t, r) \hookrightarrow \sigma^{(r)}(t, r)$$

such that existence and uniqueness of a strong solution hold. For the sake of illustration, we consider only one specific example. Indeed, the goal of this section is not to present the theory of short rate models. They are mentioned only as motivation for the introduction of the HJM modeling approach.

We choose the **Vasicek** model because of its simplicity, but for the purpose of the present discussion, a **CIR** model of the square root diffusion could have done as well. In the case of the Vasicek model, the dynamics of the short rate are given by the stochastic differential equation:

$$dr_t = (\alpha - \beta r_t)\, dt \, + \, \sigma\, dW_t. \tag{9}$$

This equation is simple enough (linear) to be solved explicitly. The solution is given by

$$r_t = e^{-\beta t} r_0 + (1 - e^{-\beta t})\frac{\alpha}{\beta} + \int_0^t e^{-\beta(t-s)} \sigma dW_s. \tag{10}$$

$\{r_t\}_{t \geq 0}$ is a Gaussian process whenever r_0 is, and at each time $t > 0$ there is a positive probability that r_t is negative. Despite this troubling feature (not only can an interest rate be negative in this model, but it is almost surely unbounded below!), this model is very popular because of its tractability and because a judicious choice of the parameters can make this probability of negative interest rate quite small. The tractability of the model is due to the fact that the random variable $\int_0^t r_s ds$ is Gaussian with mean and variance which can be explicitly computed from the parameters α, β and σ of the model, and from this fact, one gets an explicit formula for the expectation (5) giving the price of the zero coupon bonds. We get:

$$B_0(T) = e^{a(T)+b(T)r_0} \tag{11}$$

where r_0 is the current value of the short rate, and where the functions $a(T)$ and $b(T)$ are given by:

$$b(T) = -\frac{1}{\beta}\left(1 - e^{-\beta T}\right) \tag{12}$$

and

$$a(T) = \frac{4\alpha\beta - 3\sigma^2}{4\beta^3} + \frac{\sigma^2 - 2\alpha\beta}{2\beta^2}T + \frac{\sigma^2 - \alpha\beta}{\beta^3}e^{-\beta T} - \frac{\sigma^2}{4\beta^3}e^{-2\beta T}. \tag{13}$$

Alternatively, if we use the forward curve instead of the zero coupon bond curve we get:

$$f(t,T) = re^{-\beta(T-t)} + \frac{\alpha}{\beta}\left(1 - e^{-\beta(T-t)}\right) - \frac{\sigma^2}{2\beta^2}\left(1 - e^{-\beta(T-t)}\right)^2, \tag{14}$$

from which we get an expression for the initial forward curve $T \hookrightarrow f_0(T)$ by setting $t = 0$. Notice that such a forward curve converges to the constant $(2\alpha\beta - \sigma^2)/2\beta^2$ when $T \to \infty$. This limit can be given the interpretation of a *long rate* (as opposed to the short rate) when $\sigma^2 < 2\alpha\beta$. In any case, a Vasicek forward curve flattens and becomes horizontal for large maturity T. The graph of a typical example of a forward curve given by the Vasicek model is given in the left pane of Figure 1. We used the parameters $\alpha = 13.06$, $\beta = 2.5$ and $\sigma = 2$ to produce this plot. We clearly see the flattening of the curve on the right part of the plot.

Rigid Term Structures for Calibration

As we explained earlier, choosing values for the parameters of the model (α, β and σ in the Vasicek model discussed in this section) in order for the model to reproduce the observed forward curve is what is usually called calibration of the model. Since the Vasicek model depends upon three parameters, three quoted prices, say $B_0(T_1)$, $B_0(T_2)$ and $B_0(T_3)$ for three different maturities T_1, T_2 and T_3 should in principle be enough to determine these parameters. But unfortunately, the curve $T \hookrightarrow B_0(T)$ constructed from formulae (11), (12), and (13) and three parameter values derived from three bond prices does not always look like the curve produced by the market quotes, and most importantly, it changes with the choices of the three maturities T_1, T_2 and T_3. For the sake of illustration, we give in the right pane of Figure 1 the plot of the market zero-coupon forward curve on 3/28/1996, and we super-impose on the same graph the plot of the best least squares fit among the possible forward curves produced by the Vasicek model. This optimal Vasicek forward curve was obtained for the values $\alpha = 13.06$, $\beta = 2.401$ and $\sigma = 1.724$ of the parameters. The fact that a Vasicek forward curves flattens for large maturity makes it impossible to match the typical increase in T found in most practical instances.

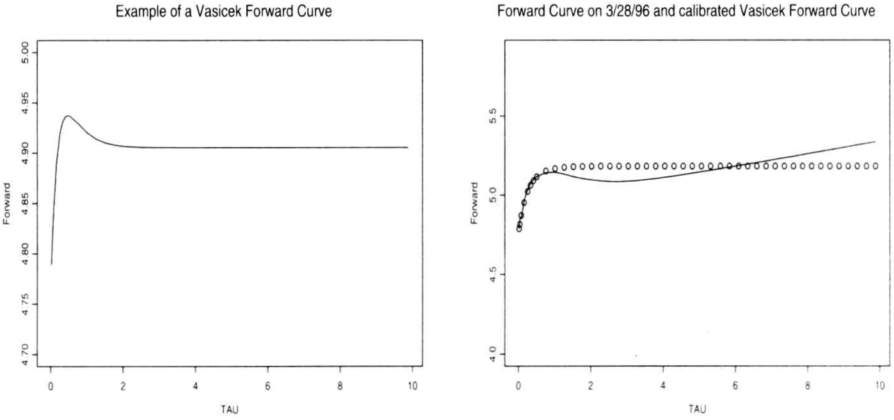


Fig. 1. Typical forward curve produced by the Vasicek model (left) and calibrated Vasicek forward curve (dotted line) to the zero-coupon forward curve on 3/28/1996

A Possible Fix

Several solutions have been proposed to the undesirable rigidity of the initial term structure curves produced by short rate models. The most popular one is to force some of the coefficients to be time dependent in order for the model to match any market forward curve $T \hookrightarrow f_0(T)$. This is especially simple and useful in the case of the Vasicek model for if the time dependent coefficients are deterministic, the solution process remains Gaussian, and closed form solutions for the values of the forward rates and zero coupon prices can still be derived. To be more specific, formula (10) becomes

$$r_t = e^{-\int_0^t \beta_s ds} r_0 + \int_0^t e^{-\int_s^t \beta_u du} \alpha_s ds + \int_0^t e^{-\int_s^t \beta_u du} \sigma_s dW_s. \quad (15)$$

and since the conditional distribution of the integral $\int_s^t f_u du$ is Gaussian, bond prices

$$B(t,T) = \mathbb{E}\{e^{-\int_t^T r_s ds} | \mathcal{F}_t\}$$

can still be derived from the expression of the Laplace transform of the Gaussian distribution.

This strategy was successfully implemented in the case of the Vasicek model (9) by Hull and White. These two authors proposed to leave the volatility σ and the mean reversion rate β constant, and to replace the parameter α by a deterministic function $t \hookrightarrow \alpha(t)$. In this case, the solution r_t is given by the formula

$$r_t = e^{-\beta t} r_0 + \int_0^t e^{-\beta(t-s)} \alpha_s ds + \sigma \int_0^t e^{-\beta(t-s)} dW_s, \quad (16)$$