Maestría en Ingeniería Matemática Facultad de Matemáticas UCM University of Oxford Università degli Studi di Firenze

# II Modelling week

# Interest Rate Curves Calibration with Monte-Carlo Simulation

Work by

A. Baena (UCM)
Y. Borhani (Univ. of Oxford)
E. Leoncini (Univ. of Florence)
R. Minguez (UCM)
J.M. Nkhaso (UCM)
A. Sanchez (UCM)
A. Vilcapoma (UCM)

Supervisors

Dr. G. Oleaga I. Colodrón

## 1 Introduction

In finance the **yield curve** is the relation between the *interest rate* and the *time* to maturity of the debt for a given borrower in a given currency. The value of an interest rate curve can be known today and can be used to obtain today's values for fixed income securities, futures, derivatives etc. Now, if we wish to know the value of these securities in the future, the interest rate curves must be simulated. Our work deals with this simulation.

We first use historical data of interest rates to estimate the parameters required to describe the future values by a Monte Carlo simulator. The underlying assumption is that the interest rates fit a certain model, for instance the log-normal model. Further work includes the generation of parameters for the Nelson-Siegel model using historical data and establishing a model that best describes the evolution of each parameter with time. This allows us to simulate future parameter values for the Nelson-Siegel model and hence gives us an alternate method for predicting future interest rate developments.

## 2 Modelling

#### 2.1 Monte Carlo Simulator assuming a Lognormal Model

Monte Carlo methods are based on random sampling and are most frequently used in problems without a deterministic solution. In finance they are used to value and analyze basic financial models through to complex instruments by simulating various sources of uncertainty affecting their value and then determining their average value over the range of resultant outcomes. We use a Monte Carlo method to simulate the evolution of interest rate curves with time, as outlined below. For this purpose we must assume a model for the evolution. The log-normal model is a suitable and convenient model because it allows us to use a variety of stochastic calculus methods that are based on this assumption.

Let us consider that the risk factors (interest rates for different maturities) follow the lognormal stochastic model with constant instantaneous mean  $\mu$  and volatility  $\sigma$ :

$$\frac{dX}{X} = \mu dt + \sigma dW \tag{2.1}$$

where dW is a Wiener process, ie

$$dW = \sqrt{dt}Z$$

where  $Z \sim N(0, 1)$ , has a standard normal distribution.

This model has an explicit solution for the stochastic variable X, namely:

$$\log (X(t)) = \log (X(t_0)) + \left(\mu - \frac{1}{2}\sigma^2\right)(t - t_0) + \sigma\sqrt{t - t_0}Z \quad Z \sim N(0, 1) .$$

This can be easily checked by Ito's Lemma:

$$d\left(\log\left(X\right)\right) = \frac{\partial}{\partial X} \left(\log\left(X\right)\right) dX + \frac{1}{2} \frac{\partial}{\partial X^2} \left(\log\left(X\right)\right) \left(dX^2\right)$$
$$= \frac{dX}{X} - \frac{1}{2} \frac{dX^2}{X^2}$$
$$= \left(\mu - \frac{1}{2} \sigma^2\right) dt + \sigma \sqrt{dt} Z$$

Then,  $\log(X(t))$  follows a brownian motion with constant drift  $\mu - \frac{1}{2}\sigma^2$  and volatity  $\sigma$ . Integrating the process between  $t_0$  and t we obtain<sup>1</sup>:

$$\log (X(t)) - \log (X(t_0)) = \left(\mu - \frac{\sigma^2}{2}\right)(t - t_0) + \sigma \sqrt{t - t_0} Z,$$

 $\mathbf{or}$ 

$$X(t) = X(t_0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)(t - t_0) + \sigma\sqrt{t - t_0}Z\right).$$
(2.2)

#### One variable: estimating $\mu$ and $\sigma$ from historical data

Suppose we have only one risk factor with its historical data. We want to estimate  $\mu$  and  $\sigma$  from the historical data.

Let us assume that we have M+1 historical observations  $x_i$ , i = 1, ..., M+1, one for each day, and we assume that

$$dt\sim 1$$
 day.

First we construct the returns

$$r_i := \frac{dx_i}{x_i} = \frac{x_{i+1} - x_i}{x_i}$$

and then use the estimator for the daily mean:

$$\bar{r} = \frac{1}{M} \sum_{i=1}^{M} r_i \sim \mu dt$$
 (2.3)

And for the daily volatility:

$$s = \sqrt{\frac{1}{M-1} \sum_{i=1}^{M} (r_i - \bar{r})^2} \sim \sigma \sqrt{dt} \quad \text{(unbiased estimator)}.$$
(2.4)

Then, we have the estimates for the annual mean and volatitity:

$$\mu = \frac{\bar{r}}{dt} \quad \sigma = \frac{s}{\sqrt{dt}} \,.$$

For a unique risk factor, the Monte Carlo simulation is now complete; we must generate a unique sample of a standard normal variable Z and plug the value into (2.2) to obtain simulations of X(t) given a value  $X(t_0)$ .

<sup>&</sup>lt;sup>1</sup>Because of the explicit solution of the SDE we can integrate it. If we change the model we-ll have to compute the solution step by step from the starting time  $t_0$  to the desired time t

#### Several variables: covariance matrix.

In the case of several variables we must take into account the stochastic dependence between samples. Let us assume that we have N risk factors that follow a lognormal model as (2.1):

$$\frac{dX_j}{X_j} = \mu_j dt + \sigma_j dW_j \,. \tag{2.5}$$

Notice that the Wiener process is indexed in order to show explicitly that we have a different stochastic variable for each risk factor.

The way to obtain the parameters for the simulation is the same as for a unique risk factor, using the returns of each variable and the mean (2.3) and volatility (2.4). We use now the notation  $x_{ij}$  for the *i* historical sample of the variable *j*. Assume again that we have M + 1 data for each variable:

$$r_{ij} = \frac{X_{i+1,j} - X_{ij}}{X_{ij}} .$$
(2.6)  

$$\bar{r}_j = \frac{1}{M} \sum_{i=1}^M r_{ij} \sim \mu_j dt .$$

$$s_j = \sqrt{\frac{1}{M-1} \sum_{i=1}^M (r_{ij} - \bar{r}_j)^2} \sim \sigma_j \sqrt{dt} .$$

Therefore, taking into account the anualised values:

$$\mu_j = \frac{1}{dt M} \sum_{i=1}^M r_{ij}$$
$$\sigma_j = \sqrt{\frac{1}{dt (M-1)} \sum_{i=1}^M (r_{ij} - \bar{r}_j)^2}$$

These are the parameters that we will use in the explicit solution for each factor:

$$X_j(t) = X_j(t_0) \exp\left(\left(\mu_j - \frac{\sigma_j^2}{2}\right)(t - t_0) + \sigma_j \sqrt{t - t_0} Z_j\right).$$

The main problem now is how to generate the N normal variables with the correlation implied by the data. This is a standard procedure that can be done in several ways. Here we need to take into account the time scaling factor in the stochastic term of the model (2.5).

Let us call R the matix of returns and  $R^*$  the centered data matrix, that is, the one that is obtained from R by subtracting the daily mean of each column  $(\mu_j dt)$ . Then:

$$\operatorname{cov}\left(r_{l}, r_{m}\right) = \frac{1}{M-1} \sum_{i=1}^{M} r_{il}^{*} r_{im}^{*} \quad (\text{unbiased estimator}) \,. \tag{2.7}$$

Notice that (due to the model selected) this is an estimate of the quantity:

$$\operatorname{cov}(r_l, r_m) \sim \sigma_l \sigma_m dt \operatorname{cov}(Z_l, Z_m)$$
.

Then our covariance matrix is given by

$$\operatorname{cov}\left(Z_{l}, Z_{m}\right) = \frac{\operatorname{cov}\left(r_{l}, r_{m}\right)}{\sigma_{l}\sigma_{m}dt} \,. \tag{2.8}$$

Notice that

$$\sigma_l \sigma_m dt = \frac{1}{(M-1)} \sqrt{\sum_{i=1}^M (r_{il} - \bar{r}_l)^2 \sum_{i=1}^M (r_{im} - \bar{r}_m)^2}.$$

#### Normal correlated samples generated by Cholesky's factorization

Let us call V the covariance matrix  $cov(Z_l, Z_m)$ . A way to generate the samples is to make a kind of Gram-Schmidt process, generating independent  $\xi_i$  normal variables:

$$Z_1 = \lambda_{11}\xi_1 \Rightarrow \lambda_{11} = 1.$$

Now we want  $\xi_2$  independent of  $\xi_1$ :

$$Z_2 = \lambda_{21}\xi_1 + \lambda_{22}\xi_2$$

where  $\lambda_{22}, \lambda_{12}$  are coefficients to be determined in such a way that  $Z_1$  and  $Z_2$  have the correct covariation:

$$V_{22} = \lambda_{22}^2 + \lambda_{21}^2 = 1$$
  
 $V_{12} = \lambda_{21} \Rightarrow \lambda_{22} = \sqrt{1 - V_{12}^2}$ 

The following step:

$$Z_3 = \lambda_{31}\xi_1 + \lambda_{32}\xi_2 + \lambda_{33}\xi_3$$

then:

$$V_{13} = \lambda_{31}$$
  

$$V_{23} = \lambda_{21}\lambda_{31} + \lambda_{22}\lambda_{32}$$
  

$$V_{33} = \lambda_{33}^2 + \lambda_{32}^2 + \lambda_{31}^2$$

and so on.

The Cholesky factorization would be:

$$Z = \Omega \xi \Rightarrow V = E \left[ Z Z^T \right] = \Omega \Omega^T \,,$$

where  $E[\cdot]$  is the expected value applied to each matix component. Computing  $\Omega$  is one way to generate N correlated normal variables by means of N independent ones.

We cannot use the Cholesky factorization of V, because it is not positive definite.

Another way to obtain a factorization of the correlation matrix is explained in the next paragraph.

#### Numbers generated with the data.

Recall that  $R^*$  is the matrix defined by the centered data:

$$r_{ij}^* = r_{ij} - \mu_j$$

and we have by definition (cf. 2.7 and 2.8) that

$$V_{ml} = \operatorname{cov} (Z_m, Z_l) = \frac{\left\{ (R^*)^T R^* \right\}_{lm}}{(M-1) \,\sigma_l \sigma_m dt}$$

So that, if we rescale the data:

$$C_{ij} = \frac{r_{ij}^*}{\sigma_j \sqrt{(M-1)\,dt}}$$

Then

$$V = C^T C$$

Now consider the output of the product

$$Z = \Omega \xi \tag{2.9}$$

where  $\xi$  is an M dimensional vector of independent standard normal variables and  $\Omega = C^T$ . If we consider the covariance matrix of the variable Z we have that

$$E\left[\left(\Omega\xi\right)\left(\Omega\xi\right)^{T}\right] = \Omega\Omega^{T} = C^{T}C = V$$

Therefore, we can generate a sample of normal correlated variables with V correlation matrix, with the formula (2.9).

#### 2.2 Fitting yield curves: the Nelson-Siegel model.

The fitting method of Nelson and Siegel [1] allows to construct the *instantaneous* forward yield curve by a family of functions consisting of a constant and the solutions of a second order differential equation with constant coefficients, when the roots of the associated polynomial are real and equal. In this case we can write:

$$r(m) = \beta_0 + \beta_1 \exp(-m/\tau) + \beta_2(m/\tau) \exp(-m/\tau) . \qquad (2.10)$$

The relationship between forward and continuous rates with maturity m is as follows. Consider the *growing factor* on the interval  $[m, m + \Delta m]$ :

$$\frac{P\left(0,m\right)}{P\left(0,m+\Delta m\right)} \quad \Delta m > 0,$$

where P(t,m) is the price at time t of a unit amount of money that will be paid in time m. When we compute the forward instantaneous in that interval for  $\Delta m \to 0$  we obtain:

$$1 + r(m)\Delta m \approx \frac{P(0,m)}{P(0,m+\Delta m)} \Rightarrow r(m) = \lim_{\Delta m \to 0} \frac{P(0,m) - P(0,m+\Delta m)}{\Delta m P(0,m+\Delta m)},$$

and then

$$r(m) = -\frac{\partial P(0,m)}{\partial m} \frac{1}{P(0,m)}$$

For continous compounding we have that:

$$P(0,m) = \exp(-mR(m)) \Rightarrow \frac{\partial P(0,m)}{\partial m} = -(mR(m))' \exp(-mR(m))$$

and we obtain:

$$r(m) = (mR(m))' \Rightarrow R(m) = \frac{1}{m} \int_0^m r(x) \, dx \, ,$$

therefore, R(m) is a kind of *mean integral* value of the forward instantaneous rates that we observe today. Taking into account this formula and the model (2.10), we can integrate and obtain:

$$R(m) = \frac{1}{m} \left( \beta_0 m - \tau \beta_1 e^{\frac{-m}{\tau}} + \beta_2 \left( \tau \left( 1 - e^{\frac{-m}{\tau}} \right) - m e^{\frac{-m}{\tau}} \right) \right)$$
$$= \beta_0 + \frac{\tau}{m} \left( \beta_1 + \beta_2 \right) \left( 1 - e^{\frac{-m}{\tau}} \right) - \beta_2 e^{\frac{-m}{\tau}}$$

In this case  $\beta_0$  is the behaviour for  $m \to \infty$ ,  $\beta_0 + \beta_1$  is the short term behavior, and  $\beta_1 + \beta_2$  defines the mid term behaviour (about two or three years).

### 3 Implementation of Montecarlo method

As mentioned previously, Monte Carlo methods require a given distribution for generating random scenarios. We propose the usage of a **lognormal distrubution** and we first need to statistically validate this choice using our historical data.

In general, interest rate samples are distributed around a central value and exhibit a drift with time. This is an important property of the data since the majority of tools used in statistical analysis depend on the stationary nature of the data over time, that is the distribution of samples around their average without drift.

In figure 1 we see a typical evolution of the interest rate in a given day. The variation is due to a unitary root in the stochastic variable. If the root is lower than 1, the variable average is stationary, with the variable itself varying in a random way without drifting away from the average value.

We conducted statistical tests to establish whether the interest rate and the first difference of its logarithm (which gives the yield) have a unit root. The data (the significance level in particular) implies that there is no unit root for the interest rates themselves, but for the first difference of the logarithm. The use of the lognormal model can therefore be justified in the latter case only.

As the data for the interest rates is not stationary, it cannot be normally distributed. Thus the equation

$$dS = \mu dt + \sigma dW \tag{3.1}$$



Figure 1: In the first column we have the histograms of the development of interest rates for different maturities (1 week, 1 month, 20 years). The second column shows the returns based on the interest rates in column 1.

D-Lag	t-adf	$\beta Y_1$	$\sigma$	t-DYlag	t-prob	AIC	F-prob
2	-3.686**	0.66740	0.01349	0.3604	0.7193	-8.572	
1	-3.841**	0.67941	0.01343	-0.3711	0.7114	-8.591	0.7193
0	-4.405**	0.66662	0.01337			-8.610	0.8755

Table 1: 1 DAY — LIR-EUR-MONEY-TN: ADF tests (T = 99, Constant;  $5\% = -2.89 \ 1\% = -3.50$ )

D-Lag	t-adf	$\beta Y_1$	$\sigma$	t-DYlag	t-prob	AIC	F-prob
2	-1.504	0.94818	0.005580	-1.731	0.0866	-10.34	
1	-1.805	0.93808	0.005637	0.3914	0.6963	-10.33	0.0866
0	-1.772	0.94036	0.005613			-10.35	0.2120

Table 2: 1 WEEK — LIR-EUR-MONEY-1W: ADF tests (T = 99, Constant; 5% = -2.89 1% = -3.50)

where W is a *Wiener process*, does not hold.

Eventually we show that historical returns follow the model

$$\frac{dS}{S} = \mu dt + \sigma dW \tag{3.2}$$

In fact we can divide it into two parts:

D-Lag	t-adf	$\beta Y_1$	$\sigma$	t-DYlag	t-prob	AIC	F-prob
2	-1.102	0.98104	0.003737	-0.2790	0.7808	-11.14	
1	-1.131	0.98069	0.003719	-1.189	0.2375	-11.16	0.7808
0	-1.233	0.97897	0.003727			-11.16	0.4805

Table 3: 1 MONTH — LIR-EUR-MONEY-1M: ADF tests (T = 99, Constant;  $5\% = -2.89 \ 1\% = -3.50$ )

• unit root test: to look for data stability. So we can summarize in tables, as we did before. From tables 4, 5, 6 we figure that we can accept the null hypotesys so we can figure that the date of the returns are stables.

D-Lag	t-adf	$\beta Y_1$	σ	t-DYlag	t-prob	AIC	F-prob
4	-5.503**	-0.65021	0.01451	0.7836	0.4353	-8.405	
3	-6.079**	-0.52145	0.01448	1.562	0.1217	-8.419	0.4353
2	-6.192**	-0.30764	0.01459	-0.03037	0.9758	-8.414	0.2238
1	-8.216**	-0.31181	0.01452	0.9226	0.3586	-8.434	0.3899
0	-11.86**	-0.19775	0.01450			-8.446	0.4258

Table 4: 1 DAY — DLIR-EUR-MONEY-TN: ADF tests (T = 96, Constant;  $5\% = -2.89 \ 1\% = -3.50$ )

D-Lag	t-adf	$\beta Y_1$	σ	t-DYlag	t-prob	AIC	F-prob
4	-4.132**	-0.17232	0.005640	-1.458	0.1484	-10.30	
3	-5.850**	-0.39837	0.005674	0.7507	0.4547	-10.29	0.1484
2	-6.667**	-0.29412	0.005661	0.6985	0.4866	-10.31	0.2650
1	-8.484**	-0.20170	0.005645	2.171	0.0324	-10.32	0.3692
0	-9.564**	0.013591	0.005756			-10.29	0.1046

Table 5: 1 WEEK — DLIR-EUR-MONEY-TN: ADF tests (T = 96, Constant;  $5\% = -2.89 \ 1\% = -3.50$ )

D-Lag	t-adf	$\beta Y_1$	σ	t-DYlag	t-prob	AIC	F-prob
4	-3.050*	0.24405	0.003293	-5.425	0.0000	-11.37	
3	-6.127**	-0.47223	0.003773	1.673	0.0978	-11.11	0.0000
2	-6.128**	-0.25784	0.003809	0.4947	0.6220	-11.10	0.0000
1	-7.615**	-0.19258	0.003794	0.4405	0.6606	-11.12	0.0000
0	-11.14**	-0.14053	0.003778			-11.14	0.0000

Table 6: 1 MONTH — DLIR-EUR-MONEY-TN: ADF tests (T = 96, Constant; 5% = -2.89 1% = -3.50)

• normal test: to verify that the returns are normal distributed. We can split the data in two parts. The data in the *short run* and those in the *long run*. For the short run we can see that the normal model doesn't fit the data. In fact the normal test refuse the null hypothesis:

 $H_0$ : does a normal distribution fit the data for the returns? as we can observe by table 7 (the behaviour is the same until the 2 years).

On the other hand data in the long run accept the null hypothesis, as we can observe by table 8 and the behaviour is quite the same for all dates after 3 years.

Observations	101
Mean	-0.00012613
Std Dev.	0.014406
$\operatorname{Skewness}$	0.16399
Excess Kurtosis	3.9402
Minimum	-0.058064
Maximum	0.058064
Asymptotic test:	$\chi^2(2) = 65.786 \ [0.0000]^{**}$
Normality test:	$\chi^2(2) = 44.356 \ [0.0000]^{**}$

Table	7:	1	DAY	normality	test

Observations	101
Mean	0.0024918
Std Dev.	0.0024918
Skewness	0.0024918
Excess Kurtosis	0.28271
Minimum	-0.027448
Maximum	0.045584
Asymptotic test:	$\chi^2(2) = 1.4467 \ [0.4851]$
Normality test:	$\chi^2(2) = 1.8847 \ [0.3897]$

Table 8: 3 YEARS normality test

We have finally to test one more thing to see if the returns fits the model. When we assume that they follow a model like in 2.1, we're assuming that the process have independent increments. If we take a daily discretization of time and we consider

$$\frac{dr_t}{r_t} = \mu dt + W_t$$
$$\frac{dr_{t+1}}{r_{t+1}} = \mu dt + W_{t+1}$$

where  $dr_t = r_{t+1} - r_t$ ,  $r_t$  is the interest rate for a fixed risk factor and  $W_t$  is a Wiener process, we are assuming that  $W_t$  and  $W_{t+1}$  are independent. So we we look at the *autocorrelation function* in fig. 2 we notice that there is not evidence of a correlation of the returns computed today with those that were given by previous days. We can observe this behaviour by the histogram of the *autocorrelation of the returns*, fig. 2 and notice that a periodic correlation with the 5<sup>th</sup> day, 10<sup>th</sup> day, ...



Figure 2: Autocorrelation of returns of different risk factors (1 week, 1 month, 20 years).

#### 3.1 Criticism of the Montecarlo method

#### Problems with the Montecarlo method

We focus now on the problems that arise using this kind of model. The first evident problem is that the curves generated with the Montecarlo method are *discontinuous* as we can observe by figure 3 Another problem is connected with the distribution. As mentioned before we noticed that in the short run data doesn't have a normal distribution while they fit it well in the long run. This is put into evidence also by fig. 4.

We can notice that while for the long run the curves approach each other, in the short run they show significant differences.

#### **Possibly solutions**

Looking at fig. 4 we notice that the density function generated by the data shows higher tails than the normal, so we could try to modified the normal distribution with the *t-Student* or other distribution with thick tails. But in this case we'll miss the power of the stochastic calculus and the more similarity with the data will change in a more difficult model, with heavier computations.

We'll try a different approach with the Nelson-Siegel model, as shown in the next paragraphs.



Figure 3: Simulated interest rate curve using Montecarlo simulator.



Figure 4: Density graphs. Here we've put the theoric density function (normal, red line) together with the estimated density function for the data.

## 4 Implementation of Nelson Siegel model

As we've done for the Montecarlo method we're going to analyze statistically the data and to get few estimates. We used the original data of interest rates and a matlab program for the calibration of the Nelson Siegel model in order to obtain *historical data for the Nelson Siegel model* (the *betas*).

#### 4.1 Stability test

We made the unit root test to look for stability and we took into account short sample size because we noticed strange variations and behaviour for the parameters estimated using the full sample size.

Integrating over time we get the graphics as shown in fig. 5 that give us the gross behaviour of the Nelson Siegel parameters.



Figure 5: Beta estimated using short sample size.

Using the unit root stationarity test we find that  $\beta_0$  and  $\tau$  are stationary, but  $\beta_1$  and  $\beta_2$  are not. Therefore we have to modelize the first difference of  $\beta_1$ and  $\beta_2$  (see fig. 6), while we-ll use the original series for  $\beta_0$  and  $\tau$ .



Figure 6: First difference of  $\beta_1$  and  $\beta_2$ .

#### 4.2 Normality test

Proceeding with the normality test we noticed that all the series of betas show a behaviour typical for normal distributions. We have few doubts only for the  $\beta_0$  parameter. Therefore we can apply the ARIMA model methodology to estimate which are the stochastic processes which generate the betas, in this case we do not make any *a priori* assumptions about the behaviour of this stochastic process.

#### 4.3 Identification of the stochastic processes that generate betas and calibration

We use in this part statistical tools to find out which would be the stochastic processes that generate the betas. We'll omit the technical details focusing our attention on few graphs to discover which are the orders of the ARIMA models which generate each parameter (see fig. 7) and giving the conclusions.

Here we can underline that the data for the betas don't show independence (as shown in the second column of graphs). In the third column of fig. 7 we put into evidence that the today-value of a parameter is strictly correlated with the value at the previous time, but is weakly connected with the other ones. This lead to the conclusion that the stochastic processes that generate the parameters are

 $\begin{array}{lll} \beta_0 & \text{ARIMA}(1,0,0) \\ \beta_1 & \text{ARIMA}(0,1,1) \\ \beta_2 & \text{ARIMA}(0,1,1) \\ \tau & \text{ARIMA}(0,0,1) \end{array}$ 

The calibration of the previous models using the historical data for the betas



Figure 7: In the first row is shown the behaviour of  $\beta_0$ ,  $D\beta_1$ ,  $D\beta_2$  and  $\tau$ . In the second column are shown the respective autocorrelation functions and in the last one the partial autocorrelation functions

leads to the calibrated models:

where we put into brackets the standard deviations for the estimate of the parameters of the model.

## 5 Conclusions

We identified the statistical models underlying the evolution of the parameters of the Nelson Siegel Model in time. We found that the parameters are not constant and, in fact, we have estimated the stochastic processes which rule the evolution of these parameters. Afterwards, we can fit the interest rate curve using the estimated parameters and, also, to forecast the parameters out-of-sample and, using them, make a prediction for the interest rate curve in the future. The



Figure 8: Comparison between historical data with the fitted curve using the Nelson Siegel model for the date  $8^{th}$  of May of 2008.



Figure 9: Forecast of the interest rate curve using the Nelson Siegel model. The T-curve is real data from  $8^{th}$  of May of 2008 and the other curves are forecasts.

advantage of using the Nelson Siegel model with estimated parameters over the Monte Carlo simulation is the absence of assumptions about the data a priori, as we did not assume the data to fit any particular model. We also found the Nelson Siegel curves for the interest rates are smoother than those generated by Monte Carlo simulation.

Regarding the Monte Carlo simulator, we found that the lognormal model is insufficient for the purpose of the simulation and that a more complex model in the form of a different distribution is required to obtain more accurate results. In particular, we need a model that accounts for the outliers in the data that do not fit a normal distribution.

### **Further Work**

All of our work was centred on a single data set. Further work would involve testing our results on different interest rate data samples to confirm their global validity as well as developing more complex models that better fit the data for the Monte Carlo simulations.

## References

 Charles R. Nelson, Andrew F. Siegel, Parsimonious Modelling of Yield Curves, Journal of Business, 1987, vol. 60, no. 4.



Figure 10: Residuals for the ARIMA models calibrated to the parameters in the Nelson Siegel model.