Analytical Solution of the Characteristic Function in the Trolle-Schwartz Model

Richard John van Gysen

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

February 10, 2019

MPhil in Mathematical Finance,
University of Cape Town.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy to the University of Cape Town. It has not before been submitted for any degree or examination.

Signed by candidate

Richard van Gysen

February 10, 2019

Abstract

In 2009, Trolle and Schwartz (2008) produced an instantaneous forward interest rate model with several stylised facts such as stochastic volatility. They derived pricing formulae in order to price bonds and bond options, which can be altered to price interest rate options such as caplets, caps and swaptions. These formulae involve implementing numerical methods for solving an ordinary differential equation (ODE). Schumann (2016) confirmed the accuracy of the pricing formulae in the Trolle and Schwartz (2008) model using Monte-Carlo simulation. Both authors used a numerical ODE solver to estimate the ODE. In this dissertation, a closed-form solution for this ODE is presented. Two solutions were found. However, these solutions rely on a simplification of the instantaneous volatility function originally proposed in the Trolle and Schwartz (2008) model. This case happens to be the stochastic volatility version of the Hull and White (1990) model. The two solutions are compared to an ODE solver for one stochastic volatility term and then extended to three stochastic volatility terms.

Acknowledgements

I would like to kindly thank my supervisors Thomas McWalter and Joerg Kienitz for challenging me with this task. They have continuously supported and encouraged me the entire way along this journey. To my dearest classmates, the endless banter and humour made the last year thoroughly enjoyable. Finally, to my loving family and friends who are always supporting me at home in Cape Town.

Spectemur Agendo - "Let us be judged by our actions."

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Interest rate modelling has evolved significantly in the last few decades. Interest rate models vary considerably in their dynamics and methods to price interest rate derivatives. The simpler the model, the easier it is to find an analytical solution. However, the model might not adequately reflect market dynamics and sometimes no matter how parameters are chosen, it may not reflect observed market values. This evolution has seen an increased complexity of models. However, many of these more accurate, complex models do not have tractable solutions. Thus, there exists a trade-off between expediency and accuracy of models. An interesting feature of interest rate modelling has been the inclusion of a volatility state variable in the diffusion term. The inclusion of such a variable is meaningful as it better reflects the stochastic nature of volatility in derivative pricing (Brigo and Mercurio, 2007). It also provides a reasonable solution to price skews observed in markets (see Corrado and Su (1997) for further discussion). Heath, Jarrow and Morton (1992) (HJM) proposed a general framework for modelling the instantaneous forward rate with n dimensional diffusion terms. Trolle and Schwartz (2008) (TS) extended this model to include n stochastic volatility variables into the diffusion terms. The forward rates are correlated with these stochastic volatility variables, which are driven by their own dynamics. The authors then provide semi-analytical solutions for simple claims. These solutions are semi-analytical as they require numerical methods to solve them and rely on the unobserved stochastic volatility variables. These unobserved variables fall under the Duffie and Kan (1996) class of affine dynamic term structure models. Trolle and Schwartz (2008) proved the pricing accuracy of their model and compared it to several other authors with stochastic volatility models. Schumann (2016) used Monte-Carlo simulation in conjunction with the TS model to price bonds and interest rate derivatives. Recently, Sitzia (2018) used the TS model to price commodity derivatives and, in that context, found an analytical solution for ordinary differential equations (ODEs) that describe the characteristic function.

The Heath-Jarrow-Morton Framework

Before the 1990s, interest rate modelling had used relatively simple models as proposed by several authors. Many models would be simple with constant coefficients in their dynamics, thus time-homogeneous, for the sake of tractability. For example, the Vasicek (1977) and Cox, Ingersoll Jr and Ross (1985) interest rate models have closed-form solutions that can price bonds and bond options. However, these models are problematic in that they generate an endogenous term structure of interest rates. This means that the initial term structure cannot fit the observed market rates regardless of parameter value choice (Brigo and Mercurio, 2007). More complex models were needed to replicate market observed rates but no general framework existed in literature yet. A major breakthrough occurred with the Heath, Jarrow and Morton (HJM) framework, who created a general and consistent framework for instantaneous forward rates. HJM originally specified a two-factor volatility model, within a discrete-time framework (see Heath, Jarrow and Morton (1989) and Heath, Jarrow and Morton (1990)). They extended this research and constructed an arbitrage-free framework for the stochastic evolution of a continuous time yield curve. Specifically, forward rates can be determined by their volatility structure (Brigo and Mercurio, 2007). Heath, Jarrow and Morton (1992) describe the forward rate by

$$f(t,T) - f(0,T) = \int_0^t \alpha(v,T,\omega)dv + \sum_{i=1}^{n} \int_0^t \sigma_i(v,T,\omega)dW_i(v), \tag{1.1}$$

with restrictions that the drift term be measurable, adapted and integrable $\int_0^T |\alpha(v,T,\omega)|dv < \infty$. Similarly, volatilities σ_i must be jointly measurable, adapted and $\int_0^T \sigma_i^2(t,T,\omega)dt < \infty$ for $1 \leqslant i \leqslant n$. Thus, n independent Brownian motions, W_i, determine the stochastic fluctuations of the entire forward rate curve from fixed initial curve $f(0,T)$ (Heath, Jarrow and Morton, 1992). This family of forward rates has a forward rate dependent on n different volatility functions $\sigma_i(v,T,\omega)$. Trolle and Schwartz (2008) describes each $\sigma_i(v,T,\omega)$ as the sensitivity of change of the forward rate to each corresponding Brownian motion, $W_i(t)$. This can be simplified to $df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t)$, where W is an n-dimensional vector $W(t) = (W_1(t),...,W_n(t))$ and $\sigma(t,T) = (\sigma_1(t,T),...,\sigma_n(t,T)$. The framework allows general stochastic terms $\alpha(t,T)$ and $\sigma(t,T)$ which can be chosen to replicate forward rates but will not necessarily be tractable. Both the Cox Ingersoll Ross and Vasiček models fall into the HJM framework, both with one volatility parameter ($n = 1$). Traditional models such as Vasicek (1977) and Cox, Ingersoll Jr and Ross (1985) can be cases of the HJM framework (see Chiarella and Kwon (2001) for proof).

Features of the Trolle-Schwartz Model

The Trolle-Schwartz (TS) instantaneous forward rate model has many interesting characteristics and stylised facts. It can be viewed as a stochastic volatility extension of the HJM model. The instantaneous forward rate model includes stochastic volatility state variables in its diffusion term and is known as a stochastic volatility model (SVM). The stochastic volatility terms in the instantaneous forward rate model are just the square-root of a stochastic variance model. Notably, each stochastic variance variable follows its own stochastic differential equation with several parameters. The first stochastic volatility interest rate model was the Heston (1993) model where the variance of bond prices followed a mean-reverting, square-root process. However, Heston (1993) primarily used this model for equity derivatives. The complete state vector for the Trolle and Schwartz (2008) model is comprised of the stochastic volatility state vectors and the term structure state variables. The idea of including a stochastic volatility term arose when Collin-Dufresne and Goldstein (2002) noted that interest rate derivatives could not be hedged by bonds alone. This led to the conclusion that the bond market is not complete and there is at least one state variance variable that drives innovations in interest rate derivatives (Collin-Dufresne and Goldstein, 2002). This was termed ”Unspanned Stochastic Volatility” (USV) and led to the creation of unspanned stochastic volatility factors in model dynamics (Collin-Dufresne and Goldstein, 2002). Li and Zhao (2006) also found similar complications in using only bonds to hedge volatility dependent cap straddles. These unspanned factors drive the interest rate volatility and therefore interest rate derivative prices, but they do not affect the term structure.

There exists much literature by other authors supporting the inclusion of unspanned volatility factors. Trolle and Schwartz (2008) state that modelling variance as a stochastic process increases performance in interest rate derivative pricing but leaves the term structure unaffected. Specifically, Trolle and Schwartz (2008) state that using three stochastic volatility factors in their model fits both interest rates and interest rate derivative prices best. A study conducted by Heidari and Wu (2003) investigated the use of a three-factor model to explain the variation of implied volatilities in the swaption market. The study concluded that these non-stochastic volatility factors only explained 59.48% of the variation in implied volatilities (Heidari and Wu, 2003). However, including three independent stochastic volatility factors in the original model explained 97.62% of the variation of the implied volatility surface (Heidari and Wu, 2003).

The TS model incorporates many of the HJM features such as fitting of the initial yield curve. However, fitting the initial yield curve provides a high-dimensional framework and can potentially be impractical. A method used by Cheyette (1992) allowed TS to reduce the dimensionality by introducing a Markovian structure. Chiarella and Kwon (2000) presented a complete stochastic model within the HJM framework, through the Hobson and Rogers (1998) technique, which showed how the stochastic dynamics can be reduced to a Markovian form. This method allows bond prices to be expressed in terms of the underlying state variables which substantially reduces the computational time needed to price interest rate derivatives (Chiarella and Kwon, 2000). A similar technique was used by TS to reduce bond prices to a similar form. The TS model uniquely defines its instantaneous forward rate volatilities as deterministic functions of time and separately uses stochastic variances as additional state variables.

Trolle and Schwartz (2008) used both swaption price and cap skew data to calibrate parameters in their state variance variables. They compared their results to Han (2007) who used only swaption prices and Jarrow, Li and Zhao (2007) who used only cap price skew data instead. It is important to note that both of these models do not incorporate correlations between interest rates and their volatilities. However, if the Han (2007) and Jarrow, Li and Zhao (2007) models incorporated non-zero correlation coefficients in the stochastic volatility term, then these competing market models would be rendered intractable. This intractability is due to the correlations interfering with the volatility dynamics under the forward measure (Trolle and Schwartz, 2008). In the TS model, the forward rate is extended to include correlations with each volatility term with coefficient ρ_i (Trolle and Schwartz, 2008). Jarrow, Li and Zhao (2007) observed a strong negative correlation between interest rates and stochastic volatility terms in their model despite setting all coefficients to zero. Andersen and Lund (1997) asserted that changes in interest rate volatility are correlated with changes in the interest rate. Casassus, Collin-Dufresne and Goldstein (2005) also stated that having non-zero correlation coefficients was necessary when fitting cap skew data. Setting $\rho_i = \pm 1$ eliminates the unspanned factor and reduces the total number of unspanned factors in the model. Fixing $\rho_i = \pm 1$ is undesirable as Trolle and Schwartz (2008) observed that multiple unspanned factors are needed to fully capture the dynamics of interest rate derivatives. Setting $\rho_i = 0$, for all $i = 1, ..., n$ eradicates all correlations and setting $n = 1$ resembles Heston (1993) dynamics. The Trolle and Schwartz (2008) model is associated with the stochastic volatility models of Han (2007) and Jarrow, Li and Zhao (2007). Under these models, the forward LIBOR rates are log-normally distributed and the swap rates are approximately log-normally distributed. However, Trolle and Schwartz (2008) imply that LIBOR and swap rates are approximately normally distributed under the forward measure which allows for the possibility of negative rates. The Gaussian distribution feature can be argued to be more desirable and reflective of real life as countries such as Japan and Germany had negative rates post the 2008 financial crisis. A further desirable characteristic of the Trolle and Schwartz (2008) model is its ability to handle radical changes to its term structure. Specifically, the Markovian feature of its forward rate volatilities allows for a variety of hump-shaped changes and is essential when fitting implied cap price skews (Trolle and Schwartz, 2008). This structure permits the unconditional volatility term structure to exhibit a hump as portrayed by Dai and Singleton (2002) and Cheyette (1995). By specifying certain instantaneous forward rate volatility functions, the stochastic volatility version of the continuous-time Ho and Lee (1986) model can be recovered by the Trolle and Schwartz (2008) model. Similarly, the model can recover the stochastic volatility version of the Hull and White (1990).

Fourier Inversion Pricing

Fourier pricing is an efficient method to price options. Option payoffs can be written as a function of probabilities of realisations of the underlying asset. Through manipulation, the payoff function can be expressed as an integral involving a characteristic function. The option price is taken as the discounted expected payoff at maturity under risk-neutral measure. Therefore, the option price can be represented as an integral of the characteristic function through Fourier inversion (see Duffie, Pan and Singleton (2000) for further discussion). Trolle and Schwartz (2008) require numerical methods such as ordinary differential equation (ODE) solvers and a Fourier inversion to price interest rate options. The model prices put options on zero-coupon bonds and requires the solution of two ODEs to compute the characteristic function. The characteristic function is used in conjunction with a Gauss-Legendre quadrature to estimate an integral. The ODEs must be solved at every point where the characteristic function is evaluated by the Gauss-Legendre quadarture to estimate the integral. Put options on zero-coupon bonds can then be manipulated to price other claims such as caplets, caps and swaptions.

Ordinary Differential Equations

ODEs have many uses in mathematics and physics and the solution will depend on the rate of change or growth (represented by its derivatives). Trolle and Schwartz (2008) use a standard fourth-order Runge-Kutta algorithm. Runge-Kutta ODE solvers require the relevant ODE, a time period and set of initial conditions to find a solution. The solver begins at an initial condition which is a starting value associated with a start time of the interval. The time interval is discretised into steps and the solution is evolved discretely over each time step. The solution at each step has an error check, using Taylor expansion, that ensures the solution falls within an error tolerance level. The time step sizes are reduced if the error of the solution falls outside the error tolerance level. The solution is propogated until the end of the time period which is the solution at the end of the time interval (For further details see Shampine and Reichelt (1997)). These ODE solvers are relatively efficient and accurate providing fast and reliable prices. However, closed-form solutions to ODEs do exist but are difficult to find. There are general methods to solve certain classes of ODEs but not all. The aim of this dissertation will be to compare closed-form solutions of the ODEs used for pricing in the Trolle and Schwartz (2008) model to a reliable and accurate ODE solver.

Structure of the Dissertation

This dissertation will follow a chronological structure. Chapter 2 will introduce the Trolle and Schwartz (2008) model, pricing formulae using Fourier inversion and the ODEs in question. Then, Chapter 3 postulates two lemmas for two closed-form solutions for the ODEs. It also discusses some numerical implementation difficulties, defines parameters and presents an ODE solver as a comparsion to the closed-form ODE solutions. Chapter 4 will provide the numerical results and comparison of the closed-form solutions and the ODE solver. Finally, Chapter 5 will conclude.

Chapter 2

Model Description

The following descriptions closely follow the original presentation by Trolle and Schwartz (2008). These include the specified dynamics, semi-analytical pricing formulae and ODE equations in mathematical form.

2.1 Risk-Neutral Dynamics

Consider the time t instantaneous forward rate, $f(t,T)$, which is the rate at which an entity may borrow or lend at a future date T, with $t \leqslant T$, for an infinitesimally-small amount of time. Furthermore, consider the stochastic variance variables $\upsilon_i(t)$, where $i = 1, ..., n$, which effect the diffusion of $f(t,T)$. The model dynamics specified by Trolle and Schwartz (2008) under risk-neutral measure $\mathbb{Q}$ are

$$df(t,T) = \mu_f(t,T)dt + \sum_{i=1}^{n} \sigma_{f,i}(t,T)\sqrt{\upsilon_i(t)}dW_i(t) \tag{2.1}$$

$$d\upsilon_i(t) = \kappa_i(\theta_i - \upsilon_i(t))dt + \sigma_i\sqrt{\upsilon_i(t)}\left(\rho_i dW_i(t) + \sqrt{1-\rho_i^2}dZ_i(t)\right), \tag{2.2}$$

for $i = 1, ..., n$ and $W_i(t), Z_i(t)$ are independent Brownian motion processes under risk-neutral measure. The TS model jointly specifies the dynamics of forward rates and the stochastic variance variables $\upsilon_i(t)$, for $i = 1, ..., n$, which drive the forward rate volatility. Thus, it has the original n spanned factors, $W_i(t)$, driving its term structure and an additional n unspanned stochastic volatility factors, $Z_i(t)$. The dynamics of each $\upsilon_i(t)$ follows a square-root, mean-reverting process where θ_i is the mean-reversion level and κ_i is the strength of mean-reversion.

Heath, Jarrow and Morton (1992) conveyed that the absence of arbitrage implies that the drift term in the TS model is given by

$$\mu_f(t,T) = \sum_{i=1}^{n} \upsilon_i(t)\sigma_{f,i}(t,T)\int_t^T \sigma_{f,i}(t,u)du. \tag{2.3}$$

Therefore, the dynamics of $f(t,T)$ under risk-neutral measure are completely determined by the initial forward curve, the forward rate volatility functions, $\sigma_{f,i}(t,T)$, and the volatility state variables $\upsilon_i(t)$ (Trolle and Schwartz, 2008). Furthermore, Trolle and Schwartz (2008) require the dynamics to be Markovian and specify the forward rate volatilities to be

$$\sigma_{f,i}(t,T) = (\alpha_{0,i} + \alpha_{1,i}(T-t))e^{-\gamma_i(T-t)}. \tag{2.4}$$

This structure affirms that the process is finite-dimensional and Markov whilst ensuring the forward rate volatilities are time-homogeneous. This volatility structure can resemble that of Cheyette (1992) except that the Trolle and Schwartz (2008) structure includes additional stochastic volatility. This formulation allows for a hump-shaped forward rate volatility structure $\sigma_{f,i}(t,T)$. The full instantaneous forward rate is specified by (2.5) in Appendix A. As previously mentioned, many traditional models lie within these model dynamics. The stochastic volatility version of the Hull and White (1990) model (as analyzed by Casassus, Collin-Dufresne and Goldstein (2005)) is recovered when $n = 1$, $\alpha_{1,1} = 0$ and define $\alpha_1 = -\alpha_{0,1}$. Similarly, setting $\gamma_1 = 0$ for $n = 1$ recovers the stochastic volatility version of the continuous-time Ho and Lee (1986) model. Given the dynamics of $f(t,T)$, its solution is given by

$$f(t,T) = f(0,T) + \sum_{i=1}^{n} \mathcal{B}_{x_i}(T-t)x_i(t) + \sum_{i=1}^{n}\sum_{j=1}^{6} \mathcal{B}_{\phi_{j,i}}(T-t)\phi_{j,i}(t), \tag{2.5}$$

where

$$\mathcal{B}_{x_i}(\tau) = (\alpha_{0,i} + \alpha_{1,i}(T-t))e^{-\gamma_i\tau} \tag{2.6}$$

$$\mathcal{B}_{\phi_{1,i}}(\tau) = \alpha_{1,i}e^{-\gamma_i\tau} \tag{2.7}$$

$$\mathcal{B}_{\phi_{2,i}}(\tau) = \frac{\alpha_{1,i}}{\gamma_i}(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}})(\alpha_{0,i} + \alpha_{1,i}(\tau))e^{-\gamma_i\tau} \tag{2.8}$$

$$\mathcal{B}_{\phi_{3,i}}(\tau) = -\left(\frac{\alpha_{0,i}\alpha_{1,i}}{\gamma_i}(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}}) + \frac{\alpha_{1,i}}{\gamma_i}(\frac{\alpha_{1,i}}{\gamma_i} + 2\alpha_{0,i})\tau + \frac{\alpha_{1,i}^2}{\gamma_i}\tau^2\right)e^{-2\gamma_i\tau} \tag{2.9}$$

$$\mathcal{B}_{\phi_{4,i}}(\tau) = \frac{\alpha_{1,i}^2}{\gamma_i}(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}})e^{-\gamma_i\tau} \tag{2.10}$$

$$\mathcal{B}_{\phi_{5,i}}(\tau) = -\frac{\alpha_{1,i}}{\gamma_i}(\frac{\alpha_{1,i}}{\gamma_i} + 2\alpha_{0,i} + 2\alpha_{1,i}\tau)e^{-2\gamma_i\tau} \tag{2.11}$$

$$\mathcal{B}_{\phi_{6,i}}(\tau) = -\frac{\alpha_{1,i}^2}{\gamma_i}e^{-2\gamma_i\tau}, \tag{2.12}$$

and state variables evolve with

$$dx_i(t) = -\gamma_i x_i(t)dt + \sqrt{\upsilon_i(t)}dW_i(t) \tag{2.13}$$
$$d\phi_{1,i}(t) = (x_i(t) - \gamma_i \phi_{1,i}(t))dt \tag{2.14}$$
$$d\phi_{2,i}(t) = (\upsilon_i(t) - \gamma_i \phi_{2,i}(t))dt \tag{2.15}$$
$$d\phi_{3,i}(t) = (\upsilon_i(t) - 2\gamma_i \phi_{3,i}(t))dt \tag{2.16}$$
$$d\phi_{4,i}(t) = (\phi_{2,i}(t) - \gamma_i \phi_{4,i}(t))dt \tag{2.17}$$
$$d\phi_{5,i}(t) = (\phi_{3,i}(t) - 2\gamma_i \phi_{5,i}(t))dt \tag{2.18}$$
$$d\phi_{6,i}(t) = (2\phi_{5,i}(t) - 2\gamma_i \phi_{6,i}(t))dt, \tag{2.19}$$

subject to $x_i(0) = \phi_{1,i}(0) = ... = \phi_{6,i}(0) = 0$. The forward rates do not depend directly on the volatility state variables but instead the dynamics are given in terms of $n \times 8$ state variables that jointly follow an affine diffusion process (Trolle and Schwartz, 2008). A notable absence of diffusion terms in $\phi_{i,1}, \phi_{i,2}, ..., \phi_{i,6}$ which are ancillary, locally deterministic state variables that help determine the path of $x_i(t)$ and $\upsilon_i(t)$. By restricting the state space of these variables, the model becomes Markovian and falls under the analytically tractable affine class of models as proposed by Duffie and Kan (1996). The model is time-inhomogeneous as the dynamics of the forward rate depend on the initial term structure of forward rates (Trolle and Schwartz, 2008).

2.2 Semi-Analytical Pricing Formulae

Trolle and Schwartz (2008) provided a semi-analytical solution of a time t price of a zero-coupon bond maturing at T,

$$\begin{aligned} P(t,T) &= e^{-\int_t^T f(t,u)du} \\ &= \tfrac{P(0,T)}{P(0,t)} \exp\left(\sum_{i=1}^{n} \beta_{x_i}(T-t)x_i(t) + \sum_{i=1}^{n}\sum_{j=1}^{6} \beta_{\phi_{j,i}}(T-t)\phi_{j,i}(t) \right), \end{aligned} \tag{2.20}$$

where

$$\beta_{x_i}(\tau) = \frac{\alpha_{1,i}}{\gamma_i}((\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}})(e^{-\gamma_i\tau} - 1) + \tau e^{-\gamma_i\tau}) \tag{2.21}$$

$$\beta_{\phi_{1,i}}(\tau) = \frac{\alpha_{1,i}}{\gamma_i}(e^{-\gamma_i\tau} - 1) \tag{2.22}$$

$$\beta_{\phi_{2,i}}(\tau) = (\frac{\alpha_{1,i}}{\gamma_i})^2(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}})((\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}})(e^{-\gamma_i\tau} - 1) + \tau e^{-\gamma_i\tau}) \tag{2.23}$$

$$\beta_{\phi_{3,i}}(\tau) = -\frac{\alpha_{1,i}}{\gamma_i^2}\Big[(\frac{\alpha_{1,i}}{2\gamma_i^2} + \frac{\alpha_{0,i}}{\gamma_i} + \frac{\alpha_{0,i}^2}{2\alpha_{1,i}})(e^{-2\gamma_i\tau} - 1) + (\frac{\alpha_{1,i}}{\gamma_i} + \alpha_{0,i})\tau e^{-2\gamma_i\tau} + \frac{\alpha_{1,i}}{2}\tau^2 e^{-2\gamma_i\tau}\Big] \tag{2.24}$$

$$\beta_{\phi_{4,i}}(\tau) = (\frac{\alpha_{1,i}}{\gamma_i})^2(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}})(e^{-\gamma_i\tau} - 1) \tag{2.25}$$

$$\beta_{\phi_{5,i}}(\tau) = -\frac{\alpha_{1,i}}{\gamma_i^2}\left((\frac{\alpha_{1,i}}{\gamma_i} + \alpha_{0,i})(e^{-2\gamma_i\tau} - 1) + \alpha_{1,i}\tau e^{-2\gamma_i\tau}\right) \tag{2.26}$$

$$\beta_{\phi_{6,i}}(\tau) = -\frac{1}{2}(\frac{\alpha_{1,i}}{\gamma_i})^2(e^{-2\gamma_i\tau} - 1). \tag{2.27}$$

Furthermore, the dynamics of $P(t,T)$ are given by,

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt + \sum_{i=1}^{n}\beta_{x_i}(T-t)\sqrt{v_i(t)}dW_i(t). \tag{2.28}$$

Trolle and Schwartz (2008) also developed a method to price bond options by taking the inverse Fourier transform method used by Duffie, Pan and Singleton (2000) and Collin-Dufresne and Goldstein (2003). They extended the work of Duffie, Pan and Singleton (2000) to HJM models using the transform,

$$\Psi(u,t,T_0,T_1) = E_t^Q[e^{-\int_t^{T_0} r_s ds}e^{u\log(P(T_0,T_1))}], \tag{2.29}$$

which holds $t < T_0 < T_1$. This transform has the solution,

$$\Psi(u,t,T_0,T_1) = \exp\left(M(T_0-t) + \sum_{i=1}^{n}N_i(T_0-t)v_i(t) + u\log(P(t,T_1)) + (1-u)\log(P(t,T_0))\right), \tag{2.30}$$

where $M(\tau)$ and $N(\tau)$ solve the following system of ODEs,

$$\frac{dM(\tau)}{d\tau} = \sum_{i=1}^{n}N_i(\tau)\kappa_i\theta_i \tag{2.31}$$

$$\begin{aligned}\frac{dN_i(\tau)}{d\tau} &= N_i(\tau)(-\kappa_i + \sigma_i\rho_i(u\beta_{x_i}(T_1-T_0+\tau) + (1-u)\beta_{x_i}(\tau))) \\ &+ \frac{1}{2}N_i(\tau)^2\sigma_i^2 + \frac{1}{2}(u^2-u)\beta_{x_i}(T_1-T_0+\tau)^2 \\ &+ \frac{1}{2}((1-u)^2 - (1-u))\beta_{x_i}(\tau)^2 \\ &+ u(1-u)\beta_{x_i}(T_1-T_0+\tau)\beta_{x_i}(\tau),\end{aligned} \tag{2.32}$$

subject to boundary conditions $M(0) = 0$ and $N(0) = 0$. Bond options can now be priced using the Fourier inversion theorem. Trolle and Schwartz (2008) used this Fourier inversion approach to price a put option on a zero-coupon bond with strike K. The price of a time t put option on a zero-coupon bond with option expiry at T_0 and maturity T_1 is,

$$\mathcal{P}(t, T_0, T_1, K) = KG_{0,1}(\log(K)) - G_{1,1}(\log(K)), \tag{2.33}$$

where

$$G_{a,b}(y) = \frac{\Psi(a,t,T_0,T_1)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\text{Im}[\Psi(a+iub,t,T_0,T_1)e^{-iuy}]}{u} du, \tag{2.34}$$

where $i = \sqrt{-1}$. A swaption can be viewed as a European put option on a coupon bond, but no analytical solution to price options on coupon bearing bonds exists yet (Trolle and Schwartz, 2008). Trolle and Schwartz (2008) use the fast and accurate stochastic duration methods developed by Munk (1999) and Wei (1997). These methods are discussed further in section 2.3.3. Trolle and Schwartz (2008) make use of a Gauss-Legendre quadrature to estimate the integral in (2.34). They use 20 points over $[0, 1000]$ and 20 more over $[1000, 8000]$. They state that this truncation and use of only 40 points is suitably accurate.

2.3 Interest Rate Derivative Instruments

Trolle and Schwartz (2008) have developed an instantaneous forward rate model that allows to price zero-coupon bonds (ZCBs) and zero-coupon bond options. These in turn can be used to find interest rates and vanilla interest rate derivatives described below. All pricing of options occurs on a nominal of one unit of currency.

2.3.1 Simple LIBOR and Swap Rates

The following rates and formulae are taken from Björk (2009). Consider the set of discrete times $t \leqslant S \leqslant T$. The simple spot rate or LIBOR spot rate at time t is

$$L(t, T) = \frac{1 - P(t, T)}{(T - t)P(t, T)}. \tag{2.35}$$

Similarly, the forward simple rate or forward LIBOR rate at time t over future period $[S, T]$ is

$$L(t; S, T) = \frac{P(t, S) - P(t, T)}{(T - t)P(t, T)}. \tag{2.36}$$

A swap is an instrument where a party exchanges a fixed payment stream at a fixed interest rate, known as the swap rate, for a payment stream at a referenced floating rate, usually LIBOR (Björk, 2009). The floating rate resets at a fixed number of equally spaced dates and the accrued net difference between the fixed and floating rates is paid at the end of the period. The time t value of a swap rate for period $[t, T_N]$, with N equally spaced time legs δ, is

$$S(t, T_N) = \frac{1 - P(t, T_N)}{\delta \sum_{j=1}^{n} P(t, T_j)}. \tag{2.37}$$

Similarly, the time t forward swap rate for the future period $[T_M, T_N]$ where $t < T_M < T_N$ is

$$S(t; T_M, T_N) = \frac{P(t, T_M) - P(t, T_N)}{\delta \sum_{j=M+1}^{n} P(t, T_j)}. \tag{2.38}$$

2.3.2 Caplets and Caps

A caplet is a call option on an underlying interest rate with strike K, known as the cap rate. The time interval $[t, T]$ can be partitioned into $t \leqslant T_1 = T_M < T_2 < ... < T_N = T$ with reset dates $T_1, ..., T_{N-1}$ and corresponding payment dates $T_2, ..., T_N$. A caplet payoff at each payment date, T_j, for tenor δ and strike K can be represented by

$$\text{Caplet}(T_j, K) = \delta(L(T_j - \delta, T_j) - K)^+. \tag{2.39}$$

This payoff can be manipulated into a scaled put option on a ZCB. Therefore, the discounted time t caplet price can be rewritten as

$$\text{Caplet}(T_j, K) = (1 + \delta K)\mathcal{P}(t, T_j - \delta, T_j, \tfrac{1}{1+\delta K}). \tag{2.40}$$

A cap is just the sum of the discounted caplet prices. Thus, the value of a time t cap maturing at T_N with strike K can be written as

$$\text{Cap}(t, T_N, K) = \sum_{j=2}^{N} \text{Caplet}(t, T_j, K). \tag{2.41}$$

Note that these equations assume that the first caplet from t to T_1 is not included. In order to price caplets and caps only the ZCB option formula (2.33) is needed. A cap is priced at-the-money-forward (ATMF) when the strike price is set to $\tilde{K} = S(t, T_1, T_N)$. This is the fair forward swap rate for initial payment T_1 and maturity T_N which sets the value of the cap to zero.

2.3.3 Swaptions and the Stochastic Duration Approach

A swap is a contract where two parties exchange a fixed interest rate, K, for a reference floating interest rate over a future period of time. Specifically a payer swap is one where the holder agrees to pay the fixed leg rate and receive the floating rate. The value of a payer swap at any time t over period T_M to T_N with payment dates $T_{M+1} < T_{M+2} < ... < T_N$ is given by

$$V(t,T_M,T_N) = \sum_{j=M+1}^{N} P(t,T_j)(L(T_j-\delta,T_j)-K)\delta$$
$$= P(t,T_M) - P(t,T_N) - K\delta \sum_{j=M+1}^{N} P(t,T_j). \quad (2.42)$$

A swaption is an option on an interest rate swap. A payer swaption is the right but not the obligation to enter a payer swap. Therefore, we can write the payoff of a payer swaption at option expiry T_M as

$$V(T_M,T_M,T_N)^+ = \left(1 - P(T_M,T_N) - K\delta \sum_{j=M+1}^{N} P(T_M,T_j)\right)^+. \quad (2.43)$$

This payoff is the same as a put option on a coupon bearing bond with coupon rate K and unit strike. For ease of notation, the value of a coupon bearing bond at time $t < T_M < T_N$ is

$$\mathrm{P}^{\mathrm{CB}}(t) = \sum_{j=M+1}^{N} P(t,T_j)Y(T_j), \quad (2.44)$$

where $Y(T_i) = K\delta$ for coupon payments at $i = M+1,...,N-1$ and $Y(T_N) = 1+K\delta$ at maturity. Munk (1999) defines the stochastic duration of a coupon bearing bond as the maturity of a zero-coupon bond which has identical relative volatility as the coupon bearing bond. The stochastic duration, $D(t)$, of $\mathrm{P}^{\mathrm{CB}}(t)$ must be found numerically as a solution to

$$\sum_{i=1}^{n} \upsilon_i(t)\beta_{x_i}(D(t))^2 = \sum_{i=1}^{n} \upsilon_i(t)\left(\sum_{j=M+1}^{N} w_j\beta_{x_j}(T_j-t)\right)^2, \quad (2.45)$$

where n is the number of stochastic volatility terms and $w_i = \frac{P(t,T_i)Y(T_i)}{\sum_{i=M+1}^{N} P(t,T_i)Y(T_i)}$. $D(t)$ is well-defined and unique when $\beta_{x_i}(\tau)$ from Section 2.2 is decreasing which is the case according to the parameter estimates of Trolle and Schwartz (2008). As suggested by Munk (1999) and Wei (1997), options on coupon bearing bonds can now be estimated by an option on a ZCB with the same stochastic duration as the coupon bearing bond. Since a swaption can be written as a put option on a coupon bearing bond, we can price swaptions using an option on a ZCB. Let the time t price of a swaption with fixed rate K be $\mathrm{Swpt}(t,T_M,T_N,K)$ with the same tenor structure as in the swap above. Then, the swaption price as stated by Munk (1999) and Wei (1997) is approximately

$$\mathrm{Swpt}(t,T_M,T_N,K) = \xi\mathcal{P}(t,T_M,t+D(t),\xi^{-1}), \quad (2.46)$$

where $\xi = \frac{\mathrm{P}^{\mathrm{CB}}(t)}{P(t,t+D(t))}$.

There exists a volatility approximation formula developed by Trolle and Schwartz (2008) that can be used to find volatility surfaces and prices swaptions in a log-normal world. Both Trolle and Schwartz (2008) and Schumann (2016) state that this method is less accurate than the stochastic duration approach. The volatility approximation approach also does not involve the ODEs in (2.31 and 2.32). Therefore, the volatility approximation method would be left out in order to focus on methods using the ODEs in the stochastic duration approach.

Chapter 3

Closed-Form Solutions and Parameters

This chapter begins by introducing the confluent hypergeometric equation along with the Kummer and Tricomi functions. Closed-form solutions to the two ODEs are defined using the Kummer and Tricomi functions. The chapter then deals with instability in the solution and defines a second closed-form solution as a result. A short discussion about how Trolle and Schwartz (2008) calibrated parameters follows the second solution. Finally, the chapter introduces the basic model which uses the ODE45 solver in MATLAB to solve the ODEs. This basic model serves as a baseline comparison for pricing accuracy and efficiency against the closed-form ODE solutions.

3.1 Features of Confluent Hypergeometric Functions

Finding a closed-form solution to the fist ODE in (2.32) requires a rather sophisticated solution and deals with some time inhomogeneous coefficients, complex numbers and non-linearity. The solution involves Kummer and Tricomi functions which are rarely used and can misbehave with ordinary computing methods. These functions are two linearly independent solutions to the confluent hypergeometric equation

$$zg''(z) + (b - z)g'(z) - ag(z) = 0. \tag{3.1}$$

The Kummer function is given by

$$M(a, b, z) = 1 + \frac{a}{b}z + \frac{a(a+1)}{b(b+1)}\frac{z^2}{2!} + \ldots = \sum_{n=0}^{\infty} \frac{(a)_n z^n}{(b)_n}, \tag{3.2}$$

while the Tricomi function is defined as

$$U(a, b, z) = \frac{\Gamma(1-b)}{\Gamma(a+1-b)}M(a, b, z) + \frac{\Gamma(b-1)}{\Gamma(a)}z^{1-b}M(a+1-b, 2-b, z). \tag{3.3}$$

The gamma function, $\Gamma(x)$, must take complex values, and $(x)_n = 1$ for $n = 0$ and $(x)_n = \prod_{i=0}^{n-1}(x+i)$ for $n \geqslant 1$. A few notable simplifications are

$$M(0,b,z) = U(0,b,z) = 1, \quad M(b,b,z) = e^z, \quad U(a,a+1,z) = z^{-a}.$$

Different sets of parameters can relate back to more familiar functions that are encountered in mathematics and physics such as Bessel, exponential, Laguerre, hyperbolic and the trigonometric functions. For further reading, see Abramowitz and Stegun (1964) and Pearson, Olver and Porter (2017). A closed-form solution for the ODEs (2.31) and (2.32) can now be defined.

Lemma 3.1. *By setting $\alpha_{1,i} = 0$ and $\alpha_i = -\alpha_{0,i}$, ODE equation (2.32) has the closed-form solution*

$$N_i(\tau) = \frac{2\gamma_i}{\sigma_i^2}\left[\beta + \mu\frac{e^{-\gamma_i\tau}}{w} + \frac{e^{-\gamma_i\tau}}{w}\frac{g'\left(\frac{e^{-\gamma_i\tau}}{w}\right)}{g\left(\frac{e^{-\gamma_i\tau}}{w}\right)}\right]. \tag{3.4}$$

The function $g(z)$ is a linear combination of Tricomi's and Kummer's confluent hypergeometric functions:

$$g(z) = k_1 M(a,b,z) + k_2 U(a,b,z),$$
$$g'(z) = k_1 \tfrac{a}{b} M(a+1,b+1,z) - k_2 a U(a+1,b+1,z).$$

The coefficients are

$$\beta = \tfrac{-c_0 \pm \sqrt{c_0^2 - 4d_0}}{2\gamma_i} \qquad c_0 = -\kappa_i + \sigma_i\rho_i\tfrac{\alpha_i}{\gamma_i}$$
$$w = \tfrac{\pm\gamma_i}{\sqrt{c_1^2 - 4d_2}} \qquad c_1 = -\sigma_i\rho_i\tfrac{\alpha_i}{\gamma_i}\left(ue^{-\gamma_i(T_1-T_0)} + (1-u)\right)$$
$$d_0 = 0 \qquad \mu = -\tfrac{1}{2}\left(1 + \tfrac{c_1 w}{\gamma_i}\right)$$
$$d_1 = 0 \qquad a = -\mu\left(\tfrac{c_0}{\gamma_i} + 1 + 2\beta\right) - \beta c_1\tfrac{w}{\gamma_i} - d_1\tfrac{w}{\gamma_i^2}$$
$$b = 2\beta + 1 + \tfrac{c_0}{\gamma_i} \qquad d_2 = \tfrac{\sigma_i^2\alpha_i^2}{4\gamma_i^2}(u^2 - u)\left[e^{-2\gamma_i(T_1-T_0)} - 2e^{-\gamma_i(T_1-T_0)} + 1\right].$$

Particularly, if $N(0) = 0$, such as in this case, then

$$k_1 = \frac{-\beta w - \mu + a\frac{U(a+1,b+1,\frac{1}{w})}{U(a,b,\frac{1}{w})}}{\frac{a}{b}M(a+1,b+1,\frac{1}{w}) + aM(a,b,\frac{1}{w})\frac{U(a+1,b+1,\frac{1}{w})}{U(a,b,\frac{1}{w})}},$$

$$k_2 = \frac{1 - k_1 M(a,b,\frac{1}{w})}{U(a,b,\frac{1}{w})}.$$

□

Proof - see (A.1).

Lemma 3.2. *If $\alpha_{1,i} = 0$ and $\alpha_i = -\alpha_{0,i}$, ODE equation (2.31) has the following solution*

$$M(\tau) = \sum_{i=1}^{n} \frac{2\kappa_i\theta_i}{\sigma_i^2}\Big(\beta\gamma_i\tau - \mu z - \log(g(z))\Big) + k_3, \tag{3.5}$$

with the same function $g(z)$ and coefficients as in (3.1). Particularly, if $N(0) = 0$, such as in this case, then

$$k_3 = \sum_{i=1}^{n} \frac{2\kappa_i\theta_i\mu}{\sigma_i^2 w}.$$

□

Proof - see (A.2).

Using Lemmas 3.1 and 3.2 for $n = 1$ produces a closed-form solution of the characteristic function to the stochastic volatility version of the Hull and White (1990) model. If n is increased, then a closed-form solution of the characteristic function is found for the Hull and White (1990) model with n stochastic volatility terms. The numerical estimate of the integral using Gauss-Legendre in (2.34) is still required. However, it does remove one numerical estimate for each of the ODEs. Although this parameterization is a certain case of the Trolle and Schwartz (2008) model, it is still significant as the Hull and White (1990) model is a popular model in practice. Setting $\alpha_{1,i} = 0$ and $\alpha_i = -\alpha_{0,i}$ is required in order to use the confluent hypergeometric functions in the proofs.

3.2 Stability Analysis

Following the results from Lemmas 3.1 and 3.2, the coefficients can be simplified and analysed further. Setting $d_0 = d_1 = 0$, presents the following coefficients:

$$\beta = \frac{-c_0 \pm \sqrt{c_0^2 - 4d_0}}{2\gamma_i} = \frac{-c_0 \pm \sqrt{c_0^2}}{2\gamma_i}$$
$$a = -\mu\big(\frac{c_0}{\gamma_i} + 1 + 2\beta\big) - \beta c_1 \frac{w}{\gamma_i} - d_1\frac{w}{\gamma_i^2} = -\mu\big(\frac{c_0}{\gamma_i} + 1 + 2\beta\big) - \beta c_1\frac{w}{\gamma_i}.$$

This means that there are two potential solutions for β denoted β_- or β_+ where

$$\beta = \begin{cases} 0, & \text{for } \beta_+ \\ \dfrac{-c_0}{\gamma_i}, & \text{for } \beta_-. \end{cases}$$

Similarly, there are two solutions for w and denoted w_- and w_+ where

$$w = \begin{cases} \dfrac{\gamma_i}{\sqrt{c_1^2 - 4d_2}}, & \text{for } w_+ \\ \dfrac{-\gamma_i}{\sqrt{c_1^2 - 4d_2}}, & \text{for } w_-. \end{cases}$$

This leads to four combinations of coefficients β and w. All combinations will be implemented in pricing and should produce the same results.

However, a numerical implementation issue occurs for the $\beta_+ w_-$ pair. Pricing in (2.33) requires $u = 0$ and $u = 1$ and both result in $d_2 = 0$. Basic cancellation means $w_- = \frac{-\gamma_i}{c_1}$ in these two cases. This causes μ to become zero because

$$\mu = -0.5(1 + \tfrac{c_1}{\gamma_i} w_-) = -0.5(1 + \tfrac{c_1}{\gamma_i}(\tfrac{-\gamma_i}{c_1})) = -0.5(1 - 1) = 0.$$

Choosing $\beta_+ = 0$ and given $\mu = 0$ under $u = 0$ or $u = 1$ means that $a = 0$. This results in $k_1 = \frac{0}{0}$ in the previous formulation for the $\beta_+ w_-$ pair. This issue only occurs for this pair and requires further attention to implement numerically. To prevent this division by zero problem, both μ and β_+ can be set to zero and a can be cancelled in each term in the numerator and denominator

$$\begin{aligned} k_1 &= \frac{a\dfrac{U(a+1, b+1, \frac{1}{w})}{U(a, b, \frac{1}{w})}}{\frac{a}{b}M(a+1, b+1, \frac{1}{w}) + aM(a, b, \frac{1}{w})\frac{U(a+1,b+1,\frac{1}{w})}{U(a,b,\frac{1}{w})}} \\ &= \frac{\dfrac{U(a+1, b+1, \frac{1}{w})}{U(a, b, \frac{1}{w})}}{\frac{1}{b}M(a+1, b+1, \frac{1}{w}) + M(a, b, \frac{1}{w})\frac{U(a+1,b+1,\frac{1}{w})}{U(a,b,\frac{1}{w})}}. \end{aligned}$$

The functions $U(a, b, z)$ and $M(a, b, z)$ are well-defined here for the parameters and $b \neq 0$. This means that all four pairs can be implemented by avoiding the problem listed above. The β and w pairs will be abbreviated with the first sign being the sign for β and the second sign being the sign for w. For example, $++$ will refer to the $\beta_+ w_+$ pair, $-+$ for the $\beta_- w_+$ pair and similarly for $+-$ and $--$.

3.3 A Second Solution

Given that the numerical instability can occur in the $\beta_+ w_-$ pair first solution, this requires further investigation of the confluent hypergeometric equations. Most combinations of a and b values, real or complex, make the $M(a, b, z)$ and $U(a, b, z)$ solutions independent. However, if the solution can be unstable in certain cases then $z^{1-b}M(a + 1 - b, 2 - b, z)$ and $z^{1-b}U(a + 1 - b, 2 - b, z)$ can be added as a second solution instead of $M(a, b, z)$ and $U(a, b, z)$ (Abramowitz and Stegun, 1964). This new solution can be more stable than the first solution as the $M(a, b, z)$ and $U(a, b, z)$ solutions may not be independent. Abramowitz and Stegun (1964) actually state eight different solutions to the confluent hypergeometric equation but

Kummer proved several of them to be equal to the original formulation. These can be found in Appendix B.1 and B.2. The new solution will be defined as:

$$\begin{aligned} g(z) &= k_1 z^{1-b} M(a+1-b, 2-b, z) + k_2 z^{1-b} U(a+1-b, 2-b, z) \\ &= k_1 z^{1-b} M(a+1-b, 2-b, z) + k_2 U(a, b, z). \end{aligned} \tag{3.6}$$

The last step comes from using one of Kummer's transformations in Abramowitz and Stegun (1964) where $z^{1-b}U(a+1-b, 2-b, z) = U(a, b, z)$. The following derivative relationships are also important to calculate $g'(z)$:

$$\begin{aligned} \tfrac{d^n}{dz^n} M(a, b, z) &= \tfrac{(a)_n}{(b)_n} M(a+n, b+n, z) \\ \tfrac{d^n}{dz^n} U(a, b, z) &= (-1)^n (a)_n U(a+n, b+n, z). \end{aligned}$$

A new $g'(z)$ can be found by using the product rule in conjunction with the two rules above:

$$\begin{aligned} g'(z) = k_1 z^{-b} \Big[(1-b) M(a+1-b, 2-b, z) + z \tfrac{a+1-b}{2-b} M(a+2-b, 3-b, z) \Big] \\ -k_2 a U(a+1, b+1, z). \end{aligned} \tag{3.7}$$

Also, new coefficients need to be determined from the initial conditions. Once again, we follow Sitzia (2018) by setting $g(\frac{1}{w}) = 1$ and $g'(\frac{1}{w}) = -\beta w - \mu$. This results in:

$$k_1 = \frac{-\beta w - \mu + a \frac{U(a+1, b+1, \frac{1}{w})}{U(a, b, \frac{1}{w})}}{(\frac{1}{w})^{-b} \Big[(1-b) M(a+1-b, 2-b, \frac{1}{w}) + \frac{1}{w} \frac{a+1-b}{2-b} M(a+2-b, 3-b, \frac{1}{w}) + \frac{a}{w} M(a+1-b, 2-b, \frac{1}{w}) \frac{U(a+1, b+1, \frac{1}{w})}{U(a, b, \frac{1}{w})} \Big]}$$

$$k_2 = \frac{1 - k_1 (\frac{1}{w})^{1-b} M(a+1-b, 2-b, \frac{1}{w})}{U(a, b, \frac{1}{w})}.$$

These new solutions still involve the same parameters and coefficients as defined in Lemmas 3.1 and 3.2 but $N_i(\tau)$ and $M(\tau)$ change as $g(z)$, $g'(z)$, k_1 and k_2 differ.

3.4 Model Parameters and Data

Trolle and Schwartz (2008) calibrated their model to weekly observations of US LIBOR/swap rates and ATMF swaption and cap volatilities from 21 August 1998 until 26 January 2007. They also used weekly observations of log-normal cap skews from 4 January 2002 until 26 January 2007. These values are all closing midquotes on Fridays and were acquired from Bloomberg. Parameter values were found calibrating to swaptions and caps simultaneously and then swaptions and caps individually. The model parameters and estimation of latent state variables were found using a Kalman filter along with a maximum likelihood estimate (MLE) (Trolle and Schwartz, 2008). The filter forecasts a path for each latent state variable and the MLE produces the most likely value for each model parameter. Both Trolle and Schwartz (2008) and Schumann (2016) used parameters under $\mathbb{Q}$ and real-world measure $\mathbb{P}$ to price. $\mathbb{P}$ was introduced by both authors for the Kalman filter to find latent state variable values. The focus here is not on estimation or calibration, but an investigation of pricing accuracy of the closed-form solution. Thus, pricing was undertaken using the risk-neutral measure, $\mathbb{Q}$, to avoid arbitrage and comparing the closed-form solution above to that of an ODE solver. Using one set of parameters ensures pricing differences come from the different methods and not from parameter estimation.

	$n = 1$		$n = 3$	
	$i = 1$	$i = 1$	$i = 2$	$i = 3$
κ_i	0.0553	0.5509	1.0187	0.1330
σ_i	0.3325	1.0497	1.4274	0.5157
$\alpha_{0,i}$	0.0045	0.0000	0.0020	-0.0097
$\alpha_{1,i}$	0.0131	0.0046	0.0265	0.0323
γ_i	0.3341	0.1777	1.1623	0.8282
ρ_i	0.4615	0.3270	0.2268	0.1777
ψ	0.0832		0.0680	

Tab. 3.1: Table of parameters under $\mathbb{Q}$ according to Trolle and Schwartz (2008).

The parameter ψ was identified in the estimation process and can be interpreted as the infinite-maturity forward rate. Trolle and Schwartz (2008) replace $f(0, T)$ with ψ so that the initial forward curve is replaced by its time-homogeneous counterpart. This results in the simplification of $\frac{P(0,T)}{P(0,t)} = e^{-\psi(T-t)}$ in (2.20). Furthermore, Trolle and Schwartz (2008) state that x_i and $\phi_{i,j}$ are 0 and $v_i(t)$ are set to 1 under $\mathbb{Q}$. They note that derivative prices are based on the actual forward rate curve and therefore independent of x_i and $\phi_{i,j}$. Finally, the last simplification made was to normalize σ_i to 1, which means $\sigma_i = 1$ for all $i = 1, 2, 3$.

The instruments to be priced are the caps and swaptions as discussed in Chapter 2. Yearly ATMF caps are priced ranging from tenors of 1-10 years. The ATMF swaption prices are written in A x B form where A is the time until option expiry and B is the underlying swap tenor. The US standard of 3 months (or a quarter year) is used between reset and payment dates. This means that $\delta = 0.25$ in Equations (2.43) and (2.40).

3.5 The Basic Model

Using the above model parameterization, option prices were found using MATLAB's ODE45 solver before implementing the closed-form solution. The use of the reliable and accurate ODE45 solver in MATLAB will serve as a baseline for accuracy. Trolle and Schwartz (2008) do not report actual prices but pricing errors and their associated surfaces instead. Thus, prices were compared to that of Schumann (2016) to ensure the implementation of the model was correct in MATLAB. Schumann (2016) used Monte-Carlo simulation for his prices and will differ slightly but they are relatively consistent. Tables 3.2 and 3.3 contain a summary of the comparative prices found by Schumann (2016) who used a QE-scheme (QE) and the ODE45 solver. Using the ODE45 solver and the above parameters under $\mathbb{Q}$ will be henceforth called the basic model. There are a few small differences in bond prices from Monte-Carlo error which are used in caplet pricing. Each caplet uses two bond prices and the bond error accumulates slightly in the caplets. Since caps are a sum of caplets, longer dated caps will have slightly higher accumulated errors then the other instruments. The focus is initially placed on the stochastic duration approach rather than the implied volatility method. The basic model will be used as a test for accuracy by comparing it to the closed-form solution for $n = 1$ and $n = 3$.

Instrument	QE	Basic Model	Abs. Difference
P(0,1)	0.920166	0.920167	0.000001
P(0,5)	0.659680	0.659680	0
P(0,10)	0.435038	0.435178	0.000140
Cap 1Y	0.001666	0.001682	0.000016
Cap 5Y	0.028856	0.030072	0.001216
Cap 10Y	0.065931	0.068817	0.002886
Swaption 3Mx1Y	0.001889	0.001897	0.000008
Swaption 2Yx3Y	0.019036	0.019360	0.000324
Swaption 5Yx5Y	0.030178	0.030370	0.000192

Tab. 3.2: Pricing differences between Schumann (2016) QE-scheme results and the basic model for $n = 1$. All options are ATMF.

Instrument	QE	Basic Model	Abs. Difference
P(0,1)	0.934264	0.934260	0.000004
P(0,5)	0.711789	0.711770	0.000019
P(0,10)	0.506479	0.506617	0.000138
Cap 1Y	0.001721	0.001862	0.000141
Cap 5Y	0.023315	0.029598	0.006283
Cap 10Y	0.062322	0.066201	0.003879
Swaption 3Mx1Y	0.001893	0.001947	0.000054
Swaption 2Yx3Y	0.015447	0.016440	0.000993
Swaption 5Yx5Y	0.025956	0.026688	0.000732

Tab. 3.3: Pricing differences between Schumann (2016) QE-scheme results and the basic model for $n = 3$. All options are ATMF.

A Gauss-Legendre quadrature with 20 points over $[0, 1000]$ and a further 20 more over $[1000, 8000]$ was used to estimate the integral in (2.34) by (Schumann, 2016). This scheme was also used in the basic model to ensure consistency and confirm accuracy. (Schumann, 2016) did suggest using a finer integration scheme with more points. Quadrature selection can be parameter dependent and experimentation of the quadrature will be considered in the results section.

Chapter 4

Results

The closed-form analysis is presented by comparing the pricing accuracy of the two closed-form solutions against that of the basic model across different interest rate options. These solutions have never been implemented in this model before and the difference in accuracy will be due to the different methods used to solve the ODEs. The basic model provides a useful baseline as it was confirmed as accurate using the stochastic duration approach in the previous chapter. The procedure is as follows:

1. Compare the first general closed-form solution to the basic model for $n = 1$ using the stochastic duration approach.

2. Compare the second closed-form solution to the basic model for $n = 1$ using the stochastic duration approach.

3. If the closed-form solutions are accurate for $n = 1$, consideration will be given to extend the model to $n = 3$.

An issue occurs with the third point on the list. Each M and N_i value can have four theoretical values given the combinations of (β, w) pairs. This means that there are sixteen theoretical combinations for $n = 2$ and sixty-four for $n = 3$. If a closed-form solution is accurate, consideration will be given for the most accurate pair and that pair will be used throughout each M and N_i value. This means that if $++$ is most accurate for $n = 1$, then $++$ will be used for each M_i and N_i value for $i = 2, 3$ as well. If the closed-form solution is inaccurate for $n = 1$ then the same pair will not be increased for $n = 2$ or $n = 3$.

4.1 Results of the First Closed-Form Solution

Instrument	$--$	$+-$	$-+$	$++$	Basic Model
Cap 1Y	0.000780	0.000780	0.000780	0.000780	0.000786
Cap 2Y	0.002213	0.002213	0.002213	0.002213	0.002240
Cap 3Y	0.003730	0.003730	0.003730	0.003730	0.003788
Cap 4Y	0.005202	0.005202	0.005202	0.005202	0.005294
Cap 5Y	0.006584	0.006584	0.006584	0.006584	0.006706
Swaption 3Mx1Y	NaN	0.000696	NaN	0.000696	0.000698
Swaption 2Yx3Y	NaN	0.002729	NaN	0.002729	0.002781
Swaption 5Yx5Y	NaN	0.002913	NaN	0.002913	0.002968

Tab. 4.1: Interest rate option pricing results of each pair compared to the basic model for $n = 1$.

The initial results in Table 4.1 above appear to be consistent for $n = 1$ relative to the basic model. All four of the pairs provide the same results when pricing to six decimal places for caps. However, the stochastic duration approach encountered an error when pricing swaptions and increased numerical precision to 64 decimal places was required to get results. This allowed accurate pricing to take place for two pairs but the other two β_- pairs resulted in NaN values (Not a Number values). This is due to the Kummer and Tricomi functions producing NaN values under certain parameterizations. The functions struggle to produce numbers under some parameters as they do not have the precision to represent the functions. It is important to note that a Kummer function written by Patrick Mousaw on MATLAB file exchange is accurate and available online. The accuracy tolerance can be adjusted in this function for more precision. However, it seems MATLAB cannot handle certain parameter cases required in this model even when extending precison using the VPA function. Schumann (2016) actually suggests using different software such as $C++$ for better precision.

The original use of Gauss-Legendre integration scheme with 20 points over $[0, 1000]$ and 20 more over $[1000, 8000]$ seems very suitable and accurate for pricing. This is confirmed by the results in Table 4.2 below.

	20:20	40:40	100:100	200:200	1000
$--$	0.000780	0.000780	0.000780	0.000780	0.000780
$+-$	0.000780	0.000780	0.000780	0.000780	0.000780
$-+$	0.000780	0.000780	0.000780	0.000780	0.000780
$++$	0.000780	0.000780	0.000780	0.000780	0.000780
Basic Model	0.000786	0.000786	0.000786	0.000786	0.000786

Tab. 4.2: Total number of quadrature points used for pricing a 1-year cap for $n = 1$.

The label '20:20' in Table 4.2 above refers to using 20 points over $[0, 1000]$ and 20 more over $[1000, 8000]$, '40:40' refers to using 40 points over $[0, 1000]$ and 40 more over $[1000, 8000]$ and so on. '1000' refers to using 1000 points over $[0, 8000]$ straight and graphs the sinusoidal shape of the characteristic function as shown in (C.1). Thus, the use of more quadrature points does not affect pricing of the closed-form solution nor the basic model. The change in quadrature would only change the 13^{th} or 14^{th} decimal place in prices and it is quite impressive that the integral can be captured by only using 40 points in total. The use of the original 40 points will be used for pricing due to its efficiency and lack of effect on accuracy.

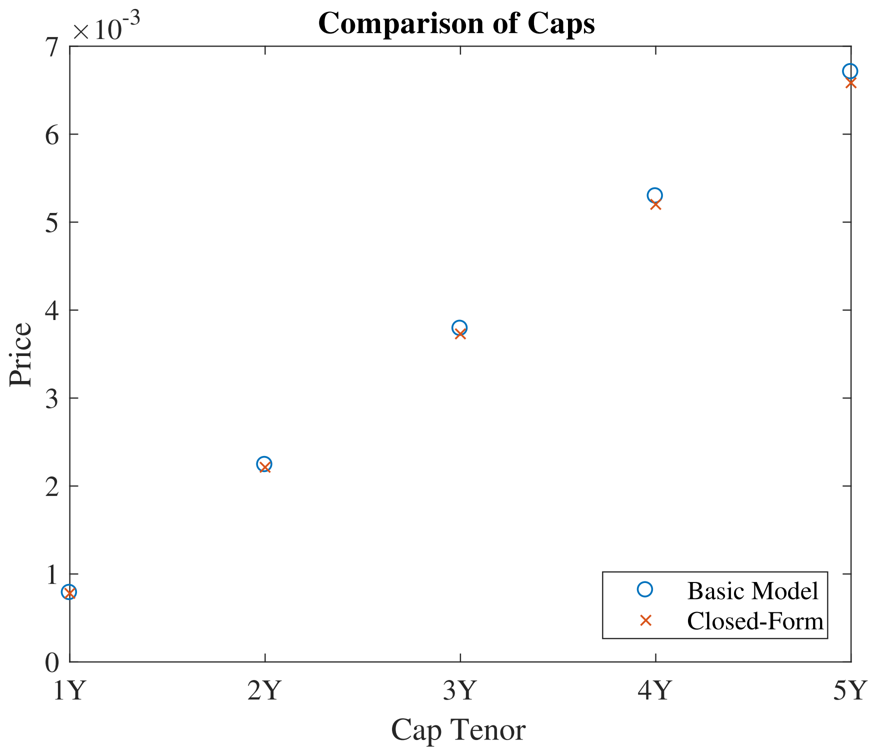


Fig. 4.1: A comparsion of cap prices of various tenors between the closed-form and basic model for $n = 1$.

Figure 4.1 above shows the basic model and closed-form are relatively accurate. They do differ more as the cap tenor increases. This is most likely due to the error tolerance of the ODE45 solver in the basic model. The error tolerance is prescribed for each caplet and adding more caplets will increase total error of a cap as shown in the Figure 4.1. Figure 4.2 identifies the relatively small absolute difference in cap prices of various tenors.

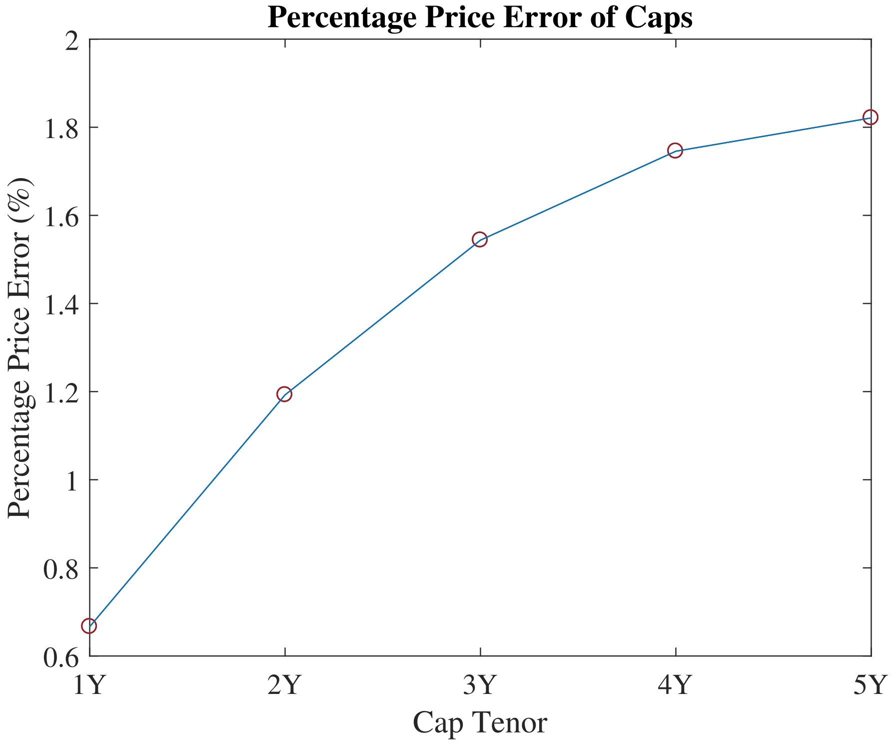


Fig. 4.2: Percentage Difference in cap prices of various tenors between the closed-form and basic model for $n = 1$.

The first closed-form solution is relatively efficient to implement and the time elapsed to price caps are recorded in Table 4.3. All four pairs take around 0.2 seconds to price caps of various tenors and use the previous cap value found. The efficient ODE45 solver in the basic model is slightly faster but not by much in absolute terms. The code for the closed-form solution can be streamlined and computational times could reduce. Also, the use of high-powered computing can reduce it even further.

Instrument	$--$	$+-$	$-+$	$++$	Basic Model
Cap 1Y	0.256	0.228	0.235	0.220	0.150
Cap 2Y	0.229	0.233	0.225	0.236	0.110
Cap 3Y	0.219	0.247	0.219	0.220	0.078
Cap 4Y	0.214	0.253	0.215	0.213	0.085
Cap 5Y	0.211	0.269	0.208	0.208	0.083

Tab. 4.3: Time to price each instrument measured in seconds for $n = 1$.

4.1.1 Extension to $n = 3$

Increasing the model to $n = 3$ prodcued some more issues. The selection of $\alpha_{0,1} = 0 = \alpha_1$ from the parameter set ensures all the coefficients are equal to zero for N_1. This creates many divisions by zero in the formulae and it would be impossible to evaluate the first closed-form solution. However, the ODE45 solver can be used to check the answer and states that $N_1 = 0$ for all values of u in this case. Therefore, $N_1 = 0$ for $n = 1$ and thus the first iteration is skipped as adding zero to $M(\tau)$ and the characteristic function has no effect.

Instrument	$--$	$+-$	$-+$	$++$	Basic Model
Cap 1Y	0.001440	0.001440	0.001440	0.001440	0.001462
Cap 2Y	0.003755	0.003755	0.003755	0.003755	0.003822
Cap 3Y	0.005988	0.005988	0.005988	0.005988	0.006065
Cap 4Y	0.008061	0.008061	0.008061	0.008061	0.008082
Cap 5Y	0.009977	0.009977	0.009977	0.009977	0.009866
Swaption 3Mx1Y	NaN	NaN	NaN	NaN	0.001173
Swaption 2Yx3Y	NaN	NaN	NaN	NaN	0.002743
Swaption 5Yx5Y	NaN	NaN	NaN	NaN	0.002112

Tab. 4.4: Interest rate option pricing results of each pair compared to the basic model for $n = 3$.

The closed-form solution performs accurately in Table 4.4 above when the model is extended to three stochastic volatility terms. Again, all four pairs result in the same price for caps. However, the NaN issue persists for swaptions and occurs in all four pairs. It seems that values in the Kummer and Tricomi functions denominator converge to zero and can't be avoided even when increasing precision. Perhaps more powerful computers or software is required. Sitzia (2018) actually states that using $C\#$ could be more precise and more efficient.

Once again, the increase of the number of quadrature points used does not affect the prices to six decimal places. The '20:20' quadrature was used for pricing in this formulation. All four pairs price a 1-year cap to 0.001440 whilst the basic model produces 0.001462. Only the 14^{th} decimal changes in closed-form pairs. However, the 8^{th} decimal place changes in the basic model under the last quadrature scheme used.

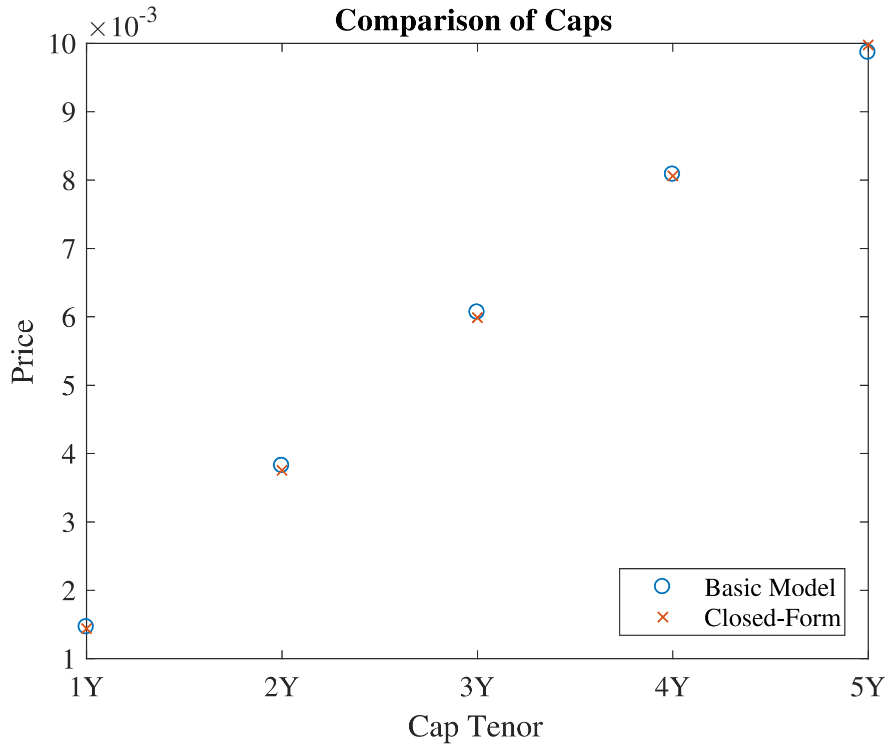


Fig. 4.3: A comparsion of cap prices of various tenors between the closed-form and basic model for $n = 3$.

The accuracy of the first closed-form solution can be observed in Figures 4.3 above and 4.4 on the next page. Figure 4.3 plots the comparative prices and Figure 4.4 on the next page highlights the small absolute difference in prices.

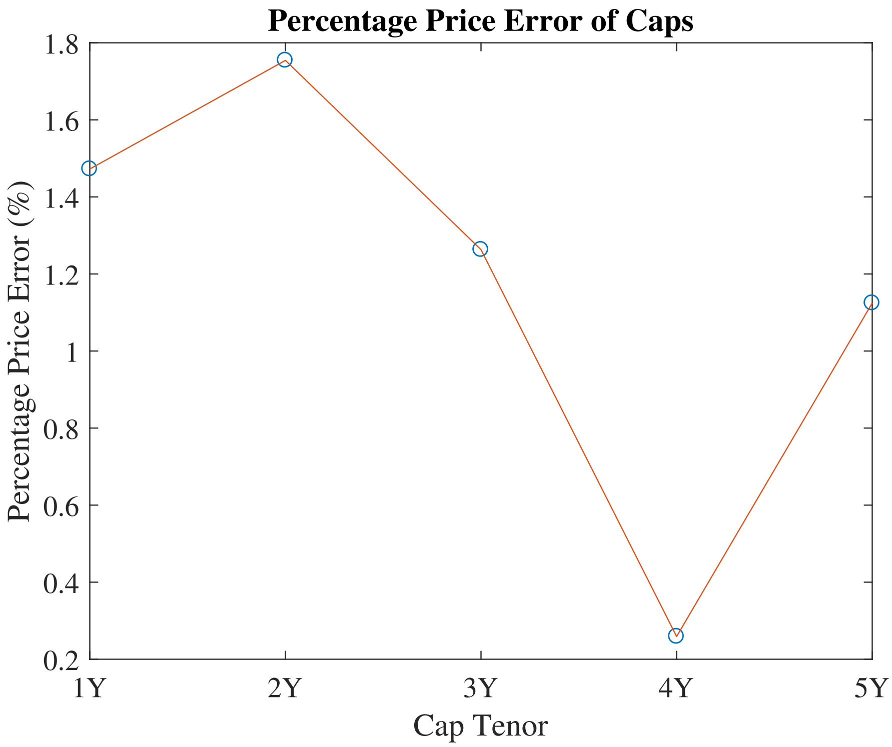


Fig. 4.4: Percentage Difference in cap prices of various tenors between the closed-form and basic model for $n = 3$.

The $n = 3$ is also relatively computationally economical to implement. All four pairs take roughly 0.4 seconds to price each cap and the times are recorded in Table 4.5 below. These times are expected to be longer than the $n = 1$ case as there are an extra two $v_i(t)$ terms. The ODE45 solver in the basic model is once again faster. Although the basic model takes half the time, the difference in time elapsed is tiny in absolute terms.

Instrument	$--$	$+-$	$-+$	$++$	Basic Model
Cap 1Y	0.406	0.387	0.380	0.393	0.180
Cap 2Y	0.415	0.414	0.413	0.426	0.156
Cap 3Y	0.403	0.398	0.399	0.398	0.138
Cap 4Y	0.391	0.388	0.388	0.390	0.152
Cap 5Y	0.384	0.379	0.376	0.380	0.139

Tab. 4.5: Time to price each instrument measured in seconds for $n = 3$.

4.2 Results of the Second Closed-Form Solution

This second closed-form solution was proposed as an alternative to the first closed-form solution. It is also an interesting experiment to see if other solutions of confluent hypergeometric functions work in this model. The second closed-form solution produces exactly the same accurate values to six decimal places as the first closed-form solution in Table 4.1. However, it unfortunately also exhibits the same NaN problem in swaption pricing as the first closed-form solution does. Similarly, the original use of Gauss-Legendre integration scheme with 20 points over $[0, 1000]$ and 20 more over $[1000, 8000]$ seems very suitable and accurate for pricing. The use of more quadrature points does not affect pricing of the second closed-form solution either and the values are the same as those of Table 4.2. This is further confirmed by the integrands of each pair tending to zero in (C.2). Since the values are exactly the same as the first closed-form solution for $n = 1$, the graphs of the comparison of cap prices and absolute difference of cap prices would be the same. Figures 4.5 and 4.6 below show the relative accuracy of the swaption prices between the second closed-form solution and the basic model. Figure 4.5 expresses visually how close the prices are and Figure 4.6 on the next page graphs the small absolute difference in prices.

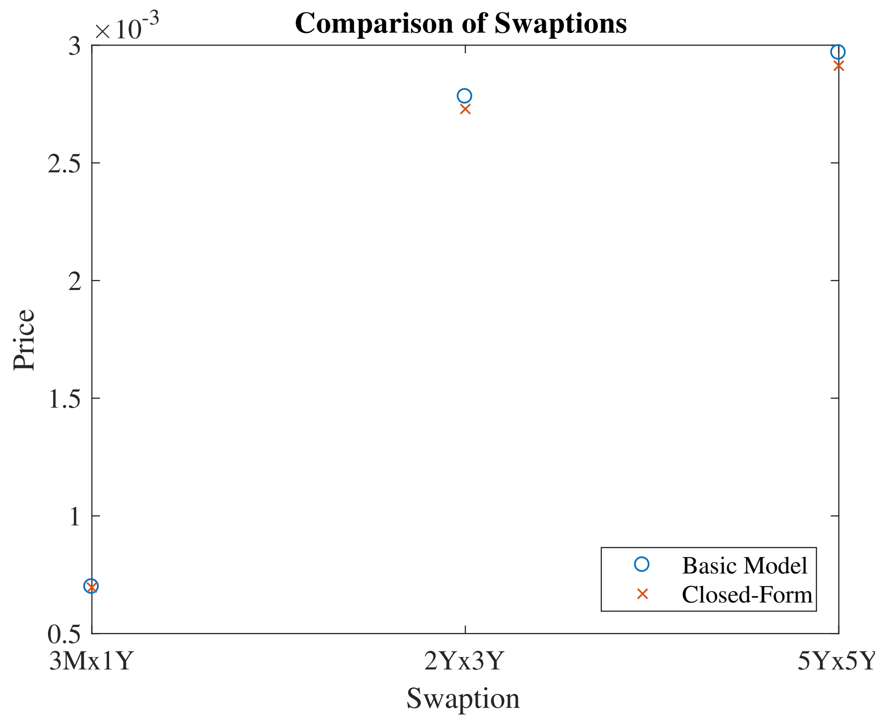


Fig. 4.5: A comparsion of swaption prices of various tenors and maturities between the closed-form and basic model for $n = 1$.

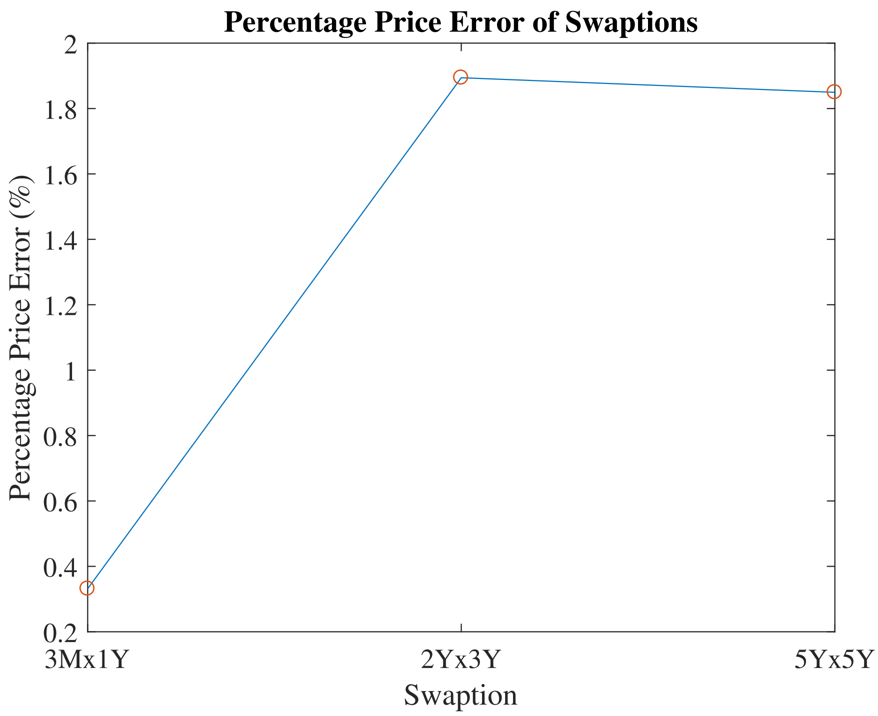


Fig. 4.6: Percentage Difference in swaption prices of various tenors and maturities between the closed-form and basic model for $n = 1$.

Similarly to the first closed-form solution, the second closed-form solution also takes around 0.2 seconds for a pair to price a cap. The results are recorded in Table 4.6 below. The ODE45 solver in the basic model was consistent once again in pricing slightly faster than the second closed-form pairs.

Instrument	$--$	$+-$	$-+$	$++$	Basic Model
Cap 1Y	0.249	0.245	0.233	0.267	0.088
Cap 2Y	0.229	0.228	0.232	0.229	0.104
Cap 3Y	0.229	0.223	0.222	0.224	0.081
Cap 4Y	0.221	0.227	0.222	0.219	0.073
Cap 5Y	0.219	0.218	0.215	0.215	0.066s

Tab. 4.6: Time to price each instrument measured in seconds for $n = 1$.

4.2.1 Extension to $n = 3$

The same issue in the first-closed form case for $n = 3$ exists here and was dealt with in the same manner. The results for $n = 3$ are exactly the same (to six decimal places) as the first closed-form solution in Table 4.4 for $n = 3$. The NaN error for swaptions persists and also can't be resolved in this second formulation. For consistency, the quadrature was checked in the same manner as the other solutions. There was no effect on prices as the number of quadrature points was increased and produced the same values for a 1-year cap in the first solution. The use of 20 points over $[0, 1000]$ and 20 more over $[1000, 8000]$ still remains efficient and suitable for accurate results.

The $n = 3$ case is also computationally inexpensive to implement as seen in Table 4.7 on the next page. As with the first closed-form case for $n = 3$, all four pairs take roughly 0.4 seconds to price each cap. The ODE45 solver in the basic model is once again faster. Although the basic model takes less than half the time, the difference is small.

Instrument	$--$	$+-$	$-+$	$++$	Basic Model
Cap 1Y	0.390	0.384	0.397	0.398	0.162
Cap 2Y	0.433	0.420	0.409	0.418	0.151
Cap 3Y	0.421	0.392	0.396	0.390	0.155
Cap 4Y	0.383	0.383	0.438	0.381	0.137
Cap 5Y	0.373	0.377	0.396	0.374	0.138

Tab. 4.7: Time to price each instrument measured in seconds for $n = 3$.

Chapter 5

Conclusion

Solving ODEs is crucial when pricing with the semi-analytical pricing formulae in the Trolle and Schwartz (2008) interest rate model. An ODE solver was the preferred method as per Trolle and Schwartz (2008) and Schumann (2016) to find estimates for the ODEs in the characteristic function. However, using the recent proof and implementation of confluent hypergeometric functions by Sitzia (2018), a closed-form solution for the ODEs and resulting characteristic function could be found. Not one, but two closed-form solutions were found and compared to MATLAB's ODE45 solver as a baseline for accuracy.

Both solutions require the TS model to be of a specific form which is the stochastic volatility version of the Hull and White (1990) model. Once this form was achieved through parameter manipulation, caps and swaptions were priced using both closed-form solutions for $n = 1$ and $n = 3$. The pricing of caps and swaptions for $n = 1$ were deemed accurate by both solutions relative to the basic model and both solutions attained the exact same prices for each selected instrument. A slight problem exists were the Kummer and Tricomi functions perform poorly in handling certain parameters. This resulted in two pairs in both solutions producing non-numbers as answers when pricing swaptions for $n = 1$ but the other two pairs in both solutions were stable and precise. The solutions continued to be accurate for $n = 3$ and again achieved the same prices for each instrument. However, no pair in the $n = 3$ case could produce a stable swaption price due to the Kummer and Tricomi precision issue. The efficiency of the solutions was compared to the ODE solver. For $n = 1$ and $n = 3$, the closed-form solutions were equally efficient as each other but slightly slower than the ODE solver. With modern computing and streamlining of the code implemented, these times can be reduced and efficiency can be improved. This dissertation provides two closed-form solutions for the TS characteristic function of a particular case. It also provides further evidence that the use of confluent hypergeometric equations developed by Sitzia (2018) to solve ODEs is accurate and efficient.

Bibliography

Abramowitz, M. and Stegun, I. (1964). Handbook of mathematical functions: With formulas, graphs, and mathematical tables applied mathematics series, *National Bureau of Standards, Washington, DC* .

Andersen, T. G. and Lund, J. (1997). Estimating continuous-time stochastic volatility models of the short-term interest rate, *Journal of Econometrics* **77**(2): 343–377.

Björk, T. (2009). *Arbitrage Theory in Continuous Time*, Oxford university press.

Brigo, D. and Mercurio, F. (2007). *Interest Rate Models-Theory and Practice: with Smile, Inflation and Credit*, Springer Science and Business Media.

Casassus, J., Collin-Dufresne, P. and Goldstein, B. (2005). Unspanned stochastic volatility and fixed income derivatives pricing, *Journal of Banking and Finance* **29**(11): 2723–2749.

Cheyette, O. (1992). Term structure dynamics and mortgage valuation, *The Journal of Fixed Income* **1**(4): 28–41.

Cheyette, O. (1995). Markov representation of the Heath-Jarrow-Morton model. Available:https://ssrn.com/abstract=6073.

Chiarella, C. and Kwon, O. K. (2000). A complete Markovian stochastic volatility model in the HJM framework, *Asia-Pacific Financial Markets* **7**(4): 293–304.

Chiarella, C. and Kwon, O. K. (2001). Classes of interest rate models under the HJM framework, *Asia-Pacific Financial Markets* **8**(1): 1–22.

Collin-Dufresne, P. and Goldstein, R. (2003). Generalizing the affine framework to HJM and random field models. Available:https://ssrn.com/abstract=410421.

Collin-Dufresne, P. and Goldstein, R. S. (2002). Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility, *The Journal of Finance* **57**(4): 1685–1730.

Corrado, C. J. and Su, T. (1997). Implied volatility skews and stock return skewness and kurtosis implied by stock option prices, *The European Journal of Finance* **3**(1): 73–85.

Cox, J., Ingersoll Jr, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates, *Econometrica* **53**(2): 385–408.

Dai, Q. and Singleton, K. J. (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* **63**(3): 415–441.

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates, *Mathematical Finance* **6**(4): 379–406.

Duffie, D., Pan, J. and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* **68**(6): 1343–1376.

Han, B. (2007). Stochastic volatilities and correlations of bond yields, *The Journal of Finance* **62**(3): 1491–1524.

Heath, D., Jarrow, R. and Morton, A. (1989). Contingent claim valuation with a random evolution of interest rates, *Technical report*, Cornell University Operations Research and Industrial Engineering.

Heath, D., Jarrow, R. and Morton, A. (1990). Bond pricing and the term structure of interest rates: A discrete time approximation, *Journal of Financial and Quantitative Analysis* **25**(4): 419–440.

Heath, D., Jarrow, R. and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica: Journal of the Econometric Society* pp. 77–105.

Heidari, M. and Wu, L. (2003). Are interest rate derivatives spanned by the term structure of interest rates?, *The Journal of Fixed Income* **13**(1): 75–86.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies* **6**(2): 327–343.

Ho, T. S. and Lee, S. (1986). Term structure movements and pricing interest rate contingent claims, *The Journal of Finance* **41**(5): 1011–1029.

Hobson, D. G. and Rogers, L. C. (1998). Complete models with stochastic volatility, *Mathematical Finance* **8**(1): 27–48.

Hull, J. and White, A. (1990). Pricing interest-rate-derivative securities, *The Review of Financial Studies* **3**(4): 573–592.

Jarrow, R., Li, H. and Zhao, F. (2007). Interest rate caps "smile" too! But can the LIBOR market models capture the smile?, *The Journal of Finance* **62**(1): 345–382.

Li, H. and Zhao, F. (2006). Unspanned stochastic volatility: Evidence from hedging interest rate derivatives, *The Journal of Finance* **61**(1): 341–378.

Munk, C. (1999). Stochastic duration and fast coupon bond option pricing in multi-factor models, *Review of Derivatives Research* **3**(2): 157–181.

Pearson, J. W., Olver, S. and Porter, M. A. (2017). Numerical methods for the computation of the confluent and Gauss hypergeometric functions, *Numerical Algorithms* **74**(3): 821–866.

Schumann, G. W. (2016). Trolle-Schwartz HJM interest rate model. University of Cape Town.

Shampine, L. F. and Reichelt, M. W. (1997). The MATLAB ODE suite, *SIAM Journal on Scientific Computing* **18**(1): 1–22.

Sitzia, N. (2018). The analytical solution of Trolle-Schwartz Model. `Available: https://ssrn.com/abstract=3121906`.

Trolle, A. B. and Schwartz, E. S. (2008). A general stochastic volatility model for the pricing of interest rate derivatives, *The Review of Financial Studies* **22**(5): 2007–2057.

Vasicek, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**(2): 177–188.

Wei, J. Z. (1997). A simple approach to bond option pricing, *Journal of Futures Markets: Futures, Options, and Other Derivative Products* **17**(2): 131–160.

Appendix A

Closed-Form Proofs

A.1 Proof of Proposition 3.1

The proof follows that of Sitzia (2018) and is applied to our ODE in (2.32). First β_{x_i} is simplified,

$$\begin{aligned}\beta_{x_i}(\tau) &= \tfrac{\alpha_{1,i}}{\gamma_i}((\tfrac{1}{\gamma_i}+\tfrac{\alpha_{0,i}}{\alpha_{1,i}})(e^{-\gamma_i\tau}-1)+\tau e^{-\gamma_i\tau})\\ &= (\tfrac{\alpha_{1,i}}{\gamma_i^2}+\tfrac{\alpha_{0,i}}{\gamma_i})(e^{-\gamma_i\tau}-1)+\tfrac{\alpha_{1,i}}{\gamma_i}\tau e^{-\gamma_i\tau}.\end{aligned}$$

Setting $\alpha_{1,i} = 0$, the model recovers the stochastic volatility version of the Hull-White model.

$$\begin{aligned}\beta_{x_i}(\tau) &= \tfrac{\alpha_{0,i}}{\gamma_i}(e^{-\gamma_i\tau}-1)\\ &= \tfrac{\alpha_i}{\gamma_i}(1-e^{-\gamma_i\tau}),\quad \text{where}\alpha_i = -\alpha_{0,i}.\end{aligned}$$

The ODE in (2.32) can be written in the form

$$N_i'(\tau) = q_0(\tau) + q_1(\tau)N_i(\tau) + q_2(\tau)N_i(\tau)^2, \tag{A.1}$$

where

$$\begin{aligned}q_0(\tau) &= \tfrac{1}{2}\tfrac{\alpha_i^2}{\gamma_i^2}(u^2-u)(1-e^{-\gamma_i(T_1-T_0+\tau)})^2+\tfrac{1}{2}\tfrac{\alpha_i^2}{\gamma_i^2}((1-u)^2-(1-u))(1-e^{-\gamma_i\tau})^2\\ &\quad+\tfrac{\alpha_i^2}{\gamma_i^2}u(1-u)(1-e^{-\gamma_i(T_1-T_0+\tau)})(1-e^{-\gamma_i\tau}),\\ q_1(\tau) &= -\kappa_i+\sigma_i\rho_i\tfrac{\alpha_i}{\gamma_i}(u(1-e^{-\gamma_i(T_1-T_0+\tau)})+(1-u)(1-e^{-\gamma_i\tau})),\\ q_2(\tau) &= \tfrac{1}{2}\sigma_i^2.\end{aligned}$$

Apply the transformation of $N_i(\tau) = \frac{-y'(\tau)}{y(\tau)q_2(\tau)}$ and note that q_2 is constant. The ODE now reduces to

$$\begin{aligned}&\left(\frac{-y'(\tau)}{y(\tau)q_2(\tau)}\right)' = q_0(\tau)+q_1(\tau)\left(\frac{-y'(\tau)}{y(\tau)q_2(\tau)}\right)+q_2(\tau)\left(\frac{-y'(\tau)}{y(\tau)q_2(\tau)}\right)^2\\ \Rightarrow\ &\frac{-y'(\tau)^2}{y(\tau)^2q_2}-\frac{-y''(\tau)}{y(\tau)q_2} = q_0(\tau)-q_1(\tau)\left(\frac{y'(\tau)}{y(\tau)q_2(\tau)}\right)+\frac{y'(\tau)^2}{y(\tau)^2q_2(\tau)}.\end{aligned}$$

This simplifies to a linear, second-order, homogeneous ODE with non-constant coefficients

$$y''(\tau) - q_1(\tau)y'(\tau) + q_0(\tau)q_2 y(\tau) = 0. \tag{A.2}$$

Further simplification is done by grouping constant coefficients and exponential coefficients

$$y''(\tau) - (c_0 + c_1 e^{-\gamma_i \tau})y'(\tau) + (d_0 + d_1 e^{-\gamma_i \tau} + d_2 e^{-2\gamma_i \tau})y(\tau) = 0, \tag{A.3}$$

where the coefficients are

$$\begin{aligned}
c_0 &= -\kappa_i + \sigma_i \rho_i \tfrac{\alpha_i}{\gamma_i}, \\
c_1 &= -\sigma_i \rho_i \tfrac{\alpha_i}{\gamma_i}(u e^{-\gamma_i(T_1 - T_0)} + (1-u)), \\
d_0 &= 0, \\
d_1 &= 0, \\
d_2 &= \tfrac{\sigma_i^2 \alpha_i^2}{4\gamma_i^2}(u^2 - u)\big[e^{-2\gamma_i(T_1 - T_0)} - 2e^{-\gamma_i(T_1 - T_0)} + 1\big].
\end{aligned}$$

Another substitution of $x = e^{-\gamma_i \tau}$ is applied and $f(x) = y(\frac{-\log(x)}{\gamma_i}) = y(\tau)$. The first two derivatives of $y(\tau)$ become

$$\begin{aligned}
\tfrac{d}{d\tau} y(\tau) &= \tfrac{dx}{d\tau}\tfrac{d}{dx} y(\tfrac{-\log(x)}{\gamma_i}) = -\gamma_i e^{-\gamma_i \tau} \tfrac{d}{dx} y(\tfrac{-\log(x)}{\gamma_i}) = -\gamma_i x f'(x), \\
\tfrac{d^2}{d\tau^2} y(\tau) &= \tfrac{dx}{d\tau}\tfrac{d}{dx}(-\gamma_i x f'(x)) = \gamma_i^2 x(f'(x) + x f''(x)).
\end{aligned}$$

The resulting ODE becomes

$$\gamma_i^2 x^2 f''(x) + \gamma_i x f'(x)(\gamma_i + c_0 + c_1 x) + f(x)(d_0 + d_1 x + d_2 x^2) = 0. \tag{A.4}$$

One final substitution is made through parameters w, , β, μ, which will be specified later to make Equation (A.4) a confluent hypergeometric equation,

$$\begin{aligned}
z &= \tfrac{x}{w}, \\
f(x) &= z^\beta e^{\mu z} g(z).
\end{aligned}$$

The derivatives change to

$$\begin{aligned}
f'(x) &= \tfrac{dz}{dx}\tfrac{d}{dz}(z^\beta e^{\mu z} g(z)) = \tfrac{1}{w} z^\beta e^{\mu z}\big[\big(\tfrac{\beta}{z} + \mu\big)g(z) + g'(z)\big], \\
f''(x) &= \tfrac{dz}{dx}\tfrac{d}{dz}\big[\tfrac{dz}{dx}\tfrac{d}{dz}\big(z^\beta e^{\mu z} g(z)\big)\big] \\
&= \tfrac{1}{w^2} z^\beta e^{\mu z}\left[\left(\big(\tfrac{\beta}{z} + \mu\big)^2 - \tfrac{\beta}{z^2}\right)g(z) + 2\big(\tfrac{\beta}{z} + \mu\big)g'(z) + g''(z)\right].
\end{aligned}$$

Placing these new derivatives into the ODE and simplifying $z^\beta e^{\mu z}$ results in

$$\begin{aligned}
0 = z^2 g''(z) + g'(z)\left[z^2\left(2\mu + \tfrac{c_1 w}{\gamma_i}\right) + z\left(2\beta + \tfrac{\gamma_i + c_0}{\gamma_i}\right)\right] + g(z)\left[z^2\left(\mu^2 + \mu\tfrac{c_1 w}{\gamma_i} + \tfrac{d_2 w^2}{\gamma_i^2}\right)\right. \\
\left. + z\left(\mu\left(\tfrac{c_0}{\gamma_i} + 1 + 2\beta\right) + \beta c_1 \tfrac{w}{\gamma_i} + d_1 \tfrac{w}{\gamma_i^2}\right) + \left(\beta^2 + \beta\tfrac{c_0}{\gamma_i} + \tfrac{d_0}{\gamma_i^2}\right)\right].
\end{aligned} \tag{A.5}$$

The following conditions are imposed:

$$
\begin{aligned}
&2\mu + \frac{c_1 w}{\gamma_i} = -1, \\
&\mu^2 + \mu\frac{c_1 w}{\gamma_i} + \frac{d_2 w^2}{\gamma_i^2} = 0, \\
&\beta^2 + \beta\frac{c_0}{\gamma_i} + \frac{d_0}{\gamma_i^2} = 0.
\end{aligned}
$$

The parameters are now given by:

$$
\begin{aligned}
\beta &= \frac{-c_0 \pm \sqrt{c_0^2 - 4d_0}}{2\gamma_i}, \\
w &= \frac{\pm\gamma_i}{\sqrt{c_1^2 - 4d_2}}, \\
\mu &= -\frac{1}{2}\big(1 + \frac{c_1 w}{\gamma_i}\big).
\end{aligned}
$$

Finally we divide by z and (A.5) is reduced to the confluent hypergeometric equation

$$zg''(z) + (b - z)g'(z) - ag(z) = 0, \tag{A.6}$$

where

$$
\begin{aligned}
a &= -\mu\big(\tfrac{c_0}{\gamma_i} + 1 + 2\beta\big) - \beta c_1 \tfrac{w}{\gamma_i} - d_1 \tfrac{w}{\gamma_i^2}, \\
b &= 2\beta + 1 + \tfrac{c_0}{\gamma_i}.
\end{aligned}
$$

The solution can now be provided using linearly independent Kummer and Tricomi functions $M(a, b, z)$ and $U(a, b, z)$,

$$g(z) = k_1 M(a, b, z) + k_2 U(a, b, z). \tag{A.7}$$

Using the fact that $\frac{d}{dz}M(a, b, z) = \frac{a}{b}M(a + 1, b + 1, z)$ and $\frac{d}{dz}U(a, b, z) = -aU(a + 1, b + 1, z)$, the first derivative of $g(z)$ can be found

$$g'(z) = k_1 \tfrac{a}{b} M(a + 1, b + 1, z) - k_2 a U(a + 1, b + 1, z).$$

The substitutions are reversed and a solution for $N(\tau)$ is established

$$N(\tau) = \tfrac{-y'(\tau)}{y(\tau)q_2(\tau)} = \gamma_i x \frac{f'(x)}{f(x)\frac{\sigma_i^2}{2}} = \tfrac{2\gamma_i}{\sigma_i^2}\left(\beta + \mu z + z\tfrac{g'(z)}{g(z)}\right). \tag{A.8}$$

The constants k_1 and k_2 need to be defined in order to match the initial conditions

$$N(0) = 0 = \tfrac{2\gamma_i}{\sigma_i^2}\left(\beta + \mu\tfrac{1}{w} + \tfrac{1}{w}\tfrac{g'(\frac{1}{w})}{g(\frac{1}{w})}\right). \tag{A.9}$$

A restriction of $g(\frac{1}{w}) = 1$ is placed for sake of simplicity and result in the following two conditions:

$$
\begin{aligned}
&k_1 M(a, b, \tfrac{1}{w}) + k_2 U(a, b, \tfrac{1}{w}) = 1, \\
&k_1 \tfrac{a}{b} M(a + 1, b + 1, \tfrac{1}{w}) - k_2 a U(a + 1, b + 1, \tfrac{1}{w}) = -\beta w - \mu.
\end{aligned}
$$

These are ultimately solved as:

$$k_1 = \frac{-\beta w - \mu + a\frac{U(a+1,b+1,\frac{1}{w})}{U(a,b,\frac{1}{w})}}{\frac{a}{b}M(a+1,b+1,\frac{1}{w}) + aM(a,b,\frac{1}{w})\frac{U(a+1,b+1,\frac{1}{w})}{U(a,b,\frac{1}{w})}},$$
$$k_2 = \frac{1 - k_1 M(a,b,\frac{1}{w})}{U(a,b,\frac{1}{w})}.$$

□

A.2 Proof of Proposition 3.2

It is adequate to integrate the right-hand side of the equation as it is independent of $M(\tau)$

$$M(\tau) = \int \sum_{i=1}^{n} \kappa_i \theta_i N_i(\tau) d\tau. \tag{A.10}$$

The same variable transformations in the proof of (3.1) are applied here so $z = \frac{e^{-\gamma_i \tau}}{w}$ and $d\tau = -\frac{1}{\gamma_i z}dz$. The integral becomes trivial to solve

$$\begin{aligned} M(\tau) &= \int \sum_{i=1}^{n} \kappa_i \theta_i N_i(\tau) d\tau \\ &= \int \sum_{i=1}^{n} \kappa_i \theta_i \frac{2\gamma_i}{\sigma_i^2}\left(\beta + \mu z + z\frac{g'(z)}{g(z)}\right) d\tau \\ &= \sum_{i=1}^{n} \frac{2\kappa_i\theta_i}{\sigma_i^2}\left(\int -\mu - \frac{g'(z)}{g(z)}dz + \beta\gamma_i\tau\right), \\ &= \sum_{i=1}^{n} \frac{2\kappa_i\theta_i}{\sigma_i^2}\left(\beta\gamma_i\tau - \mu z - \log(g(z))\right) + k_3. \end{aligned} \tag{A.11}$$

Setting $M(0) = 0$ then

$$k_3 = \sum_{i=1}^{n} \frac{2\kappa_i\theta_i\mu}{\sigma_i^2 w}.$$

□

Appendix B

Confluent Hypergeometric Functions

B.1 Eight Solutions

The following eight solutions are taken from Abramowitz and Stegun (1964) and solve the confluent hypergeometric equation $zg''(z) + (b - z)g'(z) - ag(z) = 0$. The complete solution is $g(z) = k_1 M(a, b, z) + k_2 U(a, b, z)$ where $M(a, b, z)$ and $U(a, b, z)$ are two generally linearly independent solutions. However, there exist eight solutions:

1. $g_1(z) = M(a, b, z)$.
2. $g_2(z) = z^{1-b} M(1 + a - b, 2 - b, z)$.
3. $g_3(z) = e^z M(b - a, b, -z)$.
4. $g_4(z) = z^{1-b} e^z M(1 - a, 2 - b, -z)$.
5. $g_5(z) = U(a, b, z)$.
6. $g_6(z) = z^{1-b} U(1 + a - b, 2 - b, z)$.
7. $g_7(z) = e^z U(b - a, b, -z)$.
8. $g_8(z) = z^{1-b} e^z U(1 - a, 2 - b, -z)$.

B.2 Kummer Transformations

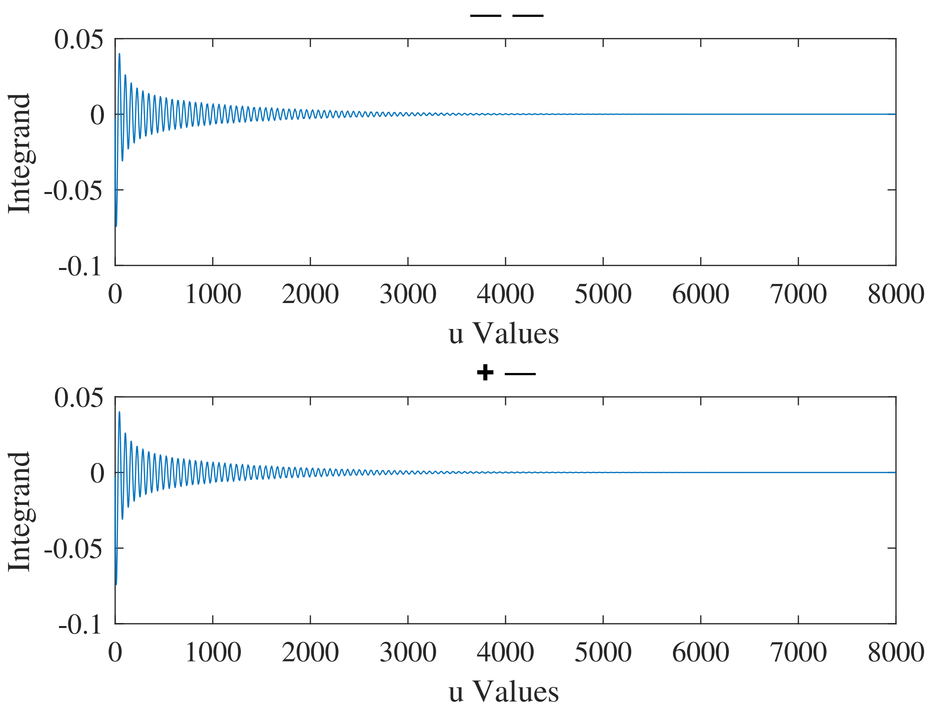
The following Kummer transformations are proven to be equal in Abramowitz and Stegun (1964):

1. $M(a, b, z) = e^z M(b - a, b, -z)$.
2. $z^{1-b} M(1 + a - b, 2 - b, z) = z^{1-b} e^z M(1 - a, 2 - b, -z)$.
3. $U(a, b, z) = z^{1-b} U(1 + a - b, 2 - b, z)$.

Appendix C

Additional Results

C.1 First Closed-Form Solution



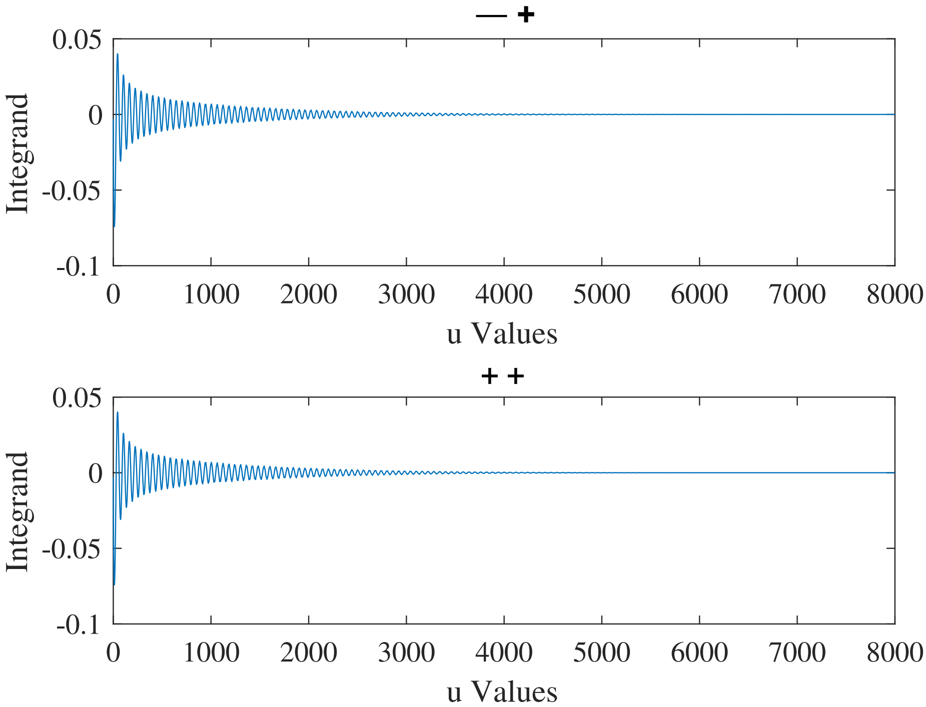
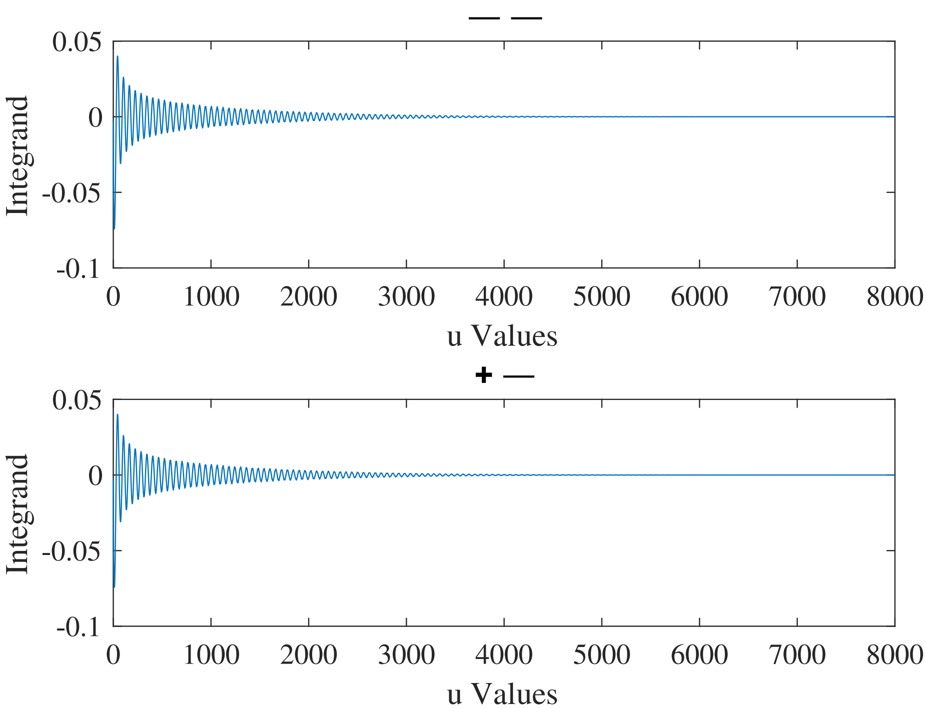


Fig. C.1: Integrand values for 1000 Gauss-Legendre values of u over $[0, 8000]$ in (2.34) for $n = 1$, $K = 0.08407$, $a = b = 1$, $t = 0$, $T_0 = 0.25$, $T_1 = 0.5$.

C.2 Second Closed-Form Solution



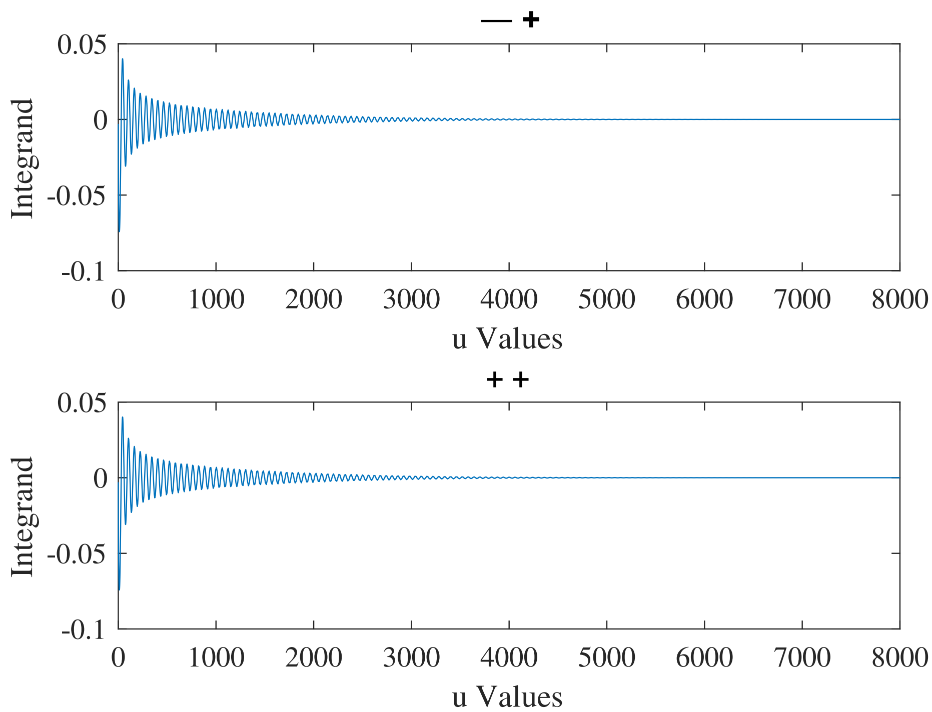


Fig. C.2: Integrand values for 1000 Gauss-Legendre values of u over $[0, 8000]$ in (2.34) for $n = 1$, $K = 0.08407$, $a = b = 1$, $t = 0$, $T_0 = 0.25$, $T_1 = 0.5$.