Quadratic Term Structure Models: Analysis and Performance*

Grégoire Leblon†
Université de Rennes 1 and CNRS

Franck Moraux
Université de Rennes 1 and CNRS

June 2009

*Part of this paper was presented at the AFFI Spring conference organized at Brest in May 2009. We thank participants for comments and discussion. Leblon and Moraux acknowledge financial support from IAE Rennes and CREM (CNRS 6211). The usual disclaimer applies.

†Corresponding author. Postal address: Université de Rennes 1, IAE de Rennes, 11 rue Jean Macé 35000 Rennes. Mail to: gregoire.leblon@univ-rennes1.fr.

Quadratic Term Structure Models: Analysis and Performance

Abstract: The family of the Affine Term Structure of interest rate has been a lot developed in the literature since the first work of Vasicek (1977) and Cox, Ingersoll and Ross (1985b). Although their performances increase, they are still facing several difficulties in their capacity to fully explain the behaviour of the Term Structure of interest rate. Some of these issues are explain by the omission of non linear relation in the affine model (Dai and Singleton, 1999). This paper is in the continuity of this reflexion. It presents, develops, applies and discuses the quadratic model both in discrete and continuous time.

JEL Codes: G10, G11.

1 Introduction

Most of the papers about the Term Structure Modelling (TSM) are relative to the family of the Affine Term Structure Models (ATSM). This family of models considers a linear relation between the log price of a bond and its states factors. Those models have been first developed by Vasicek (1977) and Cox, Ingersoll and Ross (1985b). Later, Duffie & Kan (1996) clarified the primitive assumptions underlying this framework. Since the first models, as noted by Dai & Singleton (1999), the ATSM have increased in performance but are still facing two main issues which suggest to looking for a new family of models. The first issue is that to be admissible, an ATSM needs non zero conditional correlation of its states variables. This condition is incompatible with certain structure of the bond price volatility especially the ones which do not allow negative nominal interest rate. Then, one needs to do a trade off between those two objectives. Secondly, the form of the pricing error for ATSM suggests that non linearity is omitted in this family of models.

Regarding the development of the ATSM, the others have been less developed as the Quadratic Term Structure Model (QTSM). This family, first introduce by Beaglehole & Tammey (1991) and Constantinides (1992) are now more developed in the literature especially because of the issues encountered with the ATSM. Furthermore, they are now also applied to the pricing of contingent claims (Lieppold & Wu (2002, 2003)) and to the credit risk pricing (Chen, Filipovic and Poor (2004)).

The main model analysed in this paper is in discrete time and belongs to the QTSMs family. It has been theoretically presented by Realdon (2006) and is derived from the continuous QTSM of Dai-Le-Singleton (2005). Regarding the continuous time, the discrete time allows more flexibility in the specification of the market price of risk as mention by Dai-Le-Singleton (2005). This property remains as long as the factors transition density remains Gaussian. Furthermore, as noted by Realdon (2006), when the discrete time steps converge to zero, a discrete time model converges to a continuous one. Then, the class of the models in continuous time may be seen as a particular case of the discrete one.

This paper is organized as follows. In Section 1, the model is presented and defined as it is in the paper of Realdon. In Section 2, several properties of the model are exposed in order to make its utilization and calibration easier. In section 3, the calibration model is presented and is applied to the US treasury rate in section 4.

2 The traditional framework

Quadratics models have been firstly investigated by Beaglehole & Tammey (1991) and Constantinides (1992). They combine positive rates with some tractability. Although not a lot of attention has been paid to then at the beginning, they are now more and more developed in order to solve the issues encountered with the ATSM. QTSM assumes that the instantaneous spot rate r_t is the sum of square of quadratic state variables $X_t = (X_{1,t}, X_{2,t}, ..., X_{n,t})$. Formally, the model is:

$$r_t = \alpha + \beta' X_t + X_t^{'} \gamma X_t \tag{1}$$

With α and β two $N \times 1$ vectors, γ a $N \times N$ matrix and X_t a n-dimensional state variable which is supposed to follow a diffusion process under the risk neutral probability:

$$dX_t = f(X_t)dt + \rho(X_t)dW_t \tag{2}$$

With $f(X_t)$ the drift ($N \times 1$ vector), $\rho(X_t)$ the diffusion parameter ($N \times N$ matrix) and dW_t a Wiener process in $\mathbb{R}^n$under the risk neutral probability. Regarding (1), it appears that the affine model is a particular case of the QTSM where γ is the null matrix.

The associate log bond price is also assumes to be a quadratic form of the state variables as:

$$P(t,t+\tau) = e^{-\int_{s=t}^{t+\tau} r_s ds} \quad (3)$$

$$= e^{A(\tau)+B'(\tau)X_t+X_t'C(\tau)X_t} \quad (4)$$

With $P(t,t+\tau)$ the price at t of a bond with a time to maturity equals to τ, $A(\tau)$ and $B(\tau)$ two $N \times 1$ vectors only depending of the time to maturity and $C(\tau)$ a $N \times N$ matrix also depending of τ. If $C(\tau)$ is null for all the time to maturity, the price formula is then the one of a traditional Affine model.

3 The single factor in discrete time

The quadratic model presented in this section is in discrete time. QTSM have been nearly uniquely developed in a continuous time setting or the discrete time offers, considering ATSM or QTSM, more flexibility in the definition of the market price of risk (Realdon (2006), Dai, Le, Singleton (2005)). This characteristic is important at the estimation step. Furthermore, discrete models are more suitable to macroeconomic variable which are discrete in their availability.

A single factor model, based on the one presented by Realdon (2006) is presented in this second section as it allows a easy understanding of the QTSM's behaviour.

3.1 The assumptions of the model

In the single factor model, the factor is assumed to follow a diffusion process, the short rate to be quadratic and the price of a bond $P(t,n)$ in t and of time to maturity $n\Delta t$ to be as followed:

$$
\begin{aligned}
r_t &= \alpha + \beta . x_t + \gamma . x_t^2 & (5)\\
\Delta x_t &= \kappa \left(\theta - x_t\right) \Delta t + \sigma_0 \left(\sigma_1 + x_t^\delta\right) \sqrt{\Delta t} \varepsilon_{t+1} & (6)\\
& \varepsilon_t \tilde{} N(0,1) & (7)\\
P(t,n) &= e^{A_n + B_n . x_t + C_n . x_t^2} & (8)
\end{aligned}
$$

With α, β, θ, κ and γ five constants, x_t the underlying factor of the model and ξ_t the noise term which is assumed to be normally distributed with a mean of 0 and a variance of 1.

Several first constraints must be defined here in order to have non negative short rate. Then, according to Ahn-Dittmar-Gallant (2002), the parameter must verify:

$$
\begin{aligned}
\beta &= 0 & (9)\\
\alpha &> \frac{\beta^2}{4\gamma} = 0 & (10)\\
\gamma &> 0 & (11)
\end{aligned}
$$

These constraints are necessary in order to have the parameter θ identifiable. Then, the instantaneous spot rate is a pure quadratic model and (5) can now be rewritten as:

$$
\begin{aligned}
r_t &= \alpha + \gamma . x_t^2 & (12)\\
\Delta x_t &= \kappa \left(\theta - x_t\right) \Delta t + \sigma_0 \left(\sigma_1 + x_t^\delta\right) \sqrt{\Delta t} \varepsilon_{t+1} \\
& \varepsilon_t \tilde{} N(0,1) \\
P(t,n) &= e^{A_n + B_n . x_t + C_n . x_t^2}
\end{aligned}
$$

With this new formulation, it appears that the parameter α is in fact the floor of the instantaneous spot rate. It is then easy to see that r_t is always strictly positive and at least equalled to α.

3.2 The associated bond price

It clearly appears, from the formula of the bond price, that the parameters A_n, B_n and C_n will be, in a first step, defined recursively. Let's $P(t,n)$ be the price at t of a bond of maturity n and E_t the conditional expectation at time t under the risk neutral measure. Then it appears that:

$$P(t,n) = e^{A_n + B_n . x_t + C_n . x_t^2} \tag{13}$$

And:

$$\begin{aligned} P(t,n) &= E_t \left[e^{-\sum_{i=t}^{t+n-1} r_i \Delta t} \right] && (14) \\ &= e^{-r_t \Delta t} . E_t \left[P(t+\Delta t, n-1) \right] && (15) \end{aligned}$$

This implies a first result with a recursive solution:

$$A_n + B_n x_t + C_n x_t^2 = -r_t \Delta t + \ln \left(E_t \left[e^{A_{n-1} + B_{n-1} x_{t+\Delta t} + C_{n-1} x_{t+\Delta t}^2} \right] \right) \tag{16}$$

At this step, it appears that several constraints have to be made in order to be able to closely identify the parameters A_n, B_n and C_n. Then, we must have:

$$\delta = 0 \tag{17}$$

The model is now a traditional Vasicek model because the diffusion process must be an affine function of the underlying factor. This result has already been shown in the continuous time by Leippold & Wu (2003). The model is finally:

$$
\begin{aligned}
r_t &= \alpha + \gamma . x_t^2 \\
\Delta x_t &= \kappa \left(\theta - x_t\right) \Delta t + \sigma \sqrt{\Delta t} \varepsilon_{t+1} \\
& \quad \varepsilon_t \tilde{} N(0,1) \\
P(t,n) &= e^{A_n + B_n . x_t + C_n . x_t^2}
\end{aligned} \tag{18}
$$

The complete resolution and the proof of the necessity to have an homoscedastic underlying factor may be seen in (7). The recursive solution of the parameters A_n, B_n and C_n is:

$$
\begin{aligned}
A_n &= -\alpha \Delta t + A_{n-1} + B_{n-1} \kappa \theta \Delta t + C_{n-1} \left(\kappa \theta \Delta t\right)^2 + \\
& \quad \frac{\sigma^2 \Delta t \left(B_{n-1} + 2 C_{n-1} \kappa \theta \Delta t\right)^2}{2\left(1 - 2\sigma^2 \Delta t C_{n-1}\right)} - \frac{\ln\left(1 - 2\sigma^2 \Delta t C_{n-1}\right)}{2}
\end{aligned} \tag{19}
$$

$$
B_n = \left(1 - \kappa \Delta t\right) \frac{2\kappa \theta \Delta t C_{n-1} + B_{n-1}}{1 - 2\sigma^2 \Delta t C_{n-1}} \tag{20}
$$

$$
C_n = -\gamma \Delta t + C_{n-1} \frac{\left(1 - \kappa \Delta t\right)^2}{1 - 2\sigma^2 \Delta t C_{n-1}} \tag{21}
$$

And the corresponding interest rate $R(t,n)$ of a bond valuing at time t and of time to maturity $n\Delta t$ is given by:

$$
R(t,n) = \frac{-A_n - B_n x_t - C_n x_t^2}{n \Delta t} \tag{22}
$$

4 A new formulation of the single discrete model

The recursive formulation of the model has the goodness to be easy to obtain and to implement. However, it is very costly in term of computation time and may imply an approximation error: at each step, a value is calculated and used for the next estimation.

Or these values are rounded and although these approximations have a negligible impact for the first ranks, it may be no more the case for huge maturities.

The aim of this part is then to find an indicial formulation of the parameters (for C_n and B_n) or, at least, an expression using matrix calculation (for A_n). All the demonstrations are reported in (7).

4.1 A new formulation for C_n

C_n can be then transformed into an indicial form by doing some matrix calculation. Then, its new formula is given by:

$\forall n \in \mathbb{N}^*$,

$$C_n = \gamma \Delta t \frac{\lambda_1^n - \lambda_2^n}{\lambda_2^n (1 - \lambda_1) - \lambda_1^n (1 - \lambda_2)} \tag{23}$$

With:

$$\lambda_1 = \frac{2\gamma\sigma^2\Delta t^2 + (1 - \kappa\Delta t)^2 + 1 - \Delta t\sqrt{(2\gamma\sigma^2 + \kappa^2)(2\gamma\sigma^2\Delta t^2 + (2 - \kappa\Delta t)^2)}}{2} \tag{24}$$

$$\lambda_2 = \frac{2\gamma\sigma^2\Delta t^2 + (1 - \kappa\Delta t)^2 + 1 + \Delta t\sqrt{(2\gamma\sigma^2 + \kappa^2)(2\gamma\sigma^2\Delta t^2 + (2 - \kappa\Delta t)^2)}}{2} \tag{25}$$

It is then possible to calculate this parameter for any maturity without calculating all the previous ones which is a real gain in performance during the calibration.

4.2 A new formulation for B_n

B_n can also be rewritten with a non recursive formula:

$\forall n \in \mathbb{N} \setminus \{0; 1\}$

$$B_n = \frac{2\kappa\theta\Delta t}{1-\kappa\Delta t}\left(C_n + \gamma\Delta t\right)\left[\begin{array}{c} 1 \\ +\frac{(\lambda_1\lambda_2)^{n-1}}{\left(\lambda_1^{n-1}-\lambda_2^{n-1}\right)(1-\kappa\Delta t)^{n-2}}\sum_{i=1}^{2}\left((-1)^i\frac{1-\left(\frac{1-\kappa\Delta t}{\lambda_i}\right)^{n-2}}{\lambda_i-1+\kappa\Delta t}\right)\end{array}\right] \tag{26}$$

And with:

$$B_0 = B_1 = 0 \tag{27}$$

As we can see, the indicial formulation of B_n is more complex than the C_n's one. This is due to the fact that B_n is a recursive factor that depends of both the last values of B_n and C_n.

4.3 A new formulation for A_n

This parameter cannot be easily expressed as it has been done on B_n and C_n. However, it is still possible to over perform the computation by using a matrix formulation. Then:

$$A = M \times \Phi \tag{28a}$$

With:

$$A = (A_0, A_1, ..., A_n)' \tag{29}$$

$$M = (a_{i,j})_{i\in<1;n>, j\in<1,n>} \tag{30}$$

$$a_{i,j} = 1 \text{ if } j < i \tag{31}$$

$$a_{i,j} = 0 \text{ if } j \geq i \tag{32}$$

$$\Phi = (\Phi_0, \Phi_1, ..., \Phi_n)' \tag{33}$$

And $\forall i \in< 0; n >$,

$$\Phi_i = -\alpha\Delta t + B_i\kappa\theta\Delta t + C_i\left(\kappa\theta\Delta t\right)^2 + \frac{\sigma^2\Delta t\left(B_i + 2C_i\kappa\theta\Delta t\right)^2}{2\left(1 - 2\sigma^2\Delta t C_i\right)} - \frac{\ln\left(1 - 2\sigma^2\Delta t C_i\right)}{2} \tag{34}$$

These three new formulations are used in the next parts to have a faster calibration of the model.

4.4 Properties of parameters A_n, B_n and C_n

Although it is not reported in Realdon's paper, several useful properties and characteristics may be investigate. All the results of this subsection are detailed in (7).

First, we consider the parameter C_n which is the sensibility of interest rate to the square of the state variable (22). C_n is always negative, strictly inferior to $-\gamma$ when n is strictly superior to 0 and null for the null value of n. Furthermore, C_n is strictly decreasing and converge to the value l_C:

$$l_C = \frac{\left(-2\sigma^2\gamma\Delta t - 2\kappa + \kappa^2\Delta t\right) - \sqrt{\left(-2\sigma^2\gamma\Delta t - 2\kappa + \kappa^2\Delta t\right)^2 + 8\sigma^2\gamma\Delta t}}{4\sigma^2} \tag{35}$$

Secondly, the parameter B_n is strictly negative for every n superior to 1 and null for all the other values of n. As C_n, B_n converges to l_B:

$$l_B = \frac{2\kappa\theta l_C(1 - \kappa\Delta t)}{\kappa - 2\sigma^2 l_C} \tag{36}$$

5 Model Implementation

5.1 A new method to calibrate the model

Because the instantaneous interest rate is not observable, we have to exploit the dynamics of the spot interest rate. The model is calibrated on the evolution of an interest rate with a maturity equals to $n_1\Delta t$. Using (22), the evolution of this interest rate is given by:

$$\Delta R(t,n_1) = R(t+\Delta t,n_1) - R(t,n_1) \tag{37}$$
$$= \frac{B_{n_1}\left(x_t - x_{t+\Delta t}\right) + C_{n_1}\left(x_t^2 - \left(x_{t+\Delta t}\right)^2\right)}{n_1 \Delta t} \tag{38}$$

Note that, using (6), the previous formula can be rewritten as:

$$\Delta R(t,n_1) = \frac{-B_{n_1}\left(\Delta x_t\right) - C_{n_1}\left(\left(\Delta x_t\right)^2 + 2x_t \Delta x_t\right)}{n_1 \Delta t} \tag{39}$$

The state variable x_t is here not observable. Indeed, equation (22) displays an under-identification problem. In other words, it is not possible to find the value of x_t if only one maturity is considered. To solve this issue, a solution is to consider a second spot interest rate with another time to maturity. With $n_2 \Delta \Delta t$ this second maturity, we have:

$$R(t,n_1) = \frac{-A_{n_1} - B_{n_1} x_t - C_{n_1} x_t^2}{n_1 \Delta t} \tag{40}$$
$$R(t,n_2) = \frac{-A_{n_2} - B_{n_2} x_t - C_{n_2} x_t^2}{n_2 \Delta t} \tag{41}$$

The value of the state variable at each time t is now identify with the formula (see (7)):

$$x_t = \frac{C_{n_2}\left(R(t,n_1) n_1 \Delta t + A_{n_1}\right) - C_{n_1}\left(R(t,n_2) n_2 \Delta t + A_{n_2}\right)}{C_{n_1} B_{n_2} - B_{n_1} C_{n_2}} \tag{42}$$

5.2 The GMM approach

The model is calibrated using the General Method of Moments. To avoid under-estimation, at least eight moments have to be computed. The moments used for with the GMM approach are (detailed in 7):

$$f_{1,j} = \frac{1}{N}\sum_{t=1}^{N}\left\{\Delta R(t,n_j)n_j\Delta t - \left(-B_{n_j} * M_1 - C_{n_j}\left(2x_tM_1 + M_2\right)\right)\right\} \tag{43}$$

$$f_{2,j} = \frac{1}{N}\sum_{t=1}^{N}\left\{\begin{array}{c}(\Delta R(t,n_j)n_j\Delta t)^2 - \\ ((B_{n_j} + 2x_tC_{n_j})^2M_2 + C_{n_j}^2M_4 + 2(B_{n_j} + 2 * x_tC_{n_j})C_{n_j}M_3)\end{array}\right\} \tag{44}$$

$$f_3 = \frac{1}{N}\sum_{t=1}^{N}\left\{x_{t+\Delta t} - x_t - M_1\right\} \tag{45}$$

$$f_4 = \frac{1}{N}\sum_{t=1}^{N}\left\{\left(x_{t+\Delta t} - x_t\right)^2 - M_2 + M_1^2\right\} \tag{46}$$

$$f_{5,j} = \frac{1}{N}\sum_{t=1}^{N}\left\{R(t,n_j)n_j\Delta t + A_{n_j} + B_{n_j}x_t + C_{n_j}x_t^2\right\} \tag{47}$$

$$f_{6,j} = \frac{1}{N}\sum_{t=1}^{N}\left\{\left(R(t,n_j)n_j\Delta t + A_{n_j} + B_{n_j}x_t + C_{n_j}x_t^2\right)^2\right\} \tag{48}$$

With $j \in \{1,2\}$ and where M_i is the moment of order i of Δx_t $\left(E_t\left[(\Delta x_t)^i\right]\right)$:

$$M_1 = \kappa\left(\theta - x_t\right)dt \tag{49}$$

$$M_2 = M_1^2 + \sigma^2\Delta t \tag{50}$$

$$M_3 = M_1^3 + 3M_1\left(M_2 - M_1^2\right) \tag{51}$$

$$M_4 = M_1^4 + 6M_1^2\left(M_2 - M_1^2\right) + 3\left(M_2 - M_1^2\right)^2 \tag{52}$$

The value to minimize is then the sum of the square of the moments. The choice of these moments is motivated by the wish to underline several characteristics on which we want the model to be calibrated. The first and the second moments (f_1, f_2) ensure that the evolution of the predicted spot rate on one period has the same first and second moment as the observed ones. The third and the forth moments (f_3, f_4) are used to confirm the assumptions of a Gaussian state variable guarantying an estimated process which has the good characteristics. Finally, the two last moments ensure that the estimated spot rate is closed enough to the observed one. It has also to be noted that the moment 1, 2, 5 and 6 are computed on both the first maturity and the second one. It is done due to the fact that the state variable is derived from the dynamic of these two spot rates. Then, to have consistent results, it is necessary to consider those moments on both.

6 Empirical analysis

The calibration is done on the weekly US treasury rate available from the 03/08/2001 to the 30/01/2009 and using the method describe in the previous section. The data are from the FED website and contain four different maturities: four weeks, three months, six months and one year. The first maturity ($n_1\Delta t$) (notation used in 5) is associated to the shorter available interest rate: the one month spot rate. The second maturity ($n_2\Delta t$) is associated to the second shorter available interest rate: the three months spot rate. The calibration is performed 500 times, each time with a different set of initials values to be less sensitive to the initial state.

6.1 Parameter estimation

The estimated parameters are reported in Table 1 (p34). The estimated α is, as it should be, very small and equalled to 0.0212. This value is logical regarding its interpretation as a floor value for the instantaneous interest rate. Among the five parameters, two are well estimated: α and κ with an estimated value which change only marginally when the initials values change. The others parameters are much instable especially gamma. However, the estimated values are the one providing the best results.

6.2 Assessing the performance of the model

Figures 1 and 2 (p**??**, p**??**) present some visual results of the calibration. On top of the figure 1 (p**??**), the weekly evolution of the one month and three months interest rate is represented. It appears that the model fits well the original data on most of the time period for both the three and one month time to maturity. The two graphs on the bottom of figure 1 represents the observed spot rate versus the estimated one. The blue line represents the area where the green dots should be. Regarding the graphs, it is acceptable to say that the model predicts well the two interest rate. However, on the graphs at the bottom of Figure 1, the quadratic form appears when we are plotting the observed interest rate versus the predicted one. Normally, the dot should be around the blue line if the model were correct in its assumptions.

The Figure 2 (p**??**) represents the result on the residuals. The two tops graphs illustrate the residuals across the time and the two graphs on the bottom their distribution. Those results introduce the fact that the error of prediction is bias for the one month interest rate but is nearly correct for the three months one. Furthermore, the distribution of the error fits well a normal distribution.

As a last indicator of performances, the Table 2 (p35) gives some statistical result based on the model. Those results are product on the build sample (the one month and three months interest rate) and on the output of sample (the six months and one year interest rate). From the top to the bottom, the figure 4 gives information about the number of observation, the correlation between the estimated and observed interest rate, the correlation between the error of prediction and the observed interest rate, the mean of the observed IR, the mean error of prediction and its standard deviation. Globally, the correlations are goods (all above 92% except for the one month IR) and significant considering a 5% level: the model explains well the behaviour of the interest rate. Regarding the correlation between the error and the observed IR, it appears that, for short time to maturity, only a few part of the error is not explained by the model. For the others, the model does not permit to fully explain the behaviour: the model lost of its predictive power when the maturity increases. Furthermore, the one year maturity has also a low mean interest rate which gives a mean error equals in absolute value to the mean of the interest rate itself. That also underlines an issue of this kind of model: if it calibrated when the interest rate level is high, the value of alpha will be also high which may implies issue when a strong decrease is observed.

7 Conclusion

The single factor QTSM in discrete time has been applied here to the US treasury bills rate. The model is in itself relatively simple to use and gives a good explication of the behaviour of the interest rate. However, the performance shows some issues. The first one is linked to the representation of the observed interest rate versus the predicted one . This Figure (p**??**) shows a quadratic form which should be corrected by the model. An explication could be that the model does not take into account enough state variables. It is true that generally, a good model should have at least two or three states variables. However, the discrete model presented here does not allow a simple resolution in a multivariate context or this resolution is so costly in term of time of computation that it makes it hard to use.

The two others issues encountered by the model are first the incapacity to take into account a heteroscedastic state factor and secondly the cost to do not allow negative value for the interest rate which is to do not allow too small values. The issue encountered with a threshold is always when we are closed to it.

References

[1] Ahn D.H., R. Dittmar, A. Gallant (2002): "Quadratic Term Structure Models: Theory and Evidence", *Review of Financial Studies* 15(1), 243-288.

[2] Beaglehole D., M. Tenney (1991): "General Solutions of Some Interest Rate-Contingent Claim Pricing Equations", *Journal of Fixed Income* 1, 69-83.

[3] Beaglehole D., M. Tenney (1992): "A nonlinear equilibrium model of term structures of interest rates: corrections and additions", *Journal of Financial Economics,* 32, 345-454.

[4] Chen L., D. Filipovic and H.V. Poor (2004): "Quadratic Term Structure Models for risk-free and defaultable rates", *Mathematical Finance*, 14(4), 515-536.

[5] Constantinides G. (1992): "A Theory of the Nominal Term Structure of Interest Rates", *Review of Financial Studies* 5, 531-552.

[6] Cox, Ingersoll, Ross, 1985b, "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, 385-406.

[7] Dai Q., K. Singleton (2000): "Specification Analysis of Affine Term Structure Models", *Journal of Finance*, 55, 1943-1978.

[8] Dai Q., K. Singleton (2003): "Term structure dynamics in theory and reality", *Review of Financial Studies*, 16, 631-678.

[9] Dai Q., A. Le and K. Singleton (2005): "Discrete-time Dynamic Term Structure Models with generalised Market Prices of Risk", *Working Paper.*

[10] Duffie D., R. Kan (1996): "A Yield-Factor Model of Interest Rates", *Mathematical Finance* 6, 379 – 406

[11] Leippold M., L. Wu (2002): "Asset Pricing under the quadratic Class", *Journal of Financial and Quantitative Analysis* 37(2), 271-294.

[12] Leippold M., L. Wu (2003): "Design and Estimation of Quadratic Term Structure Models", *European Finance Review* 7, 47-73.

[13] Realdon M. (2006): "Quadratic Term Structure Models in Discrete Time", *Finance Research Letters*, 3(4), 277-289.

[14] Vasicek O. (1977): "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics* 5, $177-178$

Appendix A: The recursive solution

First, in order to have an easier reading of the results, let's consider:

$$\begin{aligned}
\alpha &= \alpha \Delta t \\
\gamma &= \gamma \Delta t \\
\kappa &= \kappa \Delta t \\
\sigma_0 &= \sigma_0 \Delta t
\end{aligned}$$

From (16), and by using (6), we have then:

$$\begin{aligned}
\ln\left(P_{t,n}\right) &= -r_t\Delta t + \ln\left(E_t\left[e^{A_{n-1}+B_{n-1}(\kappa\theta+x_t(1-\kappa)+\varepsilon_{t+1})+C_{n-1}(\kappa\theta+x_t(1-\kappa)+\varepsilon_{t+1})^2}\right]\right) && (53)\\
&= -r_t\Delta t + A_{n-1} + B_{n-1}\left(\kappa\theta + x_t\left(1-\kappa\right)\right) + C_{n-1}\left(\kappa\theta + x_t\left(1-\kappa\right)\right)^2 && (54)\\
&\quad + \ln\left(E_t\left[e^{B_{n-1}\varepsilon_{t+1}+C_{n-1}\left(\varepsilon_{t+1}^2+2(\kappa\theta+x_t(1-\kappa))\varepsilon_{t+1}\right)}\right]\right) \\
&= -\alpha + A_{n-1} + B_{n-1}\kappa\theta + C_{n-1}\left(\kappa\theta\right)^2 && (55)\\
&\quad + x_t\left[-\beta + B_{n-1}\left(1-\kappa\right) + 2C_{n-1}\kappa\theta\left(1-\kappa\right)\right] \\
&\quad + x_t^2\left[-\gamma + C_{n-1}\left(1-\kappa\right)^2\right] \\
&\quad + \ln\left(E_t\left[e^{a\zeta_{t+1}+b\zeta_{t+1}^2}\right]\right)
\end{aligned}$$

With:

$$a = \sigma_0\left(\sigma_1 + x_t^{\delta}\right)\left(B_{n-1} + 2C_{n-1}\left(\kappa\theta + x_t\left(1-\kappa\right)\right)\right) \quad (56)$$

$$b = C_{n-1}\sigma_0^2\left(\sigma_1 + x_t^{\delta}\right)^2 \quad (57)$$

The term $\ln\left(E_t\left[e^{a\varepsilon_{t+1}+b\varepsilon_{t+1}^2}\right]\right)$ is easily find by using the relation:

$$E_t\left[e^{a\varepsilon_{t+1}+b\varepsilon_{t+1}^2}\right] = \int e^{a\varepsilon_{t+1}+b\varepsilon_{t+1}^2} f(\varepsilon_{t+1}) d\varepsilon_{t+1} \quad (58)$$

With $f(\varepsilon_{t+1})$ the density function of ε_{t+1} which follows a normal distribution of mean 0 and with a standard deviation equals to 1. Then:

$$f(\varepsilon_{t+1}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\varepsilon_{t+1}^2}{2}} \quad (59)$$

Using (59) in (58), we have:

$$E_t\left[e^{a\varepsilon_{t+1}+b\varepsilon_{t+1}^2}\right]=\frac{1}{\sqrt{2\pi}}\int e^{a\varepsilon_{t+1}+b\varepsilon_{t+1}^2-\frac{\varepsilon_{t+1}^2}{2}}d\varepsilon_{t+1} \tag{60}$$

Or:

$$\begin{aligned} a\varepsilon_{t+1}+\varepsilon\zeta_{t+1}^2-\frac{\varepsilon_{t+1}^2}{2\sigma^2} &= -\frac{1}{2}\left(-2a\varepsilon_{t+1}-2b\varepsilon_{t+1}^2+2\varepsilon_{t+1}^2\right) && (61)\\ &= -\frac{1}{2}\left(\varepsilon_{t+1}\left(1-2b\right)^{\frac{1}{2}}-a\left(1-2b\right)^{-\frac{1}{2}}\right)^2+\frac{a}{2}\left(1-2b\right)^{-1} && (62)\\ &= -\frac{1}{2}\left(\varepsilon_{t+1}\Gamma^{-1}-a\Gamma\right)^2+\frac{a^2\Gamma^2}{2} && (63)\end{aligned}$$

With:

$$\Gamma=\left(1-2b\right)^{-\frac{1}{2}} \tag{64}$$

From (61), the term $1-2b$ has to be strictly positive i.e. $1-2\sigma_0^2\left(\sigma_1+x_t^\delta\right)^2C_n$ must be strictly positive for each value of n. This constraint is in fact always matched and will be developed in (7). Then, (60) becomes:

$$\begin{aligned} E_t\left[e^{a\varepsilon_{t+1}+b\varepsilon_{t+1}^2}\right] &= \frac{1}{\sqrt{2\pi}}\int e^{-\frac{1}{2}\left(\varepsilon_{t+1}\Gamma^{-1}-a\Gamma\right)^2+\frac{a^2\Gamma^2}{2}}d\varepsilon_{t+1} && (65)\\ &= \Gamma e^{\frac{a^2\Gamma^2}{2}}\frac{1}{\Gamma\sqrt{2\pi}}\int e^{-\frac{1}{2}\left(\frac{\varepsilon_{t+1}\Gamma-a\Gamma^2}{\Gamma}\right)^2}d\varepsilon_{t+1} && (66)\\ &= \Gamma e^{\frac{a^2\Gamma^2}{2}} && (67)\\ &= \left(1-2C_{n-1}\sigma_0^2\left(\sigma_1+x_t^\delta\right)^2\right)^{-\frac{1}{2}} && (68)\\ &\quad \times e^{\frac{\left(B_{n-1}+2C_{n-1}\left(\kappa\theta+x_t(1-\kappa)\right)\right)^2\left(1-2C_{n-1}\sigma_0^2\left(\sigma_1+x_t^\delta\right)^2\right)^{-1}}{2}} && (69)\end{aligned}$$

We then obtain a new formulation for (55):

$$\begin{aligned} \ln\left(P_{t,n}\right) &= -\alpha+A_{n-1}+B_{n-1}\kappa\theta+C_{n-1}\left(\kappa\theta\right)^2 && (70)\\ &\quad +x_t\left[-\beta+B_{n-1}\left(1-\kappa\right)+2C_{n-1}\kappa\theta\left(1-\kappa\right)\right]\\ &\quad +x_t^2\left[-\gamma+C_{n-1}\left(1-\kappa\right)^2\right]\\ &\quad +\ln\left(\begin{array}{c}\left(1-2C_{n-1}\sigma_0^2\left(\sigma_1+x_t^\delta\right)^2\right)^{-\frac{1}{2}}\\ \times e^{\frac{\left(B_{n-1}+2C_{n-1}\left(\kappa\theta+x_t(1-\kappa)\right)\right)^2\left(1-2C_{n-1}\sigma_0^2\left(\sigma_1+x_t^\delta\right)^2\right)^{-1}}{2}}\end{array}\right)\end{aligned}$$

It is then clear that at this step, a diffusion parameter with a δ different from 0 will introduce a $\ln(P_{t,n})$ with a non linear relation of the state variable. Because this relation does not match the model's assumptions, we must have δ equals to 0 which means that σ_1 is equals to 1. Let's now write σ_0 as σ in order to simplify the notation. Then, we have:

$$\ln(P_{t,n}) = \begin{bmatrix} -\alpha + A_{n-1} + B_{n-1}\kappa\theta + C_{n-1}(\kappa\theta)^2 \\ +\frac{\sigma^2(B_{n-1}+2C_{n-1}\kappa\theta)^2}{2(1-2\sigma^2 C_{n-1})} - \frac{\ln(1-2\sigma^2 C_{n-1})}{2} \end{bmatrix} \tag{71}$$

$$+x_t \begin{bmatrix} -\beta + B_{n-1}(1-\kappa) + 2C_{n-1}\kappa\theta(1-\kappa) \\ +\frac{4\sigma^2 C_{n-1}(1-\kappa)(B_{n-1}+2\kappa\theta C_{n-1})}{2(1-2\sigma^2 C_{n-1})} \end{bmatrix} \tag{72}$$

$$+x_t^2 \left[-\gamma + C_{n-1}(1-\kappa)^2 + \frac{4\sigma^2 C_{n-1}^2 (1-\kappa)^2}{2(1-2\sigma^2 C_{n-1})} \right] \tag{73}$$

By identification, the Realdon's recursive resolution appears:

$$A_n = -\alpha + A_{n-1} + B_{n-1}\kappa\theta + C_{n-1}(\kappa\theta)^2 \tag{74}$$

$$+\frac{\sigma^2(B_{n-1}+2C_{n-1}\kappa\theta)^2}{2(1-2\sigma^2 C_{n-1})} - \frac{\ln(1-2\sigma^2 C_{n-1})}{2} \tag{75}$$

$$B_n = -\beta + B_{n-1}(1-\kappa) + 2C_{n-1}\kappa\theta(1-\kappa) + \frac{4\sigma^2 C_{n-1}(1-\kappa)(B_{n-1}+2\kappa\theta C_{n-1})}{2(1-2\sigma^2 C_{n-1})} \tag{76}$$

$$C_n = -\gamma + C_{n-1}(1-\kappa)^2 + \frac{4\sigma^2 C_{n-1}^2(1-\kappa)^2}{2(1-2\sigma^2 C_{n-1})} \tag{77}$$

However, this representation is not the simplest one. After simplification, and after having reintroduced the term in Δt, we have (19), (20) and (21):

$$\begin{aligned} A_n &= -\alpha\Delta t + A_{n-1} + B_{n-1}\kappa\theta\Delta t + C_{n-1}(\kappa\theta\Delta t)^2 \\ &\quad +\frac{\sigma^2\Delta t(B_{n-1}+2C_{n-1}\kappa\theta\Delta t)^2}{2(1-2\sigma^2\Delta t C_{n-1})} - \frac{\ln(1-2\sigma^2\Delta t C_{n-1})}{2} \\ B_n &= -\beta\Delta t + (1-\kappa\Delta t)\frac{2\kappa\theta\Delta t C_{n-1} + B_{n-1}}{1-2\sigma^2\Delta t C_{n-1}} \\ C_n &= -\gamma\Delta t + C_{n-1}\frac{(1-\kappa\Delta t)^2}{1-2\sigma^2\Delta t C_{n-1}} \end{aligned}$$

Appendix B: Properties of A_n, B_n and C_n

Properties of C_n

- **The negative value of C_n**

  It is an immediate result which can be proved by recurrence. Then, as soon as C_n becomes negative, the values associated to higher maturity are all negatives. Because the first value is zero and the second one strictly negative (11), C_n is always strictly negative except for the maturity equals to zero where its value is zero.

- **The decreasing of C_n**

  This point is easily shown by recurrence: for $n = 1$, it is clear that $C_1 < C_0$. From (21), if there is rank such as $C_n < C_{n-1}$, then:

$$
\begin{aligned}
C_{n+1} - C_n &= -\gamma\Delta t + C_n \frac{(1-\kappa\Delta t)^2}{1-2\sigma^2\Delta t C_n} + \gamma\Delta t - C_{n-1}\frac{(1-\kappa\Delta t)^2}{1-2\sigma^2\Delta t C_{n-1}} && (78)\\
&= (1-\kappa\Delta t)^2\left(\frac{C_n}{1-2\sigma^2\Delta t C_n} - \frac{C_{n-1}}{1-2\sigma^2\Delta t C_{n-1}}\right) && (79)\\
&= (1-\kappa\Delta t)^2\left(\frac{C_n - C_{n-1}}{(1-2\sigma^2\Delta t C_n)(1-2\sigma^2\Delta t C_{n-1})}\right) && (80)
\end{aligned}
$$

  Or $\forall n \in \mathbb{N}, 1 - 2\sigma^2\Delta t C_n > 0$ (previous point) and $C_n < C_{n-1}$, then $C_{n+1} < C_n$. Finally, the following property is shown:

$$\forall n \in \mathbb{N}^*, C_n - C_{n-1} < 0 \tag{81}$$

  C_n is strictly decreasing.

- **The convergence of C_n**

  If C_n has no floor, then:

$$\lim(C_n)_{n\to\infty} = -\infty = C_\infty \tag{82}$$

  Which means:

$$
\begin{aligned}
C_{\infty+1} &= -\gamma\Delta t + C_\infty \frac{(1-\kappa\Delta t)^2}{1-2\sigma^2\Delta t C_\infty}\\
&\tilde{} - \gamma\Delta t + \frac{(1-\kappa\Delta t)^2}{-2\sigma^2\Delta t} < C_\infty && (83)
\end{aligned}
$$

Which is not compatible with the decreasing of C_n. C_n is floored and decreasing: it converges to a finite value.

- **The limit of C_n**

  Let l_C be the limit of C_n, then from (21):

$$l_C = -\gamma\Delta t + l_C \frac{(1-\kappa\Delta t)^2}{1-2\sigma^2\Delta t l_C} \tag{84}$$

  Which means:

$$-2\sigma^2\Delta t l^2 + l\left(1-2\sigma^2\gamma\Delta t^2 - (1-\kappa\Delta t)^2\right) + \gamma\Delta t = 0 \tag{85}$$

  Or the following result implied two possible values for the limit:

$$\Delta = \left(1-2\sigma^2\gamma\Delta t^2 - (1-\kappa\Delta t)^2\right)^2 + 8\sigma^2\gamma\Delta t^2 > 0 \tag{86}$$

  Which are:

$$x_1 = \frac{\left(1-2\sigma^2\gamma\Delta t^2 - (1-\kappa\Delta t)^2\right) + \sqrt{\left(1-2\sigma^2\gamma\Delta t^2 - (1-\kappa\Delta t)^2\right)^2 + 8\sigma^2\gamma\Delta t^2}}{4\sigma^2\Delta t} \geq 0 \tag{87}$$

$$x_2 = \frac{\left(1-2\sigma^2\gamma\Delta t^2 - (1-\kappa\Delta t)^2\right) - \sqrt{\left(1-2\sigma^2\gamma\Delta t^2 - (1-\kappa\Delta t)^2\right)^2 + 8\sigma^2\gamma\Delta t^2}}{4\sigma^2\Delta t} \leq 0 \tag{88}$$

  Because C_n is always negative, it is clear that $l = x_2$.

Properties of B_n

- **B_n is strictly negative for every maturity above 1:**

  From (20), and due to the negativity of C_n, it is clear that when B_n becomes negative, all the value associated to higher maturity are negative. B_1 is negative so all the values of B_n associated to maturity higher than 1 are strictly negatives. All the others are equal to zero.

- **Convergence of C_n:**

  Due to the convergence of C_n, it appears that:

$$\exists N \in \mathbb{N} / \forall n > N, C_n \tilde{} l_C \tag{89}$$

With: $l_C = \lim_{n \to \infty}(C_n)$. Then: $\forall n > N$,

$$B_n \tilde{} (1 - \kappa \Delta t) \frac{2\kappa\theta\Delta t l_C + B_{n-1}}{1 - 2\sigma^2 \Delta t l_C} = \Gamma + \Theta B_{n-1} \tag{90}$$

With:

$$\Gamma = \frac{2\kappa\theta\Delta t l_C (1 - \kappa\Delta t)}{1 - 2\sigma^2 l_C \Delta t} \tag{91}$$

$$\Theta = \frac{1 - \kappa\Delta t}{1 - 2\sigma^2 \Delta t l_C} \tag{92}$$

where Γ and Θ are two constants. Or, we have:$\forall n \in \mathbb{N}^*$,

$$C_n < 0 \tag{93}$$

Then:

$$1 - \sigma^2 C_n > 1 \tag{94}$$

And because κ is capped by 1:

$$0 < 1 - \kappa\Delta t < 1 \tag{95}$$

Then:$\forall n \in \mathbb{N}^*$,

$$0 < \frac{1 - \kappa\Delta t}{1 - \sigma^2 \Delta t C_n} < 1 \tag{96}$$

This property is still available when n becomes high. Θ is strictly capped and floored by 1 and 0. So:$\forall n > N, \forall m > 0$,

$$B_{n+m} \tilde{} \Gamma \sum_{i=0}^{m} \Theta^i + \Theta^{m+1} B_{n-1} = \frac{\Gamma}{1 - \Theta} + \Theta^{m+1} \left(B_{n-1} - \frac{\Gamma}{1 - \Theta} \right) \tag{97}$$

Or $\Theta \in]0;1[$, then B_{n+m} has a finite limit when m becomes high which is:

$$\begin{aligned} \lim_{m\to\infty}(B_{n+m}) &= \frac{\Gamma}{1-\Theta} \\ &= \frac{2\kappa\theta\Delta t l_C(1_C-\kappa\Delta t)}{\kappa\Delta t-2\sigma^2 l_C\Delta t} \end{aligned} \tag{98}$$

To conclude, at least from a certain value of n, B_n is strictly decreasing and converges to $\frac{2\kappa\theta\Delta t l_C(1-\kappa\Delta t)}{\kappa\Delta t-2\sigma^2 l_C\Delta t}$.

Appendix C: New formulations for A_n, B_n and C_n

The new formulation for C_n

First, in order to have an easier reading of the results, let's consider:

$$\begin{aligned} \alpha &= \alpha\Delta t \\ \gamma &= \gamma\Delta t \\ \kappa &= \kappa\Delta t \\ \sigma_0 &= \sigma_0\Delta t \end{aligned}$$

From (21), $\forall n \in \mathbb{N}^*$,

$$\begin{aligned} C_n &= -\gamma + C_{n-1}\frac{(1-\kappa)^2}{1-2\sigma^2 C_{n-1}} \\ &= \frac{-\gamma + C_{n-1}\left[2\gamma\sigma^2 + (1-\kappa)^2\right]}{1-2\sigma^2 C_{n-1}} && (99) \\ &= \frac{d + aC_{n-1}}{c + bC_{n-1}} && (100) \end{aligned}$$

With :

$$d = -\gamma \quad (101)$$

$$a = 2\gamma\sigma^2 + (1-\kappa)^2 \quad (102)$$

$$c = 1 \quad (103)$$

$$b = -2\sigma^2 \quad (104)$$

Then:

$$\begin{aligned}
C_{n+1} &= \frac{d + aC_n}{c + bC_n} && (105)\\
&= \frac{d + a\frac{d+aC_{n-1}}{c+bC_{n-1}}}{c + b\frac{d+aC_{n-1}}{c+bC_{n-1}}} && (106)\\
&= \frac{d\,(c+a) + \left(a^2 + bd\right) C_{n-1}}{\left(c^2 + bd\right) + b\,(a+c)\, C_{n-1}} && (107)\\
&= \frac{d' + a'C_{n-1}}{c' + b'C_{n-1}} && (108)
\end{aligned}$$

With:

$$\begin{bmatrix} a' & b' \\ d' & c' \end{bmatrix} = \begin{bmatrix} a & b \\ d & c \end{bmatrix}^2 = M^2 \qquad (109)$$

For every n, C_n can then be rewrite as follows:

$$\begin{aligned}
C_n &= \frac{d_n + a_n C_0}{c_n + b_n C_0} \\
&= \frac{d_n}{c_n} && (110)
\end{aligned}$$

With:

$$\begin{bmatrix} a_n & b_n \\ d_n & c_n \end{bmatrix} = M^n \qquad (111)$$

And:

$$M = \begin{bmatrix} 2\gamma\sigma^2 + (1-\kappa)^2 & -2\sigma^2 \\ -\gamma & 1 \end{bmatrix} \qquad (112)$$

Let Tr be the trace of a matrix, det the determinant and $\mathcal{L}(M)$ the characteristic polynomial of M.

$$\begin{aligned}
Tr(M) &= 2\gamma\sigma^2 + (1-\kappa)^2 + 1 && (113)\\
\det(M) &= 2\gamma\sigma^2 + (1-\kappa)^2 - 2\gamma\sigma^2 && (114)\\
&= (1-\kappa)^2 > 0 && (115)
\end{aligned}$$

Then:

$$\mathcal{L}(M) = X^2 - Tr(M)X + \det(M) \tag{116}$$

$$= X^2 - \left[2\gamma\sigma^2 + (1-\kappa)^2 + 1\right] X + (1-\kappa)^2 \tag{117}$$

Or:

$$\Delta = \left[2\gamma\sigma^2 + (1-\kappa)^2 + 1\right]^2 - 4(1-\kappa)^2 \tag{118}$$

$$= \left(2\gamma\sigma^2 + (1-\kappa)^2 + 1 - 2(1-\kappa)\right)\left(2\gamma\sigma^2 + (1-\kappa)^2 + 1 + 2(1-\kappa)\right) \tag{119}$$

$$= \left(2\gamma\sigma^2 + \kappa^2\right)\left(2\gamma\sigma^2 + (2-\kappa)^2\right) > 0 \tag{120}$$

There are 2 different eigenvalues ((24) and (25)):

$$\lambda_1 = \frac{2\gamma\sigma^2 + (1-\kappa)^2 + 1 - \sqrt{(2\gamma\sigma^2 + \kappa^2)(2\gamma\sigma^2 + (2-\kappa)^2)}}{2}$$

$$\lambda_2 = \frac{2\gamma\sigma^2 + (1-\kappa)^2 + 1 + \sqrt{(2\gamma\sigma^2 + \kappa^2)(2\gamma\sigma^2 + (2-\kappa)^2)}}{2}$$

Let $U_1 = \begin{bmatrix} x \\ y \end{bmatrix}$ be one of the eigenvector associated to the first eigenvalue. Then:

$$MU_1 = \lambda_1 U_1 \tag{121}$$

Which means:

$$-\gamma x + y = \lambda_1 \tag{122}$$

Then:

$$U_1 = \begin{bmatrix} \frac{1-\lambda_1}{\gamma} \\ 1 \end{bmatrix} \tag{123}$$

By the way, it also appears that:

$$U_2 = \begin{bmatrix} \frac{1-\lambda_2}{\gamma} \\ 1 \end{bmatrix} \tag{124}$$

The associated transition matrix P is:

$$P = \begin{bmatrix} \frac{1-\lambda_1}{\gamma} & \frac{1-\lambda_2}{\gamma} \\ 1 & 1 \end{bmatrix} \tag{125}$$

$$P^{-1} = \frac{\gamma}{\lambda_2 - \lambda_1} \begin{bmatrix} 1 & \frac{\lambda_2-1}{\gamma} \\ -1 & \frac{1-\lambda_1}{\gamma} \end{bmatrix} \tag{126}$$

Then:

$$M = PDP^{-1} \tag{127}$$

With:

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{128}$$

Then, for every strictly positive n,

$$M^n = PD^nP^{-1} \tag{129}$$

$$= \frac{\gamma}{\lambda_2 - \lambda_1} \begin{bmatrix} \frac{1-\lambda_1}{\gamma} & \frac{1-\lambda_2}{\gamma} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} \begin{bmatrix} 1 & \frac{\lambda_2-1}{\gamma} \\ -1 & \frac{1-\lambda_1}{\gamma} \end{bmatrix} \tag{130}$$

$$= \frac{\gamma}{\lambda_2 - \lambda_1} \begin{bmatrix} \frac{\lambda_1^n(1-\lambda_1)-\lambda_2^n(1-\lambda_2)}{\gamma} & \frac{(1-\lambda_1)(1-\lambda_2)(\lambda_2^n-\lambda_1^n)}{\gamma^2} \\ \lambda_1^n - \lambda_2^n & \frac{\lambda_2^n(1-\lambda_1)-\lambda_1^n(1-\lambda_2)}{\gamma} \end{bmatrix} \tag{131}$$

From which (23) is obtained:$\forall n \in \mathbb{N}^*$,

$$C_n = \gamma \frac{\lambda_1^n - \lambda_2^n}{\lambda_2^n (1-\lambda_1) - \lambda_1^n (1-\lambda_2)}$$

The new formulation for B_n

From (20):

$$\begin{aligned} B_n &= (1-\kappa) \frac{2\kappa\theta C_{n-1} + B_{n-1}}{1 - 2\sigma^2 C_{n-1}} \\ &= \frac{C_n + \gamma}{1-\kappa} \left(2\kappa\theta + \frac{B_{n-1}}{C_{n-1}} \right) \end{aligned} \tag{132}$$

Let be V_n defined as:

$$\forall n \in \mathbb{N}^*, V_n = \frac{B_n}{C_n}$$

Then:

$$\forall n \in \mathbb{N}^*, V_n = \frac{2\kappa\theta}{1-\kappa}\left(\frac{C_n+\gamma}{C_n}\right)\left(1+\sum_{i=2}^{n-1}\left(\frac{1}{(1-\kappa)^{n-i}}\prod_{j=i}^{n-1}\frac{C_j+\gamma}{C_j}\right)\right) \tag{133}$$

Using (23), we have:

$$\prod_{j=i}^{n-1}\frac{C_j+\gamma}{C_j} = \prod_{j=i}^{n-1}\frac{\lambda_2\lambda_1^j-\lambda_1\lambda_2^j}{\lambda_1^j-\lambda_2^j} \tag{134}$$

$$= (\lambda_1\lambda_2)^{n-i}\prod_{j=i}^{n-1}\frac{\lambda_1^{j-1}-\lambda_2^{j-1}}{\lambda_1^j-\lambda_2^j} \tag{135}$$

$$= (\lambda_1\lambda_2)^{n-i}\,\frac{\lambda_1^{i-1}-\lambda_2^{i-1}}{\lambda_1^{n-1}-\lambda_2^{n-1}} \tag{136}$$

Then:

$$\sum_{i=2}^{n-1}\left(\frac{1}{(1-\kappa)^{n-i}}\prod_{j=i}^{n-1}\frac{C_j+\gamma}{C_j}\right) = \sum_{i=2}^{n-1}\left(\left(\frac{\lambda_1\lambda_2}{1-\kappa}\right)^{n-i}\frac{\lambda_1^{i-1}-\lambda_2^{i-1}}{\lambda_1^{n-1}-\lambda_2^{n-1}}\right) \tag{137}$$

$$= \left(\frac{\lambda_1\lambda_2}{1-\kappa}\right)^{n}\frac{\sum_{i=2}^{n-1}\left[\left(\frac{1-\kappa}{\lambda_1\lambda_2}\right)^i\left(\lambda_1^{i-1}-\lambda_2^{i-1}\right)\right]}{\lambda_1^{n-1}-\lambda_2^{n-1}} \tag{138}$$

$$= \frac{(\lambda_1\lambda_2)^{n-1}}{(1-\kappa)^{n-2}}\,\frac{\sum_{i=1}^{2}(-1)^i\left(\frac{1-\left(\frac{1-\kappa}{\lambda_i}\right)^{n-2}}{\lambda_i-1+\kappa}\right)}{\lambda_1^{n-1}-\lambda_2^{n-1}} \tag{139}$$

Finally, by merging (139) and (133), (26) is obtained: $\forall n \in \mathbb{N}\setminus\{0;1\}$

$$B_n = \frac{2\kappa\theta}{1-\kappa}\left(C_n+\gamma\right)\left[1+\frac{(\lambda_1\lambda_2)^{n-1}}{\left(\lambda_1^{n-1}-\lambda_2^{n-1}\right)(1-\kappa)^{n-2}}\sum_{i=1}^{2}(-1)^i\left(\frac{1-\left(\frac{1-\kappa}{\lambda_i}\right)^{n-2}}{\lambda_i-1+\kappa}\right)\right]$$

And with:

$$B_0 = B_1 = 0$$

Appendix D: Determination of the non observed state factor

From (41), we can write:

$$x_t^2 = \frac{-A_{n_2} - R(t, n_2)n_2\Delta t - B_{n_2}x_t}{C_{n_2}} \tag{140}$$

Then, by using (140) in (40), we have:

$$R(t, n_1) = \frac{-A_{n_1} - B_{n_1}x_t - C_{n_1}\frac{-A_{n_2} - R(t,n_2)n_2\Delta t - B_{n_2}x_t}{C_{n_2}}}{n_1\Delta t} \tag{141}$$

Which can be rewritten as (42):

$$x_t = \frac{C_{n_2}\left(R(t, n_1)n_1\Delta t + A_{n_1}\right) - C_{n_1}\left(R(t, n_2)n_2\Delta t + A_{n_2}\right)}{C_{n_1}B_{n_2} - B_{n_1}C_{n_2}}$$

Appendix E: The moments used for the GMM approach

Let's first define the different moment of Δx_t. Here, M_i denotes the moment of order i of Δx_t meaning:

$$M_i = E_t\left[(\Delta x_t)^i\right]$$

- $f_{1,j}$: From (22),

$$R(t+\Delta t, n_j)n_j\Delta t = -A_{n_j} - B_{n_j}x_{t+\Delta t} - C_{n_j}x_{t+\Delta t}^2$$

Using (6):

$$\begin{aligned} R(t+\Delta t, n_j)n_j\Delta t &= -A_{n_j} - B_{n_j}x_t - C_{n_j}x_t^2 && (142)\\ &\quad -B_{n_j}\Delta x_t - C_{n_j}\left(2x_t\Delta x_t + (\Delta x_t)^2\right) && (143)\\ &= R(t, n_j)n_j\Delta t - B_{n_j}\Delta x_t - C_{n_j}\left(2x_t\Delta x_t + (\Delta x_t)^2\right) && (144)\end{aligned}$$

Then:

$$E_t\left[\Delta R(t, n_j)n_j\Delta t\right] = -B_{n_j}M_1 - C_{n_j}\left(2x_tM_1 + M_2\right) \quad (145)$$

- $f_{2,j}$: From (144):

$$\begin{aligned} (\Delta R(t, n_j)n_j\Delta t)^2 &= (B_{n_j} + 2x_tC_{n_j})^2\,(\Delta x_t)^2 + C_{n_j}^2\,(\Delta x_t)^4 && (146)\\ &\quad +2(B_{n_j} + 2 * x_tC_{n_j})C_{n_j}\,(\Delta x_t)^3 && (147)\end{aligned}$$

Then:

$$\begin{aligned} E_t\left[(\Delta R(t, n_j)n_j\Delta t)^2\right] &= (B_{n_j} + 2x_tC_{n_j})^2M_2 + C_{n_j}^2M_4 && (148)\\ &\quad +2(B_{n_j} + 2 * x_tC_{n_j})C_{n_j}M_3 && (149)\end{aligned}$$

- f_3: From (6):

$$E_t\left[\Delta x_t\right] = M_1 \quad (150)$$

- f_4**:** From (6):

$$E_t\left[(\Delta x_t)^2\right] = M_2 \tag{151}$$

- f_5**:** From (22):

$$E_t\left[R(t,n_j)n_j\Delta t\right] = -A_{n_j} - B_{n_j}x_t - C_{n_j}x_t^2 \tag{152}$$

- f_6**:** From (22):

$$E_t\left[\left(R(t,n_j)n_j\Delta t + A_{n_j} + B_{n_j}x_t + C_{n_j}x_t^2\right)^2\right] = 0 \tag{153}$$

Table 1: Calibration Results

Parameter	Estimated value	S.D.
α	0.0211	$(\mathit{3.47.10}^{-7})$
γ	417.2692	$(\mathit{139.72})$
σ	4.10^{-5}	$(\mathit{2.83.10}^{-5})$
θ	0.0128	$(\mathit{0.0069})$
κ	0.0612	$(\mathit{0.0002})$

Table 2: Performance Analysis

	1M IR	**3M IR**	**6M IR**	**1yr IR**
Number of observation	392	392	392	39
Correlation $R_{\mathbf{obs}}$ and $R_{\mathbf{est}}$	85.03%	93.72%	96.45%	92.13%
Associated $p-$value	0%	0%	0%	0%
Correlation residuals and $R_{\mathbf{obs}}$	28.82%	10.25%	49.94%	93.41%
Associated $p-$value	0%	4.3%	0%	0%
Mean of the observed IR	2.405	2.477	2.608	1.654
Mean value of the residuals	0.194	0.006	-0.248	-1.276
Std of the residuals	0.816	0.535	0.416	0.645

Results are obtanied on the one month, three month, six month and one year interest rate. The information given by the table are the number of observations, the correlation between the observed interest rate and the predicted one plus its p-value, the correlation between the residuals and the observed interest rate plus its p-value, the mean of the observed interest rate, the mean value of the residuals and finally the standard deviation of the residuals.

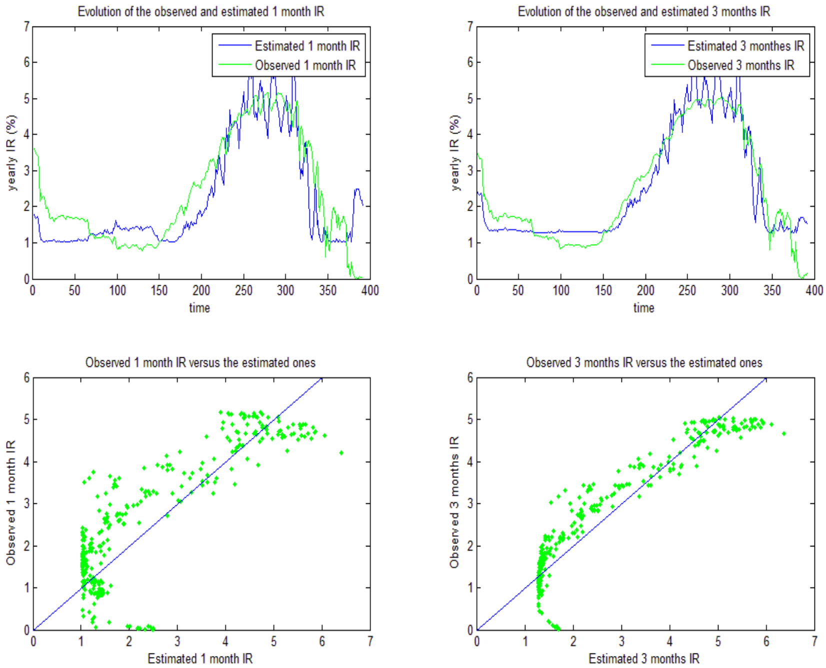

Figure 1: On top of the figure, the trajectories of the observed and predicted interest rate are plotted. On the bottom, the observed interest rate is plotted versus the estimated one. The blue line represents then the optimal position of the dots. On the right, is represented the three months interest rate and on the left the one month one.

Figure 1: Observed versus estimated paths of interest rates

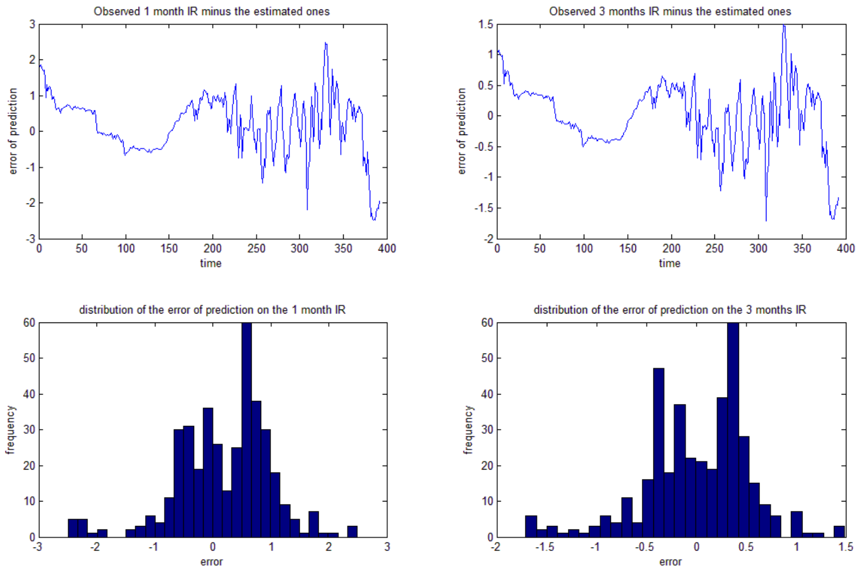

Figure 2: On top of the figure is represented the evolution of the residuals at each time and across the time. On the bottom, the distribution of the residuals is represented. On the rigth, the one month interest rate is considering and the three months one on the left.

Figure 2: Residuals