UNIVERSIDAD COMPLUTENSE DE MADRID

MODELLING WEEK 2010

Problem IV

Calibration of single-factor HJM models of interest rates

PARTICIPANTS:

ANTONIO BUENO (UCM)
JAVIER GARCÍA (UCM)
SENSHAN JI (UAB)
SANTIAGO LÓPEZ VIZCAYNO (UCM)
ALEJANDRA SÁNCHEZ (UCM)
DANIEL NEIRA VERDES-MONTENEGRO (UCM)
MARCO CAROCCIA (UDSF)DANI

COORDINATORS:

MIGUEL CARRIÓN ÁLVAREZ (BANCO SANTANDER)

GERARDO OLEAGA APADULA (UCM)

Contents

Part I

Introduction

Chapter 1

1. Introduction

The Heath-Jarrow-Morton model provides a framework for discussing arbitrage-free evolution of the interest rate curves. In this paper, we propose to explore a few issues arising the calibration of the model to real data.

First of all, we will deal with the problem of getting time series of Zero-Coupon bond prices suitable for our data-first approach. As we will see, this will lead us to the definition of a new asset unifying the desirable properties of a time-series –homocedasticity, not autocorrelated, stationarity.

After we have gone through the definition of the theoretical one-factor HJM interest rate model, we will discuss how to adapt this model to the data. Particularly, we will renounce to model the evolution of the instantaneous forward rate as it is impossible to recover it from the data.

In addition, we will have to move from an n-dimensional problem, where n correlated Brownian Motions are involved, to a one-dimensional problem, where a single Brownian Motion is would be driving the dynamic of the whole forward curve. This would lead us to conduct a Principal Component analysis of the Covariance Matrix of the forward rates.

Ideally, the resulting model will have to satisfy the no arbitrage condition and would be suitable for risk-neutral valuation of interest rate derivatives.

Part II

Survey

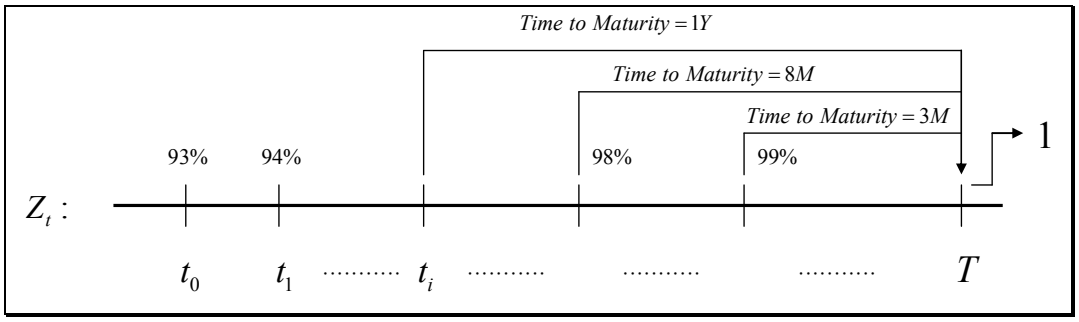
Chapter 2

2. Market and Synthetic zero coupon bonds

2.1. Synthetic asset

We are interested in studying the price of money to a certain specified period, say six months, for example. We need to know of the price of money six months, the market provides us a zero coupon bond that matures within six months. The price of this bond we can supply the interest rate to six months and therefore know the price of money six months. Such bonds are very liquid in the market and allows us to set the price of money in a way appropriate to the market, but the problem with these bonds is the following:

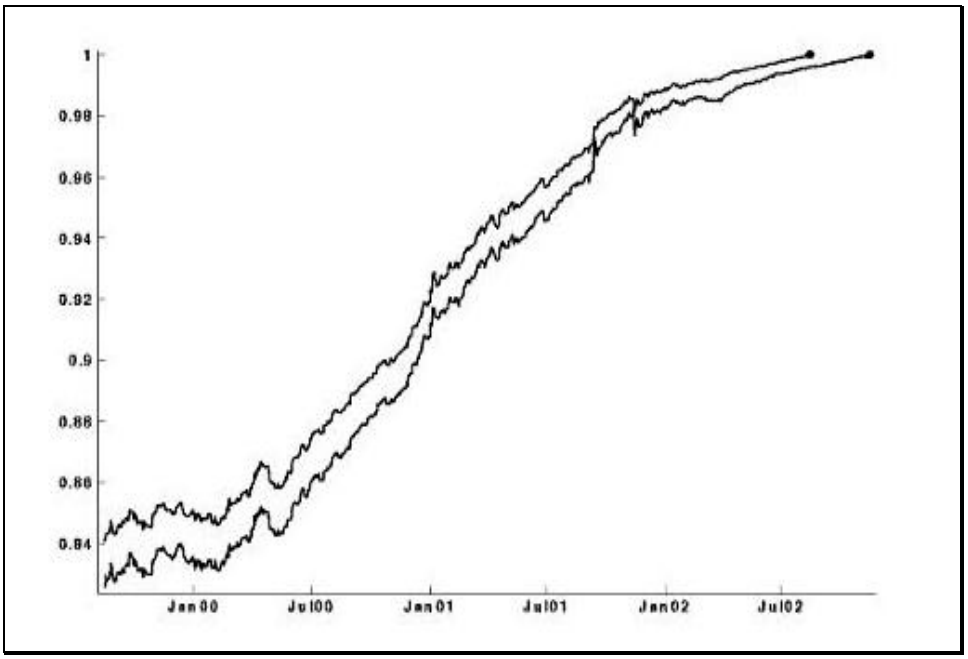
Imagine we are on January 1, 2010, so to know the price of money we observe zero coupon bond maturing on July 1, 2010. The next day, if we know the price of money six months, and we can not use the bond that matures on July 1, 2010 because he no longer has a maturity of six months, with the bonus that would not give us the price of money to six months, but the price of money at a time interval smaller.



Graph 2.1 *Synthetic asset*

This chart perfectly reflects our problem, as we move in the days, the bond we had considered going to start taking less time to maturity, so we will not be able to distill from it the money price of constant interval.

If we observe how the price of the bond evolves over the days we see that the time series of prices can never be stationary since they have to 1 when t → T is deterministic for a given time t, since we know that at maturity, the bond's value is 1; also highly correlated.

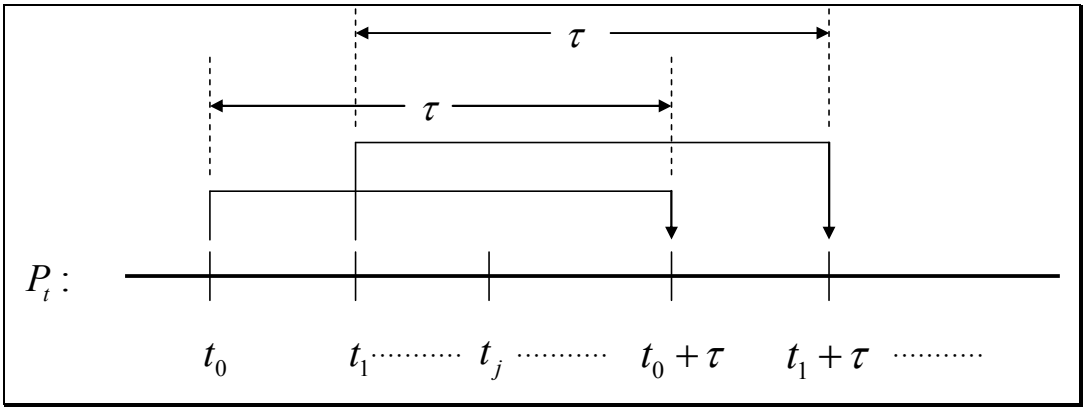


Graph 2.2 Evolution of the bond price

For this reason, we need to define a new synthetic asset that we can not find in the market (not tradeable asset). In this new asset we time to impose the time to maturity is fixed, i.e. today lacks 6 months to maturity, remain missing six months ago to maturity within a year will continue missing six months for the synthetic asset reaches maturity.

We define the active synthetic as:

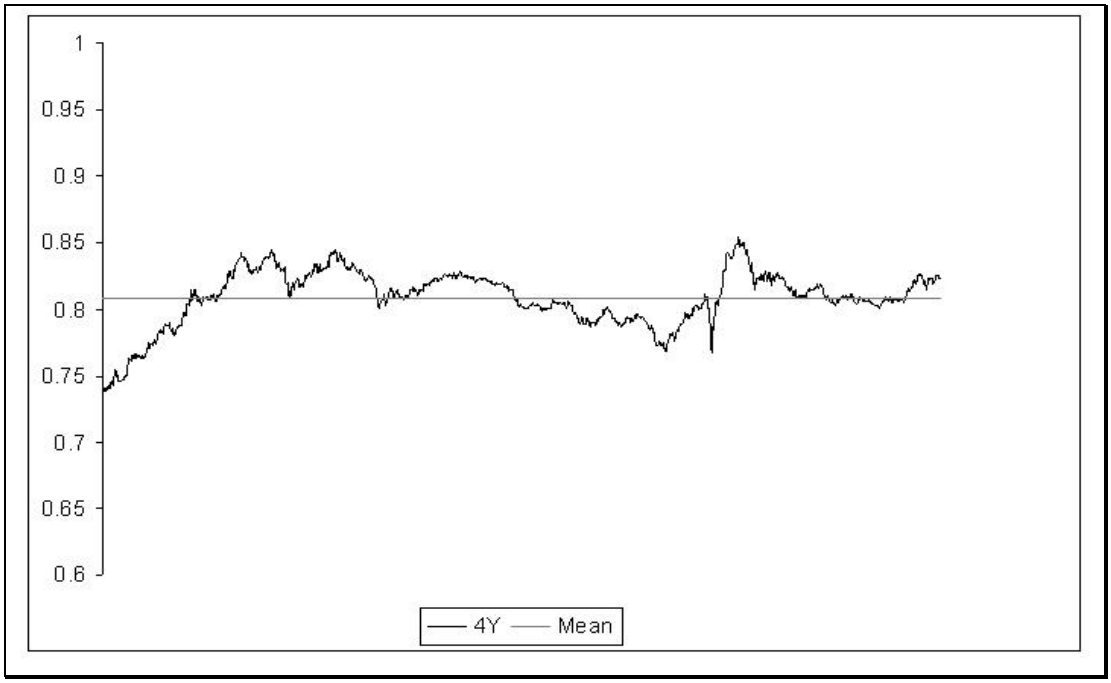
$$P(t,\tau)=Z(t,t+\tau)$$



Graph 2.3 *Synthetic variable*

The main advantages of this asset is that we can study the price of money for a specified period (eg 6 months) and we can also obtain a time series long enough to perform statistical analysis allows us to obtain valid conclusions.

If we look now to the time series synthetic asset that we have defined, we will see that this is a stationary stochastic process is not deterministic at maturity (in fact never reaches maturity). The series has no yield, homoscedastic and has no cycle. The only problem we encountered is that the series is still very autocorrelated.



Graph 2.4 *Time series synthetic asset*

2.2. Concepts and definitions

The introduction of synthetic asset we can not see directly in the market, forces us to give a brief overview of some basic financial concepts that are widely used in market interest rates and that should be well known to the zero coupon bond market, but differ slightly in the case of synthetic bond.

2.2.1. Yields

Informative measure of the market is in indication of the implied average interest rate offered by a bond. If interest rates were constant at rate , the price of the T-bond at time would be . In this particular case, can be recovered from the price:

$$r = -\ln\left(\frac{Z(t,T)}{T-t}\right).$$

The rate we derive, $R(t,T)$, is called the yield, and mapping from price to yield is one-to-one for t less than T-no information is lost

Given a discount bond price $Z(t,T)$ at time t, the yield $R(t,T)$ is given by:

$$R(t,T) = -\frac{\ln\left(Z(t,T)\right)}{T-t}$$

Given a synthetic discount bond price $P(t,\tau)$ at time t, the yield synthetic asset $R(t,\tau)$ is given by:

$$R(t,\tau) = -\frac{\ln\left(P(t,\tau)\right)}{T-t}$$

Therefore, we are going to study the yields for different τ seen in different dates from observation, $t, t+1, ..., t+N$.

2.2.2. Instantaneous rate

But what is the price of money now? The current cost of borrowing in a single number. We can do is look at the current rate for instantaneous borrowing, that is, borrowing which is paid back (nearly) instantly. The period from t to t + Δt where Δt is a small time increment, the rate we get is the yield R(t,t + Δt):

$$R(t,t+\Delta t) = -\frac{\ln\left(Z(t,t+\Delta t)\right)}{\Delta t}$$

For ever smaller time increments this value more closely approximates to R(t,t), which is the left-most point of the yield curve at time t. We call this value the instantaneous rate, or short rate, r_t which is given by both the expressions:

$$r_t = R(t,t)$$

$$\text{and } r_t = -\frac{\partial}{\partial T}\ln\left(Z(t,T)\right)$$

The period from t to T, where T is a small time increment, the rate we get is the yield:

$$R_t(\tau) = -\tfrac{1}{\tau} log P_t(T)$$

Where r_t is the instantaneos rate

$$r_t = \lim_{\tau \to 0} R_t(\tau) = -\frac{\partial}{\partial \tau}\log P_t(t+\tau)\ _{\tau=0} =$$

$$-\partial_{\overline{\partial T \log P_t(T)|_{T=t}}}$$

2.2.3. Forwards

The rate process r_t is one-to-one on how bond prices can move will not in general be enough to recover Z(t,T). Yet the instantaneous rate is a natural expression of r_t which brings back the one-to-one mapping to the prices Z(t,T) and the yields R(t,T), yet still preserves the idea of instantaneity.

Consider Forwards contracts, that is agreeing, at time t, to make a payment at a later date T_1 and receive a payment in return at an even later date T_2. Are really just striking a forward on the T_2-bond. But, what forward price should we pay?.

This deal has initial cost has initial cost $Z(t,T_2) - kZ(t,T_1)$ at time t, and will require us to make a payment of k at time T_1, and will give us a payment of one dollar at time T_2. To give the

$$k = \frac{Z(t,T_2)}{Z(t,T_1)}.$$

The corresponding (forward) yield is then

$$-\frac{\ln\left(Z(t,T_2)\right) - \ln\left(Z(t,T_1)\right)}{T_2 - T_1}$$

Were we to choose T_1 and T_2 very close together, say $T_{1=T}$ and T_2=T + Δt, then as the increment Δt became smaller this would converge to a forward rate for instantaneous borrowing:

$$f(t,T) = -\frac{\partial}{\partial T}\ln\left(Z(t,T)\right)$$

Chapter 3

3. Data Analysis

3.1. Synthetic asset

The Bank of Santander facilitated to us a series of data formed by the observations of 1528 days (from July 16, 2004 until June 6, 2010) of the prices of the zero coupon bonds with different times to maturity: ON (overnight), TN (Tomorrow next), 1W (one week), 1M,2M,3M,4M,5M,6M,9M (1,2,3,4,5,6,9 months), and 1Y,2Y,3Y,4Y,5Y,6Y,7Y,8Y,9Y,10Y (1,2,3,4,5,6,7,8,9, 10 years). Here we have an example of these:

$\tau = overnigth$ $\tau = 1\,month$

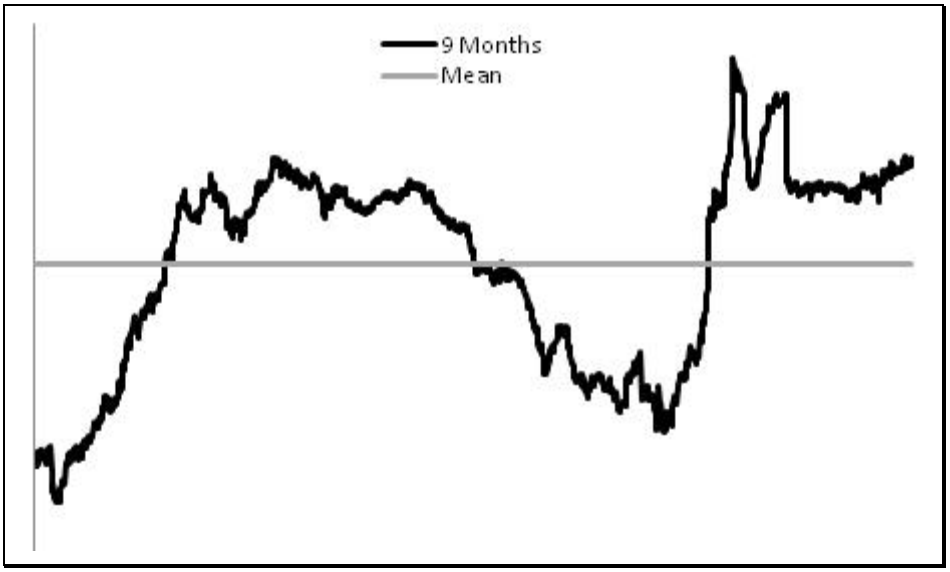
Diferents dates $t_1, t_2, \ldots$

	A	B	C	D	E
1	Date	ON	TN	1W	1M
2	16/07/2004	0,99952598	0,99936690	0,99825460	0,99434207
3	19/07/2004	0,99984441	0,99968802	0,99858110	0,99436585
4	20/07/2004	0,99984523	0,99968939	0,99859777	0,99452765
5	21/07/2004	0,99984715	0,99969404	0,99861772	0,99469278
6	22/07/2004	0,99985865	0,99942082	0,99836584	0,99439572
7	23/07/2004	0,99959742	0,99946434	0,99843800	0,99447264
8	26/07/2004	0,99987728	0,99974635	0,99873886	0,99440598
9	27/07/2004	0,99986385	0,99970856	0,99859205	0,99447733
10	28/07/2004	0,99984140	0,99967843	0,99853135	0,99455118
11	29/07/2004	0,99983647	0,99934446	0,99819011	0,99417692
12	30/07/2004	0,99952187	0,99935541	0,99820296	0,99421301
13	02/08/2004	0,99983592	0,99967104	0,99851823	0,99417015
14	03/08/2004	0,99984222	0,99968008	0,99851961	0,99432760
15	04/08/2004	0,99984194	0,99967926	0,99850922	0,99445957
16	05/08/2004	0,99984468	0,99936663	0,99821608	0,99416537
17	06/08/2004	0,99951448	0,99934911	0,99818902	0,99411435
18	09/08/2004	0,99984140	0,99968336	0,99854967	0,99417346
19	10/08/2004	0,99984687	0,99968802	0,99855432	0,99434417
20	11/08/2004	0,99984359	0,99968199	0,99854447	0,99444549
21	12/08/2004	0,99983126	0,99932284	0,99816278	0,99406302
22	13/08/2004	0,99949067	0,99932065	0,99813956	0,99399366
23	16/08/2004	0,99983208	0,99966228	0,99847696	0,99402737
24	17/08/2004	0,99982606	0,99965078	0,99846165	0,99420309
25	18/08/2004	0,99982661	0,99965105	0,99846766	0,99438950
26	19/08/2004	0,99983674	0,99933242	0,99818382	0,99411454
27	20/08/2004	0,99951202	0,99935212	0,99820351	0,99413414
28	23/08/2004	0,99983920	0,99967843	0,99851988	0,99401662
29	24/08/2004	0,99984633	0,99969021	0,99854120	0,99421630
30	25/08/2004	0,99981894	0,99963517	0,99840399	0,99408816
31	26/08/2004	0,99980470	0,99922980	0,99793030	0,99359272
32	27/08/2004	0,99941514	0,99919969	0,99792891	0,99375176
33	30/08/2004	0,99982250	0,99964229	0,99838624	0,99422445
34	31/08/2004	0,99982715	0,99965215	0,99839034	0,99387514

Graph 3.1 *Original data*

Due to the fact $Z_T(T) = 1$, $Z_t(T)$ can't be a stationary process, we are going to see it. It is said that a process is stationary if it has no trend, homoscedastic (variance of shocks in time constants) and has no stationary (movement of oscillation in a period).

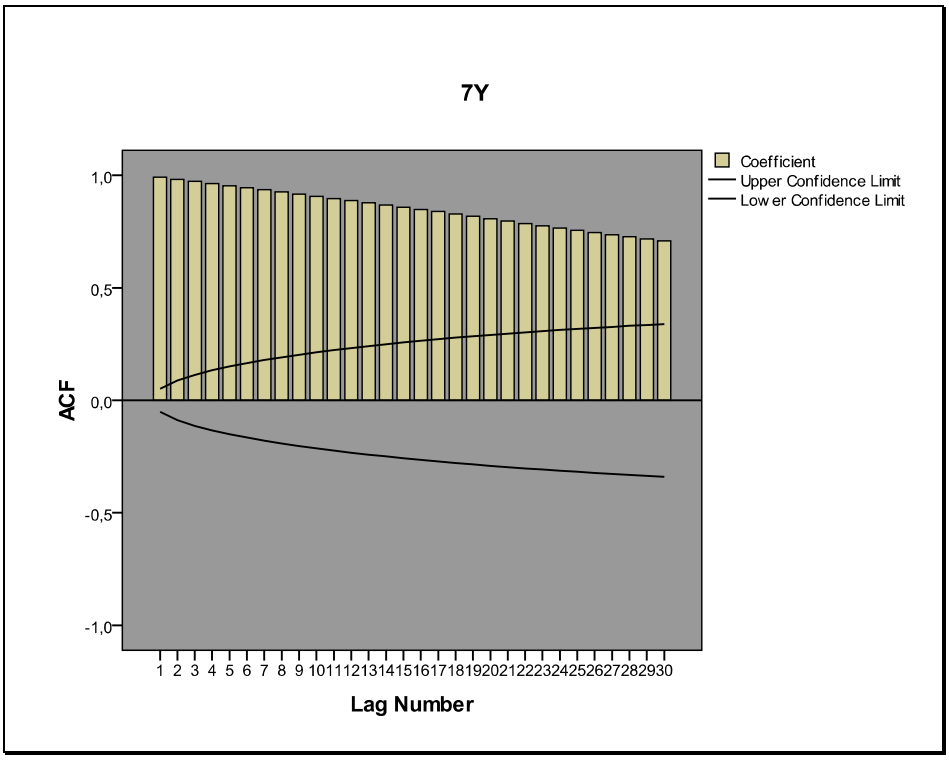
Here is the series of nine months:



Graph 3.2 *The series of nine months*

Clearly the trend series has generally increased, although decreases in some moments. Furthermore, it is homoscedastic and no seasonality. Therefore, it follows that the series is not stationary.

Looking at the graphs of autocorrelations:



Graph 3.3 *Simple autocorrelation*

Clearly shows that a auto-correlation between each day and the last.

So to work with the data we need to do some data processing to eliminate the tendency originals (which in this case we removed the stationarity since the series is homoscedastic and non-seasonal), and try to diminish the strong Autocorrelations. Thus, considering the "synthetic zero coupon bond, we calculate the following transformation:

$$X_t := -\frac{1}{\tau}\log\left(\frac{P_{t+dt}(\tau)}{P_t(\tau)}\right).$$

We have, for example:

$P(t,\tau)$

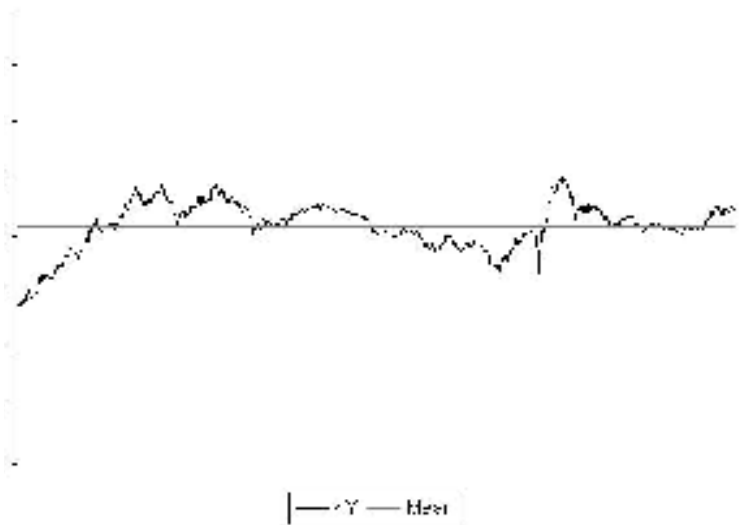
$P(t+dt,\tau)$

$-\frac{1}{\tau}\log\left(\frac{P_{t+dt}(\tau)}{P_t(\tau)}\right)$

	A	B	I	J	K	P	Q	T	U
1	Date	ON	5M	6M	9M	5Y	6Y	9Y	10Y
1490	20/04/2010	-0,07593%	-0,00105%	0,02862%	-0,05530%	0,00984%	0,00984%	0,01126%	0,01587%
1491	21/04/2010	0,00000%	0,02974%	-0,01263%	0,09082%	0,02388%	0,02498%	0,03468%	0,02847%
1492	22/04/2010	-0,02072%	-0,06291%	-0,04057%	-0,04842%	-0,07704%	-0,06355%	-0,06045%	-0,07385%
1493	23/04/2010	-4,34884%	0,02104%	0,05763%	0,02142%	-0,01082%	-0,01904%	-0,04827%	-0,00294%
1494	26/04/2010	4,36956%	0,04201%	0,02517%	0,00216%	0,02287%	0,01683%	0,04084%	0,02646%
1495	27/04/2010	0,06903%	0,00125%	-0,02846%	-0,01587%	-0,12685%	-0,11581%	-0,11413%	-0,12421%
1496	28/04/2010	0,00000%	-0,00007%	0,02117%	0,02462%	-0,06865%	-0,09923%	-0,08759%	-0,07264%
1497	29/04/2010	0,00000%	-0,05719%	-0,04068%	-0,03438%	0,09776%	0,09623%	0,09602%	0,12384%
1498	30/04/2010	-7,08205%	-0,01334%	-0,03089%	-0,02043%	-0,00992%	-0,02085%	-0,00677%	-0,03303%
1499	03/05/2010	6,94400%	0,06725%	0,02638%	0,02102%	0,00560%	0,00467%	0,00311%	0,00280%
1500	04/05/2010	-0,28305%	0,06964%	0,00358%	0,00113%	-0,12296%	-0,12834%	-0,12538%	-0,12665%
1501	05/05/2010	-0,00691%	-0,00053%	-0,02918%	0,01349%	-0,05865%	-0,03578%	-0,04730%	-0,04715%
1502	06/05/2010	0,02072%	-0,24192%	0,03802%	-0,01684%	0,00178%	0,02112%	0,00525%	0,00424%
1503	07/05/2010	-4,83887%	0,02306%	-0,07582%	-0,00833%	-0,06105%	-0,06920%	-0,07744%	-0,07872%
1504	10/05/2010	4,95622%	0,22189%	0,13075%	0,02648%	0,19308%	0,19247%	0,19215%	0,19009%
1505	11/05/2010	-0,03450%	-0,00099%	-0,04252%	0,00224%	0,03451%	0,02346%	0,01945%	0,01801%
1506	12/05/2010	-0,02072%	-0,17357%	0,00091%	-0,00293%	0,05362%	0,02157%	-0,02261%	-0,03392%
1507	13/05/2010	0,00000%	-0,06802%	-0,04596%	-0,02497%	-0,01667%	-0,01977%	-0,02872%	-0,03588%
1508	14/05/2010	-4,56968%	0,02483%	0,04081%	0,01147%	-0,00663%	-0,02034%	-0,02920%	-0,02445%
1509	17/05/2010	4,53515%	-0,00701%	0,02556%	-0,01405%	0,00594%	-0,01088%	0,00333%	0,00159%
1510	18/05/2010	0,00000%	0,00000%	-0,04191%	0,03025%	0,03122%	0,03103%	0,04998%	0,07272%
1511	19/05/2010	-0,04141%	-0,00079%	0,02030%	-0,01416%	-0,06977%	-0,07321%	-0,06340%	-0,06947%
1512	20/05/2010	0,04831%	-0,10003%	-0,07709%	-0,00352%	-0,10718%	-0,09339%	-0,10186%	-0,10268%
1513	21/05/2010	-4,99760%	0,10211%	0,07827%	0,05473%	0,01743%	0,01364%	0,01139%	0,01603%
1514	24/05/2010	4,92166%	-0,00261%	0,26669%	-0,01230%	0,05416%	0,03681%	0,03733%	0,03549%
1515	25/05/2010	0,55919%	-0,18744%	-0,28038%	-0,01631%	-0,04362%	-0,01059%	0,00159%	0,01818%
1516	26/05/2010	0,02762%	0,05010%	0,07116%	0,03578%	0,09052%	0,08366%	0,10543%	0,09249%
1517	27/05/2010	-0,06903%	0,01041%	-0,04012%	0,01182%	0,10242%	0,12413%	0,11811%	0,12055%
1518	28/05/2010	-3,86572%	-0,07339%	0,00033%	-0,04344%	0,00369%	-0,00224%	-0,00294%	0,00966%
1519	31/05/2010	3,86572%	0,06161%	0,02101%	0,00612%	0,00318%	0,00266%	0,00330%	0,00159%
1520	01/06/2010	-0,20711%	-0,02433%	-0,06229%	-0,00216%	-0,02178%	-0,02596%	-0,01914%	-0,02051%
1521	02/06/2010	-2,41612%	-0,08189%	-0,00726%	-0,02568%	0,00360%	0,03335%	0,02856%	-0,00355%
1522	03/06/2010	2,27804%	0,03008%	0,00693%	0,01717%	0,01059%	0,01041%	0,01013%	0,01008%
1523	04/06/2010	-4,76295%	0,03344%	0,02863%	-0,00033%	-0,06678%	-0,05274%	-0,05583%	-0,06372%
1524	07/06/2010	4,73535%	0,03546%	0,05747%	0,00873%	-0,00616%	-0,00676%	-0,00618%	-0,00791%
1525	08/06/2010	0,06212%	0,00131%	-0,01858%	0,00585%	0,00965%	0,00087%	0,01215%	0,02286%
1526	09/06/2010	0,10356%	0,07667%	-0,04060%	0,00049%	-0,00544%	0,01206%	0,01591%	0,01577%
1527	10/06/2010	-0,10356%	-0,04431%	0,03450%	-0,01385%	0,00282%	0,00851%	0,01674%	0,01790%
1528	11/06/2010	-4,46616%	0,06385%	-0,02929%	0,04364%	0,01271%	0,01004%	0,00681%	0,01124%

Graph 3.4 *Knowing the data*

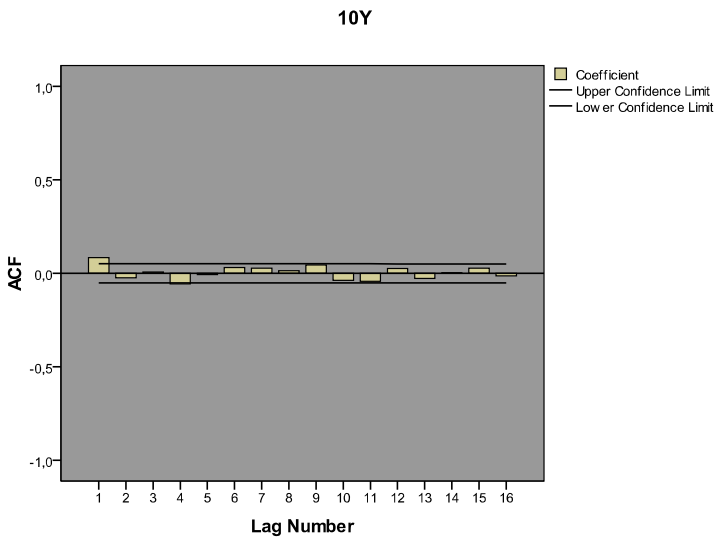
We see the graphics of one of his series:



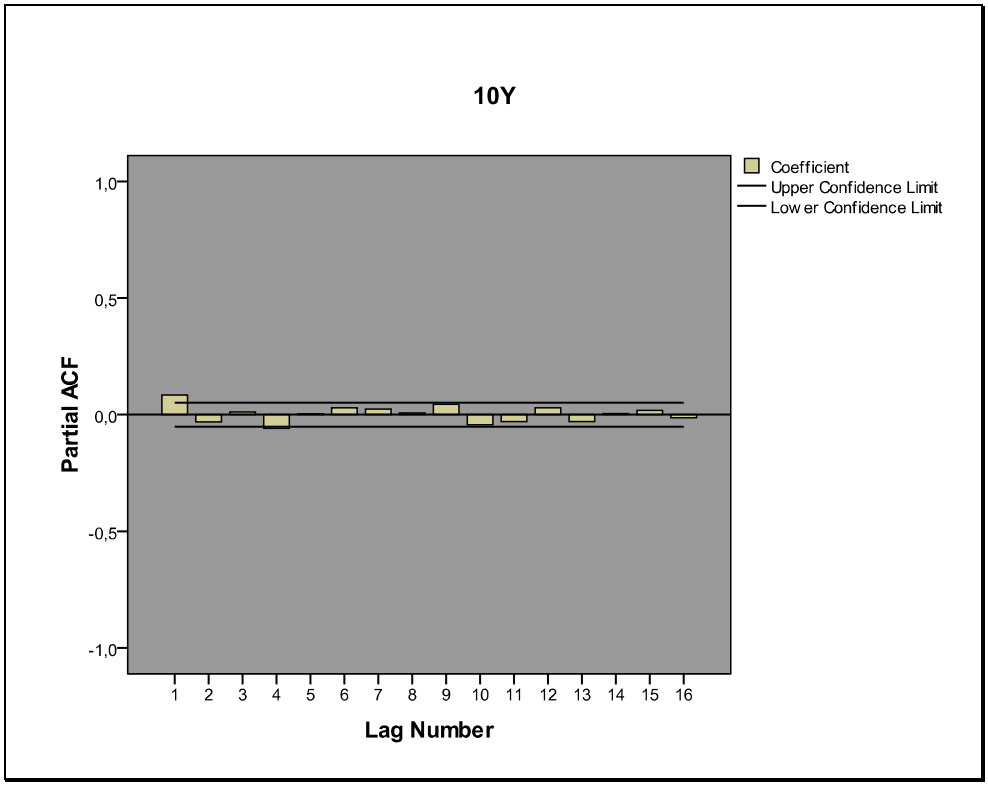
Graph 3.5 Time series

It is noted that the trend has gone, and the series still lacks homoscedastic and seasonality, so we already have stationary.

Now, let's focus on the autocorrelation functions of this time-serie:



Graph 3.6 *Simple autocorrelation*



Graph 3.7 *Partial autocorrelation*

Although we still can see a slightly significant autocorrelation on the first lag, the rest of the Partial Autocorrelation function shoes an improved behaviour, as we don't have any other significant autocorrelation. This serie is suitable to work with.

Part III

Models

Chapter 4

4. Theoretical model

4.1. Model Features

At this stage we need to choose a theoretical model that is useful for the purpose of the ongoing work. We need this model to have some features and properties that makes it suitable for our paper but that also keep things as simple as possible but not simpler than necessary.

Particularly, we want to consider stochastic models of interest rates with the following features:

- They have as few underlying stochastic factors as possible. As we will see later we can reduce our n-dimensional problem to a one-dimensional problem, relying on one single Brownian Motion to give the whole dynamic of the forward interest rate curve.
- They are consistent with absence of arbitrage opportunities ("there is no free lunch"). This will be achieved using martingales theory with the election of an appropiate numeraire.

- They can potentially accommodate any observed term structure of interest rates. We will like any observed shape to be consistent qith our model –upward sloping, downward sloping, humched-shaped, Nelson-Siegel...
- It is stationary

4.2. Heath-Jarrow-Morton Framework

The Heath–Jarrow–Morton theory ("HJM") is a general framework to model the evolution of interest rate curve - instantaneous forward rate curve in particular.

The key to these techniques is the recognition that the drifts of the no-arbitrage evolution of the instantaneous forwards can be expressed as functions of their volatilities, no drift estimation is needed.

The general parameterization of continuous stochastic evolution due to Single – factor HJM is:

$$\mathrm{d}f_t(T-t) = \epsilon_t(T-t)\mathrm{d}t + \sigma_t(T-t)\mathrm{d}W_t$$

For the purpose of modelling the bond market is suffcient for W_t to be a one dimensional Brownian Motion, but more exotic interest-rates products require higher-dimensional stochastic generators – particularly slope products that depends on the correlation structure of two different interest rates such as Swap1Y vs. CMS10Y.

As we have seen earlier, the short-rate, rt, is the rate aplicable to an infinitesemal tenor period and it defines a proccess Bt called cash account or savings account by means of

$$\frac{\mathrm{d}B_t}{B_t} = r_t \mathrm{d}t$$

Which approximates the result of earning the overnight interest rate on the amount B_t.

So the condition of absence of arbitrage ("no free lunch") mentioned earlier can be shown to the following.

First, there is a certain change of probability measure which can be expressed by saying that:

$$\mathrm{d}\tilde{W}_t = \gamma_t \mathrm{d}t + \mathrm{d}W_t$$

Is normalised centered Brown motion with respect to the new probability measure.

We call the old measure (making Wt a centered Brown motion) the "theoretical" measure, and the new measure (making $\mathrm{d}\tilde{W}_t$ a centered Brown motion) the market measure. The process γt can be interpreted as the price of risk

The discounted present value of any tradable security is a martingale with respect to market measure. So, the fact that discounted Zero-Coupon bond prices are martingale with respect to the market measure implies that the drift $\epsilon_t(\tau)$ is completely determined by the volatility structure $\sigma_t(\tau)$, as we anticipated earlier.

Choosing the Cash account B_t, the no-arbitrage condition implies:

$$\mathrm{d}f_t(T-t) = \sigma_t(T-t)[\Sigma_t(T-t)\mathrm{d}t + \mathrm{d}\tilde{W}_t]$$

Where

$$\Sigma_t(\tau) := \int_0^\tau \sigma_t(\upsilon)\mathrm{d}\upsilon$$

is the log-volatility of the process of the discounted bond price $P_t(t+\tau) = \frac{P_t(t+\tau)}{B_\tau}$

Such a model will be stationary if $\sigma_t(\tau)$ is, in fact, not just any random function, but depends on t only through a possible explicit dependence on $f_t(\tau)$ itself, for instance:

$$\Sigma_t(\tau) := \int_0^\tau \sigma_t(\upsilon)\mathrm{d}\upsilon$$

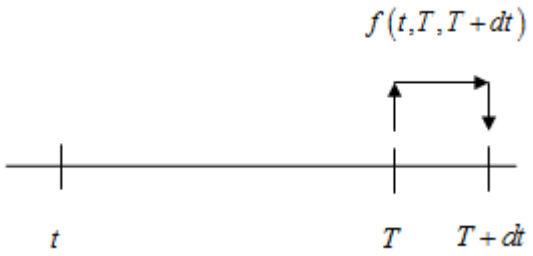
Now that we have a intenally arbitrage free model to work with, we have to analize how this framework get along with the data approach conducted.

Chapter 5

5. Practical model

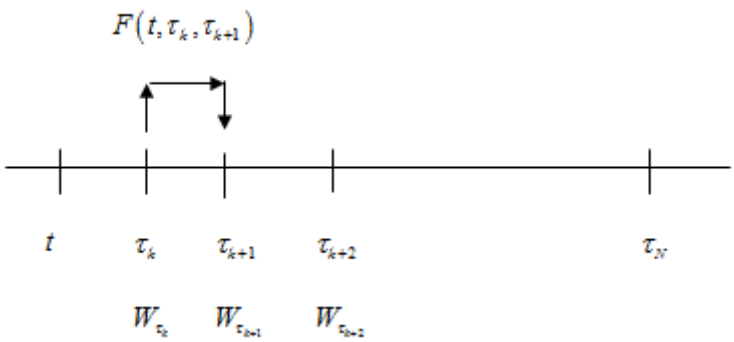
5.1. Analysis of the forward rates

The HJM model focuses in the evotion of the instantaneous forward rate. That is, the forward rate we are looking at in time t for an infinitesimal time period T – T+dT.



The problem we are facing is that instantenous forward rate is very convenient theoretical instrument but it cannot be recovered from our data as we are dealing with time-series for finite and fixed tenors. –O/N, 1M, 6M, 2Y.....

So we are constraint to study the forward rates applying between two consecutive tenors:



A new problem arises from this approach. Each forward is driven by a different Brownian motion, so that we have as many stochastic variables as tenors, so that we have a n-dimensional problem that could be intractable in practice.

The way to come through this difficulty is to conduct a Principal Component sAnalysis so that we can keep one single Brownian motion to give dynamics to the whole forward curve what would allow us to build a discretised HJM model with no arbitrage.

5.2. Principal Components Analysis.

Principal Component Analysis (PCA): We have built a row forwards for the synthetic bonds, we have 18 curves therefore forwards, each of which follows a different stochastic process with Brownian.

Our goal is now, through appropriate statistical techniques (Principal Component Analysis) to reduce the number of variables as possible, preferably to a single variable that explains the largest possible percentage of the total variance of 18 variables.

By reducing the number of variables 18-1, we loose some information, but instead we perform a univariate model where there is only one element of uncertainty (18 Brownian pass in the full model to a single Brownian to perform principal components).

The principal component analysis can be performed on the matrix of variance / covariance or correlation matrix, in our case we have done on the correlation matrix of forwards.

Each principal component is a linear combination of the initial variables, with the peculiarity that the ordered according to the eigenvalues of the correlation matrix. The first principal component corresponds to the largest eigenvalue of the correlation matrix.

As discussed above, the main components are a linear combination of the initial variables of the problem and the first principal component has the following form:

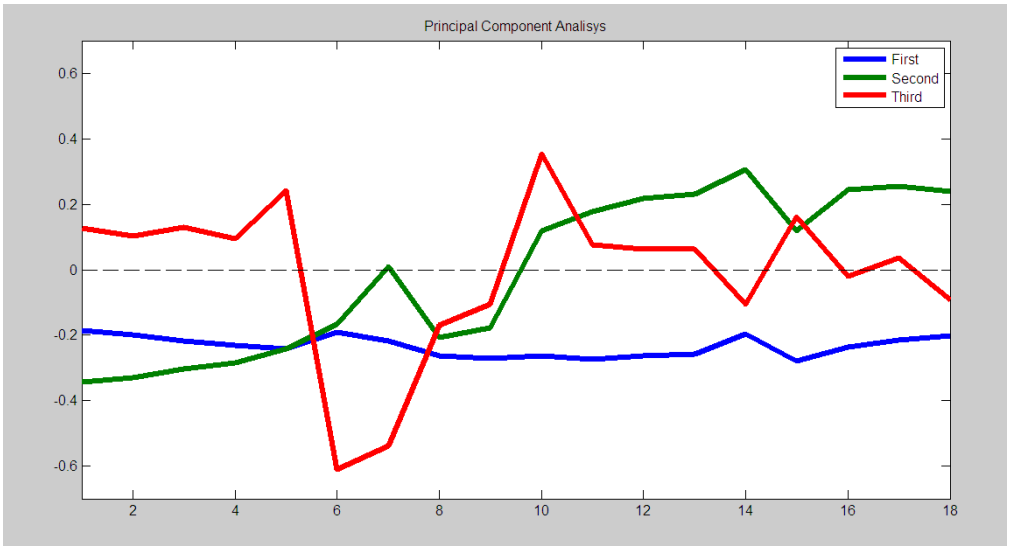
$$-0{,}184F_1 -0{,}199F_2-0{,}217F_3-0{,}231F_4-0{,}241F_5-0{,}191F_6-0{,}218F_7-$$
$$-0{,}264F_8-0{,}270F_9-0{,}263F_{10}-0{,}274F_{11}-0{,}262F_{12}-0{,}257F_{13}-0{,}196F_{14}$$
$$-0{,}279F_{15}-0{,}237F_{16}-0{,}214F_{17}-0{,}201F_{18}$$

Where is the el forward rate i (are the variables on which we make the PCA).

The main advantage of this type of analysis is that we manage to reduce the number of variables losing the least amount of information possible to explain 100% of the variance would need just 18 major components, but with only the first 3 is normal to explain about 95% of the total variance of the original data.

In our case, the principal component analysis explained how 79% of the variance with the first factor, 11% with the second and 4% in the third, getting thus explaining 94% of the total variance with only three variables.

Here are the top 3 as main components to analyze the curve forwards data inherited from our synthetic bonds:

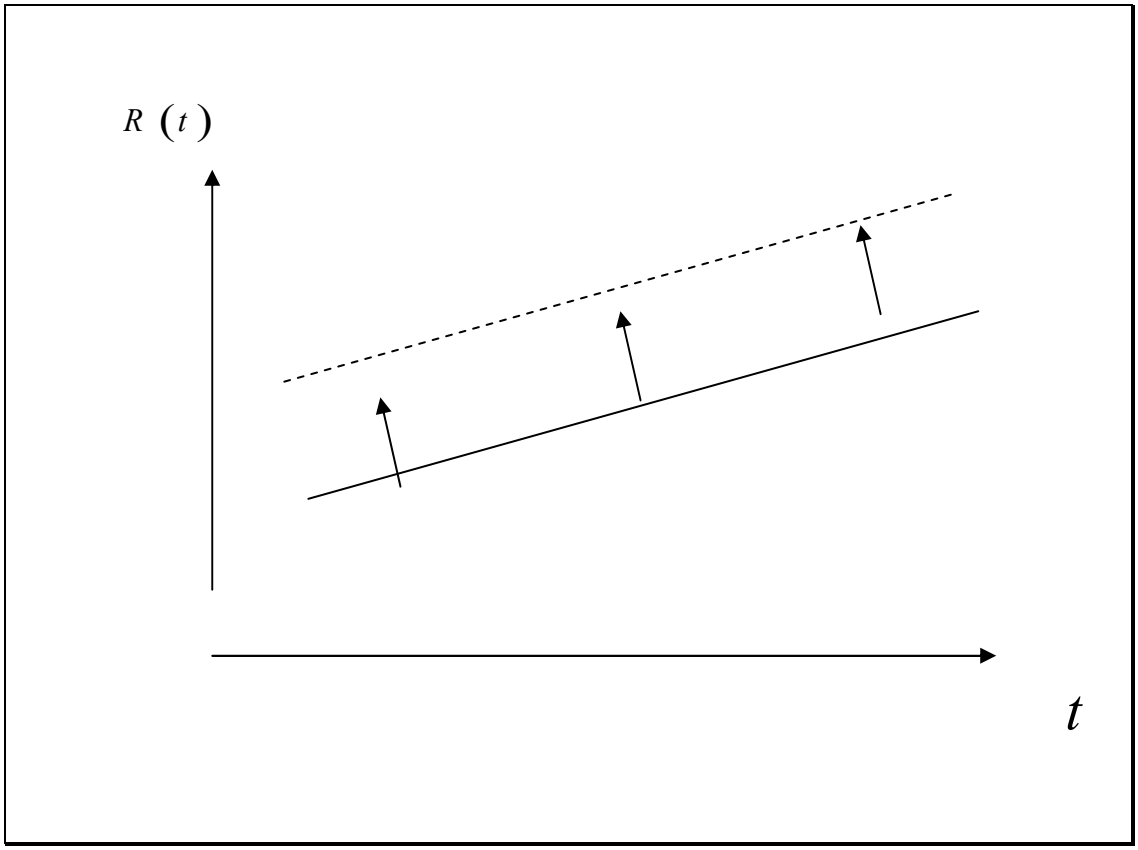


Graph 5.1 *PCA*

When we do a principal components analysis on an interest rate curve (in this case, on a bend forward rates curve), we find an empirical sense to first three principal components.

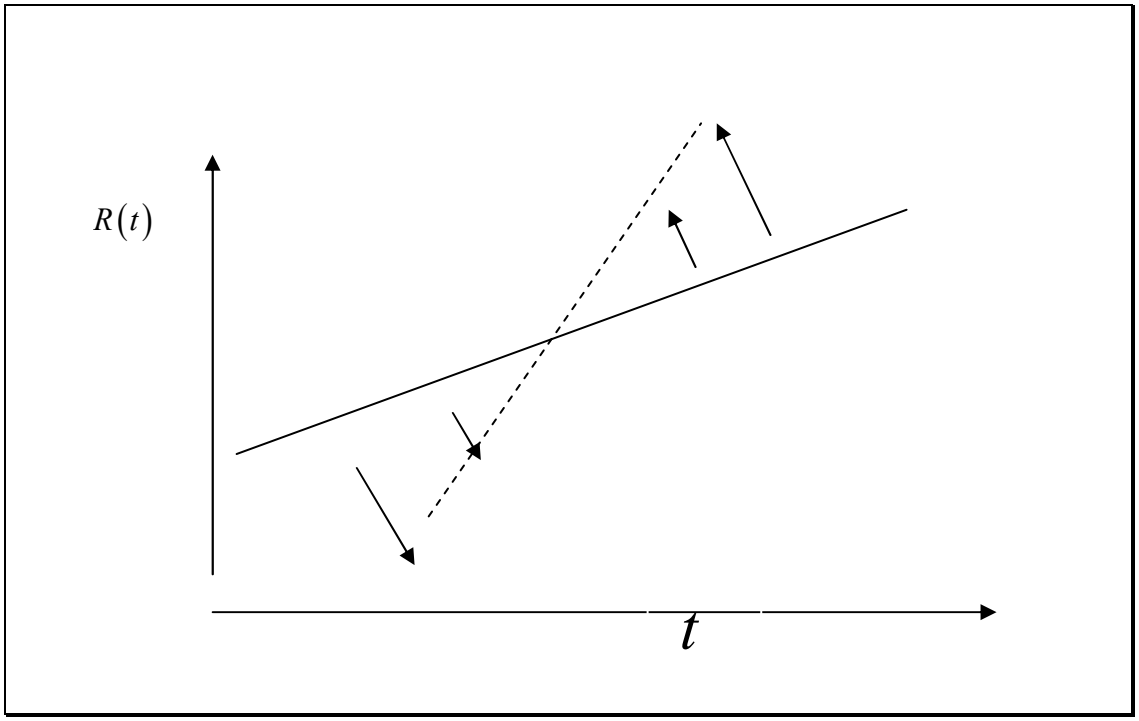
The first principal component explains parallel movements in the curve (which contained almost 80% of the variability of the curve). This type of movement correspond to increases or decreases in the rates for all maturities (forwards), we can see that this is reflected in all the coefficients of the first principal component have the same sign, if for example the first factor be increased by a point, then all would move forwards in parallel (all coefficients having the same sign).

Primer Factor- parallel shifts



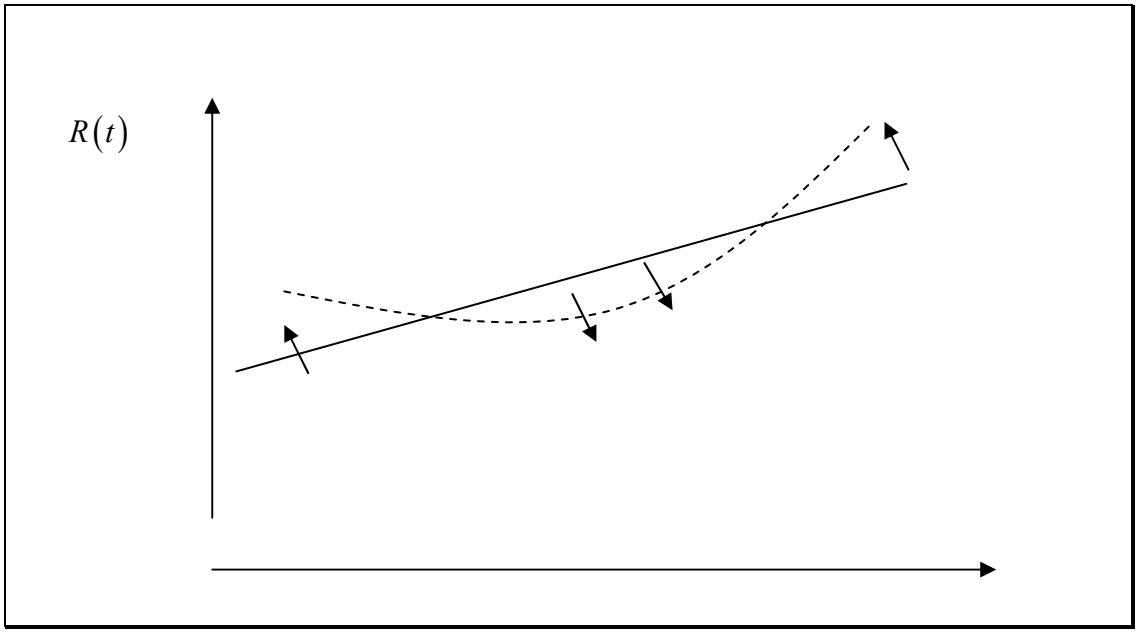
Graph 5.2 *First Factor- parallel shifts*

The second component mainly explains changes in the slope of the curve, such as a rise in short-term forwards and a fall in long-term forwards, which would produce a flattening of the curve (flat curve). We see this patent on the second principal component that forwards have negative short-and long-forwards have positive sign, with a variation on this major component would cause short-term rates to move in one direction and the types over the counter.



Graph 5.3 *Second Factor- Steepening*

The third main component is going to explain movements in the curvature of the curve (convex curve). These movements are less common and refers to similar movements in the short-term rates and long term, and movements in the opposite direction to the medium-term rates. We can also see this in our third main component, as the sign for short-term rates is the same as for long-types, while medium-term rates move in opposite direction.

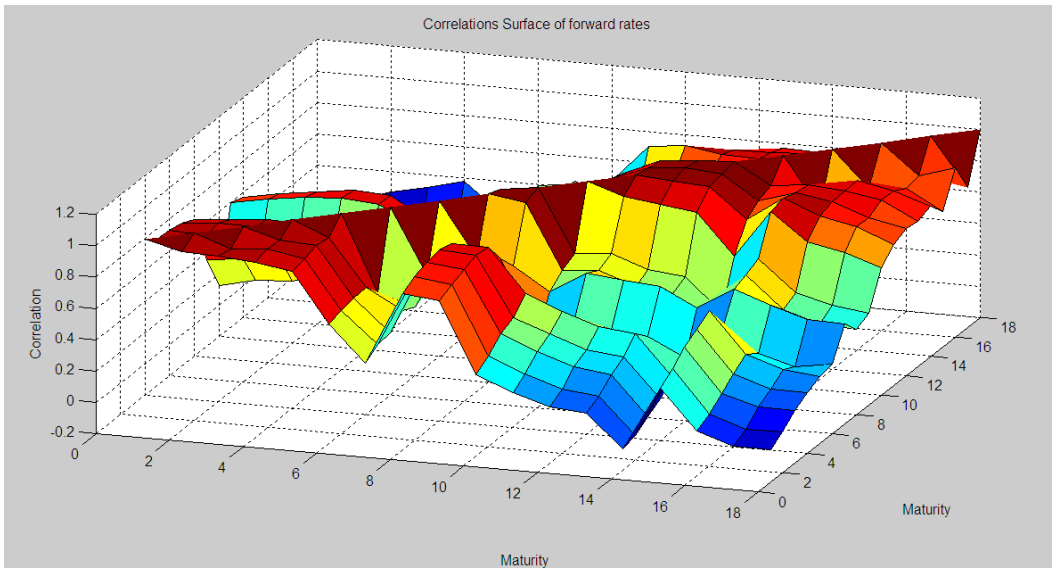


Graph 5.4 *Third Factor- Bending*

Here we see the correlation matrix of row forwards. We see among the forwards in the short term there is a fairly high correlation, and as among the forwards in the long run there is a high correlation. We can also observe how the correlation between short and long range decreases as the individual being more distant maturities.

5.3. The correlation matrix

This correlation matrix structure is entirely consistent with the market data we can see, short-term rates move in a very similar, long rates as well and there is some discrepancy between the movements in the short and long term.



Graph 5.5 *The Arbitrage Free model for the synthetic forwards*

It results that imposing the HJM condition to our synthetics assets to obtain the following arbitrage-free one-factor HJM model for our data.

$$dF(t,\tau,\tau') = \frac{\Sigma_t(\tau) - \Sigma_t(\tau')}{\tau' - \tau}\left[\left\{\gamma_t - \frac{1}{2}\left(\Sigma_t(\tau) + \Sigma_t(\tau')\right)\right\}dt + dW_t\right]$$

The model has only one risk factor as there is only one Brownin Motion involved. It can be used to risk neutral valuation setting or to model the evolution of real prices if we want to calclate P&L, VaR....In these cases an estimation of the market price of risk, , is needed.

Part IV

Conclusions

Chapter 6

6. Conclusions

The work achieved in this paper results in the construction of a one-factor interest rate model from a data-first approach. To this end, the relevant variables have been taken into account through a thoghtful statistical analysis –autocorrelations, normal tests, descriptive statistics, Principal Components analysis.

Also, the resulting model satisfies the desirable property been internally free of arbitrage, so is a valid version of the HJM Model to price interest rate derivatives –with the appropriate drft adjustments.

Further discussions can be developed on Back-testing of fitted model: comparison with historical data using descriptive statistics, both quantitative and qualitative.

It would also be interesting to conduct forescasting and scenario analysis with fitted model: simulation of future histories (Montecarlo Simulation); estimating price and risk (VaR) of simple interest-rate instruments or portfolios.

Ultimatelly, the possibilty to enrich this one-factor with a second risk factor is also a interesting line of future work. Moreover, a two-dactor model would allow us to study complex derivative depending on the behaviour of two different tenor such as slope products.

Chapter 7

7. Bibliography

- Attilio Meucci, Risk and Asset Allocation
- Baxter & Rennie, Financial Calculus
- Chatfield. The Analysis of Time Series, an Introduction
- Chatfield & Collins, Introduction to multivariate analysis
- Fabozzi. Bond Markets, Analysis and Strategies
- Filipovic. Consistency Problems for Heath-Jarrow-Morton Interest Rate Models
- Hull. Options, Futures and Other Derivatives