Exponential-Affine Diffusion Term Structure Models: Dimension, Time-Homogeneity, and Stochastic Volatility

by

João Pedro Vidal Nunes

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in Industrial and Business Studies

University of Warwick,

WARWICK
BUSINESS SCHOOL

January 2000

Contents

List of Tables

List of Figures

Acknowledgements

My first acknowledgement goes to my teacher António Gomes Mota who first motivated me to the Finance field and then "obliged" me to pursue a Ph.D. in this area of knowledge. I would also like to thank my supervisors, Dr. Les Clewlow and Professor Stewart Hodges, for their guidance, support and patience. A special word is devoted to my colleges Miguel Ferreira and George Skiadopoulos for all stimulating discussions.

This dissertation could not have been done without the constant incentive and help of my parents and sister. And, I must, most of all, acknowledge my wife, Carla, for her comprehension and for all personal and professional sacrifices that provided me the required time for this work.

This thesis is dedicated to my best and sweetest friend: my grandmother, Etelvina.

Exponential-Affine Diffusion Term Structure Models: Dimension, Time-Homogeneity, and Stochastic Volatility

by

João Pedro Vidal Nunes

Abstract

The object of study in this thesis is the most general affine term structure model characterized by Duffie and Kan (1996), which nests, as special cases, many of the interest rate models previously formulated in the literature. The purpose of the dissertation is two-fold: to derive fast and accurate pricing solutions for the general term structure framework under analysis, which enable the effective use of model' specifications yet unexplored due to their analytical intractability; and, to implement a simple and robust model' estimation methodology that enhances the model' fit to the market interest rates covariance surface.

Concerning the first (theoretical) goal, analytical exact pricing solutions, for several interest rate derivatives, are first derived under a (simpler and) nested Gaussian affine specification. Then, and as the main contribution of the present dissertation, such Gaussian formulae are transformed into first order approximate closed-form pricing solutions for the most general stochastic volatility model' formulation. These approximate solutions are shown to be both extremely fast to implement and accurate, which make them an effective alternative to the existing numerical pricing methods available.

Related to second thesis' (empirical) goal, and in order to enable the model' estimation from a panel-data of interest rate contingent claims' prices, a general equilibrium model' specification is derived under non-severe preferences' assumptions and in the context of a monetary economy. The corresponding state-space model' specification is estimated through a non-linear Kalman filter and using a panel-data of not only swap rates (as it is usual in the Finance literature) but also (for the first time) of caps and European swaptions prices. It is shown that although the model' fit to the level of the yield curve is extremely good, short-term caps and swaptions are systematically mispriced. Finally, a time-inhomogeneous HJM formulation is proposed, and the model' fit to the market interest rates covariance matrix is substantially improved.

Chapter 1

Introduction

1.1 Scope, background and motivation

The present research is devoted to the analysis of *multifactor exponential-affine diffusion (single-currency and default-free)* models for the term structure of interest rates, in terms of model' specification, concerning the methodology for model' estimation, and with the purpose of deriving better pricing solutions for several (heavily traded) interest rate contingent claims.

Stochastic term structure models are essential to price and hedge interest rate derivatives in a consistent (i.e. arbitrage-free) and aggregate manner. Although the required features for an "ideal" interest rate model are easy to define (and can be summarized as *economic realism*, i.e. foundation on theoretically sound assumptions, and *analytical tractability*, that is ability to price and hedge derivatives in "real market" time), different term structure models have been proposed in the literature. The majority of these models belongs to the *diffusion* (parametric) class, where the stochastic behavior of the model arises from a finite or infinite number of Brownian motion shocks.[1] Moreover, the wide class of diffusion term structure models can be divided into four categories, attending to the number of model' state variables under use and considering the time-homogeneity of the model' parameters. In terms of model' dimension, there exist *single-factor* (usually short-term interest rate) models and *multifactor* models (i.e. specifications with more than one state variable). Concerning the time-dependency of the model' parameters, there coexist *time-homogeneous* (or "equilibrium") models and *evolutionary* (or "no-arbitrage") models.[2] The first group

[1] The relative less significant use of *jump processes* for modelling the term structure of interest rates is, perhaps, just a consequence of the numerical elegance and simplicity provided by the standard use of continuous-time stochastic calculus in the field of Finance. On the other hand, the assumption of parametric stochastic diffusion equations for the model' factors also simplifies the model' estimation and provides, in some cases, analytical or quasi-analytical pricing solutions. Concerning the estimation of non-parametric diffusion models, see, for instance, Aït-Sahalia (1996a).

[2] The use of both "equilibrium" and "no-arbitrage" denominations can be misleading, for two reasons:

Table 1.1: Classifications for parametric diffusion term structure models

Dimension	Time-homogeneous models	Evolutionary models
Single-factor models	Merton (1970)*	Ho and Lee (1986)*
	Vasicek (1977)*	Hull and White (1990)*
	Dothan (1978)	Black, Derman and Toy (1990)
	Brennan and Schwartz (1979)	Black and Karasinski (1991)
	Constantinides and Ingersoll (1984)	
	Cox, Ingersoll and Ross (1985b)*	
	Pearson and Sun (1993)*	
Multifactor models		
2 factors	Richard (1978)*	Hull and White (1994)*
	Schaefer and Schwartz (1984)*	
	Fong and Vasicek (1991b)*	
	Chen and Scott (1992)*	
	Longstaff and Schwartz (1992a)*	
3 factors	Balduzzi et al. (1996)*	
	Chen (1996)	
n factors ($n \in \mathbb{N}$)	Langetieg (1980)*	Frachot and Lesne (1993)*
	Chen and Scott (1995b)*	
	Duffie and Kan (1996)*	
∞ factors		Heath, Jarrow and Morton (1992)
		Brace, Gatarek and Musiela (1997)
		Kennedy (1991)

*Exponential-affine interest rate models

of models is intrinsically unable to reproduce exactly the term structure of interest rates observed in the market. The second group of models takes the observed yield curve as exogenous and provides an exact fit to some market observables (such as swap rates). Table 1.1 summarizes some of the most well known diffusion (parametric) term structure models.

The single-factor models take the instantaneous default-free interest rate, $r(t)$, as the only state variable that explains the evolution of the whole yield curve. Such simplistic hypothesis offers analytical tractability, but often at the expense of unrealistic model' features.[3] Table 1.2 reproduces Duffie and Kan (1994, Table 1) and nests all one-factor models under a common framework: for the time-homogeneous models the parameters are time-independent, while for the evolutionary models the parameters are functions of time. In terms of the drift specification, most models (with $\alpha_1(t), \alpha_2(t) \neq 0$ as well as both Black et al. (1990) and Black and Karasinski (1991) models, for some volatility specifications) incorporate a *mean reversion* phenomenon (based on economic business cycles): interest rates tend to revert, over time, towards some long-run mean level. Although some empirical

firstly, because not all *time-homogeneous* models were derived from a general equilibrium formulation, and, secondly, because all *time-homogeneous* models rely upon a no-arbitrage assumption. Therefore, such alternative labels will be avoided hereafter.

[3] For instance, an intrinsic limitation of any one-factor model consists in assuming a perfect correlation amongst all interest rates (of different maturities).

Table 1.2: Single-factor models

$dr(t) = [\alpha_1(t) + \alpha_2(t) r(t) + \alpha_3(t) r(t) \ln r(t)] dt + [\beta_1(t) + \beta_2(t) r(t)]^\gamma dW(t)$						
Parameters' restrictions:	$\alpha_1(t)$	$\alpha_2(t)$	$\alpha_3(t)$	$\beta_1(t)$	$\beta_2(t)$	γ
Merton (1970)		0	0		0	1.0
Vasicek (1977)			0		0	1.0
Dothan (1978)	0	0	0	0		1.0
Brennan and Schwartz (1979)			0	0		1.0
Constantinides and Ingersoll (1984)	0	0	0	0		1.5
Cox et al. (1985b)			0	0		0.5
Pearson and Sun (1993)			0			0.5
Ho and Lee (1986)		0	0		0	1.0
Hull and White (1990)			0		0	1.0
Black et al. (1990)	0			0		1.0
Black and Karasinski (1991)	0			0		1.0

$\alpha_1(t), \alpha_2(t), \alpha_3(t), \beta_1(t), \beta_2(t) \in \Re$, $\gamma \geq 0.5$, and $W(t)$ is a standard one-dimensional Brownian motion.

studies -e.g. Chan, Karolyi, Longstaff and Sanders (1992)- show weak evidence supporting the existence of mean reversion on short rates (suggesting, perhaps, that short rates tend to a stochastic short-term mean, which would then revert towards a constant long-term level), Schlögl and Sommer (1997) have established the relevance and appropriateness of the mean reversion assumption. Concerning the parameterization of the diffusion term, the interest rate models can be classified as *Gaussian* ($\beta_2(t) = 0$) or *stochastic volatility* ($\beta_2(t) \neq 0$) models.[4] The Gaussian term structure models imply a normal transition density function for the state variables (and hence, allow the interest rates to be negative with positive probability), while for the stochastic volatility case the factors' probability distribution can be known (for example, the non-central chi-square distribution in the Cox et al. (1985b) model) or unknown (as in the Pearson and Sun (1993) model).

Because a time-homogeneous single-factor model can not match (even just) the observed yield curve (simply through the dynamics of its shortest point), two modelling improvements are possible: to introduce time-dependencies on the model' parameters (as done, for instance, by Ho and Lee (1986) and Hull and White (1990) on the Merton (1970) and Vasicek (1977) models, respectively) in order to account for the unexplained yield curve features; or, to increase the number of state variables (moving towards multifactor specifications of increasingly computational complexity) in an attempt to explain a wider range of interest rate dynamics. The first approach (and the most popular one in the financial industry) allows the "calibrated" model to automatically reproduce a set of market observables (like swap rates and cap prices), but, most frequently, at the expense of too unstable parameters

[4] Hereafter, any interest rate model with a state-variable in the specification of the diffusion term will be classified as a *stochastic volatility* model (even if such state-variable can not be interpreted as the volatility of another model' factor).

(for hedging purposes).

The second approach is justified by a Principal Component Analysis (PCA) of the term structure of interest rates, which commonly reveals -as in Litterman and Scheinkman (1991)- that two or three factors (the *level*, the *slope* and the *curvature* of the yield curve) are needed to model the dynamics of the interest rates. All two- and three-factor models listed in Table 1.1 incorporate mean reversion, have time-independent parameters and are stochastic volatility models (except the Hull and White (1994) model, which is time-inhomogeneous and Gaussian): one of the most well known two-factor specifications is the Longstaff and Schwartz (1992a) model, which, like the one-factor Cox et al. (1985b) model, was derived from a general equilibrium context. More general formulations are given by the n-factor models, that allow the use of any finite number of state variables: the Langetieg (1980) model is the most general multifactor and Gaussian time-homogeneous specification, and the Chen and Scott (1995b) model corresponds to a multifactor "square-root" (and independent) Cox et al. (1985b) process. The Duffie and Kan (1996) model is the most general time-homogeneous exponential-affine specification, including as special cases most of the other interest rate models considered so far. An exponential-affine interest rate model is one that produces pure discount bond prices as exponential-affine functions of the state variables (or, equivalently, affine[5] continuously compounded yields), which - as argued by Brown and Schaefer (1994a, Proposition 4) and Duffie and Kan (1996, page 381)- is only consistent with an affine specification for both the drift and the instantaneous variance-covariance matrix of the state variables. Because interest rates are affine in the state variables, this class of exponential-affine models is amenable to econometric analysis and analytically tractable, which explains its wide dissemination. Frachot and Lesne (1993) extend these models to specifications with time-dependent parameters, that is models with the ability to fit the observed term structure of interest rates.

All the previous models can be recast into an even more general formulation: the Heath et al. (1992) class of models. The Heath et al. (1992) framework models, simultaneously, all instantaneous forward rates (and hence involves an infinite number of factors) through a finite number of Brownian motions, requiring as inputs only the initial forward rate curve and the forward rate volatility structure. Therefore, the initial yield curve is naturally fitted by the model and, as all evolutionary models, it does not require the knowledge of market prices of interest rate risk (since all factors are directly modelled under an equivalent martingale measure). Unfortunately, most of the Heath et al. (1992) formulations are non-Markovian and, consequently, their analytical complexity inhibits its dissemination (for example, the

[5] An affine form corresponds to a constant plus a linear function.

necessary usage of non-recombining trees or of Monte Carlo simulation seriously reduces the efficiency of American option pricing). One remarkable exception is the Brace et al. (1997) model where, by modelling discretely compounded (LIBOR) rates as lognormal processes (in a Heath et al. (1992) setting but under different equivalent martingale measures), the Black (1976) pricing formulae is justified for both caps and European swaptions.[6] Kennedy (1991) extends further the Heath et al. (1992) approach to an infinite number of Brownian motion shocks, but in a Gaussian framework.

The interest rate models studied in this dissertation are all (multifactor) exponential-affine models as characterized by Duffie and Kan (1996), which correspond to the majority of the models listed in Table 1.1. All the new pricing and estimation methodologies proposed in this thesis will be derived in the context of the Duffie and Kan (1996) model and are, therefore, applicable to most of the interest rate models already present in the literature. Although the exponential-affine class of interest rate models has been (the most) widely treated in the literature, it is the author's deeply conviction that further research is still needed, in, at least, two directions. Firstly, besides few exponential-affine interest rate models with known transition probability distributions for the state variables (like all Gaussian models or the one-factor and multifactor "square-root" specifications), the majority of the formulations nested into the Duffie and Kan (1996) model only provides exact numerical (and time-consuming) pricing solutions for interest rate contingent claims. Consequently, there exist innumerous stochastic volatility exponential-affine specifications yet "unexplored", and for which this dissertation is intended to produce analytical pricing solutions that enable the effective use and empirical test of such models. Secondly, the affine specification provided by the Duffie and Kan (1996) model for continuously compounded interest rates has been typically used to estimate the model' parameters from interest rate derivatives measuring only the *level* of the yield curve (such as coupon-bearing bond prices or swap rates). By deriving analytically tractable pricing solutions for a wider range of interest rate derivative securities (like caps and swaptions), it will be possible to construct estimation methodologies that are able to incorporate into the model additional information (e.g. on the *slope* and *curvature* of the yield curve). These two directions will be pursued by considering different model dimensions and both time-homogeneous and time-inhomogeneous specifications.

[6] Moreover, forward rates remain bounded, positive and mean reverting.

1.2 Purpose, contributions and methodology

In synthesis, this dissertation is intended to pursue four research goals:

1. Derivation of *fast* and *accurate* analytical (approximate) pricing formulae for the whole class of stochastic volatility exponential-affine multifactor term structure models, as an alternative to the existent exact numerical pricing methods.

   Duffie and Kan (1996) offer a quasi-analytical pricing formula for riskless zero-coupon bonds, and price path-independent interest rate options, in a two-factor model, through an alternating directions implicit (ADI) finite-difference method. Unfortunately, such algorithm can not be easily extended to higher dimensions, for which, and according to Duffie and Kan (1994), Monte Carlo simulation appears to be the best pricing methodology available.

   Recently, Duffie, Pan and Singleton (1998) proposed exact Fourier transform pricing solutions for an affine jump-diffusion model that nests, as a special case, the Duffie and Kan (1996) framework under analysis. If the functional form of the relevant characteristic function -Duffie et al. (1998, equation B.2)- is known, then the exact Fourier transform pricing formulae are closed-form solutions (in the sense that only one-dimensional Fourier inverse integrals are involved). However, in general the characteristic function does not possess an explicit solution and must be numerically obtained from a complex-valued system of Riccati differential equations. Moreover, when the characteristic function is not known in closed-form, the optimization of both the grid size and the upper bound of integration for the computation of the inverse Fourier transforms becomes also too time-demanding for empirical purposes, since it requires the numerical evaluation of the characteristic function at each integration point.

   In summary, besides the few special cases for which the analytical solution of the characteristic function is known, for the large majority of the stochastic volatility diffusion models a simple and reliable pricing methodology is still to be found.

2. Construction of a *robust*, *simple* and *accurate* estimation method for exponential-affine multifactor models.

   The literature contains three different approaches concerning the estimation of time-homogeneous term structure models: the "time-series", the "cross-section", and the "panel-data" methodologies. The "time-series" approach -as, for instance, in Chan et al. (1992)- estimates the model' parameters (except the ones related to the investors' preferences) under objective probabilities and using only a time-series of state variables' values, that is only considering the dynamics of the yield curve. The "cross-section" methodology -as, for example, in Brown and Dybvig (1986)- fits, under an equivalent martingale measure, the current shape of the yield curve by estimating the model only through a cross-section of interest rate derivatives prices (such as today's bond prices), and therefore is prone to generate unstable parameters' estimates over time. More efficiently, the "panel-data" approach uses both time-series and cross-sectional interest rate data, by modelling the state variables dynamics under both objective and risk-neutral probability measures, yielding more stable parameters' estimates (including the ones related to the market price of interest rate risk).

Under this last methodology, some authors -e.g. Pearson and Sun (1994), Chen and Scott (1993b), or Duffie and Singleton (1997)- assume that one or more spot interest rates are observed without error, which enables the model state variables to be exactly recovered from the data without the use of filtering techniques. More realistically, other authors -such as Jegadeesh and Pennacchi (1996), Babbs and Nowman (1999), Duan and Simonato (1995) or Geyer and Pichler (1996)- explicitly account for the existence of measurement errors in the data by restating the interest rate model in a "state-space" form: the parameters are estimated and the unobservable state variables are inferred through the use of a Kalman filter. Such "panel-data" and "state-space" approach will be the one pursued in this dissertation, but with an important distinctive feature. While all the previous literature have simply applied Kalman filtering techniques to data only containing information about the *level* of the yield curve (like zero-coupon yields and swap rates), this thesis will propose (non-linear) filtering estimation methods based on a panel-data of swap rates, cap prices and swaption prices. Hence, the fit of exponential-affine term structure models to both the volatility and the correlation surfaces will be enhanced and tested.

In summary, the proposed methodology for the estimation of exponential-affine models is intended to possess three advantages over the previous approaches: *i*) robustness (because parameters' stability is improved through the simultaneous model' fit to both time-series dynamics and cross-sectional shapes of the yield curve); *ii*) accuracy (since the enlargement of the set of market observables used in the model' estimation should enhance the model ability to reproduce the market prices of basic interest rate derivatives, such as caps, floors and swaptions); and *iii*) simplicity (in the sense that the new analytical pricing solutions derived in this thesis will enable the "real time" use of additional derivatives prices in the estimation process, by providing the required measurement equations for the Kalman filter recursions).

3. Testing the dimensional requirements for exponential-affine models

   As already mentioned, the one-factor models' unrealistic assumption of perfect interest rates correlation (amongst different maturities) as well as the PCA' empirical finding of a two- or three-dimensional state space for the interest rates variance-covariance matrix make a strong argument in favor of multifactor model specifications. However, Rebonato (1998, page 70) suggests that only three state variables should not be enough to capture the *fast decorrelation phenomena* at the short-end of the yield curve, i.e. that such low-dimensional formulation would imply too high correlations between interest rates of adjacent maturities.

   In this dissertation, exponential-affine models with different (increasing) numbers of factors will be estimated and it will be empirically tested which is the minimal dimension required to reproduce the market interest rates correlation matrix (specially at the short-end of the maturity spectrum).

4. Testing the time-homogeneity assumption

   Although an evolutionary model offers the ability to *account* for the observed market prices of some set of interest rate contingent claims, it is well known that the parameters' stability is often too poor for hedging purposes. Thus, in a first stage, instead of incorporating time-dependent parameters into the model specification, the improvement of the model fit to the term structure of interest rates values, volatilities and correlations will be tried by increasing successively the number of state variables. Secondly, time-inhomogeneous HJM exponential-affine model' generalizations will also be estimated, in two stages: the time-independent parameters are still obtained from the corresponding time-homogeneous specifications (using filtering techniques); and then, the time-dependent parameters will be estimated by improving the cross-sectional fit of the model to the volatility and/or to the correlation market curves.

   Both approaches will be compared in terms of parameters' stability and pricing accuracy.

The remaining chapters of this dissertation are organized as follows. Chapter two summarizes the Duffie and Kan (1996) model under analysis (which was originally specified in terms of risk-adjusted stochastic processes for its state variables) and provides an equivalent general equilibrium formulation, under the objective probability measure and using both log and power utility functions. The proposed general equilibrium setup represents a synthesis between the consumption-based CAPM of Breeden (1979), the pure exchange economy of Lucas (1978), and the cash-in-advance one-country economy of Lucas (1982). Thus, both the short-term interest rate and the factor risk premiums are expressed in terms of the direct utility function, and as functions of the exogenous output and money supply processes. This avoids the need to solve the Hamilton-Jacobi-Bellman equation for the utility of wealth or for the endogenous consumption process, and, consequently, allows the use of a general equilibrium framework based on preference assumptions more realistic than those implied by the conventional log utility of consumption. Moreover, since a monetary economy is considered, the general equilibrium Duffie and Kan (1996) model specification that will emerge is a term structure model of *nominal* interest rates. Chapter two serves essentially one instrumental goal: it generates a (theoretically justified) functional form for the vector of market prices of risk, which will be required to estimate, in Chapter four, exponential-affine term structure models using a panel-data and state-space approach.

Chapter three is devoted to the derivation of closed-form pricing solutions, for several interest rate contingent claims, under a Gaussian (nested) version of the Duffie and Kan (1996) model. Starting from the pricing formula for default-free pure discount bonds already derived by Langetieg (1980), closed-form solutions will also be found for the prices of (short-term and long-term) interest rate futures and European (conventional and pure) futures options, European spot options on default-free (pure discount and coupon-bearing) bonds, caps, floors and European swaptions. Futures will be priced as moment generating functions, and options will be valued using the well known probabilistic change-of-numeraire technique developed in El Karoui, Lepage, Myneni, Roseau and Viswanathan (1991) or Geman, Karoui and Rochet (1995). The purpose of Chapter three is twofold. On one hand, it provides the measurement equations needed for the Kalman filter estimation of Gaussian exponential-affine term structure models, in Chapter four. On the other hand, the Gaussian exact closed-form pricing solutions produced in Chapter three are also essential to generate, in Chapter five, approximate analytical pricing formulae under a general stochastic volatility specification of the Duffie and Kan (1996) model.

In Chapter four, Gaussian time-homogeneous and exponential-affine term structure models are fitted to a panel-data of swap rates, cap prices and European swaption prices, using a non-linear Kalman filter and considering different numbers of state variables. Consistently with the previous literature, low-dimensional specifications will be able to fit extremely well the term structure of interest rates. However, even by increasing the number of factors, such models will be shown to be incapable of adequately fitting short-term cap and swaption prices. Nevertheless, it is argued that the use of such enlarged market data in the estimation process allows the model to incorporate additional information about the market interest rates covariance matrix. Concerning the time-homogeneity issue, an equivalent Gauss-Markov time-inhomogeneous HJM model, estimated in two stages, is proposed. The term structure of volatilities is then easily recovered and the pricing errors for swaptions improve substantially, in the context of stable time-homogeneous coefficients, which are still estimated through a Kalman filter approach.

Chapter five contains the main theoretical contribution of this dissertation: the derivation of approximate analytical pricing solutions under the general stochastic volatility specification of the Duffie and Kan (1996) model, which only involve one integral with respect to the maturity of the contingent claim under valuation (no matter the dimension of the interest rate model in use), and are therefore extremely easy to implement in practice. Starting by obtaining the functional form of Arrow-Debreu prices under the nested Gaussian specification developed in Chapter three, the closed-form and exact Gaussian valuation formulae of Chapter three will be converted into approximate stochastic volatility ones that involve integrals with respect not only to the maturity of the contingent claim under valuation but also to each one of the model' factors. Finally, and taking advantage of the analytical tractability provided by the nested model specification adopted, all stochastic volatility pricing formulae will be easily simplified into first order approximate ones that do not involve any integration with respect to the model' state variables. Such fast and accurate analytic approximations will be obtained for bonds, forward rate agreements, interest rate swaps, interest rate futures, European options on pure discount bonds, caps and floors, yield options, European futures options on zero-coupon bonds and on short-term interest rates, and even for European swaptions. As an accessory result, exact pricing solutions are also provided for long-term and short-term interest rate futures, under the stochastic volatility specification of the Duffie and Kan (1996) model.

Finally, Chapter six outlines possible areas of further research arising from the findings of this thesis, while Chapter seven summarizes the dissertation' conclusions.

Chapter 2

General Equilibrium Framework

2.1 Duffie and Kan (1996) model: a summary

The Duffie and Kan (1996) model imposes an exponential-affine form for the price of a riskless (unit face value) pure discount bond, that is

$$P(t,T) = \exp\left[A(\tau) + \underline{B}'(\tau) \cdot \underline{X}(t)\right], \tag{2.1}$$

where $P(t,T)$ represents the time-t price of a default-free pure discount bond expiring at time T, $\tau = T - t$ is the time-to-maturity of the zero-coupon bond, $\cdot$ denotes the inner product in $\Re^n$, and $\underline{X}(t) \in \Re^n$ is the time-t vector of state variables. In order to respect the boundary condition $P(T,T) = 1$, the time-homogeneous functions $A(\tau) \in \Re$ and $\underline{B}(\tau) \in \Re^n$ must be such that $A(0) = 0$ and $\underline{B}(0) = \underline{0}$. Moreover, the function $P(t,T)$ is assumed to be continuously differentiable in the time-to-maturity and twice continuously differentiable in the state-vector.

As in the *yield-factor model* proposed by Duffie and Kan (1996, section 5), conditional on knowing the *true* model and ignoring the existence of measurement errors, it is always possible to "observe" the state variables from a selected basis of fixed maturity spot interest rates. However, in this dissertation the state variables will be assumed to be unobservable, because the existence of market imperfections (e.g. bid-ask spreads) does not allow, in practice, all the (factor) yields to be observed without error. Although such assumption of observation errors seems to be more realistic, it also induces additional difficulties in terms of model estimation: filtering methods must be used to estimate the model' parameters and to recover the latent state variables.

Alternatively to zero-coupon bond prices, the model can be equivalently specified in terms of the riskless instantaneous spot interest rate. Because $A(\cdot)$ and $\underline{B}(\cdot)$ are continuously differentiable (since it is assumed that $P(t,T) \equiv P(\underline{X}(t);\tau) \in C^{2,1}(\mathbf{D} \times [0,\infty[)$, where $\mathbf{D} \subseteq \Re^n$ represents the admissible domain of the model' state variables), it follows from equation (2.1) that the time-t short-term interest rate $r(t)$ is an affine function of the n factors:

$$\begin{aligned} r(t) &= \lim_{\tau \to 0} \left[-\frac{\ln P(t,T)}{\tau} \right] \\ &= f + \underline{G}' \cdot \underline{X}(t), \end{aligned} \tag{2.2}$$

where $f = -\left.\frac{\partial A(\tau)}{\partial \tau}\right|_{\tau=0}$, and the i^{th} element of vector $\underline{G} \in \Re^n$ is defined as $g_i = -\left.\frac{\partial B_i(\tau)}{\partial \tau}\right|_{\tau=0}$, being $B_i(\tau)$ the i^{th} element of vector $\underline{B}(\tau)$.

Concerning the dynamics of the model' factors, Duffie and Kan (1996) start by considering a probability space $\left(\Omega, (\mathcal{F}_t)_{t>0}, \mathcal{P}\right)$, and assume that under the objective probability measure $\mathcal{P}$ the Markov process $\underline{X}(t)$ satisfies a stochastic differential equation of the generic form

$$d\underline{X}(t) = \underline{v}[\underline{X}(t)]\,dt + \sigma[\underline{X}(t)] \cdot d\underline{W^{\mathcal{P}}}(t), \tag{2.3}$$

where $\underline{v}[\underline{X}(t)] \in \Re^n$ and $\sigma[\underline{X}(t)] \in \Re^{n \times n}$ satisfy the *Lipschitz* and *growth* conditions required for a unique solution to exist for equation (2.3),[1] while $\underline{W^{\mathcal{P}}}(t) \in \Re^n$ is a standard Brownian motion under $\mathcal{P}$ generating the augmented filtration $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$. Then, they argue that it is always possible to derive a probability martingale measure $\mathcal{Q}$ equivalent to $\mathcal{P}$ (that is mutually absolutely continuous) and a standard $\mathcal{Q}$-measured Brownian motion $\underline{W^{\mathcal{Q}}}(t) \in \Re^n$ (with the same standard filtration as $\underline{W^{\mathcal{P}}}(t)$), such that

$$d\underline{X}(t) = \underline{\mu}[\underline{X}(t)]\,dt + \sigma[\underline{X}(t)] \cdot d\underline{W^{\mathcal{Q}}}(t) \tag{2.4}$$

and where $\underline{\mu}[\underline{X}(t)] \in \Re^n$ is a compatible function of $\underline{v}[\underline{X}(t)]$, $\sigma[\underline{X}(t)]$ and $P(t,T)$, in the sense that this change of drift guarantees the absence of arbitrage opportunities[2] and also preserves an exponential-affine specification for pure discount bond prices. Finally, Duffie and Kan (1996) define what they call a $(P, \underline{\mu}, \sigma)$ compatible term structure model by specifying the exponential-affine form (2.1) for $P(t,T)$ and affine formulae for both $\underline{\mu}[\underline{X}(t)]$ and $\sigma[\underline{X}(t)] \cdot \sigma[\underline{X}(t)]'$.[3] In other words, the Duffie and Kan (1996) model was

[1] And stated, for instance, in Lamberton and Lapeyre (1996, theorem 3.5.5)

[2] Meaning that the relative prices of all assets with respect to the numeraire given by a "money market account" are $\mathcal{Q}$-martingales. The time-t value of such "savings account", $\delta(t)$, corresponds to the compounded value of one monetary unit continuously reinvested, from time 0 to time t, at the short-term interest rate:

$$\delta(t) = \exp\left[\int_0^t r(s)\,ds\right].$$

[3] $\sigma[\underline{X}(t)]'$ denotes the transpose of $\sigma[\underline{X}(t)]$.

originally defined not under objective probabilities but in terms of risk-adjusted stochastic processes for its state variables, i.e. with respect to a martingale measure Q which can be understood as the probability measure obtained when a "money market account" is taken as the numeraire of the stochastic intertemporal economy underlying the model under analysis.[4]

Specifically, Duffie and Kan (1996) assume that the n state variables follow, under a martingale measure Q, a parametric Markov diffusion process, where the drift and the variance of these risk-adjusted stochastic processes also have an affine form, in order to support[5] the exponential-affine specification of equation (2.1):

$$d\underline{X}(t) = \left[a \cdot \underline{X}(t) + \underline{b}\right] dt + \Sigma \cdot \sqrt{V^D(t)} \cdot d\underline{W}^Q(t), \underline{X}(t) \in \mathbf{D}, \tag{2.5}$$

where $a, \Sigma \in \Re^{n \times n}$, $\underline{b} \in \Re^n$,

$$\sqrt{V^D(t)} = diag\left\{\sqrt{v_1(t)}, \ldots, \sqrt{v_n(t)}\right\},$$

$$v_i(t) = \alpha_i + \underline{\beta_i}' \cdot \underline{X}(t), \quad for\ i = 1, ..., n,$$

$\alpha_i \in \Re$, $\underline{\beta_i} \in \Re^n$, $\underline{W}^Q(t) \in \Re^n$ is a vector of n independent Brownian motions under measure Q, and

$$\mathbf{D} = \left\{\underline{X} \in \Re^n : \alpha_i + \underline{\beta_i}' \cdot \underline{X} \geq 0, i = 1, \ldots, n\right\} \tag{2.6}$$

is the admissible domain of the model' state variables. Notice that this model specification incorporates mean reversion[6], and accommodates both deterministic (if $\underline{\beta_i} = \underline{0}, \forall i$) or stochastic volatility (if $\exists i : \underline{\beta_i} \neq \underline{0}$) formulations. Hereafter, condition A of Duffie and Kan (1996, page 387) will be always assumed in order to ensure that a unique (strong) solution $\underline{X}(t) \in \mathbf{D}$ exists for the SDE (2.5).

Equations (2.1) -or (2.2)- and (2.5) summarize the most general stochastic volatility specification of the Duffie and Kan (1996) model (since $\underline{\beta_i}$ is not constrained to be equal to $\underline{0}$, for all i). Applying Itô's lemma, it follows that, under this general specification, the time-t price, $Y\left[\underline{X}(t), t\right] \in C^{2,1}(\mathbf{D} \times [0, \infty[)$, of an interest rate contingent claim, with a continuous "dividend yield" $i\left[\underline{X}(t), t\right]$, must satisfy the following fundamental parabolic

[4] See section 3.3 for details.

[5] As Duffie and Kan (1996, page 381) say: "...the yields are affine if, and essentially only if, the drift and diffusion functions of the stochastic differential equation for the factors are also affine". This result is equivalent to proposition 4 of Brown and Schaefer (1994a), derived in the context of one-factor affine models.

[6] $\underline{X}(t)$ mean reverts towards $a^{-1} \cdot \underline{b}$, as long as matrix a is negative definite.

partial differential equation, subject to the appropriate boundary conditions:

$$\mathcal{D}Y(\underline{x},t)+\frac{\partial Y(\underline{x},t)}{\partial t}-r(t)Y(\underline{x},t)=-i(\underline{x},t),\underline{x}\in\mathbf{D}, \tag{2.7}$$

being $\mathcal{D}$ the second-order differential operator[7]

$$\mathcal{D}Y(\underline{x},t)=\frac{\partial Y(\underline{x},t)}{\partial\underline{x}'}\cdot[a\cdot\underline{x}+\underline{b}]+\frac{1}{2}tr\left[\frac{\partial^2 Y(\underline{x},t)}{\partial\underline{x}\partial\underline{x}'}\cdot\Sigma\cdot V^D(t)\cdot\Sigma'\right],$$

with $V^D(t)=diag\{v_1(t),\ldots,v_n(t)\}$, and where the function $tr(\cdot)$ returns the trace of a square matrix. However, and as Duffie and Kan (1994) point out, PDE (2.7) can only be solved, for path-independent interest rate contingent claims, by a finite-difference method or, for large n, by Monte Carlo simulation. The only exception seems to be the valuation of default-free pure discount bonds, for which an exact quasi-closed form solution is provided by Duffie and Kan (1996). Using equations (3.9) and (3.10) of Duffie and Kan (1996)[8], first the *duration* vector $\underline{B}'(\tau)$ must be found through the solution of a system of n Riccati differential equations (for instance, by using a fifth order Runge-Kutta method),

$$\frac{\partial}{\partial\tau}\underline{B}'(\tau)=-\underline{G}'+\underline{B}'(\tau)\cdot a+\frac{1}{2}\sum_{k=1}^{n}\left[\sum_{j=1}^{n}B_j(\tau)\,\varepsilon_{jk}\right]^2\underline{\beta_k}', \tag{2.8}$$

subject to the initial condition $\underline{B}(0)=\underline{0}$, and where ε_{jk} is the j^{th}-row k^{th}-column element of matrix Σ. Then, $A(\tau)$ is obtained through the solution of a first order ordinary differential equation (for instance, by using Romberg's integration method),

$$\frac{\partial}{\partial\tau}A(\tau)=-f+\underline{B}'(\tau)\cdot\underline{b}+\frac{1}{2}\sum_{k=1}^{n}\left[\sum_{j=1}^{n}B_j(\tau)\,\varepsilon_{jk}\right]^2\alpha_k, \tag{2.9}$$

subject to the initial condition $A(0)=0$. Finally, $P(t,T)$ is given by equation (2.1). However, under this general specification of the Duffie and Kan (1996) model, even the above ODEs must be solved numerically.

The main advantage of the Duffie and Kan (1996) framework is its generality: all time-homogeneous exponential-affine models presented in the literature can be easily nested into the specification given by equations (2.2) and (2.5), through self-evident parameters'

[7] As defined in Arnold (1992, definition 2.6.1). Its relation with the infinitesimal generator of $\underline{X}(t)$, $\mathcal{A}$, is the following:

$$\mathcal{A}=\frac{\partial}{\partial t}+\mathcal{D}.$$

[8] That is substituting $Y[\underline{X}(t),t]$, in PDE (2.7), by equation (2.1), subject to the boundary condition $P(T,T)=1$.

Table 2.1: Parameters' restrictions needed to fit some term structure models into the Duffie and Kan (1996) general specification

	Vasicek (1977)	Cox et al. (1985b)	Longstaff and Schwartz (1992a)	Langetieg (1980)	Chen and Scott (1995b)
n	1	1	2		
f	0	0	0		0
$\underline{G}$	1	1			$\underline{1}$
a			diagonal		diagonal
$\underline{b}$					
Σ			I_2		diagonal
$\underline{\alpha}$	1	0	$\underline{0}$		$\underline{0}$
β	0	1	I_2	O_n	I_n

O_n and I_n denote $n \times n$ null and identity matrices, respectively.
$\underline{\alpha} \in \Re^n$ is a vector with α_i as its i^{th}-component.
$\beta \in \Re^{n \times n}$ is a matrix whose i^{th}-column is given by vector $\underline{\beta_i}$.

restrictions (Table 2.1 illustrates some examples). Therefore, the general equilibrium setup that will be constructed in this Chapter is also applicable to any of such models.

2.2 General equilibrium specification for the Duffie and Kan (1996) model

The goal of the present Chapter is to derive a Duffie and Kan (1996) model' specification under the original probability measure $\mathcal{P}$ that is compatible with the formulation given by the authors under the equivalent martingale measure $\mathcal{Q}$. This task can become useful for empirical purposes, namely for the econometric estimation of the Duffie and Kan (1996) model' parameters from a time-series of state variables' values or from a panel-data of market observables (e.g. bond prices), through Kalman filtering techniques. In fact, these parameters can also be estimated from a cross-section of bond prices, by using the risk-adjusted processes for the state variables[9], since assuming that there are no arbitrage opportunities in the bond market is equivalent to say that such interest rate contingent claims can be priced under an equivalent martingale measure $\mathcal{Q}$. However, this latter methodology should be less adequate than the time-series or panel-data approaches, because the model' parameters are assumed to be time-independent. In summary, if the Duffie and Kan (1996) model' parameters are to be estimated through a time-series or a panel-data methodology, the knowledge of the model' specification under the objective probability measure $\mathcal{P}$ is then required, and thus justifies the purpose of this Chapter. As Duffie and Kan (1994, page 578) notice: "For many applications, it will also be useful to model the distribution of processes

[9]That is through the best fit between market bond prices and those generated by the model.

under the original probability measure $\mathcal{P}$. Conversion from $\mathcal{P}$ to $\mathcal{Q}$ and back will not be dealt with here, but is an important issue, particularly from the point of view of statistical fitting of the models as well as the measurement of risk."

In order to derive the Duffie and Kan (1996) model' specification under the probability measure $\mathcal{P}$, it will be necessary to fit the model into a general equilibrium framework. This is so, because, from Girsanov's theorem, the two model specifications (under probability measures $\mathcal{P}$ and $\mathcal{Q}$) are only compatible if $\underline{\mu}[\underline{X}(t)]$ and $\underline{W}^{\mathcal{Q}}(t)$ are such that:

$$\underline{\mu}[\underline{X}(t)] = \underline{v}[\underline{X}(t)] - \sigma[\underline{X}(t)] \cdot \underline{\Lambda}[\underline{X}(t)]$$

and

$$d\underline{W}^{\mathcal{Q}}(t) = \underline{\Lambda}[\underline{X}(t)]\, dt + d\underline{W}^{\mathcal{P}}(t),$$

where $\underline{\Lambda}[\underline{X}(t)] \in \Re^n$ satisfies the Novikov's condition[10], and the Radon-Nikodym derivative $\xi_t \equiv \frac{d\mathcal{Q}}{d\mathcal{P}}$ is equal to

$$\xi_t = \exp\left\{-\int_0^t \underline{\Lambda}[\underline{X}(s)]' \cdot d\underline{W}^{\mathcal{P}}(s) - \frac{1}{2}\int_0^t \underline{\Lambda}[\underline{X}(s)]' \cdot \underline{\Lambda}[\underline{X}(s)]\, ds\right\}.$$

Hence, to go from the Duffie and Kan (1996) model specification under the original probability measure $\mathcal{P}$ -hereafter labelled as the $(P, \underline{v}, \underline{\Lambda}, \sigma)$ model- to the $(P, \underline{\mu}, \sigma)$ equivalent specification, or all the way around, it is necessary to define $\underline{\Lambda}[\underline{X}(t)]$ explicitly. For that purpose, the Duffie and Kan (1996) model will have to be fitted into a general equilibrium framework, where both the short-term interest rate and the vector of market prices of risk will be endogenously determined in the context of the underlying economy. And, unlike the majority of the general equilibrium term structure models found in the literature, the role of money is going to be explicitly considered, leading to a general equilibrium Duffie and Kan (1996) model of the term structure of *nominal* interest rates.

Next subsections are organized as follows. Subsection 2.2.1 state all the assumptions that are required to fit the Duffie and Kan (1996) model into a general equilibrium framework. In subsections 2.2.2, 2.2.3 and 2.2.4, general formulae for the equilibrium short-term interest rate and for the equilibrium factor risk premiums are derived, always in *nominal* terms: first, within the context of a production economy; then, under a consumption-based CAPM; and finally, assuming a pure exchange economy. In subsection 2.2.5, a general equilibrium Duffie and Kan (1996) model is derived under a constant relative risk aversion economy (both with power and log utility functions). Finally, section 2.3 summarizes the conclusions.

[10] $\underline{\Lambda}[\underline{X}(t)]$ can be interpreted as the time-t vector of market prices of interest rate risk.

2.2.1 General equilibrium assumptions

The following assumptions represent a synthesis between the consumption-based CAPM of Breeden (1979), the continuous-time pure exchange economy of Lucas (1978), and the cash-in-advance one-country economy of Lucas (1982), while the notation is intended to follow that used by Cox, Ingersoll and Ross (1985a):

A.1) There is a single physical good, which may be allocated to consumption or investment.

A.2) The stochastic intertemporal one-country economy that is going to be considered has a finite time horizon $\mathcal{T} = [0, T]$. Uncertainty is represented by a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$, where all the information accruing to all the agents in the economy is described by a filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$ satisfying the usual conditions: namely, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. The vector $\underline{W}^{\mathcal{P}}(t) \in \Re^n$ will represent a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, and $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$ will denote the $\mathcal{P}-$augmentation of the natural filtration generated by $\underline{W}^{\mathcal{P}}(t)$.

A.3) There are n state variables that determine the general state of the economy (both in real and monetary terms) through the following stochastic process, and under the probability measure $\mathcal{P}$:

$$d\underline{X}(t) = \left[\bar{a} \cdot \underline{X}(t) + \underline{\bar{b}}\right] dt + \Sigma \cdot \sqrt{V^D(t)} \cdot d\underline{W}^{\mathcal{P}}(t), \tag{2.10}$$

where $\bar{a} \in \Re^{n \times n}$, $\underline{\bar{b}} \in \Re^n$, and $d\underline{W}^{\mathcal{P}}(t) \in \Re^n$ is a vector of n independent Brownian increments under the objective probability measure. Hence, $\underline{\nu}\left[\underline{X}(t)\right] = a \cdot \underline{X}(t) + \underline{b}$ and $\sigma\left[\underline{X}(t)\right] = \Sigma \cdot \sqrt{V^D(t)}$. This stochastic differential equation (SDE) is intended to represent the non-risk-adjusted stochastic process followed by the state variables of the Duffie and Kan (1996) model. Thus, the diffusion is the same as in equation (2.5), and the drift was defined as another affine function of the n factors (in order to be consistent with the exponential-affine form for pure discount bond prices). The goal is precisely to determine a consistent relation between a and $\bar{a}$ as well as between $\underline{b}$ and $\underline{\bar{b}}$.

A.4) There exist m distinct production processes (or production firms) that define m investment opportunities in this economy, whose dynamics are modelled through the following SDE:

$$d\underline{S}(t) = I_S(t) \cdot \underline{\mu_S}(q, M, \underline{S}, \underline{X}, t)\, dt + I_S(t) \cdot E(q, M, \underline{S}, \underline{X}, t) \cdot d\underline{W}^{\mathcal{P}}(t). \tag{2.11}$$

The i^{th} element of $\underline{S}(t) \in \Re^m$, denoted by $S_i(t)$, represents the *nominal* value of the i^{th} production firm at time t,[11] $I_S(t) = diag\{S_1(t), \ldots, S_m(t)\}$ and therefore the production processes have stochastically constant returns to scale, $\underline{\mu_S}(q, M, \underline{S}, \underline{X}, t) \in \Re^m$ is the vector of expected rates of return on the production activities, $E(q, M, \underline{S}, \underline{X}, t) \in \Re^{m \times n}$ is assumed to be such that $E(q, M, \underline{S}, \underline{X}, t) \cdot E(q, M, \underline{S}, \underline{X}, t)'$ is positive definite, $q(t)$ denotes the time-t aggregate output of the economy, and $M(t)$ represents the time-t money supply level.

A.5) The real aggregate production output is exogenously determined[12] by the following diffusion process:

$$\frac{dq(t)}{q(t)} = \mu_q(q, \underline{X}, t)\, dt + \underline{\sigma_q}(q, \underline{X}, t)' \cdot d\underline{W}^{\mathcal{P}}(t), \tag{2.12}$$

where $\mu_q(q, \underline{X}, t) \in \Re$ is the time-t expected rate of change in the aggregate output, and $\underline{\sigma_q}(q, \underline{X}, t) \in \Re^n$ is the vector of volatilities for the rate of change in the aggregate output.

A.6) The money supply is exogenously determined by the following diffusion process:

$$\frac{dM(t)}{M(t)} = \mu_M(M, \underline{X}, t)\, dt + \underline{\sigma_M}(M, \underline{X}, t)' \cdot d\underline{W}^{\mathcal{P}}(t), \tag{2.13}$$

where $\mu_M(M, \underline{X}, t) \in \Re$ is the time-t expected growth rate of money supply, and $\underline{\sigma_M}(M, \underline{X}, t) \in \Re^n$ is the vector of volatilities for the money supply growth rate.

A.7) There are $(n - m)$ infinitely divisible financial contingent claims, whose net supply is zero, and whose *nominal* value evolves according to the following stochastic process:

$$d\underline{F}(t) = I_F(t) \cdot \underline{\mu_F}(q, M, \underline{S}, \underline{X}, t)\, dt + I_F(t) \cdot H(q, M, \underline{S}, \underline{X}, t) \cdot d\underline{W}^{\mathcal{P}}(t), \tag{2.14}$$

where the i^{th} element of $\underline{F}(t) \in \Re^{n-m}$, denoted by $F_i(t)$, represents the time-t price of the i^{th} contingent claim, $I_F(t) = diag\{F_1(t), \ldots, F_{n-m}(t)\}$, $\underline{\mu_F}(q, M, \underline{S}, \underline{X}, t) \in \Re^{n-m}$ is the vector of expected rates of return (dividend-inclusive) on the $(n - m)$ financial contingent claims, and $H(q, M, \underline{S}, \underline{X}, t) \in \Re^{(n-m) \times n}$.

A.8) There are no taxes or transaction costs, and all trades take place at equilibrium prices.

[11] Each firm's value is represented by just one (perfectly divisible) share, i.e. $S_i(t)$ can be though of as being the value of the i^{th} production firm share.

[12] This is the main difference between the pure exchange economy considered here and the Cox et al. (1985a) type of production economy.

A.9) There exists a market for instantaneous borrowing and lending at a *nominal* risk-free interest rate of $r(t)$.

A.10) There exists a fixed number of individuals, all identical in terms of their endowments and preferences, and all having homogeneous probability beliefs about future states of the world. Thus, it can be automatically assumed that markets are dynamically complete, because as said in Cox, Ingersoll and Ross (1981a, page 779): "For an economy of identical investors, prices will be set as if markets were complete, regardless of their actual scope". Moreover, each individual seeks to maximize the expected value of a time-additive and state-independent von Neumann-Morgenstern utility function for lifetime consumption, that is wishes to maximize the quantity

$$E_t \left\{ \int_t^T u\left[C(s), s\right] ds \middle| V(t) = v = \sum_{i=1}^m S_i(t) \ \ and \ \underline{X}(t) = \underline{x} \right\},$$

where t denotes the current time, T represents the terminal date, the expectation is conditional on $\mathcal{F}_t$ and computed under measure $\mathcal{P}$, $u[\cdot]$ is a von Neumann-Morgenstern period utility function, $C(s)$ represents the amount of the single physical good consumed at time s, $V(t)$ is the time-t (i.e. current) pre-decision *nominal* wealth,[13] and $\underline{x}$ denotes the current state of the economy.

A.11) The unit-velocity version of the *Quantity Theory of Money* will be assumed hereafter, that is

$$\frac{M(t)}{p(t)\,q(t)} = 1, \tag{2.15}$$

where $p(t)$ is the time-t price level for the single physical good. This working hypothesis is just a consequence of the following three underlying assumptions:

A.11.1) In the economy under analysis all agents are subject to a cash-in-advance constraint (also known as the *Clower constraint*), in the sense that all goods can be purchased only with currency accumulated in advance, i.e.

$$N(t) = p(t)\,C(t), \tag{2.16}$$

where $N(t)$ is the time-t demand for money. This constraint justifies the existence of money in the economy, because as argued by Lucas (1982, page 342): "...agents will hold non-interest-bearing units of that currency in exactly the amount needed

[13] It is being assumed that the initial endowment of the representative agent corresponds to one share of each production firm.

to cover their perfectly predictable current-period goods purchases". Instead, one could have considered, for instance, the existence of real cash balances in the direct utility function, while assuming that $q(t)$ and $M(t)$ were the only state variables, as done by Bakshi and Chen (1996). Although this procedure would be more realistic, it would also create two problems: first, the choice of state variables would not be consistent with the Duffie and Kan (1996) model specification under analysis; second, the derivation of a closed-form expression for $\underline{\Lambda}[\underline{X}(t)]$ would require the use of a log utility function, restricting the type of preferences under consideration.

A.11.2) In equilibrium, the money supply equals the demand for money:

$$M(t) = N(t). \tag{2.17}$$

A.11.3) In this pure exchange economy, all output is consumed:

$$C(t) = q(t). \tag{2.18}$$

Combining equations (2.16), (2.17), and (2.18), equation (2.15) follows immediately.

Initially, assumption A.5 will be ignored, i.e. a Cox et al. (1985a) type of production economy will be considered, but the results obtained are going to depend on the indirect utility function. Then, we will move towards the consumption-based CAPM of Breeden (1979), obtaining results that depend on the direct utility function but are still related to the endogenous consumption process. Finally, assumption A.5 will be imposed, a pure exchange economy will be completely identified, and all the relevant results will be stated in terms of the utility of consumption and as a function of the exogenous output and money supply processes (therefore avoiding the need to solve any Hamilton-Jacobi-Bellman equation).

2.2.2 Portfolio selection problem

The budget constraint

The representative agent in this economy can choose amongst three different types of investment opportunities: i) To trade the equity shares issued by the m production firms; ii) To trade $(m-n)$ financial contingent claims; and iii) To buy or sell instantaneous *nominal* risk-free zero-coupon bonds.

Hence, the representative agent must observe the following budget constraint:[14]

$$dV(t) = V(t)\underline{\omega_S}(t)' \cdot I_S^{-1}(t) \cdot d\underline{S}(t) + V(t)\underline{\omega_F}(t)' \cdot I_F^{-1}(t) \cdot d\underline{F}(t) \qquad (2.19)$$
$$+V(t)\left[1 - \underline{\omega_S}(t)' \cdot \underline{1} - \underline{\omega_F}(t)' \cdot \underline{1}\right] r(t)\,dt - p(t)\,C(t)\,dt,$$

where $\underline{\omega_S}(t) \in \Re^m$, its i^{th} element, $\omega_{S_i}(t)$, is the proportion of the current wealth invested in the i^{th} production firm, $\underline{\omega_F}(t) \in \Re^{n-m}$, its i^{th} element, $\omega_{F_i}(t)$, is the proportion of the current wealth invested in the i^{th} financial contingent claim, and $r(t)$ is the instantaneous *nominal* risk-free time-t interest rate. Considering equations (2.11) and (2.14), the above stochastic differential equation can be restated as:

$$dV(t) = \left\{\underline{\omega_S}(t)' \cdot \left[\underline{\mu_S}(t) - r(t)\underline{1}\right] V(t) + \underline{\omega_F}(t)' \cdot \left[\underline{\mu_F}(t) - r(t)\underline{1}\right] V(t) \qquad (2.20)$$
$$+V(t)\,r(t) - p(t)\,C(t)\right\} dt$$
$$+V(t)\left[\underline{\omega_S}(t)' \cdot E(t) + \underline{\omega_F}(t)' \cdot H(t)\right] \cdot d\underline{W}^{\mathcal{P}}(t).$$

The HJB equation

The individual's portfolio selection problem consists in choosing a policy for investment and consumption, i.e. choosing the controls $\left(\underline{\omega_S}(t), \underline{\omega_F}(t), C(t)\right) \equiv \left(\underline{\omega_S}, \underline{\omega_F}, C\right)$, so as to maximize the expected utility from consumption, subject to the budget constraint (2.20). In other words, the representative agent has to find $\left(\underline{\omega_S}, \underline{\omega_F}, C\right)$ such that:[15]

$$J(v, \underline{x}, t) = \max_{\left(\underline{\omega_S}, \underline{\omega_F}, C\right)} K^{\underline{\omega_S}, \underline{\omega_F}, C}(v, \underline{x}, t),$$

where

$$K^{\underline{\omega_S}, \underline{\omega_F}, C}(v, \underline{x}, t) = E_t\left\{\int_t^T u\left[C(s), s\right] ds \,\middle|\, V(t) = v \text{ and } \underline{X}(t) = \underline{x}\right\},$$

being dv given by equation (2.20) and $d\underline{x}$ given by equation (2.10).

The Hamilton-Jacobi-Bellman equation for the above stochastic optimal control problem is:

$$0 = \max_{\left(\underline{\omega_S}, \underline{\omega_F}, C\right)} \phi\left(\underline{\omega_S}, \underline{\omega_F}, C; v, \underline{x}, t\right) \qquad (2.21)$$
$$= \max_{\left(\underline{\omega_S}, \underline{\omega_F}, C\right)} \left\{u(C, t) + \left(L^{\underline{\omega_S}, \underline{\omega_F}, C} J\right)(v, \underline{x}, t)\right\},$$

[14] For clarity, all functional dependencies, except time-dependencies, will be suppressed.

[15] $J(v, \underline{x}, t)$ represents the indirect utility function of the representative agent, expressed in terms of the *nominal* wealth. Although the direct utility function is assumed to be state-independent, we can not be sure in saying the same about the indirect utility function because $r(t)$ changes stochastically.

where the Dynkin's operator is equal to[16]

$$\begin{aligned}\left(L^{\underline{\omega_S},\underline{\omega_F},C}J\right)(v,\underline{x},t) &= J_t+\left\{\underline{\omega_S}'\cdot\left[\underline{\mu_S}(t)-r(t)\underline{1}\right]v+\underline{\omega_F}'\cdot\left[\underline{\mu_F}(t)-r(t)\underline{1}\right]v\right.\\ &\quad \left.+r(t)v-p(t)C(t)\right\}J_v+J_{\underline{x}'}\cdot\left(\underline{a}\cdot\underline{x}+\underline{b}\right)\\ &\quad +\frac{v^2J_{vv}}{2}\left[\underline{\omega_S}'\cdot E(t)+\underline{\omega_F}'\cdot H(t)\right]\cdot\left[E(t)'\cdot\underline{\omega_S}+H(t)'\cdot\underline{\omega_F}\right]\\ &\quad +\frac{1}{2}tr\left[J_{\underline{x}\underline{x}'}\cdot\Sigma\cdot V^D(t)\cdot\Sigma'\right]\\ &\quad +v\left[\underline{\omega_S}'\cdot E(t)+\underline{\omega_F}'\cdot H(t)\right]\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot J_{v\underline{x}},\end{aligned}$$

with $J_{\underline{x}}=\frac{\partial J(v,\underline{x},t)}{\partial\underline{x}}$, $J_{\underline{x}\underline{x}'}=\frac{\partial^2 J(v,\underline{x},t)}{\partial\underline{x}\partial\underline{x}'}$, and $J_{v\underline{x}}=\frac{\partial^2 J(v,\underline{x},t)}{\partial v\partial\underline{x}}$, subject to the non-negativity restrictions $\omega_{S_i}\geq 0$ $(i=1,\ldots,m)$ and $C\geq 0$, as well as to the boundary condition $J(v,\underline{x},T)=0$.

Using the Kuhn-Tucker theorem, the necessary and sufficient conditions for the maximization of $\phi\left(\underline{\omega_S},\underline{\omega_F},C;v,\underline{x},t\right)$ are:

$$\phi_C=u_C(t)-p(t)J_v\leq 0, \tag{2.22}$$

$$\left[u_C(t)-p(t)J_v\right]C=0, \tag{2.23}$$

$$\begin{aligned}\underline{\phi_{\underline{\omega_S}}} &= \left[\underline{\mu_S}(t)-r(t)\underline{1}\right]vJ_v+\left[E(t)\cdot E(t)'\cdot\underline{\omega_S}+E(t)\cdot H(t)'\cdot\underline{\omega_F}\right]v^2J_{vv}\\ &\quad +vE(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot J_{v\underline{x}}\\ &\leq \underline{0},\end{aligned} \tag{2.24}$$

$$\underline{\omega_S}'\cdot\underline{\phi_{\underline{\omega_S}}}=0, \tag{2.25}$$

and

$$\begin{aligned}\underline{\phi_{\underline{\omega_F}}} &= \left[\underline{\mu_F}(t)-r(t)\underline{1}\right]vJ_v+\left[H(t)\cdot H(t)'\cdot\underline{\omega_F}+H(t)\cdot E(t)'\cdot\underline{\omega_S}\right]v^2J_{vv}\\ &\quad +vH(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot J_{v\underline{x}}\\ &= \underline{0}.\end{aligned} \tag{2.26}$$

2.2.3 Equilibrium instantaneous nominal risk-free interest rate

As in Cox et al. (1985a), equilibrium is defined by a set of stochastic processes $\left(r(t),\underline{\mu_F}(t);\underline{\omega_S},\underline{\omega_F},C\right)$ satisfying conditions (2.22) to (2.26), as well as the following market clearing conditions:

[16] In order to simplify the notation, subscripts will be used to represent derivatives.

1. In equilibrium, all wealth is invested in the physical production processes, that is $\underline{\omega_S}' \cdot \underline{1} = 1$.

2. In equilibrium, no financial contingent claims are held, i.e. $\underline{\omega_F} = \underline{0}$. That is in equilibrium the net supply or aggregate demand for each financial contingent claim is zero. This is because for each individual who demands some security, there is always another individual that creates and sells it.

The aim of this subsection is to compute, explicitly, an equilibrium formula for $r(t)$, in the context of the Duffie and Kan (1996) model. Initially, a production economy will be used, and the results obtained will be similar to those already generated by Cox et al. (1985a) and Breeden (1986). However, while these authors give equilibrium expressions for the instantaneous *real* risk-free interest rate, here their results will be adapted to the context of a monetary economy. Finally, a one-country pure exchange economy with a cash-in-advance constraint will be used, and a new equilibrium specification for the instantaneous *nominal* riskless interest rate will be obtained.

The production side of the economy: *a la* Cox et al. (1985a)

Imposing the above two market clearing conditions to (2.21), a second version for the HJB equation is obtained:

$$\max_{(\underline{\omega_S},C)} \phi\left(\underline{\omega_S}, C; v, \underline{x}, t\right) = \max_{(\underline{\omega_S},C)} \left\{u(C,t) + \left(L^{\underline{\omega_S},C} J\right)(v, \underline{x}, t)\right\} = 0, \tag{2.27}$$

where

$$\begin{aligned}\left(L^{\underline{\omega_S},C} J\right)(v, \underline{x}, t) &= J_t + \left[v\underline{\omega_S}' \cdot \underline{\mu_S}(t) - p(t)C\right] J_v + J_{\underline{x}'} \cdot \left(a \cdot \underline{x} + \underline{b}\right) \\ &+ \frac{1}{2} tr\left[J_{\underline{x}\underline{x}'} \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right] + \frac{1}{2} v^2 J_{vv} \underline{\omega_S}' \cdot E(t) \cdot E(t)' \cdot \underline{\omega_S} \\ &+ v\underline{\omega_S}' \cdot E(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}},\end{aligned}$$

with $\underline{\omega_S} \geq \underline{0}$, $C \geq 0$, and subject to $J(v, \underline{x}, T) = 0$. Similarly, conditions (2.22) to (2.26) can be rewritten as:

$$\phi_C = u_C(t) - p(t) J_v \leq 0, \tag{2.28}$$

$$\left[u_C(t) - p(t) J_v\right] C = 0, \tag{2.29}$$

$$\phi_{\underline{\omega_S}} = vJ_v \underline{\mu_S}(t) + v^2 J_{vv} E(t) \cdot E(t)' \cdot \underline{\omega_S} + vE(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}} \leq \underline{0}, \tag{2.30}$$

$$\underline{\omega_S}' \cdot \phi_{\underline{\omega_S}} = 0, \tag{2.31}$$

and

$$\underline{\omega_S}' \cdot \underline{1} = 1. \tag{2.32}$$

Following Cox et al. (1985a) and considering the Kuhn-Tucker theorem, conditions (2.30), (2.31), and (2.32) can be rewritten as a quadratic programming problem:

$$\max_{\underline{\omega_S}} \left\{ \underline{\omega_S}' \cdot \left[vJ_v \underline{\mu_S}(t) + \frac{1}{2} v^2 J_{vv} E(t) \cdot E(t)' \cdot \underline{\omega_S} + vE(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}} \right] \right\}, \tag{2.33}$$

subject to $\underline{\omega_S}' \cdot \underline{1} = 1$ and $\underline{\omega_S} \geq \underline{0}$. Moreover, using l as a Lagrange multiplier, problem (2.33) is also equivalent to

$$\begin{aligned} \max_{(\underline{\omega_S}, l)} \quad & \left\{ \underline{\omega_S}' \cdot \left[vJ_v \underline{\mu_S}(t) + \frac{1}{2} v^2 J_{vv} E(t) \cdot E(t)' \cdot \underline{\omega_S} + vE(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}} \right] \right. \\ & \left. - l \left(\underline{\omega_S}' \cdot \underline{1} - 1 \right) \right\}, \end{aligned} \tag{2.34}$$

subject to $\underline{\omega_S} \geq \underline{0}$. The corresponding Kuhn-Tucker conditions are given by

$$vJ_v^* \underline{\mu_S}(t) + v^2 J_{vv}^* E(t) \cdot E(t)' \cdot \underline{\omega_S}^* + vE(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}}^* - l\underline{1} \leq \underline{0}, \tag{2.35}$$

and

$$\begin{aligned} 0 \quad = \quad & vJ_v^* \left(\underline{\omega_S}^* \right)' \cdot \underline{\mu_S}(t) + v^2 J_{vv}^* \left(\underline{\omega_S}^* \right)' \cdot E(t) \cdot E(t)' \cdot \underline{\omega_S}^* \\ & + v \left(\underline{\omega_S}^* \right)' \cdot E(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}}^* - l \left(\underline{\omega_S}^* \right)' \cdot \underline{1}, \end{aligned} \tag{2.36}$$

where $\underline{\omega_S}^*$ denotes the optimal value of $\underline{\omega_S}$, and J^* represents the indirect utility function obtained at $\underline{\omega_S} = \underline{\omega_S}^*$.

On the other hand, if it is assumed that $\underline{\omega_F} = \underline{0}$ without considering, for the moment, that $\underline{\omega_S}' \cdot \underline{1} = 1$, instead of restrictions (2.30), (2.31), and (2.32), we would have to deal with the following two conditions:

$$\begin{aligned} \phi_{\underline{\omega_S}} \quad = \quad & vJ_v \left[\underline{\mu_S}(t) - r(t)\underline{1} \right] + v^2 J_{vv} E(t) \cdot E(t)' \cdot \underline{\omega_S} \\ & + vE(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}} \\ \leq \quad & \underline{0}, \end{aligned} \tag{2.37}$$

and

$$\underline{\omega_S}' \cdot \phi_{\underline{\omega_S}} = 0. \tag{2.38}$$

But, these last two restrictions are equivalent to another quadratic programming problem:

$$\max_{\underline{\omega_S}} \quad \left\{\underline{\omega_S}' \cdot \left[vJ_v\left(\underline{\mu_S}(t) - r(t)\,\underline{1}\right) + \frac{1}{2}v^2 J_{vv} E(t)\cdot E(t)' \cdot \underline{\omega_S} \right.\right. \tag{2.39}$$
$$\left.\left. + vE(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot J_{v\underline{x}}\right]\right\},$$

subject to $\underline{\omega_S} \geq \underline{0}$, with the associated Kuhn-Tucker conditions given by

$$0 \geq vJ_v^{**}\underline{\mu_S}(t) + v^2 J_{vv}^{**} E(t)\cdot E(t)'\cdot \underline{\omega_S}^{**} + vE(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot J_{v\underline{x}}^{**} \tag{2.40}$$
$$-vJ_v^{**} r(t)\,\underline{1},$$

and

$$0 = vJ_v^{**}\left(\underline{\omega_S}^{**}\right)'\cdot\underline{\mu_S}(t) + v^2 J_{vv}^{**}\left(\underline{\omega_S}^{**}\right)'\cdot E(t)\cdot E(t)'\cdot\underline{\omega_S}^{**} \tag{2.41}$$
$$+v\left(\underline{\omega_S}^{**}\right)'\cdot E(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot J_{v\underline{x}}^{**} - vJ_v^{**} r(t)\left(\underline{\omega_S}^{**}\right)'\cdot\underline{1},$$

where $\underline{\omega_S}^{**}$ denotes the new optimal value of $\underline{\omega_S}$, and J^{**} represents the indirect utility function obtained at $\underline{\omega_S} = \underline{\omega_S}^{**}$.

Comparing (2.35)-(2.36) with (2.40)-(2.41), it follows that if $J^{**} = J^{*}$ and $vJ_v^{**} r(t) = l$, then $\underline{\omega_S}^{**} = \underline{\omega_S}^{*}$. Therefore

$$r(t) = \frac{l}{vJ_v}. \tag{2.42}$$

Solving (2.36) for l, and since $\left(\underline{\omega_S}\right)'\cdot\underline{1} = 1$,

$$l = vJ_v\underline{\omega_S}'\cdot\underline{\mu_S}(t) + v^2 J_{vv}\underline{\omega_S}'\cdot E(t)\cdot E(t)'\cdot\underline{\omega_S} \tag{2.43}$$
$$+v\underline{\omega_S}'\cdot E(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot J_{v\underline{x}}.$$

Finally, combining (2.42) and (2.43),

$$r(t) = \underline{\omega_S}'\cdot\underline{\mu_S}(t) + v\left(\frac{J_{vv}}{J_v}\right)\underline{\omega_S}'\cdot E(t)\cdot E(t)'\cdot\underline{\omega_S} + \underline{\omega_S}'\cdot E(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot\left(\frac{J_{v\underline{x}}}{J_v}\right). \tag{2.44}$$

Equation (2.44) expresses $r(t)$ as a function of the indirect utility. This solution is similar to equation (14) of Cox et al. (1985a) and to equation (15) of Breeden (1986). However, it is not exactly equivalent since these last two equations give the equilibrium value of the short-term *real* (not *nominal*) interest rate, which is stated in terms of the *real* wealth, because both models use the single physical good as the numeraire.

Next, $r(t)$ will be derived as an explicit function of the utility of consumption, and no longer as a function of the utility of wealth.

The consumption side of the economy: *a la* Breeden (1986)

To prove that the equilibrium instantaneous interest rate is equal to minus the expected rate of change in the marginal utility of *nominal* wealth, the approach followed by Cox et al. (1985a) can be used.

Considering that in equilibrium all wealth is invested in the physical production processes, the budget constraint (2.20) is given by

$$dv = \left[v\underline{\omega_S}' \cdot \underline{\mu_S}(t) - p(t)C(t)\right]dt + v\underline{\omega_S}' \cdot E(t) \cdot dW^{\mathcal{P}}(t), \tag{2.45}$$

and

$$\begin{aligned}(LJ)(v,\underline{x},t) &= J_t + \left[v\underline{\omega_S}' \cdot \underline{\mu_S}(t) - p(t)C(t)\right]J_v + J_{\underline{x}'} \cdot (\underline{a} \cdot \underline{x} + \underline{b}) \\ &\quad + \frac{1}{2}v^2 J_{vv}\underline{\omega_S}' \cdot E(t) \cdot E(t)' \cdot \underline{\omega_S} + \frac{1}{2}tr\left[J_{\underline{x}\underline{x}'} \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right] \\ &\quad + v\underline{\omega_S}' \cdot E(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}}.\end{aligned}$$

Applying Itô's lemma to $J_v(v,\underline{x},t)$,

$$dJ_v = \mu_{J_v}(t)\,dt + \left[vJ_{vv}\underline{\omega_S}' \cdot E(t) + (J_{v\underline{x}})' \cdot \Sigma \cdot \sqrt{V^D(t)}\right] \cdot dW^{\mathcal{P}}(t), \tag{2.46}$$

where

$$\begin{aligned}\mu_{J_v}(t) &= (LJ_v)(v,\underline{x},t) \qquad\qquad (2.47)\\ &= J_{vt} + \left[v\underline{\omega_S}' \cdot \underline{\mu_S}(t) - p(t)C(t)\right]J_{vv} + J_{v\underline{x}'} \cdot (\underline{a} \cdot \underline{x} + \underline{b}) \\ &\quad + \frac{1}{2}v^2 J_{vvv}\underline{\omega_S}' \cdot E(t) \cdot E(t)' \cdot \underline{\omega_S} + \frac{1}{2}tr\left[J_{v\underline{x}\underline{x}'} \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right] \\ &\quad + v\underline{\omega_S}' \cdot E(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{vv\underline{x}}.\end{aligned}$$

But, in equilibrium $\frac{\partial\phi}{\partial v} = 0$, that is

$$\begin{aligned}0 &= J_{tv} + \underline{\omega_S}' \cdot \underline{\mu_S}(t)J_v + \left[v\underline{\omega_S}' \cdot \underline{\mu_S}(t) - p(t)C(t)\right]J_{vv} \qquad (2.48)\\ &\quad + J_{\underline{x}'v} \cdot (\underline{a} \cdot \underline{x} + \underline{b}) + vJ_{vv}\underline{\omega_S}' \cdot E(t) \cdot E(t)' \cdot \underline{\omega_S} \\ &\quad + \frac{1}{2}v^2 J_{vvv}\underline{\omega_S}' \cdot E(t) \cdot E(t)' \cdot \underline{\omega_S} + \frac{1}{2}tr\left[J_{\underline{x}\underline{x}'v} \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right] \\ &\quad + \underline{\omega_S}' \cdot E(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}} + v\underline{\omega_S}' \cdot E(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}v}.\end{aligned}$$

Hence, combining equations (2.47) and (2.48),

$$\mu_{J_v}(t) = -\underline{\omega_S}' \cdot \underline{\mu_S}(t)J_v - vJ_{vv}\underline{\omega_S}' \cdot E(t) \cdot E(t)' \cdot \underline{\omega_S} - \underline{\omega_S}' \cdot E(t) \cdot \sqrt{V^D(t)} \cdot \Sigma' \cdot J_{v\underline{x}}.$$

Consequently,

$$\begin{aligned}
-E\left(\frac{dJ_v}{J_v}\right)\frac{1}{dt} &= -\frac{\mu_{J_v}}{J_v} \\
&= \underline{\omega_S}'\cdot\underline{\mu_S}(t)+v\left(\frac{J_{vv}}{J_v}\right)\underline{\omega_S}'\cdot E(t)\cdot E(t)'\cdot\underline{\omega_S} \\
&\quad+\underline{\omega_S}'\cdot E(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot\left(\frac{J_{v\underline{x}}}{J_v}\right).
\end{aligned}$$

Comparing the above equality with equation (2.44), it follows that

$$r(t)=-\frac{\mu_{J_v}(t)}{J_v(v,\underline{x},t)}, \tag{2.49}$$

as expected.

On the other hand, if we consider condition (2.23), while assuming that $C\neq 0$, then

$$J_v(v,\underline{x},t)=\frac{u_C(t)}{p(t)}. \tag{2.50}$$

Using Itô's lemma,

$$\begin{aligned}
\mu_{J_v}(t) &= \frac{1}{p(t)}\mu_{u_C}(t)-\frac{u_C(t)}{p(t)^2}\mu_p(t)+\frac{1}{2}\frac{2u_C(t)p(t)}{p(t)^4}\underline{\sigma_p}(t)'\cdot\underline{\sigma_p}(t) \\
&\quad-\frac{1}{p(t)^2}COV\left[du_C(t),dp(t)\right],
\end{aligned} \tag{2.51}$$

where

$$du_C(t)=\mu_{u_C}(t)\,dt+\underline{\sigma_{u_C}}(t)'\cdot d\underline{W}^P(t),$$

$u_C(t)$ is the time-t marginal utility of consumption, with $\mu_{u_C}(t)\in\Re$ and $\underline{\sigma_{u_C}}(t)\in\Re^n$,

$$dp(t)=\mu_p(t)\,dt+\underline{\sigma_p}(t)'\cdot d\underline{W}^P(t),$$

$\frac{\mu_p(t)}{p(t)}\in\Re$ represents the time-t expected rate of inflation, and $\underline{\sigma_p}(t)\in\Re^n$. Combining equations (2.49), (2.50), and (2.51):

$$r(t)=-\frac{\mu_{u_C}(t)}{u_C(t)}+\left\{\frac{\mu_p(t)}{p(t)}-\frac{\underline{\sigma_p}(t)'\cdot\underline{\sigma_p}(t)}{p(t)^2}+\frac{COV\left[du_C(t),dp(t)\right]}{u_C(t)\,p(t)}\right\}. \tag{2.52}$$

From Breeden (1986, equation 19), it is known that the first term on the right-hand-side of the previous equation represents the time-t ***real*** risk-free instantaneous interest rate, which will be denominated by $k(t)$. In order to compute $\left[-\frac{\mu_{u_C}(t)}{u_C(t)}\right]$ explicitly, the following stochastic differential equation for aggregate consumption will be considered:

$$\frac{dC(t)}{C(t)} = \mu_C(t)\,dt + \underline{\sigma_C}(t)' \cdot d\underline{W^{\mathcal{P}}}(t),$$

where $\mu_C(t) \in \Re$ and $\underline{\sigma_C}(t) \in \Re^n$. Applying Itô's lemma to the marginal utility of consumption, it is possible to derive the functional form of its drift:

$$\mu_{u_C} = u_{CC}(t)\,\mu_C(t)\,C(t) + u_{Ct}(t) + \frac{1}{2}u_{CCC}(t)\,\underline{\sigma_C}(t)' \cdot \underline{\sigma_C}(t)\,C(t)^2. \tag{2.53}$$

Substituting (2.53) into the first term in the right-hand-side of (2.52), one obtains the consumption-based equilibrium equation (22) of Breeden (1986) for the *real* risk-free instantaneous interest rate:

$$\begin{aligned} k(t) &= -\frac{\mu_{u_C}(t)}{u_C(t)} \qquad (2.54)\\ &= -\frac{u_{Ct}(t)}{u_C(t)} - \frac{C(t)\,u_{CC}(t)}{u_C(t)}\mu_C(t) - \frac{1}{2}\frac{C^2(t)\,u_{CCC}(t)}{u_C(t)}\left[\underline{\sigma_C}(t)' \cdot \underline{\sigma_C}(t)\right]. \end{aligned}$$

From now on, it will be considered, as an additional assumption, that the preferences are time-separable, i.e.

A.12) $u(C,t) = e^{-\rho t}U(C)$, where ρ is the constant discount factor or time-preference parameter, $U_C > 0$ (nonsatiation assumption), and $U_{CC} < 0$ (risk aversion assumption).

Therefore, $-\frac{u_{Ct}(t)}{u_C(t)} = -\frac{-\rho e^{-\rho t}U_C}{e^{-\rho t}U_C} = \rho$, and combining equations (2.52) and (2.54),

$$r(t) = k(t) + \left\{\frac{\mu_p(t)}{p(t)} - \frac{\underline{\sigma_p}(t)' \cdot \underline{\sigma_p}(t)}{p(t)^2} + \frac{COV\left[du_C(t), dp(t)\right]}{u_C(t)\,p(t)}\right\}, \tag{2.55}$$

with

$$k(t) = \rho - \frac{C(t)\,u_{CC}(t)}{u_C(t)}\mu_C(t) - \frac{1}{2}\frac{C^2(t)\,u_{CCC}(t)}{u_C(t)}\left[\underline{\sigma_C}(t)' \cdot \underline{\sigma_C}(t)\right]. \tag{2.56}$$

The above expression for $r(t)$ is distinct both from equation (45) of Heston (1988) and from equation (69) of Cox et al. (1985b). In opposition with Heston (1988), equation (2.55) does not correspond to the well known *Fisher identity*, because

$$\begin{aligned} \frac{COV\left[du_C(t), dp(t)\right]}{u_C(t)\,p(t)} &= \frac{\underline{\sigma_{u_C}}(t)' \cdot \underline{\sigma_p}(t)}{u_C(t)\,p(t)}\\ &= \frac{C(t)\,u_{CC}(t)}{u_C(t)\,p(t)}\underline{\sigma_C}(t)' \cdot \underline{\sigma_p}(t)\\ &= \frac{C(t)\,u_{CC}(t)}{u_C(t)\,p(t)}COV\left[\frac{dC(t)}{C(t)}, dp(t)\right]\\ &\neq 0, \end{aligned}$$

that is because we are not assuming *money neutrality*[17]. In fact, Sun (1992) found a significant correlation between the price level and the growth rate of consumption, which does not support the *money neutrality* assumption. On the other hand, equation (2.55) shows two important differences when compared to equation (60) of Cox et al. (1985b). Firstly, because equation (2.55) is expressed in terms of the direct utility function, and not in terms of the utility of wealth. Secondly, because in equation (2.55) both the price level, $p(t)$, and the expected rate of inflation, $\frac{\mu_p(t)}{p(t)}$, are endogenously determined, and thus one can be sure that they will be consistent with our general equilibrium framework.

A one-country pure exchange economy

Assuming A.5, and since in equilibrium $\underline{\omega_S}' \cdot \underline{1} = 1$ as well as $\underline{\omega_F} = \underline{0}$, we move from a production economy to a Lucas (1978) type of pure exchange economy where all output is consumed, that is $C(t) = q(t)$.[18] Hence, equations (2.55) and (2.56) can be stated in terms of the exogenous aggregate output, which means that it is not necessary to solve the HJB equation (2.21) for the endogenous consumption process:

$$r(t) = k(t) + \left[\frac{\mu_p(t)}{p(t)} - \frac{\underline{\sigma_p}(t)' \cdot \underline{\sigma_p}(t)}{p(t)^2} + \frac{q(t)\, u_{qq}(t)}{u_q(t)} \underline{\sigma_q}(t)' \cdot \frac{\underline{\sigma_p}(t)}{p(t)}\right], \tag{2.57}$$

with

$$k(t) = \rho - \frac{q(t)\, u_{qq}(t)}{u_q(t)} \mu_q(t) - \frac{1}{2} \frac{q(t)^2\, u_{qqq}(t)}{u_q(t)} \left[\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right]. \tag{2.58}$$

Equation (2.58) corresponds to equation (11) of Bakshi and Chen (1997a). The next theorem rewrites the above equilibrium solution for the *nominal* short-term interest rate only in terms of the exogenous output and money supply processes.

Theorem 1 *In equilibrium, the instantaneous nominal interest rate is*

$$\begin{aligned} r(t) = \quad & \left[\rho + \mu_M(t) - \mu_q(t) - \underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t) + \sigma_{q,M}(t)\right] \\ & - \frac{q(t)\, u_{qq}(t)}{u_q(t)} \left[\mu_q(t) - \sigma_{q,M}(t) + \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right] \\ & - \frac{1}{2} \frac{q(t)^2\, u_{qqq}(t)}{u_q(t)} \left[\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right], \end{aligned} \tag{2.59}$$

where $\sigma_{q,M}(t) = COV\left[\frac{dq(t)}{q(t)}, \frac{dM(t)}{M(t)}\right]$.

---

[17] I.e. it is not assumed that the price level has no effect on the real side of the economy.

[18] Both types of economy can be made compatible through the definition of $\underline{\mu_S}(q, M, \underline{S}, \underline{X}, t)$ and $E(q, M, \underline{S}, \underline{X}, t)$ in such a way that the production economy generates an endogenous consumption process identical to the exogenously specified output process. See Heston (1988, footnote 9) or Bakshi and Chen (1997a, footnote 5).

Proof. Applying Itô's lemma to $p(t) = \frac{M(t)}{q(t)}$, all terms in equation (2.57) can be expressed as functions of only $q(t)$ and $M(t)$:

$$\mu_p(t) = \frac{\mu_M(t)\,M(t)}{q(t)} - \frac{M(t)\,\mu_q(t)}{q(t)} + \frac{M(t)\,\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)}{q(t)} - \frac{\underline{\sigma_q}(t)' \cdot \underline{\sigma_M}(t)\,M(t)}{q(t)},$$

$$\frac{\mu_p(t)}{p(t)} = \mu_M(t) - \mu_q(t) + \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t) - \underline{\sigma_q}(t)' \cdot \underline{\sigma_M}(t),$$

$$\underline{\sigma_p}(t) = \frac{1}{q(t)}\underline{\sigma_M}(t)\,M(t) - \frac{M(t)}{q(t)^2}\underline{\sigma_q}(t)\,q(t),$$

$$\frac{\underline{\sigma_p}(t)' \cdot \underline{\sigma_p}(t)}{p(t)^2} = \left[\underline{\sigma_M}(t) - \underline{\sigma_q}(t)\right]' \cdot \left[\underline{\sigma_M}(t) - \underline{\sigma_q}(t)\right],$$

and

$$\frac{\underline{\sigma_q}(t)' \cdot \underline{\sigma_p}(t)}{p(t)} = \underline{\sigma_q}(t)' \cdot \left[\underline{\sigma_M}(t) - \underline{\sigma_q}(t)\right].$$

Hence,

$$\begin{aligned} r(t) \;=\;& \rho + \left[\mu_M(t) - \mu_q(t) + \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t) - \underline{\sigma_q}(t)' \cdot \underline{\sigma_M}(t)\right] \\ & - \left[\underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t) - 2\underline{\sigma_q}(t)' \cdot \underline{\sigma_M}(t) + \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right] \\ & + \frac{q(t)\,u_{qq}(t)}{u_q(t)}\left[\mu_q(t) - \underline{\sigma_q}(t)' \cdot \underline{\sigma_M}(t) + \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right] \\ & - \frac{1}{2}\frac{q(t)^2\,u_{qqq}(t)}{u_q(t)}\left[\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right], \end{aligned}$$

which yields equation (2.59) after collecting alike terms. Following Bakshi and Chen (1997b, pages 818-819), an alternative derivation of equation (2.59) is presented in appendix 2.4.1. ■

According to (2.59), the short-term *nominal* equilibrium interest rate is increasing in: the time-preference parameter; the expected growth rate of money supply; the expected rate of change in the aggregate output (if the coefficient of relative risk aversion is greater than one); and in the volatility of the aggregate output growth rate. On the other hand, $r(t)$ is decreasing in: the volatility of the money supply growth rate; and in the covariance between the growth rates of aggregate output and money supply (again, if $-\frac{q(t)u_{qq}(t)}{u_q(t)} > 1$).

Equations (2.44) and (2.59) generate the same term structure of interest rates, because they must hold simultaneously in equilibrium. However, the use of equation (2.44) requires the existence of a closed-form solution for the indirect utility function, which has to be obtained by solving the HJB equation (2.21), or requires the assumption of restrictive preferences: namely, the use of a log utility function, as is the case in Cox et al. (1985a) and Longstaff and Schwartz (1992a). Consequently, we will try to fit the Duffie and Kan (1996) model into a general equilibrium framework with more realistic assumptions about preferences than those implied by a Bernoulli logarithmic utility function, through the use of equation (2.59) instead of equation (2.44). In fact, it turns out to be easy to work with equation (2.59) since the stochastic processes for the aggregate output and for the money supply can be exogenously specified in a suitable fashion.

2.2.4 Equilibrium Factor Risk Premiums

In order to fit the Duffie and Kan (1996) model into a general equilibrium framework, it is necessary to prove that our general equilibrium assumptions imply an affine form for $r(t)$ -as in equation (2.2)- and a risk-adjusted process for $\underline{X}(t)$ equivalent to the stochastic differential equation (2.5). But, to derive the equilibrium risk-neutral process for the model' factors (that is consistent with our general equilibrium setup), we first have to compute the risk premiums associated with each one of the non-traded state variables. Only after having derived such factor risk premiums, it is then possible to specify the equilibrium risk-adjusted drift for $d\underline{X}(t)$, by applying Girsanov's theorem or by obtaining the PDE that must be satisfied, in equilibrium, by any interest rate contingent claim.

In equilibrium, since $\underline{\omega_F} = \underline{0}$, equation (2.26) becomes:

$$vJ_v\left[\underline{\mu_F}(t) - r(t)\,\underline{1}\right] + v^2 J_{vv} H(t)\cdot E(t)'\cdot\underline{\omega_S} + vH(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot J_{v\underline{x}} = \underline{0},$$

that is

$$\left[\underline{\mu_F}(t) - r(t)\,\underline{1}\right] = -v\left(\frac{J_{vv}}{J_v}\right)H(t)\cdot E(t)'\cdot\underline{\omega_S} - H(t)\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot\left(\frac{J_{v\underline{x}}}{J_v}\right).$$

Both sides of the above equation are $n\times 1$ matrices. Taking just their i^{th}-row,

$$\mu_{F_i}(t) - r(t) = -v\left(\frac{J_{vv}}{J_v}\right)\underline{h_i}(t)'\cdot E(t)'\cdot\underline{\omega_S} - \underline{h_i}(t)'\cdot\sqrt{V^D(t)}\cdot\Sigma'\cdot\left(\frac{J_{v\underline{x}}}{J_v}\right), \tag{2.60}$$

where $\mu_{F_i}(t)$ is the expected *nominal time-t* rate of return on the i^{th} financial contingent claim, $[\mu_{F_i}(t) - r(t)]$ represents the equilibrium expected excess *nominal* rate of return (over the risk-free interest rate) generated by the i^{th} financial contingent claim, and $\underline{h_i}(t)'$ is the i^{th}-row of matrix $H(t)$.

In order to obtain $\underline{h_i}(t)$ explicitly, Itô's lemma will be applied to the value of the i^{th} financial contingent claim, $F_i(\underline{x},t)$:[19]

$$dF_i(t) = F_i(t)\,\mu_{F_i}(t)\,dt + \frac{\partial F_i(t)}{\partial \underline{x}'}\cdot\Sigma\cdot\sqrt{V^D(t)}\cdot d\underline{W^P}(t). \tag{2.61}$$

[19] It is assumed that the contractual terms of the financial contingent claim do not depend explicitly on wealth. And, again, only time-dependencies will be retained.

Comparing equations (2.14) and (2.61), it follows that

$$\begin{aligned} F_i(t)\,\underline{h_i}(t)' &= \frac{\partial F_i(t)}{\partial \underline{x}'} \cdot \Sigma \cdot \sqrt{V^D(t)} \\ &\equiv \underline{\sigma_{F_i}}(t)'. \end{aligned}$$

Thus, equation (2.60) is equivalent to:

$$\begin{aligned} \left[\mu_{F_i}(t) - r(t)\right] F_i(t) &= -v\left(\frac{J_{vv}}{J_v}\right)\frac{\partial F_i(t)}{\partial \underline{x}'} \cdot \Sigma \cdot \sqrt{V^D(t)} \cdot E(t)' \cdot \underline{\omega_S} \\ &\quad - \frac{\partial F_i(t)}{\partial \underline{x}'} \cdot \Sigma \cdot V^D(t) \cdot \Sigma' \cdot \left(\frac{J_{v\underline{x}}}{J_v}\right). \end{aligned} \tag{2.62}$$

On the other hand, equation (2.61) and the stochastic process (2.46) followed by the marginal utility of wealth, imply that

$$\begin{aligned} & COV\left[dF_i(\underline{x}, t), dJ_v(v, \underline{x}, t)\right] \\ = & \frac{\partial F_i(t)}{\partial \underline{x}'} \cdot \Sigma \cdot \sqrt{V^D(t)} \cdot \left[v J_{vv} \underline{\omega_S}' \cdot E(t) + \left(J_{v\underline{x}}\right)' \cdot \Sigma \cdot \sqrt{V^D(t)}\right]' \\ = & v J_{vv} \frac{\partial F_i(t)}{\partial \underline{x}'} \cdot \Sigma \cdot \sqrt{V^D(t)} \cdot E(t)' \cdot \underline{\omega_S} + \frac{\partial F_i(t)}{\partial \underline{x}'} \cdot \Sigma \cdot V^D(t) \cdot \Sigma' \cdot J_{v\underline{x}}. \end{aligned} \tag{2.63}$$

Comparing equations (2.62) and (2.63), a similar result to Cox et al. (1985a, equation 27) is obtained, the only difference being the fact that we are now considering expected excess equilibrium *nominal* returns instead of *real* ones:

$$\left[\mu_{F_i}(t) - r(t)\right] F_i(t) = -\frac{1}{J_v(v, \underline{x}, t)} COV\left[dF_i(\underline{x}, t), dJ_v(v, \underline{x}, t)\right]. \tag{2.64}$$

However, since a solution for the indirect utility function is not available, the above expression is of little practical use. In order to compute the equilibrium risk premiums required for the i^{th} financial contingent claim, as a function of estimable parameters, it is necessary to convert the right-hand-side of equation (2.64) in terms of the exogenously specified output and money supply processes. This is accomplished by the following theorem.

Theorem 2 *In equilibrium, the factor risk premiums on any financial contingent claim $F(t)$ satisfy*

$$\begin{aligned} \left[\mu_F(t) - r(t)\right] F(t) = &\; -\left[1 + \frac{q(t)\, u_{qq}(t)}{u_q(t)}\right] COV\left[dF(t), \frac{dq(t)}{q(t)}\right] \\ &+ COV\left[dF(t), \frac{dM(t)}{M(t)}\right]. \end{aligned} \tag{2.65}$$

Proof. As a first step, condition $J_v(v, \underline{x}, t) = \frac{u_C(t)}{p(t)}$ implies that equation (2.64) can be rewritten as

$$[\mu_{F_i}(t) - r(t)] F_i(t) = -\frac{p(t)}{u_C(t)} COV\left[dF_i(t), d\left(\frac{u_C}{p}\right)(t)\right].$$

From Itô's lemma, the diffusion of the stochastic process $d\left(\frac{u_C}{p}\right)(t)$ is given by $\frac{1}{p(t)}\underline{\sigma_{u_C}}(t)' - \frac{u_C(t)}{p(t)^2}\underline{\sigma_p}(t)'$, and therefore

$$[\mu_{F_i}(t) - r(t)] F_i(t) = -\frac{1}{u_C(t)}\underline{\sigma_{F_i}}(t)' \cdot \underline{\sigma_{u_C}}(t) + \frac{1}{p(t)}\underline{\sigma_{F_i}}(t)' \cdot \underline{\sigma_p}(t).$$

Applying again Itô's lemma while considering equations (2.10) and (2.45), it follows that

$$\underline{\sigma_{u_C}}(t)' = u_{CC}(t)\left[vC_v(t)\underline{\omega_S}' \cdot E(t) + \frac{\partial C(t)}{\partial \underline{x}'} \cdot \Sigma \cdot \sqrt{V^D(t)}\right],$$

and

$$\underline{\sigma_C}(t)' = \frac{vC_v(t)\underline{\omega_S}' \cdot E(t) + \frac{\partial C(t)}{\partial \underline{x}'} \cdot \Sigma \cdot \sqrt{V^D(t)}}{C(t)}.$$

Hence $\underline{\sigma_{u_C}}(t) = C(t)\, u_{CC}(t)\, \underline{\sigma_C}(t)$, and because $C(t) = q(t)$, then

$$[\mu_{F_i}(t) - r(t)] F_i(t) = -\frac{q(t)\, u_{qq}(t)}{u_q(t)}\underline{\sigma_{F_i}}(t)' \cdot \underline{\sigma_q}(t) + \frac{1}{p(t)}\underline{\sigma_{F_i}}(t)' \cdot \underline{\sigma_p}(t).$$

Moreover, since $\underline{\sigma_p}(t) = p(t)\,\underline{\sigma_M}(t) - p(t)\,\underline{\sigma_q}(t)$,

$$[\mu_{F_i}(t) - r(t)] F_i(t) = -\frac{q(t)\, u_{qq}(t)}{u_q(t)}\underline{\sigma_{F_i}}(t)' \cdot \underline{\sigma_q}(t) + \underline{\sigma_{F_i}}(t)' \cdot \underline{\sigma_M}(t) - \underline{\sigma_{F_i}}(t)' \cdot \underline{\sigma_q}(t).$$

Finally, applying the above equation to a general financial contingent claim with a value of $F(t)$ and an expected rate of return of $\mu_F(t)$, the equilibrium solution (2.65) follows. An alternative derivation is provided in appendix 2.4.2. ■

Thus, in order to find the equilibrium factor risk premiums (as well as the instantaneous *nominal* spot equilibrium interest rate) for the Duffie and Kan (1996) model, it is just necessary to specify an utility function as well as suitable output and money supply stochastic processes.

Before proceeding, three remarks should be made. First, equation (2.65) implies that the factor risk premiums are increasing in the conditional covariance of the contingent claim value with: *i)* the rate of change in the aggregate output (if the coefficient of relative risk aversion is greater than one); and, with *ii)* the growth rate of money supply. In other words, equation (2.65) shows that both "production risk" (i.e. technological shocks) and "monetary risk" (that is, inflationary shocks) matter. Second, from Cox et al. (1985a, equation 30) or from equation (9) of Bakshi and Chen (1997a), it is well known that the equilibrium expected excess *real* rate of return is equal to $\left(-\frac{q(t)u_{qq}(t)}{u_q(t)}\right) COV\left[dF(t), \frac{dq(t)}{q(t)}\right]$. Subtracting this "*real* risk" compensation from equation (2.65), we can now conclude that the equilibrium compensation for "*nominal* risk" must be given by $COV\left[dF(t), \frac{dM(t)}{M(t)}\right] - COV\left[dF(t), \frac{dq(t)}{q(t)}\right]$. Thirdly, equation (2.65) also shows that even in a risk-neutral economy -where $-\frac{q(t)u_{qq}(t)}{u_q(t)} = 0$, i.e. with a linear utility function- the equilibrium expected excess *nominal* rate of return on a financial contingent claim would still be non-zero (unless $COV\left[dF(t), \frac{dM(t)}{M(t)}\right] = COV\left[dF(t), \frac{dq(t)}{q(t)}\right]$). This means that in order to derive a Duffie and Kan (1996) model specification under the original probability measure $\mathcal{P}$ that is compatible with the specification given by the authors under the equivalent martingale measure $\mathcal{Q}$, it would be unrealistic to assume a zero or constant vector of market prices of risk, since such assumption would most probably be inconsistent with our general equilibrium setup.

2.2.5 The Duffie and Kan (1996) model in a constant relative risk aversion economy

An economy with a power utility function

In order to obtain the Duffie and Kan (1996) model from our general equilibrium framework, assumptions A.5, A.6, and A.12 must be further specialized.

Now an economy with decreasing absolute risk aversion will be considered, and more specifically, a power utility function will be used to characterize the preferences of the representative investor. Hence, assumption A.12 is specialized into:

A.12')

$$u(C,t) = e^{-\rho t}\frac{C^{\gamma}-1}{\gamma}, \tag{2.66}$$

where $\gamma < 1$ (and thus $u_{CC}(t) < 0$), $\gamma \neq 0$, and $(1-\gamma)$ is the Pratt's measure of relative risk aversion.[20]

Since $C(t) = q(t)$, and using (2.66), then $u(q,t) = e^{-\rho t}\frac{q^{\gamma}-1}{\gamma}$,

$$-\frac{q(t)\,u_{qq}(t)}{u_q(t)} = 1-\gamma, \tag{2.67}$$

and

$$\frac{q(t)^2\,u_{qqq}(t)}{u_q(t)} = (\gamma-1)(\gamma-2). \tag{2.68}$$

[20] $-\frac{C(t)u_{CC}(t)}{u_C(t)} = 1-\gamma$, i.e. constant relative risk aversion is being assumed.

The choice of the utility function under use was not intended to be the most general one possible but rather as general as necessary to nest, as special cases, all the affine general equilibrium interest rate frameworks presented so far in the literature (which are invariably based on the more restrictive log utility function). Nevertheless, it can be easily shown that the power utility function considered hereafter is the most general specification, under the hyperbolic absolute risk aversion class,[21] that generates constant (i.e. output independent) values for both quantities $-\frac{q(t)u_{qq}(t)}{u_q(t)}$ and $\frac{q(t)^2 u_{qqq}(t)}{u_q(t)}$ appearing in expressions (2.59) and (2.65), and therefore that supports the Duffie and Kan (1996) model under an affine specification for both the drifts and the instantaneous variances of the aggregate output and money supply processes.

In order to derive a Duffie and Kan (1996) model from our general equilibrium setup, the stochastic processes for the aggregate output and for the money supply (i.e. the functional form of $\mu_q(t)$, $\underline{\sigma_q}(t)$, $\mu_M(t)$, and $\underline{\sigma_M}(t)$) must be defined in such a way that two conditions are met: $i)$ $r(t)$ must be an affine function of the state variables; and $ii)$ $\underline{\mu}[\underline{X}(t)]$ must also be affine.

From theorem 1, condition $i)$ implies that $\mu_M(t)$, $\mu_q(t)$, $\left[\underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t)\right]$, $\left[\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right]$, and $\sigma_{q,M}(t)$ must all be affine functions of $\underline{X}(t)$. So, the drifts of the stochastic processes (2.12) and (2.13) can be defined as:

$$\mu_q(t) = \eta + \underline{\theta}' \cdot \underline{X}(t), \tag{2.69}$$

and

$$\mu_M(t) = \pi + \underline{\phi}' \cdot \underline{X}(t), \tag{2.70}$$

where $\eta, \pi \in \Re$, and $\underline{\theta}, \underline{\phi} \in \Re^n$.

Considering condition $ii)$, since $\underline{\mu}[\underline{X}(t)] = \underline{v}[\underline{X}(t)] - \sigma[\underline{X}(t)] \cdot \underline{\Lambda}[\underline{X}(t)]$ and because $\underline{v}[\underline{X}(t)]$ is defined by equation (2.10) as an affine function of the state vector, then $\underline{\mu}[\underline{X}(t)]$ can only be affine if $\sigma[\underline{X}(t)] \cdot \underline{\Lambda}[\underline{X}(t)]$ is also affine. But, because $[\mu_F(t) - r(t)]\, F(t) = \underline{\sigma_F}(t)' \cdot \underline{\Lambda}[\underline{X}(t)]$, with $\underline{\sigma_F}(t)' = \frac{\partial F(t)}{\partial \underline{x}'} \cdot \sigma[\underline{X}(t)]$, and

$$[\mu_F(t) - r(t)]\, F(t) = \frac{\partial F(t)}{\partial \underline{x}'} \cdot \left[-\gamma \Sigma \cdot \sqrt{V^D(t)} \cdot \underline{\sigma_q}(t) + \Sigma \cdot \sqrt{V^D(t)} \cdot \underline{\sigma_M}(t)\right], \tag{2.71}$$

[21] Which, accordingly to Ingersoll (1987, equation 51), can be summarized as

$$u(q,t) = e^{-\rho t} \frac{1-\gamma}{\gamma} \left(\frac{aq}{1-\gamma} + b\right)^{\gamma}, \; b > 0.$$

then

$$\sigma[\underline{X}(t)] \cdot \underline{\Lambda}[\underline{X}(t)] = -\gamma\Sigma \cdot \sqrt{V^D(t)} \cdot \underline{\sigma_q}(t) + \Sigma \cdot \sqrt{V^D(t)} \cdot \underline{\sigma_M}(t),$$

and thus $\underline{\mu}[\underline{X}(t)]$ is affine if and only if $\left[\Sigma \cdot \sqrt{V^D(t)} \cdot \underline{\sigma_q}(t)\right]$ and $\left[\Sigma \cdot \sqrt{V^D(t)} \cdot \underline{\sigma_M}(t)\right]$ are both affine functions of $\underline{X}(t)$. But, this is only possible if $\underline{\sigma_q}(t)$ and $\underline{\sigma_M}(t)$ are both equal to:

1. $\sqrt{V^D(t)}$ multiplied by some $n \times 1$ vector of parameters, since $V^D(t)$ is affine; or
2. $\left(\sqrt{V^D(t)}\right)^{-1}$ multiplied by some $n \times 1$ vector of parameters, since a constant is also an affine function; or even
3. A null $n \times 1$ vector, since zero can also be considered as an affine function.

Although all these three alternatives are possible, we will choose the first one since it represents the most general case. Thus,

$$\underline{\sigma_q}(t) = \sqrt{V^D(t)} \cdot \underline{\varphi}, \tag{2.72}$$

and

$$\underline{\sigma_M}(t) = \sqrt{V^D(t)} \cdot \underline{\chi}, \tag{2.73}$$

where $\underline{\varphi} \in \Re^n$ has φ_i as its i^{th} element, and $\underline{\chi} \in \Re^n$ contains χ_i as its i^{th} element. Equations (2.72) and (2.73) allow us to respect not only condition ii) but also condition i), since $\left[\underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t)\right] = \underline{\chi}' \cdot V^D(t) \cdot \underline{\chi}$, $\left[\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right] = \underline{\varphi}' \cdot V^D(t) \cdot \underline{\varphi}$, and $\sigma_{q,M}(t) = \underline{\chi}' \cdot V^D(t) \cdot \underline{\varphi}$ are all affine functions of $\underline{X}(t)$.

Combining equations (2.69) with (2.72), and (2.70) with (2.73), assumptions A.5 and A.6 are specialized into:

A.5')

$$\frac{dq(t)}{q(t)} = \left[\eta + \underline{\theta}' \cdot \underline{X}(t)\right] dt + \underline{\varphi}' \cdot \sqrt{V^D(t)} \cdot d\underline{W}^{\mathbb{P}}(t). \tag{2.74}$$

A.6')

$$\frac{dM(t)}{M(t)} = \left[\pi + \underline{\phi}' \cdot \underline{X}(t)\right] dt + \underline{\chi}' \cdot \sqrt{V^D(t)} \cdot d\underline{W}^{\mathbb{P}}(t). \tag{2.75}$$

To prove that our general equilibrium framework generates a Duffie and Kan (1996) model, it is only necessary to show that the assumptions A.5', A.6', and A.12' allow us to: i) Specialize equation (2.59) into equation (2.2); and ii) Define a risk-adjusted process for $\underline{X}(t)$ equivalent to equation (2.5). Next theorem verifies requirement i).

Theorem 3 *In a Duffie and Kan (1996) general equilibrium model with a power utility function, and with output and money supply processes described by assumptions A.5' and A.6', respectively, the equilibrium specification for the instantaneous nominal spot interest rate is given by:*

$$r(t) = f + \underline{G}' \cdot \underline{X}(t), \tag{2.76}$$

with

$$f = \rho + \pi - \gamma\eta + \left[\gamma\underline{\Gamma} - \underline{\chi}^2 + \frac{\gamma(1-\gamma)}{2}\underline{\varphi}^2\right]' \cdot \underline{\alpha},$$

and

$$\underline{G} = \underline{\phi} - \gamma\underline{\theta} + \beta \cdot \left[\gamma\underline{\Gamma} - \underline{\chi}^2 + \frac{\gamma(1-\gamma)}{2}\underline{\varphi}^2\right],$$

where $\underline{\varphi}^2, \underline{\chi}^2, \underline{\Gamma} \in \Re^n$ *possess* $(\varphi_i)^2$, $(\chi_i)^2$, *and* $(\chi_i\varphi_i)$ *as their* i^{th} *element, respectively.*

Proof. Substituting (2.67), (2.68), (2.69), (2.70), (2.72), and (2.73) into equation (2.59),

$$\begin{aligned} r(t) &= \rho + \pi + \underline{\phi}' \cdot \underline{X}(t) - \left[\eta + \underline{\theta}' \cdot \underline{X}(t)\right] - \underline{\chi}' \cdot V^D(t) \cdot \underline{\chi} + \underline{\chi}' \cdot V^D(t) \cdot \underline{\varphi} \\ &\quad + (1-\gamma)\left[\eta + \underline{\theta}' \cdot \underline{X}(t) - \underline{\chi}' \cdot V^D(t) \cdot \underline{\varphi} + \underline{\varphi}' \cdot V^D(t) \cdot \underline{\varphi}\right] \\ &\quad - \frac{(\gamma-1)(\gamma-2)}{2}\underline{\varphi}' \cdot V^D(t) \cdot \underline{\varphi}. \end{aligned} \tag{2.77}$$

But, because $\underline{\varphi}' \cdot V^D(t) \cdot \underline{\varphi} = \sum_{i=1}^{n} \varphi_i^2 v_i(t)$, and since $v_i(t) = \alpha_i + \underline{\beta_i}' \cdot \underline{X}(t)$, then $\underline{\varphi}' \cdot V^D(t) \cdot \underline{\varphi} = \sum_{i=1}^{n} \varphi_i^2 \alpha_i + \sum_{i=1}^{n} \varphi_i^2 \underline{\beta_i}' \cdot \underline{X}(t)$, i.e.

$$\underline{\varphi}' \cdot V^D(t) \cdot \underline{\varphi} = \left(\underline{\varphi}^2\right)' \cdot \underline{\alpha} + \left(\underline{\varphi}^2\right)' \cdot \beta' \cdot \underline{X}(t), \tag{2.78}$$

where $\underline{\varphi}^2 \in \Re^n$ has $(\varphi_i)^2$ as its i^{th} component, α_i is the i^{th} element of $\underline{\alpha}$, and β is a $n \times n$ matrix whose i^{th}-column is $\underline{\beta_i}$. Similarly, it is easy to show that

$$\underline{\chi}' \cdot V^D(t) \cdot \underline{\chi} = \left(\underline{\chi}^2\right)' \cdot \underline{\alpha} + \left(\underline{\chi}^2\right)' \cdot \beta' \cdot \underline{X}(t), \tag{2.79}$$

and

$$\underline{\chi}' \cdot V^D(t) \cdot \underline{\varphi} = \underline{\Gamma}' \cdot \underline{\alpha} + \underline{\Gamma}' \cdot \beta' \cdot \underline{X}(t), \tag{2.80}$$

where $\underline{\chi}^2 \in \Re^n$ has $(\chi_i)^2$ as its i^{th} component, and $(\chi_i\varphi_i)$ is the i^{th} element of $\underline{\Gamma} \in \Re^n$. Equations (2.78), (2.79), and (2.80) prove that assumptions A.5' and A.6' guarantee affine specifications for $\left[\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\right]$, $\left[\underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t)\right]$, and $\sigma_{q,M}(t)$.

Combining the last four equations,

$$r(t) = \left\{\rho + \pi - \eta + (1-\gamma)\eta + \left[-\underline{\chi}^2 + \underline{\Gamma} - (1-\gamma)\underline{\Gamma} + (1-\gamma)\underline{\varphi}^2\right.\right.$$

$$
-\frac{(\gamma-1)(\gamma-2)}{2}\underline{\varphi}^2\Big]' \cdot \underline{\alpha}\Big\}
$$

$$
+\Big[\underline{\phi}' - \underline{\theta}' - (\underline{\chi}^2)' \cdot \beta' + \underline{\Gamma}' \cdot \beta' + (1-\gamma)\underline{\theta}' - (1-\gamma)\underline{\Gamma}' \cdot \beta'
$$

$$
+(1-\gamma)(\underline{\varphi}^2)' \cdot \beta' - \frac{(\gamma-1)(\gamma-2)}{2}(\underline{\varphi}^2)' \cdot \beta'\Big] \cdot \underline{X}(t),
$$

and simplifying terms, equation (2.76) is obtained. ■

Equation (2.76) shows that our general equilibrium framework provides an affine form for the instantaneous spot risk-free *nominal* interest rate. Moreover, the derivation of equation (2.76) also showed that it was only possible to obtain an affine form for $r(t)$ because the drift, the variance, and the covariance of the output and money supply processes were also specified as affine functions of $\underline{X}(t)$.

Theorem 4 proves that it is possible to derive a risk-neutral process for $\underline{X}(t)$ equivalent to equation (2.5), and therefore shows that the Duffie and Kan (1996) model is in fact consistent with our type of economy.

Theorem 4 *In a Duffie and Kan (1996) general equilibrium model with a power utility function, and with output and money supply processes described by assumptions A.5' and A.6', respectively:*

1. *The risk-neutral process followed by the state variables under the equivalent martingale measure $\mathcal{Q}$ is equal to*

$$
d\underline{X}(t) = [a \cdot \underline{X}(t) + \underline{b}]\, dt + \Sigma \cdot \sqrt{V^D(t)} \cdot d\underline{W}^{\mathcal{Q}}(t),
$$

if and only if the stochastic process followed by the state variables under the original probability measure $\mathcal{P}$ is assumed to be given by:

$$
d\underline{X}(t) = [\bar{a} \cdot \underline{X}(t) + \underline{\bar{b}}]\, dt + \Sigma \cdot \sqrt{V^D(t)} \cdot d\underline{W}^{\mathcal{P}}(t), \tag{2.81}
$$

where

$$
\bar{a} = a + \Sigma \cdot \Omega^D \cdot \beta',
$$

and

$$
\underline{\bar{b}} = \underline{b} + \Sigma \cdot \Omega^D \cdot \underline{\alpha},
$$

with

$$
\Omega^D = diag\{\chi_1 - \gamma\varphi_1, \ldots, \chi_n - \gamma\varphi_n\}.
$$

2. $d\underline{W}^{\mathbb{Q}}(t) = \underline{\Lambda}[\underline{X}(t)]\,dt + d\underline{W}^{\mathbb{P}}(t)$, *with*

$$\underline{\Lambda}[\underline{X}(t)] = \sqrt{V^D(t)} \cdot (\underline{\chi} - \gamma\underline{\varphi}). \tag{2.82}$$

Proof. In order to obtain a relation between the risk-neutral and the non-risk adjusted drifts of the model' state variables, it is necessary to compute the Duffie and Kan (1996) model' factor risk premiums (under a CRRA economy). For that purpose, equations (2.71), (2.72), and (2.73) can be combined into

$$[\mu_F(t) - r(t)]\,F(t) = \frac{\partial F(t)}{\partial \underline{x}'} \cdot \left[-\gamma\Sigma \cdot V^D(t) \cdot \underline{\varphi} + \Sigma \cdot V^D(t) \cdot \underline{\chi}\right], \tag{2.83}$$

where $\left[\Sigma \cdot V^D(t) \cdot (\underline{\chi} - \gamma\underline{\varphi})\right]$ is the vector of factor risk premiums, or the vector Φ_Y in the terminology of Cox et al. (1985a). Because

$$[\mu_F(t) - r(t)]\,F(t) = \frac{\partial F(t)}{\partial \underline{x}'} \cdot \Sigma \cdot \sqrt{V^D(t)} \cdot \underline{\Lambda}[\underline{X}(t)],$$

equation (2.82) follows for the vector of market prices of risk.[22] Equation (2.83) identifies the analytical formula of the equilibrium risk premium, which makes it possible to derive the fundamental PDE for the Duffie and Kan (1996) model, under a power utility function. Since $F(t)$ is considered to be wealth-independent,

$$\begin{aligned} \mu_F(t)\,F(t) &= (LF)(\underline{x}, t) \\ &= \frac{\partial F(t)}{\partial t} + \frac{\partial F(t)}{\partial \underline{x}'} \cdot \left(a \cdot \underline{X} + \underline{b}\right) + \frac{1}{2} tr\left[\frac{\partial^2 F(t)}{\partial \underline{x} \partial \underline{x}'} \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right]. \end{aligned} \tag{2.84}$$

Combining (2.83) and (2.84), the fundamental valuation equation that must be satisfied by the equilibrium value of any financial contingent claim is obtained:

$$\begin{aligned} &\frac{\partial F(t)}{\partial \underline{x}'} \cdot \left(a \cdot \underline{X} + \underline{b}\right) + \frac{\partial F(t)}{\partial t} + \frac{1}{2} tr\left[\frac{\partial^2 F(t)}{\partial \underline{x} \partial \underline{x}'} \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right] - r(t)\,F(t) \\ =\; &\frac{\partial F(t)}{\partial \underline{x}'} \cdot \left[\Sigma \cdot V^D(t) \cdot (\underline{\chi} - \gamma\underline{\varphi})\right]. \end{aligned} \tag{2.85}$$

The right-hand-side of equation (2.85) can be simplified, providing a simple expression

[22] An alternative derivation is provided in appendix 2.4.3.

for the risk-neutral process followed by the model' state variables:

$$\Sigma \cdot V^D(t) \cdot (\underline{\chi} - \gamma\underline{\varphi}) = \Sigma \cdot \begin{bmatrix} (\chi_1 - \gamma\varphi_1)\, v_1(t) \\ \vdots \\ (\chi_n - \gamma\varphi_n)\, v_n(t) \end{bmatrix}$$
$$= \Sigma \cdot \Omega^D \cdot \underline{\alpha} + \Sigma \cdot \Omega^D \cdot \beta' \cdot \underline{X}(t)\,.$$

Thus, equation (2.85) can be rewritten as

$$0 = \frac{\partial F(t)}{\partial \underline{x}'} \cdot \left[\left(a - \Sigma \cdot \Omega^D \cdot \beta'\right) \cdot \underline{X}(t) + \left(\underline{b} - \Sigma \cdot \Omega^D \cdot \underline{\alpha}\right)\right] + \frac{\partial F(t)}{\partial t} \tag{2.86}$$
$$+\frac{1}{2} tr\left[\frac{\partial^2 F(t)}{\partial \underline{x} \partial \underline{x}'} \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right] - r(t)\, F(t)\,,$$

which, when compared with (2.7), yields equation (2.81). ■

Equations (2.76) and (2.81) completely specify our $(P, \underline{\nu}, \Lambda, \sigma)$ compatible term structure model (under a power utility function), and prove that the Duffie and Kan (1996) model can in fact be fitted into our general equilibrium framework. Equation (2.81) can now be used to estimate the model parameters from a time-series of values for the state variables.

A special case: an economy with a log utility function

Because the log utility function is just a special case of the power utility function (as γ tends to zero), the Duffie and Kan (1996) model can still be fitted into a general equilibrium setup if assumption A.12 is further specialized, maintaining all the other assumptions unchanged:

A.12")

$$u(C, t) = e^{-\rho t} \ln(C)\,. \tag{2.87}$$

Next corollary presents the equilibrium instantaneous *nominal* risk-free interest rate consistent with the above utility function.

Corollary 1 *In a Duffie and Kan (1996) general equilibrium model with a log utility function, and with output and money supply processes described by assumptions A.5' and A.6', respectively, the equilibrium specification for the instantaneous nominal spot interest rate is given by:*

$$r(t) = f + \underline{G}' \cdot \underline{X}(t)\,, \tag{2.88}$$

with

$$f = \rho + \pi - \left(\underline{\chi}^2\right)' \cdot \underline{\alpha},$$

and

$$\underline{G} = \underline{\phi} - \beta \cdot \underline{\chi}^2.$$

Proof. Equation (2.88) is simply the limit of expression (2.76) as $\gamma \to 0$. ■

Similarly, the risk-neutral process for $\underline{X}(t)$ that is consistent with assumption A.12" follows from theorem 4.

Corollary 2 *In a Duffie and Kan (1996) general equilibrium model with a log utility function, and with output and money supply processes described by assumptions A.5' and A.6', respectively:*

(a) The risk-neutral process followed by the state variables under the equivalent martingale measure Q is equal to

$$d\underline{X}(t) = [a \cdot \underline{X}(t) + \underline{b}]\, dt + \Sigma \cdot \sqrt{V^D(t)} \cdot d\underline{W}^Q(t),$$

if and only if the stochastic process followed by the state variables under the original probability measure $\mathcal{P}$ is assumed to be given by:

$$d\underline{X}(t) = \left[\tilde{a} \cdot \underline{X}(t) + \tilde{\underline{b}}\right] dt + \Sigma \cdot \sqrt{V^D(t)} \cdot d\underline{W}^{\mathcal{P}}(t), \tag{2.89}$$

where

$$\tilde{a} = a + \Sigma \cdot \Phi^D \cdot \beta',$$

and

$$\tilde{\underline{b}} = \underline{b} + \Sigma \cdot \Phi^D \cdot \underline{\alpha}.$$

with

$$\Phi^D = diag\{\chi_1, \ldots, \chi_n\}.$$

(b) $d\underline{W}^Q(t) = \underline{\Lambda}[\underline{X}(t)]\, dt + d\underline{W}^{\mathcal{P}}(t)$, *with*

$$\underline{\Lambda}[\underline{X}(t)] = \sqrt{V^D(t)} \cdot \underline{\chi}. \tag{2.90}$$

Proof. Corollary 2 is obtained from theorem 4 by taking the limit of expressions (2.81) and (2.82), as γ tends to zero. ■

Now, equations (2.88) and (2.89) completely specify a simpler but more restrictive $(P, \nu, \Lambda, \sigma)$ compatible term structure model, under a log utility function. This specification embodies as special cases several existing equilibrium term structure models, such as Cox et al. (1985b) and Longstaff and Schwartz (1992a), which were also derived under the restrictive type of preferences implied by the log utility function. Moreover, equation (2.90) is equivalent to the market prices of risk' specification estimated by Dai and Singleton (1998, equation 5), using the simulated method of moments, and considered by Lund (1997a, equation 25), through a linear Kalman filter implemented by QML estimation.

2.3 Conclusions

This Chapter was intended to bring two main contributions to the existing literature. Firstly, in theorems 1 and 2 new equilibrium specifications are given both for the *nominal* short-term interest rate and for the expected excess *nominal* return on a financial contingent claim, in the general context of a one-country monetary economy. Secondly, theorems 3 and 4 propose a general equilibrium Duffie and Kan (1996) model specification, under the original probability measure $\mathcal{P}$, that is compatible with the original model' formulation under the equivalent martingale measure $\mathcal{Q}$, and that is based on more realistic assumptions about preferences than those implied by the usual Bernoulli logarithmic utility function (since a power utility function was used). In other words, our (P, v, Δ, σ) model is a very general term structure model, not only because it is the most general in the class of the multifactor affine time-homogeneous interest rate models, but also because it relies on general assumptions about preferences. For empirical purposes, this (P, v, Δ, σ) specification is useful since it enables the econometric estimation of the Duffie and Kan (1996) model' parameters from a time-series of values for the state variables or from a panel-data of market observables.

In Chapter four, the analytical specification derived in the present Chapter for the vector of market prices of risk -equation (2.82)- will be used in order to rewrite the Duffie and Kan (1996) model into a state-space form.

2.4 Appendices

2.4.1 An alternative proof of Theorem 1

Following Bakshi and Chen (1997b, pages 818-819), formula (2.59) will be derived by assuming from the beginning a pure exchange economy. Moreover, instead of working in continuous time, we will start by considering time intervals of length Δt, and later we will take $\Delta t \rightarrow 0$. Although the same closed-form solution for $r(t)$ will be obtained in a much simpler fashion, this derivation does not provide any intuition towards the qualitative results obtained in subsection 2.2.3; and, it is also subject to a discretization error of small magnitude.

As usual, the time-t price of a default-free (unit face value) pure discount bond with maturity at time $(t + \Delta t)$ will be denominated by $P(t, t + \Delta t)$. When $\Delta t \to 0$, $r(t)$ is the yield-to-maturity of such zero coupon bond, and therefore

$$P(t, t + \Delta t) = e^{-r(t)\Delta t}.$$

On the other hand, it is well known that, in equilibrium, the loss of marginal utility of current consumption implied by the purchase of a pure discount bond at time t must be equal to the gain of expected marginal utility of future consumption implied by the obtention of a monetary unit at time $(t + \Delta t)$. And, since $P(t, t + \Delta t)$ monetary units at time t correspond to $\frac{P(t,t+\Delta t)}{p(t)}$ units of real consumption, one monetary unit received at time $(t + \Delta t)$ is equivalent to $\frac{1}{p(t+\Delta t)}$ units of real consumption, and it is assumed that $C(t) = q(t)$, then

$$\frac{P(t, t + \Delta t)}{p(t)} u_q [q(t), t] = E_t \left\{ \frac{1}{p(t + \Delta t)} u_q [q(t + \Delta t), t + \Delta t] \right\}.$$

Combining the last two equations, and considering that $u(q, t) = e^{-\rho t} U(q)$,[23]

$$1 + [\rho - r(t)] \Delta t \approx E_t \left\{ \frac{U_q [q(t + \Delta t)]}{U_q [q(t)]} \frac{p(t)}{p(t + \Delta t)} \right\}. \tag{2.91}$$

Taking the Taylor series expansion of $\frac{1}{p(t+\Delta t)}$ at $p(t)$,

$$\frac{1}{p(t + \Delta t)} = \frac{1}{p(t)} - \frac{1}{p(t)^2} \Delta p(t) + \frac{1}{2} \frac{2}{p(t)^3} [\Delta p(t)]^2 + O\left(\Delta t^{\frac{3}{2}}\right),$$

where $O\left(\Delta t^{\frac{3}{2}}\right)$ is a linear function of $\Delta t^{\frac{3}{2}}$ and higher-order terms, which are negligible. Multiplying both sides of the above equation by $p(t)$,

$$\frac{p(t)}{p(t + \Delta t)} = 1 - \frac{\Delta p(t)}{p(t)} + \left[\frac{\Delta p(t)}{p(t)} \right]^2 + O\left(\Delta t^{\frac{3}{2}}\right). \tag{2.92}$$

Taking the Taylor series expansion of $p(t + \Delta t) = \frac{M(t+\Delta t)}{q(t+\Delta t)}$ at $(M(t), q(t))$,

$$\begin{aligned} p(t + \Delta t) &= p(t) + \frac{\Delta M(t)}{q(t)} - \frac{M(t)}{q(t)^2} \Delta q(t) + \frac{M(t)}{q(t)^3} [\Delta q(t)]^2 \\ &\quad - \frac{1}{q(t)^2} \Delta q(t) \Delta M(t) + O\left(\Delta t^{\frac{3}{2}}\right), \end{aligned}$$

[23] The above mentioned discretization error arises from the fact that $\exp\{[\rho - r(t)] \Delta t\}$ is only equal to $\{1 + [\rho - r(t)] \Delta t\}$, or equivalently, terms of order higher than Δt can only be ignored, when $\Delta t \to 0$. In other words, such approximation ignores the difference between continuous and discrete compounding.

that is

$$\frac{\Delta p(t)}{p(t)} = \frac{\Delta M(t)}{M(t)} - \frac{\Delta q(t)}{q(t)} + \left[\frac{\Delta q(t)}{q(t)}\right]^2 - \frac{\Delta q(t)}{q(t)}\frac{\Delta M(t)}{M(t)} + O\left(\Delta t^{\frac{3}{2}}\right). \tag{2.93}$$

Substituting (2.93) into (2.92), and taking $\Delta t \rightarrow 0$,

$$\begin{aligned}\frac{p(t)}{p(t+\Delta t)} &= 1 - \frac{dM(t)}{M(t)} + \frac{dq(t)}{q(t)} - \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\,dt + \sigma_{q,M}(t)\,dt \\ &\quad + \underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t)\,dt + \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\,dt - 2\sigma_{q,M}(t)\,dt + O\left(dt^{\frac{3}{2}}\right),\end{aligned}$$

i.e.

$$\frac{p(t)}{p(t+\Delta t)} = 1 + \left[\underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t) - \sigma_{q,M}(t)\right]dt - \frac{dM(t)}{M(t)} + \frac{dq(t)}{q(t)} + O\left(dt^{\frac{3}{2}}\right). \tag{2.94}$$

Again, taking the Taylor series of $U_q[q(t+\Delta t)]$ at $q(t)$,

$$U_q[q(t+\Delta t)] = U_q[q(t)] + U_{qq}[q(t)]\,\Delta q(t) + \frac{1}{2}U_{qqq}[q(t)]\,[\Delta q(t)]^2 + O\left(\Delta t^{\frac{3}{2}}\right),$$

and making $\Delta t \rightarrow 0$,

$$\begin{aligned}\frac{U_q[q(t+dt)]}{U_q[q(t)]} &= 1 + \frac{U_{qq}[q(t)]}{U_q[q(t)]}dq(t) \\ &\quad + \frac{1}{2}\frac{U_{qqq}[q(t)]}{U_q[q(t)]}\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\,q(t)^2\,dt + O\left(dt^{\frac{3}{2}}\right).\end{aligned} \tag{2.95}$$

Substituting (2.95) and (2.94) into (2.91), making $\Delta t \rightarrow 0$, and considering only time-dependencies,

$$\begin{aligned}1 + [\rho - r(t)]\,dt &= E_t\left\{1 + \frac{U_{qq}(t)}{U_q(t)}dq(t) + \frac{1}{2}\frac{U_{qqq}(t)}{U_q(t)}\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\,q(t)^2\,dt\right. \\ &\quad + \left[\underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t) - \sigma_{q,M}(t)\right]dt - \frac{dM(t)}{M(t)} + \frac{dq(t)}{q(t)} \\ &\quad \left. - \frac{q(t)\,U_{qq}(t)}{U_q(t)}\sigma_{q,M}(t)\,dt + \frac{q(t)\,U_{qq}(t)}{U_q(t)}\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\,dt + O\left(dt^{\frac{3}{2}}\right)\right\}.\end{aligned}$$

Taking expectations,

$$\begin{aligned}[\rho - r(t)]\,dt &= \frac{U_{qq}(t)}{U_q(t)}\mu_q(t)\,q(t)\,dt + \frac{1}{2}\frac{U_{qqq}(t)}{U_q(t)}\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\,q(t)^2\,dt \\ &\quad + \left[\underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t) - \sigma_{q,M}(t)\right]dt - \mu_M(t)\,dt + \mu_q(t)\,dt \\ &\quad - \frac{q(t)\,U_{qq}(t)}{U_q(t)}\sigma_{q,M}(t)\,dt + \frac{q(t)\,U_{qq}(t)}{U_q(t)}\underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t)\,dt.\end{aligned}$$

Dividing both sides by dt, and because $\frac{U_{qq}(t)}{U_q(t)} = \frac{u_{qq}(t)}{u_q(t)}$ as well as $\frac{U_{qqq}(t)}{U_q(t)} = \frac{u_{qqq}(t)}{u_q(t)}$, we get exactly the same result as in equation (2.59). ■

2.4.2 An alternative proof of Theorem 2

Once again, a pure exchange economy is assumed, and time intervals of length Δt are initially considered.

In equilibrium, the factor risk premiums on any contingent claim $F(t)$ must obey to the following Euler equation:

$$E_t \left\{ \frac{U_q [q(t+\Delta t), t+\Delta t]}{U_q [q(t), t]} \frac{p(t)}{p(t+\Delta t)} \left[\frac{\Delta F(t)}{F(t)} - r(t) \Delta t \right] \right\} = 0.$$

Making $\Delta t \to 0$, and substituting $\frac{U_q[q(t+\Delta t),t+\Delta t]}{U_q[q(t),t]}$ by equation (2.95) as well as $\frac{p(t)}{p(t+\Delta t)}$ by equation (2.94), yields:

$$\begin{aligned}
0 &= E_t \left\{ \left[1 + \frac{U_{qq}(t)}{U_q(t)} dq(t) + \frac{1}{2} \frac{U_{qqq}(t)}{U_q(t)} \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t) q(t)^2 dt \right. \right. \\
&\quad + \left[\underline{\sigma_M}(t)' \cdot \underline{\sigma_M}(t) - \sigma_{q,M}(t) \right] dt - \frac{dM(t)}{M(t)} + \frac{dq(t)}{q(t)} - \frac{q(t) U_{qq}(t)}{U_q(t)} \sigma_{q,M}(t) dt \\
&\quad \left. \left. + \frac{q(t) U_{qq}(t)}{U_q(t)} \underline{\sigma_q}(t)' \cdot \underline{\sigma_q}(t) dt + O\left(dt^{\frac{3}{2}}\right) \right] \left[\left(\frac{dF(t)}{F(t)} - r(t) dt \right) \right] \right\}.
\end{aligned}$$

Multiplying both members inside the expectation operator, and taking expectations,

$$\begin{aligned}
0 &= \mu_F(t) dt - r(t) dt + \frac{U_{qq}(t)}{U_q(t)} COV \left[\frac{dF(t)}{F(t)}, dq(t) \right] dt \\
&\quad - COV \left[\frac{dF(t)}{F(t)}, \frac{dM(t)}{M(t)} \right] dt + COV \left[\frac{dF(t)}{F(t)}, \frac{dq(t)}{q(t)} \right] dt.
\end{aligned}$$

Finally, dividing both sides of the above equation by dt and rearranging terms, equation (2.65) follows. ■

2.4.3 An alternative derivation of $\Lambda[\underline{X}(t)]$ under a power utility function

Since $\underline{\mu}[\underline{X}(t)] = \underline{v}[\underline{X}(t)] - \sigma[\underline{X}(t)] \cdot \Lambda[\underline{X}(t)]$, because $\sigma[\underline{X}(t)]$ is assumed to be invertible, and using equations (2.5) and (2.10), then

$$\Lambda[\underline{X}(t)] = \left[\Sigma \cdot \sqrt{V^D(t)} \right]^{-1} \cdot \left\{ \left[a \cdot \underline{X}(t) + \underline{b} \right] - \left[a \cdot \underline{X}(t) + \underline{b} \right] \right\}.$$

Attending to equation (2.81),

$$\underline{\Lambda}[\underline{X}(t)] = \left[\sqrt{V^D(t)} \right]^{-1} \cdot \Sigma^{-1} \cdot \left[\Sigma \cdot \Omega^D \cdot \beta' \cdot \underline{X}(t) + \Sigma \cdot \Omega^D \cdot \underline{\alpha} \right].$$

Finally, since $\Sigma \cdot \Omega^D \cdot \beta' \cdot \underline{X}(t) + \Sigma \cdot \Omega^D \cdot \underline{\alpha} = \Sigma \cdot V^D(t) \cdot (\underline{\chi} - \gamma\underline{\varphi})$, then

$$\underline{\Lambda}[\underline{X}(t)] = \sqrt{V^D(t)} \cdot (\underline{\chi} - \gamma\underline{\varphi}),$$

as expected.

Chapter 3

The Gaussian Special Case

This Chapter is based on the article Nunes (1998), which supersedes a working paper with the same title and presented at the 14th AFFI Conference (Grenoble, 1997).

3.1 Introduction

In this Chapter, simple analytical pricing solutions will be derived, for several interest rate contingent claims, under a Gaussian (nested) version of the Duffie and Kan (1996) model, which still preserves the *mean reversion* and *affine yields* features of the original model.

Starting from the pricing formula for default-free pure discount bonds already derived by Langetieg (1980), closed-form solutions will be found for the prices of (short-term and long-term) interest rate futures, European (conventional and pure) interest rate futures options, European spot options on default-free (pure discount and coupon-bearing) bonds, caps, floors and European swaptions. Valuation formulae for European options will be derived through the probabilistic change-of-numeraire technique developed in El Karoui and Rochet (1989), El Karoui et al. (1991), El Karoui, Myneni and Viswanathan (1992a), and El Karoui, Myneni and Viswanathan (1992b). Hence, the pricing solutions that are going to be obtained for European options on-the-spot can be considered as just special cases of the more general ones contained in the above mentioned papers, and in this case, the only contribution that will be made is to adapt the previous results to the context of the Gaussian Duffie and Kan (1996) model specification, that is to specify and to compute the appropriate volatility parameters needed as an input for the option pricing formulae. Interest rate futures will be priced as moment generating functions -following Chen (1995)- by exploring the Gaussian specification of the state variables, and European futures option prices will be found not only for short-term interest rate futures and pure discount bond futures but also for futures on coupon-bearing bonds. A put-call parity relation for European options on coupon-bearing bond futures will be deduced.

The main contribution of the present Chapter consists in providing simple pricing formulae that do not involve a single univariate integral, under the most general Gaussian multifactor mean-reverting and affine term structure model. These analytical pricing solutions can be efficiently applied, for instance, to construct martingale control variates -as in Clewlow and Carverhill (1994)- in Monte Carlo implementations of the more generic stochastic volatility specification, or just in the context of any Gaussian time-homogeneous affine model previously derived in the literature.[1] More specifically, in the context of this dissertation, these Gaussian pricing formulae will serve two concrete purposes: they will provide the non-linear measurement equations for the Gaussian state-space models estimated in Chapter four; and, they will be used, in Chapter five, to derive analytical approximate pricing solutions under the most general stochastic volatility specification of the Duffie and Kan (1996) model.

Next sections are organized as follows. Section 3.2 presents the Gaussian version of the Duffie and Kan (1996) model, and fits it into the Langetieg (1980) model. Section 3.3 introduces another martingale probability measure, which is equivalent to Q and will be relevant for the pricing of bond options. In section 3.4 closed-form solutions are derived for the prices of European options on default-free (pure discount and coupon-bearing) bonds, being then generalized for the valuation of interest rate caps, floors and European swaptions. Section 3.5 provides pricing formulae for futures on short-term and long-term interest rates, which will be then used as an input to obtain, in section 3.6, closed-form solutions for the prices of the corresponding European options. Finally, some accessory proofs are presented in the appendix.

3.2 Nested Gaussian specification

The deterministic volatility specification of the Duffie and Kan (1996) model is obtained by imposing $\underline{\beta}_i = \underline{0}$ (for $i = 1, \ldots, n$) in equation (2.5). This nested Gaussian specification is then given by equations (2.1) and (3.1),

$$d\underline{X}(t) = [a \cdot \underline{X}(t) + \underline{b}]\, dt + S \cdot d\underline{W}^{Q}(t), \underline{X}(t) \in \Re^n, \tag{3.1}$$

[1] Although such Gaussian models possess the drawback of not precluding the existence of negative interest rates, Schlögl and Sommer (1998) have found that the *mean reversion* feature and the *dimension* of a term structure model are much more important than the distributional characteristics of its state variables. Therefore, the tractable and general Gaussian framework under analysis in the present Chapter is relevant on its own and it is not merely instrumental.

where $S = \Sigma \cdot \sqrt{U^D}$, with $\sqrt{U^D} = diag\left\{\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_n}\right\}$, and includes, as special cases, all Gaussian affine and time-homogeneous models previously derived in the literature.[2] In essence, this formulation corresponds to the Langetieg (1980) *multivariate elastic random walk model*, and thus an analytic formula exists for default-free pure discount bonds.

Proposition 1 *Under the deterministic volatility specification of the Duffie and Kan (1996) model and assuming that matrix a is non-singular, the price of a riskless zero-coupon bond is given by equation (2.1), where*

$$\underline{B}'(\tau) = \underline{G}' \cdot a^{-1} \cdot (I_n - e^{a\tau}), \tag{3.2}$$

$$\begin{aligned} A(\tau) = \quad & \tau\left(\underline{G}' \cdot a^{-1} \cdot \underline{b} - f\right) + \underline{B}'(\tau) \cdot a^{-1} \cdot \underline{b} + \frac{\tau}{2}\underline{G}' \cdot a^{-1} \cdot \Theta \cdot \left(a^{-1}\right)' \cdot \underline{G} \\ & + \underline{G}' \cdot a^{-1} \cdot (I_n - e^{a\tau}) \cdot a^{-1} \cdot \Theta \cdot \left(a^{-1}\right)' \cdot \underline{G} \\ & + \frac{1}{2}\underline{G}' \cdot a^{-1} \cdot \Delta(\tau) \cdot \left(a^{-1}\right)' \cdot \underline{G}, \end{aligned} \tag{3.3}$$

and

$$\Delta(\tau) = \int_t^T e^{a(T-v)} \cdot \Theta \cdot e^{a'(T-v)} dv, \tag{3.4}$$

with $\Theta \equiv S \cdot S'$, and $I_n \in \Re^{n \times n}$ denoting an identity matrix.

Proof. When we move towards a Gaussian specification, in which $\underline{\beta}_i = \underline{0}$ (for $i = 1, \ldots, n$), equation (2.8) is replaced by the following first order nonhomogeneous linear system of differential equations (with constant coefficients, since the matrix a is time-independent):

$$\frac{\partial}{\partial \tau}\underline{B}'(\tau) = -\underline{G}' + \underline{B}'(\tau) \cdot a. \tag{3.5}$$

In appendix 3.8.1 it is proved that the solution for this simple initial value problem is given by equation (3.2), where $e^{a\tau} = \sum_{k=0}^{\infty} a^k \frac{\tau^k}{k!}$. Although derived in a different way, this closed-form solution corresponds exactly to the *risk vector* $\nabla(t,T)' = \frac{1}{P(t,T)} \frac{\partial}{\partial \underline{X}'(t)} P(t,T)$ of Langetieg (1980, equation 25). Each one of its components measures the sensitivity of the bond rate of return (over a small time interval) to an instantaneous change of each one of the state variables, and can therefore be interpreted as a *duration* measure.

Concerning equation (3.3), because it can be shown that, under the Gaussian specification, $\sum_{k=1}^{n}\left[\sum_{j=1}^{n} B_j(\tau)\,\varepsilon_{jk}\right]^2 \alpha_k = \underline{B}'(\tau) \cdot S \cdot S' \cdot \underline{B}(\tau)$, and since $A(0) = 0$, integrating

[2] Moreover, because matrix a is not assumed to be diagonal, even a richer family of interest rate dynamics can be obtained from equation (3.1).

equation (2.9) between 0 and τ yields:

$$A(\tau) = -\tau f + \int_0^\tau \underline{B}'(u)\,du \cdot \underline{b} + \frac{1}{2}\int_0^\tau \underline{B}'(u) \cdot \Theta \cdot \underline{B}(u)\,du. \tag{3.6}$$

Finally, using solution (3.2) and straightforward integral calculus, formulae (3.3) and (3.4) are obtained.[3] ■

Remark 1 *As noticed in Langetieg (1980, footnote 20), matrix a will be singular only if one of the state variables follows a random walk. Even in such case, equations (3.2) to (3.4) can always be replaced by the more general solutions described in Lund (1994, appendix A).*

Matrix $\Delta(\tau)$ can be interpreted as the covariance of the state-vector $\underline{X}(T)$, computed under measure Q and conditional on $\mathcal{F}_t$. Next proposition offers an analytical solution, involving no single integral, under the assumption of linearly independent eigenvectors for matrix a.

Proposition 2 *Under the deterministic volatility specification of the Duffie and Kan (1996) model and assuming also that matrix a is diagonalizable,[4] the function $\Delta(\tau)$ possesses the following explicit solution:*

$$\Delta(\tau) = e^{a(T-t)} \cdot Y \cdot e^{a'(T-t)} - Y, \tag{3.7}$$

where $\Theta^{*} = Q^{-1} \cdot \Theta \cdot (Q^{-1})' \equiv \left\{\sigma^{*}_{ij}\right\}_{i,j=1,\ldots,n}$, $\Theta^{**} = \left\{\frac{\sigma^{*}_{ij}}{\lambda_i+\lambda_j}\right\}_{i,j=1,\ldots,n}$, $Y = Q \cdot \Theta^{**} \cdot Q'$, λ_i $(i = 1,\ldots,n)$ is the i^{th} eigenvalue of matrix a, and Q is a $n \times n$ matrix whose columns correspond to the eigenvectors of matrix a.

Proof. See Langetieg (1980, footnote 23) or appendix 3.8.2. ■

Remark 2 *As argued by Duan and Simonato (1995, page 26), this "assumption of diagonability does not involve an appreciable loss of generality" because the eigenvalues of a matrix are continuous functions of its elements (and thus multiple roots of the characteristic equation can be avoided by a small adjustment in the original matrix). Nevertheless, for the numerical experiments presented in this thesis, all matrix exponentials are computed using Padé approximations with scaling and squaring. For details, see Van Loan (1978).*

[3] Alternatively, the analytical solution for $A(\tau)$ also follows by combining equations (30), (32), and (39b) of Langetieg (1980), remembering that his vector $\nabla(t, T)$ corresponds to $\underline{B}(\tau)$, and considering the following additional notational conversions from Langetieg (1980) to our framework: $w = \underline{G}$, $B = a$, $a = b$, $w_0 = f$, $\Sigma = S \cdot S'$, and $\phi = 0$, since we are already working under a risk-neutral probability measure.

[4] *Meaning that matrix a possess n linearly independent eigenvectors (while the corresponding eigenvalues do not need to be distinct). Nevertheless, if the eigenvalues of matrix a are distinct (and different from zero, since a is assumed to be non-singular), then matrix a will be diagonalizable.*

3.3 Probabilistic change-of-numeraire technique: from measure $\mathcal{Q}$ to measure $\mathcal{Q}_0$

Option prices will be obtained by decomposing the option value in terms of two or more particular numeraires, and in terms of their related martingale measures. One of the numeraires will always be the pure discount bond expiring at the option's maturity[5], $P(t, T_0)$, to which is associated the forward measure that will be denominated by $\mathcal{Q}_0$. The other numeraires and their associated probability measures will be defined according to the option's underlying asset.

Since options will be priced under the reference martingale measure $\mathcal{Q}_0$, and the stochastic process (3.1) was previously defined under measure $\mathcal{Q}$, the purpose of this section is to establish the change of probability measure from $\mathcal{Q}$ to $\mathcal{Q}_0$. In what follows, we will consider a stochastic intertemporal economy with a finite time horizon $\mathcal{T} = [0, T_L]$,[6] where uncertainty is represented by a complete filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathcal{T}}, \mathcal{P}\right)$ and where all the information accruing to all the agents in the economy is described by a (right-) continuous filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$ with $\mathcal{F}_0$ containing all the $\mathcal{P}-$null sets of $\mathcal{F}$.[7] Following Harrison and Pliska (1981) and assuming that the market is arbitrage-free,[8] once the change to the martingale measure $\mathcal{Q}_0$ has been established, then it will be possible to consider the time-t forward price (for date T_0) of any attainable contingent claim V which settles at time T_0, $\frac{V_t}{P(t,T_0)}$, as a $\mathcal{Q}_0$-martingale, and thus the following fundamental pricing formula will be used:[9]

$$\frac{V_t}{P(t, T_0)} = E_{\mathcal{Q}_0}\left[\left.\frac{V_{T_0}}{P(T_0, T_0)}\right| \mathcal{F}_t\right], \ \forall t \leq T_0, \tag{3.8}$$

assuming that the terminal payoff V_{T_0} is $\mathcal{F}_{T_0}-$measurable and that the integrability condition $E_{\mathcal{Q}_0}\left[\left|\frac{V_t}{P(t,T_0)}\right|\right] < \infty$ is valid for all t.

3.3.1 Probability measure $\mathcal{Q}$ and the numeraire $\delta(t)$

First of all, it will be proved that the numeraire associated with the initial martingale measure $\mathcal{Q}$ can be thought of as being the value of a "money market account". For this purpose, $\delta(t)$ will represent the time-t value of a monetary unit invested at time 0 on the money market and continuously reinvested at the riskless instantaneous spot interest rate,

[5] The option's expiry date will be denominated by T_0, and $P(t, T_0)$ is usually called the "reference bond".
[6] T_L can be thought as the largest maturity to be used in the subsequent analysis.
[7] $(\mathcal{F}_t)_{t\in\mathcal{T}}$ will be understood as the $\mathcal{P}-$augmentation of the natural filtration generated by a standard n-dimensional Brownian motion defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$.
[8] And therefore considering that there exists, at least, one equivalent martingale probability measure.
[9] $E_{\mathcal{Q}_i}(Y|\mathcal{F}_t)$ denotes the time-t expected value of the random variable Y, computed under the probability measure $\mathcal{Q}_i$.

that is

$$\delta(t) = \exp\left[\int_0^t r(s)\,ds\right].$$

Let $P(t, T_i)$, or simply P_i, be the time-t price of a risk-free pure discount bond maturing at time T_i (and with a face value equal to one monetary unit). Using equations (2.1) and (3.1), and applying Itô's lemma to $P(t, T_i)$,

$$\frac{dP(t,T_i)}{P(t,T_i)} = r(t)\,dt + \underline{B}'(\tau_i)\cdot S\cdot d\underline{W}^Q(t), \tag{3.9}$$

where $\tau_i = T_i - t$ represents the time-to-maturity of the zero coupon bond under analysis.[10] Applying again Itô's lemma to $\ln[P(t, T_i)]$,

$$d\ln[P(t,T_i)] = \left\{r(t) - \frac{1}{2}\left[\underline{B}'(\tau_i)\cdot S\right]\cdot\left[\underline{B}'(\tau_i)\cdot S\right]'\right\}dt + \underline{B}'(\tau_i)\cdot S\cdot d\underline{W}^Q(t),$$

and solving this stochastic differential equation for $P(t, T_i)$, yields

$$\begin{aligned} P(t,T_i) &= P(0,T_i)\exp\left[\int_0^t r(u)\,du - \frac{1}{2}\int_0^t \underline{B}'(T_i-u)\cdot S\cdot S'\cdot \underline{B}(T_i-u)\,du\right. \\ &\quad \left. + \int_0^t \underline{B}'(T_i-u)\cdot S\cdot d\underline{W}^Q(u)\right]. \end{aligned}$$

Dividing both sides of the last equation by $[\delta(t)\,\delta(0)]$, and since $\delta(0) = 1$,

$$\begin{aligned} \frac{P(t,T_i)}{\delta(t)} &= \frac{P(0,T_i)}{\delta(0)}\exp\left[-\frac{1}{2}\int_0^t \underline{B}'(T_i-u)\cdot S\cdot S'\cdot \underline{B}(T_i-u)\,du\right. \\ &\quad \left. + \int_0^t \underline{B}'(T_i-u)\cdot S\cdot d\underline{W}^Q(u)\right]. \end{aligned} \tag{3.10}$$

Finally, the exponential martingale formula allows us to write the following SDE for the relative price of security $P(t, T_i)$ with respect to the numeraire $\delta(t)$:

$$d\left[\frac{P(t,T_i)}{\delta(t)}\right] = \frac{P(t,T_i)}{\delta(t)}\underline{B}'(\tau_i)\cdot S\cdot d\underline{W}^Q(t). \tag{3.11}$$

Equation (3.11) shows that $\frac{P(t,T_i)}{\delta(t)}$ is a *local martingale* under the equivalent probability measure Q. Hence, if the following *Novikov's condition* is further assumed,

$$E_Q\left\{\exp\left[\frac{1}{2}\int_0^t \underline{B}'(T_i-u)\cdot\Theta\cdot\underline{B}(T_i-u)\,du\right]\middle|\,\mathcal{F}_0\right\} < \infty, \tag{3.12}$$

then $\frac{P(t,T_i)}{\delta(t)}$ is a *martingale* and therefore, it can be concluded that the numeraire associated

[10] The functional form of the drift is a consequence of Q being a risk-neutral probability measure: by definition, the instantaneous return on all securities equals the riskless instantaneous interest rate.

with the equivalent probability measure Q can be thought of as being the value of the "savings account", $\delta(t)$. In fact, rewriting equation (3.10) for the time interval $[s,t]$ (instead of $[0,t]$) and because the random variable $\int_s^t \underline{B}'(T_i-u)\cdot S\cdot d\underline{W}^Q(u)$, conditional on $\mathcal{F}_s$, is normally distributed with mean zero and variance $\int_s^t \underline{B}'(T_i-u)\cdot\Theta\cdot\underline{B}(T_i-u)\,du$, then

$$\begin{aligned}
E_Q\left[\left.\frac{P(t,T_i)}{\delta(t)}\right|\mathcal{F}_s\right] &= \frac{P(s,T_i)}{\delta(s)}\exp\left[-\frac{1}{2}\int_s^t \underline{B}'(T_i-u)\cdot\Theta\cdot\underline{B}(T_i-u)\,du\right.\\
&\quad\left.+0+\frac{1}{2}\int_s^t \underline{B}'(T_i-u)\cdot\Theta\cdot\underline{B}(T_i-u)\,du\right]\\
&= \frac{P(s,T_i)}{\delta(s)},\ \forall s\leq t.
\end{aligned}$$

Concerning the integrability condition $E_Q\left[\left|\frac{P(t,T_i)}{\delta(t)}\right|\right]<\infty, \forall t$, using assumption (3.12) and the fact that both $P(t,T_i)$ and $\delta(t)$ are positive price processes, it is only required that the relative price $\frac{P(0,T_i)}{\delta(0)}$ is finite.

3.3.2 Probability measure Q_0 and the numeraire $P(t,T_0)$

As we shall confirm in subsection 3.5.1, the relative price process $P(t,T_0,T_i)=\frac{P(t,T_i)}{P(t,T_0)}$, for $t\leq T_0\leq T_i$, is the time-t forward price, for delivery at time T_0, of a riskless pure discount bond expiring at time T_i. Using Itô's lemma and equation (3.9),[11]

$$dP(t,T_0,T_i)=-\frac{P_i}{P_0}\frac{dP_0}{P_0}+\frac{P_i}{P_0}\frac{dP_i}{P_i}+\frac{P_i}{P_0}\left(\frac{dP_0}{P_0}\right)^2-\frac{P_i}{P_0}\frac{dP_0}{P_0}\frac{dP_i}{P_i},$$

that is

$$\begin{aligned}
\frac{dP(t,T_0,T_i)}{P(t,T_0,T_i)} &= -\left[r(t)\,dt+\underline{B}'(\tau_0)\cdot S\cdot d\underline{W}^Q(t)\right]\\
&\quad+\left[r(t)\,dt+\underline{B}'(\tau_i)\cdot S\cdot d\underline{W}^Q(t)\right]\\
&\quad+\underline{B}'(\tau_0)\cdot S\cdot S'\cdot\underline{B}(\tau_0)\,dt-\underline{B}'(\tau_i)\cdot S\cdot S'\cdot\underline{B}(\tau_0)\,dt,
\end{aligned}$$

where $\tau_0=T_0-t$ is the time-to-maturity of the option. And, since the last two terms on the right-hand-side of the above equation are both scalars,

$$\frac{dP(t,T_0,T_i)}{P(t,T_0,T_i)}=\left[\underline{B}'(\tau_i)-\underline{B}'(\tau_0)\right]\cdot S\cdot\left[-S'\cdot\underline{B}(\tau_0)\,dt+d\underline{W}^Q(t)\right].$$

Hence, if another probability measure Q_0 defined on the measurable space $(\Omega,\mathcal{F})$, and

[11] For easy of exposition, let $P(t,T_i)=P_i$, $\forall i$.

equivalent to $\mathcal{Q}$ and $\mathcal{P}$, exists, such that

$$d\underline{W}^{\mathcal{Q}_0}(t) = -S' \cdot \underline{B}(\tau_0)\, dt + d\underline{W}^{\mathcal{Q}}(t) \tag{3.13}$$

is also a vector of standard Brownian motion increments in $\Re^n$ (with the same standard filtration as $d\underline{W}^{\mathcal{Q}}(t)$ and $d\underline{W}^{\mathcal{P}}(t)$), then it will be possible to write:

$$\frac{dP(t,T_0,T_i)}{P(t,T_0,T_i)} = \left[\underline{B}'(\tau_i) - \underline{B}'(\tau_0)\right] \cdot S \cdot d\underline{W}^{\mathcal{Q}_0}(t). \tag{3.14}$$

In order for equation (3.14) to be valid, it is necessary to prove that the Radon-Nikodym derivative of $\mathcal{Q}_0$ with respect to $\mathcal{Q}$ (at any terminal time $t \leq T_0$)[12], which (from Girsanov's Theorem) is given by

$$\begin{aligned}\left(\frac{d\mathcal{Q}_0}{d\mathcal{Q}}\right)_t &\equiv \exp\left\{-\int_0^t \left[-S' \cdot \underline{B}(T_0-u)\right]' \cdot d\underline{W}^{\mathcal{Q}}(u)\right. \\ &\quad \left. -\frac{1}{2}\int_0^t \underline{B}'(T_0-u) \cdot S \cdot S' \cdot \underline{B}(T_0-u)\, du\right\},\end{aligned} \tag{3.15}$$

is a martingale. And, for that purpose, it is sufficient to assume -see, for instance, Lamberton and Lapeyre (1996, remark 4.2.3)- that the *Novikov's condition* (3.12) is satisfied with $i = 0$, i.e.

$$E_{\mathcal{Q}}\left\{\exp\left[\frac{1}{2}\int_0^t \underline{B}'(T_0-u) \cdot \Theta \cdot \underline{B}(T_0-u)\, du\right] \middle| \mathcal{F}_0\right\} < \infty. \tag{3.16}$$

In fact, combining equations (2.1) and (3.6), it follows that the growth condition (3.16) is a natural requirement for the boundedness of pure discount bond' prices.

The sufficient condition (3.16) can be illustrated if a change of numeraire is associated to the above change of probability measure (from $\mathcal{Q}$ to $\mathcal{Q}_0$). Replacing T_i by T_0 in equation (3.10), and dividing both sides by $\frac{P(0,T_0)}{\delta(0)}$, gives

$$\begin{aligned}\frac{P(t,T_0)}{P(0,T_0)}\frac{\delta(0)}{\delta(t)} &= \exp\left[-\frac{1}{2}\int_0^t \underline{B}'(T_0-u) \cdot S \cdot S' \cdot \underline{B}(T_0-u)\, du\right. \\ &\quad \left. + \int_0^t \underline{B}'(T_0-u) \cdot S \cdot d\underline{W}^{\mathcal{Q}}(u)\right],\end{aligned}$$

which when compared with (3.15) yields the usual result:

$$\left(\frac{d\mathcal{Q}_0}{d\mathcal{Q}}\right)_t = \frac{P(t,T_0)}{P(0,T_0)}\frac{\delta(0)}{\delta(t)}.$$

That is, if $P(t,T_0,T_i)$ is a martingale with respect to the probability measure $\mathcal{Q}_0$, and

[12] That is restricted to $\mathcal{F}_t$.

because $P(t,T_0,T_i)$ can be interpreted as the relative price of security $P(t,T_i)$ with respect to the numeraire $P(t,T_0)$, then it is easy to conclude that the numeraire with respect to which Q_0 is a martingale measure must be the price of a pure discount bond maturing at the option expiry date (i.e. the price of the "reference bond"). In fact,

$$\begin{aligned} E_Q\left[\left(\frac{dQ_0}{dQ}\right)_t \middle| \mathcal{F}_s\right] &= E_Q\left[\frac{P(t,T_0)}{P(0,T_0)}\frac{\delta(0)}{\delta(t)}\middle| \mathcal{F}_s\right] \\ &= \frac{\delta(0)}{P(0,T_0)}E_Q\left[\frac{P(t,T_0)}{\delta(t)}\middle| \mathcal{F}_s\right], \end{aligned}$$

and because, under assumption (3.16), $\frac{P(t,T_0)}{\delta(t)}$ is a Q-martingale, then

$$\begin{aligned} E_Q\left[\left(\frac{dQ_0}{dQ}\right)_t \middle| \mathcal{F}_s\right] &= \frac{\delta(0)}{P(0,T_0)}\frac{P(s,T_0)}{\delta(s)} \\ &= \left(\frac{dQ_0}{dQ}\right)_s, \ \forall s \leq t, \end{aligned}$$

which together with the integrability condition[13] $E_Q\left[\left|\left(\frac{dQ_0}{dQ}\right)_t\right|\right] < \infty, \forall t$, ensures that the Radon-Nikodym derivative (3.15) is indeed a martingale.

In order to simplify the notation, we shall define from now on

$$\underline{H_i}'(t) \equiv \left[\underline{B}'(\tau_i) - \underline{B}'(\tau_0)\right]\cdot S, \tag{3.17}$$

where $\underline{H_i}(t) \in \Re^n$ can be interpreted as a vector of forward price volatilities, and thus equation (3.14) becomes

$$\frac{dP(t,T_0,T_i)}{P(t,T_0,T_i)} = \underline{H_i}'(t)\cdot d\underline{W}^{Q_0}(t). \tag{3.18}$$

The vector $\underline{H_i}(t)$ is time-dependent but not state-dependent and, as it will be shown later, this state independence feature of equation (3.18) will greatly simplify the valuation of interest rate contingent claims. Notice that, from equation (3.2), a simple closed-form solution is available for $\underline{H_i}(t)$.

Applying the same reasoning, it is a simple matter to show that, in general,

$$d\underline{W}^{Q_i}(t) = -S'\cdot \underline{B}(\tau_i)\,dt + d\underline{W}^{Q}(t)$$

is a Q_i-measured vector of standard Brownian motion increments in $\Re^n$, where Q_i is the martingale measure associated with the numeraire $P(t,T_i)$, as long as condition (3.12) is

[13] That also follows from assumption (3.16).

met. Equivalently, combining the last equality with equation (3.13), it follows that

$$d\underline{W}^{Q_i}(t) = -\underline{H}_i(t)\,dt + d\underline{W}^{Q_0}(t)\,. \tag{3.19}$$

3.4 Pricing of European bond options

3.4.1 Options on pure discount bonds

Since in subsection 3.3.2 it was shown that Q_0 is a martingale measure with respect to the numeraire $P(t,T_0)$, it is now possible to use result (3.8):

$$\frac{c_t\left[P(t,T_1)\,;K;T_0\right]}{P(t,T_0)} = E_{Q_0}\left\{\left.\frac{c_{T_0}\left[P(T_0,T_1)\,;K;T_0\right]}{P(T_0,T_0)}\right|\mathcal{F}_t\right\},$$

where $c_t\left[P(t,T_1)\,;K;T_0\right]$ represents the time-t Gaussian fair price of an European[14] call on the asset $P(t,T_1)$, with a strike price of K, and expiry date at time $T_0 \leq T_1$. $P(t,T_1)$ is the time-t price of a (unit face value) riskless zero coupon bond, maturing at time T_1, and will hereafter be referred as the "underlying bond". Because the option's value at the expiry date equals its *intrinsic value*, then[15]

$$\begin{aligned} c_{T_0}\left[P(T_0,T_1)\,;K;T_0\right] &= \left[P(T_0,T_1)-K\right]^+ \\ &= \left[P(T_0,T_1)-K\right]1_\varepsilon, \end{aligned}$$

where $\varepsilon = \{\omega \in \Omega :\; P(T_0,T_1)(\omega) > K\}$ is the set of states of the world in which the option ends in-the-money and is exercised.[16] Combining the last two results,

$$\begin{aligned} c_t\left[P(t,T_1)\,;K;T_0\right] &= P(t,T_0)\,E_{Q_0}\left[\left.\frac{P(T_0,T_1)\,1_\varepsilon}{P(T_0,T_0)}\right|\mathcal{F}_t\right] \\ &\quad -KP(t,T_0)\,E_{Q_0}\left[\left.\frac{1_\varepsilon}{P(T_0,T_0)}\right|\mathcal{F}_t\right]. \end{aligned} \tag{3.20}$$

Now, it is necessary to compute the two expected values embodied in the previous

[14] From Merton (1973a, theorems 1 and 2) we know that the value of an American call on a non-dividend paying security is, in general, equal to the value of the correspondent European call. However, when we are dealing with interest rate contingent claims this property is only true if the spot interest rate is positive. Since in any Gaussian term structure model the price process of the "money market account" is not necessarily increasing, then the Q-probability of attaining negative spot interest rates is not zero, and therefore the results obtained in this subsection could only be extended to the pricing of American options if one assumes that such Q-probability is small for reasonable parameters' estimated values.

[15] As usual, $[x]^+ = \max(x,0)\,,\;\forall x \in \Re$.

[16] The dependency on the path (ω) will be omitted hereafter.

equation. Beginning with the second one, and since $P(T_0, T_0) = 1$,

$$E_{Q_0}\left[\frac{1_\varepsilon}{P(T_0,T_0)}\middle|\mathcal{F}_t\right] = E_{Q_0}\left(1_\varepsilon|\mathcal{F}_t\right).$$

In order to compute the first expected value, it is convenient to make another change of probability measure (and thus another change of numeraire). Because $P(t, T_1)$ is a non-dividend paying numeraire[17], and condition (3.16) ensures that $\frac{P(t,T_1)}{P(t,T_0)}$ is a Q_0-martingale, it is possible to define an equivalent martingale measure Q_1 with a Radon-Nikodym derivative with respect to Q_0 equal to

$$\left(\frac{dQ_1}{dQ_0}\right)_{T_0} = \frac{P(T_0,T_1)}{P(0,T_1)}\frac{P(0,T_0)}{P(T_0,T_0)}. \tag{3.21}$$

In fact, $E_{Q_0}\left[\left(\frac{dQ_1}{dQ_0}\right)_{T_0}\middle|\mathcal{F}_0\right] = 1$, as required, because $\frac{P(t,T_1)}{P(t,T_0)}$ is a Q_0-martingale. So, if Y_t represents the time-t price of any attainable contingent claim Y which settles at time T_0, the following conditional expectations formula can be applied:

$$E_{Q_1}\left[\frac{Y_{T_0}}{P(T_0,T_1)}\middle|\mathcal{F}_t\right] = \frac{E_{Q_0}\left[\left(\frac{dQ_1}{dQ_0}\right)_{T_0}\frac{Y_{T_0}}{P(T_0,T_1)}\middle|\mathcal{F}_t\right]}{E_{Q_0}\left[\left(\frac{dQ_1}{dQ_0}\right)_{T_0}\middle|\mathcal{F}_t\right]},$$

assuming that the terminal payoff Y_{T_0} is $\mathcal{F}_{T_0}$-measurable and that the integrability condition $E_{Q_1}\left[\left|\frac{Y_t}{P(t,T_1)}\right|\right] < \infty$ is valid for all t. Or, expressing the Radon-Nikodym derivative $\left(\frac{dQ_1}{dQ_0}\right)_{T_0}$ in terms of the numeraires, and since $\frac{P(t,T_1)}{P(t,T_0)}$ is a Q_0-martingale, then

$$P(t,T_0)\,E_{Q_0}\left[\frac{Y_{T_0}}{P(T_0,T_0)}\middle|\mathcal{F}_t\right] = P(t,T_1)\,E_{Q_1}\left[\frac{Y_{T_0}}{P(T_0,T_1)}\middle|\mathcal{F}_t\right]. \tag{3.22}$$

Applying relation (3.22) to the first expected value in equation (3.20),

$$c_t\left[P(t,T_1);K;T_0\right] = P(t,T_1)\,E_{Q_1}\left(1_\varepsilon|\mathcal{F}_t\right) - K P(t,T_0)\,E_{Q_0}\left(1_\varepsilon|\mathcal{F}_t\right). \tag{3.23}$$

Equation (3.23) is a standard result: it corresponds, for instance, to El Karoui and Rochet (1989, equation (2.4) with $N = 1$). Next proposition computes explicitly the two expectations involved in the last pricing formula, and also presents the corresponding result for European puts.

Proposition 3 *Under the Gaussian specification of the Duffie and Kan (1996) model, the time-t price of an European call on the riskless pure discount bond $P(t,T_1)$, with a strike*

[17]Since it is a zero coupon bond.

price equal to K, and with maturity at time T_0 (such that $t \leq T_0 \leq T_1$) is equal to

$$c_t\left[P(t,T_1);K;T_0\right] = P(t,T_1)\,\Phi\left[d_1(t)\right] - KP(t,T_0)\,\Phi\left[d_0(t)\right], \tag{3.24}$$

with

$$d_1(t) = \frac{\ln\left[\frac{P(t,T_1)}{KP(t,T_0)}\right] + \frac{V_1(\tau_0)}{2}}{\sqrt{V_1(\tau_0)}},$$

$$d_0(t) = d_1(t) - \sqrt{V_1(\tau_0)},$$

$$V_1(\tau_0) = \underline{B}'(T_1 - T_0)\cdot\Delta(T_0 - t)\cdot\underline{B}(T_1 - T_0), \tag{3.25}$$

$\tau_0 = T_0 - t$, and where Φ represents the cumulative density function of the univariate standard normal distribution. The corresponding put price is

$$p_t\left[P(t,T_1);K;T_0\right] = -P(t,T_1)\,\Phi\left[-d_1(t)\right] + KP(t,T_0)\,\Phi\left[-d_0(t)\right]. \tag{3.26}$$

Proof. From the definition of the exercise set ε, and since the expectation of an indicator function results in a probability, then

$$E_{Q_0}\left(1_\varepsilon|\,\mathcal{F}_t\right) = \Pr_{Q_0}\left[P(T_0,T_1) > K|\,\mathcal{F}_t\right],$$

where $\Pr_{Q_0}[\cdot]$ represents a probability computed under the martingale measure Q_0. This is simply the probability, under the martingale measure Q_0, that the option will be exercised at time T_0. Moreover, because $P(T_0,T_0) = 1$ and $P(T_0,T_0,T_1) = \frac{P(T_0,T_1)}{P(T_0,T_0)}$, then

$$E_{Q_0}\left(1_\varepsilon|\,\mathcal{F}_t\right) = \Pr_{Q_0}\left[P(T_0,T_0,T_1) > K|\,\mathcal{F}_t\right].$$

In order to compute the last probability it is necessary to know the probability distribution of the forward price $P(t,T_0,T_1)$ under Q_0. This information can be obtained by using equation (3.18)[18], and applying Itô's lemma to $\ln P(t,T_0,T_1)$:

$$d\ln\left[P(t,T_0,T_1)\right] = \underline{H_1}'(t)\cdot d\underline{W}^{Q_0}(t) - \frac{1}{2}\underline{H_1}'(t)\cdot\underline{H_1}(t)\,dt.$$

Hence,

$$P(t,T_0,T_1) = P(0,T_0,T_1)\exp\left[-\frac{1}{2}V_1(t) + \int_0^t \underline{H_1}'(s)\cdot d\underline{W}^{Q_0}(s)\right], \tag{3.27}$$

[18] With $i = 1$.

where[19]

$$V_1(t) = \int_0^t \underline{H_1}'(s) \cdot \underline{H_1}(s)\, ds. \tag{3.28}$$

Applying equation (3.27) to the time interval $[t, T_0]$ corresponding to the option's life,

$$\begin{aligned}
& E_{Q_0}(1_\varepsilon | \mathcal{F}_t) \\
= & \Pr_{Q_0}\left\{ P(t, T_0, T_1) \exp\left[-\frac{1}{2}V_1(T_0 - t) + \int_t^{T_0} \underline{H_1}'(s) \cdot d\underline{W}^{Q_0}(s)\right] > K \middle| \mathcal{F}_t \right\} \\
= & \Pr_{Q_0}\left\{ \int_t^{T_0} \underline{H_1}'(s) \cdot d\underline{W}^{Q_0}(s) < \ln\left[\frac{P(t, T_0, T_1)}{K}\right] - \frac{1}{2}V_1(\tau_0) \middle| \mathcal{F}_t \right\}.
\end{aligned}$$

Using, for instance, corollary (4.5.6) of Arnold (1992), and assuming that $V_1(\tau_0) < \infty$, it follows that the stochastic integral contained in the last equality is normally distributed with mean zero and variance equal to $V_1(\tau_0)$. That is[20]

$$\left[\int_t^{T_0} \underline{H_1}'(s) \cdot d\underline{W}^{Q_0}(s)\right] \backsim N^1(0, V_1(\tau_0)).$$

Therefore,

$$\begin{aligned}
E_{Q_0}(1_\varepsilon | \mathcal{F}_t) &= \Phi\left\{ \frac{\left[\ln \frac{P(t,T_0,T_1)}{K} - \frac{1}{2}V_1(\tau_0)\right] - 0}{\sqrt{V_1(\tau_0)}} \right\} \\
&= \Phi[d_0(t)]. \qquad (3.29)
\end{aligned}$$

Concerning the second expectation involved in the pricing solution (3.24),

$$\begin{aligned}
E_{Q_1}(1_\varepsilon | \mathcal{F}_t) &= \Pr_{Q_1}[P(T_0, T_1) > K | \mathcal{F}_t] \\
&= \Pr_{Q_1}[P(T_0, T_0, T_1) > K | \mathcal{F}_t],
\end{aligned}$$

where $\Pr_{Q_1}[\cdot]$ represents a probability evaluated under the martingale measure Q_1. This is just the Q_1-probability that the option will be exercised at time T_0, and Q_1 is the equivalent martingale measure obtained from Q_0 by changing the numeraire from the "reference bond" to the "underlying bond". Now, it is necessary to find the probability distribution of $P(T_0, T_0, T_1)$ not under Q_0 but under the measure Q_1.

Using (3.21),

$$\left(\frac{dQ_1}{dQ_0}\right)_{T_0} = \frac{P(T_0, T_0, T_1)}{P(0, T_0, T_1)},$$

[19] $V_1(t)$ is time-homogeneous because $\underline{H_1}(\cdot)$ only depends on the durations vectors $\underline{B}(\cdot)$, which are also time-homogeneous.

[20] The notation $X \backsim N^1(\mu, \sigma^2)$ means that the one-dimensional random variable X is normally distributed, with mean μ, and variance σ^2.

while considering equations (3.27) and (3.28),

$$\left(\frac{dQ_1}{dQ_0}\right)_{T_0} = \exp\left[-\frac{1}{2}\int_0^{T_0} \underline{H_1}'(s)\cdot\underline{H_1}(s)\,ds + \int_0^{T_0} \underline{H_1}'(s)\cdot d\underline{W}^{Q_0}(s)\right].$$

Thus, Girsanov's Theorem implies that

$$d\underline{W}^{Q_1}(t) = -\underline{H_1}(t)\,dt + d\underline{W}^{Q_0}(t),$$

as long as condition

$$\exp\left[\frac{1}{2}\int_0^{T_0} \underline{H_1}'(s)\cdot\underline{H_1}(s)\,ds\right] < \infty$$

is verified, which agrees with the more general formula (3.19). Combining the last equality with equations (3.27) and (3.28),

$$P(t,T_0,T_1) = P(0,T_0,T_1)\exp\left[\frac{1}{2}V_1(t) + \int_0^t \underline{H_1}'(s)\cdot d\underline{W}^{Q_1}(s)\right]. \tag{3.30}$$

Consequently,

$$E_{Q_1}(1_\varepsilon|\mathcal{F}_t) = \Pr_{Q_1}\left\{\int_t^{T_0} \underline{H_1}'(s)\cdot d\underline{W}^{Q_1}(s) < \ln\left[\frac{P(t,T_0,T_1)}{K}\right] + \frac{1}{2}V_1(\tau_0)\middle|\mathcal{F}_t\right\}.$$

Finally, since $\left[\int_t^{T_0} \underline{H_1}'(s)\cdot d\underline{W}^{Q_1}(s)\right] \backsim N^1(0, V_1(\tau_0))$, then

$$\begin{aligned} E_{Q_1}(1_\varepsilon|\mathcal{F}_t) &= \Phi\left\{\frac{\left[\ln\frac{P(t,T_0,T_1)}{K} + \frac{V_1(\tau_0)}{2}\right] - 0}{\sqrt{V_1(\tau_0)}}\right\} \\ &= \Phi[d_1(t)]. \end{aligned}$$

Combining the last equality with equations (3.29) and (3.23), the pricing solution (3.24) follows immediately.

In order to obtain the European put valuation formula (3.26), it is only necessary to substitute equation (3.24) into the well known put-call parity for European options on zero coupon bonds,

$$c_t[P(t,T_1);K;T_0] - p_t[P(t,T_1);K;T_0] = P(t,T_1) - KP(t,T_0),$$

yielding

$$p_t[P(t,T_1);K;T_0] = -P(t,T_1)[1-\Phi(d_1(t))] + KP(t,T_0)[1-\Phi(d_0(t))].$$

Since the normal distribution is symmetric[21], equation (3.26) is obtained.

Finally, the explicit solution for the volatility term $V_1(\tau_0)$ follows when combining expressions (3.28), (3.17) and (3.2):

$$\begin{aligned} V_1(\tau_0) &= \int_t^{T_0} \underline{G}' \cdot a^{-1} \cdot \left[e^{a(T_0-u)} - e^{a(T_1-u)}\right] \cdot S \cdot S' \\ &\quad \cdot \left[e^{a'(T_0-u)} - e^{a'(T_1-u)}\right] \cdot (a^{-1})' \cdot \underline{G} du. \end{aligned}$$

Since $\Theta \equiv S \cdot S'$,

$$\begin{aligned} V_1(\tau_0) &= \underline{G}' \cdot a^{-1} \cdot \left\{ \int_t^{T_0} \left[e^{a(T_0-u)} \cdot \Theta \cdot e^{a'(T_0-u)} - e^{a(T_0-u)} \cdot \Theta \cdot e^{a'(T_1-u)} \right.\right. \\ &\quad \left.\left. - e^{a(T_1-u)} \cdot \Theta \cdot e^{a'(T_0-u)} + e^{a(T_1-u)} \cdot \Theta \cdot e^{a'(T_1-u)} \right] du \right\} \cdot (a^{-1})' \cdot \underline{G}. \end{aligned}$$

Because the first integral on the right-hand-side of the last equation corresponds to matrix $\Delta(T_0 - t)$, and converting all the other integrals as functions also of $\Delta(T_0 - t)$, yields:

$$\begin{aligned} V_1(\tau_0) &= \underline{G}' \cdot a^{-1} \cdot \left[\Delta(T_0 - t) - \Delta(T_0 - t) \cdot e^{a'(T_1 - T_0)} \right. \\ &\quad \left. - e^{a(T_1 - T_0)} \cdot \Delta(T_0 - t) + e^{a(T_1 - T_0)} \cdot \Delta(T_0 - t) \cdot e^{a'(T_1 - T_0)} \right] \cdot (a^{-1})' \cdot \underline{G}. \end{aligned}$$

Rearranging terms, the analytical solution (3.25) follows easily. ■

Equation (10.2) of Duffie and Kan (1996) is a special case of formula (3.24), when $n = 2$ and $\alpha_i = 1$ $(for\ i = 1, 2)$. Besides the functional form of the volatility term $V_1(\tau_0)$, which is related to the particular specification of the Duffie and Kan (1996) Gaussian model under analysis, equation (3.24) is just a generalization of a standard result for Gaussian term structure models in the one-dimensional case -see, for instance, Jamshidian (1991, corollary 1). It is also similar to the Black and Scholes (1973) stock option pricing formula, because both formulae assume a log-normal distribution for the underlying asset, but it has three main differences: formula (3.24) does not assume deterministic interest rates; the volatility used to price the option is time-dependent, and thus not constant; and finally, the "pull-to-par" phenomena is taken into account, since $\underline{B}(0) = \underline{0}$ and $A(0) = 0$. However, and as it is usual in any Gaussian framework, the existence of negative interest rates is not precluded.

3.4.2 Interest rate caps, floors, collars and yield options

The results obtained so far for European options on default-free pure discount bonds can be easily generalized for European options on *nominal*[22] "money-market" forward interest

[21] That is $\Phi(-x) = 1 - \Phi(x)$.
[22] That is rates involving simple (as opposed to continuous) compounding.

rates, under the Duffie and Singleton (1997) assumption of symmetric counterparty credit risk. In fact, Duffie and Singleton (1997) have shown that, as long as the counterparties have symmetric probabilities of default, any term structure model previously formulated for government yield curves can also be used to price defaultable interest rate contingent claims, after the short-term interest rate process is adjusted for default and liquidity factors. Therefore, the symmetric credit risk assumption as well as the other four implicit hypothesis described in Duffie and Singleton (1997, section 1) will be adopted hereafter whenever the pricing of LIBOR-rate derivatives is dealt with. Note however that, since the risk-free short-term interest rate must be replaced by a risk- and liquidity-adjusted instantaneous interest rate process when the term structure model is applied to LIBOR-rate derivatives, it is not possible to price simultaneously riskless and defaultable interest rate contingent claims.

The value of an interest rate cap can be decomposed into a portfolio of caplets. The terminal payoff of a standard caplet for the compounding period $(t_{i+1} - t_i)$, with $t_{i+1} > t_i$, occurs at time t_{i+1}, and is equal to:

$$[R(t_i, t_{i+1}) - k]^+ (t_{i+1} - t_i),$$

where $R(t_i, t_{i+1})$ is the time-t_i spot interest rate (with a compounding period of $(t_{i+1} - t_i)$ years) for the period $(t_{i+1} - t_i)$, k is the cap rate, and the cap is assumed to have a unit contract size. Therefore, the time-t value of the caplet, with $t \leq t_i$, is equal to the price of an European call on the time-t forward rate for the period $(t_{i+1} - t_i)$, with a strike equal to k, with maturity at time t_{i+1}, and with a contract size of $(t_{i+1} - t_i)$. However, it is well known -see, for instance, Baxter and Rennie (1996, page 171)- that the same caplet can be valued as an European put with maturity at time t_i, with a contract size of $[1 + (t_{i+1} - t_i) k]$, with a strike price of $\frac{1}{1+(t_{i+1}-t_i)k}$, and on a pure discount bond with maturity at time t_{i+1}. That is the time-t value of the caplet corresponds to

$$[1 + (t_{i+1} - t_i) k] p_t \left[P(t, t_{i+1}); \frac{1}{1 + (t_{i+1} - t_i) k}; t_i \right], \tag{3.31}$$

which can be computed, for the deterministic volatility specification of the Duffie and Kan (1996) model, using equation (3.26).

Similarly, the time-t value of a floorlet for the compounding period $(t_{i+1} - t_i)$, with $t_{i+1} > t_i$, can be shown to be equal to the price of an European call with maturity at time t_i, with a contract size of $[1 + (t_{i+1} - t_i) k]$, with a strike price of $\frac{1}{1+(t_{i+1}-t_i)k}$, and on a pure discount bond with maturity at time t_{i+1}:

$$[1 + (t_{i+1} - t_i) k] c_t \left[P(t, t_{i+1}); \frac{1}{1 + (t_{i+1} - t_i) k}; t_i \right], \tag{3.32}$$

where k is now a floor rate. Consequently, an interest rate floor (i.e. a portfolio of floorlets) can also be valued, under the Gaussian version of the Duffie and Kan (1996) model, using proposition 3. The same can be said about the valuation of interest rate borrowing/lending collars, since their value is decomposable into a long/short cap and a short/long floor, respectively.

In order to value European yield call and put options, with settlement in arrears (i.e. with payoff generated at time t_{i+1}), on the time-t nominal forward rate for the period $(t_{i+1} - t_i)$, with a strike equal to k, with maturity at time t_i, and with a unit contract size, it is simply necessary to divide the valuation formulae previously given for caplets and floorlets by the compounding period $(t_{i+1} - t_i)$.

3.4.3 Options on coupon-bearing bonds

Rank 1 approximation

Representing by $c_t [S(t); X; T_0]$ the time-t fair price of an European call with a strike price of X, expiry date at time T_0, and on a coupon bond with present value $S(t)$, and because Q_0 is a martingale measure with respect to the numeraire $P(t, T_0)$,

$$c_t [S(t); X; T_0] = P(t, T_0) E_{Q_0} \left\{ \frac{c_{T_0} [S(T_0); X; T_0]}{P(T_0, T_0)} \middle| \mathcal{F}_t \right\}. \tag{3.33}$$

Since at time T_0 the option's *time-value* is zero, and thus its price corresponds to its *intrinsic value*, then:

$$\begin{aligned} c_{T_0} [S(T_0); X; T_0] &= [S(T_0) - X]^+ \\ &= \left[\sum_{i=1}^{N_0} k_i P(T_0, T_i) - X \right]^+, \end{aligned}$$

where $T_0 < T_1 < \ldots < T_{N_0}$, and N_0 represents the number of cash flows k_i generated by the underlying coupon bond from time T_0 and until the bond's expiry date (T_{N_0}). Defining by $\theta = \{\omega \in \Omega : S(T_0)(\omega) > X\}$ the set of states of the world in which the option ends in-the-money, the option's terminal value may be rewritten as:

$$c_{T_0} [S(T_0); X; T_0] = \left[\sum_{i=1}^{N_0} k_i P(T_0, T_i) - X \right] 1_\theta.$$

Using this result,

$$c_t\left[S(t);X;T_0\right] = \sum_{i=1}^{N_0} k_i P(t,T_0)\, E_{Q_0}\left[\left.\frac{P(T_0,T_i)\,1_\theta}{P(T_0,T_0)}\right|\mathcal{F}_t\right] \tag{3.34}$$
$$-XP(t,T_0)\,E_{Q_0}\left[\left.\frac{1_\theta}{P(T_0,T_0)}\right|\mathcal{F}_t\right].$$

Because $P(T_0,T_0)=1$, then

$$E_{Q_0}\left[\left.\frac{1_\theta}{P(T_0,T_0)}\right|\mathcal{F}_t\right] = \Pr_{Q_0}\left[S(T_0)>X|\,\mathcal{F}_t\right], \tag{3.35}$$

which corresponds to the Q_0-probability that the option will be in-the-money at time T_0. On the other hand, because $P(t,T_i)$ is a non-dividend paying numeraire, and condition (3.16) ensures that $\frac{P(t,T_i)}{P(t,T_0)}$ is a Q_0-martingale, we can define an equivalent martingale measure Q_i with a Radon-Nikodym derivative with respect to Q_0 equal to[23]

$$\left(\frac{dQ_i}{dQ_0}\right)_{T_0} = \frac{P(T_0,T_i)}{P(0,T_i)}\frac{P(0,T_0)}{P(T_0,T_0)}, \tag{3.36}$$

such that

$$P(t,T_0)\,E_{Q_0}\left[\left.\frac{Y_{T_0}}{P(T_0,T_0)}\right|\mathcal{F}_t\right] = P(t,T_i)\,E_{Q_i}\left[\left.\frac{Y_{T_0}}{P(T_0,T_i)}\right|\mathcal{F}_t\right],$$

where Y_t represents the time-t price of any attainable contingent claim Y which settles at time T_0, and subject to the appropriate integrability and measurability conditions (the proof is identical to the one already given for equation (3.22)). Hence,

$$P(t,T_0)\,E_{Q_0}\left[\left.\frac{P(T_0,T_i)\,1_\theta}{P(T_0,T_0)}\right|\mathcal{F}_t\right] = P(t,T_i)\,E_{Q_i}\left(1_\theta|\,\mathcal{F}_t\right)$$
$$= P(t,T_i)\Pr_{Q_i}\left[S(T_0)>X|\,\mathcal{F}_t\right], \tag{3.37}$$

where the second term on the right-hand-side is the Q_i-probability that the option will be in-the-money at time T_0.

Substituting (3.35) and (3.37) into (3.34), a probability valuation formula is obtained for the European call option:

$$c_t\left[S(t);X;T_0\right] = \sum_{i=1}^{N_0} k_i P(t,T_i)\Pr_{Q_i}\left[S(T_0)>X|\,\mathcal{F}_t\right] - XP(t,T_0)\Pr_{Q_0}\left[S(T_0)>X|\,\mathcal{F}_t\right]. \tag{3.38}$$

Next proposition computes explicitly all probabilities involved in the previous expression, and offers an equivalent pricing formula for the European put option.

[23] $E_{Q_0}\left[\left.\left(\frac{dQ_i}{dQ_0}\right)_{T_0}\right|\mathcal{F}_0\right]=1$, as required, because $\frac{P(t,T_i)}{P(t,T_0)}$ is assumed to be a Q_0-martingale.

Proposition 4 *Under the Gaussian specification of the Duffie and Kan (1996) model, the time-t price of an European call with a strike price equal to X, with maturity at time T_0 ($\geq t$), and on a riskless coupon-bearing bond with present value $S(t)$ and generating N_0 cash flows of value k_i at times T_i (such that $T_i \geq T_0, i = 1, \ldots, N_0$), is approximately equal to*

$$c_t\left[S(t); X; T_0\right] \cong \sum_{i=1}^{N_0} k_i P(t, T_i) \Phi\left[d_i(t)\right] - X P(t, T_0) \Phi\left[d_0(t)\right], \tag{3.39}$$

where

$$d_0(t) : \sum_{i=1}^{N_0} k_i P(t, T_i) \exp\left[-\frac{1}{2} V_i(\tau_0) - \sqrt{V_i(\tau_0)} d_0(t)\right] = X P(t, T_0), \tag{3.40}$$

$$d_i(t) = d_0(t) + \sqrt{V_i(\tau_0)},$$

and

$$V_i(\tau_0) = \underline{B}'(T_i - T_0) \cdot \Delta(T_0 - t) \cdot \underline{B}(T_i - T_0). \tag{3.41}$$

The corresponding put price is approximated by

$$p_t\left[S(t); X; T_0\right] \cong -\sum_{i=1}^{N_0} k_i P(t, T_i) \Phi\left[-d_i(t)\right] + X P(t, T_0) \Phi\left[-d_0(t)\right]. \tag{3.42}$$

Proof. From the definitions of θ and $S(T_0)$, and since $P(T_0, T_0) = 1$,

$$\Pr_{Q_0}\left[S(T_0) > X \mid \mathcal{F}_t\right] = \Pr_{Q_0}\left[\sum_{i=1}^{N_0} k_i P(T_0, T_0, T_i) > X \,\middle|\, \mathcal{F}_t\right].$$

In order to proceed it is necessary to know the probability distribution of the forward price $P(t, T_0, T_i)$ under Q_0, which is precisely given by equations (3.27) and (3.28) when T_1 is substituted by T_i:

$$P(t, T_0, T_i) = P(0, T_0, T_i) \exp\left[-\frac{1}{2} V_i(t) + \int_0^t \underline{H_i}'(s) \cdot d\underline{W}^{Q_0}(s)\right], \tag{3.43}$$

where

$$V_i(t) = \int_0^t \underline{H_i}'(s) \cdot \underline{H_i}(s)\, ds. \tag{3.44}$$

Thus, $P(T_0, T_0, T_i)$ is given by equation (3.43) when the time interval under consideration is $[t, T_0]$ instead of $[0, t]$, and therefore

$$\Pr_{Q_0}\left[S(T_0) > X \mid \mathcal{F}_t\right] = \Pr_{Q_0}\left\{\sum_{i=1}^{N_0} k_i P(t, T_0, T_i)\right.$$

$$\exp\left[-\frac{1}{2}V_i(\tau_0) - \int_t^{T_0}\left(-\underline{H_i}'(s)\right)\cdot d\underline{W}^{Q_0}(s)\right] > X \middle| \mathcal{F}_t\right\}.$$

Because $\left[\int_t^{T_0}\left(-\underline{H_i}'(s)\right)\cdot d\underline{W}^{Q_0}(s)\right] \backsim N^1(0, V_i(\tau_0))$, and since equations (3.2) and (3.17) show that the volatility vectors $\underline{H_i}'(t)$ possess an exponential form, then -as argued by El Karoui and Rochet (1989, page 22)- the random variables $\left[\int_t^{T_0}\left(-\underline{H_i}'(s)\right)\cdot d\underline{W}^{Q_0}(s)\right]$ can be thought of as being proportional, that is

$$\int_t^{T_0}\left(-\underline{H_i}'(s)\right)\cdot d\underline{W}^{Q_0}(s) \cong \sqrt{V_i(\tau_0)}Z,\ for\ i = 1,\ldots,N_0,$$

and where $Z \backsim N^1(0,1)$.[24] Combining the last two results,

$$\begin{aligned} &\Pr_{Q_0}\left[S(T_0) > X \mid \mathcal{F}_t\right] \\ \cong\ &\Pr_{Q_0}\left\{\sum_{i=1}^{N_0} k_i P(t,T_i)\exp\left[-\frac{1}{2}V_i(\tau_0) - \sqrt{V_i(\tau_0)}Z\right] > X P(t,T_0) \middle| \mathcal{F}_t\right\}. \end{aligned}$$

And, since the left-hand-side of the above inequality is a decreasing function of Z,

$$\begin{aligned} \Pr_{Q_0}\left[S(T_0) > X \mid \mathcal{F}_t\right] &\cong \Pr_{Q_0}\left[Z < d_0(t) \mid \mathcal{F}_t\right] \\ &\cong \Phi\left[d_0(t)\right], \end{aligned} \tag{3.45}$$

where $d_0(t)$ is implicitly defined by (3.40).

Concerning the other N_0 probabilities contained in the valuation formula (3.39), from the definitions of θ and $S(T_0)$, and since $P(T_0,T_0) = 1$,

$$\Pr_{Q_i}\left[S(T_0) > X \mid \mathcal{F}_t\right] = \Pr_{Q_i}\left[\sum_{i=1}^{N_0} k_i P(T_0,T_0,T_i) > X \middle| \mathcal{F}_t\right].$$

In order to derive the probability distribution of the forward price $P(t,T_0,T_i)$ under the martingale measure Q_i, relation (3.19) can be substituted into equation (3.43):

$$P(t,T_0,T_i) = P(0,T_0,T_i)\exp\left[-\frac{1}{2}V_i(t) + V_i(t) - \int_0^t\left(-\underline{H_i}'(s)\right)\cdot d\underline{W}^{Q_i}(s)\right]. \tag{3.46}$$

[24] This assumption is exactly equivalent to the *rank 1* approximation suggested by Brace and Musiela (1994a, equation 6.1), with $\sqrt{V_i(\tau_0)} = \gamma_i$, accordingly to their notation. The scalar γ_i can be obtained as $\gamma_i = \sqrt{\delta_1}\left(\underline{v_1}\right)_i$, where δ_1 is the first eigenvalue and $\left(\underline{v_1}\right)_i$ is the i^{th}-element of the first eigenvector of the $(N_0 \times N_0)$ matrix

$$\left\{COV\left[lnP(T_0,T_i), lnP(T_0,T_j) \mid \mathcal{F}_t\right]\right\}_{i,j=1,\ldots,N_0}.$$

Hence,

$$\begin{aligned}\Pr_{Q_i}[S(T_0) > X|\mathcal{F}_t] &= \Pr_{Q_i}\left\{\sum_{i=1}^{N_0} k_i P(t,T_0,T_i)\right.\\ &\quad \left.\exp\left[-\frac{1}{2}V_i(\tau_0) + V_i(\tau_0) - \int_t^{T_0}\left(-\underline{H_i}'(s)\right)\cdot d\underline{W}^{Q_i}(s)\right] > X\middle|\mathcal{F}_t\right\}.\end{aligned}$$

Moreover, since $\left[\int_t^{T_0}\left(-\underline{H_i}'(s)\right)\cdot d\underline{W}^{Q_i}(s)\right] \backsim N^1(0,V_i(\tau_0))$, because the volatility vectors $\underline{H_i}'(t)$ possess an exponential form, and assuming proportionality, that is

$$\int_t^{T_0}\left(-\underline{H_i}'(s)\right)\cdot d\underline{W}^{Q_i}(s) \cong \sqrt{V_i(\tau_0)}Z,$$

with $Z \backsim N^1(0,1)$, then

$$\begin{aligned}\Pr_{Q_i}[S(T_0) > X|\mathcal{F}_t] &\cong \Pr_{Q_i}\left\{\sum_{i=1}^{N_0} k_i P(t,T_i)\right.\\ &\quad \left.\exp\left[-\frac{1}{2}V_i(\tau_0) - \sqrt{V_i(\tau_0)}\left(Z - \sqrt{V_i(\tau_0)}\right)\right] > XP(t,T_0)\middle|\mathcal{F}_t\right\}.\end{aligned}$$

Since the left-hand-side of the inequality contained in the above probability is a decreasing function of $\left(Z - \sqrt{V_i(\tau_0)}\right)$, then

$$\begin{aligned}\Pr_{Q_i}[S(T_0) > X|\mathcal{F}_t] &\cong \Pr_{Q_i}\left[Z - \sqrt{V_i(\tau_0)} < d_0(t)\middle|\mathcal{F}_t\right]\\ &\cong \Phi[d_i(t)].\end{aligned} \tag{3.47}$$

Combining equations (3.38), (3.45), and (3.47) yields the valuation formula (3.39).

The European put valuation equation follows from the put-call parity for European options on coupon bonds, which can be stated as:

$$c_t[S(t);X;T_0] - p_t[S(t);X;T_0] = \left[S(t) - \sum_{t\le T_i\le T_0} k_i P(t,T_i)\right] - XP(t,T_0). \tag{3.48}$$

Combining the above equality with equation (3.39), considering that the present value of the underlying coupon bond is equal to $S(t) = \sum_{t<T_i<T_{N_0}} k_i P(t,T_i)$, and solving for the price of the European put,

$$p_t[S(t);X;T_0] \cong -\sum_{i=1}^{N_0} k_i P(t,T_i)\cdot[1-\Phi(d_i(t))] + XP(t,T_0)[1-\Phi(d_0(t))].$$

Since the normal distribution is symmetric, the pricing formula (3.42) is obtained.

Concerning the analytical solution of $V_i(\tau_0)$, formula (3.41) follows from equation (3.25) simply by replacing T_1 with T_i. ■

Remark 3 *Equation (3.39) is exactly equivalent to formula (4.7) of El Karoui and Rochet (1989) when $n = 1$, although the volatility structure has been adapted to the Gaussian specification of the Duffie and Kan (1996) model.*

Remark 4 *Alternatively, $c_t[S(t); X; T_0]$ and $p_t[S(t); X; T_0]$ can also be written as a sum of N_0 European call and put options, respectively, on pure discount bonds $P(t, T_i)$, with maturity at time T_0, with a contract size of k_i monetary units, and with an adjusted strike price. From the definition of $d_0(t)$,*

$$X = \sum_{i=1}^{N_0} k_i \frac{P(t, T_i) \exp\left[-\frac{1}{2} V_i(\tau_0) - \sqrt{V_i(\tau_0)} d_0(t)\right]}{P(t, T_0)}.$$

Substituting X for the above expression into equation (3.39),

$$\begin{aligned} c_t[S(t); X; T_0] &\cong \sum_{i=1}^{N_0} k_i \left\{P(t, T_i) \Phi[d_i(t)] - X_i P(t, T_0) \Phi[d_0(t)]\right\} \\ &\cong \sum_{i=1}^{N_0} k_i c_t[P(t, T_i); X_i; T_0], \end{aligned} \tag{3.49}$$

where the adjusted strikes are given by

$$X_i = \frac{P(t, T_i) \exp\left[-\frac{1}{2} V_i(\tau_0) - \sqrt{V_i(\tau_0)} d_0(t)\right]}{P(t, T_0)}.$$

Similarly,

$$\begin{aligned} p_t[S(t); X; T_0] &\cong \sum_{i=1}^{N_0} k_i \left\{-P(t, T_i) \Phi[-d_i(t)] + X_i P(t, T_0) \Phi[-d_0(t)]\right\} \\ &\cong \sum_{i=1}^{N_0} k_i p_t[P(t, T_i); X_i; T_0]. \end{aligned} \tag{3.50}$$

Lognormal approximation

As an alternative to the *rank 1* approximation presented before, Pang (1996) suggests approximating the distribution of the underlying coupon-bearing bond price to a lognormal distribution, by matching its first two moments. Although it is well known that the sum of lognormal random variables is not lognormally distributed, the price of a coupon bond weights mostly its last pure discount bond price component (i.e. the one associated with the redemption of the bond's face value and with the payment of the last coupon). Therefore, the intuition behind the lognormal approximation proposed by Pang (1996) is that the distribution of the coupon-bearing bond price should essentially depend upon the probabilistic behavior of its last component, which is lognormally distributed for the Gaussian framework under analysis.

Specifically, equation (3.33) is approximated by

$$c_t [S(t); X; T_0] \cong P(t, T_0) E_{Q_0} \left[\left(e^Y - X\right)^+ \Big| \mathcal{F}_t \right], \tag{3.51}$$

where $Y \equiv \ln [S(T_0)] \backsim N^1 (a, b^2)$ with $a = \ln \left(\frac{m}{\sqrt{1+\frac{v^2}{m^2}}} \right)$, $b^2 = \ln \left(1 + \frac{v^2}{m^2}\right)$, $m = E_{Q_0} [S(T_0)| \mathcal{F}_t]$, and $v^2 = E_{Q_0} \left\{ [S(T_0)]^2 \Big| \mathcal{F}_t \right\} - m^2$. Next proposition solves equation (3.51) and computes explicitly both the mean and the variance of the time-T_0 underlying coupon-bearing bond price.

Proposition 5 *Under the Gaussian specification of the Duffie and Kan (1996) model, the time-t price of an European call with a strike price equal to X, with maturity at time T_0 ($\geq t$), and on a riskless coupon-bearing bond with present value $S(t)$ and generating N_0 cash flows of value k_i at times T_i (such that $T_i \geq T_0, i = 1, \ldots, N_0$), can be approximated by*

$$c_t [S(t); X; T_0] \cong \sum_{i=1}^{N_0} k_i P(t, T_i) \Phi \left[\frac{a - \ln(X)}{b} + b \right] - X P(t, T_0) \Phi \left[\frac{a - \ln(X)}{b} \right], \tag{3.52}$$

where

$$a = \ln \left(\frac{m}{\sqrt{1 + \frac{v^2}{m^2}}} \right), \qquad b^2 = \ln \left(1 + \frac{v^2}{m^2} \right),$$

$$m = \sum_{i=1}^{N_0} k_i \frac{P(t, T_i)}{P(t, T_0)}, \tag{3.53}$$

and

$$v^2 = \sum_{j=1}^{N_0} \sum_{i=1}^{N_0} k_j k_i \frac{P(t, T_j) P(t, T_i)}{[P(t, T_0)]^2} \exp [M_{ij}(t)] - m^2, \tag{3.54}$$

with

$$M_{ij}(t) = \underline{B}'(T_j - T_0) \cdot \Delta(T_0 - t) \cdot \underline{B}(T_i - T_0). \tag{3.55}$$

The corresponding put price is approximately equal to

$$p_t [S(t); X; T_0] \cong X P(t, T_0) \Phi \left[\frac{\ln(X) - a}{b} \right] - \sum_{i=1}^{N_0} k_i P(t, T_i) \Phi \left[\frac{\ln(X) - a}{b} - b \right]. \tag{3.56}$$

Proof. Under the assumption $Y \backsim N^1 (a, b^2)$, equation (3.51) can be rewritten under an equivalent integral form:

$$c_t\left[S(t);X;T_0\right]\cong P(t,T_0)\int_{\ln(X)}^{\infty}\left(e^Y-X\right)\frac{1}{b\sqrt{2\pi}}\exp\left[-\frac{(Y-a)^2}{2b^2}\right]dY.$$

Solving this integral explicitly, a pricing solution, similar to the one contained in Pang (1996, proposition 5), is obtained:

$$c_t\left[S(t);X;T_0\right]\cong P(t,T_0)\left\{e^{a+\frac{b^2}{2}}\Phi\left[\frac{a-\ln(X)}{b}+b\right]-X\Phi\left[\frac{a-\ln(X)}{b}\right]\right\}. \tag{3.57}$$

In order to compute m and v^2, it is convenient to rewrite the price of the underlying coupon bond at the option's expiry date using equation (3.43):

$$S(T_0)=\sum_{i=1}^{N_0}k_i\frac{P(t,T_i)}{P(t,T_0)}\exp\left[-\frac{1}{2}V_i(\tau_0)+\int_t^{T_0}\underline{H_i}'(s)\cdot d\underline{W}^{Q_0}(s)\right]. \tag{3.58}$$

Taking expectations,

$$m=\sum_{i=1}^{N_0}k_i\frac{P(t,T_i)}{P(t,T_0)}\exp\left[-\frac{1}{2}V_i(\tau_0)\right]E_{Q_0}\left\{\exp\left[\int_t^{T_0}\underline{H_i}'(s)\cdot d\underline{W}^{Q_0}(s)\right]\middle|\mathcal{F}_t\right\},$$

because $\left[\int_t^{T_0}\underline{H_i}'(s)\cdot d\underline{W}^{Q_0}(s)\right]\backsim N^1(0,V_i(\tau_0))$, and since the expectation on the right-hand-side of the last equality corresponds to the moment generating function of the normal random variable inside the exponential, formula (3.53) follows immediately. Using again equation (3.58),

$$\begin{aligned}
[S(T_0)]^2 &= \sum_{j=1}^{N_0}\sum_{i=1}^{N_0}k_jk_i\frac{P(t,T_j)P(t,T_i)}{[P(t,T_0)]^2}\exp\left[-\frac{1}{2}V_j(\tau_0)-\frac{1}{2}V_i(\tau_0)\right]\\
&\quad\exp\left\{\int_t^{T_0}\left[\underline{H_j}'(s)+\underline{H_i}'(s)\right]\cdot d\underline{W}^{Q_0}(s)\right\},
\end{aligned}$$

and taking expectations,

$$E_{Q_0}\left\{[S(T_0)]^2\middle|\mathcal{F}_t\right\}=\sum_{j=1}^{N_0}\sum_{i=1}^{N_0}k_jk_i\frac{P(t,T_j)P(t,T_i)}{[P(t,T_0)]^2}\exp\left[M_{ij}(t)\right], \tag{3.59}$$

with

$$\begin{aligned}
M_{ij}(t) &= -\frac{1}{2}V_j(\tau_0)-\frac{1}{2}V_i(\tau_0)+\frac{1}{2}\int_t^{T_0}\left\|\left[\underline{H_j}(s)+\underline{H_i}(s)\right]\right\|^2ds\\
&= \int_t^{T_0}\underline{H_j}'(s)\cdot\underline{H_i}(s)\,ds,
\end{aligned}$$

where $\|\cdot\|$ denotes the Euclidean norm in $\Re^n$. Finally, using definitions (3.2) and (3.17), while expressing all the resulting integrals in terms of the matrix $\Delta(T_0 - t)$, equation (3.55) follows.

Because $e^{a+\frac{b^2}{2}} = m$, combining equations (3.57) and (3.53) yields the pricing solution (3.52). Similarly, substituting formula (3.52) into the put-call parity (3.48), equation (3.56) arises. ■

Comparison

In order to compare the performance of both *rank 1* and *lognormal* approximations, the three-factor affine and Gaussian model estimated by Babbs and Nowman (1999, Table 2) will be used. For this purpose, such model was converted into the specification offered by equations (2.2) and (3.1), i.e.

$$
\begin{aligned}
f &= 0.0701, \underline{G} = -\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, a = diag\{-0.6553, -0.0705, -0.0525\}, \\
\underline{b} &= \begin{bmatrix} -0.003385 \\ 0.002186 \\ -0.001947 \end{bmatrix}, \Sigma = \begin{bmatrix} 0.0214 & 0 & 0 \\ -0.017755 & 0.006479 & 0 \\ 0.014267 & -0.004647 & 0.007882 \end{bmatrix}, \underline{\alpha} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},
\end{aligned}
$$

and the state variables' values

$$
\underline{X}(t) = \begin{bmatrix} -0.005475 & 0.006897 & -0.001374 \end{bmatrix}'
$$

were defined in order to produce continuously compounded spot interest rates around 7%, for maturities up to 21 years (the maximum deviation from the flat yield curve level of 7% is less than one basis point).

Table 3.1 values, for different strikes, European options with a maturity of 0.5 years, on a coupon bond with a maturity of 2.5 years, a face value of one monetary unit, semi-annually coupons, and a coupon rate of 7% per year. The *rank 1* approximate option prices are computed from proposition 4, and the *lognormal* approximation is implemented using proposition 5, but only the corresponding percentage pricing errors are presented. The exact price of each option was estimated through standard Monte Carlo simulation, using the usual Euler discretization of equation (3.1), dividing the option's maturity into 100 time steps, generating independent normal variates through the Box-Muller algorithm, running 1,000,000 simulations, and computing analytically the option's terminal payoff

Table 3.1: Valuation of European options with a maturity of 0.5 years, on a unit face value coupon-bearing bond paying 4 semi-annually coupons of 0.035 each after the option's expiry date, using the Babbs and Nowman (1999, Table 2) model

	Standard Monte Carlo (exact option prices)			Percentage Absolute Pricing Errors			
				Rank 1 Appr.		Lognormal Appr.	
Strikes	Call	Put	Std. errors	Call	Put	Call	Put
0.858164	0.134811	0	0.000014	0.013%	NA	0.013%	NA
0.875589	0.117986	0	0.000014	0.015%	NA	0.015%	NA
0.884433	0.109446	0	0.000014	0.016%	NA	0.016%	NA
0.902390	0.092107	0	0.000014	0.018%	NA	0.018%	NA
0.911506	0.083305	0	0.000014	0.022%	NA	0.022%	NA
0.930013	0.065435	0	0.000014	0.026%	NA	0.026%	NA
0.939407	0.056364	0	0.000014	0.032%	NA	0.032%	NA
0.958481	0.037956	0	0.000014	0.047%	NA	0.047%	NA
0.968162	0.028683	0.000083	0.000014	0.059%	2.410%	0.059%	1.205%
0.987820	0.011569	0.001935	0.000011	0.069%	0.362%	0.043%	0.258%
0.997798	0.005489	0.005489	0.000008	0.018%	0.018%	0.091%	0.091%
1.017854	0.000518	0.019884	0.000002	0.193%	0.010%	0.579%	0.015%
1.028032	0.000093	0.029287	0.000001	0.000%	0.000%	1.075%	0.003%
1.048696	0.000001	0.049147	0	0.000%	0.000%	0.000%	0.000%
1.059183	0	0.059272	0	NA	0.000%	NA	0.000%
1.080472	0	0.079829	0	NA	0.000%	NA	0.000%
1.091277	0	0.090262	0	NA	0.000%	NA	0.000%
1.113211	0	0.111442	0	NA	0.000%	NA	0.000%
1.124344	0	0.122191	0	NA	0.000%	NA	0.000%
1.146943	0	0.144012	0	NA	0.000%	NA	0.000%

Monte Carlo: 1,000,000 simulations with 100 time steps per option maturity.
The strike in bold is the forward price of the underlying coupon bond for the option's expiry date (i.e. the ATM strike).
"NA" stands for not available.
Percentage absolute pricing errors are equal to the approximate price -as given by propositions 4 or 5- over the Monte Carlo estimate, minus one.

from proposition 1. Besides the Monte Carlo price estimates, the standard errors for the call prices' estimates are also shown.

The results show that, for short maturity European options on short-term coupon bonds, the pricing differences between the two approximations are negligeable, irrespective of the option's moneyness. Similarly, table 3.2 prices, for different strikes, European calls and puts expiring in 5 years, on a coupon bond with the same features as before, except that its maturity is now equal to 6 years. Again, the same conclusions prevail. That is, for (short-term or long-term) European options on coupon-bearing bonds expiring close to the option's maturity, both *rank 1* and *lognormal* approximations produce similar pricing errors.

However, when the maturity of the underlying coupon-bearing bond is significantly longer than the life of the option, the performance of both approximations seems to depend on the option's moneyness. Table 3.3 values, for different strikes, European options with

Table 3.2: Valuation of European options with a maturity of 5 years, on a unit face value coupon-bearing bond paying 2 semi-annually coupons of 0.035 each after the option's expiry date, using the Babbs and Nowman (1999, Table 2) model

	Standard Monte Carlo (exact option prices)			Percentage Absolute Pricing Errors			
				Rank 1 Appr.		Lognormal Appr.	
Strikes	Call	Put	Std. errors	Call	Put	Call	Put
0.859012	0.098487	0	0.000017	0.019%	NA	0.019%	NA
0.876453	0.086196	0	0.000016	0.021%	NA	0.021%	NA
0.885307	0.079957	0	0.000016	0.023%	NA	0.023%	NA
0.903282	0.067289	0	0.000016	0.025%	NA	0.025%	NA
0.912406	0.060859	0	0.000015	0.028%	NA	0.028%	NA
0.930931	0.047804	0	0.000015	0.033%	NA	0.033%	NA
0.940335	0.041178	0	0.000014	0.039%	NA	0.039%	NA
0.959427	0.027776	0.000051	0.000014	0.043%	1.961%	0.043%	1.961%
0.969119	0.021132	0.000235	0.000013	0.033%	1.277%	0.033%	1.277%
0.988796	0.009289	0.00225	0.000010	0.140%	0.578%	0.140%	0.578%
0.998783	0.005029	0.005029	0.000008	0.358%	0.358%	0.358%	0.358%
1.018859	0.000854	0.015003	0.000003	1.054%	0.067%	1.171%	0.067%
1.029048	0.000253	0.021582	0.000002	1.581%	0.019%	1.581%	0.019%
1.049731	0.000010	0.035917	0	0.990%	0.000%	0.990%	0.000%
1.060229	0	0.043306	0	NA	0.000%	NA	0.000%
1.081539	0	0.058324	0	NA	0.000%	NA	0.000%
1.092355	0	0.065946	0	NA	0.000%	NA	0.000%
1.114311	0	0.08142	0	NA	0.000%	NA	0.000%
1.125454	0	0.089273	0	NA	0.000%	NA	0.000%
1.148076	0	0.105216	0	NA	0.000%	NA	0.000%

Monte Carlo: 1,000,000 simulations with 100 time steps per option maturity.
The strike in bold is the forward price of the underlying coupon bond for the option's expiry date (i.e. the ATM strike).
"NA" stands for not available.
Percentage absolute pricing errors are equal to the approximate price -as given by propositions 4 or 5- over the Monte Carlo estimate, minus one.

a maturity of 5 years on a coupon bond with a maturity of 15 years, a unit face value and a coupon rate of 7% per year (with coupons paid semi-annually). It shows that, for both calls and puts on long-term coupon bonds, both approximations perform equally well for in-the-money contracts. However, for (near and) at-the-money options, the *lognormal* approximation produces lower pricing errors, while for deep out-of-the-money options, the *rank 1* approximation performs clearly better.

This pattern is confirmed in table 3.4, where European calls and puts with a maturity of just 1 year are priced on the same coupon bond as in table 3.3, but now with a life of 21 years.

In summary, both approximations produce similar results when the maturity of the underlying coupon-bearing bond is near the option's expiry date, and also for in-the-money options on long-term bonds. The *lognormal* approximation seems to perform better for

Table 3.3: Valuation of European options with a maturity of 5 years, on a unit face value coupon-bearing bond paying 20 semi-annually coupons of 0.035 each after the option's expiry date, using the Babbs and Nowman (1999, Table 2) model

	Standard Monte Carlo (exact option prices)			Percentage Absolute Pricing Errors			
				Rank 1 Appr.		Lognormal Appr.	
Strikes	Call	Put	Std. errors	Call	Put	Call	Put
0.852268	0.098920	0.001187	0.000067	0.098%	8.256%	0.163%	13.564%
0.869572	0.087593	0.002056	0.000066	0.114%	4.864%	0.198%	8.414%
0.878356	0.081998	0.002651	0.000065	0.123%	3.810%	0.216%	6.677%
0.896190	0.071033	0.004255	0.000062	0.148%	2.468%	0.253%	4.254%
0.905243	0.065703	0.005305	0.000061	0.163%	2.017%	0.272%	3.374%
0.923623	0.055453	0.008009	0.000058	0.204%	1.411%	0.303%	2.098%
0.932952	0.050575	0.009706	0.000056	0.229%	1.195%	0.314%	1.638%
0.951895	0.041413	0.013894	0.000052	0.297%	0.885%	0.319%	0.950%
0.961510	0.037166	0.016424	0.000050	0.334%	0.749%	0.307%	0.688%
0.981032	0.029412	0.022428	0.000045	0.415%	0.548%	0.235%	0.308%
0.990942	0.025925	0.025925	0.000042	0.463%	0.463%	0.170%	0.174%
1.010860	0.019808	0.033846	0.000037	0.576%	0.337%	0.010%	0.006%
1.020968	0.017143	0.038305	0.000035	0.647%	0.287%	0.134%	0.060%
1.041490	0.012579	0.048204	0.000030	0.795%	0.207%	0.477%	0.124%
1.051905	0.010661	0.053625	0.000028	0.882%	0.177%	0.694%	0.136%
1.073048	0.007491	0.065357	0.000023	1.095%	0.124%	1.241%	0.144%
1.083779	0.006212	0.071640	0.000021	1.159%	0.101%	1.642%	0.141%
1.105563	0.004175	0.084956	0.000017	1.293%	0.064%	2.587%	0.127%
1.116618	0.003383	0.091956	0.000015	1.360%	0.050%	3.133%	0.116%
1.139062	0.002168	0.106558	0.000012	1.568%	0.032%	4.382%	0.089%

Monte Carlo: 1,000,000 simulations with 100 time steps per option maturity.
The strike in bold is the forward price of the underlying coupon bond for the option's expiry date (i.e. the ATM strike).
Percentage absolute pricing errors are equal to the approximate price -as given by propositions 4 or 5- over the Monte Carlo estimate, minus one.

at-the-money options on coupon bonds with a long remaining life after the contract expiry date, while the *rank 1* approximation is better suited for pricing such options on long-term bonds at deep out-of-the-money strikes.

3.4.4 European swaptions

Under the Duffie and Singleton (1997) assumption of symmetric counterparty credit risk, the approximate Gaussian solutions previously derived for European options on coupon-bearing bonds can be generalized for the pricing of European swaptions.

Let the time-t price of an European payer swaption maturing at time $T_u \equiv t + u\delta$, with a strike equal to x, and on a forward swap with a unitary principal and settled in arrears at times $T_{u+i} \equiv T_u + i\delta, i = 1, \ldots, m$, be denoted by $Payerswpn_t(x, \delta, u, m)$. Since this option gives the right to enter, at time T_u, into the underlying swap paying the pre-established

Table 3.4: Valuation of European options with a maturity of 1 year, on a unit face value coupon-bearing bond paying 40 semi-annually coupons of 0.035 each after the option's expiry date, using the Babbs and Nowman (1999, Table 2) model

	Standard Monte Carlo (exact option prices)			Percentage Absolute Pricing Errors			
				Rank 1 Appr.		Lognormal Appr.	
Strikes	Call	Put	Std. errors	Call	Put	Call	Put
0.874994	0.104892	0.000322	0.000055	0.092%	30.124%	0.109%	35.404%
0.892760	0.088901	0.000895	0.000054	0.134%	13.296%	0.154%	15.419%
0.901777	0.080987	0.001389	0.000053	0.163%	9.503%	0.182%	10.583%
0.910886	0.073186	0.002082	0.000052	0.201%	7.061%	0.213%	7.445%
0.920087	0.065554	0.003028	0.000051	0.250%	5.416%	0.246%	5.317%
0.929381	0.058152	0.004291	0.000049	0.313%	4.241%	0.277%	3.752%
0.938769	0.051044	0.005936	0.000047	0.392%	3.369%	0.304%	2.628%
0.948251	0.044296	0.008029	0.000045	0.492%	2.715%	0.325%	1.806%
0.957829	0.037970	0.010634	0.000043	0.616%	2.200%	0.337%	1.204%
0.977277	0.026801	0.017597	0.000037	0.944%	1.438%	0.287%	0.443%
0.987149	0.022034	0.022035	0.000034	1.153%	1.153%	0.213%	0.213%
1.006990	0.014277	0.032778	0.000028	1.660%	0.720%	0.105%	0.046%
1.017060	0.011215	0.039104	0.000025	1.979%	0.568%	0.348%	0.100%
1.037503	0.006573	0.053523	0.000019	2.723%	0.334%	1.126%	0.138%
1.047878	0.004902	0.061525	0.000016	3.101%	0.249%	1.714%	0.137%
1.068941	0.002579	0.078840	0.000011	3.994%	0.131%	3.296%	0.108%
1.079630	0.001818	0.088046	0.000010	4.510%	0.093%	4.345%	0.090%
1.101331	0.000854	0.107315	0.000006	5.386%	0.043%	7.026%	0.056%
1.112344	0.000570	0.117299	0.000005	5.614%	0.027%	8.947%	0.043%
1.134702	0.000240	0.137816	0.000003	5.833%	0.009%	13.750%	0.024%

Monte Carlo: 1,000,000 simulations with 100 time steps per option maturity.
The strike in bold is the forward price of the underlying coupon bond for the option's expiry date (i.e. the ATM strike).
Percentage absolute pricing errors are equal to the approximate price -as given by propositions 4 or 5- over the Monte Carlo estimate, minus one.

fixed rate x against the receivement of a floating rate, its terminal payoff corresponds to[25]

$$Payerswpn_{T_u}(x,\delta,u,m) = \left[w(T_u,T_{u+m}) - x\right]^+ \delta \sum_{i=1}^{m} P(T_u,T_{u+i}),$$

where $w(T_u,T_{u+m})$ is the time-T_u spot swap rate for a $m\delta$-year swap with a compounding period of δ years. Because, by definition, a spot swap rate equals the fixed rate of the interest rate swap for which the value of the contract is zero, then $w(T_u,T_{u+m}) \equiv \frac{1-P(T_u,T_{u+m})}{\delta\sum_{i=1}^{m} P(T_u,T_{u+i})}$, and therefore (as shown, for instance, by Brace and Musiela (1994a, page 266)):

$$Payerswpn_{T_u}(x,\delta,u,m) = \left\{1 - \left[P(T_u,T_{u+m}) + x\delta\sum_{i=1}^{m} P(T_u,T_{u+i})\right]\right\}^+.$$

[25] In the London market, the discounting is made using not spot yields but rather the spot swap rate. That is, the European payer swaptions' terminal payoff is equal to $[w(T_u,T_{u+m}) - x]^+ \delta\sum_{i=1}^{m}[1 + w(T_u,T_{u+m})\delta]^{-i}$. Nevertheless, Pang (1996, Table 6) has shown that, in this case, a good approximation can be obtained by replacing the spot swap rate by spot rates.

That is, an European payer swaption can be valued as an European put with the same expiry date, with a strike price equal to 1, and on a coupon-bearing bond corresponding to the fixed leg of the underlying interest rate swap. Hence,

$$Payerswpn_t(x, \delta, u, m) = p_t[S(t); 1; T_u], \tag{3.60}$$

where the put price is computed from equations (3.42) or (3.56), but with $N_0 = m$ and $k_i = 1_{\{i=m\}} + x\delta$.

Similarly, it can also be shown that the time-t price of an European receiver swaption maturing at time $T_u \equiv t + u\delta$, with a strike equal to x, and on a forward swap with a unitary principal and settled in arrears at times $T_{u+i} \equiv T_u + i\delta, i = 1, \ldots, m$, can be computed as

$$Receiverswpn_t(x, \delta, u, m) = c_t[S(t); 1; T_u], \tag{3.61}$$

where the call price is given by (3.39) or (3.52), but with $N_0 = m$ and $k_i = 1_{\{i=m\}} + x\delta$.

3.5 Pricing of interest rate futures

3.5.1 Futures on default-free bonds

The hypothesis of continuous marking to market will be assumed hereafter and whenever futures contracts are involved.

Starting with the valuation of futures on riskless pure discount bonds (e.g. futures on Treasury Bills), let the time-t price of a futures contract for delivery at time T_f and on a default-free zero-coupon bond with maturity at time T_1 (such that $t \leq T_f \leq T_1$) be denoted by $FP(t, T_f, T_1)$. It is well known -see for instance Cox, Ingersoll and Ross (1981b, equation 46)- that a futures price is just the expectation of the spot price on the delivery date, under the martingale measure Q, and therefore:

$$FP(t, T_f, T_1) = E_Q[P(T_f, T_1)| \mathcal{F}_t]. \tag{3.62}$$

Or using the exponential-affine formula (2.1),

$$FP(t, T_f, T_1) = \exp[A(T_1 - T_f)] E_Q\left\{\exp\left[\underline{B}'(T_1 - T_f) \cdot \underline{X}(T_f)\right] | \mathcal{F}_t\right\}.$$

Moreover, the last expectation is just the moment generating function of the random variable $[\underline{B}'(T_1 - T_f) \cdot \underline{X}(T_f)]$, with a coefficient of $+1$.

On the other hand, since matrix a is time-homogeneous and assuming that matrix a is also nonsingular, Arnold (1992, corollary 8.2.4) provides the following *strong* solution for equation (3.1), with $t \geq t_0$:

$$\underline{X}(t) = e^{a(t-t_0)} \cdot \underline{X}(t_0) + \left[e^{a(t-t_0)} - I_n\right] \cdot a^{-1} \cdot \underline{b} + \int_{t_0}^{t} e^{a(t-v)} \cdot S \cdot d\underline{W}^{Q}(v).$$

Consequently,[26] $\underline{X}(t) \backsim N^n(\underline{u}(t-t_0), \Delta(t-t_0))$, where

$$\underline{u}(t-t_0) = e^{a(t-t_0)} \cdot \underline{X}(t_0) + \left[e^{a(t-t_0)} - I_n\right] \cdot a^{-1} \cdot \underline{b}, \tag{3.63}$$

and thus

$$\begin{aligned} FP(t, T_f, T_1) &= \exp\left[A(T_1 - T_f) + \underline{B}'(T_1 - T_f) \cdot \underline{u}(T_f - t)\right. \\ &\left. + \frac{1}{2}\underline{B}'(T_1 - T_f) \cdot \Delta(T_f - t) \cdot \underline{B}(T_1 - T_f)\right]. \end{aligned} \tag{3.64}$$

Proposition 6 expresses the futures price (3.64) in terms of the corresponding forward price.

Proposition 6 *Under the deterministic volatility specification of the Duffie and Kan (1996) model, the time-t price, $FP(t, T_f, T_1)$, of a futures contract for delivery at time T_f and on a pure discount bond with maturity at time T_1 ($t \leq T_f \leq T_1$) is equal to*

$$FP(t, T_f, T_1) = \frac{P(t, T_1)}{P(t, T_f)} \exp[-J(t)], \tag{3.65}$$

where

$$\begin{aligned} J(t) &= \underline{G}' \cdot a^{-1} \cdot \left\{\Theta \cdot (a^{-1})' \cdot [\underline{B}(T_1 - t) - \underline{B}(T_1 - T_f) - \underline{B}(T_f - t)]\right. \\ &\left. - \Delta(T_f - t) \cdot \underline{B}(T_1 - T_f)\right\}. \end{aligned}$$

Proof. Combining equations (3.63) and (3.64), as well as considering equations (3.3), (3.4) and (3.6),

$$\begin{aligned} FP(t, T_f, T_1) &= \exp\left\{(T_1 - T_f)(\underline{G}' \cdot a^{-1} \cdot \underline{b} - f) + \underline{B}'(T_1 - T_f) \cdot a^{-1} \cdot \underline{b}\right. \\ &+ \frac{1}{2}\int_{T_f}^{T_1} \underline{B}'(T_1 - v) \cdot \Theta \cdot \underline{B}(T_1 - v)\, dv \\ &+ \underline{G}' \cdot a^{-1} \cdot \left[I_n - e^{a(T_1 - T_f)}\right] \cdot e^{a(T_f - t)} \cdot \underline{X}(t) \\ &\left. + \underline{B}'(T_1 - T_f) \cdot \left[e^{a(T_f - t)} - I_n\right] \cdot a^{-1} \cdot \underline{b} + \frac{1}{2}\int_t^{T_f} \varphi(v)\, dv\right\}, \end{aligned}$$

[26] The notation $\underline{Y} \backsim N^d(\underline{\mu}, C)$ will be used to state that the random variable $\underline{Y} \in \Re^d$ has a d-dimensional normal distribution, with mean $\underline{\mu} \in \Re^d$, and covariance matrix $C \in \Re^{d \times d}$.

where

$$\begin{aligned}\varphi(v) &= \underline{G}' \cdot a^{-1} \cdot \left[e^{a(T_f-v)} - e^{a(T_1-v)}\right] \cdot \Theta \cdot \left[e^{a'(T_f-v)} - e^{a'(T_1-v)}\right] \cdot \left(a^{-1}\right)' \cdot \underline{G} \\ &= \left[\underline{B}'(T_1-v) - \underline{B}'(T_f-v)\right] \cdot \Theta \cdot \left[\underline{B}(T_1-v) - \underline{B}(T_f-v)\right].\end{aligned}$$

Hence,

$$\begin{aligned}&FP(t,T_f,T_1) \\ =\; &\exp\Big\{(T_1-T_f)\left(\underline{G}' \cdot a^{-1} \cdot \underline{b} - f\right) \\ &+\underline{G}' \cdot a^{-1} \cdot \left[e^{a(T_f-t)} - e^{a(T_1-t)}\right] \cdot \left[\underline{X}(t) + a^{-1} \cdot \underline{b}\right] \\ &+\frac{1}{2}\int_{T_f}^{T_1} \underline{B}'(T_1-v) \cdot \Theta \cdot B(T_1-v)\,dv + \frac{1}{2}\int_t^{T_f} \underline{B}'(T_1-v) \cdot \Theta \cdot B(T_1-v)\,dv \\ &+\frac{1}{2}\int_t^{T_f} \underline{B}'(T_f-v) \cdot \Theta \cdot B(T_f-v)\,dv - \int_t^{T_f} \underline{B}'(T_f-v) \cdot \Theta \cdot B(T_1-v)\,dv\Big\},\end{aligned}$$

i.e.

$$\begin{aligned}&FP(t,T_f,T_1) \\ =\; &\exp\Big\{\left[(T_1-t) - (T_f-t)\right]\left(\underline{G}' \cdot a^{-1} \cdot \underline{b} - f\right) \\ &+\left[\underline{B}'(T_1-t) - \underline{B}'(T_f-t)\right] \cdot \left[\underline{X}(t) + a^{-1} \cdot \underline{b}\right] \\ &+\frac{1}{2}\int_t^{T_1} \underline{B}'(T_1-v) \cdot \Theta \cdot B(T_1-v)\,dv - \frac{1}{2}\int_t^{T_f} \underline{B}'(T_f-v) \cdot \Theta \cdot B(T_f-v)\,dv \\ &+\int_t^{T_f} \underline{B}'(T_f-v) \cdot \Theta \cdot B(T_f-v)\,dv - \int_t^{T_f} \underline{B}'(T_f-v) \cdot \Theta \cdot B(T_1-v)\,dv\Big\}.\end{aligned}$$

Using again equations (3.3) and (3.6), then

$$FP(t,T_f,T_1) = \frac{P(t,T_1)}{P(t,T_f)} \exp\left[-J(t)\right], \tag{3.66}$$

where

$$J(t) = \int_t^{T_f} \underline{B}'(T_f-v) \cdot \Theta \cdot \left[\underline{B}(T_1-v) - \underline{B}(T_f-v)\right] dv. \tag{3.67}$$

In order to find an explicit solution for $J(t)$, equation (3.2) yields

$$\begin{aligned}J(t) &= \underline{G}' \cdot a^{-1} \cdot \left\{\int_t^{T_f} \left[I - e^{a(T_f-v)}\right] \cdot \Theta \cdot \left[e^{a'(T_f-v)} - e^{a'(T_1-v)}\right] dv\right\} \cdot \left(a^{-1}\right)' \cdot \underline{G} \\ &= \underline{G}' \cdot a^{-1} \cdot \Theta \cdot \left\{\int_t^{T_f} \left[e^{a'(T_f-v)} - e^{a'(T_1-v)}\right] dv\right\} \cdot \left(a^{-1}\right)' \cdot \underline{G} \\ &\quad -\underline{G}' \cdot a^{-1} \cdot \left[\int_t^{T_f} e^{a(T_f-v)} \cdot \Theta \cdot e^{a'(T_f-v)} dv\right] \cdot \left(a^{-1}\right)' \cdot \underline{G}\end{aligned}$$

$$+\underline{G}' \cdot a^{-1} \cdot \left[\int_t^{T_f} e^{a(T_f - v)} \cdot \Theta \cdot e^{a'(T_1 - v)} dv \right] \cdot (a^{-1})' \cdot \underline{G}.$$

Solving the first integral in the right-hand-side of the last equation, and expressing the other two as functions of $\Delta(T_f - t)$,

$$\begin{aligned} J(t) &= -\underline{G}' \cdot a^{-1} \cdot \Theta \cdot (a')^{-1} \cdot \left\{ I - e^{a'(T_1 - T_f)} - e^{a'(T_f - t)} + e^{a'(T_1 - t)} \right\} \cdot (a^{-1})' \cdot \underline{G} \\ &\quad -\underline{G}' \cdot a^{-1} \cdot \Delta(T_f - t) \cdot (a^{-1})' \cdot \underline{G} \\ &\quad +\underline{G}' \cdot a^{-1} \cdot \Delta(T_f - t) \cdot e^{a'(T_1 - T_f)} \cdot (a^{-1})' \cdot \underline{G}. \end{aligned}$$

Finally, applying, once more, equation (3.2), expression (3.65) follows. ■

Remark 5 *The pricing formula (3.65) is similar to (and nests), for instance, the one derived by Chen (1995, equation 14) using a two-factor Ornstein-Uhlenbeck process, being the only difference the specification of $J(t)$. In both cases, and as shown by El Karoui et al. (1991, equation 49) or Jamshidian (1993, equation 1.6),*[27]

$$J(t) = \langle \ln P(t, T_f), \ln P(t, T_1) \rangle_{T_f} - \langle \ln P(t, T_f), \ln P(t, T_f) \rangle_{T_f},$$

which is confirmed by identity (3.67).

Remark 6 *The first term in the right-hand-side of (3.65) is just the time-t forward price for time T_f of a pure discount bond with expiry date T_1, that is $P(t, T_f, T_1)$. In fact, since this forward contract will only be settled at time T_f, it is equivalent to a portfolio composed by a long European call and a short European put, both on the pure discount bond $P(t, T_1)$, with expiry date T_f, and with strike prices equal to the forward price $P(t, T_f, T_1)$. Moreover, the value of such portfolio must be zero, because the initial investment in a forward contract is also null. Therefore,*

$$c_t [P(t, T_1); P(t, T_f, T_1); T_f] - p_t [P(t, T_1); P(t, T_f, T_1); T_f] = 0,$$

and using the put-call parity for European options on zero coupon bonds,

$$P(t, T_1) - P(t, T_f, T_1) P(t, T_f) = 0,$$

it follows that $P(t, T_f, T_1) = \frac{P(t, T_1)}{P(t, T_f)}$, as expected.

Concerning the valuation of coupon-bearing bond futures, the time-t price, $FS(t, T_f)$,

[27] *The process $\langle X, X \rangle_T$ denotes the quadratic variation of the process $(X_s)_{t \le s \le T}$.*

of a futures contract, for delivery at date T_f ($\geq t$), on the default-free coupon bond $S(t)$ is the expectation under measure Q of the underlying bond value at the future time T_f:

$$\begin{aligned} FS(t,T_f) &= E_Q\left[S(T_f)|\mathcal{F}_t\right] \\ &= \sum_{i=1}^{N_f} k_i E_Q\left[P(T_f,T_i)|\mathcal{F}_t\right], \end{aligned} \tag{3.68}$$

where $T_f < T_i$ (for $i = 1,\ldots,N_f$), and N_f represents the number of cash flows k_i ($i = 1,\ldots,N_f$) paid by the underlying coupon bond from the futures' expiry date and until the maturity date of the underlying bond. In practice, however, the underlying of all Treasury bond futures is not a traded but rather a theoretical bond,[28] and the party with the short position has the (*quality*) option of choosing, amongst all the deliverable bonds,[29] the one (*cheapest-to-deliver*) to be delivered on the delivery day. Moreover, for some futures contracts, the seller can also choose the delivery day during the delivery month, that is the party with the short position possess also a *timing* option.[30] Valuation formula (3.68) ignores the existence of both *delivery* options (i.e. *quality* and *timing* options), and takes as the underlying of the futures contract the *cheapest-to-deliver* bond, which is assumed to be known.[31] On the delivery day, the invoice price paid by the party with the long position corresponds to the futures settlement price times a conversion factor[32] plus the accrued interest on the delivered bond. Assuming that the futures settlement price times the conversion factor of the *cheapest-to-deliver* bond converges to the quoted price of such bond at the futures' expiry date, equation (3.68) follows.

Combining equations (3.68) and (3.62), we can conclude that the relation between futures on coupon bonds and futures on pure discount bonds is the same as the well known relation between the corresponding underlying instruments in the spot market: the price of a futures contract on a coupon-bearing bond is equal to the summation of the prices of futures on zero-coupon bonds with delivery dates corresponding to the moments where cash

[28] For instance, in the case of the US T-Bond Futures, traded at the Chicago Board of Trade, the theoretical underlying bond is an issue with a face value of US\$100,000 and a coupon rate of 8% (6%, after and including the March 2000 contracts).

[29] For the T-Bond Futures contracts, deliverable bonds are all Treasury bonds that, on the first day of the delivery month, possess a maturity of at least 15 years and are not callable for at least 15 years.

[30] Again, for the T-Bond futures contracts, delivery can occur at any business day during the delivery month. However, in the case, for instance, of the Bund Futures (traded at LIFFE), delivery must take place on the tenth calendar day of the delivery month, and thus there exists no *timing* option.

[31] In practice and before the maturity of the futures contract, the *cheapest-to-deliver* can be predicted as the deliverable bond with the highest *implied-repo-rate*. On delivery, it is the one that maximizes the difference between the futures' *invoice amount* and the gross price of the bond.

[32] The conversion factor of each deliverable bond is the cash price, per unit of face value, of such bond, at the futures' expiry date, that would produce an yield-to-maturity equal to the coupon rate of the underlying theoretical bond. Hence, each conversion factor corrects the futures price in order to adjust the invoice amount received by the seller as a function of the deliverable bond used for delivery.

flows are paid by the coupon bond, and with contract sizes equal to the value of such cash flows, that is

$$FS(t,T_f) = \sum_{i=1}^{N_f} k_i FP(t,T_f,T_i). \tag{3.69}$$

Next proposition simply generalizes proposition 6.

Proposition 7 *Under the Gaussian specification of the Duffie and Kan (1996) model and ignoring all delivery options, the time-t price, $FS(t,T_f)$, of a futures contract, for delivery at date T_f ($\geq t$), on a default-free coupon-bearing bond generating N_f cash flows of value k_i at times T_i (such that $T_i > T_f, i = 1,\ldots,N_f$), is equal to*

$$FS(t,T_f) = \sum_{i=1}^{N_f} k_i \frac{P(t,T_i)}{P(t,T_f)} \exp\left[-J_i(t)\right], \tag{3.70}$$

where

$$\begin{aligned} J_i(t) = \quad & \underline{G}' \cdot a^{-1} \cdot \left\{ \Theta \cdot \left(a^{-1}\right)' \cdot \left[\underline{B}(T_i - t) - \underline{B}(T_i - T_f) - \underline{B}(T_f - t)\right] \right. \\ & \left. -\Delta(T_f - t) \cdot \underline{B}(T_i - T_f) \right\}. \end{aligned}$$

Proof. The pricing formula (3.70) is obtained by combining equations (3.69) and (3.65), with 1 replaced by i. ■

3.5.2 Short-term interest rate futures

This subsection considers the valuation of futures on short-term nominal money-market forward interest rates. This is the case, for instance, of the widely traded Eurodollar futures contract, where the underlying nominal interest rate is the LIBOR of the USD for a three months period. In what follows, all interest rates and all bond prices are assumed to be risk-adjusted along the lines of Duffie and Singleton (1997).

Let $FR(t,T_f,T_1)$ denote the time-t price of a futures contract with maturity at time T_f ($\geq t$) and on the nominal interest rate for the period $(T_1 - T_f)$, with $T_1 \geq T_f$. By convention, the futures price is quoted on an annualized basis, and therefore the terminal futures price corresponds to

$$FR(T_f,T_f,T_1) = 100\left[1 - R(T_f,T_1)\right],$$

where $R(T_f,T_1) = \frac{1}{T_1 - T_f}\left[\frac{1}{P(T_f,T_1)} - 1\right]$ is the time-T_f nominal spot interest rate for the period $(T_1 - T_f)$.

Using again Cox et al. (1981b, equation 46) as well as the exponential-affine formula (2.1),

$$FR(t, T_f, T_1) = 100 + \frac{100}{T_1 - T_f} - \frac{100}{T_1 - T_f} \exp\left[-A(T_1 - T_f)\right] \\ E_Q\left\{\exp\left[-\underline{B}'(T_1 - T_f) \cdot \underline{X}(T_f)\right] \mid \mathcal{F}_t\right\}. \tag{3.71}$$

The expectation appearing in the right-hand-side of the last equation is the moment generating function of the random variable $[\underline{B}'(T_1 - T_f) \cdot \underline{X}(T_f)]$, with a coefficient of -1. And because $\underline{X}(T_f) \backsim N^n(\underline{u}(T_f - t), \Delta(T_f - t))$, then

$$E_Q\left\{\exp\left[-\underline{B}'(T_1 - T_f) \cdot \underline{X}(T_f)\right] \mid \mathcal{F}_t\right\} \\ = \exp\left[-\underline{B}'(T_1 - T_f) \cdot \underline{u}(T_f - t) + \frac{1}{2}\underline{B}'(T_1 - T_f) \cdot \Delta(T_f - t) \cdot \underline{B}(T_1 - T_f)\right]. \tag{3.72}$$

Combining equations (3.2), (3.6), (3.63), (3.71), and (3.72), proposition 8 follows.

Proposition 8 *Under the deterministic volatility specification of the Duffie and Kan (1996) model, the time-t price, $FR(t, T_f, T_1)$, of a futures contract with maturity at time T_f and on the nominal interest rate for the period $(T_1 - T_f)$, with $T_1 \geq T_f \geq t$, is equal to*

$$FR(t, T_f, T_1) = 100\left\{1 - \frac{1}{T_1 - T_f}\left[\frac{P(t, T_f)}{P(t, T_1)} e^{L(t)} - 1\right]\right\}, \tag{3.73}$$

where:

$$L(t) = \underline{G}' \cdot a^{-1} \cdot \left\{\Theta \cdot (a^{-1})' \cdot \left[\underline{B}(T_1 - t) - \underline{B}(T_1 - T_f) - \underline{B}(T_f - t)\right] \right. \\ \left. - e^{a(T_1 - T_f)} \cdot \Delta(T_f - t) \cdot \underline{B}(T_1 - T_f)\right\}.$$

Proof. The derivation of equation (3.73) is similar to the proof presented before for proposition 6. Using equations (3.72), (3.63), (3.2), and (3.6), it can be shown that

$$\exp\left[-A(T_1 - T_f)\right] E_Q\left\{\exp\left[-\underline{B}'(T_1 - T_f) \cdot \underline{X}(T_f)\right] \mid \mathcal{F}_t\right\} \\ = \exp\left\{\left[(T_f - t) - (T_1 - t)\right]\left(\underline{G}' \cdot a^{-1} \cdot \underline{b} - f\right) \right. \\ - \left[\underline{B}'(T_1 - t) - \underline{B}'(T_f - t)\right] \cdot \left[\underline{X}(t) + a^{-1} \cdot \underline{b}\right] \\ - \frac{1}{2}\int_t^{T_1} \underline{B}'(T_1 - v) \cdot \Theta \cdot B(T_1 - v)\, dv + \frac{1}{2}\int_t^{T_f} \underline{B}'(T_f - v) \cdot \Theta \cdot B(T_f - v)\, dv \\ \left. + \int_t^{T_f} \underline{B}'(T_1 - v) \cdot \Theta \cdot B(T_1 - v)\, dv - \int_t^{T_f} \underline{B}'(T_1 - v) \cdot \Theta \cdot B(T_f - v)\, dv\right\}.$$

Combining this result with equation (3.71), and using again equation (3.6), the pricing solution (3.73) is obtained but with

$$L(t) = \int_t^{T_f} \underline{B}'(T_1 - v) \cdot \Theta \cdot [\underline{B}(T_1 - v) - \underline{B}(T_f - v)] \, dv.$$

Finally, using equation (3.2), all the resulting integrals can be either computed analytically or expressed as functions of matrix $\Delta(T_f - t)$. ■

3.6 Pricing of European interest rate futures options

3.6.1 Futures options on pure discount bonds

This subsection only considers options with stock-style margining, also known as *conventional* futures options (using the terminology of Duffie (1989)): that is contracts with premium paid at the beginning of the options' life.

Applying result (3.8) to the time-t price, $c_t[FP(t,T_f,T_1);K_f;T_0]$, of an European call on the asset $FP(t,T_f,T_1)$, with a strike price of K_f, and expiry date at time T_0, and representing by $\varkappa = \{\omega \in \Omega : FP(T_0,T_f,T_1)(\omega) > K_f\}$ the set of states of the world in which the option ends in-the-money, then

$$\begin{aligned} c_t[FP(t,T_f,T_1);K_f;T_0] &= P(t,T_0) E_{Q_0}[FP(T_0,T_f,T_1) 1_{\varkappa} | \mathcal{F}_t] \\ &\quad - P(t,T_0) K_f E_{Q_0}(1_{\varkappa} | \mathcal{F}_t). \end{aligned}$$

Furthermore, using the law of iterative expectations, and because $E_{Q_0}(1_{\varkappa} | \mathcal{F}_t)$ corresponds to the Q_0-probability that the option will be exercised at time T_0,

$$\begin{aligned} c_t[FP(t,T_f,T_1);K_f;T_0] &= P(t,T_0) E_{Q_0}[FP(T_0,T_f,T_1) | \mathcal{F}_t] E_{Q_0}(\eta 1_{\varkappa} | \mathcal{F}_t) \\ &\quad - P(t,T_0) K_f \Pr_{Q_0}[FP(T_0,T_f,T_1) > K_f | \mathcal{F}_t], \end{aligned}$$

where

$$\eta = \frac{FP(T_0,T_f,T_1)}{E_{Q_0}[FP(T_0,T_f,T_1) | \mathcal{F}_t]} \tag{3.74}$$

is a Radon-Nikodym derivative such that $E_{Q_0}(\eta | \mathcal{F}_t) = 1$.

In order to identify the change of numeraire associated with the above Radon-Nikodym derivative, it is useful to notice that η can also be written as

$$\eta = \frac{h(T_0,T_f,T_1)}{h(t,T_f,T_1)} \frac{P(t,T_0)}{P(T_0,T_0)},$$

where

$$h(t,T_f,T_1) \equiv P(t,T_0)\, E_{Q_0}\left[FP(T_0,T_f,T_1)|\,\mathcal{F}_t\right] \tag{3.75}$$

represents the time-t Q_0-expected discounted futures price. Comparing this result with equations (3.21) or (3.36), it is obvious that the Radon-Nikodym derivative η can be though of as representing the change from the numeraire $P(t,T_0)$ to a new numeraire $h(t,T_f,T_1)$: as Chen (1992, page 100) points out, $h(t,T_f,T_1)$ is the time-t price of an asset that pays the futures price $FP(T_0,T_f,T_1)$ at time T_0. Hence, it can be said that $\eta \equiv \left(\frac{dQ_h}{dQ_0}\right)_{T_0}$, where Q_h is the equivalent martingale measure associated with the numeraire $h(t,T_f,T_1)$, and therefore

$$\begin{aligned} E_{Q_0}\left(\eta 1_x|\,\mathcal{F}_t\right) &= E_{Q_h}\left(1_x|\,\mathcal{F}_t\right) \\ &= \Pr_{Q_h}\left[FP(T_0,T_f,T_1) > K_f|\,\mathcal{F}_t\right], \end{aligned}$$

which corresponds to the Q_h-probability that the option will be exercised at time T_0.

Finally, all the above results can be combined in the following probabilistic valuation formula:

$$\begin{aligned} c_t\left[FP(t,T_f,T_1)\,;K_f;T_0\right] &= h(t,T_f,T_1)\Pr_{Q_h}\left[FP(T_0,T_f,T_1) > K_f|\,\mathcal{F}_t\right] \\ &\quad - P(t,T_0)\,K_f\Pr_{Q_0}\left[FP(T_0,T_f,T_1) > K_f|\,\mathcal{F}_t\right]. \end{aligned} \tag{3.76}$$

Proposition 9 offers explicit solutions for the price $h(t,T_f,T_1)$ as well as for both probabilities contained in the right-hand-side of equation (3.76). It follows the structure of Chen (1992, equation 9), Chen (1995, equation 15) or Schöbel (1990, equation 4.14), which nests as special cases.

Proposition 9 *Under the deterministic volatility specification of the Duffie and Kan (1996) model, the time-t premium of an European conventional call on the asset $FP(t,T_f,T_1)$, with a strike price of K_f, and expiry date at time T_0 (such that $t \leq T_0 \leq T_f \leq T_1$), is equal to*

$$c_t\left[FP(t,T_f,T_1)\,;K_f;T_0\right] = h(t,T_f,T_1)\,\Phi\left[d_1^f(t)\right] - P(t,T_0)\,K_f\Phi\left[d_0^f(t)\right], \tag{3.77}$$

where

$$d_1^f(t) = \frac{\ln\left[\frac{h(t,T_f,T_1)}{P(t,T_0)K_f}\right] + \frac{\sigma_h^2(t)}{2}}{\sigma_h(t)},$$

$$d_0^f(t) = d_1^f(t) - \sigma_h(t),$$

$$h(t,T_f,T_1) = P(t,T_0)\,FP(t,T_f,T_1)\exp\left[I(t)\right], \tag{3.78}$$

$$
\begin{aligned}
I(t) = \quad & \underline{G}' \cdot a^{-1} \cdot \Theta \cdot (a')^{-1} \cdot [\underline{B}(T_1 - t) + \underline{B}(T_f - T_0) - \underline{B}(T_1 - T_0) - \underline{B}(T_f - t)] \\
& + \underline{G}' \cdot a^{-1} \cdot \Delta(T_0 - t) \cdot [\underline{B}(T_f - T_0) - \underline{B}(T_1 - T_0)],
\end{aligned}
$$

and

$$
\sigma_h^2(t) = [\underline{B}'(T_1 - T_0) - \underline{B}'(T_f - T_0)] \cdot \Delta(T_0 - t) \cdot [\underline{B}(T_1 - T_0) - \underline{B}(T_f - T_0)]. \quad (3.79)
$$

The corresponding put price is

$$
p_t[FP(t, T_f, T_1); K_f; T_0] = -h(t, T_f, T_1)\,\Phi\left[-d_1^f(t)\right] + P(t, T_0)\,K_f \Phi\left[-d_0^f(t)\right]. \quad (3.80)
$$

Proof. The derivation of an explicit formula for $h(t, T_f, T_1)$ involves three steps: first, $FP(T_0, T_f, T_1)$ will be stated as a function of $FP(t, T_f, T_1)$; second, its time-t Q_0-expected value will be found; and finally, identity (3.75) will be used. From equation (3.66),

$$
FP(T_0, T_f, T_1) = \frac{P(T_0, T_1)}{P(T_0, T_f)} \exp\left\{\int_{T_0}^{T_f} \underline{B}'(T_f - v) \cdot \Theta \cdot [\underline{B}(T_f - v) - \underline{B}(T_1 - v)]\, dv\right\}.
$$

On the other hand, because $\frac{P(T_0, T_1)}{P(T_0, T_f)} = \frac{P(T_0, T_0, T_1)}{P(T_0, T_0, T_f)}$ and using results (3.43)-(3.44),

$$
\begin{aligned}
& \frac{P(T_0, T_1)}{P(T_0, T_f)} \\
= \; & \frac{P(t, T_0, T_1) \exp\left[-\frac{1}{2}\int_t^{T_0} \underline{H_1}'(v) \cdot \underline{H_1}(v)\, dv + \int_t^{T_0} \underline{H_1}'(v) \cdot d\underline{W}^{Q_0}(v)\right]}{P(t, T_0, T_f) \exp\left[-\frac{1}{2}\int_t^{T_0} \underline{H_f}'(v) \cdot \underline{H_f}(v)\, dv + \int_t^{T_0} \underline{H_f}'(v) \cdot d\underline{W}^{Q_0}(v)\right]}, \quad (3.81)
\end{aligned}
$$

with $\underline{H_f}'(v) = [\underline{B}'(T_f - v) - \underline{B}'(T_0 - v)] \cdot S$. Combining the last two formulae with expression (3.17), then

$$
\begin{aligned}
FP(T_0, T_f, T_1) \quad = \quad & \frac{P(t, T_1)}{P(t, T_f)} \exp\left\{\int_t^{T_0} [\underline{B}'(T_1 - v) - \underline{B}'(T_f - v)] \cdot S \cdot d\underline{W}^{Q_0}(v)\right\} \\
& \exp\left\{\int_{T_0}^{T_f} \underline{B}'(T_f - v) \cdot \Theta \cdot [\underline{B}(T_f - v) - \underline{B}(T_1 - v)]\, dv\right\} \\
& \exp\left\{-\frac{1}{2}\int_t^{T_0} \left[\underline{H_1}'(v) \cdot \underline{H_1}(v) - \underline{H_f}'(v) \cdot \underline{H_f}(v)\right] dv\right\}.
\end{aligned}
$$

For $FP(T_0, T_f, T_1)$ to be expressed as a function of $FP(t, T_f, T_1)$, it is just necessary to take the first term and to adjust the third term in the right-hand-side of the above equality, according to equation (3.66):

$$
FP(T_0, T_f, T_1) \quad = \quad FP(t, T_f, T_1) \exp\left\{\int_t^{T_0} [\underline{B}'(T_1 - v) - \underline{B}'(T_f - v)] \cdot S \cdot d\underline{W}^{Q_0}(v)\right\}
$$

$$\exp\left\{\int_t^{T_0} \underline{B}'(T_f - v)\cdot\Theta\cdot[\underline{B}(T_1 - v) - \underline{B}(T_f - v)]\,dv\right\}$$
$$\exp\left\{-\frac{1}{2}\int_t^{T_0}\left[\underline{H_1}'(v)\cdot\underline{H_1}(v) - \underline{H_f}'(v)\cdot\underline{H_f}(v)\right]dv\right\}.$$

Concerning the second step, it is worth noticing that

$$\left\{\int_t^{T_0}\left[\underline{B}'(T_1 - v) - \underline{B}'(T_f - v)\right]\cdot S\cdot d\underline{W}^{Q_0}(v)\right\} \backsim N^1\left(0, \sigma_h^2(t)\right),$$

with[33]

$$\begin{aligned}
\sigma_h^2(t) &= \int_t^{T_0}\left[\underline{B}'(T_1 - v) - \underline{B}'(T_f - v)\right]\cdot\Theta\cdot[\underline{B}(T_1 - v) - \underline{B}(T_f - v)]\,dv \\
&= \int_t^{T_0}\left[\underline{H_1}'(v) - \underline{H_f}'(v)\right]\cdot\left[\underline{H_1}(v) - \underline{H_f}(v)\right]dv.
\end{aligned} \tag{3.82}$$

Therefore, the expectation of the second term in the right-hand-side of the last expression obtained for $FP(T_0, T_f, T_1)$ corresponds to a moment generating function (with a coefficient of 1), and thus

$$\begin{aligned}
&E_{Q_0}\left[FP(T_0, T_f, T_1)|\mathcal{F}_t\right] \\
&= FP(t, T_f, T_1)\exp\left\{\int_t^{T_0}\underline{H_f}'(v)\cdot\left[\underline{H_f}(v) - \underline{H_1}(v)\right]dv\right\} \\
&\quad \exp\left\{\int_t^{T_0}\underline{B}'(T_f - v)\cdot\Theta\cdot[\underline{B}(T_1 - v) - \underline{B}(T_f - v)]\,dv\right\}.
\end{aligned}$$

Finally, using identity (3.75), and combining the last two terms in the right-hand-side of the last expression, a general solution is derived for $h(t, T_f, T_1)$:

$$\begin{aligned}
h(t, T_f, T_1) = \; & P(t, T_0)\,FP(t, T_f, T_1) \\
& \exp\left\{\int_t^{T_0}\underline{B}'(T_0 - v)\cdot\Theta\cdot[\underline{B}(T_1 - v) - \underline{B}(T_f - v)]\,dv\right\}.
\end{aligned} \tag{3.83}$$

In appendix 3.8.4 it is shown how to convert equation (3.83) into the explicit solution (3.78).

In order to obtain an explicit formula for the Q_0-exercise probability, it will be useful to express $FP(T_0, T_f, T_1)$ in terms of $h(t, T_f, T_1)$. Multiplying and dividing $FP(T_0, T_f, T_1)$ by the factor $\exp\left[\frac{\sigma_h^2(t)}{2}\right]$, and following the same steps used in deriving formula (3.83),

$$\begin{aligned}
FP(T_0, T_f, T_1) &= FP(t, T_f, T_1)\exp\left\{\int_t^{T_0}\left[\underline{B}'(T_1 - v) - \underline{B}'(T_f - v)\right]\cdot S\cdot d\underline{W}^{Q_0}(v)\right\} \\
&\quad \exp\left\{\int_t^{T_0}\underline{B}'(T_0 - v)\cdot\Theta\cdot[\underline{B}(T_1 - v) - \underline{B}(T_f - v)]\,dv\right\}\exp\left[-\frac{\sigma_h^2(t)}{2}\right].
\end{aligned}$$

[33] In appendix 3.8.3 it is shown that equation (3.82) is equivalent to the analytical solution (3.79).

Using equation (3.83),

$$FP(T_0, T_f, T_1) = \frac{h(t, T_f, T_1)}{P(t, T_0)} \exp\left[-\frac{\sigma_h^2(t)}{2}\right] \exp\left\{\int_t^{T_0} \left[\underline{B}'(T_1 - v) - \underline{B}'(T_f - v)\right] \cdot S \cdot d\underline{W}^{\mathcal{Q}_0}(v)\right\}, \tag{3.84}$$

and thus

$$\Pr_{\mathcal{Q}_0}\left[FP(T_0, T_f, T_1) > K_f | \mathcal{F}_t\right] = \Pr_{\mathcal{Q}_0}\left\{\int_t^{T_0} \left[\underline{B}'(T_1 - v) - \underline{B}'(T_f - v)\right] \cdot S \cdot d\underline{W}^{\mathcal{Q}_0}(v) < \ln\left[\frac{h(t, T_f, T_1)}{P(t, T_0) K_f}\right] - \frac{\sigma_h^2(t)}{2} \middle| \mathcal{F}_t\right\}.$$

Because $\left\{\int_t^{T_0} \left[\underline{B}'(T_1 - v) - \underline{B}'(T_f - v)\right] \cdot S \cdot d\underline{W}^{\mathcal{Q}_0}(v)\right\} \backsim N^1\left(0, \sigma_h^2(t)\right)$, the standard normal probability distribution function can be applied to the standardized right-hand-side of the last inequality:

$$\begin{aligned}
\Pr_{\mathcal{Q}_0}\left[FP(T_0, T_f, T_1) > K_f | \mathcal{F}_t\right] &= \Phi\left\{\frac{\ln\left[\frac{h(t, T_f, T_1)}{P(t, T_0) K_f}\right] - \frac{\sigma_h^2(t)}{2}}{\sigma_h(t)}\right\} \\
&= \Phi\left[d_0^f(t)\right].
\end{aligned}$$

Concerning the evaluation of the $\mathcal{Q}_h$-exercise probability, and in order to find the probability distribution of $FP(T_0, T_f, T_1)$ under $\mathcal{Q}_h$, the Radon-Nikodym derivative η will be computed explicitly. Combining equations (3.74), (3.75), and (3.83),

$$\eta = \frac{FP(T_0, T_f, T_1)}{FP(t, T_f, T_1)} \exp\left\{-\int_t^{T_0} \underline{B}'(T_0 - v) \cdot \Theta \cdot \left[\underline{B}(T_1 - v) - \underline{B}(T_f - v)\right] dv\right\},$$

using equation (3.66),

$$\begin{aligned}
\eta &= \frac{\frac{P(T_0, T_1)}{P(T_0, T_f)}}{\frac{P(t, T_1)}{P(t, T_f)}} \exp\left\{\int_t^{T_0} \underline{B}'(T_f - v) \cdot \Theta \cdot \left[\underline{B}(T_1 - v) - \underline{B}(T_f - v)\right] dv\right\} \\
&\quad \exp\left\{-\int_t^{T_0} \underline{B}'(T_0 - v) \cdot \Theta \cdot \left[\underline{B}(T_1 - v) - \underline{B}(T_f - v)\right] dv\right\},
\end{aligned}$$

and considering equation (3.81),

$$\begin{aligned}
\eta &= \exp\left\{\int_t^{T_0} \left[\underline{B}'(T_f - v) - \underline{B}'(T_0 - v)\right] \cdot \Theta \cdot \left[\underline{B}(T_1 - v) - \underline{B}(T_f - v)\right] dv\right\} \\
&\quad \exp\left\{-\frac{1}{2}\int_t^{T_0} \left[\underline{H_1}'(v) \cdot \underline{H_1}(v) - \underline{H_f}'(v) \cdot \underline{H_f}(v)\right] dv\right\}
\end{aligned}$$

$$
\begin{aligned}
&\exp\left\{\int_t^{T_0}\left[\underline{B}'(T_1-v)-\underline{B}'(T_f-v)\right]\cdot S\cdot d\underline{W}^{Q_0}(v)\right\} \\
&= \exp\left\{-\frac{\sigma_h^2(t)}{2}+\int_t^{T_0}\left[\underline{B}'(T_1-v)-\underline{B}'(T_f-v)\right]\cdot S\cdot d\underline{W}^{Q_0}(v)\right\} \qquad (3.85)
\end{aligned}
$$

Applying Girsanov's Theorem to the above equality, it is possible to define by $\underline{W}^{Q_h}(t)$ a standard Brownian motion in $\Re^n$ under the probability measure Q_h, provided that:

$$\exp\left[\frac{1}{2}\sigma_h^2(t)\right] < \infty,$$

and

$$d\underline{W}^{Q_h}(t) = -S'\cdot\left[\underline{B}(T_1-t)-\underline{B}(T_f-t)\right]dt + d\underline{W}^{Q_0}(t).$$

Moreover, combining this last relation with equation (3.84), the probability distribution of $FP(T_0,T_f,T_1)$ under Q_h is obtained,

$$
\begin{aligned}
FP(T_0,T_f,T_1) &= \frac{h(t,T_f,T_1)}{P(t,T_0)}\exp\left[\frac{\sigma_h^2(t)}{2}\right] \\
&\quad \exp\left\{\int_t^{T_0}\left[\underline{B}'(T_1-v)-\underline{B}'(T_f-v)\right]\cdot S\cdot d\underline{W}^{Q_h}(v)\right\},
\end{aligned}
$$

and therefore,

$$
\begin{aligned}
\Pr_{Q_h}\left[FP(T_0,T_f,T_1)>K_f\mid\mathcal{F}_t\right] &= \Pr_{Q_h}\left\{\int_t^{T_0}\left[\underline{B}'(T_1-v)-\underline{B}'(T_f-v)\right]\cdot S\cdot d\underline{W}^{Q_h}(v)\right. \\
&\quad \left. < \ln\left[\frac{h(t,T_f,T_1)}{P(t,T_0)K_f}\right]+\frac{\sigma_h^2(t)}{2}\,\middle|\,\mathcal{F}_t\right\}.
\end{aligned}
$$

Finally, because $\left\{\int_t^{T_0}\left[\underline{B}'(T_1-v)-\underline{B}'(T_f-v)\right]\cdot S\cdot d\underline{W}^{Q_h}(v)\right\} \backsim N^1\left(0,\sigma_h^2(t)\right)$,

$$
\begin{aligned}
\Pr_{Q_h}\left[FP(T_0,T_f,T_1)>K_f\mid\mathcal{F}_t\right] &= \Phi\left\{\frac{\ln\left[\frac{h(t,T_f,T_1)}{P(t,T_0)K_f}\right]+\frac{\sigma_h^2(t)}{2}}{\sigma_h(t)}\right\} \\
&= \Phi\left[d_1^f(t)\right].
\end{aligned}
$$

Combining the previous results, equation (3.77) is derived. And, substituting equation (3.77) into the put-call parity for European options on pure discount bond futures, which, according to Chen (1995, page 364), can be stated as

$$c_t\left[FP(t,T_f,T_1);K_f;T_0\right]-p_t\left[FP(t,T_f,T_1);K_f;T_0\right]=h(t,T_f,T_1)-K_fP(t,T_0),$$

the analytical solution -(3.80)- for European put options follows immediately. ■

3.6.2 Futures options on coupon-bearing bonds

Denoting by $c_t[FS(t,T_f);X_f;T_0]$ the time-t fair price of an European call on the coupon-bearing bond future $FS(t,T_f)$, with a strike price of X_f, and expiry date at time T_0, and using result (3.8),

$$c_t[FS(t,T_f);X_f;T_0] = P(t,T_0)\,E_{Q_0}\{[FS(T_0,T_f)-X_f]\,1_\delta|\,\mathcal{F}_t\},$$

where $\delta=\{\omega\in\Omega:\ FS(T_0,T_f)(\omega)>X_f\}$ represents the set of states of the world in which the option ends in-the-money. Applying relation (3.69), that is ignoring the existence of delivery options,

$$\begin{aligned} c_t[FS(t,T_f);X_f;T_0] &= \sum_{i=1}^{N_f} k_i P(t,T_0)\,E_{Q_0}[FP(T_0,T_f,T_i)\,1_\delta|\,\mathcal{F}_t] \\ &\quad -P(t,T_0)\,X_f E_{Q_0}(1_\delta|\,\mathcal{F}_t). \end{aligned}$$

Using the law of iterative expectations, and because $E_{Q_0}(1_\delta|\,\mathcal{F}_t)$ corresponds to the Q_0-probability that the option will be exercised at time T_0, then

$$\begin{aligned} c_t[FS(t,T_f);X_f;T_0] = &\sum_{i=1}^{N_f} k_i P(t,T_0)\,E_{Q_0}[FP(T_0,T_f,T_i)|\,\mathcal{F}_t]\,E_{Q_0}(\eta_i 1_\delta|\,\mathcal{F}_t) \\ &-P(t,T_0)\,X_f \Pr_{Q_0}[FS(T_0,T_f)>X_f|\,\mathcal{F}_t], \end{aligned}$$

where

$$\eta_i = \frac{FP(T_0,T_f,T_i)}{E_{Q_0}[FP(T_0,T_f,T_i)|\,\mathcal{F}_t]},\ for\ i=1,\ldots,N_f, \tag{3.86}$$

is again a Radon-Nikodym derivative such that $E_{Q_0}(\eta_i|\,\mathcal{F}_t)=1,\forall i$. If the time-$t$ Q_0-expected discounted zero-coupon bond futures prices are represented by

$$h(t,T_f,T_i)\equiv P(t,T_0)\,E_{Q_0}[FP(T_0,T_f,T_i)|\,\mathcal{F}_t],$$

then it can be shown that

$$\eta_i = \frac{h(T_0,T_f,T_i)}{h(t,T_f,T_i)}\frac{P(t,T_0)}{P(T_0,T_0)},$$

which means that the Radon-Nikodym derivative η_i represents the change from the numeraire $P(t,T_0)$ to a new numeraire $h(t,T_f,T_i)$. That is $\eta_i\equiv\left(\frac{dQ_{h_i}}{dQ_0}\right)_{T_0}$, where Q_{h_i} is the equivalent martingale measure associated with the numeraire $h(t,T_f,T_i)$, and therefore

$$E_{Q_0}(\eta_i 1_\delta|\,\mathcal{F}_t) = E_{Q_{h_i}}(1_\delta|\,\mathcal{F}_t)$$

$$= \Pr_{Q_{h_i}} \left[FS(T_0, T_f) > X_f | \mathcal{F}_t \right],$$

which corresponds to the Q_{h_i}-probability that the option will be exercised at time T_0.

In summary,

$$c_t \left[FS(t, T_f); X_f; T_0 \right] = \sum_{i=1}^{N_f} k_i h(t, T_f, T_i) \Pr_{Q_{h_i}} \left[FS(T_0, T_f) > X_f | \mathcal{F}_t \right] \qquad (3.87)$$
$$- P(t, T_0) X_f \Pr_{Q_0} \left[FS(T_0, T_f) > X_f | \mathcal{F}_t \right].$$

The following proposition transforms the last probabilistic valuation formula into an explicit solution, using the *rank 1* approximation described in subsection 3.4.3.[34]

Proposition 10 *Under the Gaussian specification of the Duffie and Kan (1996) model, the time-t price of an European conventional call with a strike price of X_f, expiry date at time T_0 ($> t$), and on a futures contract, for delivery at date T_f ($> T_0$), on a default-free coupon-bearing bond generating N_f cash flows of value k_i at times T_i ($> T_f, i = 1, \ldots, N_f$), is approximately equal to*

$$c_t \left[FS(t, T_f); X_f; T_0 \right] \cong \sum_{i=1}^{N_f} k_i h(t, T_f, T_i) \Phi \left[d_i^f(t) \right] - P(t, T_0) X_f \Phi \left[d_0^f(t) \right], \qquad (3.88)$$

where

$$d_0^f(t) : \sum_{i=1}^{N_f} k_i h(t, T_f, T_i) \exp \left[-\frac{\sigma_{h_i}^2(t)}{2} - \sigma_{h_i}(t) d_0^f(t) \right] = P(t, T_0) X_f, \qquad (3.89)$$

$$d_i^f(t) = d_0^f(t) + \sigma_{h_i}(t),$$

$$h(t, T_f, T_i) = P(t, T_0) FP(t, T_f, T_i) \exp \left[I_i(t) \right], \qquad (3.90)$$

$$I_i(t) = \underline{G}' \cdot a^{-1} \cdot \Theta \cdot (a')^{-1} \cdot \left[\underline{B}(T_i - t) + \underline{B}(T_f - T_0) - \underline{B}(T_i - T_0) - \underline{B}(T_f - t) \right]$$
$$+ \underline{G}' \cdot a^{-1} \cdot \Delta(T_0 - t) \cdot \left[\underline{B}(T_f - T_0) - \underline{B}(T_i - T_0) \right],$$

and

$$\sigma_{h_i}^2(t) = \left[\underline{B}'(T_i - T_0) - \underline{B}'(T_f - T_0) \right] \cdot \Delta(T_0 - t) \cdot \left[\underline{B}(T_i - T_0) - \underline{B}(T_f - T_0) \right]. \qquad (3.91)$$

[34] The *lognormal* approximation of Pang (1996) could have also been easily adapted to the present context.

The corresponding approximate put price is

$$p_t\left[FS\left(t,T_f\right);X_f;T_0\right] \cong -\sum_{i=1}^{N_f} k_i h\left(t,T_f,T_i\right)\Phi\left[-d_i^f(t)\right] + P\left(t,T_0\right)X_f\Phi\left[-d_0^f(t)\right]. \quad (3.92)$$

Proof. Following exactly the same steps as in the derivation of proposition 9,[35] equation (3.90) is obtained.

Similarly, the probability distribution of $FP\left(T_0,T_f,T_i\right)$ under Q_0 is given by equations (3.84) and (3.82), when 1 is replaced by i:

$$\begin{aligned} FP\left(T_0,T_f,T_i\right) &= \frac{h\left(t,T_f,T_i\right)}{P\left(t,T_0\right)}\exp\left[-\frac{\sigma_{h_i}^2(t)}{2}\right] \\ &\quad \exp\left\{-\int_t^{T_0}\left[\underline{B}'\left(T_f-v\right)-\underline{B}'\left(T_i-v\right)\right]\cdot S\cdot d\underline{W}^{Q_0}(v)\right\}, \end{aligned} \quad (3.93)$$

where[36]

$$\sigma_{h_i}^2(t) = \int_t^{T_0}\left[\underline{B}'\left(T_f-v\right)-\underline{B}'\left(T_i-v\right)\right]\cdot\Theta\cdot\left[\underline{B}\left(T_f-v\right)-\underline{B}\left(T_i-v\right)\right]dv. \quad (3.94)$$

Hence, combining equations (3.69) and (3.93),

$$\begin{aligned} &\Pr_{Q_0}\left[FS\left(T_0,T_f\right)>X_f\middle|\mathcal{F}_t\right] \\ &= \Pr_{Q_0}\left\{\sum_{i=1}^{N_f}k_i\frac{h\left(t,T_f,T_i\right)}{P\left(t,T_0\right)}\exp\left[-\frac{\sigma_{h_i}^2(t)}{2}\right.\right. \\ &\quad \left.\left.-\int_t^{T_0}\left(\underline{B}'\left(T_f-v\right)-\underline{B}'\left(T_i-v\right)\right)\cdot S\cdot d\underline{W}^{Q_0}(v)\right]>X_f\middle|\mathcal{F}_t\right\}. \end{aligned}$$

Moreover, since $\left\{\int_t^{T_0}\left[\underline{B}'\left(T_f-v\right)-\underline{B}'\left(T_i-v\right)\right]\cdot S\cdot d\underline{W}^{Q_0}(v)\right\}\backsim N^1\left(0,\sigma_{h_i}^2(t)\right)$, and using a *rank 1* approximation, i.e. assuming that

$$\int_t^{T_0}\left[\underline{B}'\left(T_f-v\right)-\underline{B}'\left(T_i-v\right)\right]\cdot S\cdot d\underline{W}^{Q_0}(v)\cong\sigma_{h_i}(t)Z,\ for\ i=1,\ldots,N_f,$$

where $Z\backsim N^1(0,1)$, then

$$\begin{aligned} &\Pr_{Q_0}\left[FS\left(T_0,T_f\right)>X_f\middle|\mathcal{F}_t\right] \\ &\cong \Pr_{Q_0}\left\{\sum_{i=1}^{N_f}k_i h\left(t,T_f,T_i\right)\exp\left[-\frac{\sigma_{h_i}^2(t)}{2}-\sigma_{h_i}(t)Z\right]>P\left(t,T_0\right)X_f\middle|\mathcal{F}_t\right\}. \end{aligned}$$

---

[35] Or, simply by substituting 1 by i in equation (3.78).

[36] It can be easily shown that the integral equation (3.94) yields the explicit expression (3.91).

Finally, because the left-hand-side of the above inequality is a decreasing function of the standard normal random variable Z,

$$\begin{aligned}\Pr_{Q_0}\left[FS(T_0,T_f) > X_f \middle| \mathcal{F}_t\right] &\cong \Pr_{Q_0}\left[Z < d_0^f(t) \middle| \mathcal{F}_t\right] \\ &\cong \Phi\left[d_0^f(t)\right].\end{aligned}$$

In order to compute

$$\Pr_{Q_{h_i}}\left[FS(T_0,T_f) > X_f \middle| \mathcal{F}_t\right] = \Pr_{Q_{h_i}}\left[\sum_{i=1}^{N_f} k_i FP(T_0,T_f,T_i) > X_f \middle| \mathcal{F}_t\right], \tag{3.95}$$

it is necessary to obtain the probability distribution of $FP(T_0,T_f,T_i)$ under the equivalent martingale measure Q_{h_i}. For this purpose, the Radon-Nikodym derivative η_i, defined in equation (3.86), must be computed explicitly following the same methodology as in the derivation of proposition 9, that is by making $1 = i$ in equation (3.85):

$$\eta_i = \exp\left\{-\frac{\sigma_{h_i}^2(t)}{2} - \int_t^{T_0}\left[\underline{B}'(T_f - v) - \underline{B}'(T_i - v)\right] \cdot S \cdot d\underline{W}^{Q_0}(v)\right\}.$$

Applying Girsanov's Theorem to the above equality, and as long as

$$\exp\left[\frac{\sigma_{h_i}^2(t)}{2}\right] < \infty,$$

it follows that

$$d\underline{W}^{Q_{h_i}}(t) = S' \cdot \left[\underline{B}(T_f - t) - \underline{B}(T_i - t)\right] dt + d\underline{W}^{Q_0}(t),$$

where $\underline{W}^{Q_{h_i}}(t)$ is a standard Brownian motion in $\Re^n$ under the probability measure Q_{h_i}. Then, combining this last relation with equation (3.93), the probability distribution of $FP(T_0,T_f,T_i)$ under Q_{h_i} is obtained:

$$\begin{aligned}FP(T_0,T_f,T_i) &= \frac{h(t,T_f,T_i)}{P(t,T_0)}\exp\left[-\frac{\sigma_{h_i}^2(t)}{2} + \sigma_{h_i}^2(t)\right] \\ &\quad \exp\left\{-\int_t^{T_0}\left[\underline{B}'(T_f - v) - \underline{B}'(T_i - v)\right] \cdot S \cdot d\underline{W}^{Q_{h_i}}(v)\right\}.\end{aligned}$$

Substituting the above equality into expression (3.95), while assuming that

$$\int_t^{T_0}\left[\underline{B}'(T_f - v) - \underline{B}'(T_i - v)\right] \cdot S \cdot d\underline{W}^{Q_{h_i}}(v) \cong \sigma_{h_i}(t)\, Z, \; for\ i = 1, \ldots, N_f,$$

where $Z \backsim N^1(0,1)$, gives:

$$\begin{aligned}
\Pr_{Q_{h_i}}\left[FS(T_0,T_f) > X_f \middle| \mathcal{F}_t\right] &\cong \Pr_{Q_{h_i}}\left\{\sum_{i=1}^{N_f} k_i h(t,T_f,T_i) \exp\left[-\frac{\sigma^2_{h_i}(t)}{2} - \sigma_{h_i}(t)(Z - \sigma_{h_i}(t))\right]\right. \\
&\quad \left. > P(t,T_0) X_f \middle| \mathcal{F}_t\right\}.
\end{aligned}$$

And, because the left-hand-side of the last inequality is a decreasing function of $(Z - \sigma_{h_i}(t))$,

$$\begin{aligned}
\Pr_{Q_{h_i}}\left[FS(T_0,T_f) > X_f \middle| \mathcal{F}_t\right] &\cong \Pr_{Q_{h_i}}\left[Z - \sigma_{h_i}(t) < d_0^f(t) \middle| \mathcal{F}_t\right] \\
&\cong \Phi\left[d_i^f(t)\right],
\end{aligned}$$

which proves equation (3.88).

Concerning the derivation of the put price solution (3.92), result (3.8) and the law of iterative expectations imply that

$$\begin{aligned}
p_t\left[FS(t,T_f); X_f; T_0\right] &= -\sum_{i=1}^{N_f} k_i h(t,T_f,T_i) E_{Q_0}\left(\eta_h 1_{\delta^c} \middle| \mathcal{F}_t\right) \\
&\quad + P(t,T_0) X_f E_{Q_0}\left(1_{\delta^c} \middle| \mathcal{F}_t\right),
\end{aligned}$$

where $\delta^c = \{\omega \in \Omega : FS(T_0,T_f)(\omega) < X_f\}$ represents the set of states of the world in which the European put ends in-the-money. Using the definition of the new exercise set, and the symmetry of the normal distribution,

$$\begin{aligned}
E_{Q_0}\left(\eta_h 1_{\delta^c} \middle| \mathcal{F}_t\right) &= E_{Q_{h_i}}\left(1_{\delta^c} \middle| \mathcal{F}_t\right) \\
&= \Pr_{Q_{h_i}}\left[FS(T_0,T_f) < X_f \middle| \mathcal{F}_t\right] \\
&\cong 1 - \Phi\left[d_i^f(t)\right] \\
&\cong \Phi\left[-d_i^f(t)\right].
\end{aligned}$$

Similarly for the second expectation,

$$\begin{aligned}
E_{Q_0}\left(1_{\delta^c} \middle| \mathcal{F}_t\right) &= \Pr_{Q_0}\left[FS(T_0,T_f) < X_f \middle| \mathcal{F}_t\right] \\
&\cong 1 - \Phi\left[d_0^f(t)\right] \\
&\cong \Phi\left[-d_0^f(t)\right],
\end{aligned}$$

which proves the analytical approximate solution (3.92). ■

Remark 7 *The put-call parity for European options on coupon-bearing bond futures can be* *obtained by subtracting equation (3.92) from formula (3.88):*

$$c_t\left[FS\left(t,T_f\right);X_f;T_0\right]-p_t\left[FS\left(t,T_f\right);X_f;T_0\right]=\sum_{i=1}^{N_f}k_ih\left(t,T_f,T_i\right)-P\left(t,T_0\right)X_f.$$

3.6.3 Futures options on short-term interest rates

This subsection is devoted to the valuation of European futures options on short-term *nominal* "money-market" forward interest rates, and makes use of the symmetric credit risk assumption of Duffie and Singleton (1997). Next proposition considers the case of futures options with stock-style margining.

Proposition 11 *Under the deterministic volatility specification of the Duffie and Kan (1996) model, the time-t premium of an European conventional call on the futures contract $FR(t,T_f,T_1)$, with a strike price equal to K_R, and expiring at time T_0 (such that $t\leq T_0\leq T_f\leq T_1$), is equal to*

$$\begin{aligned}c_t\left[FR\left(t,T_f,T_1\right);K_R;T_0\right]= & \frac{100P\left(t,T_0\right)}{T_1-T_f}\left\{-\frac{P\left(t,T_f\right)}{P\left(t,T_1\right)}e^{L(T_0)+\rho(t)}\Phi\left[-d_1^R\left(t\right)\right]\right. \\ & \left.+\left[1+\left(T_1-T_f\right)\frac{100-K_R}{100}\right]\Phi\left[-d_0^R\left(t\right)\right]\right\},\end{aligned}\tag{3.96}$$

where

$$d_0^R\left(t\right)=\frac{\ln\left[\frac{\frac{P\left(t,T_f\right)}{P\left(t,T_1\right)}}{1+\left(T_1-T_f\right)\frac{100-K_R}{100}}\right]+L\left(T_0\right)+\rho\left(t\right)-\frac{\sigma_R^2(t)}{2}}{\sigma_R\left(t\right)},$$

$$d_1^R\left(t\right)=d_0^R\left(t\right)+\sigma_R\left(t\right),$$

$$\sigma_R^2\left(t\right)=\left[\underline{B}'\left(T_1-T_0\right)-\underline{B}'\left(T_f-T_0\right)\right]\cdot\Delta\left(\tau_0\right)\cdot\left[\underline{B}\left(T_1-T_0\right)-\underline{B}\left(T_f-T_0\right)\right],$$

and

$$\rho\left(t\right)=\left[\underline{B}'\left(T_1-T_0\right)-\underline{B}'\left(T_f-T_0\right)\right]\cdot\Delta\left(\tau_0\right)\cdot\underline{B}\left(T_1-T_0\right).$$

The time-t premium of the corresponding European conventional put option is given by

$$\begin{aligned}p_t\left[FR\left(t,T_f,T_1\right);K_R;T_0\right]= & \frac{100P\left(t,T_0\right)}{T_1-T_f}\left\{\frac{P\left(t,T_f\right)}{P\left(t,T_1\right)}e^{L(T_0)+\rho(t)}\Phi\left[d_1^R\left(t\right)\right]\right. \\ & \left.-\left[1+\left(T_1-T_f\right)\frac{100-K_R}{100}\right]\Phi\left[d_0^R\left(t\right)\right]\right\}.\end{aligned}\tag{3.97}$$

Proof. Using equation (3.73), the (intrinsic) terminal value of the call option is

$$c_{T_0}\left[FR\left(T_0,T_f,T_1\right);K_R;T_0\right]=\frac{100}{T_1-T_f}\left\{\left[1+\left(T_1-T_f\right)\frac{100-K_R}{100}\right]-\frac{\exp\left[L\left(T_0\right)\right]}{P\left(T_0,T_f,T_1\right)}\right\}^+,\tag{3.98}$$

where $P(T_0, T_f, T_1) \equiv \frac{P(T_0,T_1)}{P(T_0,T_f)}$ is the time-T_0 forward price, for delivery at date T_f, of a zero-coupon bond expiring at time T_1. And, applying result (3.8),

$$\frac{c_t\left[FR(t, T_f, T_1); K_R; T_0\right]}{P(t, T_0)} = E_{Q_0}\left\{c_{T_0}\left[FR(T_0, T_f, T_1); K_R; T_0\right] \mid \mathcal{F}_t\right\}. \tag{3.99}$$

In order to compute the last expectation, it is necessary to use the stochastic process followed by the forward price $P(T_0, T_f, T_1)$ under the risk-neutral measure Q_0, which, from equation (3.81), can be written as

$$P(T_0, T_f, T_1) = \frac{P(t, T_1)}{P(t, T_f)} \exp\left[\nu(t) - z\right], \tag{3.100}$$

where

$$\begin{aligned}
\nu(t) &= \underline{G}' \cdot a^{-1} \cdot \left\{ e^{a(T_1 - T_0)} \cdot \Delta(\tau_0) \cdot \left[I_n - \frac{1}{2} e^{a'(T_1 - T_0)} \right] \right. \\
&\quad \left. + e^{a(T_f - T_0)} \cdot \Delta(\tau_0) \cdot \left[\frac{1}{2} e^{a'(T_f - T_0)} - I_n \right] \right\} \cdot (a^{-1})' \cdot \underline{G},
\end{aligned}$$

and

$$z = \int_t^{T_0} \left[\underline{B}'(T_f - u) - \underline{B}'(T_1 - u)\right] \cdot S \cdot d\underline{W}^{Q_0}(u).$$

Because $z \backsim N^1\left(0, \sigma_R^2(t)\right)$, and combining equations (3.98) (3.99) and (3.100), yields

$$\begin{aligned}
& c_t\left[FR(t, T_f, T_1); K_R; T_0\right] \\
= & \frac{100 P(t, T_0)}{T_1 - T_f} \int_{-\infty}^{\infty} dz \frac{1}{\sigma_R(t)\sqrt{2\pi}} \exp\left[-\frac{z^2}{2\sigma_R^2(t)}\right] 1_{\{z < z^*\}} \\
& \left\{ \left[1 + (T_1 - T_f) \frac{100 - K_R}{100}\right] - \exp\left[L(T_0) - \nu(t)\right] \frac{P(t, T_f)}{P(t, T_1)} e^z \right\},
\end{aligned}$$

with

$$z^* = \ln\left[\frac{1 + (T_1 - T_f)\frac{100 - K_R}{100}}{\frac{P(t,T_f)}{P(t,T_1)}}\right] - L(T_0) + \nu(t).$$

Solving the last integral explicitly and defining $\rho(t) \equiv \frac{\sigma_R^2(t)}{2} - \nu(t)$, equation (3.96) is easily obtained. The put option solution (3.97) can be also derived along the same lines. ■

Remark 8 *If the maturity date of the futures option is the same as the delivery date of the underlying futures contract (as is the case, for instance, of the Quarterly Eurodollar futures options traded at the International Money Market Division of the Chicago Mercantile Exchange), equations (3.96) and (3.97) are still applicable but with T_0 replaced by T_f, i.e. with $L(T_0) + \rho(t) = \sigma_R^2(t)$.*

Remark 9 *If $T_0 = T_f$, then equation (3.98) can be rewritten as*

$$c_{T_f}\left[FR(T_f,T_f,T_1);K_R;T_f\right] = 100\left\{\frac{100-K_R}{100} - \frac{1}{T_1-T_f}\left[\frac{P(T_f,T_f)}{P(T_f,T_1)}-1\right]\right\}^+,$$

and

$$p_{T_f}\left[FR_G(T_f,T_f,T_1);K_R;T_f\right] = 100\left\{\frac{1}{T_1-T_f}\left[\frac{P(T_f,T_f)}{P(T_f,T_1)}-1\right] - \frac{100-K_R}{100}\right\}^+,$$

where $R(t,T_f,T_1) \equiv \frac{1}{T_1-T_f}\left[\frac{P(t,T_f)}{P(t,T_1)}-1\right]$ is the time-t nominal forward rate for the time period $(T_1 - T_f)$. Therefore, equations (3.96) and (3.97), when $T_0 = T_f$, are also the pricing solutions for European puts and calls, respectively, on the nominal forward interest rate $R(t,T_f,T_1)$, with a strike equal to $\frac{100-K_R}{100}$, with a contract size of 100, and with settlement at the option's expiry date (instead of settlement in arrears, as was the case in subsection 3.4.2).

All the valuation formulae derived so far in this subsection are only valid for futures options with stock-style margining. However, the short-term interest rate futures options traded at the London International Financial Futures Exchange (LIFFE) have futures-style margining requirements, that is are *pure* futures options according to Duffie (1989). This means that the option premium is not paid at the time of purchase, but only when the contract is exercised. Moreover, option positions are marked-to-market daily, in exactly the same way as the underlying futures contract. Next proposition takes these features into account.

Proposition 12 *Under the deterministic volatility specification of the Duffie and Kan (1996) model, the time-t premium of a pure European futures call on the futures contract $FR(t,T_f,T_1)$, with a strike price equal to K_R, and maturity at date T_0 (such that $t \leq T_0 \leq T_f \leq T_1$), is equal to*

$$\begin{aligned} & Fc_t\left[FR(t,T_f,T_1);K_R;T_0\right] \\ = & \frac{100}{T_1-T_f}\left\{\left[\frac{100-K_R}{100}\right]\Phi\left[-d_0^{FR}(t)\right] - \frac{P(t,T_f)}{P(t,T_1)}e^{L(t)}\Phi\left[-d_1^{FR}(t)\right]\right\}, \end{aligned} \tag{3.101}$$

where

$$d_0^{FR}(t) = \frac{\ln\left[\frac{\frac{P(t,T_f)}{P(t,T_1)}}{1+(T_1-T_f)\frac{100-K_R}{100}}\right] + L(t) - \frac{\sigma_R^2(t)}{2}}{\sigma_R(t)},$$

and

$$d_1^{FR}(t) = d_0^{FR}(t) + \sigma_R(t).$$

The time-t premium of the corresponding pure European futures put is given by

$$
\begin{aligned}
& Fp_t\left[FR\left(t,T_f,T_1\right);K_R;T_0\right] & (3.102)\\
= & \frac{100}{T_1-T_f}\left\{\frac{P\left(t,T_f\right)}{P\left(t,T_1\right)}\exp\left[L\left(t\right)\right]\Phi\left[d_1^{FR}\left(t\right)\right]-\left[1+\left(T_1-T_f\right)\frac{100-K_R}{100}\right]\Phi\left[d_0^{FR}\left(t\right)\right]\right\}.
\end{aligned}
$$

Proof. A *pure* futures option behaves just like a futures contract on the corresponding *conventional* futures option, and thus its price must be equal to the expectation, under measure Q, of the terminal underlying spot price. In the case of a *pure* European futures call,

$$
Fc_t\left[FR\left(t,T_f,T_1\right);K_R;T_0\right]=E_Q\left\{c_{T_0}\left[FR\left(T_0,T_f,T_1\right);K_R;T_0\right]\middle|\mathcal{F}_t\right\},
$$

or, using equation (3.98),

$$
\begin{aligned}
& Fc_t\left[FR\left(t,T_f,T_1\right);K_R;T_0\right] & (3.103)\\
= & \frac{100}{T_1-T_f}E_Q\left\{\left[\left(1+\left(T_1-T_f\right)\frac{100-K_R}{100}\right)-\frac{e^{L(T_0)}}{P\left(T_0,T_f,T_1\right)}\right]^+\middle|\mathcal{F}_t\right\}.
\end{aligned}
$$

In order to proceed, it is necessary to know the stochastic process followed by the forward price $P\left(T_0,T_f,T_1\right)$ under the martingale measure Q. Using equation (3.13), formula (3.100) can be rewritten as

$$
P\left(T_0,T_f,T_1\right)=\frac{P\left(t,T_1\right)}{P\left(t,T_f\right)}\exp\left[\eta\left(t\right)-w\right], \tag{3.104}
$$

where

$$
\begin{aligned}
\eta\left(t\right) = & \underline{G}'\cdot a^{-1}\left\{\left[e^{a\left(T_f-T_0\right)}-e^{a\left(T_1-T_0\right)}-e^{a\left(T_f-t\right)}+e^{a\left(T_1-t\right)}\right]\cdot a^{-1}\cdot\Theta\right.\\
& \left.-\frac{1}{2}e^{a\left(T_1-T_0\right)}\cdot\Delta\left(\tau_0\right)\cdot e^{a'\left(T_1-T_0\right)}+\frac{1}{2}e^{a\left(T_f-T_0\right)}\cdot\Delta\left(\tau_0\right)\cdot e^{a'\left(T_f-T_0\right)}\right\}\cdot\left(a^{-1}\right)'\cdot\underline{G},
\end{aligned}
$$

and

$$
w=\int_t^{T_0}\left[\underline{B}'\left(T_f-u\right)-\underline{B}'\left(T_1-u\right)\right]\cdot S\cdot d\underline{W}^Q\left(u\right).
$$

Combining equations (3.103) and (3.104), and because $w\backsim N^1\left(0,\sigma_R^2\left(t\right)\right)$, then

$$
\begin{aligned}
& Fc_t\left[FR\left(t,T_f,T_1\right);K_R;T_0\right]\\
= & \frac{100}{T_1-T_f}\int_{-\infty}^{\infty}dw\frac{1}{\sigma_R\left(t\right)\sqrt{2\pi}}\exp\left[-\frac{w^2}{2\sigma_R^2\left(t\right)}\right]1_{\{w<w^*\}}\\
& \left\{\left[1+\left(T_1-T_f\right)\frac{100-K_R}{100}\right]-\exp\left[L\left(T_0\right)-\eta\left(t\right)\right]\frac{P\left(t,T_f\right)}{P\left(t,T_1\right)}e^w\right\},
\end{aligned}
$$

where

$$w^* = \ln\left[\frac{1+(T_1-T_f)\frac{100-K_R}{100}}{\frac{P(t,T_f)}{P(t,T_1)}}\right] - L(T_0) + \eta(t)\,.$$

Finally, computing explicitly the above integral, and because $L(T_0) - \eta(t) + \frac{\sigma_R^2(t)}{2} = L(t)$, the closed-form solution (3.101) is easily obtained. The put option solution (3.102) can be easily derived by combining formula (3.101) with the following well known put-call parity:

$$Fc_t\left[FR(t,T_f,T_1);K_R;T_0\right] - Fp_t\left[FR(t,T_f,T_1);K_R;T_0\right] = FR(t,T_f,T_1) - K_R.$$

■

Remark 10 *Equations (3.101) and (3.102) can also be applied to value pure American futures options, because, and as shown by Chen and Scott (1993a), the price of a pure American futures option before expiration will always exceed its intrinsic value, and therefore early exercise should not occur.*

3.7 Conclusions

This Chapter considered a deterministic volatility version of the Duffie and Kan (1996) model, which constitutes the most general Gaussian multifactor time-homogeneous and affine term structure model. Under such tractable framework, exact pricing solutions were derived for several European-style interest rate contingent claims, such as: interest rate futures, European options on zero-coupon bonds, caps and floors, European yield options, and European interest rate futures options. For European options on coupon-bearing bonds (and thus, for European swaptions), two types of approximations were described and compared: if the maturity of the underlying bond is close to the expiry date of the option, both approximations produce excellent results; if not, the *lognormal* approximation is best suited for pricing at-the-money contracts, while the *rank 1* approximation produces lower pricing errors for deep out-of-the-money strikes.

The Gaussian analytical pricing solutions obtained in the present chapter will be necessary to produce, in Chapter five, expedite approximate explicit pricing formulae for the same contracts, but under the more general stochastic volatility specification of the Duffie and Kan (1996) model. Before that, and in the next Chapter, propositions 1, 2, 3, and 4 will be used to fit a Gaussian affine state-space model to a panel-data of swap rates, cap prices, and swaption quotes. These propositions will generate the measurement equations required for the implementation of the Kalman filter recursions.

3.8 Appendices

3.8.1 A closed-form solution for $\underline{B}'(\tau)$

Following Edwards and Penney (1993, section 5.7, equation 43'), the unique solution of the initial value problem

$$\frac{d}{dt}\underline{y}(t) = A \cdot \underline{y}(t) + \underline{c}(t), \qquad \underline{y}(0) = \underline{y}_0,$$

with $\underline{y}(t) \in \Re^n$, $A \in \Re^{n \times n}$, and $\underline{c}(t) \in \Re^n$, is given by

$$\underline{y}(t) = e^{At} \cdot \underline{y}_0 + e^{At} \cdot \int_0^t e^{-As} \cdot \underline{c}(s)\, ds,$$

where e^{At} is the fundamental matrix of the associated homogeneous system[37]. That is $\underline{y}(t)$ is obtained by adding the solution of the correspondent homogenous system (known as the complementary function) to a particular solution of the nonhomogeneous system.

Applying the above result to the initial value problem given by equation (2.8) subject to the initial condition $\underline{B}(0) = \underline{0}$, it follows that

$$\underline{B}(\tau) = e^{a'\tau} \cdot \underline{0} + e^{a'\tau} \cdot \int_0^{T-t} e^{-a's} \cdot (-\underline{G})\, ds.$$

Since $\frac{d}{dt}\left(e^{At}\right) = A \cdot e^{At}$, because $e^{O_n} = I_n$, and assuming that matrix a is of full rank, then

$$\underline{B}(\tau) = e^{a'\tau} \cdot (a')^{-1} \cdot \left(e^{-a'\tau} - I_n\right) \cdot \underline{G}.$$

Finally, because $A^p \cdot e^{At} = e^{At} \cdot A^p$ for any integer p, equation (3.2) follows.

3.8.2 A closed-form solution for $\Delta(T-t)$

Because the eigenvectors of matrix a are assumed to be linearly independent, then $a = Q \cdot \Lambda \cdot Q^{-1}$ and $e^{a(T-t)} = Q \cdot \exp[\Lambda(T-t)] \cdot Q^{-1}$, where Q is a $n \times n$ matrix with columns corresponding to the eigenvectors of matrix a, and $\Lambda = diag\{\lambda_1, \ldots, \lambda_n\}$. Therefore, equation (3.4) becomes

$$\Delta(T-t) = \int_t^T Q \cdot e^{\Lambda(T-v)} \cdot \Theta^{\bullet} \cdot e^{\Lambda(T-v)} \cdot Q' dv,$$

where $\Theta^{\bullet} = Q^{-1} \cdot \Theta \cdot (Q^{-1})' \equiv \left\{\sigma^{\bullet}_{ij}\right\}_{i,j=1,\ldots,n}$. Moreover, because $e^{\Lambda(T-v)}$ is a diagonal matrix with the i^{th} element of the main diagonal equal to $e^{\lambda_i(T-v)}$, and following Langetieg

[37] I.e. $\frac{d}{dt}\underline{y}(t) = A \cdot \underline{y}(t)$.

(1980, footnote 23),

$$e^{\Lambda(T-v)} \cdot \Theta^{*} \cdot e^{\Lambda(T-v)} = \left\{ e^{(\lambda_i+\lambda_j)(T-v)} \sigma^{*}_{ij} \right\}_{i,j=1,\dots,n}.$$

Using this last result, the integral $\Delta(T-t)$ can now be computed analytically:

$$\begin{aligned}
\Delta(T-t) &= \left[Q \cdot \left\{ -e^{(\lambda_i+\lambda_j)(T-v)} \frac{\sigma^{*}_{ij}}{\lambda_i+\lambda_j} \right\}_{i,j=1,\dots,n} \cdot Q' \right]_t^T \\
&= -Q \cdot \left\{ \frac{\sigma^{*}_{ij}}{\lambda_i+\lambda_j} \right\}_{i,j=1,\dots,n} \cdot Q' \\
&\quad + Q \cdot \left\{ e^{(\lambda_i+\lambda_j)(T-t)} \frac{\sigma^{*}_{ij}}{\lambda_i+\lambda_j} \right\}_{i,j=1,\dots,n} \cdot Q' \\
&= -Q \cdot \Theta^{**} \cdot Q' + Q \cdot e^{\Lambda(T-t)} \cdot \Theta^{**} \cdot e^{\Lambda(T-t)} \cdot Q',
\end{aligned}$$

where $\Theta^{**} = \left\{ \frac{\sigma^{*}_{ij}}{\lambda_i+\lambda_j} \right\}_{i,j=1,\dots,n}$. Finally, using the change of notation $Y = Q \cdot \Theta^{**} \cdot Q'$,

$$\Delta(T-t) = Q \cdot e^{\Lambda(T-t)} \cdot Q^{-1} \cdot Q \cdot \Theta^{**} \cdot Q' \cdot (Q')^{-1} \cdot e^{\Lambda(T-t)} \cdot Q' - Y,$$

and equation (3.7) follows.

3.8.3 A closed-form solution for $\sigma_h^2(t)$

Starting from equation (3.82), using result (3.2),

$$\begin{aligned}
\sigma_h^2(t) &= \underline{G}' \cdot a^{-1} \cdot \left[\int_t^{T_0} e^{a(T_f-v)} \cdot \Theta \cdot e^{a'(T_f-v)} dv - \int_t^{T_0} e^{a(T_f-v)} \cdot \Theta \cdot e^{a'(T_1-v)} dv \right. \\
&\quad \left. - \int_t^{T_0} e^{a(T_1-v)} \cdot \Theta \cdot e^{a'(T_f-v)} dv + \int_t^{T_0} e^{a(T_1-v)} \cdot \Theta \cdot e^{a'(T_1-v)} dv \right] \cdot (a^{-1})' \cdot \underline{G},
\end{aligned}$$

and expressing all the four previous integrals in terms of $\Delta(T_0-t)$,

$$\begin{aligned}
\sigma_h^2(t) &= \underline{G}' \cdot a^{-1} \cdot \left[e^{a(T_f-T_0)} \cdot \Delta(\tau_0) \cdot e^{a'(T_f-T_0)} - e^{a(T_f-T_0)} \cdot \Delta(\tau_0) \cdot e^{a'(T_1-T_0)} \right. \\
&\quad \left. - e^{a(T_1-T_0)} \cdot \Delta(\tau_0) \cdot e^{a'(T_f-T_0)} + e^{a(T_1-T_0)} \cdot \Delta(\tau_0) \cdot e^{a'(T_1-T_0)} \right] \cdot (a^{-1})' \cdot \underline{G}.
\end{aligned}$$

Finally, combining the last result with equation (3.2), formula (3.79) is derived.

3.8.4 A closed-form solution for $h(t, T_f, T_1)$

Using equation (3.2) and simplifying,

$$\int_t^{T_0} \underline{B}'(T_0-v) \cdot \Theta \cdot [\underline{B}(T_1-v) - \underline{B}(T_f-v)]\, dv$$

$$
\begin{aligned}
= \quad & \underline{G}' \cdot a^{-1} \cdot \Theta \cdot \left\{ \int_t^{T_0} \left[e^{a'(T_f - v)} - e^{a'(T_1 - v)} \right] dv \right\} \cdot \left(a^{-1}\right)' \cdot \underline{G} \\
& - \underline{G}' \cdot a^{-1} \cdot \left\{ \int_t^{T_0} e^{a(T_0 - v)} \cdot \Theta \cdot \left[e^{a'(T_f - v)} - e^{a'(T_1 - v)} \right] dv \right\} \cdot \left(a^{-1}\right)' \cdot \underline{G}.
\end{aligned}
$$

Solving the first integral on the right-hand-side, and expressing the second one as a function of $\Delta\,(T_0 - t)$,

$$
\begin{aligned}
& \int_t^{T_0} \underline{B}'\,(T_0 - v) \cdot \Theta \cdot \left[\underline{B}\,(T_1 - v) - \underline{B}\,(T_f - v) \right] dv \\
= \quad & - \underline{G}' \cdot a^{-1} \cdot \Theta \cdot (a')^{-1} \cdot \left\{ e^{a'(T_f - T_0)} - e^{a'(T_1 - T_0)} - e^{a'(T_f - t)} + e^{a'(T_1 - t)} \right\} \cdot \left(a^{-1}\right)' \cdot \underline{G} \\
& - \underline{G}' \cdot a^{-1} \cdot \Delta\,(T_0 - t) \cdot \left[e^{a'(T_f - T_0)} - e^{a'(T_1 - T_0)} \right] \cdot \left(a^{-1}\right)' \cdot \underline{G}.
\end{aligned}
$$

Combining the above expression with (3.83), and using again equation (3.2), yields solution (3.78).

Chapter 4

A Gaussian State-Space Formulation

This Chapter is based on the paper Nunes and Clewlow (1999), presented at the 9th Annual Derivatives Securities Conference (Boston, 1999) and at the 26th Annual Meeting of the European Finance Association (Helsinki, 1999).

4.1 Introduction

Based on the dimensionality of the data under consideration, it is possible to distinguish three different approaches towards the estimation of time-homogeneous term structure models: the "time-series", the "cross-section", and the "panel-data" approaches. In the first case -as, for instance, in Chan et al. (1992) or Andersen and Lund (1997)- the model' parameters are estimated using a time-series of state variables (or of their observable proxies) values, and therefore only the dynamics of the yield curve are captured. In the second case -for example, in Brown and Dybvig (1986) or Brown and Schaefer (1994b)- the model is fitted to a cross-section of market observables (e.g. prices of coupon-bearing bonds for several maturities), and thus only the current shape of the yield curve is incorporated. Obviously, these two more traditional methodologies can be said to be inefficient in the sense that they do not make use of the full set of market (spatial) information available.[1]

The "panel-data" approach has the advantage of using simultaneously time-series and cross-sectional data (enhancing, therefore, the efficiency of the parameters' estimates). It uses, in a consistent way, both objective and risk-neutral model dynamics (yielding estimates for the parameters related to the market price of risk) and can be further divided into two

[1] Furthermore, while the first approach is "incomplete" since it does not provide estimates for the parameters defining the investor's preferences, the second ("implied") method is not theoretically justified because it can yield a different set of (supposedly fixed) parameter's values for each time period.

categories, depending on the assumptions made about measurement errors. In the first category, one or more points of the yield curve are assumed to be observable without error -see, for instance, Pearson and Sun (1994), Chen and Scott (1993b), or Duffie and Singleton (1997)- and consequently the model' factors can be identified exactly without the use of filtering techniques. Although the *zero measurement errors* assumption allows the model to be inverted in such a way that the state variables can be expressed as deterministic functions of the observed data, in practice, such hypothesis is contradicted by the existence of several market imperfections (such as transaction costs, liquidity premiums, taxation effects, or simply the non-synchronous arrival of market information). More realistically, the second category explicitly considers the existence of measurement errors in the data by representing the term structure model in a *state-space* form, and treats the state variables as truly unobservable by estimating the model' parameters through the use of a Kalman filter. The present Chapter will hereafter focus on this latter "panel-data", "state-space" approach.

During the current decade, several authors have used Kalman filtering techniques to estimate, mainly, *time-homogeneous* term structure models belonging to the *exponential-affine* class characterized by Duffie and Kan (1996). Traditionally, the model specification fitted to the panel-data of market observables under analysis has been either a restricted[2] version of the Langetieg (1980) model, or the multi-factor version of the Cox et al. (1985b) model: see, for example, Jegadeesh and Pennacchi (1996), Berardi (1997), or Babbs and Nowman (1999) for the Gaussian case; and Duan and Simonato (1995), Chen and Scott (1995a), or Geyer and Pichler (1996) for the "square-root" specification. In any case, the *time-homogeneous* assumption is required because, in order to fit the model both through time and cross-sectionally, it is necessary to know the stochastic differential equation followed by the state variables under the objective probabilities as well as with respect to an equivalent martingale measure.[3] On the other hand, the intensive use of the *exponential-affine* class of models in the empirical literature is mainly a consequence of its analytical tractability: it offers exact closed-form solutions for bond prices (or spot yields), which have traditionally been used to estimate the model' parameters.[4]

[2]For instance, the double-decay model proposed by Beaglehole and Tenney (1991). As it will become clear latter, restrictions are needed in order to transform the over-parameterized Langetieg (1980) model into an identifiable specification that ensures the non-singularity of the model' information matrix.

[3]As a possible exception, Bhar and Chiarella (1996) estimate a one-factor Markovian Heath-Jarrow-Morton (HJM) model using a non-linear Kalman filter, although their model is fitted not to a panel-data but only to one time-series of bond prices.

[4]Honoré (1998) stays outside this simple affine class, since it estimates, using a "panel-data" approach, non-linear term structure models with no closed-form solution for bond prices. However, by imposing appropriate linear restrictions on the measurement errors, no Kalman filtering technique is required to estimate the parameters.

In a first stage, the term structure model that will be employed in this Chapter is the Dai and Singleton (1998) general Gaussian canonical formulation (with unobservable state variables), which in essence can be made equivalent to the Langetieg (1980) model through an affine invariant transformation that ensures model identifiability. Although the framework is not new, this Gaussian time-homogeneous exponential-affine model will be estimated using a panel-data consisting simultaneously of swap rates, at-the-money (ATM) cap prices, and ATM European swaption prices. That is, the distinctive feature and main contribution of the empirical work presented in this Chapter consists in applying Kalman filtering techniques (for the first time, to the author's knowledge) to market data containing not only information about the level of the yield surface but also information about the volatility and the correlation surfaces. As Rebonato (1998, page 372) clearly summarizes: "...swaps and FRAs price the level of (different portions of) the yield curve; caps and floors price the volatility of (i.e. the diagonal elements of the covariance matrix among) the different forward rates; swaptions assign a price both to the diagonal and to the off-diagonal elements of the same covariance matrix...". Consequently, the purposed enlargement of the set of market observables, from which the factors are filtered and the parameters are estimated, is intended to capture a wider range of market sources of uncertainty, enabling the estimated model to price (accurately) also a broader class of interest rate contingent claims.

The motivation to use a Gaussian framework (besides its well known analytical appealing features) and an enlarged data set is twofold. On one hand, Babbs and Nowman (1999) have shown that a simple two-factor affine Gaussian model is able to fit remarkably well the term structure of (USD) interest rates. It would be surely interesting to test whether such a simple model is also able to explain the behavior of the term structures of interest rate volatilities and correlations. On the other hand, the suggestion put forward by Rogers and Stummer (1994, page 28) concerning the use of other liquid derivatives prices in the estimation stage, besides the traditional ones related to the level of the yield curve, is also considered.

As in Babbs and Nowman (1999), the empirical results obtained in this Chapter also suggest that a simple affine and time-homogeneous Gaussian model, with two or three factors, can fit the yield curve under a one basis point precision level. However, when the same model is also fitted to both cap and European swaption prices, it is never possible to reproduce the hump observed at the short-end of the volatility curve and swaptions are heavily mispriced. To test whether such inability to explain the market cap and swaption prices is simply the result of modelling the volatility term structure as being time-homogeneous, an equivalent Gauss-Markov Heath et al. (1992) model is proposed, where the diffusion term (of the instantaneous forward rate stochastic differential equation) is augmented by a time-dependent parameter. In order to retain the most explanatory power within the time-homogeneous framework, such arbitrage-free model is then estimated in two stages: first, the time-independent parameter's values are taken directly from the Kalman filter estimation of the corresponding "equilibrium" model specification; secondly, the time-inhomogeneous parameter is only used to improve the cross-sectional fit of the term structures of volatilities and/or of interest rate correlations. It will be shown that the exact pricing of swaps and caps can now be achieved, while European swaptions have much lower pricing errors, in the context of stable time-homogeneous diffusion coefficients.

In summary, the purpose of this Chapter is to address the following four empirical questions:

1. The ability of simple (Gaussian) time-homogenous affine models to fit not only the term structure of interest rates, but also the market volatility and correlation functions.

   It is well known that the selected term structure model suffers from strong theoretical limitations: besides the extensively reported possibility of attaining negative interest rates,[5] also the term structure of interest rate volatilities is not time but rather maturity dependent. And, although, for most currencies, the shape of the volatility curve seems to be similar over time, it can never be considered as constant. Nevertheless, it will be tested whether increasing successively the number of model state variables (instead of incorporating time-dependent parameters into the model) will allow a better description of the ***structural*** behavior of the market volatility and correlation functions, while avoiding the fitting of data noise.

2. The dimensionality requirements for "non-PC based" factor-models.

   Rebonato (1998, page 70) suggests that the 3-factor usual "prescription" (*level*, *slope*, and *curvature*) offered by a Principal Component Analysis (PCA) underestimates the number of state variables needed to fit the market correlation structure, even for models where the factors are not taken to be the first three principal components, because the observed ***fast decorrelation phenomena*** between adjacent forward rates in the short end of the yield curve is not conveniently reproduced. In this Chapter, by increasing successively the number of model' factors, it will be tested empirically which is the dimension required for the model correlation function to move from

[5] Rogers (1996) identifies the potential pricing implications of this model "deficiency".

the traditional sigmoid shape to the more realistic exponential-decaying behavior. Moreover, it will be shown analytically that the result obtained previously by Cooper and Rebonato (1995) in the context of a "simple" two-factor model can indeed be extended to an arbitrary large number of state variables.

3. The efficiency of linear *versus* non-linear filtering methods.

   Since the model state variables are not affine functions of the swap rates, the cap prices, or the swaption prices, the model parameters will have to be estimated through the use of a non-linear Kalman filter. Furthermore, because exact non-linear filtering techniques would be numerically over-intensive for the multi-factor model specifications implemented in this Chapter, an approximate non-linear filtering method will be employed, which, of course, is sub-optimal in the sense that it does not provide exact asymptotic properties for the parameter's estimates. But, linearizing the data (for instance, obtaining spot rates from swap rates) in order to be able to apply exact filtering techniques, can also be argued to imply not only a loss of market information (swap rates contain more information than spot rates of equal maturity) but also an additional bias in the final parameter's estimates (depending on the accuracy of the linearization method employed). In this Chapter, and wherever possible, the two approaches will be compared.

4. The relevance of the time-homogeneous assumption.

   In the final part of this Chapter, the Dai and Singleton (1998) Gaussian model will be converted into an equivalent Gauss-Markov HJM formulation, but with a time-inhomogeneous diffusion component, which will be fitted (cross-sectionally) to the same data set. Both "equilibrium" (i.e. state-space) and "HJM" specifications will be compared in terms of parameter's stability and pricing accuracy.

The remainder of this Chapter is organized as follows. Section 4.2 presents the interest rate "equilibrium" model specification under analysis, and reviews the analytical pricing formulae provided by such Gaussian framework for swaps, caps, and European swaptions. In section 4.3, the interest rate model is specified in a state-space form, the Kalman filtering estimation technique is described, and a Monte Carlo study is presented. Section 4.4 describes the US LIBOR-rate derivatives data set under analysis, while section 4.5 is devoted to the estimation of the state-space model, for different numbers of state variables, and through both linear and non-linear Kalman filtering algorithms. Then, section 4.6 purposes a similar time-inhomogeneous Gauss-Markov HJM model, which is also calibrated to the same data set. Section 4.7 summarizes the main conclusions of the Chapter. In the appendix, the empirical analysis described in sections 4.5 and 4.6 is replicated for another data set containing UK swap rates, cap prices, and European swaption prices.

4.2 Model description

4.2.1 Dai and Singleton (1998) Gaussian formulation

As in Duffie and Kan (1996), Dai and Singleton (1998) also formulate a general (stochastic volatility) time-homogeneous exponential-affine model, but starting from objective probabilities, i.e. from measure $\mathcal{P}$. The time-t short-term interest rate, $r(t)$, is still given by equation (2.2), i.e. it is still described by an affine function of the model' factors, and, under the original probability measure $\mathcal{P}$, the state variables evolve through time according to the following Markov diffusion process:

$$d\underline{X}(t) = K \cdot [\underline{\theta} - \underline{X}(t)]\, dt + \Sigma \cdot \sqrt{V^D(t)} \cdot d\underline{W}^{\mathcal{P}}(t), \tag{4.1}$$

where $K, \Sigma \in \Re^{n \times n}$, $\underline{\theta} \in \Re^n$, and with matrix $\sqrt{V^D(t)}$ defined as in equation (2.5), that is

$$\sqrt{V^D(t)} = diag\left\{\sqrt{v_1(t)}, \ldots, \sqrt{v_n(t)}\right\},$$

$$v_i(t) = \alpha_i + \underline{\beta}_i' \cdot \underline{X}(t), \quad for\ i = 1, \ldots, n,$$

$\alpha_i \in \Re$, and $\underline{\beta}_i \in \Re^n$. The vector $d\underline{W}^{\mathcal{P}}(t) \in \Re^n$ contains n independent and $\mathcal{P}$-measurable Brownian motion increments.

By assuming a particular analytical specification for the vector of market prices of risk,

$$\underline{\Lambda}(t) = \sqrt{V^D(t)} \cdot \underline{\lambda}, \tag{4.2}$$

where $\underline{\lambda} \in \Re^n$, Dai and Singleton (1998) are able to define an equivalent martingale probability measure $\mathcal{Q}$ on the same measurable space $(\Omega, \mathcal{F})$, such that

$$d\underline{W}^{\mathcal{Q}}(t) = \underline{\Lambda}(t)\, dt + d\underline{W}^{\mathcal{P}}(t) \tag{4.3}$$

is, under measure $\mathcal{Q}$, also a vector of standard Brownian motion increments in $\Re^n$ (with the same standard filtration as $d\underline{W}^{\mathcal{P}}(t)$). Combining equations (4.1), (4.2), and (4.3), a risk-adjusted stochastic process can be obtained for the vector of state variables, and interest rate contingent claims can be priced.

The functional form adopted for $\underline{\Lambda}(t)$ is analytically convenient, because it supports an affine risk-neutral drift for the state variables. Moreover, equation (4.2) is also theoretically justified because the Dai and Singleton (1998) specification under analysis can be nested into the general equilibrium Duffie and Kan (1996) formulation derived in Chapter two. In fact, using Theorem 4, that is working under preferences described by a power utility function, it follows that $\bar{a} = -K$, $\underline{b} = K \cdot \underline{\theta}$, and $\underline{\lambda} = \underline{\chi} - \gamma\underline{\varphi}$, i.e. $\underline{\lambda}$ is defined in terms of the Pratt's measure of relative risk aversion as well as in terms of the diffusions for the exogenously specified output and money supply stochastic processes. Therefore, the framework that will be used for the empirical analysis presented in this Chapter can be viewed as a particular case of the general equilibrium formulation derived in Chapter two.

Under this general affine formulation, and attending to the total number of model factors (n) as well as to the number of state variables (m) that are included in the diffusion term of equation (4.1), Dai and Singleton (1998) define the restrictions that must be imposed to the model parameters in order for the model to be both admissible[6] and just-identified. For each pair (n, m), such restrictions define what Dai and Singleton (1998) denominate as the $\mathbb{A}_m(n)$ canonical models.

Since this Chapter restricts its attention to the (simpler) Gaussian case, it will adopt from now on the Dai and Singleton (1998) $\mathbb{A}_0(n)$ canonical model, which is obtainable from (4.1), (4.2), and (4.3) by imposing two types of restrictions. First, and in order for the instantaneous volatility of (4.1) to be deterministic, it is necessary that

$$\underline{\beta_i} = \underline{0}, \tag{4.4}$$

for $i = 1, \ldots, n$. Secondly, following Dai and Singleton (1998, definition III.1), and in order for the model to be just-identified, the minimal number of restrictions that must be imposed to the parameters is:

$$K_{ij} = 0, \ for \ i < j, \tag{4.5}$$

where K_{ij} is the i^{th}-row j^{th}-column element of matrix K,[7]

$$\underline{\theta} = \underline{0}, \tag{4.6}$$

$$\Sigma = I_n, \tag{4.7}$$

[6] That is, for a strong solution to exist for (4.1).

[7] Matrix K can be either upper or lower triangular. In this Chapter, the second hypothesis will be assumed.

where $I_n \in \Re^{n \times n}$ is an identity matrix, and

$$\underline{\alpha} = \underline{1}, \tag{4.8}$$

where $\underline{\alpha} \in \Re^n$ is a vector with α_i as its i^{th}-component.

In synthesis, the Gaussian term structure model adopted hereafter is completed defined by the short-term interest rate equation (2.2), by the stochastic process

$$d\underline{X}(t) = -K^{\Delta} \cdot \underline{X}(t)\,dt + d\underline{W}^{\mathcal{P}}(t), \tag{4.9}$$

where K^{Δ} equals matrix K subject to restriction (4.5), and by the relation

$$d\underline{W}^{\mathcal{Q}}(t) = \underline{\lambda} dt + d\underline{W}^{\mathcal{P}}(t). \tag{4.10}$$

This is clearly a parsimonious version of the Langetieg (1980) model considered in Chapter three, with only $1 + 2n + \frac{n(n+1)}{2}$ parameters. Nevertheless, it involves the maximal number of parameters that ensures identifiability, and it nests several other Gaussian formulations previously used in the "panel-data" literature: for instance, in appendix 4.8.1 it is shown how to convert the Babbs and Nowman (1999) model into the $\mathbb{A}_0(n)$ canonical formulation.

4.2.2 Pricing of LIBOR-rate derivatives

Next propositions simply summarize the already known Gaussian pricing formulae for the contingent claims that will constitute the panel-data sample used in the empirical analysis of this Chapter, and follow from the analytical solutions derived in Chapter three for the more general deterministic volatility Duffie and Kan (1996) model. Since the $\mathbb{A}_0(n)$ formulation is going to be applied to the valuation of LIBOR-rate derivatives (and not to the pricing of default-free interest rate contingent claims), the Duffie and Singleton (1997) assumption of *symmetric counterparty credit risk* will be implicitly used hereafter.[8]

Proposition 13 *Under the $\mathbb{A}_0(n)$ canonical formulation and assuming that matrix K^{Δ} is non-singular, the time-t price of a (unit face value) pure discount bond with expiry date $T(\geq t)$ is given by*

$$P(\underline{X}(t);\tau) = \exp\left[A(\tau) + \underline{B}'(\tau) \cdot \underline{X}(t)\right], \tag{4.11}$$

where

$$\underline{B}'(\tau) = \underline{G}' \cdot \left(K^{\Delta}\right)^{-1} \cdot \left(e^{-K^{\Delta}\tau} - I_n\right), \tag{4.12}$$

[8] Meaning that the model' short-term interest rate will be regarded not as a riskless rate but rather as a default- and liquidity-adjusted rate.

$$A(\tau) = -\tau f + \underline{G}' \cdot \left(K^{\Delta}\right)^{-1} \cdot \left[\frac{\tau}{2} I_n + \left(e^{-K^{\Delta}\tau} - I_n\right) \cdot \left(K^{\Delta}\right)^{-1} + \frac{1}{2}\Delta(\tau)\right] \cdot \left(\left(K^{\Delta}\right)'\right)^{-1} \cdot \underline{G}, \tag{4.13}$$

$$\Delta(\tau) = \left[K^{\Delta} + \left(K^{\Delta}\right)'\right]^{-1} \cdot \left\{I_n - e^{-\left[K^{\Delta} + \left(K^{\Delta}\right)'\right]\tau}\right\}, \tag{4.14}$$

and $\tau = T - t$ represents the time-to-maturity of the zero-coupon bond.

Proof. Proposition 13 follows from proposition 1 by imposing the parameter's restrictions (4.5) to (4.8), and adopting a straightforward change of notation. ■

Remark 11 *Matrix K^{Δ} is hereafter assumed to be non-singular. This assumption is intuitive and not very severe, because the principal diagonal elements of K^{Δ} are expected to be strictly positive (and, therefore, non-zero).*

Remark 12 *Although $\Delta(\tau)$ involves integrals of matrix exponentials under the general Langetieg (1980) model (see expression (3.4)), equation (4.14) provides a simple integral-free analytical solution as a consequence of restrictions (4.7) and (4.8).*

The time-t *spot swap rate* for an interest rate swap (hereafter labeled by IRS) settled in arrears at times $T_i = t + i\delta, i = 1, \ldots, m$ will be denoted by[9]

$$IRS\left(\underline{X}(t); \delta, m\right) = \frac{1 - P\left(\underline{X}(t); m\delta\right)}{\delta \sum_{i=1}^{m} P\left(\underline{X}(t); i\delta\right)}. \tag{4.15}$$

It is simply the fixed-rate for which the present value of the swap is equal to zero.

Proposition 14 *Under the $\mathbb{A}_0(n)$ canonical formulation, the time-t price of a forward cap on a unitary principal, with a cap rate of k, and settled in arrears at times $T_i = t + i\delta, i = 2, \ldots, m$ is*

$$\begin{aligned} Cap\left(\underline{X}(t); k, \delta, m\right) = & \sum_{i=1}^{m-1} \left\{P\left(\underline{X}(t); i\delta\right) \Phi\left[\sigma_{i\delta,(i+1)\delta} - d(i)\right]\right. \\ & \left. - (1 + \delta k) P\left(\underline{X}(t); (i+1)\delta\right) \Phi\left[-d(i)\right]\right\}, \end{aligned} \tag{4.16}$$

with

$$d(i) = \frac{\ln\left[\frac{P(\underline{X}(t);(i+1)\delta)(1+\delta k)}{P(\underline{X}(t);i\delta)}\right] + \frac{\sigma^2_{i\delta,(i+1)\delta}}{2}}{\sigma_{i\delta,(i+1)\delta}},$$

and

$$\sigma^2_{i\delta,(i+1)\delta} = \underline{B}'(\delta) \cdot \Delta(i\delta) \cdot \underline{B}(\delta).$$

[9] Without loss of generality, and in order to simplify the notation, it will be assumed that $T_i - T_{i-1} = \delta, \forall i$.

Proof. Equation (4.16) results from proposition 3 and formula (3.31), that is follows from valuing the cap as a portfolio of $(m-1)$ European put options (caplets) on zero-coupon bonds. In particular, $\sigma^2_{i\delta,(i+1)\delta}$ represents the time-t variance of the time-T_i log-price of a time-T_{i+1} maturity pure discount bond, and is obtained from (3.25). ■

Remark 13 *The time-t price of the corresponding at-the-money forward cap will be designated by* $ATMCap(\underline{X}(t);\delta,m)$, *and is obtained from (4.16) by defining k as being equal to the market time-t forward swap rate with settlement in arrears at times* $T_i = t + i\delta, i = 2,\ldots,m$.

Proposition 15 *Under the* $\mathbb{A}_0(n)$ *canonical formulation, the time-t price of an European payer swaption maturing at time* $T_u = t + u\delta$, *with a strike equal to x, and on a forward swap with a unitary principal and settled in arrears at times* $T_{u+i} = T_u + i\delta, i = 1,\ldots,m$, *is*

$$\begin{aligned} Payerswpn(\underline{X}(t);x,\delta,u,m) \cong\ & P(\underline{X}(t);u\delta)\,\Phi(-h) \\ & -\sum_{i=1}^{m} k_i P(\underline{X}(t);(u+i)\delta)\,\Phi\left(-h-\sigma_{u\delta,(u+i)\delta}\right), \end{aligned} \tag{4.17}$$

where h is the solution of

$$\sum_{i=1}^{m} k_i P(\underline{X}(t);(u+i)\delta)\exp\left(-\frac{\sigma^2_{u\delta,(u+i)\delta}}{2}-\sigma_{u\delta,(u+i)\delta}h\right) = P(\underline{X}(t);u\delta),$$

$$k_i = 1_{\{i=m\}} + x\delta,$$

and

$$\sigma^2_{u\delta,(u+i)\delta} = \underline{B}'(i\delta)\cdot\Delta(u\delta)\cdot\underline{B}(i\delta).$$

Proof. Equation (4.17) follows from proposition 4 and equation (3.60), i.e. values an European payer swaption as an European put with the same expiry date, with a strike price equal to 1, and on a coupon-bearing bond corresponding to the fixed leg of the underlying interest rate swap, using a *rank 1* approximation. ■

Remark 14 *Instead of proposition 4, one could have priced European swaptions using the lognormal approximation of proposition 5. Bearing in mind the empirical results obtained in subsection 3.4.3, and because the swaption contracts contained in the data sets to be used in the following sections are mainly on short-maturity swaps, it seems, a priori, to be indifferent the type of approximation to consider. Nevertheless, because the price of each swaption will be linearized with respect to the model state-vector (in order to implement the Extended Kalman Filter, in section 4.3), and since the rank 1 approximation involves less severe non-linearities, proposition 4 was the one to be selected.*

Remark 15 *The time-t price of the corresponding at-the-money European swaption will be denominated by $ATMswpn(\underline{X}(t); \delta, u, m)$, and follows from (4.17) by defining x as being equal to the market time-t forward swap rate with settlement in arrears at times $T_{u+i} = T_u + i\delta, i = 1, \ldots, m$.*

4.3 State-space formulation and Kalman filter recursions

The term structure model described in section 4.2 must be rewritten in a state-space form, to allow the explicit treatment of measurement errors. Moreover, in order for the model parameters to be estimated by maximizing the log-likelihood function of the observed market data, the (unobservable) state-variables must be recovered using Kalman filtering techniques. This section presents the model state-space formulation as well as both linear and non-linear Kalman filter algorithms, and is based on Harvey (1989, chapter 3).

4.3.1 Transition equation

Let the panel-data of market observables be composed by the observation of M interest rate contingent claim's market values at N discrete and equally spaced time-periods $t_k\,(k = 1, \ldots, N)$, such that $t_k - t_{k-1} = h, \forall k$.[10] To evolve the vector of state-variables through time (say, from t_{k-1} to t_k), equation (4.9) can be solved explicitly yielding a simple Gaussian VAR(1) process (see, for instance, Lund (1997a, equation 4)):

$$\underline{X}_k = F \cdot \underline{X}_{k-1} + \underline{v}_k, \tag{4.18}$$

where

$$F = e^{-K^{\Delta} h}, \tag{4.19}$$

$$\underline{v}_k = \int_{t_{k-1}}^{t_k} e^{-K^{\Delta}(t_k - s)} \cdot d\underline{W}^{\mathcal{P}}(s),$$

and $\underline{X}_k \equiv \underline{X}(t_k)$. Using Fubini's theorem and Itô's isometry, it follows that

$$\underline{v}_k \backsim N^n(\underline{0}, \Delta(h)). \tag{4.20}$$

4.3.2 Measurement equation

Equations (4.15), (4.16), and (4.17) show that there exists a non-linear relationship between the model' state-variables and the observed swap rates, cap prices, or swaption prices.

[10] The analysis could be easily generalized for unequally spaced time-periods, but at the expense of additional notation.

Therefore, the measurement equations (and, consequently, the Kalman filter algorithms) considered in this Chapter will generally be non-linear. However, when fitting the model (just) to the level of the yield curve (using swap rates), it is possible to transform the observed data (using, for instance, linear interpolation and bootstrap methods) into spot interest rates, with the advantage of providing a linear measurement equation. Both cases are now formalized.

Linear Kalman filter

According to equation (4.11), there exists a linear relationship between spot interest rates and model' factors:

$$R(t, t+\tau) = -\frac{A(\tau)}{\tau} - \frac{B'(\tau)}{\tau} \cdot \underline{X}(t),$$

where $R(t, t+\tau)$ represents the time-t spot interest rate, with continuous compounding, prevailing during τ-years.

Assuming that the observed data is composed by M continuously compounded spot interest rates (for different constant maturities τ_i, $i = 1, \ldots, M$), at each one of the N time-periods, and that the measurement errors are additive and normally distributed, the usual linear measurement equation is obtained:

$$\underline{R}_k = \underline{A} + H' \cdot \underline{X}_k + \underline{\varepsilon}_k, \tag{4.21}$$

with

$$\underline{\varepsilon}_k \backsim N^M\left(\underline{0}, U^D\right), \tag{4.22}$$

where the i^{th} row of $\underline{R}_k \in \Re^M$, $\underline{A} \in \Re^M$, $H' \in \Re^{M \times n}$, and $\underline{\varepsilon}_k \in \Re^M$ is given, respectively, by the market spot rate $R(t_k, t_k + \tau_i)$, by $-\frac{A(\tau_i)}{\tau_i}$, by $-\frac{B'(\tau_i)}{\tau_i}$, and by the time-$t_k$ measurement error for the i^{th} maturity. Concerning the specification of the measurement errors' covariance matrix, the elements of $\underline{\varepsilon}_k$ will be assumed to be cross-sectionally and serially uncorrelated as well as homoskedastic:

$$U^D = \sigma_\varepsilon^2 I_M, \tag{4.23}$$

where $I_M \in \Re^{M \times M}$ is an identity matrix, and $\sigma_\varepsilon^2 \in \Re^+$ is the common variance of all measurement errors. Although it can be argued, as in Geyer and Pichler (1996, page 6), that the variance of the measurement errors should depend on the maturity of the spot rate (and hence that matrix U^D should involve M new parameters, instead of just one), the constant variance assumption will be used for computational reasons: it allows the fitting of higher-dimensional models (larger n) to larger data-sets (larger M), by reducing the number of parameters to be estimated. Similarly, the presence of first order serial correlation amongst the measurement errors is also commonly found in the literature -see, for instance, Brown and Schaefer (1994b, table 5)- and it would require the state-space formulation to be augmented as suggested by Berardi (1997, equations 17 and 18). But, again, such modification would require the additional estimation of a diagonal matrix of autocorrelation coefficients, which would, in turn, increase the computational complexity of the estimation procedure proposed. Therefore, although assumption (4.23) might prove to be unrealistic, it will be used in order to allow the model fit to an enlarged data set (that is, for values of M higher than it is usual in the literature).

In summary, the transition equation (4.18) and the measurement equation (4.21) define a *linear Gaussian* state-space model, where the state-variables can be inferred using the standard linear Kalman filter, and the $2 + 2n + \frac{n(n+1)}{2}$ model' parameters can be estimated through (exact) Maximum Likelihood (ML).

Non-linear Kalman filter

When fitting the model to swap rates, cap prices and/or swaption prices, it is still possible to obtain a measurement equation relating the vector of market observables with the contemporaneous (unobservable) vector of model' state-variables, although such relation is no longer linear.

In general terms, let the observed data be composed, at each one of the N time-periods, by M_1 swap rates, M_2 cap prices, and M_3 European swaption prices, such that

$$M = M_1 + M_2 + M_3.$$

The vector $\underline{R}_k \in \Re^M$ now represents the time-t_k market values of all interest rate contingent claims, i.e.

$$\underline{R}_k = \begin{bmatrix} \underline{R_1}_k \\ \underline{R_2}_k \\ \underline{R_3}_k \end{bmatrix},$$

where the j^{th} element of $\underline{R_1}_k \in \Re^{M_1}$ is the observed time-t_k spot swap rate with m_j^{IRS} resetting periods of δ^{IRS}-years each, the j^{th} element of $\underline{R_2}_k \in \Re^{M_2}$ is the observed time-t_k price of an ATM forward-start cap with m_j^{cap} resetting periods of δ^{cap}-years each, and the j^{th} element of $\underline{R_3}_k \in \Re^{M_3}$ is the observed time-t_k price of an ATM European swaption with u_j^{swpn} periods to maturity (of δ^{swpn}-years each) and on a forward swap with m_j^{swpn} resetting periods (also of δ^{swpn}-years each).[11] Similarly, let $\underline{Z}_k(\underline{X}_k) \in \Re^M$ denote the corresponding model values for the same set of interest rate derivatives, that is

$$\underline{Z}_k(\underline{X}_k) = \begin{bmatrix} \underline{Z_1}_k(\underline{X}_k) \\ \underline{Z_2}_k(\underline{X}_k) \\ \underline{Z_3}_k(\underline{X}_k) \end{bmatrix},$$

where the j^{th} element of $\underline{Z_1}_k(\underline{X}_k) \in \Re^{M_1}$ is given by $IRS\left(\underline{X}_k; \delta^{IRS}, m_j^{IRS}\right)$, the j^{th} element of $\underline{Z_2}_k(\underline{X}_k) \in \Re^{M_2}$ is equal to $ATMCap\left(\underline{X}_k; \delta^{cap}, m_j^{cap}\right)$, and the j^{th} element of $\underline{Z_3}_k(\underline{X}_k) \in \Re^{M_3}$ corresponds to $ATMswpn\left(\underline{X}_k; \delta^{swpn}, u_j^{swpn}, m_j^{swpn}\right)$. The (non-linear) measurement equation is therefore

$$\underline{R}_k = \underline{Z}_k(\underline{X}_k) + \underline{\varepsilon}_k, \tag{4.24}$$

where the vector $\underline{\varepsilon}_k \in \Re^M$ of additive measurement errors still satisfies (4.22), but its covariance matrix is now block-diagonal in order to accommodate specific variances for each market segment:

$$U^D = \begin{bmatrix} \sigma_{\varepsilon_1}^2 I_{M_1} & O_{M_1 \times M_2} & O_{M_1 \times M_3} \\ O_{M_2 \times M_1} & \sigma_{\varepsilon_2}^2 I_{M_2} & O_{M_2 \times M_3} \\ O_{M_3 \times M_1} & O_{M_3 \times M_2} & \sigma_{\varepsilon_3}^2 I_{M_3} \end{bmatrix}, \tag{4.25}$$

with O and I representing, respectively, null and identity matrices, with dimensions given by the corresponding subscripts, and $\sigma_{\varepsilon_1}^2, \sigma_{\varepsilon_2}^2, \sigma_{\varepsilon_3}^2 \in \Re^+$.

The state-space model defined by equations (4.18) and (4.24) is non-linear (in the measurement equation), and involves $4 + 2n + \frac{n(n+1)}{2}$ parameters. Since the use of an exact non-linear filter -along the lines of Kitagawa (1987)- would be numerically infeasible (for the multi-factor model specifications to be considered in this thesis), an approximate non-linear Kalman filter will be used instead and the model' parameters will be estimated by Quasi-Maximum Likelihood (QML).[12] The next question concerns the choice of the "optimal" non-linear approximate filtering method to employ, which, as Jazwinski (1970, page 361) points out, depends on the nature of the specific problem under consideration. For example,[13] Claessens and Pennacchi (1996), in the context of defaultable bond pricing,

[11] For simplicity, the contract specifications δ^{IRS} and m_j^{IRS} ($j = 1, \ldots, M_1$), δ^{cap} and m_j^{cap} ($j = 1, \ldots, M_2$), and δ^{swpn}, u_j^{swpn} and m_j^{swpn} ($j = 1, \ldots, M_3$) are assumed to be constant for all the N time-periods.

[12] Of course, there is a price to be paid by such simplification: as argued by Lund (1997b), the state-variables' filtered estimates can be biased (as well as inefficient) and the parameters' estimators may be inconsistent. Nevertheless, for empirical purposes, the small sample properties of the model parameters' estimators are, perhaps, more relevant, and for this reason a Monte Carlo study is presented at the end of the present section.

[13] See Tanizaki (1996) for alternative non-linear filters.

have used the Extended Kalman Filter (EKF) described in Harvey (1989, section 3.7.2), and Lund (1997b) implemented, for multi-factor Gaussian term structure models, the Iterated Extended Kalman Filter (IEKF) as given in Jazwinski (1970, theorem 8.2). Empirical evidence reported in Jazwinski (1970, chapter 9) suggests that the IEKF should be more effective than the EKF in dealing with significant measurement non-linearities. Moreover, the EKF technique can be thought as being nested into the more general IEKF because, for each time-period (and for each candidate set of model' parameters), the IEKF method "updates" the filtered vector of state-variables through the numerical solution of a generalized least squares (GLS) problem, as in Lund (1997b, equation 16), while the EKF simply involves one iteration of such optimization algorithm. However, this greater generality of the IEKF also makes it more numerically involved, particularly for the high-dimensional models implemented in this dissertation: as the number of model' factors is increased, not only the number of optimizing variables in each one of the N GLS problems is higher, but also, in the author's experience, the number of (Gauss-Newton) iterations required becomes larger. Consequently, and only based on computational-time reasons, the simpler EKF technique will be used, with its finite sample properties tested through a Monte Carlo study, at the end of this section.

The EKF is based on the linearization of the state-space model, and subsequent use of the standard Kalman filter recursions. In the present case, it is only necessary to linearize the measurement equation around $\hat{\underline{X}}_{k|k-1}$, i.e. around the forecast of $\underline{X}_k$ based on the information available at time t_{k-1}. Using a first-order Taylor series expansion for the vector of model contingent claims values,

$$\underline{Z}_k\left(\underline{X}_k\right) \cong \underline{Z}_k\left(\hat{\underline{X}}_{k|k-1}\right) + \hat{Z}'_k \cdot \left(\underline{X}_k - \hat{\underline{X}}_{k|k-1}\right),$$

where

$$\hat{Z}'_k = \left.\frac{\partial \underline{Z}_k\left(\underline{X}_k\right)}{\partial \underline{X}'_k}\right|_{\underline{X}_k = \hat{\underline{X}}_{k|k-1}}, \tag{4.26}$$

equation (4.24) can be approximated by:

$$\underline{R}_k \cong \left[\underline{Z}_k\left(\hat{\underline{X}}_{k|k-1}\right) - \hat{Z}'_k \cdot \hat{\underline{X}}_{k|k-1}\right] + \hat{Z}'_k \cdot \underline{X}_k + \underline{\varepsilon}_k.$$

Note that matrix $\hat{Z}_k$ is easily analytically computed, by differentiating equations (4.15), (4.16), and (4.17).[14]

[14]Formulae summarized in appendix 4.8.2.

4.3.3 Log-likelihood function

Following, for instance, Lund (1997a, section 3) and Claessens and Pennacchi (1996, appendix B), the (exact/quasi) log-likelihood function of the observed market data $(\underline{R}_1, \ldots, \underline{R}_N)$, conditional on the vector of all model parameters $\underline{\Psi} \in \Re^p$,[15] can be obtained, for both (linear/non-linear) Kalman filters, by the following prediction error decomposition:

$$\ln L(\underline{R}_1, \ldots, \underline{R}_N; \underline{\Psi}) = -\frac{MN}{2} \ln(2\pi) - \frac{1}{2} \sum_{k=1}^{N} \left[\ln |\Omega_k| + \underline{\mu}_k' \cdot \Omega_k^{-1} \cdot \underline{\mu}_k \right], \tag{4.27}$$

where the vectors of prediction errors, $\underline{\mu}_k$, and their covariance matrices, Ω_k, are iteratively computed through standard Kalman filter "prediction" and "update" recursions.

In the "prediction" step, the mean of the unobservable state-vector $\underline{X}_k$, conditional on the information available at time t_{k-1},

$$\hat{\underline{X}}_{k|k-1} = F \cdot \hat{\underline{X}}_{k-1}, \tag{4.28}$$

is obtained, as well as its mean square error (MSE) matrix

$$P_{k|k-1} = F \cdot P_{k-1} \cdot F' + \Delta(h). \tag{4.29}$$

Then, in the "update" stage, $\underline{X}_k$ is estimated based on the information available at time t_k,

$$\underline{X}_k = \hat{\underline{X}}_{k|k-1} + P_{k|k-1} \cdot \hat{Z}_k \cdot \Omega_k^{-1} \cdot \underline{\mu}_k, \tag{4.30}$$

and the corresponding MSE matrix is computed:

$$P_k = \left[P_{k|k-1}^{-1} + \hat{Z}_k \cdot \left(U^D\right)^{-1} \cdot \hat{Z}_k' \right]^{-1}. \tag{4.31}$$

In all the formulae, $\hat{Z}_k$ is given by equation (4.26), for the EKF, or simply by H, for the linear case. Similarly, the prediction error vector is

$$\underline{\mu}_k = \underline{R}_k - \left(\underline{A} + H' \cdot \hat{\underline{X}}_{k|k-1} \right), \tag{4.32}$$

for the standard Kalman filter, or

$$\underline{\mu}_k = \underline{R}_k - \underline{Z}_k \left(\hat{\underline{X}}_{k|k-1} \right), \tag{4.33}$$

[15] The vector $\underline{\Psi}$, includes not only the "structural" model parameters f, $\underline{G}$, K^{Δ}, and λ, but also the variance(s) of the measurement errors. Its dimension, p, corresponds to the total number of parameters to be estimated.

for the EKF. Finally, the inverse and determinant of Ω_k can be efficiently obtained through the formulae contained in Harvey (1989, page 108):

$$\Omega_k^{-1} = \left(U^D\right)^{-1} - \left(U^D\right)^{-1} \cdot \hat{Z}_k' \cdot P_k \cdot \hat{Z}_k \cdot \left(U^D\right)^{-1}, \tag{4.34}$$

and[16]

$$|\Omega_k| = \left|U^D\right| \left|P_{k|k-1}\right| \left|P_k^{-1}\right|. \tag{4.35}$$

Assuming that the stochastic process (4.9) is stationary, the above mentioned Kalman filter recursions are initialized at the first two unconditional moments of the state-variables vector:

$$\hat{\underline{X}}_0 = \underline{0}, \tag{4.36}$$

and

$$vec(P_0) = \left(I_{n^2} - F \otimes F\right)^{-1} \cdot vec(\Delta(h)), \tag{4.37}$$

where $I_{n^2} \in \Re^{n^2 \times n^2}$ denotes an identity matrix.

4.3.4 Optimization algorithm

The model parameters $\underline{\Psi}$ are estimated by maximizing the log-likelihood function (4.27), through a quasi-Newton method involving backtracking line searches (following Dennis and Schnabel (1996, section 6.3)). The iteration rule of the optimization process can be represented as

$$\underline{\Psi}^{s+1} = \underline{\Psi}^s + \rho_s \left[Q\left(\underline{\Psi}^s\right) + \gamma_s I_p\right]^{-1} \cdot \underline{g}\left(\underline{\Psi}^s\right), \tag{4.38}$$

where $\underline{\Psi}^s$ represents the vector of parameters' estimates at the s^{th} iteration, $\underline{g}\left(\underline{\Psi}^s\right) \in \Re^p$ is the gradient of the log-likelihood function evaluated at the s^{th} iteration, $Q\left(\underline{\Psi}^s\right) \in \Re^{p\times p}$ is the estimator of the parameters' asymptotic covariance matrix evaluated at $\underline{\Psi}^s$, and $I_p \in \Re^{p\times p}$ is an identity matrix. The step length ρ_s is obtained, at each iteration, by a backtracking line search procedure that ensures global convergence, while γ_s is chosen, at each iteration, in order to guarantee that $\left[Q\left(\underline{\Psi}^s\right) + \gamma_s I_p\right]$ is positive definite.

Following Harvey (1989, equation 3.4.70), the i^{th} element of the gradient vector has the following analytical form:[17]

$$g_i\left(\underline{\Psi}\right) = \sum_{k=1}^{N} \frac{\partial l_k\left(\underline{\Psi}\right)}{\partial \Psi_i},$$

[16] Since U^D is a diagonal matrix, $\left(U^D\right)^{-1}$ and $\left|U^D\right|$ can be further simplified.

[17] For simplicity, the obvious dependence of Ω_k and $\underline{\mu}_k$ on $\underline{\Psi}$ has been suppressed.

with

$$\frac{\partial l_k(\underline{\Psi})}{\partial \Psi_i} = -\frac{1}{2} tr\left[\Omega_k^{-1} \cdot \frac{\partial \Omega_k}{\partial \Psi_i} \cdot \left(I_M - \Omega_k^{-1} \cdot \underline{\mu}_k \cdot \underline{\mu}_k'\right)\right] - \frac{\partial \underline{\mu}_k'}{\partial \Psi_i} \cdot \Omega_k^{-1} \cdot \underline{\mu}_k, \tag{4.39}$$

and where Ψ_i is the i^{th} element of $\underline{\Psi}$. As suggested by Berndt, Hall, Hall and Hausman (1974, page 655), matrix Q is taken to be not the Hessian (with opposite sign) of the log-likelihood function (as in the Newton-Raphson method), but the covariance matrix of $\underline{\Psi}$. Under the exact linear case, the ML estimator $\underline{\Psi}_{ML}$ of $\underline{\Psi}$ is asymptotically normal, with mean $\underline{\Psi}$, and its asymptotic covariance matrix is given by the Fisher's information matrix $IA(\underline{\Psi}_{ML})$. Harvey (1989, equation 3.4.69) provides the following explicit approximation for the i^{th}-row j^{th}-column element of $IA(\underline{\Psi})$:

$$IA_{ij}(\underline{\Psi}) \cong \sum_{k=1}^{N} \left\{ \frac{1}{2} tr\left[\Omega_k^{-1} \cdot \frac{\partial \Omega_k}{\partial \Psi_i} \cdot \Omega_k^{-1} \cdot \frac{\partial \Omega_k}{\partial \Psi_j}\right] + \frac{\partial \underline{\mu}_k'}{\partial \Psi_i} \cdot \Omega_k^{-1} \cdot \frac{\partial \underline{\mu}_k}{\partial \Psi_j} \right\}. \tag{4.40}$$

For the EKF, the QML estimator $\underline{\Psi}_{QML}$ of $\underline{\Psi}$ is still asymptotically normal, approximately consistent, and its asymptotic covariance matrix can be approximated by:[18]

$$IA^{-1}(\underline{\Psi}_{QML}) \cdot J(\underline{\Psi}_{QML}) \cdot IA^{-1}(\underline{\Psi}_{QML}), \tag{4.41}$$

where the i^{th}-row j^{th}-column element of the "outer-product-of-the-gradient" matrix $J(\underline{\Psi})$ is given by

$$J_{ij}(\underline{\Psi}) = \sum_{k=1}^{N} \frac{\partial l_k(\underline{\Psi})}{\partial \Psi_i} \frac{\partial l_k(\underline{\Psi})}{\partial \Psi_j}.$$

In summary, for the standard Kalman filter, $Q(\underline{\Psi}^s) = IA(\underline{\Psi}^s)$, yielding a modified method of scoring, while for the non-linear case, $Q(\underline{\Psi}^s) = IA^{-1}(\underline{\Psi}^s) \cdot J(\underline{\Psi}^s) \cdot IA^{-1}(\underline{\Psi}^s)$.

As in Duan and Simonato (1995, page 11), the stopping criteria is based on a maximum absolute difference smaller than 10^{-4} in both the log-likelihood function value and in the parameters' vector. Finally, note that the system derivatives $\frac{\partial \underline{\mu}_k}{\partial \Psi_i}$ and $\frac{\partial \Omega_k}{\partial \Psi_i}$ can be computed analytically using the recursions derived in Harvey (1989, section 3.4.6). In appendix 4.8.3 these formulae are simply adapted to the EKF.

4.3.5 Monte Carlo study

In order to test the performance of the EKF in finite samples of swap rates, cap prices, and swaption prices, a Monte Carlo study was conducted on the two-factor model estimated, using only spot interest rates, by Babbs and Nowman (1999, Table 2). The true parameter

[18]Ignoring the possible autocorrelation amongst the scores $\frac{\partial l_k(\underline{\Psi})}{\partial \underline{\Psi}}$, as observed by Lund (1997b, page 14).

Table 4.1: Monte Carlo study of a two-factor Babbs and Nowman (1999) model

Parameter	True value	Mean	Std. Error
f	0.072800	0.072815	0.000054
G_1	-0.003950	-0.003821	0.000135
G_2	-0.010206	-0.010195	0.000014
K_{11}	0.552900	0.551486	0.005910
K_{21}	0.743020	0.742189	0.005919
K_{22}	0.065200	0.065534	0.000894
λ_1	-0.084900	-0.087222	0.002841
λ_2	0.096300	0.094165	0.003140
σ_{ε_1}	0.001285	0.001286	0.000025
σ_{ε_2}	0.000072	0.000072	0.000002
σ_{ε_3}	0.000183	0.000186	0.000003

Each simulated data set contains 200 daily cross-sections of:
6 swap rates (for 0.5, 2, 3, 5, 7, and 10 years),
7 ATM forward-start cap prices (for 1, 2, 3, 4, 5, 7, and 10 years),
8 ATM European swaption prices (for 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 0.5×10, 1×4, and 2×8 years).
True values are obtained from Babbs and Nowman (1999, Table 2).
Means and standard errors are for 500 Monte Carlo simulations.

values (second column of table 4.1) were obtained, through appendix 4.8.1, by transforming the two-factor Babbs and Nowman (1999) model into an equivalent $\mathbb{A}_0(2)$ one. In addition, the true standard deviation of swap rates' measurement errors is set equal to the average one estimated by Babbs and Nowman (1999) for continuously compounded spot yields (i.e. 12.85 basis points). For cap and swaption prices, the true standard deviations of measurements errors are arbitrarily set equal to an equivalent spread of 1 b.p. in terms of a one-year forward-start cap (with quarterly compounding) or in terms of a 0.5×2 swaption (with semi-annually compounding), respectively (and assuming a flat continuously compounded yield curve of 5%).[19]

Using such "true" model parameters, the data was simulated and the QML estimator was applied 500 times. For each replication, first the vector of state variables was evolved through time using the Euler discretization of equation (4.9), with 1,000 subdivision per day. Then, for $N = 200$ days,[20] a panel-data of $M_1 = 6$ swap rates (with maturities of 0.5,[21] 2, 3, 5, 7, and 10 years), of $M_2 = 7$ ATM forward-start cap prices (with maturities

[19] That is,

$$\sigma_{\varepsilon_2} = 0.01\% \times 0.25 \times \left(e^{-5\% \times 0.5} + e^{-5\% \times 0.75} + e^{-5\% \times 1}\right),$$

and

$$\sigma_{\varepsilon_3} = 0.01\% \times 0.5 \times \left(e^{-5\% \times 1} + e^{-5\% \times 1.5} + e^{-5\% \times 2} + e^{-5\% \times 2.5}\right).$$

[20] The configuration of the simulated data sets, in terms of the number of cross-sections and contracts/maturities considered, is intended to reproduce the features of the US real data set that will be used in the following sections.

[21] This is simply a 6-month LIBOR spot rate.

of 1, 2, 3, 4, 5, 7, and 10 years), and of $M_3 = 8$ ATM European swaptions (with swaption maturity $\times$ underlying IRS length equal to 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 0.5×10, 1×4, and 2×8 years) was simulated assuming additive, independent, and normally distributed measurement errors with zero mean and common standard deviations for each one of the three blocks of derivatives.[22] Semi-annually compounding was assumed for both IRSs and swaptions ($\delta^{IRS} = \delta^{swpn} = 0.5$), while quarterly compounding was used for caps ($\delta^{cap} = 0.25$). Finally, for each simulated panel-data set, the model' parameters were estimated through QML, using the EKF algorithm.

The last two columns of table 4.1 show the sample mean and the standard error of the parameter's QML estimates obtained from the 500 Monte Carlo simulations. The overall conclusions are that the mean estimates are very close to the corresponding "true" parameters' values, and that the more significant standard errors are, as is commonly found, related to the market price of risk parameters.

4.4 Data description and PCA

The raw data set used for the empirical analysis presented in the following sections consists of Eurodollar rates (for 3, 6, and 12 months), US swap rates (for 2, 3, 5, 7, 10, and 15 years), US ATM forward-start cap flat yield volatilities (for 1, 2, 3, 4, 5, 7, and 10 years), and US ATM European swaption flat yield volatilities (for 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 0.5×10, 1×4, 2×4 years).[23] These are daily average (between bid and ask) quotes, from 21/06/95 to 30/05/96, which yield only $N = 189$ cross-sections of complete market observations, after clearing the data from all the missing data points. Although the data set spans a smaller time-horizon than it is usual in the Kalman filter literature, it should be noticed that the spatial dimension of the sample under consideration is similar as a result of the use of additional derivatives quotes (per cross-section) in the model estimation.[24]

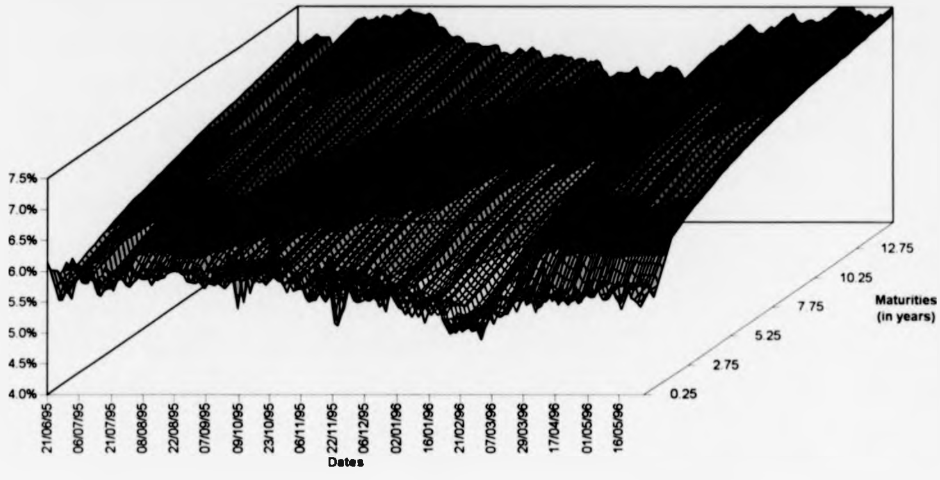
The Eurodollar rates and the swap rates were transformed into discount factors for all the quarterly maturities comprised between 0.25 and 15 years, using linear interpolation of swap rates and bootstrapping of pure discount bond prices. These discount factors are used to compute cap and swaption prices from their quoted flat yield volatilities, and also allow the application of linear Kalman filtering techniques on continuously compounded spot rates. Figure 4-1 shows the daily evolution, between 21/06/95 and 30/05/96, of the

[22] Each independent and univariate normal variate was generated from uniform variates (between 0 and 1) transformed through the Box-Muller algorithm.

[23] Data kindly provided by Martin Cooper, Tokai Bank Europe.

[24] Moreover, the use, in appendix 4.8.5, of a larger data set (containing 784 cross-sections) of UK LIBOR-rate derivatives' quotes does not change the empirical findings to be obtained from the smaller US data set.

Figure 4-1: Term structures of continuously compounded US spot interest rates (from 21/06/95 to 30/05/96, and for quarterly maturities from 0.25 to 15 years)

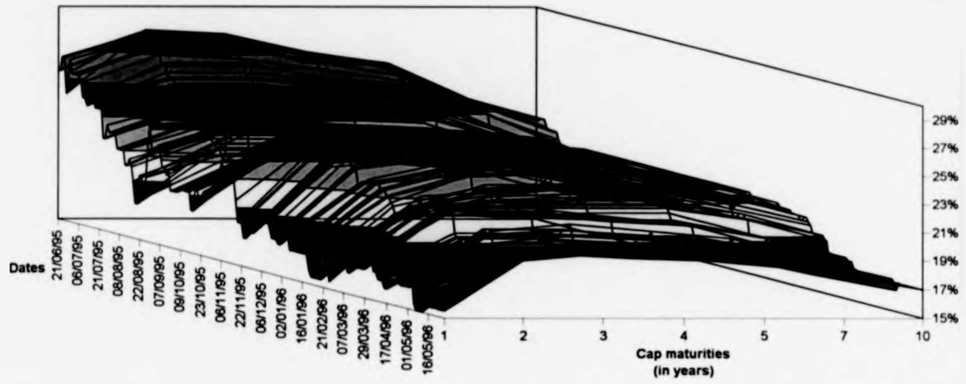


continuously compounded spot interest rate for several maturities. During the period under analysis, the yield curve is predominantly upward sloping, although, for much of the sample, it also presents a negative slope in the money-market maturities.

Figure 4-2 presents the flat yield volatility surface during the same time-period, and for all the cap maturities. Although such flat yield volatilities can be understood as cumulative averages of caplet yield (i.e. "forward forward") volatilities, the usual hump between the one- and the two-years maturities is still evident. Because the measurement equation (4.24) is stated in terms of cap prices (instead of volatilities quotes), all the flat yield volatilities were converted into forward-start cap prices, applying the usual Black formula for each caplet (as in Clewlow and Strickland (1998a, equation 6.15)), and using as strike price the corresponding forward swap rate with quarterly compounding. Similarly, the ATM European swaption flat yield volatility quotes were also converted into option prices, using the "market" assumption of log-normally distributed forward swap rates (as in Rebonato (1998, page 17)).

In order to introduce the empirical work presented in the next sections, a Principal Component Analysis (PCA) was applied to the full time-series (from 21/06/95 to 30/05/96) of continuously compounded spot interest rates for the following maturities: 1, 2, 3, 4, 5, 6, 7, 10, 12, and 15 years. The goal is to identify the state-space dimension of the interest rate

Figure 4-2: US caps' flat yield volatility quotes (from 21/06/95 to 30/05/96, and for all sample cap maturities)



data under consideration, that is to find the minimum number of non-trivial factors needed to reproduce almost all the data variance structure. Then, section 4.5 will test whether the interest rate model under consideration with such pre-specified number of state-variables is able to capture the dynamics of the yield curve, while also pricing accurately caps and swaptions.

The PCA was performed not on the interest rate levels but rather on the daily interest rate changes, since the latter were checked to be stationary. Therefore, the data matrix under use can be represented by

$$\Delta R = \left\{ (\Delta R)_{tj} = \Delta R_j(t) \right\}_{\substack{t=2,\ldots,189 \\ j=1,\ldots,10}},$$

where $\Delta R_j(t)$ denotes the t^{th} observation of the daily change in the spot rate for the j^{th} maturity considered. Then, the eigenvalues and eigenvectors associated to the sample covariance matrix of ΔR were computed, being the eigenvectors scaled to unit length: that is, the eigenvectors (or "loadings") matrix $P \in \Re^{10\times 10}$ is an orthogonal matrix. Finally, each factor (or "principal component") is obtained as a vector of linear combinations between

Table 4.2: Principal Component Analysis of daily US spot rate changes (from 21/06/95 to 30/05/96)

	ΔZ_1	ΔZ_2	ΔZ_3	ΔZ_4	ΔZ_5	ΔZ_6	ΔZ_7
Eigenvalues	4.68E-6	1.09E-6	2.94E-7	6.81E-8	1.60E-8	1.03E-8	5.03E-9
Explained variance	75.9%	17.7%	4.8%	1.1%	0.3%	0.2%	0.1%
Cum. Expl. var.	75.9%	93.6%	98.4%	99.5%	99.8%	99.9%	100.0%
Original variables:				Eigenvectors			
1 year (ΔR_1)	-0.156	0.985	0.070	-0.023	0.014	-0.004	-0.008
2 years (ΔR_2)	-0.344	-0.019	-0.154	0.847	-0.186	0.265	0.185
3 years (ΔR_3)	-0.356	-0.032	-0.489	-0.051	0.645	-0.188	0.073
4 years (ΔR_4)	-0.353	-0.035	-0.360	-0.222	0.055	0.046	0.056
5 years (ΔR_5)	-0.350	-0.040	-0.230	-0.402	-0.543	0.282	0.038
6 years (ΔR_6)	-0.337	-0.053	-0.022	-0.111	-0.317	-0.150	-0.246
7 years (ΔR_7)	-0.324	-0.067	0.193	0.187	-0.089	-0.594	-0.540
10 years (ΔR_8)	-0.298	-0.075	0.418	-0.096	-0.014	-0.290	0.626
12 years (ΔR_9)	-0.297	-0.078	0.411	-0.085	0.125	0.048	0.230
15 years (ΔR_{10})	-0.297	-0.083	0.411	-0.069	0.355	0.593	-0.398

Each column corresponds to one of the first seven principal components.
Eigenvalues and (orthogonal) eigenvectors computed from sample covariance matrix.

the loadings and the original data:

$$\Delta Z = \Delta R \cdot P,$$

where $\Delta Z = \left\{ (\Delta Z)_{tj} = \Delta Z_j (t) \right\}_{\substack{t=2,\dots,189 \\ j=1,\dots,10}}$ is usually called the "scores" matrix, and $\Delta Z_j (t)$ represents the time-t value of the j^{th} principal component.[25]

Table 4.2 presents the first 7 eigenvalues (in a strictly decreasing order) and the corresponding eigenvectors (matrix P) of the sample covariance matrix. In addition, the third line represents the proportion of all the original variables variability that is explained by each principal component (ΔZ_j), and the fourth line shows the cumulative explanatory power of retaining successive factors.[26] As usual, the first principal component (ΔZ_1) explains about three quarters of all the variance in the data (75.9%), and the first three factors, taken together, account for almost all the variation in ΔR (98.4%). Moreover, the analysis of the loading coefficients yields the usual interpretation for the first three principal components: the first one is related to the *level* of the interest rates; the second one represents a *slope* factor; and, the third one accounts for the *curvature* of the yield curve. In summary, this PCA study seems to suggest that two or three factors should be enough to model (almost)

[25] Since $P' \cdot P = I_{10}$, then

$$\Delta R' = P \cdot \Delta Z',$$

which possesses a clear analogy with the HJM modelling approach.

[26] These percentages were computed using two well known facts: the variance of each principal component is equal to the associated eigenvalue; and, the trace of the sample covariance matrix equals the sum of all eigenvalues.

Table 4.3: Estimation of the Dai and Singleton (1998) Gaussian model through EKF and using US swap rates

Parameter	One-factor model		Two-factor model		Three-factor model		Four-factor model	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
f	0.05281	0.10572	0.05345	0.01859	0.03701	0.02669	0.04256	0.03762
G_1	0.01471	0.00060	0.01431	0.00295	0.02731	0.00383	0.04299	0.00581
G_2			0.01102	0.00081	0.01147	0.00154	0.01088	0.00197
G_3					-0.00651	0.00064	-0.00661	0.00904
G_4							-0.00093	0.04881
K_{11}	0.00038	0.00025	1.59903	0.12614	1.52706	0.16018	3.24036	0.25591
K_{21}			1.79234	0.19440	3.78846	0.62220	5.80152	0.97495
K_{31}					0.06564	0.78805	0.96620	33.72856
K_{41}							-4.34072	6.71122
K_{22}			0.06975	0.00813	0.47193	0.03826	0.19189	0.00324
K_{32}					0.04460	0.10355	-0.11524	2.25478
K_{42}							0.31037	0.91186
K_{33}					-0.02231	0.00317	-0.13354	3.28023
K_{43}							0.42981	0.24934
K_{44}							-0.13612	3.25229
λ_1	-0.19969	0.00256	-0.03771	0.48669	-0.07049	0.43438	-0.05570	0.59323
λ_2			-0.35163	0.54710	-0.22626	0.81372	0.30898	1.07913
λ_3					0.10271	0.09630	0.52387	17.56387
λ_4							2.23828	2.48789
σ_{ε_1}	0.00145	0.00005	0.00042	0.00002	0.00022	0.00002	0.00014	0.00002
ln L	7872		9183		9675		9968	
BIC	7862		9162		9641		9965	

Data: US swap rates for 0.5, 2, 3, 5, 7, 10 and 15 years (from 21/06/95 to 30/05/96).

all the uncertainty in the data.

4.5 State-space model estimation

4.5.1 Fitting the yield curve

In the first instance, the non-linear model defined by equations (4.18) and (4.24) was estimated through the EKF presented in subsection 4.3.2, using only swap interest rates (for maturities of 0.5, 2, 3, 5, 7, 10 and 15 years[27]). Table 4.3 presents the QML parameters estimates, their standard errors (computed from (4.41)), the optimal value of the log-likelihood function, and the BIC information criterion for one-, two-, three- and four-factor model specifications. As usual, the less accurately estimated parameters are those related to the market price of risk. Moreover, and as found by Geyer and Pichler (1996, page 9), for models with more than three factors, the standard errors of some QML estimates become too

[27] I.e. $M = M_1 = 7$ and $M_2 = M_3 = 0$.

Table 4.4: Goodness of fit to market US swap rates of the models estimated in table 4.3

IRS	One-factor model		Two-factor model		Three factor model		Four factor model	
maturities:	Mean Error	MAE	Mean Error	MAE	Mean Error	MAE	Mean Error	MAE
6 months	0.0224%	0.2622%	-0.0020%	0.0231%	-0.0006%	0.0042%	-0.0002%	0.0019%
2 years	0.0104%	0.0393%	0.0125%	0.0326%	0.0044%	0.0117%	0.0025%	0.0091%
3 years	0.0161%	0.0489%	0.0039%	0.0249%	0.0011%	0.0102%	-0.0012%	0.0070%
5 years	0.0341%	0.0576%	-0.0165%	0.0217%	0.0086%	0.0120%	0.0057%	0.0101%
7 years	0.0072%	0.0702%	0.0040%	0.0172%	0.0052%	0.0151%	0.0076%	0.0088%
10 years	0.0138%	0.0876%	0.0096%	0.0274%	0.0029%	0.0166%	0.0034%	0.0072%
15 years	0.0060%	0.0768%	-0.0045%	0.0420%	-0.0026%	0.0191%	0.0002%	0.0041%

Errors are differences between model and market swap rates.
(from 21/06/95 to 30/05/96).
MAE are mean absolute errors.

large. Nevertheless, the BIC criterion (as well as the likelihood ratio tests[28], not presented in table 4.3) always rejects the nested models in favor of the higher-dimensional ones.

Table 4.4 shows the mean and mean absolute differences, for each maturity and over the whole sample, between the swap rates generated by the model parameters' estimates presented in table 4.3 and the market quotes. The results clearly show that with only two or three state variables it is already possible to obtain average pricing errors of less than one basis point. Therefore, and as in Babbs and Nowman (1999), it can be said that a low-dimensional (with two or three factors) affine and time-homogeneous Gaussian model can fit extremely well the term structure of interest rates (as predicted by the principal component analysis conducted in section 4.4). In addition, table 4.3 also indicates that higher dimensional specifications yield too unstable parameter estimates, although the fit to the market observables can be marginally improved. Bearing in mind that the PCA conducted before identified two/three non-trivial factors, such parameters' instability can be even understood as a symptom of model over-fitting, since only seven long term rates are being used to estimate the model.

In order to compare the efficiency of the proposed "sub-optimal" non-linear Kalman filter with the standard linear one, the linear Gaussian state-space model defined by equations (4.18) and (4.21) was also estimated using continuously compounded spot yields with the same maturities as the swap rates considered in tables 4.3 and 4.4. Because one-factor models produce too high pricing errors while models with more than three factors seem difficult to identify from the data, table 4.5 only presents two- and three-factor model parameters' estimates and standard errors, obtained using a standard linear Kalman filter

[28]Since all the four Gaussian models are nested into each other, they can also be compared through a likelihood ratio test equal to twice the difference between the log-likelihood function value of the general and restricted model specifications. Such test is asymptotically chi-squared distributed, with degrees of freedom equal to the number of parameter restrictions (under the null hypothesis of valid parameter restrictions).

Table 4.5: Estimation of the Dai and Singleton (1998) Gaussian model using a linear Kalman filter and US spot yields

	Two-factor model		Three-factor model	
Parameter	Estimate	Std. Error	Estimate	Std. Error
f	0.06180	0.04970	0.02221	0.03543
G_1	0.01260	0.00179	0.02143	0.00182
G_2	0.01120	0.00076	0.01232	0.00081
G_3			-0.00395	0.00034
K_{11}	1.70000	0.09220	1.45261	0.08750
K_{21}	2.01000	0.25000	3.50533	0.32327
K_{31}			0.04009	0.43263
K_{22}	0.05670	0.00266	0.27238	0.01549
K_{32}			0.03654	0.04125
K_{33}			-0.07565	0.00163
λ_1	0.04330	0.59100	-0.62949	0.49081
λ_2	-0.22800	0.69100	0.04174	0.94216
λ_3			-1.74463	0.06364
σ_ε	0.00053	0.00001	0.00030	0.00001
ln L	8970		9434	
BIC	8949		9400	

Data: US spot rates for 0.5, 2, 3, 5, 7, 10 and 15 years (from 21/06/95 to 30/05/96).

and ML estimation. Table 4.6 contains the average differences between the model and the market values for the spot interest rates used in the estimation process, as well as for swap rates of identical maturities. It can be observed that, although the model fit to the market yields is very good (average absolute errors lower then 3 b.p. can be obtained using three state variables), the linear Kalman filter consistently overestimates the market swap rates: the average errors for swap rates are much higher than those presented in table 4.4, and positive for all maturities. In other words, the proposed EKF generates a better fit to the observed term structure of interest rates, at least for the finite sample under consideration.

In summary, the empirical analysis developed so far produced three conclusions:

i) Only two or three state variables are required to fit the yield curve, while preserving parameters' stability;

ii) Model specifications with more than three factors generate unstable parameters' estimates;

iii) Although only the standard linear Kalman filter can produce exact asymptotic parameter estimates properties, a non-linear filter seems to perform better in a finite sample of swap rates.

Table 4.6: Goodness of fit to US market yields and swap rates of the models estimated in table 4.5

	Two-factor model		Three-factor model			
	Spot rates	Swap rates	Spot rates		Swap rates	
Maturities:	MAE	MAE	Mean Error	MAE	Mean Error	MAE
6 months	0.0307%	0.1362%	-0.0023%	0.0104%	0.1379%	0.1379%
2 years	0.0383%	0.0707%	0.0127%	0.0188%	0.0588%	0.0591%
3 years	0.0303%	0.0457%	-0.0062%	0.0170%	0.0420%	0.0428%
5 years	0.0299%	0.0327%	-0.0150%	0.0174%	0.0371%	0.0375%
7 years	0.0204%	0.0534%	0.0063%	0.0187%	0.0577%	0.0578%
10 years	0.0353%	0.0720%	0.0069%	0.0284%	0.0623%	0.0625%
15 years	0.0566%	0.0651%	-0.0013%	0.0210%	0.0628%	0.0628%

Errors are differences between model and market rates (from 21/06/95 to 30/05/96). MAE are mean absolute errors.

4.5.2 Fitting cap and swaption prices

Having established that a low-dimensional Gaussian affine and time-homogeneous model is able to reproduce remarkably well the term structure of interest rates, the question is now to test whether such model can also fit the market covariance matrix amongst interest rates.

Firstly, the non-linear state-space model of equations (4.18) and (4.24) was estimated using the same swap rates as in the previous subsection plus cap prices for 1, 2, 3, 4, 5, 7, and 10 years. That is, $M_1 = M_2 = 7$ and $M_3 = 0$. Table 4.7 summarizes the QML parameters' estimates and corresponding standard errors for two- and three-factor model specifications. Comparing with the results previously given in table 4.3, it is possible to verify that the enlargement of the data set produced some significant changes in the parameters' values: the most evident ones are related to the vector λ. Table 4.8 contains the average differences, computed over the all sample and for all the contracts used in the estimation stage, between the market swap rates or cap prices and the corresponding values generated by the model specifications of table 4.7: for swap rates, the pricing errors are given in absolute differences; for cap prices, both percentage pricing errors and absolute percentage differences are shown. The obvious conclusion is that, although it is still possible to fit well the market swap rates, unfortunately the Gaussian affine and time-homogeneous models under consideration are simply unable to reproduce the market prices of caps, mainly for the one-year maturity. In other words, these type of term structure models can not completely accommodate the hump observed in the short-end of the volatility curve (see figure 4-2), although the PCA conducted in section 4.4 had suggested that a two- or a three-factor model would be enough to capture the market interest rate variances.

Figure 4-3 presents the term structure of instantaneous forward rate volatilities generated with the two-factor specification of table 4.7, and compares it with the one that would

Table 4.7: Estimation of the Dai Singleton (1998) Gaussian model through EKF and using US swap rates and cap prices

	Two-factor model		Three-factor model	
Parameter	Estimate	Std. Error	Estimate	Std. Error
f	0.05307	0.00154	0.05461	0.00976
G_1	0.00693	0.00222	0.02086	0.00332
G_2	0.01186	0.00041	0.00763	0.00114
G_3			-0.00732	0.00050
K_{11}	1.44448	0.08162	1.38958	0.13229
K_{21}	1.63235	0.13632	3.60433	0.82152
K_{31}			-1.27773	0.62075
K_{22}	0.07538	0.00494	0.43464	0.04509
K_{32}			-0.07915	0.08683
K_{33}			-0.02597	0.00802
λ_1	-0.00089	0.14208	0.16467	0.44022
λ_2	-0.31355	0.17238	0.03396	1.20631
λ_3			0.11307	0.40702
σ_{ε_1}	0.00053	0.00002	0.00024	0.00001
σ_{ε_2}	0.00215	0.00010	0.00257	0.00014
ln L	16403		16833	
BIC	16382		16799	

Data: US swap rates for 0.5, 2, 3, 5, 7, 10 and 15 years
and cap prices for 1, 2, 3, 4, 5, 7, and 10 years
(from 21/06/95 to 30/05/96).

have been obtained if the model had been estimated using just swap rates (i.e. under the two-factor model parameters of table 4.3). In both cases, the time-t instantaneous variance of the continuously compounded forward rate for maturity at time $T (\geq t)$ is given by

$$\sigma_f^2(t,T) = \underline{G}' \cdot e^{\left[K^{\Delta} + (K^{\Delta})'\right](T-t)} \cdot \underline{G}, \tag{4.42}$$

as can be easily derived by applying Itô's lemma to the forward rate $f(t,T) = -\frac{\partial \ln P(\underline{X}(t), T-t)}{\partial T}$ while using expressions (4.9), (4.11) and (4.12). Notice that although the model volatility curve is still different from the one implied by the market price of caps, figure 4-3 shows that the use of cap prices in the estimation process improved substantially its shape. Hence, even though the market prices of caps can not be correctly reproduced by the model, its use for estimation purposes incorporates additional and valuable information: it can, for instance, help to distinguish the expectations of future rates from the expectations of future volatility, implicit in the same yield curve.

The inability to fit exactly the market prices of caps could simply derive from the fact that the interest rate model under consideration produces time-homogeneous (instantaneous or average) term structures of interest rate or of pure discount bond price volatilities.[29]

[29] As shown by equation (4.42) or by the last equation under proposition 14, for the term structure of

Table 4.8: Goodness of fit to market US swap and cap values of the models estimated in table 4.7

	Two-factor model			Three-factor model		
	Swap rates	Cap prices		Swap rates	Cap prices	
Maturities:	MAE	MPE	MAPE	MAE	MPE	MAPE
6 months	0.0379%			0.0067%		
1 year		37.3521%	44.9284%		35.0458%	43.8762%
2 years	0.0404%	3.0542%	9.0578%	0.0140%	2.4747%	9.4391%
3 years	0.0364%	0.1584%	6.3533%	0.0098%	1.9781%	7.6934%
4 years		0.2068%	4.9100%		2.3089%	6.7546%
5 years	0.0376%	-0.9797%	3.8795%	0.0148%	0.5456%	5.3433%
7 years	0.0357%	0.5612%	3.1039%	0.0169%	0.7624%	4.2350%
10 years	0.0399%	0.2418%	3.1891%	0.0200%	-0.2353%	3.4227%
15 years	0.0361%			0.0155%		

Errors are differences between model and market values.
MAE are mean absolute errors.
MPE are mean percentage errors, i.e. average of errors divided by market values.
MAPE are mean absolute percentage errors.

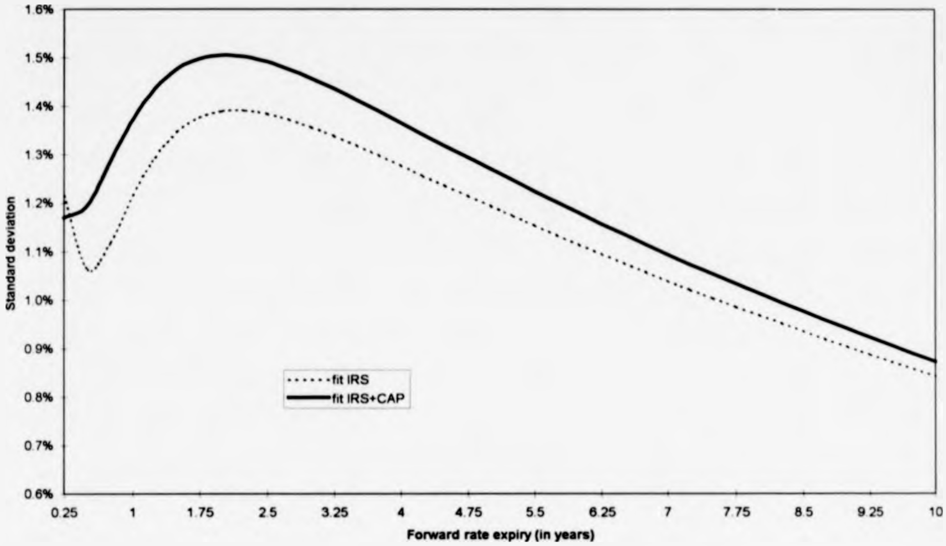
In fact, figure 4-2 clearly shows that the shape of the volatility curve, although similar over time, is not exactly constant. Therefore, in section 4.6 it will be tested whether such modelling insufficiency can be suppressed in the context of a time-inhomogeneous framework.

Since it is not possible to reproduce accurately the principal diagonal elements of the market interest rate covariance matrix (specially for short maturities) using the Gaussian state-space model under analysis, it seems unrealistic to expect a reasonable model fit to the market prices of European swaptions. Nevertheless, two- and three-factor non-linear state-space models were estimated using the same panel-data as in table 4.7 plus ATM European swaption prices for the following (short) option maturities and swap lengths: 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 0.5×10, 1×4, 2×4 years.[30] Table 4.9 summarizes the estimated QML parameter values and standard errors, and table 4.10 presents the average pricing errors with respect to all the contracts used in the estimation of the models, for all the maturities, and over the all sample. Comparing the models' estimates and standard errors with the corresponding ones of table 4.7, two comments can be made: again, the use of additional derivatives prices produced important changes in the parameters' estimates; and, this new enlargement of the data set improved the standard errors of the estimates.

In terms of pricing accuracy, and as predicted before, table 4.10 shows that the Gaussian time-homogeneous model under analysis can not be successfully fitted to the market swaption prices (even using three state variables, i.e. the specification with highest BIC). One

instantaneous forward rate volatilities or of pure discount bond price average volatilities, respectively.
[30] Hence $M_1 = M_2 = 7$ and $M_3 = 8$.

Figure 4-3: Term structures of instantaneous forward rate volatilities (for two-factor models fitted only to US swap rates or also to US cap prices)



possible explanation for this model inability, as put forward by Cooper and Rebonato (1995), is the intrinsic limitation of low-dimensional models to capture the empirically observed fast *decorrelation phenomena* among interest rates of adjacent maturities: because the model interest rate correlations are higher than the ones implicit in the (positive) difference between cap and swaption volatility quotes, swaptions tend to be overpriced. Cooper and Rebonato (1995, appendix 1) justified the last statement by showing that, for a "simple PC-based two-factor model", the slope of the correlation function between changes in instantaneous forward rates tends to zero as the difference in their maturities goes to zero: i.e., the correlation function tends to be flat at the origin, and thus possesses the wrong sigmoid shape, instead of an exponential-decaying one. However, it is shown in appendix 4.8.4 that for the general n-dimensional model specification

$$df(t,T) = \alpha(t,T)\,dt + \underline{\gamma}'(t,T)\cdot d\underline{W}^{Q}(t), \tag{4.43}$$

where $\alpha(t,T)$ is the drift of the forward rate process[31], and $\underline{\gamma}(t,T) \in \Re^n$ is a possibly

[31] Its functional form is irrelevant for the present discussion, but can always be obtained through the usual HJM no-arbitrage condition.

Table 4.9: Estimation of the Dai and Singleton (1998) Gaussian model through EKF and using US swap rates, cap prices and European swaption prices

	Two-factor model		Three-factor model	
Parameter	Estimate	Std. Error	Estimate	Std. Error
f	0.05186	0.00019	0.05589	0.00024
G_1	0.00215	0.00157	0.00580	0.00138
G_2	0.01294	0.00039	0.00383	0.00091
G_3			-0.00458	0.00081
K_{11}	0.76810	0.07229	1.46678	0.10521
K_{21}	0.98134	0.06380	2.63609	0.27048
K_{31}			-6.40563	1.34490
K_{22}	0.11809	0.01103	0.30821	0.02604
K_{32}			-0.82420	0.17999
K_{33}			0.00606	0.00202
λ_1	0.01522	0.02918	0.01438	0.02375
λ_2	-0.33874	0.02860	0.10440	0.03286
λ_3			0.12661	0.10153
σ_{ε_1}	0.00099	0.00004	0.00109	0.00005
σ_{ε_2}	0.00116	0.00004	0.00056	0.00003
σ_{ε_3}	0.00898	0.00055	0.00878	0.00054
ln L	22779		23398	
BIC	22758		23364	

Data: US swap rates for 0.5, 2, 3, 5, 7, 10 and 15 years, cap prices for 1, 2, 3, 4, 5, 7, and 10 years, and swaption prices for 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 0.5×10, 1×4, 2×4 years (from 21/06/95 to 30/05/96).

time-inhomogeneous but state-independent diffusion term, it is always the case that:

$$\left.\frac{\partial \rho(T_1, T_2)}{\partial T_2}\right|_{T_2 = T_1} = 0, \tag{4.44}$$

being $\rho(T_1, T_2)$ the correlation between the changes of the instantaneous forward rates with maturities T_1 and $T_2 (\geq T_1)$. Consequently, the flat slope at the short-end of the correlation function (and thus, its sigmoid shape) is not an exclusive of low-dimensional models, but rather a common feature of all the interest rate models represented by equation (4.43), no matter their number of state variables. This fact, and the already observed instability of some parameters' estimates for high-dimensional model specifications, suggest that the accuracy of European swaption pricing shall not be significantly improved by simply increasing the number of model' factors.

Figure 4-4 presents the correlation function between the short rate and instantaneous forward rates of different maturities for all the model specifications shown in table 4.3 (and computed from equation (4.70) - appendix 4.8.4). It can be observed that by increasing the number of model' factors, not only the short-end sigmoid shape did not disappeared, but also the long-term behavior of the correlation function became implausible. Similarly,

Table 4.10: Goodness of fit to market US swap, cap and swaption values of the models estimated in table 4.9

	Two-factor model						
	Swap rates		Cap prices		Swaption prices		
Maturity	Mean Error	MAE	MPE	MAPE	Maturity	MPE	MAPE
0.5 years	0.005%	0.093%			0.5x2	13.096%	39.399%
1 year			40.259%	45.563%	0.5x3	15.165%	41.747%
2 years	-0.008%	0.056%	3.761%	8.068%	0.5x4	16.905%	43.625%
3 years	-0.024%	0.071%	-0.027%	4.344%	0.5x5	18.501%	45.284%
4 years			0.036%	2.349%	0.5x7	21.602%	48.137%
5 years	-0.029%	0.091%	-0.968%	1.922%	0.5x10	16.417%	45.704%
7 years	-0.003%	0.089%	0.691%	0.913%	1x4	14.462%	35.709%
10 years	0.001%	0.081%	-0.085%	1.542%	2x4	19.932%	32.565%
15 years	-0.025%	0.066%					

	Three-factor model						
	Swap rates		Cap prices		Swaption prices		
Maturity	Mean Error	MAE	MPE	MAPE	Maturity	MPE	MAPE
0.5 years	0.008%	0.067%			0.5x2	26.104%	50.802%
1 year			12.834%	28.571%	0.5x3	22.574%	48.314%
2 years	0.003%	0.066%	-1.087%	3.613%	0.5x4	17.070%	44.830%
3 years	-0.019%	0.079%	0.581%	1.427%	0.5x5	12.989%	42.886%
4 years			1.185%	1.222%	0.5x7	10.968%	44.225%
5 years	-0.039%	0.101%	-0.487%	0.750%	0.5x10	9.468%	46.013%
7 years	-0.021%	0.104%	0.119%	0.465%	1x4	13.121%	35.327%
10 years	-0.009%	0.099%	0.092%	0.352%	2x4	13.652%	28.034%
15 years	-0.012%	0.085%					

Errors are differences between model and market values.
MAE are mean absolute errors.
MPE (MAPE) are mean (absolute) percentage errors.

figure 4-5 describes the same correlation function but for the two-factor model specifications estimated in tables 4.3, 4.7, and 4.9. It can be noticed how the enlargement of the data set used in the parameters' estimation process effectively changes the level (but not the shape) of the correlation function, which, for maturities larger than one year, already possesses the right exponential-decaying form.

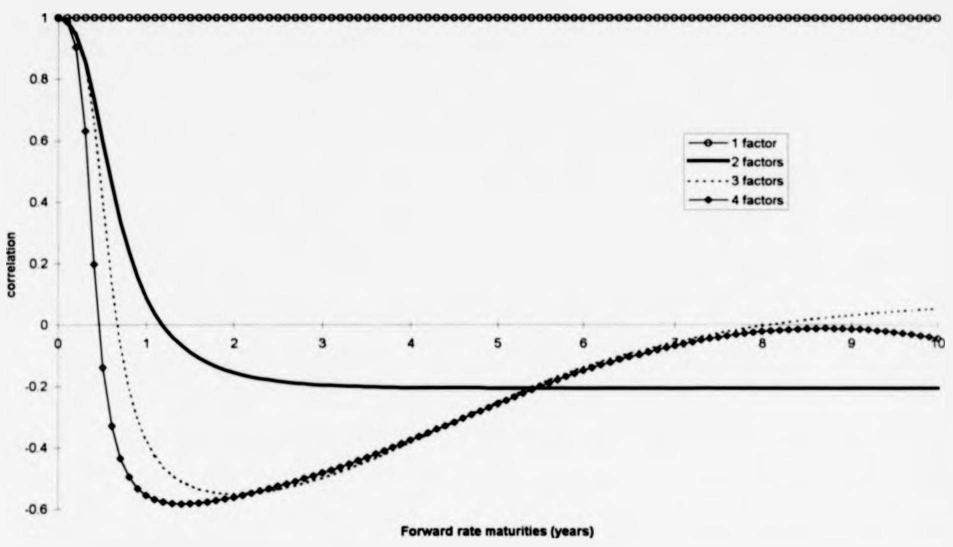
In summary, the empirical evidence presented in this subsection suggests that:

i) The Gaussian time-homogeneous and affine framework under consideration is incapable of fitting cap prices for short maturities;

ii) Swaptions can not also be correctly priced, even by increasing the dimension of the model;

iii) Although it is not possible to exactly recover both cap and swaption prices, it is still important to consider these derivatives during the estimation stage, because valuable additional information, concerning the interest rates covariance structure, is incorporated into the model (as indicated by figures 4-3 and 4-5).

Figure 4-4: Correlation between the short rate and instantaneous forward rates of different maturities (for one-, two-, three-, and four-factor models fitted only to US swap rates)

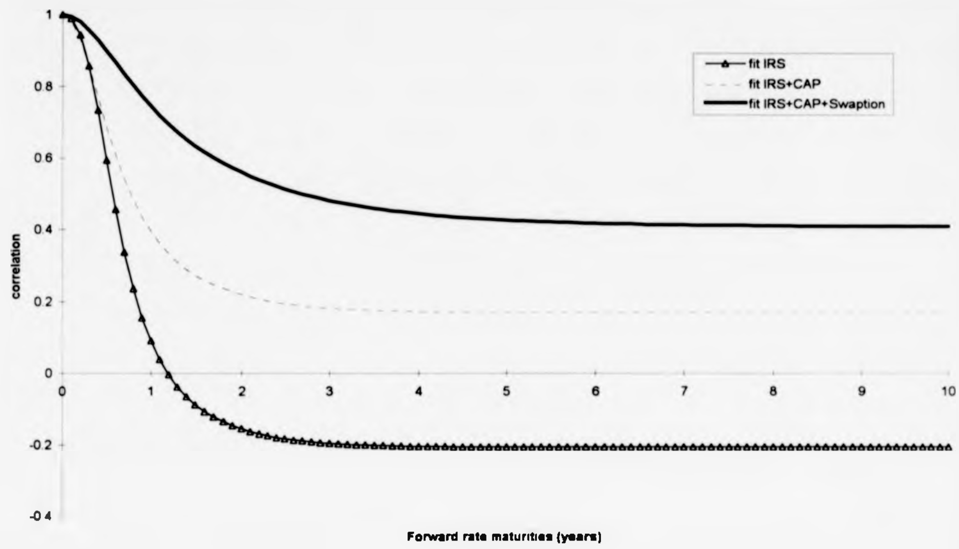


It is well known that swaption prices depend both on the diagonal and off-diagonal elements of the interest rates covariance matrix. On the other hand, the analytical and empirical results presented in this subsection suggest that there is little room to improve the fit to the off-diagonal elements, in the context of the Gaussian time-homogeneous models under analysis. Bearing also in mind that the *rank 1* approximation of proposition 15 only involves forward rates variances, and that imperfect terminal forward rates correlation can be enhanced simply by using time-inhomogeneous forward rates instantaneous volatilities (see Rebonato (1998, page 81)), next section will test whether the model fit to swaption prices can be improved only by forcing an (almost) exact fit to cap prices (i.e., by recovering just the principal diagonal elements of the interest rates covariance matrix).

4.6 The time-homogeneity assumption

Section 4.5 has shown that multifactor Gaussian exponential-affine and time-homogeneous models can fit remarkably well (with just two or three factors) the yield curve but can not reproduce accurately both cap and swaption prices. Three possible explanations can be put forward for these results. Firstly, it can be the case that the tractable exponential-affine class under analysis is simply not rich enough to fit the complex interest rates covariance

Figure 4-5: Correlation between the short rate and instantaneous forward rates of different maturities (for two-factor models fitted only to US swap rates or also to US cap prices or also to US swaption prices)



matrix observed in the market. Secondly, it is also possible that an affine stochastic volatility model could do a better job than the simple Gaussian one under consideration. Finally, perhaps a better fit (at least) to the principal diagonal elements of the market covariance matrix can be achieved, even staying inside the Gaussian exponential-affine class, if the model could simply generate a time-inhomogeneous term structure of volatilities.

The last and simpler hypothesis will be tested in the present section, while the other two will await further research. For this purpose, this section considers a Gauss-Markov HJM model that is equivalent to the "equilibrium" specification defined by equations (2.2), (4.9) and (4.10),[32] except that now the diffusion term $\underline{\gamma}(t,T)$ possesses both a time-homogeneous and a time-dependent term. The cast of the "equilibrium" model into an HJM framework does not represent, by itself, a major improvement because the state-space model was shown (in section 4.5) to already fit well the term structure of interest rates. However, the time-inhomogeneous diffusion component is now intended to allow the shape of the term structure of volatilities to change (slightly) over time. Nevertheless, because the time-homogeneous component is still expected to retain the most explanatory power, such time-inhomogeneous Gaussian HJM model will be estimated in two stages: first, the time-independent parameters

[32] Of course, if and only if the initial forward rate curve can be reproduced by the "equilibrium" model.

are estimated under the state-space formulation by applying a non-linear Kalman filter to an historical panel-data of swap rates, cap prices and swaption prices; then, the time-dependent diffusion term is used to better fit the HJM model to the current cross-section of cap and swaption prices, i.e. to *account* for the covariance structure that can not be fully *explained* by the time-homogeneous part. This estimation methodology is in the spirit of the one employed by Scott (1995) in the context of a multifactor Cox et al. (1985b) model fitted only to the term structure of interest rates.

4.6.1 Equivalent time-inhomogeneous Gauss-Markov HJM formulation

In terms of pure discount bond prices, the HJM model that will be used hereafter is the same as given in equation (4.67), but with an additional time-dependent function $h : \Re^{+} \rightarrow \Re$ in the diffusion term:

$$\frac{dP(t,T)}{P(t,T)} = r(t)\,dt + \underline{v}'(t,T) \cdot d\underline{W}^{Q}(t), \tag{4.45}$$

where

$$\underline{v}(t,T) = h(t)\,\underline{B}(\tau).$$

The duration vector $\underline{B}(\tau)$ is still given by expression (4.12), but the time-t price of the T-maturity zero coupon bond, $P(t,T)$, is no longer time-homogeneous (although still log-normally distributed). Notice that for a well behaved function $h(t)$, $\underline{B}(0) = \underline{0}$ implies that $\underline{v}(t,t) = \underline{0}$, which is consistent with the "pull-to-par" phenomena. Moreover, the diffusion term of equation (4.45) possesses the separable form that, according to Carverhill (1994), ensures the Markovian nature of the model.[33]

In order to define the HJM model in terms of instantaneous forward interest rates, Itô's lemma can be applied to $\ln P(t,T)$, yielding

$$\begin{aligned}\ln P(t,T) &= \ln P(0,T) + \int_0^t \left[r(u) - \frac{h^2(u)}{2}\underline{B}'(T-u)\cdot\underline{B}(T-u)\right]du \\ &\quad + \int_0^t h(u)\,\underline{B}'(T-u)\cdot d\underline{W}^{Q}(u).\end{aligned}$$

Then, using the relation $f(t,T) = -\frac{\partial \ln P(\underline{X}(t);\tau)}{\partial T}$,

$$f(t,T) = f(0,T) + \int_0^t h^2(u)\frac{\partial}{\partial T}\underline{B}'(T-u)\cdot\underline{B}(T-u)\,du - \int_0^t h(u)\frac{\partial}{\partial T}\underline{B}'(T-u)\cdot d\underline{W}^{Q}(u).$$

[33] In fact, it would be a simple exercise to show that the pure discount bond price process (4.45) is consistent with a Markovian model specification given by equation (2.2), and by the following modification of equations (4.9) and (4.10):

$$d\underline{X}(t) = -\left[\underline{\lambda} + K^{\Delta}\cdot\underline{X}(t)\right]dt + h(t)\,d\underline{W}^{Q}(t).$$

That is, the time-inhomogeneous Gaussian HJM model is defined by equation (4.43), with

$$\underline{\gamma}(t,T) = -h(t)\frac{\partial}{\partial T}\underline{B}(\tau), \tag{4.46}$$

and

$$\alpha(t,T) = h^2(t)\frac{\partial}{\partial T}\underline{B}'(\tau)\cdot\underline{B}(\tau). \tag{4.47}$$

Of course, for the model to be arbitrage-free it is still necessary to show that the no-arbitrage condition (4.69) is satisfied. But this is clearly the case because

$$\begin{aligned} -h(t)\,\underline{B}(\tau) &= -\int_t^T h(t)\frac{\partial}{\partial s}\underline{B}(s-t)\,ds \\ &= \int_t^T \underline{\gamma}(t,s)\,ds. \end{aligned}$$

Comparing equations (4.68) and (4.46), it follows that if

i) $h(t) = 1, \forall t$; and if

ii) the initial forward rate curve corresponds to

$$f(0,T) = -\frac{\partial}{\partial T}A(T) - \frac{\partial}{\partial T}\underline{B}'(T)\cdot\underline{X}(0),$$

then the HJM model under analysis is exactly equivalent to the "equilibrium" one defined by equations (2.2), (4.9) and (4.10).

4.6.2 Calibration to cap and European swaption prices

Concerning the pricing of caps, equation (4.16) is still valid but with two differences: the pure discount bond prices are taken as given from the market (as a result of the HJM nature of the model); and, the volatility of the terminal log-price of each underlying zero coupon bond is now time-inhomogeneous (because $\underline{v}(t,T)$ is also time-inhomogeneous).

Proposition 16 *Under the HJM formulation of equation (4.46), the time-0 price of a forward cap on a unitary principal, with a cap rate of k, and settled in arrears at times $i\delta^{cap}, i = 2, \ldots, m^{cap}$ is*

$$\begin{aligned} Cap_0\left(k,\delta^{cap},m^{cap}\right) &= \sum_{i=1}^{m^{cap}-1}\left\{P\left(0,i\delta^{cap}\right)\Phi\left[\sigma_{i\delta^{cap},(i+1)\delta^{cap}} - d(i)\right]\right. \\ &\quad \left. -\left(1+\delta^{cap}k\right)P\left(0,(i+1)\delta^{cap}\right)\Phi\left[-d(i)\right]\right\}, \end{aligned} \tag{4.48}$$

where

$$d(i) = \frac{\ln\left[\frac{P(0,(i+1)\delta^{cap})(1+\delta^{cap}k)}{P(0,i\delta^{cap})}\right] + \frac{\sigma^2_{i\delta^{cap},(i+1)\delta^{cap}}}{2}}{\sigma_{i\delta^{cap},(i+1)\delta^{cap}}},$$

$P(0, i\delta^{cap}), i = 1, \ldots, m^{cap}$, *are obtained from market data, and*

$$\sigma^2_{i\delta^{cap},(i+1)\delta^{cap}} = \int_0^{i\delta^{cap}} h^2(v) \left\|\underline{B}((i+1)\delta^{cap} - v) - \underline{B}(i\delta^{cap} - v)\right\|^2 dv, \qquad (4.49)$$

with $\|\cdot\|$ *denoting the Euclidean norm in* $\Re^n$.

Proof. Equation (4.48) follows immediately from (4.16). The functional form of the volatility term

$$\sigma^2_{i\delta^{cap},(i+1)\delta^{cap}} = VAR[\ln P(i\delta^{cap}, (i+1)\delta^{cap}) | \mathcal{F}_0]$$

is easily derived from (4.45) by applying Itô's lemma to the log-forward price $\ln \frac{P(t,(i+1)\delta^{cap})}{P(t,i\delta^{cap})}$. ■

Similarly, the price of an European payer swaption can still be obtained from a modified version of equation (4.17), which is consistent with the observed yield curve and produces time-dependent volatilities.

Proposition 17 *Under the HJM formulation of equation (4.46), the time-0 price of an European payer swaption maturing at time* $u^{swpn}\delta^{swpn}$, *with a strike equal to* x, *and on a forward swap with a unitary principal and settled in arrears at times* $(u^{swpn} + i)\delta^{swpn}, i = 1, \ldots, m^{swpn}$, *is*

$$\begin{aligned} & Payerswpn_0(x, \delta^{swpn}, u^{swpn}, m^{swpn}) & (4.50)\\ \cong\ & P(0, u^{swpn}\delta^{swpn}) \Phi(-h) \\ & - \sum_{i=1}^{m^{swpn}} k_i P(0, (u^{swpn} + i)\delta^{swpn}) \Phi\left(-h - \sigma_{u^{swpn}\delta^{swpn},(u^{swpn}+i)\delta^{swpn}}\right), \end{aligned}$$

where h *is the solution of*

$$\begin{aligned} P(0, u^{swpn}\delta^{swpn}) = & \sum_{i=1}^{m^{swpn}} k_i P(0, (u^{swpn} + i)\delta^{swpn}) \\ & \exp\left(-\frac{\sigma^2_{u^{swpn}\delta^{swpn},(u^{swpn}+i)\delta^{swpn}}}{2} - \sigma_{u^{swpn}\delta^{swpn},(u^{swpn}+i)\delta^{swpn}} h\right), \end{aligned}$$

$$k_i = 1_{\{i=m^{swpn}\}} + x\delta^{swpn},$$

$P(0, (u^{swpn} + i)\delta^{swpn}), i = 0, \ldots, m^{swpn}$, *are obtained from market data, and*

$$\sigma^2_{u^{swpn}\delta^{swpn},(u^{swpn}+i)\delta^{swpn}} \qquad (4.51)$$

$$= \int_0^{u^{swpn}\delta^{swpn}} h^2(v) \left\| \underline{B}\left((u^{swpn}+i)\,\delta^{swpn}-v\right) - \underline{B}\left(u^{swpn}\delta^{swpn}-v\right)\right\|^2 dv.$$

Proof. Equation (4.50) follows from (4.17), while attending to the HJM nature of the model under analysis. Equation (4.51), that is the functional form of

$$\sigma^2_{u^{swpn}\delta^{swpn},(u^{swpn}+i)\delta^{swpn}} = VAR\left[\ln P\left(u^{swpn}\delta^{swpn},(u^{swpn}+i)\,\delta^{swpn}\right)\middle|\,\mathcal{F}_0\right],$$

is obtained from (4.45). ■

For computational purposes, $h(t)$ will hereafter be assumed to be a piecewise constant function:

$$h(t) = h_j \Leftarrow l_{j-1}\delta^{cap} \le t < l_j\delta^{cap},\ for\ j = 1,\ldots,\theta, \tag{4.52}$$

where l_j are breakpoints defined in terms of an integer number of cap compounding periods (δ^{cap}), with $l_0 = 0$, and θ is the number of possible values (h_j) for function $h(t)$. Under this simple specification of $h(t)$, equations (4.49) and (4.51) can now be solved explicitly:

Proposition 18 *Under assumption (4.52), with $\alpha,\beta \in \Re$, and as long as $l_1\delta^{cap} \le \alpha$,*[34]

$$\begin{aligned}\sigma^2_{\alpha,\alpha+\beta} = {} & \sum_{s=1}^{j} h_s^2 \underline{B}'(\beta)\cdot e^{-K^{\Delta}(\alpha-l_s\delta^{cap})}\cdot\Delta\left((l_s-l_{s-1})\,\delta^{cap}\right)\cdot e^{-\left(K^{\Delta}\right)'(\alpha-l_s\delta^{cap})}\cdot\underline{B}(\beta)\\ & + h_{j+1}^2\underline{B}'(\beta)\cdot\Delta\left(\alpha-l_j\delta^{cap}\right)\cdot\underline{B}(\beta),\end{aligned} \tag{4.53}$$

where

$$j:\ \sup_{j=1,\ldots,\theta-1}\left\{l_j : l_j\delta^{cap} \le \alpha\right\}.$$

Proof. This result follows from (4.49) or (4.51) by using the relation

$$\Delta(T-t) = \int_t^T e^{-\left[K^{\Delta}+\left(K^{\Delta}\right)'\right](T-s)}ds,$$

and straightforward algebra. ■

Taken as given the time-homogeneous model' parameters (which are obtained from the estimation of the equilibrium model specification), the HJM time-inhomogeneous model can now be fitted cross-sectionally to only cap prices or to both cap and European swaption prices, by minimizing the sum of squared percentage differences between model and market prices, with respect to h_j ($j = 1,\ldots,\theta$). For this purpose, it was used the same quasi-Newton optimization method, with backtracking line search, as in subsection 4.3.4 (but now, both gradient and Hessian are computed numerically through finite differences).

[34] *Because, in the sample, the first caplet maturity is $\delta^{cap} = 0.25$ years, this restriction can be rewritten as $l_1 \le 1$.*

Table 4.11: Calibration of a two-factor time-inhomogeneous HJM model to US cap prices, using the time-homogeneous coefficients of table 4.7

		Estimated $h(t)$		Fit to cap prices		
j	l_j	Mean(h_j)	Stdev(h_j)	Maturities	MPE	MAPE
1	1	0.903201	0.162100	1 year	-0.0002%	0.0006%
2	2	0.379732	0.360098	2 years	0.0005%	0.0056%
3	4	1.282043	0.127453	3 years	0.0022%	0.0155%
4	6	0.748490	0.242804	4 years	-0.0100%	0.0229%
5	8	1.176523	0.087393	5 years	0.0111%	0.0247%
6	10	0.265532	0.262661	7 years	-0.0040%	0.0155%
7	12	1.105435	0.109279	10 years	0.0003%	0.0051%
8	16	1.042507	0.139551			
9	20	1.134969	0.123257			
10	28	0.437323	0.208148			
11	36	1.573459	0.111231			
12	44	1.222173	0.125760			

Data: cap prices for 1, 2, 3, 4, 5, 7, and 10 years (from 21/06/95 to 30/05/96).
l_j is the j^{th} breakpoint of function $h(t)$.
h_j is the value of $h(t)$ for the time-interval $[l_{j-1}\delta^{cap}, l_j\delta^{cap}[$.
MPE (MAPE) are mean (absolute) percentage errors.

Initially, the two-factor equilibrium specification estimated in table 4.7 was converted into the HJM time-inhomogeneous model given by equations (4.43), (4.46) and (4.47), and fitted cross-sectionally just to same set of cap prices as used in table 4.7. The breakpoints of function $h(t)$ were defined exactly in the same way as in Brace et al. (1997, table 4.3): that is $\theta = 12$, $l_1 = 1$, $l_2 = 2$, $l_3 = 4$, $l_4 = 6$, $l_5 = 8$, $l_6 = 10$, $l_7 = 12$, $l_8 = 16$, $l_9 = 20$, $l_{10} = 28$, $l_{11} = 36$, and $l_{12} = 44$. Table 4.11 shows the mean and standard errors of the 189 cross-sectionally estimates of h_j ($j = 1, \ldots, 12$), as well as the quality of the model fit to the market cap prices: an almost exact fit to the cap prices (with absolute average percentage pricing errors lower then 3 b.p.) can be achieved. Note that this result could even be further improved if function $h(t)$ was defined to be piecewise constant on each δ^{cap} interval: i.e., $l_j = j\delta^{cap}$ (for $j = 1, \ldots, 40$). In fact, assuming that caplet prices, or equivalently caplet price volatilities, are observable in δ^{cap} time-intervals, then equation (4.53) implies that h_i can be found by exactly fitting the volatility $\sigma_{i\delta^{cap},(i+1)\delta^{cap}}$ (after having obtained h_s by matching $\sigma_{s\delta^{cap},(s+1)\delta^{cap}}$, $s = 1, \ldots, i-1$):

$$\begin{aligned}\sigma^2_{i\delta^{cap},(i+1)\delta^{cap}} &= \sum_{s=1}^{i-1} h_s^2 \underline{B}'\left(\delta^{cap}\right) \cdot e^{-K^{\Delta}\left(i\delta^{cap}-s\delta^{cap}\right)} \cdot \Delta\left(\delta^{cap}\right) \cdot e^{-\left(K^{\Delta}\right)'\left(i\delta^{cap}-s\delta^{cap}\right)} \cdot \underline{B}\left(\delta^{cap}\right) \\ &\quad + h_i^2 \underline{B}'\left(\delta^{cap}\right) \cdot \Delta\left(\delta^{cap}\right) \cdot \underline{B}\left(\delta^{cap}\right).\end{aligned}$$

In order to test whether caps and swaptions can be priced simultaneous and consistently just by introducing time-dependencies into the forward interest rate variances, the two-

Table 4.12: Calibration of a two-factor time-inhomogeneous HJM model to US cap and swaption prices, using the time-homogeneous coefficients of table 4.9

		Estimated $h(t)$		Fit to cap prices		Fit to swaption prices		
j	l_j	Mean(h_j)	Stdev(h_j)	Maturity	MAPE	Maturity	MPE	MAPE
1	1	0.661332	0.479054	1 year	17.532%	0.5x2	4.234%	12.342%
2	2	0.776754	0.778324	2 years	15.200%	0.5x3	7.488%	11.094%
3	4	0.655437	0.300730	3 years	7.143%	0.5x4	9.380%	11.229%
4	6	0.758925	0.560693	4 years	2.169%	0.5x5	9.817%	11.488%
5	8	0.315750	0.593199	5 years	1.367%	0.5x7	9.473%	11.972%
6	10	1.840157	1.089353	7 years	0.995%	0.5x10	3.875%	16.688%
7	12	0.015941	0.111595	10 years	0.079%	1x4	0.184%	17.159%
8	16	0.036696	0.153154			2x4	1.180%	8.279%
9	20	0.708231	0.687708					
10	28	0.754327	0.684637					
11	36	0.738773	0.511101					
12	44	2.045992	1.701756					

Data: cap prices for 1, 2, 3, 4, 5, 7, and 10 years, and swaption prices for 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 0.5×10, 1×4, 2×4 years (from 21/06/95 to 30/05/96).
l_j is the j^{th} breakpoint of function $h(t)$.
h_j is the value of $h(t)$ for the time-interval $[l_{j-1}\delta^{cap}, l_j\delta^{cap}[$.
MPE (MAPE) are mean (absolute) percentage errors.

factor equilibrium specification estimated in table 4.9 was transformed into the arbitrage-free model of equations (4.43), (4.46) and (4.47). This new HJM model was fitted cross-sectionally to the same cap and swaption prices as considered in table 4.9, and using the same time-discretization of function $h(t)$ as before. The results are shown in table 4.12, and three conclusions can be made. First, although the pricing errors for European swaptions have improved significantly (compare columns MPE and MAPE, for swaption prices, of tables 4.10 and 4.12), they are still too large to be "tolerable". Second, the fit to the market prices of caps is reasonable and much better than the one provided by the equivalent equilibrium model, but nonetheless worse than the one obtained in table 4.11 (i.e. without calibrating the HJM model also to swaption prices). Finally, the standard errors of the h_j estimates are much higher than the ones reported previously in table 4.11, which probably means that the simultaneous fitting of cap and swaption prices is a too heavy burden for the function $h(t)$ alone. In fact, the calibration to the market swaption prices depends entirely on function $h(t)$: substituting equation (4.46) into formula (4.66), the time-homogeneous interest rate correlation function (4.70) is again obtained.

4.7 Conclusions

Recent literature on panel-data estimation of state-space term structure models (e.g. Babbs and Nowman (1999)) has shown that a low-dimensional Gaussian affine and time-homogeneous interest rate model can fit remarkably well the yield curve. The main purpose of this Chapter was to test whether such a simple framework could also reproduce well enough the market interest rates covariance matrix; and if not, to suggest alternative modelling solutions.

In a first stage, the Dai and Singleton (1998) general Gaussian canonical specification, with unobservable state variables, was represented in a state-space form and estimated using a panel-data not only of swap rates but also of ATM cap prices, and of ATM European swaption prices. For this purpose, a non-linear Kalman filter (the EKF) was used, its finite sample properties being successfully diagnosed through a Monte Carlo study. Hence, one of the contributions of this Chapter is the application of Kalman filtering estimation techniques to an enlarged panel-data set containing information not only about the level of the yield curve but also about its volatility and correlation surfaces.

As in Babbs and Nowman (1999), the empirical results presented in this dissertation also show that a simple affine and time-homogeneous Gaussian model, with only two or three state variables, can successfully reproduce the term structure of interest rates. Additionally, it also seems that the model fit of the yield curve is better achieved by a non-linear Kalman filter that possesses no exact asymptotic properties, than by the usual linear one that makes no efficient use of all market data.

However, even by increasing the number of model' factors, it is not possible to price accurately short-maturity caps and European swaptions. Moreover, for high-dimensional model specifications some parameters' estimates became too unstable, while the fit to the market interest rates covariance matrix did not improve significantly. But, even though the "equilibrium" model specification under consideration can not reproduce adequately market cap and swaption prices, it is still important to consider such market data in the estimation of the model' parameters, because additional information about the market covariance matrix seems to be incorporated into the model: in particular, it was shown that the model term structure of volatilities and correlation function are affected by the set of market observables to which the model is fitted.

The model inability to fit the observed hump at the short-end of the volatility curve can be attributed to the time-homogeneous nature of the model, while the mispricing of swaptions has been associated (e.g. Cooper and Rebonato (1995)) with the fact that low-dimensional models are not able to produce interest rate correlation functions of exponential-decaying shape. However, it was shown, both analytically and empirically, that the sigmoid shape at the short-end of the correlation function is independent of the model' dimension. Therefore, a better fit to both cap and swaption prices was tried in the context of a time-inhomogeneous framework, and not by using higher-dimensional time-homogeneous specifications.

In a second stage, a Gaussian HJM model, equivalent to the previous "equilibrium" specification but with a time-inhomogeneous diffusion term, was proposed and estimated in two stages. First, the time-independent diffusion parameters were estimated from the previous "equilibrium" specification, i.e. by using a Kalman filter approach (which ensures their stability over time). Then, the time-dependent component was used to better fit, cross-sectionally, the principal diagonal elements of the interest rates covariance matrix. Exact pricing of swaps and caps can be easily achieved, while European swaption pricing errors are significantly (but not satisfactorily) reduced, in the context of stable time-homogeneous diffusion coefficients.

The inability of affine Gaussian models to fit the market interest rates covariance matrix opens the question of whether better results could be obtained, still under the affine class of term structure models, but using more general stochastic volatility formulations. However, such stochastic volatility models are, in general, unable to produce analytic solutions for swaps, caps, and European swaptions, which makes it virtually impossible to fit the model to a panel-data of such interest rate contingent claims' quotes, using the methodology described in this Chapter. Next Chapter proposes fast and accurate approximate closed-form solutions for these derivatives, and therefore opens the possibility of fitting affine stochastic volatility models to the term structures of interest rates, volatilities, and correlations.

4.8 Appendices

4.8.1 $\mathbb{A}_0(n)$ formulation of the Babbs and Nowman (1999) model

The Babbs and Nowman (1999) model can be rewritten, with self-evident notational changes, as:

$$r(t) = \mu - \underline{1}' \cdot \underline{Y}(t),$$

$$d\underline{Y}(t) = -\xi^D \cdot \underline{Y}(t)\,dt + \kappa \cdot d\underline{W}^{\mathcal{P}}(t),$$

and

$$d\underline{W}^{\mathcal{Q}}(t) = \underline{\Theta}dt + d\underline{W}^{\mathcal{P}}(t),$$

where $\underline{Y}(t) \in \Re^n$ denotes the state variables vector, and $\mu \in \Re$, $\underline{\Theta} \in \Re^n$, $\kappa \in \Re^{n \times n}$ and $\xi^D = diag\{\xi_1, \ldots, \xi_n\}$ are all model parameters.

Applying the *invariant affine transformation*

$$\underline{X}(t) = \kappa^{-1} \cdot \underline{Y}(t),$$

equations (2.2), (4.9), and (4.10) follow but subject to:

$$f = \mu, \tag{4.54}$$

$$\underline{G} = -\kappa' \cdot \underline{1}, \tag{4.55}$$

$$K^{\Delta} = \kappa^{-1} \cdot \xi^{D} \cdot \kappa, \tag{4.56}$$

and

$$\underline{\lambda} = \underline{\Theta}. \tag{4.57}$$

If matrix κ is further assumed to be (lower) triangular -as it is the case in the empirical analysis conducted by Babbs and Nowman (1999)- then, equations (4.55) and (4.56) yield a unique solution, and all the initial parameters are exactly recovered under the $\mathbb{A}_0(n)$ formulation.

4.8.2 Analytical formulae for $\hat{Z}_k$

The availability of a closed-form solution for

$$\hat{Z}_k = \left.\begin{bmatrix} \frac{\partial Z_{1_k}(X_k)}{\partial X_k} \\ \frac{\partial Z_{2_k}(X_k)}{\partial X_k} \\ \frac{\partial Z_{3_k}(X_k)}{\partial X_k} \end{bmatrix}\right|_{X_k = \hat{X}_{k|k-1}}$$

will greatly enhance the numerical efficiency of the Kalman filter recursions described in subsection 4.3.3. This appendix simply states such analytical solution.

Differentiating equation (4.15) with respect to $\underline{X}_k$, it is possible to write the j^{th} column of matrix $\frac{\partial Z_{1_k}(X_k)}{\partial X_k}$ as

$$\begin{aligned} &\underline{\left(\frac{\partial Z_{1_k}(\underline{X}_k)}{\partial \underline{X}_k}\right)_j} \\ \equiv\; - &\left[\delta^{IRS} \sum_{i=1}^{m_j^{IRS}} P\left(\underline{X}_k; i\delta^{IRS}\right)\right]^{-2} \\ &\left[\underline{B}\left(m_j^{IRS}\delta^{IRS}\right) P\left(\underline{X}_k; m_j^{IRS}\delta^{IRS}\right) \delta^{IRS} \sum_{i=1}^{m_j^{IRS}} P\left(\underline{X}_k; i\delta^{IRS}\right)\right. \end{aligned} \tag{4.58}$$

$$+ \left(1 - P\left(\underline{X}_k; m_j^{IRS}\delta^{IRS}\right)\right) \delta^{IRS} \sum_{i=1}^{m_j^{IRS}} \underline{B}\left(i\delta^{IRS}\right) P\left(\underline{X}_k; i\delta^{IRS}\right) \Bigg] .$$

Similarly, differentiating equation (4.16) with respect to $\underline{X}_k$, it can be shown that the j^{th} column of matrix $\frac{\partial Z_{2_k}(\underline{X}_k)}{\partial \underline{X}_k}$ is equal to

$$\left(\frac{\partial Z_{2_k}(\underline{X}_k)}{\partial \underline{X}_k}\right)_j \equiv \sum_{i=1}^{m_j^{cap}-1} \left\{\underline{B}\left(i\delta^{cap}\right) P\left(\underline{X}_k; i\delta^{cap}\right) \Phi\left[\sigma_{i\delta^{cap},(i+1)\delta^{cap}} - d_{k,j}(i)\right]\right. \tag{4.59}$$
$$\left. - \left(1 + \delta^{cap} w_{k,j}^{cap}\right) \underline{B}\left((i+1)\delta^{cap}\right) P\left(\underline{X}_k; (i+1)\delta^{cap}\right) \Phi\left[-d_{k,j}(i)\right]\right\},$$

with

$$d_{k,j}(i) = \frac{\ln\left[\frac{P(\underline{X}_k;(i+1)\delta^{cap})\left(1+\delta^{cap} w_{k,j}^{cap}\right)}{P(\underline{X}_k;i\delta^{cap})}\right] + \frac{\sigma^2_{i\delta^{cap},(i+1)\delta^{cap}}}{2}}{\sigma_{i\delta^{cap},(i+1)\delta^{cap}}},$$

and where $w_{k,j}^{cap}$ denotes the market time-t_k forward swap rate with settlement in arrears at times $t_k + i\delta^{cap}, i = 2, \ldots, m_j^{cap}$.

Finally, differentiating equation (4.17) with respect to $\underline{X}_k$, it follows that the j^{th} column of matrix $\frac{\partial Z_{3_k}(\underline{X}_k)}{\partial \underline{X}_k}$ corresponds to

$$\left(\frac{\partial Z_{3_k}(\underline{X}_k)}{\partial \underline{X}_k}\right)_j \equiv \underline{B}\left(u_j^{swpn}\delta^{swpn}\right) P\left(\underline{X}_k; u_j^{swpn}\delta^{swpn}\right) \Phi\left(-d_{k,j}\right) \tag{4.60}$$
$$- \sum_{i=1}^{m_j^{swpn}} \underline{B}\left(\left(u_j^{swpn} + i\right)\delta^{swpn}\right) x_i^{k,j} P\left(\underline{X}_k; \left(u_j^{swpn} + i\right)\delta^{swpn}\right)$$
$$\Phi\left(-d_{k,j} - \sigma_{u_j^{swpn}\delta^{swpn},\left(u_j^{swpn}+i\right)\delta^{swpn}}\right),$$

where $d_{k,j}$ is the solution of

$$P\left(\underline{X}(t); u\delta\right) = \sum_{i=1}^{m_j^{swpn}} x_i^{k,j} P\left(\underline{X}_k; \left(u_j^{swpn} + i\right)\delta^{swpn}\right)$$
$$\exp\left(-\frac{\sigma^2_{u_j^{swpn}\delta^{swpn},\left(u_j^{swpn}+i\right)\delta^{swpn}}}{2} - d_{k,j}\sigma_{u_j^{swpn}\delta^{swpn},\left(u_j^{swpn}+i\right)\delta^{swpn}}\right),$$

$$x_i^{k,j} = 1_{\left\{i=m_j^{swpn}\right\}} + w_{k,j}^{swpn}\delta^{swpn},$$

and $w_{k,j}^{swpn}$ represents the market time-t_k forward swap rate with settlement in arrears at times $t_k + \left(u_j^{swpn} + i\right)\delta^{swpn}, i = 1, \ldots, m_j^{swpn}$.

4.8.3 Derivatives' recursions

As shown by Harvey (1989, section 3.4.6) for the standard linear Kalman filter, the system derivatives $\frac{\partial \mu_k}{\partial \Psi_i}$ and $\frac{\partial \Omega_k}{\partial \Psi_i}$ (for $i = 1, \ldots, p$) can be evaluated analytically by running, for each k, $6 \times p$ recursions in parallel with the Kalman filter iterations presented in subsection 4.3.3. Using straightforward change of notation, those formulae can be used for the linear state-space model represented by equations (4.18) and (4.21). The purpose of this appendix is to adapt such derivatives' recursions for the EKF.

Differentiating (4.33) with respect to the i^{th} model parameter,

$$\frac{\partial \underline{\mu}_k}{\partial \Psi_i} = - \left. \frac{\partial \underline{Z}_k(\underline{X}_k)}{\partial \Psi_i} \right|_{\underline{X}_k = \hat{\underline{X}}_{k|k-1}} - \hat{Z}_k' \cdot \frac{\partial \hat{\underline{X}}_{k|k-1}}{\partial \Psi_i}, \tag{4.61}$$

where $\left. \frac{\partial \underline{Z}_k(\underline{X}_k)}{\partial \Psi_i} \right|_{\underline{X}_k = \hat{\underline{X}}_{k|k-1}}$ can be computed using finite differences. Substituting H by $\hat{Z}_k$ (as given by equation (4.26)), when moving from the linear case -Harvey (1989, equation 3.4.72)- to the EKF,

$$\begin{aligned} \frac{\partial \Omega_k}{\partial \Psi_i} &= \frac{\partial \hat{Z}_k'}{\partial \Psi_i} \cdot P_{k|k-1} \cdot \hat{Z}_k + \hat{Z}_k' \cdot \frac{\partial P_{k|k-1}}{\partial \Psi_i} \cdot \hat{Z}_k \\ &\quad + \hat{Z}_k' \cdot P_{k|k-1} \cdot \frac{\partial \hat{Z}_k}{\partial \Psi_i} + \frac{\partial U^D}{\partial \Psi_i}, \end{aligned} \tag{4.62}$$

where

$$\frac{\partial \hat{Z}_k}{\partial \Psi_i} = \left. \frac{\partial^2 \underline{Z}_k(\underline{X}_k)}{\partial \underline{X}_k \partial \Psi_i} \right|_{\underline{X}_k = \hat{\underline{X}}_{k|k-1}} + \sum_{l=1}^{n} \left. \frac{\partial^2 \underline{Z}_k(\underline{X}_k)}{\partial \underline{X}_k \partial (\underline{X}_k)_l} \right|_{\underline{X}_k = \hat{\underline{X}}_{k|k-1}} \cdot \frac{\partial \left(\hat{\underline{X}}_{k|k-1} \right)_l}{\partial \Psi_i}. \tag{4.63}$$

The matrices $\left. \frac{\partial^2 \underline{Z}_k(\underline{X}_k)}{\partial \underline{X}_k \partial (\underline{X}_k)_l} \right|_{\underline{X}_k = \hat{\underline{X}}_{k|k-1}}$ can be obtained analytically by differentiating twice the pricing formulae (4.15), (4.16), and (4.17),[35] while $\left. \frac{\partial^2 \underline{Z}_k(\underline{X}_k)}{\partial \underline{X}_k \partial \Psi_i} \right|_{\underline{X}_k = \hat{\underline{X}}_{k|k-1}}$ can be computed using finite differences.

The derivative recursions for the "prediction" step, $\frac{\partial \hat{\underline{X}}_{k|k-1}}{\partial \Psi_i}$ and $\frac{\partial P_{k|k-1}}{\partial \Psi_i}$, are exactly the same as for the standard Kalman filter (because the transition equation is still linear), and the recursions for the "update" step, $\frac{\partial \hat{\underline{X}}_k}{\partial \Psi_i}$ and $\frac{\partial P_k}{\partial \Psi_i}$, are easily obtained by replacing H by $\hat{Z}_k$ in Harvey (1989, equations 3.4.74a and 3.4.74b).

Finally, differentiating equations (4.36) and (4.37), it is possible to initialize the derivative recursions:

$$\frac{\partial \hat{\underline{X}}_0}{\partial \Psi_i} = \underline{0}, \tag{4.64}$$

[35] Or, simply by differentiation of the formulae presented in appendix 4.8.2.

and

$$vec\left(\frac{\partial P_0}{\partial \Psi_i}\right) = \left(I_{n^2} - F \otimes F\right)^{-1} \cdot \left[\frac{\partial (F \otimes F)}{\partial \Psi_i} \cdot \left(I_{n^2} - F \otimes F\right)^{-1} \cdot vec(\Delta(h)) + vec\left(\frac{\partial \Delta(h)}{\partial \Psi_i}\right)\right]. \tag{4.65}$$

4.8.4 Correlation function

Using equation (4.43), the time-t correlation coefficient between changes in instantaneous forward rates of maturities $T_1\,(\geq t)$ and $T_2\,(\geq T_1)$,

$$\rho(T_1, T_2) = \frac{COV\left[df(t, T_1), df(t, T_2) | \mathcal{F}_t\right]}{\sqrt{VAR\left[df(t, T_1) | \mathcal{F}_t\right]}\sqrt{VAR\left[df(t, T_2) | \mathcal{F}_t\right]}},$$

can be written as:

$$\rho(T_1, T_2) = \frac{\underline{\gamma}'(t, T_1) \cdot \underline{\gamma}(t, T_2)}{\sqrt{\underline{\gamma}'(t, T_1) \cdot \underline{\gamma}(t, T_1)}\sqrt{\underline{\gamma}'(t, T_2) \cdot \underline{\gamma}(t, T_2)}}. \tag{4.66}$$

Using straightforward, but tedious, differential calculus, the slope of the correlation function can be found to be

$$\frac{\partial \rho(T_1, T_2)}{\partial T_2} = \frac{\underline{\gamma}'(t, T_1) \cdot \frac{\partial}{\partial T_2}\underline{\gamma}(t, T_2) - \underline{\gamma}'(t, T_1) \cdot \underline{\gamma}(t, T_2) \cdot \frac{\underline{\gamma}'(t, T_2) \cdot \frac{\partial}{\partial T_2}\underline{\gamma}(t, T_2)}{\underline{\gamma}'(t, T_2) \cdot \underline{\gamma}(t, T_2)}}{\sqrt{\underline{\gamma}'(t, T_1) \cdot \underline{\gamma}(t, T_1)}\sqrt{\underline{\gamma}'(t, T_2) \cdot \underline{\gamma}(t, T_2)}}.$$

Hence, expression (4.44) is verified, no matter the dimension of $\underline{\gamma}(t, T)$.

The last general result is also valid for the Gaussian interest rate model under analysis, because this model is nested into the general specification (4.43). In fact, applying Itô's lemma to (4.11), while using equations (4.9) and (4.10), yields

$$\frac{dP(\underline{X}(t); \tau)}{P(\underline{X}(t); \tau)} = r(t)\, dt + \underline{B}'(\tau) \cdot d\underline{W}^{Q}(t). \tag{4.67}$$

And since $f(t, T) = -\frac{\partial \ln P(\underline{X}(t); \tau)}{\partial T}$, then $\underline{\gamma}(t, T) = -\frac{\partial}{\partial T}\underline{B}(\tau)$. Or using equation (4.12),

$$\underline{\gamma}(t, T) = \underline{G}' \cdot e^{-K^{\Delta}(T-t)}, \tag{4.68}$$

which is a time-homogeneous but state-independent function. From the no-arbitrage condition

$$\alpha(t, T) = \underline{\gamma}'(t, T) \cdot \int_t^T \underline{\gamma}(t, s)\, ds, \tag{4.69}$$

it is obvious that the drift term is also time-homogeneous but state-independent, and therefore it can be concluded that the specification (4.43) nests the model defined by equations (2.2), (4.9) and (4.10). Moreover, for this latter particular model, the function (4.66) is specialized into

$$\rho(T_1, T_2) = \frac{\underline{G}' \cdot e^{-K^{\Delta}(T_1 - t)} \cdot e^{-(K^{\Delta})'(T_2 - t)} \cdot \underline{G}}{\sqrt{\underline{G}' \cdot e^{-[K^{\Delta}+(K^{\Delta})'](T_1 - t)} \cdot \underline{G}}\sqrt{\underline{G}' \cdot e^{-[K^{\Delta}+(K^{\Delta})'](T_2 - t)} \cdot \underline{G}}}, \quad (4.70)$$

and thus it possesses a completely time-homogeneous nature.

4.8.5 UK data set

This appendix simply replicates the empirical analysis presented before for US data. The purpose it to show the robustness of the conclusions obtained so far, by testing them using a different currency (GBP) and a much larger data set (containing 784 cross-sections of LIBOR-rate derivatives' quotes, instead of just 189 periods as before).

Table 4.13 presents the summary statistics of the new data set, which was obtained from Datastream. It contains daily middle quotes, from 01/01/96 to 31/12/98, of UK swap rates (for 2, 3, 4, 5, 7, and 10 years), UK ATM forward-start cap flat yield volatilities (for 1, 2, 3, 4, 5, 7, and 10 years), and UK ATM European swaption flat yield volatilities (for 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 1×2, 1×3, and 2×4 years).[36] Using also money-market GBP LIBOR rates, discount factors were computed for all quarterly maturities from 0.25 years to 10 years (through linear interpolation of swap rates and bootstrapping of discount factors), and all cap and swaption flat yield volatility quotes were converted into option prices. Attending to the large standard deviations of the caps and swaptions quotes, it can be inferred that, within the sample, the prices of these derivatives presented large variations.

In table 4.14, one-, two-, and three-factor models are estimated from a panel-data containing only UK swap rates, and using the Extended Kalman Filter described in subsection 4.3.2. The one-factor model is poorly fitted to the data: not only the standard errors of the parameters estimates are significantly high, but also the single state-variable exhibits a mean-fleeting behavior ($K_{11} < 0$). On the other hand, the two- and three-factor models possess realistic parameters' estimates, with much lower standard errors, as well as an extremely small estimate for the standard deviation of the swaps measurement errors. Table 4.15 summarizes the average and absolute average pricing errors between model' and market' swap rates, for all the sample and with respect to each specification estimated in table 4.14. As for the US data set, with only three state variables it is possible to obtain mean absolute pricing errors consistently lower than 1 basis point, i.e. it is possible to obtain an

[36] Hence, $M_1 = 6$, $M_2 = 7$, and $M_3 = 8$.

extremely good fit to the level of the UK yield curve through a low-dimensional and affine Gaussian model. Figure 4-6 exemplifies such remarkable fit, by comparing the market UK 10-year swap rate with the one generated by the three-factor model.

Table 4.16 estimates two- and three-factor Gaussian and non-linear state-space models, using not only UK swap rates but also UK ATM cap prices. In table 4.17 average pricing errors are computed for all the contracts used in the estimation stage, and for both models estimated in table 4.16. As was the case for the US data set, it is not possible to obtain a good model' fit to short-term UK cap prices (mainly for the one-year maturity).

Table 4.18 contains the two- and three-factor parameters' estimates (and standard errors) obtained when fitting the non-linear state-space model -defined by equations (4.18) and (4.24)- to the all UK data set: that is, to swap rates, ATM cap prices, and ATM European swaption prices. The corresponding pricing errors are given in table 4.19. As for the US data set, the Gaussian time-homogeneous and affine model under analysis can not be successfully fitted to swaption market prices (the average absolute percentage errors vary from 27.5% to 59% for the three-factor model). However, and in opposition to the results previously shown in table 4.10, UK swaptions are underpriced by the Gaussian model for a significant part of the sample. Figure 4-7 compares the short-term interest rate estimated from two-factor models fitted just to swap rates, also to cap prices, or even also to swaption prices: the three time-series only differ slightly in terms of levels, but the pattern is essentially the same.

Table 4.20 converts the three-factor time-homogeneous model previously estimated in table 4.16 into the equivalent HJM time-inhomogeneous specification of equations (4.43), (4.46) and (4.47). The time-independent parameters are taken directly from table 4.16, and then the time-dependent step-function $h(t)$ is calibrated, cross-sectionally, to ATM cap prices, using the same breakpoints as in table 4.11. Similarly to the results obtained from the US data set, table 4.20 shows that an affine and time-inhomogeneous Gaussian model can fit remarkably well UK ATM cap prices. In table 4.21, the same three-factor model is calibrated cross-sectionally to both cap and swaption prices (using the time-independent parameters estimated in table 4.18). Although the fit to swaption prices has improved significantly (when compared with table 4.19), it is still far from being acceptable; moreover, the calibration to swaption prices deteriorates the model' fit to cap prices (that is, both types of derivatives can not be simultaneously and consistently priced through the HJM Gaussian affine and time-inhomogeneous model under consideration).

Table 4.13: Descriptive statistics of UK daily data (middle quotes from 01/01/96 to 31/12/98)

	Mean	Std. Deviation	Maximum	Minimum
IRSs:				
2 years	6.931%	0.436%	7.700%	5.445%
3 years	6.989%	0.415%	7.520%	5.450%
4 years	7.033%	0.471%	7.735%	5.440%
5 years	7.063%	0.546%	7.905%	5.430%
7 years	7.117%	0.696%	8.225%	5.370%
10 years	7.187%	0.844%	8.505%	5.340%
ATM Caps:				
1 year	13.298%	3.537%	23.500%	8.000%
2 years	16.416%	3.818%	25.500%	10.200%
3 years	17.363%	3.535%	27.250%	11.700%
4 years	17.614%	3.217%	27.000%	12.700%
5 years	17.584%	3.201%	27.000%	12.700%
7 years	16.792%	2.832%	25.000%	12.700%
10 years	15.792%	2.499%	24.000%	12.700%
ATM European Swaptions:				
0.5 x 2 years	15.380%	3.085%	23.050%	10.050%
0.5 x 3 years	15.184%	2.775%	22.700%	10.300%
0.5 x 4 years	15.005%	2.559%	22.650%	10.500%
0.5 x 5 years	14.746%	2.477%	22.550%	10.450%
0.5 x 7 years	14.126%	2.165%	22.350%	10.600%
1 x 2 years	16.076%	2.911%	24.900%	11.100%
1 x 3 years	15.677%	2.568%	24.100%	11.050%
2 x 4 years	14.813%	1.934%	22.500%	12.000%

Table 4.14: Estimation of the Dai and Singleton (1998) Gaussian model through EKF and using UK swap rates

	One-factor model		Two-factor model		Three-factor model	
Parameter	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
f	0.04800	0.14908	-0.00727	0.02997	0.10287	0.00897
G_1	0.00265	0.00153	0.00796	0.00067	0.01882	0.00464
G_2			0.00647	0.00030	-0.00246	0.00053
G_3					-0.00760	0.00038
K_{11}	-0.21905	0.00822	0.52187	0.01367	2.15277	0.35992
K_{21}			0.29348	0.05428	0.08451	0.19125
K_{31}					-3.22786	0.61846
K_{22}			-0.03310	0.00195	0.02770	0.01007
K_{32}					0.35266	0.02624
K_{33}					0.15665	0.01979
λ_1	1.53493	11.38829	-0.80272	0.28222	-0.28683	0.27208
λ_2			-0.22068	0.19807	0.03405	0.08330
λ_3					1.47093	0.41244
σ_{ε_1}	0.00245	0.00005	0.00025	0.00001	0.00011	0.00000
ln L	25791		34997		37363	
BIC	25778		34970		37320	

Data: UK swap rates for 2, 3, 4, 5, 7, and 10 years
(from 01/01/96 to 31/12/98)

Table 4.15: Goodness of fit to market values of UK swap rates of the models estimated in table 4.14

IRS	One-factor model		Two-factor model		Three-factor model	
maturities:	Mean Error	MAE	Mean Error	MAE	Mean Error	MAE
2 years	0.0438%	0.3776%	0.0029%	0.0216%	-0.0002%	0.0044%
3 years	0.0098%	0.1972%	-0.0040%	0.0142%	0.0016%	0.0087%
4 years	-0.0075%	0.1006%	-0.0048%	0.0177%	-0.0021%	0.0058%
5 years	-0.0084%	0.0796%	0.0021%	0.0164%	0.0003%	0.0083%
7 years	0.0018%	0.1391%	0.0067%	0.0105%	0.0007%	0.0089%
10 years	0.0364%	0.2109%	-0.0020%	0.0181%	0.0003%	0.0055%

Errors are differences between model and market rates.
MAE are mean absolute errors.

Table 4.16: Estimation of the Dai and Singleton (1998) Gaussian model through EKF and using UK swap rates and cap prices

	Two-factor model		Three-factor model	
Parameter	Estimate	Std. Error	Estimate	Std. Error
f	0.03653	0.00333	0.07211	0.00433
G_1	0.01336	0.00034	0.01864	0.00158
G_2	0.00767	0.00012	-0.00384	0.08579
G_3			-0.00761	0.04331
K_{11}	0.55027	0.01304	2.43093	0.05359
K_{21}	0.23624	0.04407	0.14758	31.97113
K_{31}			-2.82987	1.69573
K_{22}	-0.05886	0.00207	0.07487	4.59832
K_{32}			0.40735	0.02957
K_{33}			0.07500	4.60007
λ_1	-1.46033	0.21866	-2.94644	0.40787
λ_2	-0.62398	0.10398	-0.11406	40.57106
λ_3			3.59065	1.38954
σ_{ε_1}	0.00274	0.00004	0.00221	0.00003
σ_{ε_2}	0.00068	0.00001	0.00045	0.00001
ln L	60982		63612	
BIC	60955		63568	

Data: UK swap rates for 2, 3, 4, 5, 7, and 10 years
and cap prices for 1, 2, 3, 4, 5, 7, and 10 years
(from 01/01/96 to 31/12/98)

Table 4.17: Goodness of fit to UK market swap and cap values of the models estimated in table 4.16

	Two-factor model			Three-factor model		
	Swap rates	Cap prices		Swap rates	Cap prices	
Maturities:	Mean Error	MPE	MAPE	Mean Error	MPE	MAPE
1 year		30.2137%	38.2200%		11.5536%	20.5462%
2 years	-0.1423%	-2.3440%	5.8254%	-0.1712%	-2.9056%	4.9609%
3 years	-0.0348%	0.0626%	3.3545%	-0.0227%	0.5547%	2.1808%
4 years	0.0386%	0.1906%	1.5472%	0.0468%	0.3918%	1.4210%
5 years	0.0920%	0.0993%	1.8006%	0.0880%	-0.0115%	0.9323%
7 years	0.1363%	0.5024%	1.1618%	0.1102%	0.2406%	0.9632%
10 years	0.1097%	0.0951%	0.6719%	0.0650%	0.0902%	0.3407%

Errors are differences between model and market values.
MAE are mean absolute errors.
MPE are mean percentage errors, i.e. average of errors divided by market values.
MAPE are mean absolute percentage errors.

Table 4.18: Estimation of the Dai and Singleton (1998) Gaussian model through EKF and using UK swap rates, cap prices and European swaption prices

	Two-factor model		Three-factor model	
Parameter	Estimate	Std. Error	Estimate	Std. Error
f	0.04020	0.00101	0.07454	0.00361
G_1	0.01220	0.00033	0.01738	0.00185
G_2	0.00830	0.00012	-0.00518	0.03516
G_3			-0.00759	0.02409
K_{11}	0.51300	0.01208	2.42198	0.05445
K_{21}	0.12100	0.04393	-0.10983	15.62059
K_{31}			-3.36222	0.58377
K_{22}	-0.05490	0.00207	0.07678	1.88410
K_{32}			0.40612	0.02902
K_{33}			0.07739	1.88838
λ_1	-1.42000	0.10149	-2.79182	0.30653
λ_2	-0.35500	0.08815	0.14080	17.98735
λ_3			3.87048	0.77834
σ_{ε_1}	0.00280	0.00004	0.00225	0.00003
σ_{ε_2}	0.00069	0.00001	0.00046	0.00001
σ_{ε_3}	0.00576	0.00013	0.00595	0.00014
ln L	90100		92472	
BIC	90073		92429	

Data: UK swap rates for 2, 3, 4, 5, 7, and 10 years, cap prices for 1, 2, 3, 4, 5, 7, and 10 years, and swaption prices for 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 1×2, 1×3, and 2×4 (from 01/01/96 to 31/12/98).

Table 4.19: Goodness of fit to UK market swap, cap, and swaption values of the models estimated in table 4.18

Two-factor model							
	Swap rates		Cap prices		Swaption prices		
Maturity	Mean Error	MAE	MPE	MAPE	Maturity	MPE	MAPE
1 year			31.59%	38.98%	0.5x2	-21.36%	32.39%
2 years	-0.145%	0.172%	-2.27%	5.71%	0.5x3	-20.01%	37.40%
3 years	-0.057%	0.139%	0.12%	3.56%	0.5x4	-15.89%	43.45%
4 years	0.005%	0.178%	0.31%	1.43%	0.5x5	-12.25%	48.59%
5 years	0.050%	0.231%	0.28%	1.83%	0.5x7	-5.47%	57.25%
7 years	0.084%	0.293%	0.67%	1.30%	1x2	-28.79%	34.11%
10 years	0.047%	0.370%	0.10%	0.69%	1x3	-22.61%	33.67%
					2x4	-19.11%	35.56%

Three-factor model							
	Swap rates		Cap prices		Swaption prices		
Maturities	Mean Error	MAE	MPE	MAPE	Maturities	MPE	MAPE
1 year			13.1434%	20.8738%	0.5x2	-7.15%	34.82%
2 years	-0.1686%	0.1723%	-3.2629%	5.1096%	0.5x3	-5.89%	40.23%
3 years	-0.0496%	0.0822%	0.5032%	2.1446%	0.5x4	-5.21%	45.94%
4 years	0.0060%	0.1028%	0.5415%	1.2687%	0.5x5	-5.04%	50.40%
5 years	0.0406%	0.1522%	0.1647%	0.9994%	0.5x7	-3.00%	58.97%
7 years	0.0619%	0.2273%	0.2772%	1.0675%	1x2	-15.82%	27.46%
10 years	0.0342%	0.3185%	0.0894%	0.3430%	1x3	-10.94%	32.30%
					2x4	-10.42%	35.45%

Errors are differences between model and market values.
MAE are mean absolute errors.
MPE (MAPE) are mean (absolute) percentage errors.

Table 4.20: Calibration of a three-factor time-inhomogeneous HJM model to UK cap prices, using the time-homogeneous coefficients of table 4.16

Estimated $h(t)$				Fit to cap prices		
j	l_j	Mean(h_j)	Stdev(h_j)	Maturities	MPE	MAPE
1	1	0.843219	0.210747	1 year	-0.0005%	0.0023%
2	2	0.215131	0.190007	2 years	0.0044%	0.0216%
3	4	1.653994	0.261137	3 years	-0.0119%	0.0689%
4	6	0.866117	0.659903	4 years	0.0278%	0.0768%
5	8	1.668803	0.386090	5 years	0.0224%	0.1325%
6	10	0.432386	0.364093	7 years	-0.1353%	0.1942%
7	12	0.967360	0.487924	10 years	0.0905%	0.1414%
8	16	1.277202	0.746938			
9	20	0.388136	0.441380			
10	28	0.198016	0.241229			
11	36	0.281669	0.335690			
12	44	0.305548	0.384403			

Data: cap prices for 1, 2, 3, 4, 5, 7, and 10 years (from 01/01/96 to 31/12/98).
l_j is the j^{th} breakpoint of function $h(t)$.
h_j is the value of $h(t)$ for the time-interval $[l_{j-1}\delta^{cap}, l_j\delta^{cap}[$.
MPE (MAPE) are mean (absolute) percentage errors.

Table 4.21: Calibration of a three-factor time-inhomogeneous HJM model to UK cap and swaption prices, using the time-homogeneous coefficients of table 4.18

		Estimated $h(t)$		Fit to cap prices		Fit to swaption prices		
j	l_j	Mean(h_j)	Stdev(h_j)	Maturity	MAPE	Maturity	MPE	MAPE
1	1	0.291149	0.379624	1 year	22.365%	0.5x2	2.914%	12.958%
2	2	1.687360	0.834543	2 years	10.760%	0.5x3	-0.613%	5.514%
3	4	0.943314	0.419254	3 years	5.815%	0.5x4	-2.960%	6.416%
4	6	0.937911	0.882382	4 years	1.990%	0.5x5	-4.410%	9.400%
5	8	0.410927	0.493058	5 years	0.563%	0.5x7	-4.365%	15.047%
6	10	0.967105	0.883354	7 years	0.174%	1x2	-7.398%	8.641%
7	12	0.498955	0.604351	10 years	0.064%	1x3	-7.005%	8.339%
8	16	1.052603	0.725080			2x4	-7.454%	11.087%
9	20	1.103483	0.568870					
10	28	0.426619	0.359125					
11	36	0.584936	0.441412					
12	44	1.292506	1.413902					

Data: cap prices for 1, 2, 3, 4, 5, 7, and 10 years, and swaption prices for 0.5×2, 0.5×3, 0.5×4, 0.5×5, 0.5×7, 1×2, 1×3, and 2×4 (from 01/01/96 to 31/12/98).
l_j is the j^{th} breakpoint of function $h(t)$.
h_j is the value of $h(t)$ for the time-interval $[l_{j-1}\delta^{cap}, l_j\delta^{cap}[$.
MPE (MAPE) are mean (absolute) percentage errors.

Figure 4-6: Market *versus* three-factor model 10-year UK swap rate

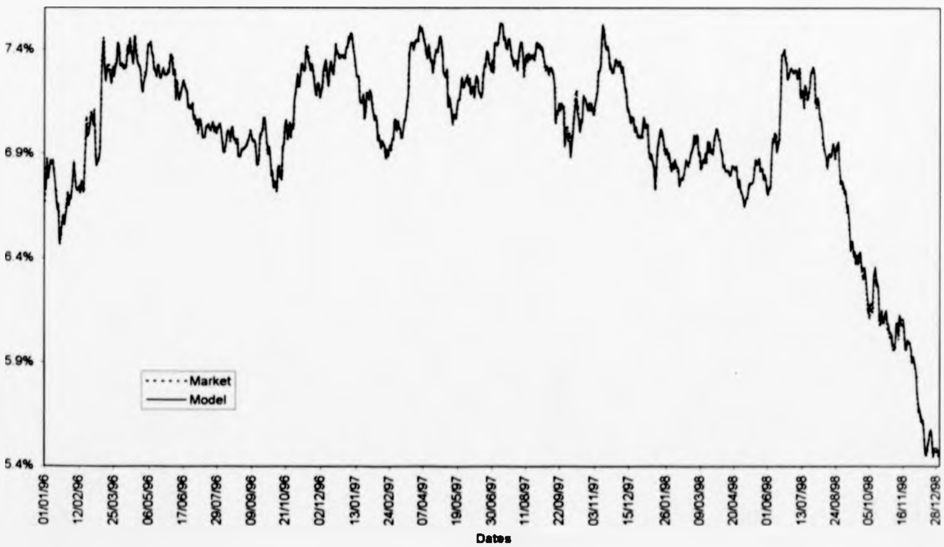


Figure 4-7: UK short-term interest rate estimated by two-factor models fitted to only swap rates, also to cap prices, or even also to swaption prices

Chapter 5

The Stochastic Volatility General Case

This Chapter is based on the article Nunes, Clewlow and Hodges (1999), which supersedes a working paper entitled "Interest Rate Derivatives in a Duffie and Kan Model with Stochastic Volatility: Application of Green's Functions" and presented at the following 1998 meetings: FORC Summer Seminars (University of Warwick), EFMA-FMA Conference (Lisbon), AFFI Meeting (Lille), EFA Conference (Paris), SFA Conference (Florida), and QMF Meeting (Sydney).

5.1 Introduction

Under its most general stochastic volatility specification, Duffie and Kan (1996) derived a quasi-closed form pricing formula for default-free pure discount bonds (involving the numerical solution of the Riccati differential equation (2.8)), and priced path-independent interest rate options, in a two-factor model, through an alternating directions implicit (ADI) finite-difference method. However, such algorithm can not be easily extended to higher dimensions, for which, and according to Duffie and Kan (1994), Monte Carlo simulation appears to be the best pricing methodology available. Consequently, the *expedite* and *accurate* analytic approximate pricing solutions that will be proposed in this Chapter are intended to provide more efficient pricing and calibration alternative tools for this general affine class of term structure models, specially for high dimensional formulations (e.g. three-factor models).

Recently, Duffie et al. (1998) proposed exact Fourier transform pricing solutions for an affine jump-diffusion model that nests, as a special case, the Duffie and Kan (1996) framework under analysis.[1] Although such exact formulae are also applicable to the Duffie and Kan (1996) model, the advantage of the approximate pricing solutions derived in this dissertation is the fact that they are, in general, much faster to implement than the corresponding exact ones obtainable from Duffie et al. (1998). In fact, if the functional form of the relevant characteristic function -Duffie et al. (1998, equation B.2)- is known, then the exact Fourier transform pricing formulae are "explicit" or closed-form solutions (in the sense that only one-dimensional Fourier inverse integrals are involved). However, in general the characteristic function does not possess an explicit solution and must be numerically obtained from a complex-valued system consisting of one unidimensional ODE and another n-dimensional Riccati differential equation, where n denotes the number of state variables. Because the computation, for instance, of the exact price of an European option on a pure discount bond requires two inverse Fourier transforms (and thus two one-dimensional integrals; one for each exercise probability), and since the characteristic function must be evaluated numerically at each integration point, then, for n state variables, the Fourier transform exact solution involves a $2m(n+1)$ integration problem,[2] where m is the chosen number of steps in the numerical integration, whereas the corresponding first order approximate formulae proposed in this Chapter will only include one time-integral (no matter the order of n). For the valuation of a cap (or a floor), the difference between the two (exact and approximate) solutions, in terms of computational effort, is even multiplied by the number of component caplets (or floorlets). Consequently, in the general case where no closed-form solution exists for the characteristic function, the proposed approximate pricing formulae will be shown to be much faster to implement than the exact Fourier transform ones. Moreover, when the characteristic function is not known in closed-form, the optimization of both the grid size and the upper bound of integration for the computation of the inverse Fourier transforms becomes also very time-demanding, since it requires the numerical evaluation of the characteristic function.

By imposing a deterministic volatility specification to the Duffie and Kan (1996) formulation, the Langetieg (1980) *multivariate elastic random walk model* was obtained in Chapter three. This type of Gaussian multifactor affine models has received an extensive treatment in the literature, and exact closed-form pricing formulae have been derived for several interest rate contingent claims, among others, by El Karoui et al. (1991), Jamshidian (1993), Brace and Musiela (1994a), and Nunes (1998). The purpose of this Chapter is to use the Gaussian valuation solutions derived in Chapter three in order to obtain approximate closed-form pricing formulae, under the stochastic volatility specification of the

[1] As already suggested in Chen and Scott (1995b, page 58).

[2] In other words, the computational burden grows linearly with the number of model' factors.

Duffie and Kan (1996) model, for several European-style interest rate contingent claims[3], namely for: default-free bonds, FRAs, IRSs, short-term and long-term interest rate futures, European spot and futures options on zero-coupon bonds, interest rate caps and floors, European (*conventional* and *pure*) futures options on short-term interest rates, and even for European swaptions.[4]

In order to derive the above mentioned stochastic volatility approximate pricing solutions, first, the functional form of an Arrow-Debreu price, for the Gaussian specification of the Duffie and Kan (1996) model, will be obtained in a slightly more general form than the one given by Beaglehole and Tenney (1991, page 73). Then, each stochastic volatility approximate analytic solution will be expressed in terms of the previously derived Gaussian Arrow-Debreu state price, in terms of the corresponding Gaussian exact pricing formula derived in Chapter three, and in terms of the model' parameters imposing stochastic volatility. The resulting first order approximate pricing formula will include one integral with respect to each one of the model' state variables, and another integral with respect to the time-to-maturity of the contingent claim under valuation. Hence, the methodology employed in this Chapter follows, up to this point, the work of Chen (1996), although his "general" and "special" three-factor model specifications are different from the ones used here.

However, this type of multidimensional integral solutions would have to be computed numerically through repeated one-dimensional integration or by using Monte Carlo integration, which does not necessarily represent any improvement in terms of efficiency when compared with the existing exact numerical solutions. Consequently, because the practical usefulness of these multidimensional integral approximations may be questionable, a different approach is pursued: to reduce the dimensionality of the problem analytically. Unlike in Chen (1996) and as the main contribution of the present work, all the stochastic volatility first order approximate closed-form solutions will be simplified into equivalent pricing formulae that do not involve any integration with respect to the model' factors. Such simplification will be allowed by the tractability of the chosen nested Gaussian specification, and will increase enormously the numerical efficiency of the stochastic volatility pricing approximations: only one time-integral is involved, irrespective of the model' dimension. Therefore, such first order analytic approximations will be shown to be extremely fast, as well as accurate.

To the author's knowledge, although the use of approximations involving Arrow-Debreu securities is common in Finance, the derivation of factor-integral independent pricing solu-

[3]That is derivatives with only one future admissible payoff.

[4]The valuation of LIBOR-rate derivatives will be based on the Duffie and Singleton (1997) assumption of symmetric counterparty credit risk.

tions (in the context of the most general multifactor affine term structure model) represents an original result. In addition, exact pricing formulae (involving the numerical solution of Riccati differential equations) are also found for long-term and short-term interest rate futures, under the stochastic volatility specification of the Duffie and Kan (1996) model.

Next sections are organized as follows. Section 5.2 derives an analytical solution for Arrow-Debreu state prices under the deterministic volatility specification of the Duffie and Kan (1996) model. Then, section 5.3 provides a series expansion pricing equation for a generic interest rate derivative, under the stochastic volatility specification, and based on the previously derived Gaussian Arrow-Debreu state-prices. Section 5.4 simplifies the previous pricing solution for any "exponential-affine" interest rate contingent claim, and yields a general first order explicit approximation only involving one time-integral. Such explicit approximate stochastic volatility pricing formula is then applied to different contracts: bonds, FRAs, IRSs, bond futures, and short-term interest rate futures. Similarly, in section 5.5 the global series expansion pricing equation is converted into an explicit first order approximation (only involving one time-integral) for a generic European and path-independent interest rate option. This explicit generic solution is then specialized to options on pure discount bonds, caps and floors, swaptions, yield options, futures options on zero-coupon bonds, and options on short-term interest rate futures. Finally, section 5.6 summarizes the main conclusions. All accessory proofs are relegated to the appendix, while the more illustrative ones are kept in the text.

5.2 Arrow-Debreu Prices under the Gaussian Specification

In this section a closed form solution will be derived for an Arrow-Debreu state price, under the Gaussian specification corresponding to equations (2.2) and (3.1). The formula that will be obtained is equivalent to the one given by Beaglehole and Tenney (1991, page 73), with just two differences: the short-term interest rate is not constrained to be one of the model' factors; and, proposition 2 ensures that no single integral is involved when all the eigenvectors of matrix a are assumed to be linearly independent.

Proposition 19 *Let $G[\underline{X}(T), T; \underline{X}(t), t]$ represent the value, at time t (and in state $\underline{X}(t)$), of a unit payoff occurring at time T ($\geq t$) and in state $\underline{X}(T)$. Under the deterministic volatility specification of the Duffie and Kan (1996) model,*[5] *the Arrow-Debreu price $G[\underline{X}(T), T; \underline{X}(t), t] \equiv \varphi[\underline{X}(t), t] \in C^{2,1}(\Re^n \times [0, T])$ possesses the following analytical*

[5] *That is conditional on $\underline{X}(t) \in \Re^n$ being a strong solution of SDE (3.1).*

form:

$$\begin{aligned} & G\left[\underline{X}(T), T; \underline{X}(t), t\right] \\ = \; & P_G(t,T)\,\frac{\exp\left\{-\frac{1}{2}\left[\underline{X}(T)-\underline{M}(\tau)\right]'\cdot\Delta^{-1}(\tau)\cdot\left[\underline{X}(T)-\underline{M}(\tau)\right]\right\}}{\sqrt{(2\pi)^n\left|\Delta(\tau)\right|}}, \end{aligned} \tag{5.1}$$

where

$$\underline{M}(\tau)=a^{-1}\cdot\left(e^{a\tau}-I_n\right)\cdot\left[\underline{b}+\Theta\cdot\left(a^{-1}\right)'\cdot\underline{G}\right]+e^{a\tau}\cdot\underline{X}(t)-\Delta(\tau)\cdot\left(a^{-1}\right)'\cdot\underline{G},$$

and $P_G(t,T)$ denotes a pure discount bond price computed under proposition 1, i.e. under the deterministic volatility specification of the Duffie and Kan (1996) model.[6]

Proof. The Arrow-Debreu security $G\left[\underline{X}(T), T; \underline{X}(t), t\right]$, as any other contingent claim, is, under the deterministic volatility specification of the Duffie and Kan (1996) model and subject to existence conditions (as stated, for instance, in Friedman (1975, theorem 4.5)), the solution of the following Kolmogorov's backward equation

$$\begin{aligned} 0 \;=\; & \mathcal{D}_G G\left[\underline{X}(T), T; \underline{x}, t\right]+\frac{\partial G\left[\underline{X}(T), T; \underline{x}, t\right]}{\partial t} \\ & -r(t)\, G\left[\underline{X}(T), T; \underline{x}, t\right], \end{aligned} \tag{5.2}$$

$\underline{x}\in\Re^n$, subject to a specific boundary condition

$$G\left[\underline{X}(T), T; \underline{x}, T\right]=\delta\left[\underline{x}-\underline{X}(T)\right], \underline{X}(T)\in\Re^n, \tag{5.3}$$

where $\mathcal{D}_G$ is a second-order differential operator under the nested deterministic volatility specification of the Duffie and Kan (1996) model,[7] i.e.

$$\begin{aligned} \mathcal{D}_G G\left[\underline{X}(T), T; \underline{x}, t\right] \;=\; & \frac{\partial G\left[\underline{X}(T), T; \underline{x}, t\right]}{\partial \underline{x}'}\cdot\left(a\cdot\underline{x}+\underline{b}\right) \\ & +\frac{1}{2} tr\left\{\frac{\partial^2 G\left[\underline{X}(T), T; \underline{x}, t\right]}{\partial \underline{x}\partial \underline{x}'}\cdot\Theta\right\}, \end{aligned} \tag{5.4}$$

[6] *Hereafter, the subscripts G and S will be used to distinguish between contingent claims' prices computed under the Gaussian or stochastic volatility specifications of the Duffie and Kan (1996) model, respectively.*

[7] Its relation with the Gaussian infinitesimal generator of $\underline{X}(t)$, $\mathcal{A}_G$, is the following:

$$\mathcal{A}_G=\frac{\partial}{\partial t}+\mathcal{D}_G.$$

$tr(A)$ represents the trace of A, and $\delta[\cdot]$ is the Dirac delta function. Similarly, the Fourier transform of $G[\underline{X}(T), T; \underline{X}(t), t]$,

$$\tilde{G}[\underline{\phi}, T; \underline{X}(t), t] = \int_{\underline{X}(T)\in \Re^n} d\underline{X}(T) \frac{\exp[i\underline{\phi}' \cdot \underline{X}(T)]}{\sqrt{(2\pi)^n}} G[\underline{X}(T), T; \underline{X}(t), t], \tag{5.5}$$

with $\underline{\phi} \in \Re^n$ and $i^2 = -1$, is the solution of the following PDE

$$\mathcal{D}_G \tilde{G}[\underline{\phi}, T; \underline{x}, t] + \frac{\partial \tilde{G}[\underline{\phi}, T; \underline{x}, t]}{\partial t} - r(t)\tilde{G}[\underline{\phi}, T; \underline{x}, t] = 0, \tag{5.6}$$

$\underline{x} \in \Re^n$, subject to the boundary condition

$$\tilde{G}[\underline{\phi}, T; \underline{x}, T] = \frac{1}{\sqrt{(2\pi)^n}} \exp[i\underline{\phi}' \cdot \underline{x}]. \tag{5.7}$$

Substituting the trial solution

$$\tilde{G}[\underline{\phi}, T; \underline{X}(t), t] = \frac{1}{\sqrt{(2\pi)^n}} \exp\left[\tilde{G}1(\tau; \underline{\phi}) + \underline{\tilde{G}2}'(\tau; \underline{\phi}) \cdot \underline{X}(t)\right], \tag{5.8}$$

with $\tilde{G}1(0; \underline{\phi}) = 0$ and $\underline{\tilde{G}2}(0; \underline{\phi}) = i\underline{\phi}$, into equations (5.6) and (5.7), the last PDE can be split into one n-dimensional ODE for $\underline{\tilde{G}2}(\tau; \underline{\phi}) \in \Re^n$,

$$\frac{\partial}{\partial \tau}\underline{\tilde{G}2}'(\tau; \underline{\phi}) = -\underline{G}' + \underline{\tilde{G}2}'(\tau; \underline{\phi}) \cdot a,$$

and into another one-dimensional ODE for $\tilde{G}1(\tau; \underline{\phi}) \in \Re$,

$$\frac{\partial}{\partial \tau}\tilde{G}1(\tau; \underline{\phi}) = -f + \underline{\tilde{G}2}'(\tau; \underline{\phi}) \cdot \underline{b} + \frac{1}{2}\underline{\tilde{G}2}'(\tau; \underline{\phi}) \cdot \Theta \cdot \underline{\tilde{G}2}(\tau; \underline{\phi}).$$

The first n-dimensional ODE, subject to the terminal condition $\underline{\tilde{G}2}(0; \underline{\phi}) = i\underline{\phi}$, has the solution

$$\underline{\tilde{G}2}'(\tau; \underline{\phi}) = i\underline{\phi}' \cdot e^{a\tau} + \underline{B}'(\tau), \tag{5.9}$$

where the Gaussian duration vector $\underline{B}'(\tau)$ is given by proposition 1. To solve the last one-dimensional ODE, subject to $\tilde{G}1(0; \underline{\phi}) = 0$, result (5.9) can be used, yielding, after simplifications:

$$\begin{aligned} \tilde{G}1(\tau; \underline{\phi}) &= A(\tau) + i\underline{\phi}' \cdot a^{-1} \cdot (e^{a\tau} - I_n) \cdot \underline{b} - \frac{1}{2}\underline{\phi}' \cdot \Delta(\tau) \cdot \underline{\phi} \\ &\quad + i\underline{\phi}' \cdot \left[a^{-1} \cdot (e^{a\tau} - I_n) \cdot \Theta - \Delta(\tau)\right] \cdot (a^{-1})' \cdot \underline{G}, \end{aligned} \tag{5.10}$$

where $A(\tau)$ is computed under proposition 1, and $\Delta(\tau)$ is given by equations (3.4) or (3.7).

Substituting solutions (5.9) and (5.10) into equation (5.8), and inverting equation (5.5), yields

$$
\begin{aligned}
G\left[\underline{X}(T), T; \underline{X}(t), t\right] &= P_G(t,T)\frac{1}{(2\pi)^n}\int_{\underline{\phi}} d\underline{\phi}\exp\left[-i\underline{\phi}'\cdot\underline{X}(T)\right] \\
&\quad \exp\left[i\underline{\phi}'\cdot\underline{M}(\tau)-\frac{1}{2}\underline{\phi}'\cdot\Delta(\tau)\cdot\underline{\phi}\right].
\end{aligned}
$$

Since the second exponential inside the integral is just the characteristic function of a normal n-dimensional random variable with mean $\underline{M}(\tau)$ and variance $\Delta(\tau)$, equation (6) of Shephard (1991a) implies the closed form solution (5.1). ■

Remark 16 *The fundamental solution (5.1) corresponds simply to the product between the time-t price of a pure discount bond with maturity at time T, and the probability density function of $\underline{X}(T)$, conditional on $\underline{X}(t)$, under the equivalent martingale probability measure obtained when such zero-coupon bond is taken as the numeraire. This result is in line with corollary 2 of Jamshidian (1991), which was obtained in the context of a one-factor Gaussian term structure model.*

Remark 17 *The fundamental solution (5.1) corresponds to an Arrow-Debreu state price and not precisely to a Green's function, in the mathematical sense of the term. Nevertheless, both terms are often used interchangeably in the Finance literature.*

5.3 Series Expansion Solution for the Stochastic Volatility Specification

5.3.1 A general result

This section provides the theoretical background needed to produce approximate pricing formulae under the stochastic volatility specification, from the Gaussian "Green's function" derived earlier, and using the corresponding exact solutions already found (in Chapter three) for the deterministic volatility version of the Duffie and Kan (1996) model. The result obtained is a very general one in the sense that it can be applied to any interest rate contingent claim.

Theorem 5 *Let $V_G\left[\underline{X}(t), t\right] \in C^{2,1}(\Re^n \times [0,\infty[)$ and $V_S\left[\underline{X}(t), t\right] \in C^{2,1}(\mathbf{D} \times [0,\infty[)$ be the time-t prices, for the same contingent claim with maturity at time $T(\geq t)$, computed under the Gaussian and the stochastic volatility specifications of the Duffie and Kan (1996) model, respectively. Assuming that the terminal payoff function and the dividend yield* *process are of the same form for both* $V_S[\underline{X}(t),t]$ *and* $V_G[\underline{X}(t),t]$, *when* $\underline{X}\in\mathbf{D}$, *but identically zero if* $\underline{X}\notin\mathbf{D}$, *then:*

$$V_S[\underline{X}(t),t]=\sum_{p\geq 0}\frac{1}{2^p}V_p[\underline{X}(t),t], \tag{5.11}$$

where $V_0[\underline{X}(t),t]\equiv V_G[\underline{X}(t),t]$, *and*[8]

$$\begin{aligned} V_{p+1}[\underline{X}(t),t] &= \int_t^T dl\int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}(l)\,G[\underline{X}(l),l;\underline{X}(t),t] \\ & \quad tr\left\{\frac{\partial^2 V_p[\underline{X}(l),l]}{\partial\underline{X}(l)\,\partial\underline{X}'(l)}\cdot\Sigma\cdot W^D(l)\cdot\Sigma'\right\}, \end{aligned} \tag{5.12}$$

for $p\geq 0$, *with* $W^D(t)=diag\left\{\underline{\beta_1}'\cdot\underline{X}(t),\ldots,\underline{\beta_n}'\cdot\underline{X}(t)\right\}$.

Proof. Under the most general specification of the Duffie and Kan (1996) model, the time-t value $V_S[\underline{X}(t),t]$ of any contingent claim with terminal payoff $H[\underline{X}(T)]$ and continuous "dividend yield" $i[\underline{X}(t),t]$ is the solution of the following initial value problem:

$$-i(\underline{x},t)=\mathcal{D}_S V_S(\underline{x},t)+\frac{\partial V_S(\underline{x},t)}{\partial t}-r(t)\,V_S(\underline{x},t), \tag{5.13}$$

$\underline{x}\in\mathbf{D}$, subject to

$$V_S[\underline{X}(T),T]=H[\underline{X}(T)],\underline{X}(T)\in\mathbf{D}, \tag{5.14}$$

where $\mathcal{D}_S$ is a second-order differential operator under the "stochastic volatility" specification, i.e.

$$\begin{aligned} \mathcal{D}_S V_S(\underline{x},t) &= \frac{\partial V_S(\underline{x},t)}{\partial\underline{x}'}\cdot(a\cdot\underline{x}+\underline{b}) \\ & \quad +\frac{1}{2}tr\left[\frac{\partial^2 V_S(\underline{x},t)}{\partial\underline{x}\partial\underline{x}'}\cdot\Sigma\cdot V^D(t)\cdot\Sigma'\right]. \end{aligned} \tag{5.15}$$

In what follows it will be always assumed that both $(a\cdot\underline{x}+\underline{b})$ and $\left(\Sigma\cdot\sqrt{V^D(t)}\right)$ satisfy condition A of Duffie and Kan (1996, page 387),[9] and that the functions $H:\mathbf{D}\to\Re$ and $i:\mathbf{D}\times[0,T]\to\Re$ verify enough technical regularity conditions -namely, growth conditions in $\underline{x}$; see, for instance, equations (7.3) and (7.4) of Friedman (1964, theorems 12 and 16)- for a unique solution to exist for (5.13)-(5.14). Because

$$\Sigma\cdot V^D(t)\cdot\Sigma'=\Theta+\Sigma\cdot W^D(t)\cdot\Sigma', \tag{5.16}$$

[8] *The author wishes to thank Qiang Dai for deriving the elegant recursive relation (5.12).*
[9] In order to ensure that a strong solution exists for the SDE (2.5).

it is possible to rewrite equation (5.13) as:

$$
\begin{aligned}
& -i\left[\underline{x},t\right]-\frac{1}{2}tr\left\{\frac{\partial^2 V_S\left[\underline{x},t\right]}{\partial\underline{x}\partial\underline{x}'}\cdot\Sigma\cdot W^D\left(t\right)\cdot\Sigma'\right\} \\
= \; & \mathcal{D}_G V_S\left[\underline{x},t\right]+\frac{\partial V_S\left[\underline{x},t\right]}{\partial t}-r\left(t\right)V_S\left[\underline{x},t\right],\underline{x}\in\mathbf{D}.
\end{aligned}
\tag{5.17}
$$

On the other hand, since the Gaussian *Arrow-Debreu state price*, $G\left[\underline{X}\left(T\right),T;\underline{x},t\right]$, solves the initial value problem (5.2)-(5.3), it follows that the *exact* solution of the initial value problem (5.17)-(5.14) can be written as:

$$
\begin{aligned}
V_S\left[\underline{X}\left(t\right),t\right] = & \int_{\underline{X}(T)\in\mathbf{D}} d\underline{X}\left(T\right)G\left[\underline{X}\left(T\right),T;\underline{X}\left(t\right),t\right]H\left[\underline{X}\left(T\right)\right] \\
& +\int_t^T dl\int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}\left(l\right)G\left[\underline{X}\left(l\right),l;\underline{X}\left(t\right),t\right] \\
& \left\{i\left[\underline{X}\left(l\right),l\right]+\frac{1}{2}tr\left[\frac{\partial^2 V_S\left[\underline{X}\left(l\right),l\right]}{\partial\underline{X}\left(l\right)\partial\underline{X}'\left(l\right)}\cdot\Sigma\cdot W^D\left(l\right)\cdot\Sigma'\right]\right\}.
\end{aligned}
\tag{5.18}
$$

In fact, substituting solution (5.18) into the right-hand side of equation (5.17) and using standard differential calculus, yields

$$
\begin{aligned}
& -i\left[\underline{X}\left(t\right),t\right]-\frac{1}{2}tr\left\{\frac{\partial^2 V_S\left[\underline{X}\left(t\right),t\right]}{\partial\underline{X}\left(t\right)\partial\underline{X}'\left(t\right)}\cdot\Sigma\cdot W^D\left(t\right)\cdot\Sigma'\right\} \\
= & \int_{\underline{X}(T)\in\mathbf{D}} d\underline{X}\left(T\right)H\left[\underline{X}\left(T\right)\right] \\
& \left\{\mathcal{D}_G G\left[\underline{X}\left(T\right),T;\underline{X}\left(t\right),t\right]+\frac{\partial G\left[\underline{X}\left(T\right),T;\underline{X}\left(t\right),t\right]}{\partial t}-r\left(t\right)G\left[\underline{X}\left(T\right),T;\underline{X}\left(t\right),t\right]\right\} \\
& +\int_t^T dl\int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}\left(l\right)\left\{i\left[\underline{X}\left(l\right),l\right]+\frac{1}{2}tr\left[\frac{\partial^2 V_S\left[\underline{X}\left(l\right),l\right]}{\partial\underline{X}\left(l\right)\partial\underline{X}'\left(l\right)}\cdot\Sigma\cdot W^D\left(l\right)\cdot\Sigma'\right]\right\} \\
& \left\{\mathcal{D}_G G\left[\underline{X}\left(l\right),l;\underline{X}\left(t\right),t\right]+\frac{\partial G\left[\underline{X}\left(l\right),l;\underline{X}\left(t\right),t\right]}{\partial t}-r\left(t\right)G\left[\underline{X}\left(l\right),l;\underline{X}\left(t\right),t\right]\right\} \\
& -\int_{\underline{X}(t)\in\mathbf{D}} d\underline{X}\left(t\right)G\left[\underline{X}\left(t\right),t;\underline{X}\left(t\right),t\right] \\
& \left\{i\left[\underline{X}\left(t\right),t\right]+\frac{1}{2}tr\left[\frac{\partial^2 V_S\left[\underline{X}\left(t\right),t\right]}{\partial\underline{X}\left(t\right)\partial\underline{X}'\left(t\right)}\cdot\Sigma\cdot W^D\left(t\right)\cdot\Sigma'\right]\right\}.
\end{aligned}
$$

Using identity (5.3), the last term on the right-hand side of the last equation cancels with the corresponding left-hand side, and therefore

$$
\begin{aligned}
0 = & \int_{\underline{X}(T)\in\mathbf{D}} d\underline{X}\left(T\right)H\left[\underline{X}\left(T\right)\right] \\
& \left\{\mathcal{D}_G G\left[\underline{X}\left(T\right),T;\underline{X}\left(t\right),t\right]+\frac{\partial G\left[\underline{X}\left(T\right),T;\underline{X}\left(t\right),t\right]}{\partial t}-r\left(t\right)G\left[\underline{X}\left(T\right),T;\underline{X}\left(t\right),t\right]\right\} \\
& +\int_t^T dl\int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}\left(l\right)\left\{i\left[\underline{X}\left(l\right),l\right]+\frac{1}{2}tr\left[\frac{\partial^2 V_S\left[\underline{X}\left(l\right),l\right]}{\partial\underline{X}\left(l\right)\partial\underline{X}'\left(l\right)}\cdot\Sigma\cdot W^D\left(l\right)\cdot\Sigma'\right]\right\}
\end{aligned}
$$

$$\left\{\mathcal{D}_G G\left[\underline{X}(l), l; \underline{X}(t), t\right] + \frac{\partial G\left[\underline{X}(l), l; \underline{X}(t), t\right]}{\partial t} - r(t) G\left[\underline{X}(l), l; \underline{X}(t), t\right]\right\},$$

which is a true proposition, as implied by equation (5.2). Concerning the boundary condition, the evaluation of solution (5.18) at $t = T$,

$$V_S\left[\underline{X}(T), T\right] = \int_{\underline{X}(T)\in\mathbf{D}} d\underline{X}(T) G\left[\underline{X}(T), T; \underline{X}(T), T\right] H\left[\underline{X}(T)\right],$$

combined with definition (5.3), generates exactly the terminal payoff function (5.14).

Assuming that, when the same contingent claim is valued under the nested Gaussian specification of the Duffie and Kan (1996) model, the terminal payoff and the continuous dividend processes are still equal to zero, for $\underline{X} \notin \mathbf{D}$, and given by $H\left[\underline{X}(T)\right]$ and $i\left[\underline{X}(t), t\right]$, respectively, for $\underline{X} \in \mathbf{D}$,[10] then the corresponding time-t Gaussian price $V_G\left[\underline{X}(t), t\right]$ of the contingent claim can be obtained as the solution of

$$-i(\underline{x}, t) = \mathcal{D}_G V_G(\underline{x}, t) + \frac{\partial V_G(\underline{x}, t)}{\partial t} - r(t) V_G(\underline{x}, t), \tag{5.19}$$

$\underline{x} \in \mathbf{D}$, subject to

$$V_G\left[\underline{X}(T), T\right] = H\left[\underline{X}(T)\right], \underline{X}(T) \in \mathbf{D}. \tag{5.20}$$

And, using again results (5.2)-(5.3), such solution can be represented by an integral equation (see, for instance, Jamshidian (1991, equation 37)):

$$\begin{aligned} V_G\left[\underline{X}(t), t\right] &= \int_{\underline{X}(T)\in\mathbf{D}} d\underline{X}(T) G\left[\underline{X}(T), T; \underline{X}(t), t\right] H\left[\underline{X}(T)\right] \\ &\quad + \int_t^T dl \int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}(l) G\left[\underline{X}(l), l; \underline{X}(t), t\right] i\left[\underline{X}(l), l\right]. \end{aligned} \tag{5.21}$$

Combining equations (5.18) and (5.21),[11]

$$\begin{aligned} V_S\left[\underline{X}(t), t\right] &= V_G\left[\underline{X}(t), t\right] + \frac{1}{2}\int_t^T dl \int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}(l) \\ &\quad G\left[\underline{X}(l), l; \underline{X}(t), t\right] tr\left\{\frac{\partial^2 V_S\left[\underline{X}(l), l\right]}{\partial \underline{X}(l)\, \partial \underline{X}'(l)} \cdot \Sigma \cdot W^D(l) \cdot \Sigma'\right\}. \end{aligned} \tag{5.22}$$

[10] Because this dissertation only deals with European-style interest rate contingent claims -that is $i\left[\underline{X}(t), t\right] = 0$, $\forall t$- the only relevant assumption is the one concerning the terminal payoff function.

[11] As pointed out by Qiang Dai, the integral equation (5.22) can also be stated as

$$\begin{aligned} V_S\left[\underline{X}(t), t\right] &= V_G\left[\underline{X}(t), t\right] \\ &\quad + \int_t^T dl \int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}(l) G\left[\underline{X}(l), l; \underline{X}(t), t\right] \left[\mathcal{D}_S - \mathcal{D}_G\right] V_S\left[\underline{X}(l), l\right], \end{aligned}$$

where $\left[\mathcal{D}_S - \mathcal{D}_G\right] V_S\left[\underline{X}(l), l\right]$ can be understood as a *perturbation term*.

Finally, replacing repeatedly $V_S[\underline{X}(l), l]$ by the right-hand side of (5.22) evaluated at $t = l$ yields the series expansion (5.11)-(5.12). ■

Remark 18 *The series expansion pricing formula (5.11) is similar to equation (1.21) of Chen (1996). As Chen (1996, page 19) notices, all the terms* $\frac{1}{2}V_1, \frac{1}{4}V_2, \ldots$ *are strictly decreasing in magnitude, and therefore a good approximation should be obtained by only retaining the first few terms in the expansion.*

Remark 19 *The series expansion pricing formula (5.11) only depends on the Gaussian Arrow-Debreu state-price* G*, on the corresponding exact pricing formula under the deterministic volatility specification* V_G*, and on the "stochastic volatility parameters"* $\underline{\beta_i}$ $(i = 1, \ldots, n)$*, through matrix* W^D*.*

Remark 20 *Intuitively, equation (5.11) arises essentially because the Gaussian specification is nested into the more general stochastic volatility one. More formally, because the stochastic volatility and Gaussian instantaneous variances of the model' factors are related through the identity (5.16).*

Remark 21 *The recursive relation (5.12) shows that the* p^{th} *order approximating term,* $V_p[\underline{X}(t), t]$*, involves p time-integrals and p factor-integrals (on* **D***), and therefore its numerical computation would require the use of repeated one-dimensional or Monte Carlo integration. Next sections will simplify such general result by extending the integration with respect to the state variables to the all n-dimensional Euclidean space.*

5.3.2 Asymptotic properties

Next two propositions describe the limiting behavior of the general pricing solution (5.11) as the stochastic volatility model tends to its nested Gaussian specification, and when the series expansion (5.11) is truncated, while the domain of integration, in (5.12), is expanded from **D** to $\Re^n$.

Proposition 20 *The limit of the series expansion (5.11), as the perturbed parameters tend to zero, exists and is well defined:*

$$\lim_{\beta \to O_n} \sum_{p \geq 0} \frac{1}{2^p} V_p[\underline{X}(t), t] = V_G[\underline{X}(t), t], \qquad (5.23)$$

where $O_n \in \Re^{n \times n}$ *is a null matrix, and* $\beta \in \Re^{n \times n}$ *is a matrix whose* i^{th}*-column is given by vector* $\underline{\beta_i}$*.*

Proof. Because $W^D(l) \to O_n$ as $\beta \to O_n$, then $V_p[\underline{X}(t), t] \to 0$ as $\beta \to O_n$, for $p \geq 1$. ■

Remark 22 *The limit (5.23) is well behaved in the sense that* $V_G\left[\underline{X}(t), t\right]$ *is the solution of the initial value problem (5.13)-(5.14) when* $\beta = O_n$.

In practice, it is usually impossible to obtain analytically series terms of order higher than the first, and the series expansion (5.11) must be truncated, which induces a "truncation" error. Moreover, even for the first order term to be computed explicitly (that is without involving any factor-integral) it is almost always necessary to extend the integration bounds from $\mathbf{D}$ to $\Re^n$, introducing an "integration" error. The following proposition shows that the "integration" and "truncation" errors involved in the first order explicit solutions (with extended integration bounds) proposed hereafter are of order *strictly smaller* than the perturbed parameters.[12]

Proposition 21 *Under the* $\mathbb{A}_m(n)$ *canonical formulation of Dai and Singleton (1998, definition III.1), let* $\beta = \lambda\bar{\beta}$, *where* $\lambda \in \Re^+$ *is a common scale for the perturbed parameters,*

$$\bar{\beta} = \begin{bmatrix} I_{m\times m} & \bar{\beta}^{DB}_{m\times(n-m)} \\ O_{(n-m)\times m} & O_{(n-m)\times(n-m)} \end{bmatrix}$$

and $\bar{\beta}^{DB}$ *is a matrix of positive constants.*[13] *Let also the series* $\left\{\tilde{U}_p\left[\underline{X}(t), t\right], p \geq 0\right\}$ *be defined by* $\tilde{U}_0\left[\underline{X}(t), t\right] = V_G\left[\underline{X}(t), t\right]$ *and, for* $p \geq 1$,

$$\begin{aligned}\tilde{U}_p\left[\underline{X}(t), t\right] = & \int_t^T dl \int_{\underline{X}(l)\in\Re^n} d\underline{X}(l)\, G\left[\underline{X}(l), l; \underline{X}(t), t\right] \\ & tr\left\{\frac{\partial^2 \tilde{U}_{p-1}\left[\underline{X}(l), l\right]}{\partial \underline{X}(l)\,\partial \underline{X}'(l)} \cdot \Sigma \cdot \bar{W}^D(l) \cdot \Sigma'\right\},\end{aligned}$$

with $\bar{W}^D(l) = diag\left\{\underline{\bar{\beta}_1}' \cdot \underline{X}(l), \ldots, \underline{\bar{\beta}_n}' \cdot \underline{X}(l)\right\}$, *and where* $\underline{\bar{\beta}_i}$ *denotes the* i^{th}*-column of matrix* $\bar{\beta}$.

If $\left|\tilde{U}_p\left[\underline{X}(t), t\right]\right| < \infty$ *for* $p > 1$, *then for every* $c_0 \in \Re^+$ *there exists a* $\lambda_0 \in \Re^+$ *such that*

$$\left|V_S\left[\underline{X}(t), t\right] - V_G\left[\underline{X}(t), t\right] - \frac{1}{2}\tilde{V}_1\left[\underline{X}(t), t\right]\right| \leq c_0\left|\lambda\right| \qquad for\ \left|\lambda\right| \leq \lambda_0, \tag{5.24}$$

where

$$\begin{aligned}\tilde{V}_1\left[\underline{X}(t), t\right] = & \int_t^T dl \int_{\underline{X}(l)\in\Re^n} d\underline{X}(l)\, G\left[\underline{X}(l), l; \underline{X}(t), t\right] \\ & tr\left\{\frac{\partial^2 V_G\left[\underline{X}(l), l\right]}{\partial \underline{X}(l)\,\partial \underline{X}'(l)} \cdot \Sigma \cdot W^D(l) \cdot \Sigma'\right\}.\end{aligned}$$

[12]The author wishes to thank Qiang Dai for showing how to generalize proposition 21 from the more restrictive $\mathbb{A}_n(n)$ specification to the more general $\mathbb{A}_m(n)$ canonical form.

[13]*Dai and Singleton (1998, definition III.1) normalize* λ *to unity.*

Proof. See appendix 5.7.1. ■

Remark 23 *The cast of proposition 21 under the Dai and Singleton (1998) canonical form does not represent any loss of generality because any exponential-affine model already proposed in the literature can always be nested under the* $\mathbb{A}_m(n)$ *specification, through an appropriate invariant transformation.*

Remark 24 *Along the same lines, it can also be shown that, without approximating the integration domain, the asymptotic "truncation" error for a fixed partial sum of the series (5.11) would be of the order of the first omitted term, that is:*

$$V_S\left[\underline{X}(t), t\right] - \sum_{p=0}^{k} \frac{1}{2^p} V_p\left[\underline{X}(t), t\right] = O\left(\lambda^{k+1}\right).$$

However, if one is to go beyond the first order approximation term, attention should also be paid for the corresponding "integration" error.

5.3.3 Invariant affine transformations and nested models

Before actually applying Theorem 5, it is usually necessary to perform an affine invariant transformation (along the lines of Dai and Singleton (1998)), in order to ensure: *i*) the existence of a *strong* solution for the stochastic differential equation (3.1), which is satisfied by the state vector when zeroing off the parameters $\underline{\beta_i}$ $(i = 1, \ldots, n)$;[14] and *ii*) that the nested Gaussian specification is close enough to the general stochastic volatility one.[15]

In order to illustrate the analysis, let us consider the stochastic volatility specification defined by equations (2.2) and (2.5). The problem is that if one tries to apply directly Theorem 5 to such stochastic volatility specification, by simply imposing that $\underline{\beta_i} = \underline{0}$ $(for\ i = 1, \ldots, n)$, in some cases, the resulting Gaussian nested formulation that is obtained is too far apart from the original general stochastic volatility model. Moreover, if $\alpha_i \leq 0$ for some i, then Theorem 5 can not even be used.

However, by redefining the vector of state variables through an invariant affine transformation

$$\underline{\tilde{X}}(t) = \underline{X}(t) - \underline{u}, \tag{5.25}$$

where $\underline{\tilde{X}}(t), \underline{u} \in \Re^n$, and applying Itô's lemma, an exactly equivalent[16] stochastic volatility

[14] For instance, it is impossible, *a priori*, to nest a Gaussian specification into a multifactor CIR model, and thus it would seem impossible to apply Theorem 5 to such stochastic volatility formulation. It will be shown shortly that this is not the case.

[15] The closer is $V_G\left[\underline{X}(t), t\right]$ to $V_S\left[\underline{X}(t), t\right]$, the less important should be the neglected approximating terms $\frac{1}{2^p} V_p\left[\underline{X}(t), t\right]$, $p > k$, where k is the order of a truncated series (5.11). In this Chapter $k = 1$.

[16] In the sense that all interest rate contingent claims' prices and price probability distributions remain unchanged.

formulation follows:

$$r(t) = \tilde{f} + \underline{G}' \cdot \underline{\tilde{X}}(t), \tag{5.26}$$

with

$$\tilde{f} = f + \underline{G}' \cdot \underline{u}, \tag{5.27}$$

and

$$d\underline{\tilde{X}}(t) = \left[a \cdot \underline{\tilde{X}}(t) + \underline{\tilde{b}}\right] dt + \Sigma \cdot \sqrt{\tilde{V}^D(t)} \cdot d\underline{W}^Q(t), \tag{5.28}$$

where

$$\underline{\tilde{b}} = a \cdot \underline{u} + \underline{b}, \tag{5.29}$$

$$\sqrt{\tilde{V}^D(t)} = diag\left\{\sqrt{\tilde{v}_1(t)}, \ldots, \sqrt{\tilde{v}_n(t)}\right\}, \tag{5.30}$$

$$\tilde{v}_i(t) = \tilde{\alpha}_i + \underline{\beta_i}' \cdot \underline{\tilde{X}}(t), \tag{5.31}$$

and

$$\tilde{\alpha}_i = \alpha_i + \underline{\beta_i}' \cdot \underline{u}, \quad i = 1, ..., n. \tag{5.32}$$

The advantage of this transformed stochastic volatility specification is that $\underline{u}$ can be defined in such a way that Theorem 5 is applicable (i.e. $\tilde{\alpha}_i > 0$ for all i), and that the Gaussian nested specification, obtained with $\underline{\beta_i} = \underline{0}$ (for $i = 1, \ldots, n$), is close enough to the more general stochastic volatility one.

Alternative transformations, distinguished by different definitions of $\underline{u} \in \Re^n$, will be used for the numerical examples presented in this Chapter.[17] The preferred transformation consists in matching the first two time-t conditional moments of the new state vector (evaluated at the maturity date of the derivative under valuation, which is denominated by $T\,(\geq t)$, where t is the current pricing date), between the nested and the general specifications of the Duffie and Kan (1996) model. In appendix 5.7.2 it is shown that, no matter how $\underline{u}$ is defined, the conditional mean of $\underline{\tilde{X}}(T)$ is always the same for both Gaussian and stochastic volatility specifications. Furthermore, it is also shown that the transformation

$$\underline{u} = \underline{X}(t) \tag{5.33}$$

approximates the conditional Gaussian and stochastic volatility covariance matrices of $\underline{\tilde{X}}(T)$, at least for short maturity derivatives. This is precisely the same type of transformation as taken by Leblanc and Scaillet (1998, page 360) in order to ensure that the

[17] Although all pricing formulae are stated under the stochastic volatility specification (2.2)-(2.5), if an affine invariant transformation is used, it is understood that $\underline{X}(t)$, f, $\underline{b}$, and α_i are implicitly replaced by $\underline{\tilde{X}}(t)$, $\tilde{f}$, $\underline{\tilde{b}}$, and $\tilde{\alpha}_i$, respectively.

stationary distributions (under Q) of the state variables, for both general and nested models, have the same first two moments. Additional transformations of the form

$$\underline{u} = \eta \underline{X}(t), \tag{5.34}$$

where $\eta \in \Re$ but $\eta \neq 1$, and

$$\underline{u} = -a^{-1} \cdot \underline{b}, \tag{5.35}$$

will be also considered. In the last case (equation (5.35)), the unconditional mean of the state vector is used to minimize the stochastic volatility effects that can arise from the drift of the state process. On average, the numerical experiments presented in this Chapter will suggest that transformation (5.33) yields the lowest pricing errors for the proposed first order approximations.

Although other candidates for nested specifications exist (such as the multifactor CIR model or the three-factor Chen (1996) benchmark model), the Gaussian Langetieg (1980) specification was selected as the *bare* model from which each pricing solution for the *full* Duffie and Kan (1996) model is expanded, for two reasons:

i) Firstly, because the chosen nested specification must possess analytically tractable closed-form pricing solutions in order to yield explicit first order approximating terms. In other words, the chosen nested Gaussian specification provides an analytical solution for Arrow-Debreu state prices, which will allow all factor-integrals to be transformed into expectations with respect to a Gaussian kernel. It is exactly this feature that enables the first order approximating term $V_1 [\underline{X}(t), t]$ just to involve one time-integral (no matter the dimension of the interest rate model under consideration).[18]

ii) Secondly, and as shown in appendix 5.7.2, the selected nested model possesses the advantage that its first conditional moment for the state vector is automatically equal to the one given by the general stochastic volatility model.

5.4 Pricing of Exponential-Affine Derivatives

5.4.1 Explicit stochastic volatility approximation

When the Gaussian price of the interest rate contingent claim under valuation can be expressed as an exponential-affine function of the vector of state variables, the general

[18] The price to pay for such simplicity is that perhaps another *bare* model could provide a better zeroth order approximation. However, the computation of the corresponding first order approximating term would be too time-consuming for practical purposes.

stochastic volatility valuation equation (5.11) can be easily converted into a first order approximation that is "explicit" in the sense that it does not involve any factor-integral. Corollary 3 proposes a first order approximate and analytical pricing solution (only involving one time-integral) for exponential-affine derivatives, by extending the bounds of integration, in equation (5.12) and for $p = 0$, to the n-dimensional Euclidean space. It also provides bounds for the approximation error involved in extending the domain of integration, and contains the exact analytical solution of the first order approximating term for the univariate case ($n = 1$).

Corollary 3 *Under the assumptions of Theorem 5, let*

$$V_G\left[\underline{X}(t), t\right] = \exp\left[\varphi(t,T) + \underline{\psi}'(t,T) \cdot \underline{X}(t)\right], \tag{5.36}$$

with $\varphi(t,T) \in \Re$ and $\underline{\psi}(t,T) \in \Re^n$, denote the time-$t$ price of a contingent claim computed under the Gaussian version of the Duffie and Kan (1996) model.

1. *Approximating $\mathbf{D}$ by $\Re^n$, a first order analytical stochastic volatility approximate solution is obtained from (5.11), with*[19]

$$\begin{aligned} V_1\left[\underline{X}(t), t\right] \cong & \int_t^T dl P_G(t,l) \exp\left[\varphi(l,T) + \frac{1}{2}\underline{\psi}'(l,T) \cdot \Delta(l-t)\right. \\ & \left.\cdot \underline{\psi}(l,T) + \underline{\psi}'(l,T) \cdot \underline{M}(l-t)\right] \\ & \left[\sum_{k=1}^n \left(\underline{\psi}'(l,T) \cdot \underline{\varepsilon_k}\right)^2 \underline{\beta_k}'\right] \cdot \left[\Delta(l-t) \cdot \underline{\psi}(l,T) + \underline{M}(l-t)\right] \end{aligned} \tag{5.37}$$

and where $\underline{\varepsilon_k}$ is the k^{th} column of matrix Σ.

2. *For the unidimensional case ($n = 1$), the integration over $\mathbf{D}$ can be solved analytically, and the following solution becomes exact:*

$$\begin{aligned} V_1\left[\underline{X}(t), t\right] = & \int_t^T dl P_G(t,l) \exp\left[\varphi(l,T) + \frac{1}{2}\underline{\psi}'(l,T) \cdot \Delta(l-t)\right. \\ & \left.\cdot \underline{\psi}(l,T) + \underline{\psi}'(l,T) \cdot \underline{M}(l-t)\right] \sum_{k=1}^n \left[\underline{\psi}'(l,T) \cdot \underline{\varepsilon_k}\right]^2 \\ & \left\{\sqrt{\frac{\underline{\beta_k}' \cdot \Delta(l-t) \cdot \underline{\beta_k}}{2\pi}} e^{-\frac{1}{2}\frac{\left(\alpha_k + \underline{\beta_k}' \Delta(l-t) \underline{\mu}(t)\right)^2}{\underline{\beta_k}' \Delta(l-t) \underline{\beta_k}}}\right. \end{aligned} \tag{5.38}$$

[19] *It can be easily checked that the p^{th} order ($p > 1$) approximating term is still exponential-affine, modulo a p^{th}-degree polynomial pre-factor. For higher accuracy, it can be computed analytically (up to a numerical p-dimensional integration over time). However, the examples presented in this Chapter suggest that a first order "explicit" approximation should be enough for the valuation of simple "exponential-affine" derivatives.*

$$+\left[\underline{\beta_k}'\cdot\Delta(l-t)\cdot\underline{\mu}(l)\right]\Phi\left[\frac{\alpha_k+\underline{\beta_k}'\cdot\Delta(l-t)\cdot\underline{\mu}(l)}{\sqrt{\underline{\beta_k}'\cdot\Delta(l-t)\cdot\underline{\beta_k}}}\right]\Bigg\},$$

with

$$\underline{\mu}(l)=\underline{\psi}(l,T)+\Delta^{-1}(l-t)\cdot\underline{M}(l-t)\,, \tag{5.39}$$

and where Φ represents the cumulative density function of the univariate standard normal distribution.

3. *The exact value of the first order approximation term can be bounded from above and from bellow, using the following inequalities:*[20]

$$\begin{aligned}V_1\left[\underline{X}(t),t\right]\leq\quad&\int_t^T dlP_G(t,l)\exp\Big[\varphi(l,T)+\frac{1}{2}\underline{\psi}'(l,T)\cdot\Delta(l-t)\\&\cdot\underline{\psi}(l,T)+\underline{\psi}'(l,T)\cdot\underline{M}(l-t)\Big]\sum_{k=1}^{n}\left[\underline{\psi}'(l,T)\cdot\underline{\varepsilon_k}\right]^2\\&\sqrt{\underline{\beta_k}'\cdot\left[\Delta(l-t)+\Delta(l-t)\cdot\underline{\mu}(l)\cdot\underline{\mu}'(l)\cdot\Delta(l-t)\right]\cdot\underline{\beta_k}}\\&\prod_{j=1}^{n}\left\{\Phi\left[\frac{\alpha_j+\underline{\beta_j}'\cdot\Delta(l-t)\cdot\underline{\mu}(l)}{\sqrt{\underline{\beta_j}'\cdot\Delta(l-t)\cdot\underline{\beta_j}}}\right]\right\}^{\frac{1}{2^j+1}},\end{aligned} \tag{5.40}$$

$$\begin{aligned}V_1\left[\underline{X}(t),t\right]\geq\quad&-\int_t^T dlP_G(t,l)\exp\Big[\varphi(l,T)+\frac{1}{2}\underline{\psi}'(l,T)\cdot\Delta(l-t)\\&\cdot\underline{\psi}(l,T)+\underline{\psi}'(l,T)\cdot\underline{M}(l-t)\Big]\sum_{k=1}^{n}\left[\underline{\psi}'(l,T)\cdot\underline{\varepsilon_k}\right]^2\alpha_k\\&\prod_{j=1}^{n}\Phi\left[\frac{\alpha_j+\underline{\beta_j}'\cdot\Delta(l-t)\cdot\underline{\mu}(l)}{\sqrt{\underline{\beta_j}'\cdot\Delta(l-t)\cdot\underline{\beta_j}}}\right].\end{aligned} \tag{5.41}$$

Proof. In order to eliminate the factor-integral from equation (5.12) for $p=0$, the first order approximating term will be represented as an expectation with respect to a Gaussian kernel, and then such expectation will be computed explicitly. Because equation (5.36) implies that

$$\begin{aligned}&tr\left\{\frac{\partial^2V_G\left[\underline{X}(l),l\right]}{\partial\underline{X}(l)\,\partial\underline{X}'(l)}\cdot\Sigma\cdot W^D(l)\cdot\Sigma'\right\}\\=\quad&V_G\left[\underline{X}(l),l\right]\left[\sum_{k=1}^{n}\left(\underline{\psi}'(l,T)\cdot\underline{\varepsilon_k}\right)^2\underline{\beta_k}'\right]\cdot\underline{X}(l)\,,\end{aligned}$$

[20] *These loose bounds are used in the examples presented in this section simply to emphasize that the approximation error involved in assuming $\mathbf{D}=\Re^n$ is negligeable. In proposition 21, such error has already been shown to be of smaller (asymptotic) order than the perturbed parameters.*

equation (5.12) yields, for $p = 0$, the following functional form for the first order approximating term:

$$V_1\left[\underline{X}(t),t\right] = \int_t^T dl \int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}(l)\, G\left[\underline{X}(l),l;\underline{X}(t),t\right] V_G\left[\underline{X}(l),l\right] \sum_{k=1}^{n}\left(\underline{\psi}'(l,T)\cdot\underline{\varepsilon_k}\right)^2 \underline{\beta_k}'\cdot\underline{X}(l). \tag{5.42}$$

Approximating $\mathbf{D}$ by $\Re^n$, using the analytical solution (5.1) for the Gaussian Arrow-Debreu prices, and rearranging terms:

$$\begin{aligned} V_1\left[\underline{X}(t),t\right] \cong & \int_t^T dl P_G(t,l)\exp\left[\varphi(l,T)-\frac{1}{2}\underline{M}'(l-t)\cdot\Delta^{-1}(l-t)\right. \\ & \left.\cdot\underline{M}(l-t)+\frac{1}{2}\underline{\mu}'(l)\cdot\Delta(l-t)\cdot\underline{\mu}(l)\right] \\ & \int_{\underline{X}(l)\in\Re^n} d\underline{X}(l)\frac{\left[\sum_{k=1}^{n}\left(\underline{\psi}'(l,T)\cdot\underline{\varepsilon_k}\right)^2\underline{\beta_k}'\right]\cdot\underline{X}(l)}{\sqrt{(2\pi)^n\left|\Delta(l-t)\right|}} \\ & \exp\left\{-\frac{1}{2}\left[\underline{X}(l)-\Delta(l-t)\cdot\underline{\mu}(l)\right]'\cdot\Delta^{-1}(l-t)\right. \\ & \left.\cdot\left[\underline{X}(l)-\Delta(l-t)\cdot\underline{\mu}(l)\right]\right\}. \end{aligned}$$

The factor-integral contained in the last expression for the first order approximating term can be interpreted as the expectation of the random variable $\left[\sum_{k=1}^{n}\left(\underline{\psi}'(l,T)\cdot\underline{\varepsilon_k}\right)^2\underline{\beta_k}'\right]\cdot \underline{X}(l)$, conditional on $\underline{X}(t)$, under some equivalent probability measure with respect to which $\underline{X}(l)$ is normally distributed with mean $\Delta(l-t)\cdot\underline{\mu}(l)$ and covariance $\Delta(l-t)$, i.e. $\underline{X}(l) \frown N^n\left(\Delta(l-t)\cdot\underline{\mu}(l),\Delta(l-t)\right)$. Computing the expectation explicitly,

$$\begin{aligned} V_1\left[\underline{X}(t),t\right] \cong & \int_t^T dl P_G(t,l)\exp\left[\varphi(l,T)-\frac{1}{2}\underline{M}'(l-t)\cdot\Delta^{-1}(l-t)\right. \\ & \left.\cdot\underline{M}(l-t)+\frac{1}{2}\underline{\mu}'(l)\cdot\Delta(l-t)\cdot\underline{\mu}(l)\right] \\ & \left[\sum_{k=1}^{n}\left(\underline{\psi}'(l,T)\cdot\underline{\varepsilon_k}\right)^2\underline{\beta_k}'\right]\cdot\Delta(l-t)\cdot\underline{\mu}(l), \end{aligned}$$

and simplifying terms, the factor-integral independent solution (5.37) arises.

Items 2 and 3 of Corollary 3 are derived in appendix 5.7.3. ■

Remark 25 *Because the exact analytical solution (5.38) is only valid for one-factor models and the focus of this dissertation is on multifactor frameworks, the approximate solution (5.37) will be used hereafter. In fact, the numerical experiments implemented in this section show that the pricing errors resulting from extending* $\mathbf{D}$ *to* $\Re^n$ *in computing* $V_1\left[\underline{X}(t),t\right]$ *are small enough to be neglected.*

Remark 26 *The "explicit" approximation (5.37) is very fast to implement since it only involves one time-integral, and can be easily computed using, for example, Romberg's integration method (on a closed interval).*

For the remaining of this section Corollary 3 will be specialized for different types of "exponential-affine" interest rate contingent claims, by nesting each Gaussian price into formula (5.36).

5.4.2 Bonds, FRAs and IRSs

A first order explicit approximation

The following proposition offers a first order approximate explicit solution for the price of a zero-coupon bond.

Proposition 22 *Under the stochastic volatility specification of the Duffie and Kan (1996) model, the time-t price $P_S(t,T)$ of a default-free pure discount bond with maturity at time T $(T \geq t)$ can be approximated by the following first order solution:*

$$P_S(t,T) \cong P_G(t,T) + \frac{1}{2} V_1 \left[\underline{X}(t), t\right], \tag{5.43}$$

where $P_G(t,T)$ is the corresponding exact Gaussian bond price computed under proposition 1, and $V_1\left[\underline{X}(t), t\right]$ is given by equation (5.37) with $\varphi(l,T) = A(T-l)$, and $\underline{\psi}(l,T) = \underline{B}(T-l)$.

Proof. This result simply follows from Corollary 3, by comparing equations (2.1) and (5.36). ■

The analytical results obtained so far in this section can be further used to value all interest rate contingent claims whose price can be decomposed into a portfolio of pure discount bonds (as it is the case, for instance, of a coupon-bearing bond).

Moreover, under the Duffie and Singleton (1997) assumption of symmetric counterparty credit risk, proposition 22 can also be used to value forward rate agreements and interest rate swaps. Following, for instance, Baxter and Rennie (1996, section 5.6), the time-t fixed rate corresponding to a zero FRA value, under the stochastic volatility Duffie and Kan (1996) model, is equal to the forward interest rate

$$\frac{1}{t_2 - t_1}\left[\frac{P_S(t,t_1)}{P_S(t,t_2)} - 1\right],$$

where t_1 and t_2 are the maturity dates of the FRA contract and of its underlying borrowing/lending operation, respectively $(t \leq t_1 \leq t_2)$. Similarly, the time-t fixed rate corre-

Table 5.1: Pricing of pure discount bonds and swaps using the same parameter values as in the three-factor CIR model of Schlögl and Sommer (1998, Figure 5), for different affine invariant transformations

PDB's Expiry (years)	Exact price (P_S)	Percentage Pricing Errors				
		$\underline{u} = -a^{-1} \cdot \underline{b}$		$\underline{u} = \underline{X}(t)$		
		$\frac{P_G - P_S}{P_S}$	$\frac{P_G + \frac{V_1}{2} - P_S}{P_S}$	$\frac{P_G - P_S}{P_S}$	$\frac{P_G + \frac{V_1}{2} - P_S}{P_S}$	MAPE[b]
0.5	0.970442	0.0000%	0.0000%	0.0000%	0.0000%	0.0001%
1	0.941755	0.0000%	0.0000%	0.0000%	0.0000%	0.0007%
1.5	0.913916	-0.0001%	0.0000%	0.0000%	0.0000%	0.0023%
2	0.886900	0.0000%	0.0000%	0.0001%	0.0000%	0.0049%
2.5	0.860686	0.0000%	0.0000%	0.0001%	0.0000%	0.0090%
			...			
18	0.338659	0.1220%	-0.0035%	0.0502%	-0.0027%	0.8489%
18.5	0.328568	0.1303%	-0.0037%	0.0530%	-0.0027%	0.8963%
19	0.318776	0.1380%	-0.0044%	0.0552%	-0.0031%	0.9442%
19.5	0.309272	0.1461%	-0.0049%	0.0579%	-0.0032%	0.9929%
20	0.300049	0.1543%	-0.0053%	0.0600%	-0.0033%	1.0417%
IRS[a]	6.1045%	-0.0972%	0.0029%	-0.0398%	0.0020%	
Time	44.71s	1.05s		1.16s		

P_S is the exact stochastic volatility price, computed from equations (2.8) and (2.9).
P_G is the exact Gaussian price, computed from proposition 1.
$P_G + 0.5V_1$ is the first order approximate stochastic vol. price, given by proposition 22.
$\underline{X}(t)$ is the current state-vector, a and $\underline{b}$ are model' parameters, and $\underline{u}$ defines the affine transformation under use.
[a]20-years swap rate with semiannually compounding.
[b]$MAPE = \frac{1}{2}\frac{\max(|MinimumV_1 - EstimatedV_1|, |MaximumV_1 - EstimatedV_1|)}{P_S}$ is maximum absolute percentage error for the V_1 estimate. Maximum/Minimum V_1 are computed from Corollary 3.

sponding to a zero present value for a forward-start "plain-vanilla" IRS, under the stochastic volatility Duffie and Kan (1996) model, is equal to the forward swap rate

$$\frac{P_S(t, t_0) - P_S(t, t_m)}{\sum_{k=1}^{m}(t_k - t_{k-1})\, P_S(t, t_k)},$$

where the swap starts at time t_0, and generates m cash flows at times t_k $(k = 1, \ldots, m)$, with $t_k \geq t_0 \geq t$. All the risky zero-coupon bond prices involved in the last two formulae can be quickly computed using proposition 22.

Now, the relevant (empirical) question is to verify the accuracy of the proposed first order approximation. That is to test whether the approximating terms of order higher than the first are small enough to be neglected, as predicted before.

Examples

Table 5.1 prices (unit face value) pure discount bonds and a 20-year swap rate (with semi-annually compounding), for different affine invariant transformations, using the three-factor

CIR model of Schlögl and Sommer (1998, Figure 5) where:

$$
\begin{aligned}
f &= 0, \underline{G} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}', \underline{X}(t) = 0.02 \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}', a = diag\{-0.1, -0.15, -0.2\} \\
\underline{b} &= \begin{bmatrix} 0.002607 & 0.003 & 0.003426 \end{bmatrix}', \Sigma = diag\{0.03, 0.04, 0.05\}, \underline{\alpha} = \underline{0}, \beta = I_3,
\end{aligned}
$$

being $\underline{\alpha} \in \Re^n$ a vector with α_i as its i^{th} element. Exact stochastic volatility zero-coupon bond prices are computed using the exact numerical solution of equations (2.8) and (2.9), through an adaptive stepsize fifth-order Runge-Kutta method (for $\underline{B}(\tau)$), and Romberg's integration (for $A(\tau)$).[21] Approximate stochastic volatility prices are given by the "explicit" first order formula (5.43), where its second term is implemented using Romberg's integration method on a closed interval, and the associated Gaussian pure discount bond prices are computed from proposition 1. For each affine transformation, instead of the zero and first order approximate prices, the corresponding percentage pricing errors are presented. Throughout this Chapter, the CPU time is always shown in seconds (except if stated otherwise), and all computations are made running Pascal programs on a Pentium 233MHz with 32MB of RAM memory.

For the transformation (5.33), the lower and upper bounds of the first order approximating term are computed according to equations (5.41) and (5.40), respectively. Based on these, the maximum absolute percentage error arising from assuming that $\mathbf{D} = \Re^n$ in computing $V_1[\underline{X}(t), t]$ is presented.

The overall conclusion is that the proposed approximation is very accurate: all invariant transformations produce pricing errors for the IRS smaller than a tenth of a basis point. In other words, the neglected approximating terms (of order higher than the first) seem to constitute an irrelevant part of the (stochastic volatility) pure discount bond price. Moreover, the use of the approximate formula is also faster since it avoids the solution of the Riccati equations (2.8) through Runge-Kutta methods: the swap rate was forty times faster to compute using the explicit first order approximation! Notice also that the first order approximation is always more accurate than the zeroth order one.

Table 5.2 presents the same empirical analysis as before, but using the following $\mathbb{A}_2(3)$

[21] Although the exact analytical solution of Chen and Scott (1995b, page 54) is also available, in general, the stochastic volatility Duffie and Kan (1996) model does not produce exact closed-form solutions. Therefore, the efficiency of the explicit first order approximations shall be compared against the exact numerical solutions.

Table 5.2: Pricing of pure discount bonds and swaps using the parameters corresponding to an $\mathbb{A}_2(3)$ model, for different affine invariant transformations

PDB's Expiry (years)	Exact price (P_S)	Percentage Pricing Errors				
		$\underline{u} = 0.5\underline{X}(t)$		$\underline{u} = \underline{X}(t)$		
		$\frac{P_G - P_S}{P_S}$	$\frac{P_G + \frac{V_1}{2} - P_S}{P_S}$	$\frac{P_G - P_S}{P_S}$	$\frac{P_G + \frac{V_1}{2} - P_S}{P_S}$	MAPE[b]
0.5	0.975063	0.0000%	0.0000%	0.0000%	0.0000%	0.0000%
1	0.947129	-0.0004%	0.0000%	0.0002%	0.0000%	0.0005%
1.5	0.919966	-0.0018%	0.0000%	0.0005%	0.0000%	0.0019%
2	0.893223	-0.0046%	0.0000%	0.0013%	0.0000%	0.0050%
2.5	0.866454	-0.0092%	0.0000%	0.0022%	0.0000%	0.0097%
			...			
18	0.172002	-10.7382%	0.9599%	-5.4174%	0.4076%	6.3996%
18.5	0.159963	-11.8327%	1.0796%	-6.0239%	0.4639%	7.1025%
19	0.148605	-12.9962%	1.2052%	-6.6727%	0.5235%	7.8528%
19.5	0.137907	-14.2292%	1.3364%	-7.3651%	0.5866%	8.6517%
20	0.127847	-15.5326%	1.4697%	-8.1011%	0.6516%	9.4977%
IRS[a]	8.8908%	3.9567%	-0.3314%	1.9450%	-0.1424%	
Time	474.06s	1.53s		2.92s		

P_S is the exact stochastic volatility price, computed from equations (2.8) and (2.9).
P_G is the exact Gaussian price, computed from proposition 1.
$P_G + 0.5V_1$ is the first order approximate stochastic vol. price, given by proposition 22.
$\underline{X}(t)$ is the current state-vector and $\underline{u}$ defines the affine transformation under use.
[a]20-years swap rate with semiannually compounding.
[b]$MAPE = \frac{1}{2}\frac{\max(|MinimumV_1 - EstimatedV_1|,|MaximumV_1 - EstimatedV_1|)}{P_S}$ is maximum absolute percentage error of the V_1 estimate. Maximum/Minimum V_1 are computed from Corollary 3.

model:[22]

$$f = -0.00394, \underline{G} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}', a = \begin{bmatrix} -2.78 & -0.41238 & 1386.106 \\ 0 & 0.02138 & 39.9 \\ 0 & 0.000741 & -2.2328 \end{bmatrix}, \underline{\alpha} = \underline{0},$$

$$b = \begin{bmatrix} -6.1e-18 \\ 0.002445 \\ 9.49e-05 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & -1 & -252 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \beta = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.00237 & 0 \\ 1 & 0 & 6.45e-05 \end{bmatrix},$$

where the state variables' values, $\underline{X}(t) = \begin{bmatrix} 0.01 & 0.03 & 0.0001 \end{bmatrix}'$, were defined in order to have an upward slopping yield curve (the spot rates with continuous compounding vary from 4.564%, for three months, to 10.285%, for 20 years). The first order explicit approximation is still fast to implement and accurate (although the pricing errors are higher for longer maturities). As before, the pricing errors resulting from extending $\mathbf{D}$ to $\Re^n$ in computing $V_1[\underline{X}(t), t]$ are small (at least for short maturities).

[22]This model specification was borrowed from a previous version (Table IV) of the Dai and Singleton (1998) paper.

5.4.3 Bond futures

An exact pricing solution

Under the stochastic volatility specification, and as Duffie and Kan (1996) did for pure discount bonds, it is also possible to find an exponential-affine exact pricing formula for futures on zero-coupon bonds that involves maturity-dependent functions satisfying Riccati differential equations. Once again, the hypothesis of continuous marking to market will be assumed, whenever futures contracts are involved.

Proposition 23 *Under the stochastic volatility specification of the Duffie and Kan (1996) model, the time-t price, $FP_S(t, T_f, T_1)$, of a futures contract for delivery at time T_f and on a pure discount bond with maturity at time T_1 ($t \leq T_f \leq T_1$) is equal to*

$$FP_S(t, T_f, T_1) = \frac{P_S(t, T_1)}{P_S(t, T_f)} \exp\left[C(t, T_f, T_1) + \underline{D}'(t, T_f, T_1) \cdot \underline{X}(t)\right], \tag{5.44}$$

where $\underline{D}(t, T_f, T_1) \in \Re^n$ is the solution of

$$\begin{aligned} \frac{\partial \underline{D}'(t, T_f, T_1)}{\partial t} = & -\underline{D}'(t, T_f, T_1) \cdot a \\ & -\sum_{k=1}^{n} \underline{B}'(T_f - t) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \cdot \left[\underline{B}(T_f - t) - \underline{B}(T_1 - t)\right] \underline{\beta_k}' \\ & -\frac{1}{2} \sum_{k=1}^{n} \underline{D}'(t, T_f, T_1) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \\ & \cdot \left[2\underline{B}(T_1 - t) - 2\underline{B}(T_f - t) + \underline{D}(t, T_f, T_1)\right] \underline{\beta_k}' \end{aligned} \tag{5.45}$$

subject to $\underline{D}(T_f, T_f, T_1) = \underline{0}$, and $C(t, T_f, T_1) \in \Re$ is obtained from

$$\begin{aligned} \frac{\partial C(t, T_f, T_1)}{\partial t} = & -\underline{D}'(t, T_f, T_1) \cdot \underline{b} \\ & -\sum_{k=1}^{n} \underline{B}'(T_f - t) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \cdot \left[\underline{B}(T_f - t) - \underline{B}(T_1 - t)\right] \alpha_k \\ & -\frac{1}{2} \sum_{k=1}^{n} \underline{D}'(t, T_f, T_1) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \\ & \cdot \left[2\underline{B}(T_1 - t) - 2\underline{B}(T_f - t) + \underline{D}(t, T_f, T_1)\right] \alpha_k \end{aligned} \tag{5.46}$$

subject to $C(T_f, T_f, T_1) = 0$.

Proof. See appendix 5.7.4. ■

Remark 27 *Equation (5.45) can be solved numerically through Runge-Kutta methods, while equation (5.46) seems to only require univariate integration algorithms. However, both*

(5.45) and (5.46) involve the solution of Riccati equations similar to (2.8) at each evaluation point, since they are both functions of the duration vectors $\underline{B}(T_f - t)$ and $\underline{B}(T_1 - t)$. Therefore, the following explicit approximation should provide significant efficiency gains.

A first order explicit approximation

Next proposition proposes an approximate stochastic volatility pricing solution that is easier to implement than the exact numerical one offered by proposition 23.

Proposition 24 *Under the stochastic volatility specification of the Duffie and Kan (1996) model, the time-t price, $FP_S(t, T_f, T_1)$, of a futures contract for delivery at time T_f and on a pure discount bond with maturity at time T_1 $(t \leq T_f \leq T_1)$ can be approximated by*

$$FP_S(t, T_f, T_1) \cong FP_G(t, T_f, T_1) + \frac{1}{2} V_1 \left[\underline{X}(t), t \right], \tag{5.47}$$

where $FP_G(t, T_f, T_1)$ is computed from proposition 6, and $V_1 \left[\underline{X}(t), t \right]$ has the "explicit" solution given by equation (5.37) but with $T = T_f$, $\varphi(l, T) = A(T_1 - l) - A(T_f - l) - J(l)$, and $\underline{\psi}(l, T) = \underline{B}(T_1 - l) - \underline{B}(T_f - l)$.

Proof. This result follows from Corollary 3, by comparing equations (3.65) and (5.36). ■

To value futures on coupon-bearing bonds, and following expression (3.69), it is just necessary to consider the summation of the prices of futures on zero-coupon bonds with maturity dates corresponding to the moments where cash flows are paid by the underlying coupon bond, and with contract sizes equal to the value of such cash flows. That is[23]

$$FS_S(t, T_f) = \sum_{i=1}^{N_f} k_i FP_S(t, T_f, T_i),$$

where $T_f < T_i$ $(\forall i)$, $FS_S(t, T_f)$ represents the stochastic volatility time-t price of a futures contract for delivery at time T_f, on a coupon-bearing bond paying N_f cash flows k_i $(i = 1, \ldots, N_f)$ from the futures' expiry date and until the bond's maturity date (T_{N_f}), and $FP_S(t, T_f, T_i)$ is computed under proposition 24.

Example

Table 5.3 values futures with a maturity of 6 months on (unit face value) zero-coupon bonds with maturities ranging from 1 year to 20.5 years, using the three-factor CIR model of Schlögl and Sommer (1998, Figure 5), as well as a futures contract with a maturity of 6

[23] Ignoring the existence of quality and/or timing options.

Table 5.3: Pricing of bond futures with a maturity of 6 months using the same parameter values as in the three-factor CIR model of Schlögl and Sommer (1998, Figure 5), for different affine invariant transformations

PDB's Expiry (years)	Exact price (FP_S)	Percentage Pricing Errors: $\underline{u}$: VAR match $\frac{FP_G - FP_S}{FP_S}$	$\frac{FP_G + \frac{V_1}{2} - FP_S}{FP_S}$	$\underline{u} = \underline{X}(t)$ $\frac{FP_G - FP_S}{FP_S}$	$\frac{FP_G + \frac{V_1}{2} - FP_S}{FP_S}$	MAPE[b]
1	0.970434	0.0000%	0.0000%	0.0000%	0.0000%	0.0003%
1.5	0.941742	0.0000%	0.0000%	0.0000%	0.0000%	0.0009%
2	0.913899	0.0001%	0.0001%	0.0001%	0.0001%	0.0019%
2.5	0.886883	0.0001%	0.0001%	0.0001%	0.0001%	0.0030%
3	0.860667	0.0002%	0.0002%	0.0002%	0.0002%	0.0043%
			...			
18.5	0.338553	0.0520%	0.0520%	0.0526%	0.0526%	0.0264%
19	0.328463	0.0545%	0.0542%	0.0551%	0.0548%	0.0262%
19.5	0.318670	0.0571%	0.0568%	0.0577%	0.0574%	0.0264%
20	0.309167	0.0595%	0.0592%	0.0598%	0.0598%	0.0268%
20.5	0.299944	0.0620%	0.0617%	0.0623%	0.0623%	0.0269%
FCBB[a]	121.7132	0.0265%	0.0263%	0.0268%	0.0268%	
Time	1499.02s	1.76s		1.76s		

FP_G is the exact Gaussian price, computed from proposition 6.
FP_S is the exact stochastic volatility price, computed from proposition 23.
$FP_G + 0.5V_1$ is the first order approximate stochastic vol. price, given by proposition 24.
$\underline{X}(t)$ is the current state-vector and $\underline{u}$ defines the affine transformation under use.
[a]6-month future on a benchmark bond with a maturity of 20.5 years, a semi-annual coupon of 8% per annum, and a face value of 100. Delivery options are ignored.
[b]$MAPE = \frac{1}{2}\frac{\max(|MinimumV_1 - EstimatedV_1|,|MaximumV_1 - EstimatedV_1|)}{FP_S}$ is maximum absolute percentage error for the V_1 estimate. Maximum/Minimum V_1 are given by Corollary 3.

months, on a theoretical coupon-bearing bond with a maturity of 20 years at the futures expiry date, with a semi-annual coupon rate of 8% per annum, and with a face value of 100. No provision is made for the existence of delivery options. Exact stochastic volatility futures prices were obtained from proposition 23, by using an adaptive stepsize fifth-order Runge-Kutta method for equation (5.45), and Romberg's integration method for equation (5.46). Approximate stochastic volatility futures prices were computed through the first order "explicit" solution obtained in proposition 24, and the corresponding Gaussian futures price resulted from proposition 6.

A new transformation (denominated by "VAR match") is also tested where $\underline{u}$ is defined in order for the variance of the state variables, at the futures' expiry date and conditional on the current value of the state vector, to be equal between the nested Gaussian and the multifactor CIR models.[24] As before, different affine invariant transformations produce similar

[24]Writing the multifactor CIR model, under measure Q, as $r(t) = \sum_{j=1}^{n} X_j(t)$ with

$$dX_j(t) = [k_j\theta_j - (k_j + \lambda_j) X_j(t)]\, dt + \sigma_j \sqrt{X_j(t)} dW_j^Q(t), j = 1, \ldots, n,$$

and applying the invariant transformation (5.25), it can be shown that the matching of the factor Gaussian results: the proposed first order explicit stochastic volatility approximation is still accurate and extremely fast to implement. Moreover, the pricing errors arising from approximating the integration domain in computing $V_1[\underline{X}(t), t]$ are again negligeable.

5.4.4 Short-term interest rate futures

An exact pricing solution

As in proposition 23, for futures on zero-coupon bonds, it is also possible to obtain an equivalent exact numerical solution for futures on short-term *nominal* "money-market" forward interest rates.

Proposition 25 *Under the stochastic volatility specification of the Duffie and Kan (1996) model, the time-t price, $FR_S(t, T_f, T_1)$, of a futures contract with maturity at time T_f and on the nominal interest rate for the period $(T_1 - T_f)$, with $T_1 \geq T_f \geq t$, is equal to*

$$\begin{aligned} FR_S(t, T_f, T_1) = \quad & 100\left\{1 - \frac{1}{T_1 - T_f}\left[\frac{P_S(t, T_f)}{P_S(t, T_1)}\right.\right. \\ & \left.\left. \exp\left(E(t, T_f, T_1) + \underline{F}'(t, T_f, T_1) \cdot \underline{X}(t)\right) - 1\right]\right\}, \end{aligned} \tag{5.48}$$

where $\underline{F}(t, T_f, T_1) \in \Re^n$ is the solution of

$$\begin{aligned} \frac{\partial \underline{F}'(t, T_f, T_1)}{\partial t} = \quad & -\underline{F}'(t, T_f, T_1) \cdot a \\ & -\sum_{k=1}^{n} \underline{B}'(T_1 - t) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \cdot [\underline{B}(T_1 - t) - \underline{B}(T_f - t)]\, \underline{\beta_k}' \\ & -\frac{1}{2}\sum_{k=1}^{n} \underline{F}'(t, T_f, T_1) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \\ & \cdot [2\underline{B}(T_f - t) - 2\underline{B}(T_1 - t) + \underline{F}(t, T_f, T_1)]\, \underline{\beta_k}' \end{aligned} \tag{5.49}$$

subject to $\underline{F}(T_f, T_f, T_1) = \underline{0}$, and $E(t, T_f, T_1) \in \Re$ is obtained from

$$\begin{aligned} \frac{\partial E(t, T_f, T_1)}{\partial t} = \quad & -\underline{F}'(t, T_f, T_1) \cdot \underline{b} \\ & -\sum_{k=1}^{n} \underline{B}'(T_1 - t) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \cdot [\underline{B}(T_1 - t) - \underline{B}(T_f - t)]\, \alpha_k \\ & -\frac{1}{2}\sum_{k=1}^{n} \underline{F}'(t, T_f, T_1) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \end{aligned} \tag{5.50}$$

---

and multifactor CIR variances for maturity $T\,(\geq t)$ is obtained if

$$u_j = \frac{\theta_j + [2X_j(t) - \theta_j]\, e^{-(k_j + \lambda_j)(T-t)}}{e^{-(k_j + \lambda_j)(T-t)} + 1}, j = 1, \ldots, n,$$

where u_j is the j^{th} element of vector $\underline{u}$.

$$\cdot \left[2\underline{B}(T_f - t) - 2\underline{B}(T_1 - t) + \underline{E}(t, T_f, T_1)\right] \alpha_k$$

subject to $E(T_f, T_f, T_1) = 0$.

Proof. The derivation of the above exact numerical result is identical to the proof of proposition 23. ■

A first order explicit approximation

Proposition 8 allows the next result to be extracted from Corollary 3.

Proposition 26 *Under the stochastic volatility specification of the Duffie and Kan (1996) model, the time-t price,* $FR_S(t, T_f, T_1)$*, of a futures contract with maturity at time* T_f *and on the nominal interest rate for the period* $(T_1 - T_f)$*, with* $T_1 \geq T_f \geq t$*, can be approximated by*

$$FR_S(t, T_f, T_1) \cong FR_G(t, T_f, T_1) - \frac{50}{T_1 - T_f} V_1\left[\underline{X}(t), t\right], \tag{5.51}$$

where $FR_G(t, T_f, T_1)$ *is computed from proposition 8, and* $V_1\left[\underline{X}(t), t\right]$ *has the "explicit" solution given by equation (5.37) but with* $T = T_f$*,* $\varphi(l, T) = A(T_f - l) - A(T_1 - l) + L(l)$*, and* $\underline{\psi}(l, T) = \underline{B}(T_f - l) - \underline{B}(T_1 - l)$.

Proof. This result follows from Corollary 3, by comparing equation (5.36) with $\left[1 + (T_1 - T_f)\frac{100 - FR_G(t, T_f, T_1)}{100}\right]$, where $FR_G(t, T_f, T_1)$ is given by equation (3.73). ■

Example

Table 5.4 prices three-month Eurodollar futures contracts, with maturities varying from one month to 9 years, and based on the $\mathbb{A}_1(3)_{DS}$ model of Dai and Singleton (1998, Table II), where

$$f = 0, \underline{G} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}', a = \begin{bmatrix} -0.33458 & 0 & 0 \\ 0.878876 & -0.226 & 0 \\ -9.190106 & 17.4 & -17.4 \end{bmatrix}, \underline{\alpha} = \underline{0},$$

$$\underline{b} = \begin{bmatrix} 0.005475 \\ 0.012350 \\ 0.021683 \end{bmatrix}, \Sigma = \begin{bmatrix} 0.088431 & 0 & 0 \\ 0 & 1 & -0.0943 \\ 0.377599 & -3.42 & 1 \end{bmatrix}, \beta = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and the factor' values, $\underline{X}(t) = \begin{bmatrix} 0.01 & 0.12 & 0.11 \end{bmatrix}'$, were defined in order to have a downward slopping yield curve (the spot rates with continuous compounding vary from 11.396%,

Table 5.4: Pricing of 3-month Eurodollar futures using the parameters corresponding to the $\mathbb{A}_1(3)_{DS}$ Dai and Singleton (1998, Table II) model, for different affine invariant transformations

Futures' Maturity (years)	Exact price (FR_S)	Percentage Pricing Errors $\underline{u} = -\underline{a}^{-1} \cdot \underline{b}$: $\frac{FR_G}{FR_S} - 1$	$\underline{u} = -\underline{a}^{-1} \cdot \underline{b}$: $\frac{FR_G - \frac{50V_1}{0.25}}{FR_S} - 1$	$\underline{u} = \underline{X}(t)$: $\frac{FR_G}{FR_S} - 1$	$\underline{u} = \underline{X}(t)$: $\frac{FR_G - \frac{50V_1}{0.25}}{FR_S} - 1$	MAPE[a]
1/12	88.3744	0.0008%	0.0008%	0.0000%	0.0000%	0.0001%
2/12	88.4027	0.0007%	0.0008%	-0.0001%	-0.0001%	0.0001%
3/12	88.4491	0.0007%	0.0008%	-0.0001%	-0.0001%	0.0002%
4/12	88.4981	0.0006%	0.0008%	-0.0001%	-0.0001%	0.0003%
5/12	88.5459	0.0005%	0.0007%	-0.0001%	-0.0001%	0.0004%
			...			
8	89.1890	-0.0022%	-0.0008%	0.0024%	0.0002%	0.0035%
8.25	89.1738	-0.0022%	-0.0008%	0.0026%	0.0003%	0.0035%
8.5	89.1586	-0.0022%	-0.0008%	0.0027%	0.0004%	0.0035%
8.75	89.1436	-0.0021%	-0.0008%	0.0028%	0.0004%	0.0035%
9	89.1288	-0.0021%	-0.0008%	0.0029%	0.0005%	0.0035%
Time	43393s		29s		40.97s	

FR_G is the exact Gaussian price, computed from proposition 8.
FR_S is the exact stochastic volatility price, computed from proposition 25.
$FR_G - \frac{50V_1}{0.25}$ is the first order approximate stochastic vol. price, given by proposition 26.
$\underline{X}(t)$ is the current state-vector, $\underline{a}$ and $\underline{b}$ are model' parameters, and $\underline{u}$ defines the affine transformation under use.
[a] $MAPE = \frac{50}{0.25} \frac{\max(|MinimumV_1 - EstimatedV_1|, |MaximumV_1 - EstimatedV_1|)}{FR_S}$ is maximum absolute percentage error for V_1 estimate. Maximum/Minimum V_1 are given by Corollary 3.

for 6 months, to 10.392%, for 20 years). The Gaussian prices are computed under proposition 8, while the approximate stochastic volatility ones are obtained from proposition 26. Exact numerical stochastic volatility futures prices result from proposition 25 (using a fifth-order Runge-Kutta method for equation (5.49), and Romberg's integration for expression (5.50)).

For all the invariant transformations tested, the pricing errors are almost inexistent. Again, the error induced by the extension of the integration domain from $\mathbf{D}$ to $\Re^n$ is very small, and the proposed first order approximation is much faster than the exact numerical solution as well as always more accurate than the zeroth order one.

5.5 Pricing of European Interest Rate Options

5.5.1 Explicit stochastic volatility approximation

Besides the already considered exponential-affine derivatives, it is also possible to obtain explicit first order pricing solutions for several European interest rate options, such as: options on pure discount bonds, caps and floors, yield options, and (*conventional* or *pure*) futures options on zero-coupon bonds and on short-term interest rates. Next Corollary establish the general first order explicit solution which can be applied (specialized) to any of the specific option contracts described before.

Corollary 4 *Let the time-t price of an European option, with maturity at date $T_0 (\geq t)$, and computed under the Gaussian specification of the Duffie and Kan (1996) model, be represented by:*

$$V_G[\underline{X}(t), t] = \theta q \left\{ \exp\left[U(t,\cdot) + \underline{Q}'(t,\cdot) \cdot \underline{X}(t)\right] \Phi[\theta d_1(t)] - K \exp\left[S(t,T_0) + \underline{T}'(t,T_0) \cdot \underline{X}(t)\right] \Phi[\theta d_0(t)] \right\}, \tag{5.52}$$

with

$$d_1(t) = \frac{\ln\left\{\frac{\exp[U(t,\cdot)+\underline{Q}'(t,\cdot)\cdot\underline{X}(t)]}{K\exp[S(t,T_0)+\underline{T}'(t,T_0)\cdot\underline{X}(t)]}\right\} + \frac{\sigma^2(t)}{2}}{\sigma(t)}, \tag{5.53}$$

$$d_0(t) = d_1(t) - \sigma(t), \tag{5.54}$$

$$\sigma^2(t) = \underline{Q}'(T_0,\cdot) \cdot \Delta(T_0 - t) \cdot \underline{Q}(T_0,\cdot), \tag{5.55}$$

and where $\theta \in \{-1, 1\}$, $q, K, U(t,\cdot) \in \Re$, $S(t,T_0) \in \Re$ is such that $S(T_0,T_0) = 0$, $\underline{Q}(t,\cdot) \in \Re^n$, and $\underline{T}(t,T_0) \in \Re^n$ satisfies $\underline{T}(T_0,T_0) = \underline{0}$.[25]

*Under the assumptions of Theorem 5, and approximating **D** by $\Re^n$, the corresponding price of the same option contract but for the stochastic volatility version of the Duffie and Kan (1996) model can be approximated by the first order analytical solution obtained from (5.11) with:*

$$V_1[\underline{X}(t), t] \cong q \int_t^{T_0} dl \left[V_{11}(l) + V_{12}(l) + V_{13}(l)\right]. \tag{5.56}$$

For $i = 1, 2$:

$$\begin{aligned} V_{1i}(l) = \; & \theta\left[(2-i) - K(i-1)\right] P_G(t,l) \sqrt{\frac{\left|\Omega^{-1}(l) \cdot \Psi^{-1}(l)\right|}{\left|\Delta(l-t) \cdot \Delta(T_0 - l)\right|}} \\ & \exp\left[F_i(l) - \frac{1}{2}\underline{M}'(l-t) \cdot \Delta^{-1}(l-t) \cdot \underline{M}(l-t) + \frac{1}{2}\underline{\mu c_i}'(l)\right. \\ & \cdot \Omega^{-1}(l) \cdot \underline{\mu c_i}(l) - \frac{1}{2}\underline{MC_i}'(T_0 - l) \cdot \Delta^{-1}(T_0 - l) \cdot \underline{MC_i}(T_0 - l) \\ & \left. + \frac{1}{2}\underline{N_i}'(l) \cdot \Psi^{-1}(l) \cdot \underline{N_i}(l)\right] \left\{ \left[\sum_{k=1}^{n} \left(\underline{D_i}'(l) \cdot \underline{\varepsilon_k}\right)^2 \underline{\beta_k}'\right] \right. \\ & \cdot \Omega^{-1}(l) \cdot \underline{\mu c_i}(l) \Phi\left[\theta \frac{\underline{H}'(T_0) \cdot \Psi^{-1}(l) \cdot \underline{N_i}(l) - K^*}{\sqrt{\underline{H}'(T_0) \cdot \Psi^{-1}(l) \cdot \underline{H}(T_0)}}\right] \\ & + \frac{\lambda_i(l)}{\theta} \sqrt{\frac{\underline{H}'(T_0) \cdot \Psi^{-1}(l) \cdot \underline{H}(T_0)}{2\pi}} \end{aligned} \tag{5.57}$$

---

[25] *The functions $U(t,\cdot)$ and $\underline{Q}(t,\cdot)$ may involve other maturities than just the current time, depending on the contract specifications.*

$$\exp\left[-\frac{1}{2}\frac{\left(K^{*}-\underline{H}'(T_0)\cdot\Psi^{-1}(l)\cdot\underline{N}_i(l)\right)^2}{\underline{H}'(T_0)\cdot\Psi^{-1}(l)\cdot\underline{H}(T_0)}\right]+\lambda_i(l)$$

$$\left[\underline{H}'(T_0)\cdot\Psi^{-1}(l)\cdot\underline{N}_i(l)\right]\Phi\left[\theta\frac{\underline{H}'(T_0)\cdot\Psi^{-1}(l)\cdot\underline{N}_i(l)-K^{*}}{\sqrt{\underline{H}'(T_0)\cdot\Psi^{-1}(l)\cdot\underline{H}(T_0)}}\right]\Bigg\},$$

where

$$F_i(l)=(2-i)\,U(l,\cdot)+(i-1)\,S(l,T_0),$$

$$\underline{D}_i(l)=(2-i)\,\underline{Q}(l,\cdot)+(i-1)\,\underline{T}(l,T_0),$$

$$\underline{H}(l)=\underline{Q}(l,\cdot)-\underline{T}(l,T_0),$$

$$K^{*}=\ln(K)-U(T_0,\cdot),$$

$$\begin{aligned}\underline{MC}_i(T_0-l)= &\quad a^{-1}\cdot\left[e^{a(T_0-l)}-I_n\right]\cdot\left[\underline{b}+\Theta\cdot\left(a^{-1}\right)'\cdot\underline{G}\right]\\ &-\Delta(T_0-l)\cdot\left[\left(a^{-1}\right)'\cdot\underline{G}-(2-i)\,\underline{Q}(T_0,\cdot)\right],\end{aligned}$$

$$\Omega(l)=\Delta^{-1}(l-t)+e^{a'(T_0-l)}\cdot\Delta^{-1}(T_0-l)\cdot e^{a(T_0-l)},$$

$$\begin{aligned}\underline{\mu c}_i(l)= &\quad \Delta^{-1}(l-t)\cdot\underline{M}(l-t)+\underline{D}_i(l)\\ &-e^{a'(T_0-l)}\cdot\Delta^{-1}(T_0-l)\cdot\underline{MC}_i(T_0-l),\end{aligned}$$

$$\Psi(l)=\Delta^{-1}(T_0-l)\cdot\left[I_n-e^{a(T_0-l)}\cdot\Omega^{-1}(l)\cdot e^{a'(T_0-l)}\cdot\Delta^{-1}(T_0-l)\right],$$

$$\underline{N}_i(l)=\Delta^{-1}(T_0-l)\cdot\left[\underline{MC}_i(T_0-l)+e^{a(T_0-l)}\cdot\Omega^{-1}(l)\cdot\underline{\mu c}_i(l)\right],$$

$$\lambda_i(l)=\frac{\underline{C}_i'(l)\cdot\underline{1}}{\underline{H}'(T_0)\cdot\underline{1}},$$

and

$$\underline{C}_i'(l)=\left[\sum_{k=1}^{n}\left(\underline{D}_i'(l)\cdot\underline{\varepsilon}_k\right)^2\underline{\beta}_k'\right]\cdot\Omega^{-1}(l)\cdot e^{a'(T_0-l)}\cdot\Delta^{-1}(T_0-l).$$

For $i=3$:

$$\begin{aligned}V_{13}(l)= &\quad \frac{KP_G(t,l)}{\sqrt{2\pi\sigma^2(l)}}\sqrt{\frac{|\varphi^{-1}(l)|}{|\Delta(l-t)|}}\exp\left\{S(l,T_0)-\frac{[d_0^{*}(l)]^2}{2\sigma^2(l)}\right.\\ &\left.-\frac{1}{2}\underline{M}'(l-t)\cdot\Delta^{-1}(l-t)\cdot\underline{M}(l-t)+\frac{1}{2}\underline{m}'(l)\cdot\varphi^{-1}(l)\cdot\underline{m}(l)\right\}\\ &\left\{\sum_{k=1}^{n}\left[\underline{H}'(l)\cdot\underline{\varepsilon}_k\right]^2\underline{\beta}_k'\right\}\cdot\varphi^{-1}(l)\cdot\underline{m}(l),\end{aligned} \tag{5.58}$$

where

$$d_0^*(l) = U(l,\cdot) - S(l,T_0) - \ln(K) - \frac{\sigma^2(l)}{2},$$

$$\underline{m}(l) = \Delta^{-1}(l-t)\cdot \underline{M}(l-t) + \underline{T}(l,T_0) - \frac{d_0^*(l)}{\sigma^2(l)}\underline{H}(l),$$

and

$$\varphi(l) = \Delta^{-1}(l-t) + \frac{1}{\sigma^2(l)}\underline{H}(l)\cdot \underline{H}'(l).$$

Proof. See appendix 5.7.5. ■

5.5.2 Options on pure discount bonds

A first order explicit approximation

Proposition 3 allows Corollary 4 to be specialized for the valuation of European options on pure discount bonds under the stochastic volatility specification of the Duffie and Kan (1996) model.

Proposition 27 *Under the stochastic volatility specification of the Duffie and Kan (1996) model, the time-t price of an European call on the riskless pure discount bond $P_S(t,T_1)$, with a strike price equal to K, and with maturity at time T_0 (such that $t \leq T_0 \leq T_1$) can be approximated by*

$$c_t^S[P_S(t,T_1);K;T_0] \cong c_t^G[P_G(t,T_1);K;T_0] + \frac{1}{2}V_1[\underline{X}(t),t], \tag{5.59}$$

where $c_t^G[P_G(t,T_1);K;T_0]$ is computed from equation (3.24), and $V_1[\underline{X}(t),t]$ has the "explicit" solution given by equation (5.56) but with $q=1$, $U(t,\cdot)=A(T_1-t)$, $\underline{Q}(t,\cdot)=\underline{B}(T_1-t)$, $S(t,T_0)=A(T_0-t)$, $\underline{T}(t,T_0)=\underline{B}(T_0-t)$, and $\theta=1$. The corresponding stochastic volatility put price can be approximated by

$$p_t^S[P_S(t,T_1);K;T_0] \cong p_t^G[P_G(t,T_1);K;T_0] + \frac{1}{2}V_1[\underline{X}(t),t], \tag{5.60}$$

where $p_t^G[P_G(t,T_1);K;T_0]$ is obtained from equation (3.26), and $V_1[\underline{X}(t),t]$ is similarly computed but with $\theta=-1$.

Proof. Comparing equations (3.24) and (3.26) with the general Gaussian option price (5.52), proposition 27 follows immediately. ■

Caps, floors, yield options, and swaptions

The result obtained in proposition 27, for European options on default-free pure discount bonds, can be easily generalized for the valuation of caps, floors, collars and European yield options, under the Duffie and Singleton (1997) assumption of symmetric counterparty credit risk. In order to value caps, floors and collars, equations (3.31) and (3.32) can be used: the difference to Chapter three is that the corresponding European call and put options on zero-coupon bonds are now valued under proposition 27 (instead of proposition 3). The valuation of European yield options with settlement in arrears can be made using the stochastic volatility approximate pricing formulae for caplets and floorlets, after adjusting them for the compounding period of the underlying interest rate (as described at the end of subsection 3.4.2).

Unfortunately, although an approximate stochastic volatility solution can also be derived for European options on coupon-bearing bonds (and thus for European swaptions as well), based on the Gaussian *rank 1* (proposition 4) or *lognormal* (proposition 5) approximations, such first order stochastic volatility approximation can not be made explicit (i.e. the integration with respect to the model' state variables can not be eliminated from the final solution). However, because an analytical stochastic volatility first order solution exists for European options on pure discount bonds, it is always possible to price European swaptions using the *stochastic duration* approximation suggested by Wei (1997) and Munk (1998).[26]

In summary, the first order stochastic volatility explicit approximation derived for European options on pure discount bonds can be applied to a wide variety of effectively traded interest rate options.

Examples

Tables 5.5 to 5.7 price a five-year interest rate floor (with quarterly compounding), for different strikes and different invariant transformations, using the three-factor CIR model of Schlögl and Sommer (1998, Figure 5). The floor value is divided into 19 European calls:

$$Floor_0 = (1 + 0.25k) \sum_{i=1}^{19} c_0 \left[P(0, 0.25(i+1)); (1+0.25k)^{-1}; 0.25i \right],$$

where k is the floor rate and $c_0[S; X; T]$ denotes the time-0 price of an European call on the asset S, with a strike X, and with maturity at time T. The exact multifactor CIR call prices are computed using the analytical Fourier transforms' approach of Chen and Scott (1995b). The Duffie et al. (1998) pricing methodology is also implemented by

[26] In essence, an European call on a coupon-bearing bond, with strike X and maturity T, is approximated by ξ times an European call, with strike $\frac{X}{\xi}$ and maturity T, on a pure discount bond with expiry equal to the *stochastic duration* of the coupon bond. The constant ξ is the forward price of the coupon-bearing bond for its *stochastic duration*.

Table 5.5: Pricing of a five-year ATM floor with quarterly compounding using the three-factor CIR model of Schlögl and Sommer (1998, Figure 5), for different affine invariant transformations

Call Expiry (years)	Money-ness[a]	Exact MCIR price(c_0^S)	Duffie *et al.* (1998)	Percentage Pricing Errors			
				$\underline{u} = -a^{-1} \cdot \underline{b}$		$\underline{u} = \underline{X}(t)$	
				$\frac{c_0^G - c_0^S}{c_0^S}$	$\frac{c_0^G + \frac{V_1}{2} - c_0^S}{c_0^S}$	$\frac{c_0^G - c_0^S}{c_0^S}$	$\frac{c_0^G + \frac{V_1}{2} - c_0^S}{c_0^S}$
0.25	-0.0001%	0.046329	-0.0005%	-0.677%	0.0408%	0.107%	0.0443%
0.5	-0.0002%	0.063169	0.0001%	-0.535%	0.0768%	0.213%	0.0839%
0.75	-0.0003%	0.074636	0.0003%	-0.391%	0.1115%	0.322%	0.1228%
1	-0.0003%	0.083181	-0.0021%	-0.247%	0.1454%	0.434%	0.1634%
1.25	-0.0002%	0.089795	-0.0060%	-0.101%	0.1784%	0.549%	0.2044%
1.5	-0.0002%	0.095013	-0.0061%	0.045%	0.2103%	0.661%	0.2405%
				...			
3.5	0.0002%	0.111222	0.0831%	1.205%	0.4277%	1.453%	0.3930%
3.75	0.0003%	0.111500	0.0861%	1.347%	0.4499%	1.554%	0.4142%
4	0.0003%	0.111559	0.0866%	1.486%	0.4709%	1.657%	0.4367%
4.25	0.0003%	0.111429	0.0849%	1.624%	0.4908%	1.760%	0.4599%
4.5	0.0002%	0.111134	0.0815%	1.761%	0.5096%	1.862%	0.4836%
4.75	0.0002%	0.110696	0.0767%	1.896%	0.5273%	1.964%	0.5074%
Floor		1.890898	0.0456%	1.130%	0.3363%	1.186%	0.3742%
Time			43625.8s	71.35s		56.9s	

The floor rate is set equal to the 5-year forward swap rate (with quarterly compounding): $k = 6.0456\%$. Floor prices are for \$100 of Notional Value.

[a]Difference between forward price of underlying PDB and strike $(1 + 0.25k)^{-1}$, over the strike.

Exact MCIR prices, c_0^S, are computed from Chen and Scott (1995b) formulae.

Duffie et al. (1998) approach implemented by evaluating numerically the characteristic function and inverting each Fourier transform through a 10-point Gaussian quadrature.

c_0^G is the exact Gaussian price, computed from proposition 3.

$c_0^G + 0.5V_1$ is the first order approximate stochastic vol. price, given by proposition 27.

$\underline{X}(t)$ is the current state-vector, a and $\underline{b}$ are model' parameters, and $\underline{u}$ defines the affine transformation under use.

computing the characteristic function not analytically but rather numerically,[27] using a 10-point Gaussian quadrature to invert each Fourier transform. Gaussian call prices are obtained from proposition 3, and first order approximate stochastic volatility prices are computed using proposition 27 (where equation (5.56) was implemented using Romberg's integration on an open interval).

Table 5.8 values an at-the-money five-year interest rate floor (with quarterly compounding), using the same $\mathbf{A}_2(3)$ specification as in table 5.2. Because, in this case, no closed-form solution exists for European options on pure discount bonds, the exact price of each call was estimated through standard Monte Carlo simulation, using the usual Euler discretization of equation (2.5) with $1,000$ time steps per year, independent normal variates generated

[27]Because, in general, the analytical form of the relevant characteristic function is unknown, this procedure enables the assessment of the computational time involved in this pricing methodology

Table 5.6: Pricing of a five-year OTM floor with quarterly compounding using the three-factor CIR model of Schlögl and Sommer (1998, Figure 5), for different affine invariant transformations

Call Expiry (years)	Money-ness[a]	Exact MCIR price(c_0^S)	Percentage Pricing Errors: Duffie et al. (1998)	$\underline{u}$: VAR match $\frac{c_0^G - c_0^S}{c_0^S}$	$\underline{u}$: VAR match $\frac{c_0^G + \frac{V_1}{2} - c_0^S}{c_0^S}$	$\underline{u} = \underline{X}(t)$ $\frac{c_0^G - c_0^S}{c_0^S}$	$\underline{u} = \underline{X}(t)$ $\frac{c_0^G + \frac{V_1}{2} - c_0^S}{c_0^S}$
0.25	-0.258%	0.000436	-0.0899%	49.699%	-3.9281%	50.258%	-4.0793%
0.5	-0.258%	0.003148	0.1813%	30.826%	-0.1865%	31.369%	-0.3581%
0.75	-0.258%	0.006929	-0.1872%	24.868%	0.5117%	25.540%	0.3875%
1	-0.258%	0.010796	-0.4358%	22.015%	0.8221%	22.776%	0.7036%
1.25	-0.258%	0.014413	-0.0053%	20.377%	1.0055%	21.223%	0.8916%
1.5	-0.258%	0.017674	-0.0031%	19.350%	1.1382%	20.276%	1.0261%
...							
3.5	-0.257%	0.032466	-0.0028%	17.566%	1.7256%	18.936%	1.6001%
3.75	-0.257%	0.033306	-0.0054%	17.616%	1.7737%	19.020%	1.6454%
4	-0.257%	0.033997	-0.0081%	17.688%	1.8181%	19.121%	1.6869%
4.25	-0.257%	0.034556	-0.0107%	17.779%	1.8590%	19.236%	1.7247%
4.5	-0.257%	0.034998	-0.0132%	17.884%	1.8966%	19.361%	1.7591%
4.75	-0.257%	0.035336	-0.0155%	18.002%	1.9311%	19.493%	1.7902%
Floor		0.450374	-0.0158%	18.282%	1.5733%	19.542%	1.4487%
Time			42486.2s	56.25s		53.77s	

The floor rate is set equal to $k = 5\%$ ($< 6.0456\%$).
Floor prices are for \$100 of Notional Value.
[a]Difference between forward price of underlying PDB and strike $(1 + 0.25k)^{-1}$, over the strike.
Exact MCIR prices, c_0^S, are computed from Chen and Scott (1995b) formulae.
Duffie et al. (1998) approach implemented by evaluating numerically the characteristic function and inverting each Fourier transform through a 10-point Gaussian quadrature.
c_0^G is the exact Gaussian price, computed from proposition 3.
$c_0^G + 0.5V_1$ is the first order approximate stochastic vol. price, given by proposition 27.
$\underline{X}(t)$ is the current state-vector and $\underline{u}$ defines the affine transformation under use.

through the Box-Muller algorithm, 200,000 simulations, and the numerical solution of equations (2.8) and (2.9) in order to compute the option' terminal payoff. Besides the Monte Carlo price estimate, the percentage of its standard error on the mean price is also shown.

In general terms, all the previous examples show that: *i*) the first order stochastic volatility approximation is still accurate and fast to implement for interest rate options; *ii*) the pricing errors increase with the maturity of the contingent claim and are higher for out-of-the-money options; *iii*) the first order approximating term improves significantly the zeroth order approximation; and *iv*) the proposed transformation (5.33) yields, on average, the best results.

Finally, using again the same $\mathbf{A}_2(3)$ specification, table 5.9 prices a 6-month European call on a 5-year coupon-bearing bond (with a 6% annual coupon and a face value of 100), for different strikes, through the approximation of Wei (1997) and Munk (1998). Once more, the proposed first order stochastic volatility explicit approximation is fast and accurate.

Table 5.7: Pricing of a five-year ITM floor with quarterly compounding using the three-factor CIR model of Schlögl and Sommer (1998, Figure 5), for different affine invariant transformations

Call Expiry (years)	Money-ness[a]	Exact MCIR price(c_0^S)	Duffie *et al.* (1998)	Percentage Pricing Errors $\underline{u} = -a^{-1} \cdot \underline{b}$		$\underline{u} = \underline{X}(t)$	
				$\frac{c_0^G - c_0^S}{c_0^S}$	$\frac{c_0^G + \frac{V_1}{2} - c_0^S}{c_0^S}$	$\frac{c_0^G - c_0^S}{c_0^S}$	$\frac{c_0^G + \frac{V_1}{2} - c_0^S}{c_0^S}$
0.25	0.235%	0.228866	-0.0001%	-0.696%	-0.6960%	-0.142%	-0.0053%
0.5	0.235%	0.230697	-0.0005%	-3.096%	-3.0955%	-0.454%	0.0052%
0.75	0.235%	0.233313	0.0027%	-5.776%	-5.7762%	-0.700%	0.0243%
1	0.235%	0.235610	0.0094%	-8.271%	-8.2710%	-0.877%	0.0454%
1.25	0.235%	0.237351	0.0126%	-10.504%	-10.5022%	-1.007%	0.0670%
1.5	0.235%	0.238519	0.0001%	-12.489%	-12.4785%	-1.101%	0.0884%
				...			
3.5	0.235%	0.233367	0.0001%	-22.749%	-21.8815%	-1.280%	0.2410%
3.75	0.235%	0.231575	0.0001%	-23.611%	-22.5237%	-1.268%	0.2571%
4	0.235%	0.229631	-0.0066%	-24.417%	-23.0907%	-1.253%	0.2725%
4.25	0.235%	0.227557	0.0000%	-25.173%	-23.5915%	-1.236%	0.2871%
4.5	0.235%	0.225372	0.0000%	-25.885%	-24.0341%	-1.216%	0.3010%
4.75	0.235%	0.223092	0.0000%	-26.558%	-24.4251%	-1.194%	0.3142%
Floor		4.517682	0.0010%	-16.669%	-16.1096%	-1.080%	0.1610%
Time			41782s	56.63s		56.79s	

The floor rate is set equal to $k = 7\%$ ($> 6.0456\%$).

Floor prices are for \$100 of Notional Value.

[a]Difference between forward price of underlying PDB and strike $(1 + 0.25k)^{-1}$, over the strike.

Exact MCIR prices, c_0^S, are computed from Chen and Scott (1995b) formulae.

Duffie et al. (1998) approach implemented by evaluating numerically the characteristic function and inverting each Fourier transform through a 10-point Gaussian quadrature.

c_0^G is the exact Gaussian price, computed from proposition 3.

$c_0^G + 0.5V_1$ is the first order approximate stochastic vol. price, given by proposition 27.

$\underline{X}(t)$ is the current state-vector, a and $\underline{b}$ are model' parameters, and $\underline{u}$ defines the affine transformation under use.

5.5.3 Futures options on pure discount bonds

A first order explicit approximation

Next proposition applies Corollary 4 to the valuation of European options on pure discount bond futures with stock-style margining.

Proposition 28 *Under the stochastic volatility specification of the Duffie and Kan (1996) model, the time-t premium of an European conventional call on the asset $FP_S(t, T_f, T_1)$, with a strike price of K_f, and expiry date at time T_0 (such that $t \leq T_0 \leq T_f \leq T_1$), can be approximated by*

$$c_t^S [FP_S(t, T_f, T_1); K_f; T_0] \cong c_t^G [FP_G(t, T_f, T_1); K_f; T_0] + \frac{1}{2} V_1 [\underline{X}(t), t], \tag{5.61}$$

Table 5.8: Pricing of a five-year ATM floor with quarterly compounding using an $\mathbb{A}_2(3)$ model

Call Expiry (years)	Money-ness[a]	Standard Monte Carlo price (c_0^S)	Standard Monte Carlo % std. error	Percentage Pricing Errors: Duffie *et al.* (1998)	$u = \underline{X}(t)$: $\frac{c_0^G - c_0^S}{c_0^S}$	$u = \underline{X}(t)$: $\frac{c_0^G + \frac{V_1}{2} - c_0^S}{c_0^S}$
0.25	0.2016%	0.197400	0.1104%	0.2258%	0.0246%	0.1986%
0.5	0.1363%	0.157100	0.1925%	-0.1155%	0.4464%	0.0483%
0.75	0.1286%	0.166500	0.2131%	-0.2253%	2.2049%	0.1358%
1	0.1317%	0.177400	0.2201%	-0.2563%	3.5054%	0.2368%
1.25	0.1299%	0.182500	0.2271%	-0.5878%	4.0310%	-0.0240%
1.5	0.1196%	0.180800	0.2363%	-0.5354%	4.5817%	0.0256%
			...			
3.5	-0.1128%	0.112500	0.3523%	-0.0485%	1.8958%	0.4934%
3.75	-0.1473%	0.104700	0.3690%	0.4500%	1.6021%	1.1187%
4	-0.1821%	0.097700	0.3850%	0.7523%	1.0691%	1.5610%
4.25	-0.2169%	0.091600	0.3998%	0.6557%	0.1064%	1.6086%
4.5	-0.2519%	0.086000	0.4150%	0.4969%	-0.9362%	1.5973%
4.75	-0.2869%	0.080700	0.4306%	0.4609%	-1.8656%	1.7165%
Floor		2.710338		-0.1013%	2.7013%	0.3778%
Time		60 hours		45543.21s	67.29s	

The floor rate is set equal to the 5-year forward swap rate (with quarterly compounding): $k = 6.3933\%$. Floor prices are for \$100 of Notional Value.
[a]Difference between forward price of underlying PDB and strike, divided by strike price.
Monte Carlo: 200,000 simulations with 1,000 time steps per year; % std. error is standard error divided by option price estimate.
Duffie et al. (1998) approach implemented by evaluating numerically the characteristic function and inverting each Fourier transform through a 10-point Gaussian quadrature.
c_0^G is the exact Gaussian price, computed from proposition 3.
$c_0^G + 0.5V_1$ is the first order approximate stochastic vol. price, given by proposition 27.
$\underline{X}(t)$ is the current state-vector and $\underline{u}$ defines the affine transformation under use.

where $c_t^G[FP_G(t,T_f,T_1);K_f;T_0]$ is computed from equation (3.77), and $V_1[\underline{X}(t),t]$ has the "explicit" solution given by equation (5.56) but with $q=1$, $U(t,\cdot) = A(T_0-t)+A(T_1-t)-A(T_f-t)-J(t)+I(t)$, $\underline{Q}(t,\cdot) = \underline{B}(T_0-t)+\underline{B}(T_1-t)-\underline{B}(T_f-t)$, $S(t,T_0) = A(T_0-t)$, $\underline{T}(t,T_0) = \underline{B}(T_0-t)$, $K = K_f$, and $\theta = 1$. The corresponding stochastic volatility put price can be approximated by

$$p_t^S[FP_S(t,T_f,T_1);K_f;T_0] \cong p_t^G[FP_G(t,T_f,T_1);K_f;T_0] + \frac{1}{2}V_1[\underline{X}(t),t], \qquad (5.62)$$

where $p_t^G[FP_G(t,T_f,T_1);K_f;T_0]$ is obtained from equation (3.80), and $V_1[\underline{X}(t),t]$ is similarly computed but with $\theta=-1$.

Proof. Comparing equations (3.77) and (3.80) with the general Gaussian option price (5.52), proposition 28 is obtained. ■

Table 5.9: Pricing of a 6-month European call on a 5-year coupon-bearing bond (CBB), using an $\mathbb{A}_2(3)$ model

call on CBB (6% annual coupon)				1st-order approximation (with $\underline{u} = \underline{X}(t)$)			
Strike	Money-	Standard	Monte Carlo	call on PDB		call on CBB	
(X)	ness[a]	price	%std.error	$\frac{X}{\zeta}$	$c_0^G + \frac{V_1}{2}$	$\zeta\left(c_0^G + \frac{V_1}{2}\right)$	% error
99	1.668%	2.064355	0.2162%	0.7662	0.016007	2.068404	0.1962%
99.5	1.157%	1.720938	0.2410%	0.7700	0.013353	1.725422	0.2605%
100	0.652%	1.409296	0.2698%	0.7739	0.010941	1.413725	0.3143%
100.6515	0.000%	1.054243	0.3151%	0.7790	0.008192	1.058508	0.4045%
101	-0.345%	0.888957	0.3439%	0.7816	0.006912	0.893177	0.4747%
101.5	-0.836%	0.682297	0.3920%	0.7855	0.005311	0.686210	0.5734%
102	-1.322%	0.510303	0.4505%	0.7894	0.003978	0.514039	0.7320%
102.5	-1.803%	0.371407	0.5220%	0.7933	0.002899	0.374629	0.8676%
Time		63148s			14.83s		

[a]Difference between 6-month forward price of CBB (100.6515) and strike, divided by the strike.
Monte Carlo: 200,000 simulations with 1,000 time steps per year; % std. error is standard error divided by option price estimate.
ξ is the forward price of the CBB for its stochastic duration: 129.2149.
The stochastic duration of the CBB is the maturity of a PDB with the same instantaneous variance of relative price changes: 4.460377 years.
c_0^G and V_1 are computed from propositions 3 and 27, respectively.
The European call on the CBB, with strike X, is approximated by ξ times an European call, with strike $\frac{X}{\xi}$, on a PDB with maturity equal to the stochastic duration of the underlying CBB.

Example

Table 5.10 prices European futures calls on pure discount bonds, for different strikes, with $(T_0 - t, T_f - t, T_1 - t) = (0.25, 0.5, 2.5)$, and using the three-factor CIR model of Schlögl and Sommer (1998, Figure 5). Exact stochastic volatility prices are obtained through standard Monte Carlo simulation, as described in table 5.8 (although now the terminal option payoff is computed from proposition 23). The Gaussian or zeroth order price is given by equation (3.77), and the first order approximation is obtained from proposition 28.

Again, the accuracy of the first order explicit stochastic volatility solution is acceptable (pricing errors of about one standard error of the Monte Carlo estimate), while its computational time is significantly lower than the one taken by the exact numerical result.

5.5.4 Options on short-term interest rate futures

A first order explicit approximation

Next result generalizes propositions 11 and 12 to the stochastic volatility specification of the Duffie and Kan (1996) model.

Table 5.10: Pricing of 3-month European calls on 6-month pure discount bond futures with a maturity of 2.5 years for the underlying bond, using the three-factor CIR model of Schlögl and Sommer (1998, Figure 5)

				$\underline{u} = \underline{X}(t)$			
Strike	Money-	Standard Monte Carlo		Gaussian model		SV model	
	ness[a]	price (c_0^S)	%std.error	c_0^G	$\frac{c_0^G - c_0^S}{c_0^S}$	$c_0^G + \frac{V_1}{2}$	$\frac{c_0^G + \frac{V_1}{2} - c_0^S}{c_0^S}$
0.88	-0.776%	0.007449	0.1788%	0.007398	-0.6789%	0.007436	-0.1798%
0.8825	-0.494%	0.005510	0.2177%	0.005467	-0.7860%	0.005497	-0.2270%
0.885	-0.212%	0.003846	0.2691%	0.003820	-0.6663%	0.003835	-0.2916%
0.88688	0.000%	0.002806	0.3194%	0.002798	-0.2828%	0.002796	-0.3511%
0.8875	0.070%	0.002507	0.3387%	0.002505	-0.0742%	0.002498	-0.3696%
0.89	0.352%	0.001510	0.4358%	0.001530	1.3483%	0.001503	-0.4456%
0.8925	0.633%	0.000830	0.5778%	0.000865	4.2088%	0.000827	-0.4024%
0.895	0.915%	0.000412	0.7962%	0.000450	9.2744%	0.000411	-0.2209%
Time		64236s		17.63s			

[a]Difference between underlying futures price and strike price, divided by strike price.
Monte Carlo: 200,000 simulations with 1,000 time steps per year; % std. error is standard error divided by option price estimate.
c_0^G is the exact Gaussian price, computed from proposition 9.
$c_0^G + 0.5V_1$ is the first order approximate stochastic vol. price, given by proposition 28.
$\underline{X}(t)$ is the current state-vector and $\underline{u}$ defines the affine transformation under use.

Proposition 29 *Under the stochastic volatility specification of the Duffie and Kan (1996) model, the time-t premium of an European option on the futures contract $FR_S(t, T_f, T_1)$, with a strike price equal to K_R, and expiring at time T_0 (such that $t \leq T_0 \leq T_f \leq T_1$), can be approximated by a first order solution where $V_G[\underline{X}(t), t]$ is computed under propositions 11 or 12, and $V_1[\underline{X}(t), t]$ has the "explicit" solution given by equation (5.56) but with $q = \frac{100}{T_1 - T_f}$, $U(t, \cdot) = A(T_f - t) - A(T_1 - t) + \phi[A(T_0 - t) + L(T_0) + \rho(t)] + (1 - \phi) L(t)$, $\underline{Q}(t, \cdot) = \underline{B}(T_f - t) - \underline{B}(T_1 - t) + \phi \underline{B}(T_0 - t)$, $S(t, T_0) = \phi A(T_0 - t)$, $\underline{T}(t, T_0) = \phi \underline{B}(T_0 - t)$, and $K = 1 + (T_1 - T_f)\frac{100 - K_R}{100}$. For conventional futures options $\phi = 1$, while for pure futures options $\phi = 0$. For puts $\theta = 1$, and for calls $\theta = -1$.*

Proof. Comparing equations (3.96), (3.97), (3.101) and (3.102) with the general Gaussian option price (5.52), proposition 29 follows. ■

Remark 28 *Following remarks 9 and 10, proposition 29 can also be used to value European yield options with settlement at the option's expiry date or pure American short-term interest rate futures options, under the stochastic volatility specification of the Duffie and Kan (1996) model.*

Example

Table 5.11 prices 6-month pure puts on 3-month Eurodollar futures (with $T_f = T_0$), for different strikes, and using the $\mathbf{A}_1(3)_{DS}$ model of Dai and Singleton (1998, Table II). Exact

Table 5.11: Pricing of 6-month pure put options on 3-month Eurodollar futures (also with a maturity of 6 months) using the $\mathbf{A}_1(3)_{DS}$ Dai and Singleton (1998, Table II) model

Strike	Money-ness[a]	Standard Monte Carlo price (Fp_0^S)	% std. error	Percentage Pricing Errors (affine transformation: $\underline{u} = \underline{X}(t)$) $\frac{Fp_0^G - Fp_0^S}{Fp_0^S}$	$\frac{Fp_0^G + \frac{V_1}{2} - Fp_0^S}{Fp_0^S}$
88.00	-0.668%	0.186360	0.4908%	-1.3784%	0.2939%
88.25	-0.386%	0.268631	0.4101%	-1.2087%	-0.1089%
88.50	-0.104%	0.373441	0.3465%	-1.0245%	-0.3221%
88.592	0.000%	0.417997	0.3265%	-0.9686%	-0.3780%
88.75	0.178%	0.502075	0.2955%	-0.8788%	-0.4452%
89.00	0.461%	0.654135	0.2541%	-0.7286%	-0.4710%
89.25	0.743%	0.827910	0.2202%	-0.5620%	-0.4158%
89.50	1.025%	1.021050	0.1920%	-0.4181%	-0.3395%
Time		34306s		23.34s	

[a]Difference between underlying futures price and strike price, divided by strike price.
Monte Carlo: 200,000 simulations with 1,000 time steps per year; % std. error is standard error divided by option price estimate.
Fp_0^G is the exact Gaussian price, computed from proposition 12
$Fp_0^G + 0.5V_1$ is the first order approximate stochastic volatility price, given by proposition 29.
$\underline{X}(t)$ is the current state-vector and $\underline{u}$ defines the affine transformation under use.

stochastic volatility option prices are obtained through standard Monte Carlo simulation (with terminal option payoff computed from proposition 25), Gaussian prices are given by equation (3.102), and the first order approximation follows from proposition 29. As before, the proposed approximation is fast and accurate.

5.6 Conclusions

The main purpose and contribution of this Chapter consisted in providing (approximate) pricing formulae, under the most general multifactor, mean-reverting, time-homogeneous, and affine term structure model, that only involve one integral with respect to the maturity of the contingent claim under valuation, and are therefore extremely easy to implement in practice.

Firstly, the functional form for Arrow-Debreu prices under the Gaussian nested version of the Duffie and Kan (1996) model was derived. Then, the exact Gaussian valuation formulae derived in Chapter three were converted into approximate stochastic volatility ones that involved integrals with respect not only to the maturity of the contingent claim under valuation but also to each one of the model' factors. Finally, and taking advantage of the analytical tractability provided by the "special" model specification adopted, all stochastic volatility pricing formulae were simplified into first order approximate ones that do not involve any integration with respect to the model' state variables.

Such factor-integral independent stochastic volatility valuation formulae were derived for a wide range of interest rate contingent claims: bonds, FRAs, IRSs, interest rate futures, European options on pure discount bonds, caps and floors, yield options, and European futures options on zero-coupon bonds and on short-term interest rates. The empirical results presented in this Chapter, for different parameter' configurations, have shown that the proposed approximations are extremely fast to implement as well as accurate. In fact, because there is no need to integrate numerically with respect to each state variable, the numerical efficiency of these pricing formulae is still good for high dimensional model specifications. An additional advantage of the first order explicit approximate stochastic volatility pricing formulae proposed in this dissertation is that they can be easily differentiated with respect to each state variable, and thus enable the implementation of dynamic hedging strategies. As an accessory result, exact pricing solutions were obtained for long-term and short-term interest rate futures, under the "general" specification of the Duffie and Kan (1996) model.

In terms of practical applicability, the proposed explicit approximate stochastic volatility pricing formulae constitute efficient tools to estimate (using, for instance, the non-linear Kalman filter approach described in Chapter four) exponential-affine term structure models, based on market information about LIBOR rates, FRAs, short-term interest rate futures and futures options, swaps, caps, floors, and even European swaptions.

5.7 Appendices

5.7.1 Proof of Proposition 21

To prove proposition 21 is equivalent to verify that

$$V_S\left[\underline{X}(t),t\right]-V_G\left[\underline{X}(t),t\right]-\frac{1}{2}\tilde{V}_1\left[\underline{X}(t),t\right]=o(\lambda).$$

In order to highlight the dependencies on the perturbation parameter λ, the previous error term can be further rewritten as:

$$\begin{aligned}
&V_S\left[\underline{X}(t),t\right]-V_G\left[\underline{X}(t),t\right]-\frac{1}{2}\tilde{V}_1\left[\underline{X}(t),t\right]\\
=\;&\frac{1}{2}\left\{V_1\left[\underline{X}(t),t\right]-\tilde{V}_1\left[\underline{X}(t),t\right]\right\}+\sum_{p\geq 2}\frac{1}{2^p}V_p\left[\underline{X}(t),t\right]\\
=\;&\frac{\lambda}{2}\left\{U_1\left[\underline{X}(t),t\right]-\tilde{U}_1\left[\underline{X}(t),t\right]\right\}+\sum_{p\geq 2}\left(\frac{\lambda}{2}\right)^p U_p\left[\underline{X}(t),t\right],
\end{aligned}$$

where $U_0[\underline{X}(t),t] = V_G[\underline{X}(t),t]$ and, for $p \geq 1$,

$$
\begin{aligned}
U_p[\underline{X}(t),t] &= \int_t^T dl \int_{\underline{X}(l)\in\mathbf{D}} d\underline{X}(l)\, G[\underline{X}(l),l;\underline{X}(t),t] \\
&\quad tr\left\{\frac{\partial^2 U_{p-1}[\underline{X}(l),l]}{\partial\underline{X}(l)\,\partial\underline{X}'(l)}\cdot\Sigma\cdot W^D(l)\cdot\Sigma'\right\}.
\end{aligned}
$$

Because the $\mathbb{A}_m(n)$ canonical specification allows definition (2.6) to be restated as

$$
\mathbf{D} = \{\underline{X}\in\Re^n : \alpha_i + \lambda X_i > 0, i = 1,\ldots,m\},
$$

it follows that:

$$
\begin{aligned}
&\lim_{\lambda\to 0}\left|\frac{V_S[\underline{X}(t),t] - V_G[\underline{X}(t),t] - \frac{1}{2}\tilde{V}_1[\underline{X}(t),t]}{\lambda}\right| \qquad (5.63)\\
&= \lim_{\lambda\to 0}\left|\int_t^T dl \int_{\underline{X}(l)\in\Re^n} d\underline{X}(l)\left(\prod_{j=1}^m 1_{\{\lambda X_j(l)>-\alpha_j\}} - 1\right) G[\underline{X}(l),l;\underline{X}(t),t]\right.\\
&\quad \left.\frac{1}{2}tr\left\{\frac{\partial^2 V_G[\underline{X}(l),l]}{\partial\underline{X}(l)\,\partial\underline{X}'(l)}\cdot\Sigma\cdot W^D(l)\cdot\Sigma'\right\} + \sum_{p\geq 2}\frac{1}{2^p}\lambda^{p-1}U_p[\underline{X}(t),t]\right|.
\end{aligned}
$$

Since $\lim_{\lambda\to 0}\prod_{j=1}^m 1_{\{\lambda X_j(l)>-\alpha_j\}} = 1$ and $\lim_{\lambda\to 0} U_p[\underline{X}(t),t] = \tilde{U}_p[\underline{X}(t),t]$ as long as $\alpha_j > 0$, for $j = 1,\ldots,m$,[28] then the limit (5.63) is zero if $\left|\tilde{U}_p[\underline{X}(t),t]\right| < \infty$ for $p \geq 2$. ■

5.7.2 Conditional Mean and Covariance of $\tilde{\underline{X}}(T)$

For $T \geq t$, and assuming that a^{-1} exists, equation (5.28) can be rewritten under the following integral form:

$$
\begin{aligned}
\tilde{\underline{X}}(T) &= e^{a(T-t)}\cdot\tilde{\underline{X}}(t) + \left[e^{a(T-t)} - I_n\right]\cdot a^{-1}\cdot\tilde{b} \qquad (5.64)\\
&\quad + \int_t^T e^{a(T-v)}\cdot\Sigma\cdot\sqrt{\tilde{V}^D(v)}\cdot d\underline{W}^Q(v).
\end{aligned}
$$

Clearly, under the boundedness conditions

$$
\int_t^T e^{a(T-v)}\cdot\Sigma\cdot\tilde{U}^D\cdot\Sigma'\cdot e^{a'(T-v)}dv < \infty,
$$

[28] Although Dai and Singleton (1998, definition III.1) normalize α_j to zero for the first m factors, an invariant transformation, along the lines of Dai and Singleton (1998, definition A.1), can always yield the desired condition.

where $\tilde{U}^D = diag\{\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n\}$, and

$$E_Q^S\left[\left(\int_t^T e^{a(T-v)} \cdot \Sigma \cdot \tilde{V}^D(v) \cdot \Sigma' \cdot e^{a'(T-v)} dv\right) \middle| \underline{\tilde{X}}(t)\right] < \infty,$$

the Itô's integral contained in the right-hand side of equation (5.64) is a martingale, and therefore the conditional mean of the state vector is the same for the Gaussian and for the stochastic volatility specifications of the Duffie and Kan (1996) model:

$$\begin{aligned} E_Q^G\left[\underline{\tilde{X}}(T)\middle|\underline{\tilde{X}}(t)\right] &= E_Q^S\left[\underline{\tilde{X}}(T)\middle|\underline{\tilde{X}}(t)\right] \\ &= e^{a(T-t)} \cdot \underline{\tilde{X}}(t) + \left[e^{a(T-t)} - I_n\right] \cdot a^{-1} \cdot \underline{\tilde{b}}, \end{aligned} \tag{5.65}$$

where $E_Q^G\left[\cdot\middle|\underline{\tilde{X}}(t)\right]$ and $E_Q^S\left[\cdot\middle|\underline{\tilde{X}}(t)\right]$ denote a conditional expectation (on $\underline{\tilde{X}}(t)$), computed under the martingale measure Q, for the Gaussian or stochastic volatility versions of the Duffie and Kan (1996) model, respectively.

Using again equation (5.64), the second conditional moment of the state vector under the nested deterministic volatility specification is

$$COV^G\left[\underline{\tilde{X}}(T)\middle|\underline{\tilde{X}}(t)\right] = \int_t^T e^{a(T-v)} \cdot \Sigma \cdot \tilde{U}^D \cdot \Sigma' \cdot e^{a'(T-v)} dv. \tag{5.66}$$

For the general stochastic volatility formulation, the conditional covariance matrix corresponds to:

$$COV^S\left[\underline{\tilde{X}}(T)\middle|\underline{\tilde{X}}(t)\right] = \int_t^T e^{a(T-v)} \cdot \Sigma \cdot E_Q^S\left[\tilde{V}^D(v)\middle|\underline{\tilde{X}}(t)\right] \cdot \Sigma' \cdot e^{a'(T-v)} dv, \tag{5.67}$$

with

$$\begin{aligned} & E_Q^S\left[\tilde{V}^D(v)\middle|\underline{\tilde{X}}(t)\right] \\ = \; & diag\left\{\tilde{\alpha}_1 + \underline{\beta_1}' \cdot E_Q^S\left[\underline{\tilde{X}}(v)\middle|\underline{\tilde{X}}(t)\right], \ldots, \tilde{\alpha}_n + \underline{\beta_n}' \cdot E_Q^S\left[\underline{\tilde{X}}(v)\middle|\underline{\tilde{X}}(t)\right]\right\}. \end{aligned}$$

If the following crude approximation is made,

$$\tilde{V}^D(v) \cong \tilde{V}^D(t), \forall v \in [t, T], \tag{5.68}$$

then

$$\underline{u} = \underline{X}(t) \Rightarrow COV^G\left[\underline{\tilde{X}}(T)\middle|\underline{\tilde{X}}(t)\right] = COV^S\left[\underline{\tilde{X}}(T)\middle|\underline{\tilde{X}}(t)\right].$$

5.7.3 Proof of Corollary 3

Combining equation (5.42) with the Gaussian *Arrow-Debreu state price* (5.1), and with the definition (2.6) of the state variables' domain under the general Duffie and Kan (1996) model, yields:

$$
\begin{aligned}
V_1\left[\underline{X}(t), t\right] &= \int_t^T dlP_G(t,l)\exp\left[\varphi(l,T) + \frac{1}{2}\underline{\psi}'(l,T)\cdot\Delta(l-t)\right. \\
&\quad \left.\cdot\underline{\psi}(l,T) + \underline{\psi}'(l,T)\cdot\underline{M}(l-t)\right]\left[\sum_{k=1}^{n}\left(\underline{\psi}'(l,T)\cdot\underline{\varepsilon_k}\right)^2\right]U,
\end{aligned}
\tag{5.69}
$$

with

$$
\begin{aligned}
U &= \int_{\underline{X}(l)\in\Re^n} d\underline{X}(l)\,\frac{\left[\underline{\beta_k}'\cdot\underline{X}(l)\right]\prod_{j=1}^{n}1_{\left\{\underline{\beta_j}'\cdot\underline{X}(l)\geq-\alpha_j\right\}}}{\sqrt{(2\pi)^n\left|\Delta(l-t)\right|}} \\
&\quad \exp\left\{-\frac{1}{2}\left[\underline{X}(l)-\Delta(l-t)\cdot\underline{\mu}(l)\right]'\cdot\Delta^{-1}(l-t)\cdot\left[\underline{X}(l)-\Delta(l-t)\cdot\underline{\mu}(l)\right]\right\}.
\end{aligned}
$$

If $n=1$, then

$$
U = \int_{-\alpha_k}^{\infty} dy\,\frac{y}{\sqrt{2\pi\underline{\beta_k}'\cdot\Delta(l-t)\cdot\underline{\beta_k}}}\exp\left[-\frac{1}{2}\frac{\left(y-\underline{\beta_k}'\cdot\Delta(l-t)\cdot\underline{\mu}(l)\right)^2}{\underline{\beta_k}'\cdot\Delta(l-t)\cdot\underline{\beta_k}}\right],
$$

and the exact solution (5.38) follows.

In order to set an upper bound for $V_1\left[\underline{X}(t), t\right]$, Schwarz inequality can be applied:

$$
\begin{aligned}
U^2 &\leq \underline{\beta_k}'\cdot E\left[\underline{X}(l)\cdot\underline{X}'(l)\middle|\mathcal{F}_t\right]\cdot\underline{\beta_k} \\
&\quad \int_{\underline{X}(l)\in\Re^n} d\underline{X}(l)\,\frac{\prod_{j=1}^{n}1^2_{\left\{\underline{\beta_j}'\cdot\underline{X}(l)\geq-\alpha_j\right\}}}{\sqrt{(2\pi)^n\left|\Delta(l-t)\right|}} \\
&\quad \exp\left\{-\frac{1}{2}\left[\underline{X}(l)-\Delta(l-t)\cdot\underline{\mu}(l)\right]'\cdot\Delta^{-1}(l-t)\cdot\left[\underline{X}(l)-\Delta(l-t)\cdot\underline{\mu}(l)\right]\right\}.
\end{aligned}
\tag{5.70}
$$

Because the factor-integral on the right-hand side of (5.70) is surely positive, another application of Schwarz inequality can be made and square roots can be taken:

$$
\begin{aligned}
U^2 &\leq \underline{\beta_k}'\cdot E\left[\underline{X}(l)\cdot\underline{X}'(l)\middle|\mathcal{F}_t\right]\cdot\underline{\beta_k}\sqrt{\Pr\left[\underline{\beta_1}'\cdot\underline{X}(l)\geq-\alpha_1\right]} \\
&\quad \left(\int_{\underline{X}(l)\in\Re^n} d\underline{X}(l)\,\frac{\prod_{j=2}^{n}1^4_{\left\{\underline{\beta_j}'\cdot\underline{X}(l)\geq-\alpha_j\right\}}}{\sqrt{(2\pi)^n\left|\Delta(l-t)\right|}}\right. \\
&\quad \left.\exp\left\{-\frac{1}{2}\left[\underline{X}(l)-\Delta(l-t)\cdot\underline{\mu}(l)\right]'\cdot\Delta^{-1}(l-t)\cdot\left[\underline{X}(l)-\Delta(l-t)\cdot\underline{\mu}(l)\right]\right\}\right)^{\frac{1}{2}}.
\end{aligned}
$$

Repeating successively the same reasoning, it can be shown that

$$U^2 \leq \underline{\beta_k}' \cdot E\left[\underline{X}(l) \cdot \underline{X}'(l) \middle| \mathcal{F}_t\right] \cdot \underline{\beta_k} \prod_{j=1}^{n} \left\{\Pr\left[\underline{\beta_j}' \cdot \underline{X}(l) \geq -\alpha_j\right]\right\}^{\frac{1}{2^j}}. \tag{5.71}$$

Finally, imposing a zero lower bound to U, computing explicitly the previous expectation and all the probabilities, and taking square roots from both sides of inequality (5.71), inequality (5.40) arises.

The lower bound (5.41) follows from (5.69), imposing

$$\left[\underline{\beta_k}' \cdot \underline{X}(l)\right] \prod_{j=1}^{n} 1_{\left\{\underline{\beta_j}' \cdot \underline{X}(l) \geq -\alpha_j\right\}} = -\alpha_k \prod_{j=1}^{n} 1_{\left\{\underline{\beta_j}' \cdot \underline{X}(l) \geq -\alpha_j\right\}}, \forall \underline{X}(l),$$

and assuming the independence amongst the events $\left\{\underline{\beta_j}' \cdot \underline{X}(l) \geq -\alpha_j\right\}$, for all j. ■

5.7.4 Proof of Proposition 23

Using equations (5.13) and (5.14), considering the zero-endowment nature of futures contracts, and the well known convergence of the terminal futures price to its underlying spot price, it follows that the futures price $FP_S(t, T_f, T_1)$ is the solution of the following initial value problem:

$$0 = \mathcal{D}_S FP_S(t, T_f, T_1) + \frac{\partial FP_S(t, T_f, T_1)}{\partial t}, \tag{5.72}$$

subject to

$$FP_S(T_f, T_f, T_1) = P_S(T_f, T_1). \tag{5.73}$$

Clearly, solution (5.44) satisfies the boundary condition (5.73). Moreover, substituting (5.44) into the PDE (5.72), rearranging terms as well as adding and subtracting the time-t instantaneous interest rate, $r(t)$,

$$\begin{aligned}
0 \;=\; & \left\{\underline{B}'(T_1 - t) \cdot [a \cdot \underline{X}(t) + \underline{b}] + \left[\frac{\partial A(T_1 - t)}{\partial t} + \frac{\partial \underline{B}'(T_1 - t)}{\partial t} \cdot \underline{X}(t)\right]\right. \\
& \left. + \frac{1}{2} tr\left[\underline{B}(T_1 - t) \cdot \underline{B}'(T_1 - t) \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right] - r(t)\right\} \\
& - \left\{\underline{B}'(T_f - t) \cdot [a \cdot \underline{X}(t) + \underline{b}] + \left[\frac{\partial A(T_f - t)}{\partial t} + \frac{\partial \underline{B}'(T_f - t)}{\partial t} \cdot \underline{X}(t)\right]\right. \\
& \left. + \frac{1}{2} tr\left[\underline{B}(T_f - t) \cdot \underline{B}'(T_f - t) \cdot \Sigma \cdot V^D(t) \cdot \Sigma'\right] - r(t)\right\} \\
& + \left\{\underline{D}'(t, T_f, T_1) \cdot [a \cdot \underline{X}(t) + \underline{b}] + \left[\frac{\partial C(t, T_f, T_1)}{\partial t} + \frac{\partial \underline{D}'(t, T_f, T_1)}{\partial t}\right.\right. \\
& \left. \cdot \underline{X}(t)\right] - \frac{1}{2} \underline{B}'(T_1 - t) \cdot \Sigma \cdot V^D(t) \cdot \Sigma' \cdot \underline{B}(T_1 - t) + \frac{1}{2} \underline{B}'(T_f - t) \\
& \cdot \Sigma \cdot V^D(t) \cdot \Sigma' \cdot \underline{B}(T_f - t) + \frac{1}{2}\left[\underline{B}(T_1 - t) - \underline{B}(T_f - t) + \underline{D}(t, T_f, T_1)\right]'
\end{aligned}$$

$$\Sigma \cdot V^D(t) \cdot \Sigma' \cdot \left[\underline{B}(T_1 - t) - \underline{B}(T_f - t) + \underline{D}(t, T_f, T_1)\right]\} .$$

The first two terms on the right-hand-side of the previous equation are equal to zero, since they are just the PDEs satisfied by the pure discount bond prices $P_S(t, T_1)$ and $P_S(t, T_f)$, respectively. Therefore, simplifying some terms, and since $\Sigma \cdot V^D(t) \cdot \Sigma' = \sum_{k=1}^{n} \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \left[\alpha_k + \underline{\beta_k}' \cdot \underline{X}(t)\right]$, then the right-hand-side of the last equation can be rewritten as an affine function of $\underline{X}(t)$:

$$\begin{aligned}
0 \;=\; & \left\{ \underline{D}'(t, T_f, T_1) \cdot a + \frac{\partial \underline{D}'(t, T_f, T_1)}{\partial t} + \sum_{k=1}^{n} \underline{B}'(T_f - t) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \right. \\
& \cdot \left[\underline{B}(T_f - t) - \underline{B}(T_1 - t)\right] \underline{\beta_k}' + \frac{1}{2} \sum_{k=1}^{n} \underline{D}'(t, T_f, T_1) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \\
& \cdot \left[2\underline{B}(T_1 - t) - 2\underline{B}(T_f - t) + \underline{D}(t, T_f, T_1)\right] \underline{\beta_k}' \Big\} \cdot \underline{X}(t) \\
& + \left\{ \underline{D}'(t, T_f, T_1) \cdot \underline{b} + \frac{\partial C(t, T_f, T_1)}{\partial t} + \sum_{k=1}^{n} \underline{B}'(T_f - t) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \right. \\
& \cdot \left[\underline{B}(T_f - t) - \underline{B}(T_1 - t)\right] \alpha_k + \frac{1}{2} \sum_{k=1}^{n} \underline{D}'(t, T_f, T_1) \cdot \underline{\varepsilon_k} \cdot \underline{\varepsilon_k}' \\
& \cdot \left[2\underline{B}(T_1 - t) - 2\underline{B}(T_f - t) + \underline{D}(t, T_f, T_1)\right] \alpha_k \Big\} .
\end{aligned}$$

The previous PDE can now be split into the n-dimensional Riccati differential equation (5.45) and into the first order ODE (5.46). ■

5.7.5 Proof of Corollary 4

Using equations (5.52) to (5.55), the functional form of the "gamma matrix" $\frac{\partial^2 V_p[\underline{X}(t), t]}{\partial \underline{X}(t) \partial \underline{X}'(t)}$ can be computed, and it can be shown that equations (5.11) and (5.12) yield the following first order approximation:

$$V_S[\underline{X}(t), t] \cong V_G[\underline{X}(t), t] + \frac{q}{2} \int_t^{T_0} dl \left[V_{11}(l) + V_{12}(l) + V_{13}(l)\right], \tag{5.74}$$

where[29]

$$\begin{aligned}
V_{11}(l) \;=\; & \theta \exp\left[U(l, \cdot)\right] \int_{\underline{X}(l)} d\underline{X}(l)\, G\left[\underline{X}(l), l; \underline{X}(t), t\right] \\
& \exp\left[\underline{Q}'(l, \cdot) \cdot \underline{X}(l)\right] \Phi\left[\theta d_1(l)\right] \left[\sum_{k=1}^{n} \left(\underline{Q}'(l, \cdot) \cdot \underline{\varepsilon_k}\right)^2 \underline{\beta_k}'\right] \cdot \underline{X}(l),
\end{aligned} \tag{5.75}$$

[29] In this appendix, all factor-integrals refer to integration over $\Re^n$.

$$V_{12}(l) = -\theta K \exp\left[S(l,T_0)\right] \int_{\underline{X}(l)} d\underline{X}(l)\, G\left[\underline{X}(l), l; \underline{X}(t), t\right] \tag{5.76}$$

$$\exp\left[\underline{T}'(l,T_0)\cdot \underline{X}(l)\right] \Phi\left[\theta d_0(l)\right] \left[\sum_{k=1}^{n} \left(\underline{T}'(l,T_0)\cdot \underline{\varepsilon_k}\right)^2 \underline{\beta_k}'\right] \cdot \underline{X}(l),$$

and

$$V_{13}(l) = \frac{K e^{S(l,T_0)}}{\sigma(l)\sqrt{2\pi}} \int_{\underline{X}(l)} d\underline{X}(l)\, G\left[\underline{X}(l), l; \underline{X}(t), t\right] e^{\underline{T}'(l,T_0)\cdot \underline{X}(l)} \tag{5.77}$$

$$\exp\left\{-\frac{1}{2}\left[d_0(l)\right]^2\right\} \left\{\sum_{k=1}^{n} \left[\left(\underline{Q}'(l,\cdot) - \underline{T}'(l,T_0)\right)\cdot \underline{\varepsilon_k}\right]^2 \underline{\beta_k}'\right\} \cdot \underline{X}(l).$$

The next step consists in eliminating all factor-integrals from the above equations. Beginning with $V_{13}(l)$, using the definition (5.1) of Gaussian Arrow-Debreu prices, and because $d_0(l)$ can be written as an explicit function of $\underline{X}(l)$,

$$d_0(l) = \frac{1}{\sigma(l)} \left\{\left[\underline{Q}'(l,\cdot) - \underline{T}'(l,T_0)\right] \cdot \underline{X}(l) + d_0^*(l)\right\}, \tag{5.78}$$

then

$$V_{13}(l) = \exp\left\{S(l,T_0) - \frac{\left[d_0^*(l)\right]^2}{2\sigma^2(l)} - \frac{1}{2}\underline{M}'(l-t)\cdot \Delta^{-1}(l-t)\cdot \underline{M}(l-t)\right.$$

$$\left. + \frac{1}{2}\underline{m}'(l)\cdot \varphi^{-1}(l)\cdot \underline{m}(l)\right\} \frac{K P_G(t,l)}{\sigma(l)\sqrt{2\pi}} \sqrt{\frac{\left|\varphi^{-1}(l)\right|}{\left|\Delta(l-t)\right|}}$$

$$\int_{\underline{X}(l)} d\underline{X}(l) \left\{\sum_{k=1}^{n} \left[\left(\underline{Q}'(l,\cdot) - \underline{T}'(l,T_0)\right)\cdot \underline{\varepsilon_k}\right]^2 \underline{\beta_k}'\right\} \cdot \underline{X}(l)$$

$$\frac{\exp\left\{-\frac{1}{2}\left[\underline{X}(l) - \varphi^{-1}(l)\cdot \underline{m}(l)\right]' \cdot \varphi(l)\cdot \left[\underline{X}(l) - \varphi^{-1}(l)\cdot \underline{m}(l)\right]\right\}}{\sqrt{(2\pi)^n \left|\varphi^{-1}(l)\right|}}.$$

But, the last integral is just the expectation of the random variable $\left\{\sum_{k=1}^{n} \left[\left(\underline{Q}'(l,\cdot) - \underline{T}'(l,T_0)\right)\cdot \underline{\varepsilon_k}\right]^2 \underline{\beta_k}'\right\} \cdot \underline{X}(l)$, with $\underline{X}(l) \backsim N^n\left(\varphi^{-1}(l)\cdot \underline{m}(l), \varphi^{-1}(l)\right)$. Computing such expected value explicitly, the factor-integral independent analytical formula (5.58) is finally obtained.

In order to simplify $V_{11}(l)$, it is convenient to express $\Phi\left[\theta d_1(l)\right]$ as a n-dimensional integral with respect to $\underline{X}(T_0)$. Evaluating (5.52) at $t = T_0$,

$$V_G\left[\underline{X}(T_0), T_0\right] = q\left\{\theta \exp\left[U(T_0,\cdot) + \underline{Q}'(T_0,\cdot)\cdot \underline{X}(T_0)\right] - \theta K\right\}^+,$$

and using result (5.21),

$$
\begin{aligned}
V_G\left[\underline{X}(t), t\right] &= \theta q \int_{\underline{X}(T_0)} d\underline{X}(T_0)\, G\left[\underline{X}(T_0), T_0; \underline{X}(t), t\right] 1_\varepsilon \\
&\quad \left\{\exp\left[U(T_0,\cdot) + \underline{Q}'(T_0,\cdot)\cdot \underline{X}(T_0)\right] - K\right\},
\end{aligned}
$$

where $\varepsilon = \left\{\underline{X}(T_0) : \theta \underline{Q}'(T_0,\cdot)\cdot\underline{X}(T_0) \geq \theta K^*\right\}$. Solving the above integral equation, and comparing each term with (5.52), it can be shown that:

$$
\begin{aligned}
\Phi\left[\theta d_1(l)\right] &= \int_{\underline{X}(T_0)} d\underline{X}(T_0) \frac{1}{\sqrt{(2\pi)^n \left|\Delta(T_0 - l)\right|}} 1_\varepsilon \\
&\quad \exp\left\{-\frac{1}{2}\left[\underline{X}(T_0) - \underline{M}(T_0 - l) - \Delta(T_0 - l)\cdot \underline{Q}(T_0,\cdot)\right]'\right. \\
&\quad \left. \cdot\Delta^{-1}(T_0 - l)\cdot\left[\underline{X}(T_0) - \underline{M}(T_0 - l) - \Delta(T_0 - l)\cdot\underline{Q}(T_0,\cdot)\right]\right\}.
\end{aligned} \tag{5.79}
$$

Combining this last result with (5.75),

$$
\begin{aligned}
V_{11}(l) &= \theta P_G(t,l)\sqrt{\frac{\left|\Omega^{-1}(l)\right|}{\left|\Delta(l-t)\right|}} \exp\left[U(l,\cdot) - \frac{1}{2}\underline{M}'(l-t)\cdot\Delta^{-1}(l-t)\right. \\
&\quad \left.\cdot \underline{M}(l-t) + \frac{1}{2}\underline{\mu_1}'(l)\cdot\Omega^{-1}(l)\cdot\underline{\mu_1}(l)\right] \int_{\underline{X}(T_0)} d\underline{X}(T_0)\frac{1}{\sqrt{(2\pi)^n\left|\Delta(T_0-l)\right|}} 1_\varepsilon \\
&\quad \exp\left\{-\frac{1}{2}\left[\underline{X}(T_0) - \underline{MC_1}(T_0 - l)\right]'\cdot\Delta^{-1}(T_0 - l)\cdot\left[\underline{X}(T_0)\right.\right. \\
&\quad \left.\left. - \underline{MC_1}(T_0 - l)\right]\right\} \int_{\underline{X}(l)} d\underline{X}(l) \frac{\left[\sum_{k=1}^n \left(\underline{Q}'(l,\cdot)\cdot\underline{\varepsilon_k}\right)^2 \underline{\beta_k}'\right]\cdot\underline{X}(l)}{\sqrt{(2\pi)^n\left|\Omega^{-1}(l)\right|}} \\
&\quad \exp\left\{-\frac{1}{2}\left[\underline{X}(l) - \Omega^{-1}(l)\cdot\underline{\mu_1}(l)\right]'\cdot\Omega(l)\cdot\left[\underline{X}(l) - \Omega^{-1}(l)\cdot\underline{\mu_1}(l)\right]\right\},
\end{aligned}
$$

where

$$
\begin{aligned}
\underline{\mu_1}(l) &= \Delta^{-1}(l-t)\cdot\underline{M}(l-t) + \underline{Q}(l,\cdot) \\
&\quad + e^{a'(T_0 - l)}\cdot\Delta^{-1}(T_0 - l)\cdot\left[\underline{X}(T_0) - \underline{MC_1}(T_0 - l)\right].
\end{aligned}
$$

Because the integral with respect to $\underline{X}(l)$ is just the expectation of $\left[\sum_{k=1}^n \left(\underline{Q}'(l,\cdot)\cdot\underline{\varepsilon_k}\right)^2 \underline{\beta_k}'\right]$ $\cdot\underline{X}(l)$, with $\underline{X}(l) \frown N^n\left(\Omega^{-1}(l)\cdot\underline{\mu_1}(l), \Omega^{-1}(l)\right)$, then, and after some linear algebra manipulations,

$$
\begin{aligned}
V_{11}(l) &= \theta P_G(t,l)\sqrt{\frac{\left|\Omega^{-1}(l)\cdot\Psi^{-1}(l)\right|}{\left|\Delta(l-t)\cdot\Delta(T_0 - l)\right|}} \exp\left[U(l,\cdot) - \frac{1}{2}\underline{M}'(l-t)\right. \\
&\quad \cdot\Delta^{-1}(l-t)\cdot\underline{M}(l-t) + \frac{1}{2}\underline{\mu c_1}'(l)\cdot\Omega^{-1}(l)\cdot\underline{\mu c_1}(l) - \frac{1}{2}\underline{MC_1}'(T_0 - l) \\
&\quad \left.\cdot\Delta^{-1}(T_0 - l)\cdot\underline{MC_1}(T_0 - l) + \frac{1}{2}\underline{N_1}'(l)\cdot\Psi^{-1}(l)\cdot\underline{N_1}(l)\right]
\end{aligned}
$$

$$\int_{\underline{X}(T_0)} d\underline{X}(T_0) \frac{1}{\sqrt{(2\pi)^n |\Psi^{-1}(l)|}} 1_\epsilon \left[\sum_{k=1}^{n} \left(\underline{Q}'(l,\cdot) \cdot \underline{\varepsilon_k} \right)^2 \underline{\beta_k}' \right]$$
$$\cdot \Omega^{-1}(l) \cdot \left[\underline{\mu c_1}(l) + e^{a'(T_0-l)} \cdot \Delta^{-1}(T_0 - l) \cdot \underline{X}(T_0) \right]$$
$$\exp\left\{ -\frac{1}{2} \left[\underline{X}(T_0) - \Psi^{-1}(l) \cdot \underline{N_1}(l) \right]' \cdot \Psi(l) \cdot \left[\underline{X}(T_0) - \Psi^{-1}(l) \cdot \underline{N_1}(l) \right] \right\}.$$

Noticing that the expectation of an indicator function results in a probability,

$$\begin{aligned} V_{11}(l) &= \theta P_G(t,l) \exp\left[U(l,\cdot) - \frac{1}{2} \underline{M}'(l-t) \cdot \Delta^{-1}(l-t) \cdot \underline{M}(l-t) \right. \\ &\quad + \frac{1}{2} \underline{\mu c_1}'(l) \cdot \Omega^{-1}(l) \cdot \underline{\mu c_1}(l) - \frac{1}{2} \underline{MC_1}'(T_0 - l) \cdot \Delta^{-1}(T_0 - l) \\ &\quad \left. \cdot \underline{MC_1}(T_0 - l) + \frac{1}{2} \underline{N_1}'(l) \Psi^{-1}(l) \underline{N_1}(l) \right] \sqrt{\frac{|\Omega^{-1}(l) \cdot \Psi^{-1}(l)|}{|\Delta(l-t) \cdot \Delta(T_0 - l)|}} \\ &\quad \left\{ \eta + \left[\sum_{k=1}^{n} \left(\underline{Q}'(l,\cdot) \cdot \underline{\varepsilon_k} \right)^2 \underline{\beta_k}' \right] \cdot \Omega^{-1}(l) \cdot \underline{\mu c_1}(l) \right. \\ &\quad \left. \Pr\left[\theta \underline{Q}'(T_0,\cdot) \cdot \underline{X}(T_0) \geq \theta K^* \right] \right\}, \end{aligned} \tag{5.80}$$

where

$$\begin{aligned} \eta &= \int_{\underline{X}(T_0)} d\underline{X}(T_0) \frac{\left[\underline{C_1}'(l) \cdot \underline{X}(T_0) \right] 1_{\{\theta \underline{Q}'(T_0,\cdot) \cdot \underline{X}(T_0) \geq \theta K^*\}}}{\sqrt{(2\pi)^n |\Psi^{-1}(l)|}} \\ &\quad \exp\left\{ \frac{1}{2} \left[\underline{X}(T_0) - \Psi^{-1}(l) \cdot \underline{N_1}(l) \right]' \cdot \Psi(l) \cdot \left[\underline{X}(T_0) - \Psi^{-1}(l) \cdot \underline{N_1}(l) \right] \right\}, \end{aligned}$$

and $\Pr(A)$ denotes the probability of occurrence of the event A.

Because $\underline{X}(T_0) \backsim N^n\left(\Psi^{-1}(l) \cdot \underline{N_1}(l), \Psi^{-1}(l) \right)$ implies that the random variable $\left[\underline{Q}'(T_0, \cdot) \cdot \underline{X}(T_0) \right]$ possesses a univariate normal distribution with mean $\underline{Q}'(T_0,\cdot) \cdot \Psi^{-1}(l) \cdot \underline{N_1}(l)$ and variance $\underline{Q}'(T_0,\cdot) \cdot \Psi^{-1}(l) \cdot \underline{Q}(T_0,\cdot)$, the probability contained in equation (5.80) corresponds to

$$\Pr\left[\theta \underline{Q}'(T_0,\cdot) \cdot \underline{X}(T_0) \geq \theta K^* \right] = \Phi \left[\theta \frac{\underline{Q}'(T_0,\cdot) \cdot \Psi^{-1}(l) \cdot \underline{N_1}(l) - K^*}{\sqrt{\underline{Q}'(T_0,\cdot) \cdot \Psi^{-1}(l) \cdot \underline{Q}(T_0,\cdot)}} \right]. \tag{5.81}$$

Concerning the term η, and for reasons of analytical tractability, $\underline{C_1}(l)$ is going to be approximated by the vector $\left[\lambda_1(l) \underline{Q}(T_0,\cdot) \right]$, where $\lambda_1(l)$ is chosen as to minimize the Euclidean distance between the two vectors:

$$\begin{aligned} \eta &\cong \int_{\underline{X}(T_0)} d\underline{X}(T_0) \frac{\lambda_1(l) \left[\underline{Q}'(T_0,\cdot) \cdot \underline{X}(T_0) \right] 1_{\{\theta \underline{Q}'(T_0,\cdot) \cdot \underline{X}(T_0) \geq \theta K^*\}}}{\sqrt{(2\pi)^n |\Psi^{-1}(l)|}} \\ &\quad \exp\left\{ \frac{1}{2} \left[\underline{X}(T_0) - \Psi^{-1}(l) \cdot \underline{N_1}(l) \right]' \cdot \Psi(l) \cdot \left[\underline{X}(T_0) - \Psi^{-1}(l) \cdot \underline{N_1}(l) \right] \right\}, \end{aligned}$$

with[30]

$$\lambda_1(l) : \min_{\lambda_1(l)} \left\| \underline{C_1}(l) - \lambda_1(l)\,\underline{Q}(T_0,\cdot) \right\|.$$

The last integral is equal to the expectation of the random variable $\lambda_1(l)\left[\underline{Q}'(T_0,\cdot)\cdot\underline{X}(T_0)\right]$ $1_{\{\theta\underline{Q}'(T_0,\cdot)\underline{X}(T_0)\geq\theta K^*\}}$, subject to $\underline{X}(T_0) \backsim N^n\left(\Psi^{-1}(l)\cdot\underline{N_1}(l), \Psi^{-1}(l)\right)$. To evaluate such expectation, it is simpler to use the density of the random variable $\theta\left[\underline{Q}'(T_0,\cdot)\cdot\underline{X}(T_0)\right] \equiv y$, because the integral under consideration becomes one-dimensional:

$$\begin{aligned}
\eta \;&\cong\; \int_{\theta K^*}^{\infty} dy \frac{\lambda_1(l)}{\theta} y \frac{1}{\sqrt{2\pi\theta^2\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{Q}(T_0,\cdot)}} \\
&\quad \exp\left\{-\frac{1}{2}\frac{\left[y-\theta\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{N_1}(l)\right]^2}{\theta^2\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{Q}(T_0,\cdot)}\right\} \\
&\cong\; \frac{\lambda_1(l)}{\theta}\sqrt{\frac{\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{Q}(T_0,\cdot)}{2\pi}} \\
&\quad \exp\left\{-\frac{1}{2}\frac{\left[K^*-\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{N_1}(l)\right]^2}{\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{Q}(T_0,\cdot)}\right\} + \lambda_1(l) \\
&\quad \left[\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{N_1}(l)\right]\Phi\left[\theta\frac{\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{N_1}(l)-K^*}{\sqrt{\underline{Q}'(T_0,\cdot)\cdot\Psi^{-1}(l)\cdot\underline{Q}(T_0,\cdot)}}\right].
\end{aligned}$$

Combining this last result with equations (5.80) and (5.81) yields the "explicit" solution (5.57) for $i = 1$.

Following exactly the same steps as for $V_{11}(l)$, equation (5.57) can also be derived for $i = 2$. Alternatively, such "explicit" formula for $V_{12}(l)$ also arises by comparing the analytical forms of $V_{11}(l)$ and $V_{12}(l)$ under equations (5.75) and (5.76), as well as the definitions of $d_1(l)$ and $d_0(l)$. In fact, $V_{12}(l)$ can be obtained from $-KV_{11}(l)$ when $U(l,\cdot)$, $\underline{Q}(l,\cdot)$, and $\underline{M}(T_0-l)$ are replaced by $S(l,T_0)$, $\underline{T}(l,T_0)$, and $\left[\underline{M}(T_0-l)-\Delta(T_0-l)\cdot\underline{Q}(T_0,\cdot)\right]$, respectively. Performing these substitutions in equation (5.57) with $i = 1$ yields equation (5.57) for $i = 2$. ■

[30] Note that in the univariate case ($n = 1$), this is not an approximation but an exact result. However, the focus of this dissertation is on the multivariate case.

Chapter 6

A Stochastic Volatility State-Space Formulation For Further Research

6.1 Introduction

Chapter four proposes a Gaussian HJM affine and time-inhomogeneous multifactor model, estimated in two stages: first, the time-independent parameters are obtained from an "equivalent" state-space formulation by applying a non-linear Kalman filter to a panel-data of swap rates, caps' prices and European swaptions' prices (hence, ensuring the stability of the time-homogeneous parameters); then, the time-dependent model' volatility components (that is, function $h(t)$, as defined in equation (4.45)) are calibrated cross-sectionally to caps' prices and/or to European swaptions' prices. This estimation methodology enables the model to recover both the market term structures of interest rates and of volatilities, but it does not provide a satisfactory fit to the observed prices of European swaptions.

As shown in Chapter four, the main reason for the above mentioned model' inability to reproduce swaption prices is the completely time-homogeneous nature of the correlation function amongst interest rates, that arises from the Gaussian feature of the model. Analytically, the argument can be stated as follows: because equation (4.70) does not depend on $h(t)$, then the function $h(t)$ can never improve the model' fit to the market correlation function. In fact, tables 4.12 and 4.21 show that the calibration of $h(t)$ to both caps' and swaptions' prices can only improve the model' fit to the latter by deteriorating the fit to the former.

One possible solution towards the simultaneous and consistent pricing of caps and swaptions could be to estimate the Gaussian HJM model only cross-sectionally, that is not to back-up the model' time-homogeneous parameters from an "equivalent" equilibrium specification. Instead, not only $h(t)$ but also the parameters $\underline{G}$ and K^{Δ} would be calibrated to a cross-section of caps' and swaptions' prices: the burden of recovering the market correlation function would be given to both $\underline{G}$ and K^{Δ}, while function $h(t)$ would mainly fit the term structure of interest rate volatilities. The main problem with this line of attack is that the stability of the time-homogeneous parameters would be lost, as $\underline{G}$ and K^{Δ} would most certainly differ from day to day. In other words, although such approach makes sense for pricing exotics consistently with the market prices of liquid plain-vanilla options, it is not so much appropriate for hedging purposes.

Alternatively, the research already conducted in the present dissertation opens the possibility of implementing a new approach towards the simultaneous fitting of market caps' and swaptions' prices, while still preserving model' parameters stability over time. Specifically, the state-space formulation described in Chapter four can be adapted, using the approximate analytical pricing solutions derived in Chapter five, in order to define a stochastic volatility (instead of Gaussian) HJM affine multifactor model, with time-inhomogeneous diffusion components and still estimated in two stages. The reason why such new stochastic volatility HJM model could, at least theoretically, improve the fit to the market interest rate correlation function (and therefore, to swaptions' prices) relies on the fact that the model' correlation function would now be state-dependent. Using the stochastic volatility specification of equations (2.2) and (2.5), it can be easily shown that the time-t correlation coefficient between changes in instantaneous forward rates of maturities $T_1\,(\geq t)$ and $T_2\,(\geq T_1)$ is equal to

$$
\begin{aligned}
\rho(T_1,T_2) \;=\; & \left[\frac{\partial}{\partial T_1}\underline{B_S}(T_1-t)'\cdot\Sigma\cdot V^D(t)\cdot\Sigma'\cdot\frac{\partial}{\partial T_2}\underline{B_S}(T_2-t)\right] \\
& \left[\frac{\partial}{\partial T_1}\underline{B_S}(T_1-t)'\cdot\Sigma\cdot V^D(t)\cdot\Sigma'\cdot\frac{\partial}{\partial T_1}\underline{B_S}(T_1-t)\right]^{-\frac{1}{2}} \\
& \left[\frac{\partial}{\partial T_2}\underline{B_S}(T_2-t)'\cdot\Sigma\cdot V^D(t)\cdot\Sigma'\cdot\frac{\partial}{\partial T_2}\underline{B_S}(T_2-t)\right]^{-\frac{1}{2}},
\end{aligned} \tag{6.1}
$$

where $\underline{B_S}(T-t)$ satisfies the Riccati differential equation (2.8), and the subscript "S" is used to clearly distinguish the stochastic volatility duration vector from the corresponding Gaussian one (which is given by equation (3.2)). Because the correlation function depends on the state-vector, through matrix $V^D(t)$, it seems realistic to expect a better fit to swaptions' prices, at least by increasing the number of model' factors.

Next sections outline the theoretical implementation of such new stochastic volatility HJM affine and time-inhomogeneous model (estimated in two stages). For reasons to be explained in subsection 6.2.1, the empirical testing of such modelling approach will await for further research.

6.2 State-space specification

Instead of adopting the Duffie and Kan (1996) general stochastic volatility formulation, corresponding to equations (2.2) and (2.5), it is necessary to specify the state-space model in terms of the more restrictive and nested Dai and Singleton (1998) $\mathbf{A}_m(n)$ canonical form ($0 < m \leq n$), which embodies the minimal number of parameters' restrictions needed to ensure that the term structure model is both admissible and just-identified. Following Dai and Singleton (1998, definition III.1), the stochastic volatility model under consideration is given by equations (2.2), (4.1), (4.2), and (4.3), subject to the following restrictions:

$$G_j \geq 0, \qquad j = m+1, \ldots, n, \tag{6.2}$$

where G_j is the j^{th} element of vector $\underline{G} \in \Re^n$,

$$K_{ij} = 0, \qquad i = 1, \ldots, m, \qquad j = m+1, \ldots, n, \tag{6.3}$$

$$K_{ij} \leq 0, \qquad j = 1, \ldots, m, \qquad j \neq i, \tag{6.4}$$

where K_{ij} is the i^{th}-row j^{th}-column element of matrix K,

$$\sum_{j=1}^{m} K_{ij}\theta_j > 0, \qquad i = 1, \ldots, m, \tag{6.5}$$

$$\theta_j = 0, \qquad j = m+1, \ldots, n, \tag{6.6}$$

$$\theta_j \geq 0, \qquad j = 1, \ldots, m, \tag{6.7}$$

where θ_j is the j^{th} element of vector $\underline{\theta} \in \Re^n$,

$$\Sigma = I_n, \tag{6.8}$$

where $I_n \in \Re^{n \times n}$ is an identity matrix,

$$\alpha_j = 0, \qquad j = 1, \ldots, m, \tag{6.9}$$

$$\alpha_j = 1, \qquad j = m+1, \ldots, n, \tag{6.10}$$

where α_j is the j^{th} element of vector $\underline{\alpha} \in \Re^n$, and

$$\beta = \begin{bmatrix} I_{m \times m} & \beta^{DB}_{m \times (n-m)} \\ O_{(n-m) \times m} & O_{(n-m) \times (n-m)} \end{bmatrix}, \tag{6.11}$$

being β^{DB} a matrix of non-negative constants (the subscripts denote the dimension of each matrix).

The above $\mathbf{A}_m(n)$ canonical specification can be easily nested into the more general Duffie and Kan (1996) stochastic volatility formulation of equations (2.2) and (2.5), through the following identities:

$$a = -K, \tag{6.12}$$

and

$$\underline{b} = K \cdot \underline{\theta}. \tag{6.13}$$

For the remaining of this section, the non-linear state-space model described in section 4.3 is simply adapted from the Gaussian $\mathbf{A}_0(n)$ canonical specification to the more general stochastic volatility $\mathbf{A}_m(n)$ form.

6.2.1 Transition equation

Solving equation (4.1) explicitly, using restriction (6.8), considering equally spaced time-periods (that is $t_k - t_{k-1} = h, \forall k$), and assuming that matrix K is non-singular, the following non-Gaussian transition equation is obtained:

$$\underline{X}_k = \underline{J} + F \cdot \underline{X}_{k-1} + \underline{\upsilon}_k, \tag{6.14}$$

where

$$\underline{J} = \left(I_n - e^{-Kh}\right) \cdot \underline{\theta}, \tag{6.15}$$

$$F = e^{-Kh}, \tag{6.16}$$

and

$$\underline{\upsilon}_k = \int_{t_{k-1}}^{t_k} e^{-K(t_k - s)} \cdot \sqrt{V^D(s)} \cdot d\underline{W}^{\mathcal{P}}(s), \tag{6.17}$$

with $\underline{X}_k \equiv \underline{X}(t_k)$. Notice that the error vector $\underline{\upsilon}_k$ is no longer Gaussian (as was the case in section 4.3), and its covariance matrix is now different from $\Delta(h)$.

In order construct the mean square prediction error matrix, at each Kalman filter recursion (see equation (4.29)), it is still necessary to specify the covariance matrix of the error vector $\underline{\upsilon}_k$, for which two approaches are possible. The first and usual approach -see, for instance, Duan and Simonato (1995) or Chen and Scott (1995a)- is based on the conditional covariance matrix of the model' state-vector. Considering such approach and using definition (6.17), it follows that:

$$
\begin{aligned}
Q_k &\equiv COV^S\left(\underline{v}_k | \underline{X}_{k-1}\right) \\
&= \int_{t_{k-1}}^{t_k} e^{-K(t_k-u)} \cdot E_{\mathcal{P}}^S\left[V^D(u) | \underline{X}_{k-1}\right] \cdot e^{-K'(t_k-u)} du,
\end{aligned}
$$

where

$$
\begin{aligned}
& E_{\mathcal{P}}^S\left[V^D(s) | \underline{X}_{k-1}\right] \\
= \; & diag\left\{\alpha_1 + \underline{\beta_1}' \cdot E_{\mathcal{P}}^S\left[\underline{X}(s) | \underline{X}_{k-1}\right], \ldots, \alpha_n + \underline{\beta_n}' \cdot E_{\mathcal{P}}^S\left[\underline{X}(s) | \underline{X}_{k-1}\right]\right\}.
\end{aligned}
$$

Adopting approximation[1] (5.68) and considering equally spaced time-periods, a proxy is then obtained for the error vector' covariance matrix:

$$
Q_k \approx \int_0^h e^{-K(h-u)} \cdot V_{k-1}^D \cdot e^{-K'(h-u)} du, \tag{6.18}
$$

where

$$
V_{k-1}^D = diag\left\{\alpha_1 + \underline{\beta_1}' \cdot \underline{X}_{k-1}, \ldots, \alpha_n + \underline{\beta_n}' \cdot \underline{X}_{k-1}\right\}. \tag{6.19}
$$

Moreover, because $\underline{X}_{k-1}$ is unobservable, Q_k is evaluated at $\hat{\underline{X}}_{k-1}$. In summary, the first approach consists in replacing $\Delta(h)$ by Q_k, as given by equations (6.18) and (6.19), for all the Kalman filter recursions of subsection 4.3.3.

There exist, however, two essential problems with this approach, which justify the postponement of the corresponding empirical analysis for further research:

1. The *inconsistency* and *inefficiency* of the QML model' parameters estimator.

   It is well known -see Lund (1997a, subsection 4.3)- that the QML estimator generated by a standard Kalman filter for a non-Gaussian state-space model is both *inconsistent* and *inefficient*. This is so because although Q_k depends on the lagged unobservable state-vector, $\underline{X}_{k-1}$, the standard Kalman filter can only provide, for stochastic volatility state-space models, not $\hat{\underline{X}}_{k-1}$ but rather the linear projection of $\underline{X}_{k-1}$ on the linear subspace generated by the observed market panel data.

   As Lund (1997a) suggests, two possible solutions for this problem can be either the use of more computationally involved estimation methodologies (such as Markov-Chain Monte Carlo methods and the Simulated Method of Moments) or the adoption of the unrealistic, but simpler, *zero measurement errors assumption* (which would not

[1] Notice that such approximation ensures that $V^D(s)$ is a non-anticipating function. Of course, the smaller is the time-step h, the more acceptable should be such approximation, which constitutes an additional reason for using daily observations in the empirical analysis.

require the filtering of the model' factors). Bearing in mind the additional computation burden involved by the measurement equations of stochastic volatility state-space models, further research is needed in order to optimize the trade-off between the numerical complexity and the model' realism associated with the above mentioned potential solutions.

2. The *positive definiteness* of matrix Q_k.

   An additional difficulty associated to this first approach is to ensure that the covariance matrix Q_k is positive definite, for all k. Because the standard Kalman filter recurrence relations do not incorporate the admissibility restrictions contained in equations (6.2) to (6.11), it is usually necessary to constrain all elements of $\underline{X}_{k-1}$ to be non-negative. However, such constraints will most probably introduce an additional bias in the QML parameters' estimates.

The second approach was proposed by Lund (1997a), and is based on the unconditional covariance matrix of the state-vector. That is, in order to obtain *consistency* for the QML parameters' estimator, Lund (1997a) proposes that Q_k is replaced, for all k, by the unconditional covariance matrix of the error term $\underline{v}_k$:

$$\begin{aligned} Q &\equiv COV^S(\underline{v}_k) \\ &= \int_0^h e^{-K(h-u)} \cdot V^D \cdot e^{-K'(h-u)} du, \end{aligned} \tag{6.20}$$

where

$$V^D = diag\left\{\alpha_1 + \underline{\beta_1}' \cdot \underline{\theta}, \ldots, \alpha_n + \underline{\beta_n}' \cdot \underline{\theta}\right\}. \tag{6.21}$$

Again, the standard Kalman filter recursions described in subsection 4.3.3 are still valid, but now with $\Delta(h)$ replaced by Q. Notice that such procedure implicitly ignores the *conditional heteroskedasticity* properties of the stochastic differential equation (4.1).

The main advantage of this second approach is that, because matrix Q is state-independent (and, therefore, positive definite), the standard Kalman filter possess the MMSLE (minimum mean square linear estimator) property and, hence, the resulting QML parameters' estimator is surely *consistent*. However, and as Lund (1997a, page 16) notices, "...we should expect that the cost of...ensuring consistency is further loss of efficiency". In fact, the Monte Carlo experiments conducted by Lund (1997a) did not find any improvement of his QML estimator' finite sample properties over the usual "conditionally heteroskedastic" one.

In summary, the empirical analysis of the stochastic volatility state-space model described in the present section is postponed until further research is conducted towards the implementation of a more *consistent* and *efficient* estimation methodology.

6.2.2 Non-linear measurement equation

Considering a panel data of swap rates, cap and swaption prices, the measurement equation of the stochastic volatility state-space model is still given by formulae (4.24) and (4.25), being the former linearized using expression (4.26). Each element of vector $\underline{Z}_k$ is defined as in subsection 4.3.2, but now computed from the approximate stochastic volatility solutions proposed in Chapter five.

Swap rates are still obtained from equation (4.15), but the corresponding stochastic volatility pure discount bond prices are given by proposition 22.

Similarly, cap prices are now computed under proposition 27. That is, under the $\mathbb{A}_m(n)$ canonical formulation, the time-t (stochastic volatility) price of a forward cap on a unitary principal, with a cap rate of k, and settled in arrears at times $T_i = t + i\delta, i = 2, \ldots, v$ is

$$\begin{aligned} Cap^S\left(\underline{X}(t); k, \delta, v\right) &\cong (1+\delta k) \sum_{i=1}^{v-1} \Big\{ (1+\delta k)^{-1} P_G\left(\underline{X}(t); i\delta\right) \Phi\left[\sigma_{i\delta,(i+1)\delta} - d(i)\right] \\ &\quad - P_G\left(\underline{X}(t); (i+1)\delta\right) \Phi\left[-d(i)\right] + \frac{1}{2} V_1^i\left[\underline{X}(t), t\right] \Big\}, \end{aligned} \tag{6.22}$$

with

$$d(i) = \frac{\ln\left[\frac{P_G(\underline{X}(t);(i+1)\delta)(1+\delta k)}{P_G(\underline{X}(t);i\delta)}\right] + \frac{\sigma^2_{i\delta,(i+1)\delta}}{2}}{\sigma_{i\delta,(i+1)\delta}}, \tag{6.23}$$

$$\sigma^2_{i\delta,(i+1)\delta} = \underline{B_G}(\delta)' \cdot \Delta(i\delta) \cdot \underline{B_G}(\delta), \tag{6.24}$$

and where $V_1^i\left[\underline{X}(t), t\right]$ is given (for $i = 1, \ldots, v-1$) by equation (5.56) with $q = 1$, $K = (1+\delta k)^{-1}$, $U(t,\cdot) = A_G((i+1)\delta)$, $\underline{Q}(t,\cdot) = \underline{B_G}((i+1)\delta)$, $S(t,T_0) = A_G(i\delta)$, $\underline{T}(t,T_0) = \underline{B_G}(i\delta)$, and $\theta = -1$. In all the previous and the following formulae, the subscript "G" refers to the Gaussian specification of the Duffie and Kan (1996) model (or equivalently, to the nested $\mathbb{A}_0(n)$ canonical formulation). Hence, functions $P_G(\cdot)$, $\underline{B_G}(\cdot)$ and $A_G(\cdot)$ are all computed under proposition 1.

Finally, European swaptions are valued as European options on coupon-bearing bonds, but using a *stochastic duration* approximation and proposition 27. That is, under the $\mathbb{A}_m(n)$ canonical formulation, it can be easily shown that the time-t (stochastic volatility) price of an European payer swaption maturing at time $T_u = t + u\delta$, with a strike equal to x, and on a forward swap with a unitary principal and settled in arrears at times $T_{u+i} = T_u + i\delta, i = 1, \ldots, v$, is

$$Payerswpn\left(\underline{X}(t); x, \delta, u, v\right) \cong \xi\left\{-P_G\left(\underline{X}(t); D_S\right)\Phi(-d)\right. \tag{6.25}$$

$$+\xi^{-1} P_G\left(\underline{X}(t); u\delta\right) \Phi\left(\sigma_{u\delta, D_S} - d\right) + \frac{1}{2} V_1\left[\underline{X}(t), t\right]\Bigg\},$$

with

$$\xi = \sum_{i=1}^{v} k_i \frac{P_S\left(\underline{X}(t); (u+i)\delta\right)}{P_S\left(\underline{X}(t); D_S\right)}, \tag{6.26}$$

$$k_i = 1_{\{i=v\}} + x\delta, \tag{6.27}$$

$$d = \frac{\ln\left[\frac{P_G(\underline{X}(t); D_S)\xi}{P_G(\underline{X}(t); u\delta)}\right] + \frac{\sigma^2_{u\delta, D_S}}{2}}{\sigma_{u\delta, D_S}}, \tag{6.28}$$

$$\sigma^2_{u\delta, D_S} = \underline{B_G}\left(D_S - u\delta\right)' \cdot \Delta(u\delta) \cdot \underline{B_G}\left(D_S - u\delta\right), \tag{6.29}$$

and where $V_1\left[\underline{X}(t), t\right]$ is given by equation (5.56) with $q = 1$, $K = \xi^{-1}$, $U(t, \cdot) = A_G(D_S)$, $\underline{Q}(t, \cdot) = \underline{B_G}(D_S)$, $S(t, T_0) = A_G(u\delta)$, $\underline{T}(t, T_0) = \underline{B_G}(u\delta)$, and $\theta = -1$. The constant D_S corresponds to the *stochastic duration* of the price-process $\sum_{i=1}^{v} k_i P_S\left(\underline{X}(t); (u+i)\delta\right)$, i.e. it is equal to the time-to-maturity of a zero-coupon bond with the same instantaneous variance of relative price changes, and it is implicitly defined as the solution of the following equation:

$$\left\|\sqrt{V^D(t)} \cdot \underline{B_S}(D_S)\right\|^2 = \frac{\left\|\sqrt{V^D(t)} \cdot \sum_{i=1}^{v} \underline{B_S}\left((u+i)\delta\right) k_i P_S\left(\underline{X}(t); (u+i)\delta\right)\right\|^2}{\left[\sum_{i=1}^{v} k_i P_S\left(\underline{X}(t); (u+i)\delta\right)\right]^2}. \tag{6.30}$$

The stochastic volatility duration vector $\underline{B_S}(\cdot)$ and the discount factors $P_S\left(\underline{X}(t); \cdot\right)$ can be numerically computed from equations (2.8) and (2.9), or, alternatively, can be approximated through proposition 22.

6.2.3 Filtering and estimation methods

The parameters of the non-Gaussian state-space model under analysis can still be estimated by maximizing the log-likelihood function (4.27), through the optimization algorithm described in subsection 4.3.4.

For that purpose, the Kalman filter recursions (required for constructing the prediction-error decomposition formula (4.27)) are still given by equations (4.28) to (4.35), but with three modifications: the Gaussian covariance matrix $\Delta(h)$ must be replaced by matrices Q_k or Q, as given by formulae (6.18) or (6.20), respectively; at each iteration, and before computing the vector $\underline{Z}_k$, the affine invariant transformation (5.33) must be applied to $\hat{\underline{X}}_{k|k-1}$ (in order to enhance the accuracy of the pricing approximations described in subsection 6.2.2); and, the "prediction" equation (4.28) must be augmented by the vector $\underline{J}$, that is

$$\hat{\underline{X}}_{k|k-1} = \underline{J} + F \cdot \hat{\underline{X}}_{k-1}. \tag{6.31}$$

Again, the Kalman filter recursions can be initialized at the first two unconditional moments of the model' state variables, and, therefore, equations (4.36) and (4.37) must be replaced by

$$\hat{\underline{X}}_0 = \underline{\theta}, \tag{6.32}$$

and

$$vec\left(P_0\right) = \left(I_{n^2} - F \otimes F\right)^{-1} \cdot vec\left(Q\right), \tag{6.33}$$

respectively.

6.3 HJM specification

The "equilibrium" stochastic volatility model described in the previous sections can be restated in an HJM framework, ensuring automatically a perfect fit to the observed term structure of interest rates. Considering the SDE (2.5), applying Itô's lemma to the pure discount bond price process $P(t,T)$, and using the identity $f(t,T) = -\frac{\partial \ln P(\underline{X}(t);\tau)}{\partial T}$, it is possible to obtain the following "equivalent" HJM model' specification:

$$df\left(t,T\right) = \left[\frac{\partial}{\partial T}\underline{B_S}\left(\tau\right)' \cdot \Sigma \cdot V^D\left(t\right) \cdot \Sigma' \cdot \underline{B_S}\left(\tau\right)\right] dt - \frac{\partial}{\partial T}\underline{B_S}\left(\tau\right)' \cdot \Sigma \cdot \sqrt{V^D\left(t\right)} \cdot d\underline{W}^Q\left(t\right), \tag{6.34}$$

where

$$V^D\left(t\right) = diag\left\{\alpha_1 + \underline{\beta_1}' \cdot \underline{X}\left(t\right), \ldots, \alpha_n + \underline{\beta_n}' \cdot \underline{X}\left(t\right)\right\},$$

and $\underline{X}(t)$ satisfies (2.5). Clearly, the model is arbitrage-free since the drift of the process (6.34) respects the HJM no-arbitrage condition:

$$\begin{aligned} & \frac{\partial}{\partial T}\underline{B_S}\left(\tau\right)' \cdot \Sigma \cdot V^D\left(t\right) \cdot \Sigma' \cdot \underline{B_S}\left(\tau\right) \\ = & \left[\frac{\partial}{\partial T}\underline{B_S}\left(\tau\right)' \cdot \Sigma \cdot \sqrt{V^D\left(t\right)}\right] \cdot \int_t^T \left[\frac{\partial}{\partial u}\underline{B_S}\left(u-t\right)' \cdot \Sigma \cdot \sqrt{V^D\left(u\right)}\right]' du. \end{aligned}$$

Moreover, because formulations (6.34) and (2.5) are "equivalent", all the stochastic volatility pricing formulae derived in Chapter five (as well as the ones described in subsection 6.2.2) are still valid and applicable, subject to the following restriction: the state-vector $\underline{X}(0)$ must be such that

$$f\left(0,T\right) = -\frac{\partial}{\partial T}A_S\left(T\right) - \frac{\partial}{\partial T}\underline{B_S}\left(T\right)' \cdot \underline{X}\left(0\right), \forall T. \tag{6.35}$$

As for the Gaussian HJM model of equation (4.45), and in order to ensure the model' fit to the term structure of interest rate volatilities, the diffusion of the SDE (6.34) can be augmented by a time-dependent function $h : \Re^{+} \longrightarrow \Re$, yielding the following stochastic volatility HJM affine and time-inhomogeneous multifactor model:

$$\begin{aligned} df(t,T) &= \left[h(t)^2 \frac{\partial}{\partial T}\underline{B_S}(\tau)' \cdot \Sigma \cdot V^D(t) \cdot \Sigma' \cdot \underline{B_S}(\tau)\right] dt \\ &\quad - h(t) \frac{\partial}{\partial T}\underline{B_S}(\tau)' \cdot \Sigma \cdot \sqrt{V^D(t)} \cdot d\underline{W}^{\mathbb{Q}}(t). \end{aligned} \tag{6.36}$$

Again, it can be shown that the pricing solutions described in subsection 6.2.2 are still applicable, subject to condition (6.35), as long as equation (5.55) is replaced by

$$\sigma^2(t) = \underline{Q}'(T_0, \cdot) \cdot \left[\int_t^{T_0} h(u)^2 e^{a(T_0-u)} \cdot \Theta \cdot e^{a'(T_0-u)} du\right] \cdot \underline{Q}(T_0, \cdot). \tag{6.37}$$

The estimation of the stochastic volatility HJM model (6.36) can be done, as described in Chapter four for the nested Gaussian HJM specification, in two stages: first, the time-homogeneous parameters ($\underline{G}$, K and β) are estimated from the (underlying) state-space model of section 6.2; and, then function $h(t)$ is calibrated, cross-sectionally, to the market prices of caps and/or European swaptions. It is precisely the first step of this joint estimation method that will await for further research.

Chapter 7

Conclusions

This dissertation was devoted to the study of the class of multifactor affine term structure models characterized by Duffie and Kan (1996), and produced two main contributions: the derivation of new analytical pricing solutions that are faster to implement than the existing numerical methods; and, the formulation of a state-space and panel-data estimation methodology, based on a non-linear Kalman filter, that enables the model' fit not only to the level of the yield curve (as was always the case in the previous literature) but also to the interest rates market covariance matrix.

Firstly, a general equilibrium specification (under objective probabilities) was proposed for the Duffie and Kan (1996) model. Such equivalent specification provided a theoretically founded functional form for the vector of market prices of risk, which enables the estimation of the Duffie and Kan (1996) model' parameters through a time-series or a panel-data approach. This was the first theoretical result of the present thesis.

Secondly, exact closed-form pricing solutions were derived under a Gaussian (nested) specification of the Duffie and Kan (1996) model and for several European-style interest rate contingent claims, namely for: futures on bonds and on short-term interest rates, European options on pure discount bonds, caps and floors, European yield options, and European futures options on bonds and on short-term interest rates. For European options on coupon-bearing bonds and for European swaptions, two approximations were proposed and compared (both in terms of options' moneyness as well as in terms of the maturity' length of the underlying asset). The derivation of these Gaussian exact analytical formulae (second theoretical contribution of the present dissertation) is useful and relevant for, at least, four reasons: *i*) all the pricing solutions can be directly applied to several well-known term structure models previously proposed in the literature, which can be re-stated as nested cases of the more general Gaussian affine specification considered in this thesis; *ii*) the Gaussian exact pricing formulae can also be used as control variates in Monte Carlo implementations of the more general stochastic volatility Duffie and Kan (1996) model' formulation; *iii*) these Gaussian exact analytical solutions provide the measurement equations needed to fit Gaussian affine state-space models not only to the term structure of interest rates but also to the corresponding term structure of volatilities as well as of interest rate correlations; and *iv*) finally, such Gaussian exact solutions can constitute the zeroth order terms from which first order approximate and analytical stochastic volatility pricing formulae are constructed.

Thirdly, using the general equilibrium specification derived in this thesis for the vector of market prices of risk and the Gaussian pricing formulae also obtained in this dissertation, a Gaussian time-homogeneous and affine state-space model was fitted to a panel-data of (US and UK) swap rates, cap prices, and European swaption quotes, through a non-linear Kalman filter algorithm. Therefore, the third theoretical contribution of the present dissertation consisted in implementing a state-space and panel-data estimation methodology that allows the model' fit not only to the level of the yield curve but also to the market interest rates covariance surface (which is implicit in the market prices of non-linear derivatives, such as caps and swaptions). Consistently with the previous literature, and for both data-sets used, an extremely good model' fit to market swap rates was obtained using low-dimensional specifications. However, it was also demonstrated that such state-space affine models are unable to fit short-maturity caps and European swaptions. Moreover, the model' inability to reproduce the market (exponentially-decaying) interest rates correlation function was shown, both empirically and analytically, not to depend on the model' dimensionality. Nevertheless, it was argued that it is still important to consider such cap and swaption market data in the model' estimation, because additional information is incorporated into the model in terms of the term structures of interest rates volatilities and correlations. In order to improve the model' fit to caps and swaptions prices, while preserving the stability of the parameters' estimates, this thesis proposes an HJM and time-inhomogeneous formulation of the affine Gaussian Duffie and Kan (1996) model, which is then estimated in two stages: in a first stage, the model' time-homogeneous parameters' estimates are obtained directly from the corresponding ("equilibrium") non-linear state-space model; then, the time-dependent model' parameters are simply calibrated to the current cross-section of cap and/or swaption prices. This two-step procedure still allows the structural behavior of the interest rate data to be captured by the model' time-homogeneous component, yields an almost exact fit to cap prices, and reduces substantially the mispricing of swaptions.

Finally, the main theoretical contribution of this dissertation consisted in deriving approximate analytical pricing solutions, under the most general stochastic volatility specification of the Duffie and Kan (1996) model, for several European-style interest rate contingent claims, namely for: bonds, long-term and short-term interest rate futures, European options on zero-coupon bonds, caps and floors, European yield options, European futures options on pure discount bonds and on short-term interest rates, and even European swaptions. All these stochastic volatility approximate closed-form solutions were constructed from the exact corresponding Gaussian formulae also derived in this thesis, and are extremely fast and easy to implement since they only involve (no matter the model' dimension) one integral with respect to the time-to-maturity of the interest rate contingent claim under analysis. Moreover, asymptotic error bounds were derived for the approximations suggested in this thesis, and the numerical examples presented in this dissertation also confirm the accuracy of the proposed first order stochastic volatility approximate solutions. Therefore, such approximate pricing formulae constitute an effective and valuable pricing alternative methodology to the existing numerical methods: for high-dimensional models (where Monte Carlo simulation or finite-difference schemes become too time-consuming) and/or when the relevant characteristic function for the pricing problem in hands does not possess a known analytical solution (which makes it difficult to use Fourier transform methods), the Arrow-Debreu approximate pricing methodology proposed in this thesis seems to be, until now, the only and the best approach available.

Bibliography

Aït-Sahalia, Y., 1996a, Nonparametric Pricing of Interest Rate Derivative Securities, *Econometrica* 64, 527–560.

Aït-Sahalia, Y., 1996b, Testing Continuous Time Models of the Spot Interest Rate, *Review of Financial Studies* 9, 385–426.

Amin, K. and A. Morton, 1994, Implied Volatility Functions in Arbitrage-Free Term Strucuture Models, *Journal of Financial Economics* 35, 141–180.

Andersen, T. and J. Lund, 1997, Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate, *Journal of Econometrics* 77, 343–377.

Andreasen, J., P. Dufresne, and W. Shi, 1996, An Arbitrage Term Structure Model of Interest Rates with Stochastic Volatility, Working paper, Aarhus University.

Arnold, L., 1992, *Stochastic Differential Equations: Theory and Applications* (Krieger Publishing Company, Florida).

Babbs, S., 1987, The Valuation of Multivariate Contingent Claims in Continuous-Time Models, Working paper, County NatWest Limited and Imperial College.

Babbs, S., 1991, A Family of Ito Process Models for the Term Structure of Interest Rates, Preprint 91/24, Financial Options Research Centre, University of Warwick.

Babbs, S., 1993, Generalized Vasicek Models of the Term Structure, Working paper, Midland Global Markets.

Babbs, S., 1994, Pricing Short-Term Interest Rate Futures under Deterministic Bond Price Volatilities, Working paper, Midland Global Markets and University of Warwick.

Babbs, S., 1995a, Floating Rate Flow Valuation Under Generalized Vasicek With Application to LIBOR Set In Arrears Swaps, Working paper, First National Bank of Chicago and University of Warwick.

Babbs, S., 1995b, Interest Rate Cap and Floor Valuation Under Generalized Vasicek, Working paper, First National Bank of Chicago and University of Warwick.

Babbs, S. and K. Nowman, 1999, Kalman Filtering of Generalized Vasicek Term Structure Models, *Journal of Financial and Quantitative Analysis* 34, 115–130.

Bakshi, G. and D. Madan, 1998, Spanning and Derivative-Security Valuation, Working paper, University of Maryland.

Bakshi, G. and Z. Chen, 1996, Inflation, Asset Prices, and the Term Structure of Interest Rates in Monetary Economies, *Review of Financial Studies* 9, 241–275.

Bakshi, G. and Z. Chen, 1997a, An Alternative Valuation Model for Contingent Claims, *Journal of Financial Economics* 44, 123–165.

Bakshi, G. and Z. Chen, 1997b, Equilibrium Valuation of Foreign Exchange Claims, *Journal of Finance* 52, 799–826.

Balduzzi, P., S. Das, S. Foresi, and R. Sundaram, 1996, A Simple Approach to Three-Factor Affine Term Structure Models, *Journal of Fixed Income*, December, 43–53.

Bams, D. and P. Schotman, 1998, Direct Estimation of the Risk Neutral Factor Dynamics of Affine Term Structure Models, Working paper, University of Maastricht.

Barone, E. and L. Mengoni, 1997, Futures-Style Options on Euro-Deposit Futures: Nihil Sub Sole Novi?, *European Financial Management* 3, 99–126.

Barone, E., D. Cuoco, and E. Zautzik, 1991, Term Structure Estimation Using the Cox, Ingersoll and Ross Model: The Case of Italian Treasury Bonds, *Journal of Fixed Income* 1, 87–95.

Bates, D., 1993, Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in PHLX Deutschemark Options, Working paper, The Wharton School, University of Pennsylvania.

Baxter, M. and A. Rennie, 1996, *Financial Calculus: An Introduction to Derivative Pricing* (Cambrigde University Press, Cambridge).

Beaglehole, D. and M. Tenney, 1991, General Solutions of some Interest Rate-Contingent Claim Pricing Equations, *Journal of Fixed Income*, September, 69–83.

Berardi, A., 1997, Term Structure, Non-Neutral Inflation and Economic Growth: A Three-Factor Model, Working paper, London Business School.

Berndt, E., B. Hall, R. Hall, and J. Hausman, 1974, Estimation and Inference in Nonlinear Structural Models, *Annals of Economic and Social Measurement* 3/4, 653–665.

Bhar, R. and C. Chiarella, 1996, The Estimation of the Heath-Jarrow-Morton Model by Use of Kalman Filtering Techniques, Working paper, University of Technology, Sydney.

Black, F., 1976, The Pricing of Commodity Contracts, *Journal of Financial Economics* 3, 167–179.

Black, F. and M. Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637–654.

Black, F. and P. Karasinski, 1991, Bond and Option Pricing When Short Rates Are Lognormal, *Financial Analyst Journal*, July-August, 52–59.

Black, F., E. Derman, and W. Toy, 1990, A One-Factor Model of Interest Rates and its Application to Treasury Bond Options, *Financial Analyst Journal*, January-February, 33–39.

Bliss, R. and P. Ritchken, 1995, Empirical Tests of Two State-Variable HJM Models, Working paper 95-13, Federal Reserve Bank of Atlanta.

Boyle, P., 1977, Options: A Monte Carlo Approach, *Financial Economics* 4, 323–338.

Brace, A. and M. Musiela, 1994a, A Multifactor Gauss Markov Implementation of Heath, Jarrow, and Morton, *Mathematical Finance* 4, 259–283.

Brace, A. and M. Musiela, 1994b, Swap Derivatives in a Gaussian HJM Framework, Working paper, School of Mathematics, UNSW.

Brace, A., D. Gatarek, and M. Musiela, 1997, The Market Model of Interest Rate Dynamics, *Mathematical Finance* 7, 127–155.

Breeden, D., 1979, An Intertemporal Capital Pricing Model with Stochastic Investment Opportunities, *Journal of Financial Economics* 7, 265–296.

Breeden, D., 1986, Consumption, Production, Inflation and Interest Rates: A Synthesis, *Journal of Financial Economics* 16, 3–39.

Brennan, M. and E. Schwartz, 1979, A Continuous Time Approach to the Pricing of Bonds, *Journal of Banking and Finance* 3, 133–155.

Briys, E., M. Crouhy, and R. Schöbel, 1991, The Pricing of Default-Free Interest Rate Cap, Floor, and Collar Agreements, *Journal of Finance* 46, 1879–1892.

Brown, R. and S. Schaefer, 1994a, Interest Rate Volatility and the Shape of the Term Structure, Working paper 177, Institute of Finance and Accounting, London Business School.

Brown, R. and S. Schaefer, 1994b, The Term Structure of Real Interest Rates and the Cox, Ingersoll, and Ross Model, *Journal of Financial Economics* 35, 3–42.

Brown, S. and F. Dybvig, 1986, The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates, *Journal of Finance* 41, 617–630.

Bühler, W., M. Uhrig, U. Walter, and Th. Weber, 1995, An Empirical Comparison of Alternative Models for Valuing Interest Rate Options, Working paper 95-11, Universität Mannheim.

Bühler, W., M. Uhrig, U. Walter, and Th. Weber, 1997, Hull/White or Heat/Jarrow/Morton Type Models? An Empirical Comparison of Models for Valuing Interest Rate Options, Working paper 97-06, Universität Mannheim.

Canabarro, E., 1993, Comparing the Dynamic Accuracy of Yield-Curve Based Interest Rate Contingent Claim Pricing Models, *Journal of Financial Engineering* 2, 365–401.

Canabarro, E., 1995, Where Do One-Factor Interest Rate Models Fail?, *Journal of Fixed Income*, September, 31–52.

Carr, P. and R.-R. Chen, 1997, Valuing Bond Futures and the Quality Option, Working paper, Morgan Stanley and Rutgers University.

Carverhill, A., 1994, When Is the Short Rate Markovian?, *Mathematical Finance* 4, 305–312.

Carverhill, A. and L. Clewlow, 1990, Flexible Convolution, *Risk* 3, 25–29.

Chan, K., A. Karolyi, F. Longstaff, and A. Sanders, 1992, An Empirical Comparison of Alternative Models of Short Term Interest Rates, *Journal of Finance* 47, 1209–1227.

Chen, L., 1994, Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure and its Application in Pricing of Interest Rate Derivatives, Working paper, Federal Reserve Board.

Chen, L., 1996, *Interest Rate Dynamics, Derivatives Pricing and Risk Management*, Lecture Notes in Economics and Mathematical Systems 435 (Spinger-Verlag, New York).

Chen, R.-R., 1992, Exact Solutions for Futures and European Futures Options on Pure Discount Bonds, *Journal of Financial and Quantitative Analysis* 27, 97–107.

Chen, R.-R., 1995, A Two-Factor Preference Model for Interest Rate Sensitive Claims, *Journal of Futures Markets* 15, 345–372.

Chen, R.-R. and L. Scott, 1992, Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure, *Review of Financial Studies* 5, 613–636.

Chen, R.-R. and L. Scott, 1993a, Pricing Interest Rate Futures Options with Futures-Style Margining, *Journal of Futures Markets* 13, 15–22.

Chen, R.-R. and L. Scott, 1993b, Maximum Likelihood Estimation for a Multifactor General Equilibrium Model of the Term Structure of Interest Rates, *Journal of Fixed Income*, December, 14–31.

Chen, R.-R. and L. Scott, 1995a, Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests from a Kalman Filter Model, Working paper, University of Georgia.

Chen, R.-R. and L. Scott, 1995b, Interest Rate Options in Multifactor Cox-Ingersoll-Ross Models of the Term Structure, *Journal of Derivatives*, Winter, 53–72.

Cheng, S., 1991, On the Feasibility of Arbitrage-Based Option Pricing When Stochastic Bond Price Processes Are Involved, *Journal of Economic Theory* 53, 185–198.

Cheyette, O., 1995, Markov Representation of the Heath-Jarrow-Morton Model, Working paper, BARRA Inc.

Claessens, S. and G. Pennacchi, 1996, Estimating the Likelihood of Mexican Default from the Market Prices of Brady Bonds, *Journal of Financial and Quantitative Analysis* 31, 109–126.

Clewlow, L. and A. Carverhill, 1994, Quicker on the Curves, *Risk Magazine* 7.

Clewlow, L. and C. Strickland, 1998a, *Implementing Derivatives Models*, Financial Engineering (John & Wiley and Sons).

Clewlow, L. and C. Strickland, 1998b, Pricing Interest Rate Exotics in Multi-Factor Gaussian Interest Rate Models, Working paper, Financial Options Research Centre, University of Warwick.

Clewlow, L., K. Pang, and C. Strickland, 1996, Efficient Pricing of Caps and Swaptions in a Multi-Factor Gaussian Interest Rate Model, Working paper, Financial Options Research Centre, University of Warwick.

Constantinides, G., 1992, A Theory of the Nominal Term Structure of Interest Rates, *Review of Financial Studies* 5, 531–552.

Constantinides, G. and J. Ingersoll, 1984, Optimal Bond Trading with Personal Taxes, *Journal of Financial Economics* 13, 299–335.

Cooper, I. and R. Rebonato, 1995, The Limitations of Simple Two-Factor Interest Rate Models, Working paper 200-95, Institute of Finance and Accounting, London Business School.

Cox, D. and H. Miller, 1995, *The Theory of Stochastic Processes* (Chapman & Hall).

Cox, J., J. Ingersoll, and S. Ross, 1981a, A Re-Examination of Traditional Hypotheses About the Term Structure of Interest Rates, *Journal of Finance* 36, 769–799.

Cox, J., J. Ingersoll, and S. Ross, 1981b, The Relation Between Forward Prices and Futures Prices, *Journal of Financial Economics* 9, 321–346.

Cox, J., J. Ingersoll, and S. Ross, 1985a, An Intertemporal General Equilibrium Model of Asset Prices, *Econometrica* 53, 363–384.

Cox, J., J. Ingersoll, and S. Ross, 1985b, A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385–407.

Dai, Q. and K. Singleton, 1998, Specification Analysis of Affine Term Structure Models, Working paper, New York University and Stanford University.

Davis, R., 1973, Numerical Inversion of a Characteristic Function, *Biometrika* 60, 415–417.

Dennis, J. and R. Schnabel, 1996, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (SIAM).

Dothan, U., 1978, On the Term Structure of Interest Rates, *Journal of Financial Economics* 6, 59–69.

Duan, J.-C., 1994, Maximum Likelihood Estimation Using Price Data of the Derivative Contract, *Mathematical Finance* 4, 155–167.

Duan, J.-C. and J.-G. Simonato, 1995, Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter, Working paper, McGill University.

Duffie, D., 1989, *Futures Markets* (Prentice-Hall, Englewood Cliffs, New Jersey).

Duffie, D., 1992, *Dynamic Asset Pricing Theory* (Princeton University Press, Princeton, New Jersey).

Duffie, D. and K. Singleton, 1997, An Econometric Model of the Term Structure of Interest-Rate Swap Yields, *Journal of Finance* 52, 1287–1321.

Duffie, D. and R. Kan, 1994, Multi-Factor Term Structure Models, *Phil. Trans. R. Soc. Lond.* 347, 577–586.

Duffie, D. and R. Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6, 379–406.

Duffie, D., J. Pan, and K. Singleton, 1998, Transform Analysis and Option Pricing for Affine Jump - Diffusions, Working paper, Graduate School of Business, Stanford University.

Dybvig, P., 1989, Long Forward Rates Can Never Fall, Working paper, Yale School of Management.

Dybvig, P. and L. Rogers, 1995, Recovery of Preferences from Observed Wealth in a Single Realization, Working paper, Washington University and University of Bath.

Economou, E., 1983, *Green's Functions in Quantum Physics*, Springer Series in Solid-State Sciences (Springer-Verlag).

Edwards, C. and D. Penney, 1993, *Elementary Differential Equations with Boundary Value Problems*, third ed. (Prentice-Hall).

El Karoui, N. and J.-C. Rochet, 1989, A Pricing Formula for Options on Coupon Bonds, Working paper 72, SEEDS.

El Karoui, N. and V. Lacoste, 1992, Multifactor Models of the Term Structure of Interest Rates, Working paper, Université de Paris VI.

El Karoui, N., C. Lepage, R. Myneni, N. Roseau, and R. Viswanathan, 1991, The Valuation and Hedging of Contingent Claims with Markovian Interest Rates, Working paper, Université de Paris VI.

El Karoui, N., R. Myneni, and R. Viswanathan, 1992a, Arbitrage Pricing and Hedging of Interest Rate Claims with State Variables: I Theory, Working paper, Université de Paris VI and Stanford University.

El Karoui, N., R. Myneni, and R. Viswanathan, 1992b, Arbitrage Pricing and Hedging of Interest Rate Claims with State Variables: II Applications, Working paper, Université de Paris VI and Stanford University.

Flesaker, B., 1993a, Arbitrage Free Pricing of Interest Rate Futures and Forward Contracts, *Journal of Futures Markets* 13, 77–91.

Flesaker, B., 1993b, Testing the Heat-Jarrow-Morton/Ho-Lee Model of Interest Rate Contingent Claims Pricing, *Journal of Financial and Quantitative Analysis* 28, 483–495.

Fong, H. and O. Vasicek, 1991a, Fixed-Income Volatility Management, *Journal of Portfolio Management*, Summer, 41–46.

Fong, H. and O. Vasicek, 1991b, Interest Rate Volatility as a Stochastic Factor, Working paper, Gifford Fong Associates.

Fong, H. and O. Vasicek, 1992, Omission Impossible, *Risk* 5, 62–65.

Frachot, A. and J.-P. Lesne, 1993, Econometrics of Linear Factor Models of Interest Rates, Working paper, Banque de France.

Frachot, A., D. Janci, and V. Lacoste, 1992, Factor Analysis of the Term Structure: A Probabilistic Approach, Working paper 21, Direction Générale Des Études, Banque de France.

Freidlin, M., 1985, *Functional Integration and Partial Differential Equations*, Annals of Mathematics Studies 109 (Princeton University Press, Princeton, New Jersey).

Friedman, A., 1964, *Partial Differential Equations of Parabolic Type* (Prentice-Hall, Englewood Cliffs, N. J.).

Friedman, A., 1975, *Stochastic Differential Equations and Applications*, Vol. 1 (Academic Press, New York).

Friend, I. and M. Blume, 1975, The Demand for Risky Assets, *American Economic Review* 65, 900–922.

Geman, H., N. El Karoui, and J.-C. Rochet, 1995, Changes of Numeraire, Changes of Probability Measure and Option Pricing, *Journal of Applied Probability* 32, 443–458.

Geyer, A. and S. Pichler, 1996, A State-Space Approach to Estimate and Test Muti-Factor Cox-Ingersoll-Ross Models of the Term Structure, Working paper, University of Economics Vienna.

Gibbons, M. and K. Ramaswamy, 1993, A Test of the Cox, Ingersoll and Ross Model of the Term Structure, *Review of Financial Studies* 6, 619–658.

Gihman, I. and A. Skorohod, 1972, *Stochastic Differential Equations* (Springer-Verlag, New York).

Goldstein, R., 1997, Identifying the Class of Term Structure Models Possessing Closed-Form Solutions for Bond and Bond-Option Prices: An Expectation Approach, Working paper, Ohio State University.

Goldstein, R. and F. Zapatero, 1996, General Equilibrium with Constant Relative Risk Aversion and Vasicek Interest Rates, *Mathematical Finance* 6, 331–340.

Hamilton, J., 1994, *Time Series Analysis* (Princeton University Press, Princeton, New Jersey).

Hansen, A. and P. Jorgensen, 1998, Exact Analytical Valuation of Bonds When Spot Interest Rates Are Log-Normal, Working paper, University of Aarhus.

Harrison, J. and D. Krepps, 1979, Martingales and Arbitrage in Multiperiod Securities Markets, *Journal of Economic Theory* 20, 381–408.

Harrison, J. and S. Pliska, 1981, Martingales and Stochastic Integrals in the Theory of Continuous Trading, *Stochastic Processes and Their Applications* 11, 215–260.

Harrison, J. and S. Pliska, 1983, A Stochastic Calculus Model of Continuous Trading: Complete Markets, *Stochastic Processes and Their Applications* 15, 313–316.

Harvey, A., 1989, *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge University Press).

Heath, D., R. Jarrow, and A. Morton, 1990, Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation, *Journal of Financial and Quantitative Analysis* 25, 419–440.

Heath, D., R. Jarrow, and A. Morton, 1992, Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation, *Econometrica* 60, 77–105.

Heston, S., 1988, Testing Continuous Time Models of the Term Structure of Interest Rates, Working paper, Carnegie-Melon University.

Heston, S., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 327–343.

Ho, T. and S. Lee, 1986, Term Structure Movements and Pricing Interest Rate Contingent Claims, *Journal of Finance* 41, 1011–1029.

Honoré, P., 1998, Panel-Data Estimation of Non-Linear Term-Structure Models, Working paper, The Aarhus School of Business.

Howitt, P., 1974, Stability and the Quantity Theory, *Journal of Political Economy* 82, 133–151.

Huang, C., 1987, An Intertemporal General Equilibrium Asset Pricing Model: The Case of Diffusion Information, *Econometrica* 55, 117–142.

Hull, J., 1997, *Options, Futures and Other Derivatives*, third ed. (Prentice-Hall).

Hull, J. and A. White, 1990, Pricing Interest Rate Derivative Securities, *Review of Financial Studies* 3, 573–592.

Hull, J. and A. White, 1994, Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models, *Journal of Derivatives*, Winter, 37–48.

Hunt, P., J. Kennedy, and A. Pelsser, 1997, Markov-Functional Interest Rate Models, Working paper, ABN-Amro Bank and University of Oxford.

Ingersoll, J., 1987, *Theory of Financial Decision Making*, Studies in Financial Economics (Rowman & Littlefield).

Inui, K. and M. Kijima, 1998, A Markovian Framework in Multi-Factor Heath-Jarrow-Morton Models, *Journal of Financial and Quantitative Analysis* 33, 423–440.

Jalali, P. and H. Kazemi, 1998, A New Approach to the Valuation of Interest Rate Derivatives: Arrow-Debreu Prices Implicit in the Term Structure of Interest Rates, Working paper, University of Massachusetts.

Jamshidian, F., 1989, An Exact Bond Option Pricing Formula, *Journal of Finance* 44, 205–209.

Jamshidian, F., 1990, The Preference-Free Determination of Bond and Option Prices from the Spot Interest Rate, *Advances in Futures and Options Research* 4, 51–67.

Jamshidian, F., 1991, Bond and Option Evaluation in the Gaussian Interest Rate Model, *Research in Finance* 9, 131–170.

Jamshidian, F., 1992, Option and Futures Evaluation with Stochastic Interest Rates, Working paper, Fuji International Finance.

Jamshidian, F., 1993, Option and Futures Evaluation with Deterministic Volatilities, *Mathematical Finance* 3, 149–159.

Jazwinski, A., 1970, *Stochastic Processes and Filtering Theory* (Academic Press).

Jeffrey, A., 1995, Single Factor Heath-Jarrow-Morton Term Structure Models Based on Markov Spot Interest Rate Dynamics, *Journal of Financial and Quantitative Analysis* 30, 619–642.

Jegadeesh, N. and G. Pennacchi, 1996, The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures, *Journal of Money, Credit, and Banking* 28, 426–446.

Jong, F., 1998, Time-Series and Cross-Section Information in Affine Term Structure Models, Working paper, University of Amsterdam and CEPR.

Jong, P., 1988, The Likelihood for a State Space Model, *Biometrika* 75, 165–169.

Jong, P., 1991, The Diffuse Kalman Filter, *Annals of Statistics* 19, 1073–1083.

Ju, N., 1997, Fourier Transformation, Martingale, and the Pricing of Average-Rate Derivatives, Working paper, Haas School of Business, University of California.

Judge, G., W. Griffiths, R. Hill, and T. Lee, 1980, *The Theory and Practice of Econometrics* (John Wiley & Sons).

Kappi, J., 1997, Testing Affine Yield-Factor Models, *Journal of Financial Engineering* 6, 331–356.

Karatzas, I., 1989, Optimization Problems in the Theory of Continuous Trading, *SIAM Journal of Control and Optimization* 27, 1221–1259.

Karatzas, I. and S. Shreve, 1975, *Brownian Motion and Stochastic Calculus* (Springer-Verlag).

Karlin, S. and H. Taylor, 1975, *A First Course in Stochastic Processes*, second ed. (Academic Press).

Kennedy, D., 1991, The Term Structure of Interest Rates as a Gaussian Random Field, Working paper, University of Cambridge.

Kennedy, D., 1997, Characterizing Gaussian Models of the Term Structure of Interest Rates, *Mathematical Finance* 7, 107–118.

Kitagawa, G., 1987, Non-Gaussian State-Space Modeling of Nonstationary Time Series, *Journal of the American Statistical Association* 82, 1032–1063.

Kloeden, P. and E. Platen, 1992, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, New York).

Koopman, S. and J. Durbin, 1998, Fast Filtering and Smoothing for Multivariate State Space Models, Working paper, Tilburg University and London School of Economics and Political Science.

Lamberton, D. and B. Lapeyre, 1996, *Introduction to Stochastic Calculus Applied to Finance* (Chapman & Hall/CRC, New York).

Langetieg, T., 1980, A Multivariate Model of the Term Structure, *Journal of Finance* 35, 71–97.

Leblanc, B. and O. Scaillet, 1995, Options on Forward and Futures Contracts in the Affine Term Structure Model, *Advances in Futures and Options Research* 8, 241–261.

Leblanc, B. and O. Scaillet, 1998, Path Dependent Options on Yields in the Affine Term Structure Model, *Finance and Stochastics* 2, 349–367.

Lieu, D., 1990, Option Pricing with Futures-Style Margining, *Journal of Futures Markets* 10, 327–338.

Litterman, R. and J. Scheinkman, 1991, Common Factors Affecting Bond Returns, *Journal of Fixed Income* 1, 54–61.

Longstaff, F. and E. Schwartz, 1992a, Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, *Journal of Finance* 47, 1259–1282.

Longstaff, F. and E. Schwartz, 1992b, A Two Factor Interest Rate Model and Contingent Claims Valuation, *Journal of Fixed Income*, December, 16–23.

Lu, B., 1997, Affine Models of the Term Structure: Specification Testing and Bond Portfolio Hedging, Working paper, Yale University.

Lucas, R., 1978, Asset Pricing in an Exchange Economy, *Econometrica* 46, 1426–1446.

Lucas, R., 1982, Interest Rates and Currency Prices in a Two-Country World, *Journal of Monetary Economics* 10, 335–359.

Lund, J., 1994, Econometric Analysis of Continuous-Time Arbitrage-Free Models of the Term Structure of Interest Rates, Working paper, The Aarhus School of Business.

Lund, J., 1997a, Econometric Analysis of Continuous-Time Arbitrage-Free Models of the Term Structure of Interest Rates, Working paper, The Aarhus School of Business.

Lund, J., 1997b, Non-Linear Kalman Filtering Techniques for Term Structure Models, Working paper, The Aarhus School of Business.

Maghsoodi, Y., 1996, Solution of the Extended CIR Term Structure and Bond Option Valuation, *Mathematical Finance* 6, 89–109.

Marsh, T., 1980, Equilibrium Term Structure Models: Test Methodology, *Journal of Finance* 35, 421–438.

Marshall, P., 1990, State Space Models with Diffuse Initial Conditions, *Journal of Time Series Analysis* 13, 411–414.

Merton, R., 1969, Lifetime Portfolio Selection under Uncertainty, *Review of Economics and Statistics* 51, 247–257.

Merton, R., 1970, A Dynamic General Equilibrium Model of the Asset Markets and its Application to the Pricing of the Capital Structure of the Firm, Working paper, Sloan School of Management, Massachusetts Institute of Technology.

Merton, R., 1971, Optimal Consumption and Portfolio Rules in a Continuous Time Model, *Journal of Economic Theory* 3, 373–413.

Merton, R., 1973a, The Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* 4, 141–183.

Merton, R., 1973b, An Intertemporal Capital Asset Pricing Model, *Econometrica* 41, 867–887.

Merton, R., 1994, *Continuous-Time Finance* (Blackwell).

Moraleda, J. and A. Pelsser, 1997, Forward vs Spot Interest-Rate Models of the Term Structure: An Empirical Comparison, Working paper, Erasmus University Rotterdam.

Munk, C., 1998, Stochastic Duration and Fast Coupon Bond Option Pricing in Multi-Factor Models, Working paper, Odense University.

Munnik, J. and P. Schotman, 1994, Cross-Sectional Versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market, *Journal of Banking and Finance* 18, 997–1025.

Murdock, J., 1991, *Perturbations: Theory and Methods* (John Wiley & Sons, New York).

Musiela, M. and M. Rutkowski, 1998, *Martingale Methods in Financial Modelling*, Applications of Mathematics: Stochastic Modelling and Applied Probability 36 (Springer-Verlag).

Nayfeh, A., 1973, *Perturbation Methods* (John Wiley & Sons, New York).

Nunes, J., 1998, Interest Rate Options in a Duffie-Kan Model with Deterministic Volatility, *Revista de Mercados e Activos Financeiros* 1, 63–101.

Nunes, J. and L. Clewlow, 1999, Kalman Filtering of Gaussian Exponential-Affine LIBOR-Rate Models Using Caps and Swaptions, Working paper, Warwick Business School, University of Warwick.

Nunes, J., L. Clewlow, and S. Hodges, 1999, Interest Rate Derivatives in a Duffie and Kan Model with Stochastic Volatility: An Arrow-Debreu Pricing Approach, *Review of Derivatives Research* 3, 5–66.

Øksendal, B., 1995, *Stochastic Differential Equations: An Introduction with Applications*, fourth ed. (Springer-Verlag).

Pang, K., 1996, Can We Price Caps and Swaptions Consistently?, Working paper, Financial Options Research Centre, University of Warwick.

Pang, K. and S. Hodges, 1995, Non-Negative Affine Yield Models of the Term Structure, Working paper, Financial Options Research Centre, University of Warwick.

Pearson, N. and T.-S. Sun, 1993, An Empirical Examination of the Cox, Ingersoll and Ross Model of the Term Structure of Interest Rates Using the Method of Maximum Likelihood, *Journal of Finance* 54, 929–959.

Pearson, N. and T.-S. Sun, 1994, Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, Ross Model, *Journal of Finance* 49, 1279–1304.

Press, W., B. Flannery, S. Teukolsky, and W. Vetterling, 1994, *Numerical Recipes in Pascal: The Art of Scientific Computing* (Cambrigde University Press, Cambridge).

Protter, P., 1990, *Stochastic Integration and Differential Equations* (Springer-Verlag).

Rady, S. and K. Sandmann, 1993, The Direct Approach to Debt Option Pricing, Discussion paper b-212, London School of Economics and University of Bonn.

Rebonato, R., 1998, *Interest-Rate Option Models: Understanding, Analysing and Using Models for Exotic Interest-Rate Options*, second ed. (John Wiley & Sons).

Revuz, D. and M. Yor, 1991, *Continuous Martingales and Brownian Motion* (Springer-Verlag).

Richard, S., 1978, An Arbitrage Model of the Term Structure of Interest Rates, *Journal of Financial Economics* 6, 33–57.

Rogers, C., 1996, Gaussian Errors, *Risk* 9, 42–45.

Rogers, L. and W. Stummer, 1994, How Well Do One-Factor Models Fit Bond Prices?, Working paper, School of Mathematical Sciences, University of Bath.

Sandmann, K. and D. Sondermann, 1997, A Note on the Stability of Lognormal Interest Rate Models and the Pricing of Eurodollar Futures, *Mathematical Finance* 7, 119–125.

Sandmann, K., D. Sondermann, and K. Miltersen, 1994, Closed Form Term Structure Derivatives in a Heath-Jarrow-Morton Model with Log-Normal Annually Compounded Interest Rates, Working paper b-285, University of Bonn and Odense University.

Schaefer, S. and E. Schwartz, 1984, A Two Factor Model of the Term Structure: An Approximate Analytical Solution, *Journal of Financial and Quantitative Analysis* 19, 413–424.

Schlögl, E. and D. Sommer, 1997, Factor Models and the Shape of the Term Structure, Discussion paper b-395, University of Bonn.

Schlögl, E. and D. Sommer, 1998, Factor Models and the Shape of the Term Structure, *Journal of Financial Engineering* 7, 79–88.

Schöbel, R., 1990, Options on Short Term Interest Rate Futures, Working paper, Universität Lüneburg.

Schöbel, R. and J. Zhu, 1998, Stochastic Volatility With an Ornstein-Uhlenbeck Process: An Extension, Working paper, Tubingen University.

Scott, L., 1995, The Valuation of Interest Rate Derivatives in a Multi-Factor Term Structure Model With Deterministic Components, Working paper, University of Georgia.

Scott, L., 1997, Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods, *Mathematical Finance* 7, 413–426.

Shephard, N., 1991a, From Characteristic Function to Distribution Function: A Simple Framework for the Theory, *Econometric Theory* 7, 519–529.

Shephard, N., 1991b, Numerical Integration Rules for Multivariate Inversions, *Journal of Statistical Computation and Simulation* 39, 37–46.

Singh, M., 1995, Estimation of Multifactor Cox, Ingersoll and Ross Term Structure Model: Evidence on Volatility Structure and Parameter Stability, *Journal of Fixed Income*, September, 8–28.

Stapleton, R. and M. Subrahmanyam, 1993, The Analysis and Valuation of Interest Rate Options, *Journal of Banking and Finance* 17, 1079–1095.

Stapleton, R. and M. Subrahmanyam, 1998, Arbitrage Restrictions and Multi-Factor Models of the Term Structure of Interest Rates, Working paper, Lancaster University and New York University.

Steenkiste, R. and S. Foresi, 1999, Arrow-Debreu Prices for Affine Models, Working paper, Salomon Smith Barney and Goldman Sachs Asset Management.

Strickland, C., 1993, Interest Rate Volatility and the Term Structure of Interest Rates, Preprint 93/37, Financial Options Research Centre, University of Warwick.

Strickland, C., 1994, A Comparison of Models for Pricing Interest Rate Derivative Securities, Preprint 94/47, Financial Options Research Centre, University of Warwick.

Sun, T.-S., 1992, Real and Nominal Interest Rates: A Discrete-Time Model and Its Continuous-Time Limit, *Review of Financial Studies* 5, 581–611.

Sundaram, R., 1997, Equivalent Martingale Measures and Risk-Neutral Pricing: An Expository Note, *Journal of Derivatives*, Fall, 85–98.

Tanizaki, H., 1996, *Nonlinear Filters* (Springer Verlag).

Tsiang, S., 1956, Liquidity Preference and Loanable Funds Theories, Multiplier and Velocity Analysis: A Synthesis, *American Economic Review* 46, 539–564.

Van Loan, C., 1978, Computing Integrals Involving the Matrix Exponential, *IEEE Transactions on Automatic Control* 23, 395–404.

Vasicek, O., 1977, An Equilibrium Characterization of the Term Structure of Interest, *Journal of Financial Economics* 5, 177–188.

Vila, J.-L. and T. Zariphopoulou, 1995, Optimal Consumption and Portfolio Choice with Borrowing Constraints, Working paper, Goldman Sachs International and University of Wisconsin-Madison.

Watson, M. and R. Engle, 1983, Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying Coefficient Regression Models, *Journal of Econometrics* 23, 385–400.

Wei, J., 1997, A Simple Approach to Bond Option Pricing, *Journal of Futures Markets* 17, 131–160.

Williams, D., 1991, *Probability with Martingales* (Cambridge University Press, Cambridge).

Wilmott, P., 1998, *Derivatives: The Theory and Practice of Financial Engineering* (John Wiley & Sons, Chichester).

Wilmott, P., J. Dewynne, and S. Howison, 1996, *Option Pricing: Mathematical Models and Computation* (Oxford Financial Press, Oxford).