Yield Curve Modeling under Cyclical Influence

Mats Levander

January 2010

Abstract

There are few models for long-term yield forecasting and especially in the Real World Measure. This thesis uses Principal Component Analysis (PCA) to analyze the yield curves and gives an update of precedent studies. The conclusion is still that the first three components is enough to describe the variation of the yield curve. For simulation of the yield curves PCA and a semi parametric approach are evaluated. These models fail to yield plausible simulations. Therefore a new model is developed. The model is influenced by a business cycle and a relationship is derived from historical data between the yield curve and the cycle. Nelson-Siegel's model and Ornstein-Uhlenbeck processes are some of the features used in the model. Tools are also introduced to cope with the non-negativity problems in the aftermath of the financial crisis. The model will be used as input to calculate the Swap Breakage Exposure upon an early termination of an interest rate swap.

Acknowledgements

I would like to thank Nils Hallerström at PK AirFinance for giving me this opportunity and supporting me by sharing his wide knowledge of the financial markets. I would also like to thank the whole PK AirFinance Office in Luxembourg for baring with me during this period, and especially Christophe Beaubron, Frank Withofs and Bo Persson for guidance and discussions. Finally, I would like to thank my supervisor at KTH, Henrik Hult, for his support and ideas during the course of this thesis.

Stockholm, January 2010

Mats Levander

Contents

Chapter 1

Introduction

1.1 PK AirFinance

PK AirFinance started in 1979 when the Swedish bank PKbanken set up a subsidiary, PK Finans, in Luxemburg. In 1983 PK Finans made their first aircraft financing - a DC-10-30 to Martinair. A couple of years later a separate Air Group was established. Crédit Lyonnais bought 85% of the Air Group in 1988 and it was renamed to Crédit Lyonnais/PK AirFinance. Under the following years the company expanded, e.g. it set up offices in New York and Tokyo, started an Asset Management Group and a first version of a deal analysis model was launched. In 1991 Crédit Lyonnais bought the remaining 15% of Air group.

In 2000 the company was bought by GECAS (General Electric Commercial Aviation Services) and renamed to PK AirFinance. An office was set up in Toulouse in 2001 because of the proximity with Airbus. As of today PK AirFinance is a leading provider and arranger of asset based loans that are secured by commercial jet aircraft. The company's customers are airlines, aircraft traders, lessors, investors, financial institutions and manufacturers worldwide.

PK AirFinance's business activity is to; *invest for its own account in aircraft backed loans and securities.* This is done by providing or purchasing; senior[1] loans or finance leases, junior loans, and minor equity investments in special purpose companies that own aircraft and guarantees.

The parent company, GECAS, is the world's largest aircraft lessor. PK AirFinance has tailor-made financing deals involving more than 75 Airlines worldwide. At the end of 2008 the company's portfolio volume reached USD 7.7 billion. As of today PK AirFinance has 15 employees. The deal

[1] a senior loan has higher priority than a junior loan in a repayment structure

analysis model, SAFE[2], was created in 1993. This tool has been subject to continuous refinement during the years.

1.2 SAFE

PK AirFinance (PK Air) uses SAFE to analyze, price, underwrite and to continuously monitor deals throughout their term in the portfolio. The model is made up by several modules that contribute to the output of the model. The model has a wide range of applicability, e.g. for determining risk effectiveness, value creation and portfolio management. The graphical user interface is in an Excel environment supported by C++ and databases. In this interface PK Air can alter many different parameters in the analysis of new or existing deals. Examples of parameters are jurisdiction, margins, type of lease, aircraft type and many others. Some of these parameters are analyzed in different subroutines. For example they model the future aircraft value, the probability of default of the obligor and the industrial cycle in subroutines.

To put it short, the model boils down to calculation of the distribution of the Net Present Value (NPV) of a deal. In this calculation PK Air generates many different scenarios of the outcome of the NPV. The risk in the deals is when the NPV turns negative and PK Air will experience losses. To value different deals PK Air uses different metrics such as Risk Reward Ratio (Expected Loss divided by Expected Present Value), Average Downside Risk (the average of the negative outcomes in the NPV scenarios) and Value At Risk (VaR).

The fundamentals of PK Air's Asset Backed Financing are

- Default Risk
- Asset Value
- Deal Structure

Default Risk is the risk that the airline will default (not being able to meet its payment obligations). The probability of default is linked to the credit ratings of the airline. Another thing that affects the probability of default is the Aircraft Value Cycle which is explained thoroughly in Section 1.3, briefly it can be said to adjust for the economical cycles.

Asset value risk is the risk that is attached to the uncertainty of an aircraft's future value. PK Air has found a way of predicting the aircraft's

[2]Statistical Aircraft Finance Evaluation

future values. This model is also adjusted for the Aircraft Value Cycle and is in addition affected by a projected inflation path.

The deal structure is the way the deal is set up. This describes the outstanding loan amount or book value over the tenor, fees, rent and interest rate margins, priority ranking over the collateral (aircraft, deposits, guarantees etc.), rights to upside in the aircraft, and swap breakage exposure.

There are numerous variations of the these structures, and when the counter party defaults the pay-off function of the lease or the loan will depend on how the deal was set up. This pay-off function is a function of the asset's resale price.

Different deal structure examples are now presented, and the pay-offs are illustrated under the assumption that the borrower/lessee immediately defaults and that the loans are non-recourse[3].

- Lease of aircraft bought for USD 20mm
- Senior loan USD 20mm
- Shared Pari-Passu loan USD 10mm each
- Senior Shared loan USD 15mm and a Junior Shared loan USD 5mm

First we have the lease contract. If the lessor buys an aircraft for USD 20mm and leases it to an airline the pay-off function takes the shape of *Lease USD 20mm* in Figure 1.1. Upon default the lessor repossess the aircraft and may sell it on the market since they own it. Therefore the lessor can at most experience a loss of USD 20mm but the upside is unlimited in theory.

A Senior loan is conducted by lending USD 20mm to the airline to finance their aircraft with the aircraft as collateral. If the airline defaults, they have to sell the aircraft in the market to repay their debt. The best scenario is that the lender gets his USD 20mm back, and the worst is loosing all USD 20mm. The lender is not entitled to any upside since they do not own any aircraft in this deal.

If two lenders enter a shared Pari-Passu[4] loan, lending out USD 10mm each to the airline. This results in the payoff function of the Senior loan being "halfed", e.g. if the aircraft is sold for USD 10mm both will experience losses of USD 5mm.

[3] non-recourse means that you can only go after the aircraft to recover your money, not the owner

[4] Pari-Passu is latin and may be translated as "with equal force", for a loan it means that you have equal recovery rights

In a shared Senior-Junior loan settlement the pay-off shape changes. If we have the Senior loan of say USD 15mm and the other lessor has USD 5mm as Junior, we will only experience losses after a sale price of less than USD 15mm, since the Junior tranche experiences the first loss of USD 5mm. If we instead have the Junior tranche the pay-off will have the shape of *Junior USD 5mm*.

There is a vast amount of hybrids of these kind of structures but these lie outside of the scope of this thesis.

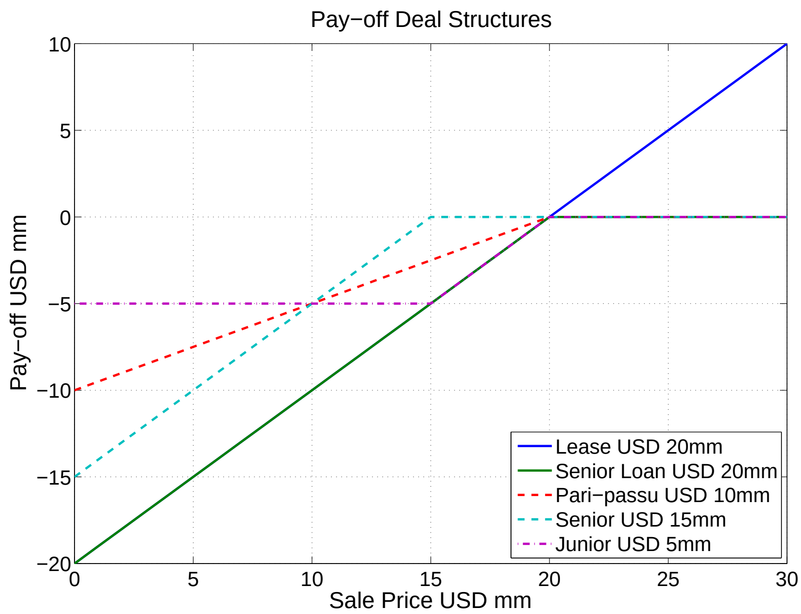

Figure 1.1: Examples of Deal Structure Pay-offs upon default

Another part of the deal structure is the Repayment Schedule. It is the agreement on how the borrower will repay its debt to the lender, i.e. the points in time and the magnitude of the cash flows from the borrower to the lender. As an example a fictitious schedule is shown in Figure 1.2 (additional fees and margins has been left out in this example). The total outstanding debt is 25 million USD that PK Air has lent to a borrower. To finance this a lender would typically take a loan of 25 million USD in the floating rate market, and to hedge themselves they will swap a fixed rate for floating matching the cash flows. The fixed rate is calculated as an effective rate so that the net present value of the cash flows is zero, see Section 3.6

for an example. The repayment schedule is as follows: the borrower repays 0.5 million USD every 6 months and after 15.5 years he pays the remaining residual of 10 million USD. PK Air has in turn matched this repayment setting with their lender.

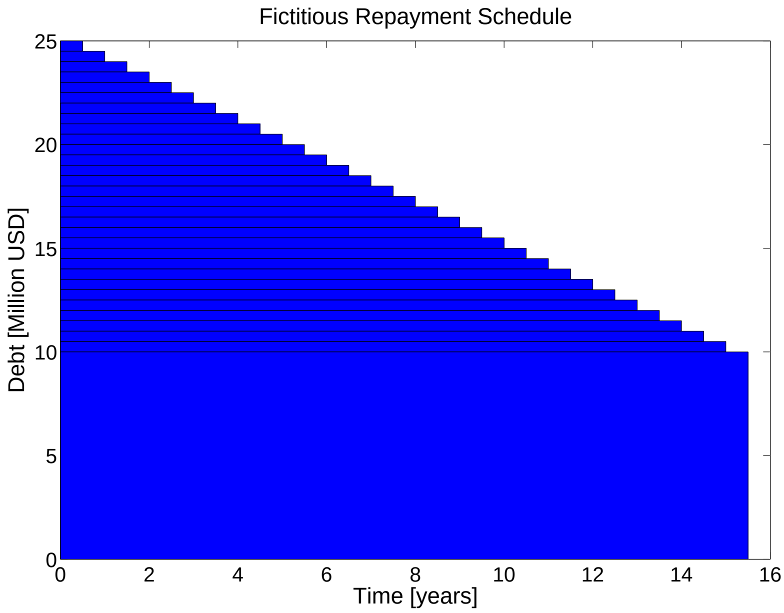

Figure 1.2: Fictitious Repayment Schedule

If PK Air's borrower defaults or there is an early repayment option of the loan PK Air will not receive the planned cash flows. Instead they have to finance their upcoming repayments with the present market setting, that is with the yield curve at that time. Depending on the yield curve, there might be a gain or loss from this refinancing procedure when the net present value and the corresponding fixed rate is calculated. Therefore the Exposure at Default is uncertain and dependent on the propagation of the yield curve. A fictitious illustration of repayment schedule outcome simulations with percentiles is shown in Figure 1.3. It shows that the uncertainty is greatest in the middle of the repayment period if we are at time zero today. In the beginning and in the end of period the uncertainty diminishes to become certain at the endpoints.

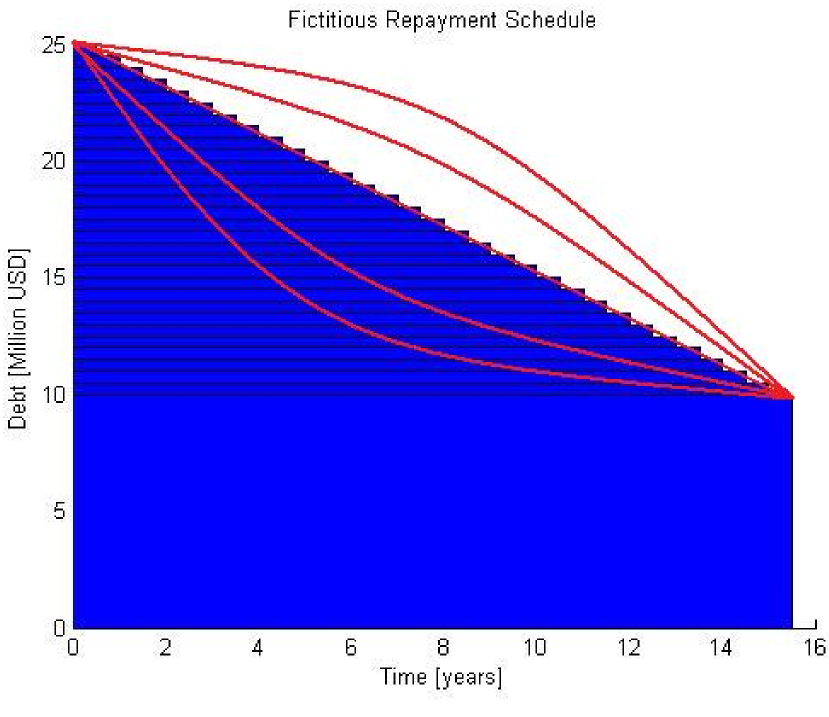

Figure 1.3: Fictitious Repayment Schedule Simulation

At today's date PK Air uses a model that uses a worst case scenario, the yield curve simulation decays exponentially with time. PK Air is not pleased with this way of modeling and wants to improve this simulation. They wish to have a model that is tractable, has few parameters, is fast to simulate and it should be influenced by the cycle described in the next Section.

1.3 The Aircraft Value Cycle

1.3.1 Introduction

The common view is that the Aircraft business is cyclical. For example the delivery of aircraft, aircraft orders, air traffic and many other factors vary cyclically with time. The same thing holds for aircraft values. The outcome of buying, selling, leasing and/or financing aircraft will be very dependent on the timing of an investment or divestment compared to the aircraft value cycle.

1.3.2 Historical evidence

PK Air has built up a database with historical resale prices of aircraft, which at the moment contains around 4000 sales transactions since 1970. By looking at specific data for each aircraft type PK Air has been able to analyze depreciation patterns, cyclicality and volatility statistically. In Figure 1.4 the resale price of the Boeing 737 300 series expressed in constant 1984 USD value is shown. The pattern around the trend line is cyclical.

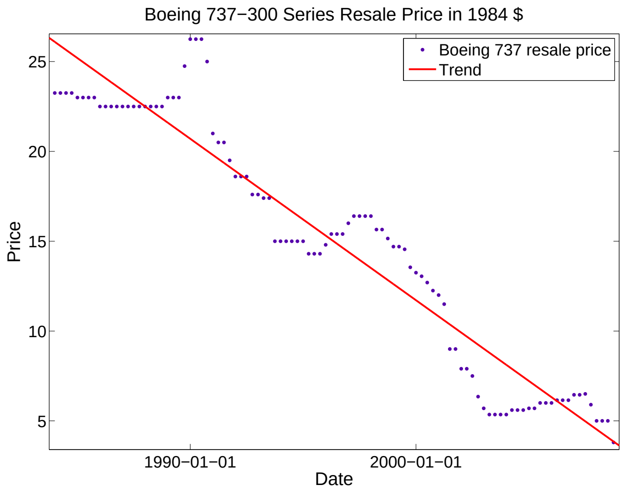

Figure 1.4: Boeing 737-300 Resale Price in constant 1984 USD over time

PK Air has from their broad study of different aircraft type values found out that the cyclical variation is almost perfectly correlated in time between different types.

1.3.3 Why this behavior?

Classic economic theory is built upon the interaction between supply and demand. When the demand for an asset is higher than the supply the price will rise to find an equilibrium and vice versa. The supply and demand mechanisms for aircraft seems not to be in equilibrium for most of the time. The explanation to this is the lag when the aircraft manufacturing industry tries to adapt to the current market situation. Their pace of manufacturing is directly dependent on the predictions of future traffic made by the airlines. The manufacturing time is normally 12-18 months. This explains the adjustment lag; by the time of delivery of the aircraft, the market might have changed and therefore the predictions and so on. This makes the traffic very hard to predict in the short term.

The demand for aircraft is estimated by air traffic, which can be expressed as revenue passengers miles per year. The relationship between production (revenue passenger miles; RPM) for one aircraft, load factor LF, block speed V, utilisation hours U and number of seats S is defined as:

$$RPM = LF \cdot V \cdot U \cdot S \tag{1.1}$$

The time series of the RPM growth are shown in Figure 1.5

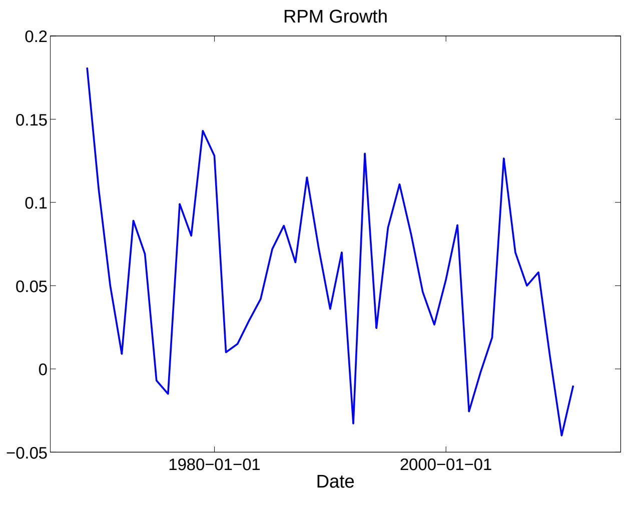

Figure 1.5: RPM Growth 1968 to 2009

This relationship also holds for airlines and the global air traffic market. Explanation of the parameters:

1. The load factor has been improved due to management system and new scheduling techniques. On the one hand a high load is good for

the airline's profit, on the other hand opportunity costs may arise from high load factors. For example the airline might have to reject passengers in peak times and vice versa if there is a trough.

2. The block speed was improved greatly during the 1960's when the aircraft obtained pressurised cabins, and later by first the turbine engines and then the jet engines. Today the block speed is considered as constant.

3. Utilisation has improved over the years due to longer range aircraft and optimized scheduling. There are limiting factors though, in general people do not fancy flights departing or arriving in the middle of the night or early morning. On average the utilisation is rising since the long haul traffic is increasing faster than the short haul traffic. The reason is that long haul planes spends more hours per day than the short haul planes.

For the short term airlines are capable to adjust load factor and utilisation to shifts in demand and supply. There is a however a point in time when they reach the maximum of these parameters, and they must therefore expand by buying more seats or the other way around if the times are tough. Fortunately, aircraft are globally traded assets and may therefore easily be transfered between operators.

1.3.4 Amplitude of Cyclical Swings

In Figure 1.6 the annual growth of physical installed seats (delivered less scrapped) is shown and the cumulative difference between RPM and $Seat$ $growth$ adjusted for the trend is shown in Figure 1.7.

We can see the cyclical behavior in the graph which is the driving process for the aircraft value swings. Amplitudes greater than zero is a measure of the pent-up relative capacity shortage in the global aircraft fleet, and an amplitude less than zero describes a pent-up surplus. Let us denote this measure pent-up relative capacity shortage/surplus ($PURCS$). In the peaks, airlines have high load factors and the aircraft is in use many hours per day. There will be a demand for aircraft so prices will go up. In troughs, the converse holds.

1.3.5 Prediction of the cycle

To know whether to invest or not we need an prediction about the future. The ($PURCS$) cycle is affected by; traffic growth and seat growth. If we

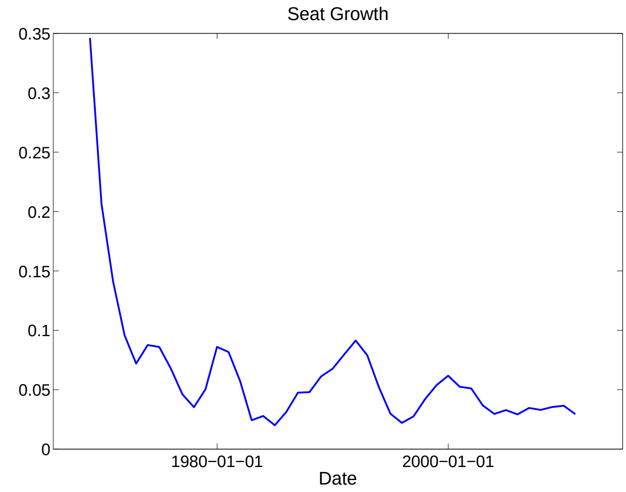

Figure 1.6: Seat Growth 1968 to 2009

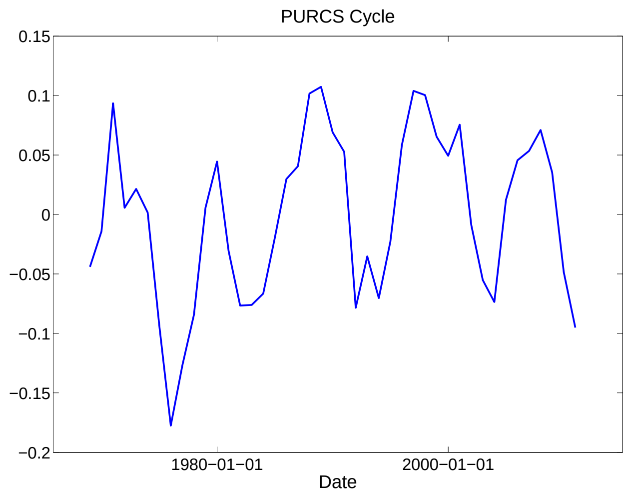

Figure 1.7: PURCS Cycle 1968 to 2009

could predict those for the time horizon of a cyclical swing (approximately 5 years) we would know when the peaks or troughs appear.

Looking at the RPM growth in Figure 1.5 above we could see that the volatility is high. The notoriously hard predicted GDP Growth drives the short term variations. For prediction PK Air has two in-house teams, which consists of members from both the parent company GECAS and PK Air. The first team is responsible for predicting the RPM Growth by predicting the parameters in equation (1.1). The second team predicts Seat Growth that is defined as

$$Seat \quad Growth = F + P - C - R \tag{1.2}$$

where F is current fleet seats, P is produced seats, C is converted seats (origins from aircraft being converted from passenger to cargo aircraft (the converse does not happen)) and R is retired seats.

In the simulations in the model, RPM Growth and Seat Growth are not simulated independently, instead PK Air starts out in the latest known peak or trough and simulates the coming peaks and troughs. In the simulation there are stochastic variables representing the uncertainty in the occurrence (time) and the level of the next peak or trough. The predictions are revised and the PURCS Cycle is updated and improved on a quarterly basis.

Chapter 2

Literature Study

There is an enormous amount of models and research on interest rate modeling. Most of the papers focus on short rate models, e.g. Hull-White, Ho-Lee, Vasicěk and Heath Jarrow Morton (HJM) and Cox Ingersoll Ross (CIR). All these models is priced under the risk neutral measure. The parameters in the models (eg. reversion speed, jump frequencies) implicitly embed, after being calibrated to market prices, a component related to risk aversion and therefore contain drift terms different from the objective measure. The drift terms creates a gap between the risk-neutral and the real world evolution of the yield curve [17]. Using these short rate models in long horizons results in evolutions that becomes totally dominated by the no-arbitrage drift terms and do hardly possess any virtual resemblance to the real-world evolution. For example Vasicěk and CIR cannot even recover an arbitrary exogenous yield curve, and the market price of risk is specified (proportional or proportional to the square root of the short rate) have been chosen to allow analytical tractability.
Another problem with these models is that they are often supposed to be used with short time horizons such as days or weeks, e.g. for banks that trade with interest rates. PK Air's typical deals have an average life of around 9 years, i.e. the time horizon is much longer than for banks. In addition their simulations are done in the real world measure, not in a risk-neutral measure, and they price the risk at the end of their simulations. To limit the scope of this study and to adapt to PK Air's time horizon, I have chosen in agreement with my academic supervisor and PK Air, not to look at the short rate models. Instead the focus is on Principal Component Analysis (PCA) and curve fitting techniques to understand and model the movements of the yield curve.

Litterman and Scheinkman [11] investigated the common factors that affect the returns on U.S. government bonds and similar securities. They compared duration analysis with PCA for hedging purposes. The duration concept is used to describe parallel shifts of the yield curve, but in real life, the yields do not always move in parallel shifts. With PCA they found three factors affecting the returns of the bonds and that they explained 96 % of the variation. The factors are referred to as the level, the steepness and the curvature. The first factor explained 89.5% of the variation, the steepness 8.5% and the curvature 2.0%. They concluded that this PCA representation is better than traditional duration hedging.

Gloria Soto [21] evaluated the performance and stability of PCA to explain term structure movements. Comparisons were done with typical one- and two-factor interest rate models. She focused on the Spanish government debt market from January 1992 to December 1999. The result was that the PCA with with three factors (level, slope, curvature) outperformed the one- and two-factor models. There were however some concerns with the stability of the PCA over time. The model's performance deteriorated significantly when the model is estimated from the most recent data in contrast to long-sample data. This points towards the use of factor models with dynamic volatility structures.

Scherer and Avellaneda [20] used PCA to study the Brady Bond Debt of Argentina, Brazil, Mexico and Venezuela from 1994 to 1999. They found that the two first principal components explains up to 90% of the variation. The analysis was split up into windows to show how the components vary with time and tried to link those to different economic events. Because of the short series of data, they looked at daily yield changes, since their results with weekly or monthly observations were dominated by noise and lacked structure. For the analysis they looked at absolute yield changes. They also verified that the results were invariant to valuation basis, that is if they e.g. used relative changes. Their opinion was that PCA on long observation windows was best for mature markets with stable economic cycles, which is compatible with the mathematical definition of stationarity. This was not the case for this market and they therefore used 120-day windows. They conclude that there was a high risk for under-estimating the behavior of the market if considering only a static PCA over the whole period.

Wesley Phoa [16] also compared traditional interest rate management such as duration management with PCA on the US Treasury market. The PCA was performed with daily changes in yield on Constant Maturity Treasuries. He pointed out that the result will only be meaningful if a consistent set of yields is used and therefore he used Constant Maturity Treasury yields. An alternative approach was to use historical swap rates, since these are par yields by definition. The results said that the most important shifts of the yield curve were the level and the slope. These two shifts can be estimated quite precisely and robust over time. The curvature shifts tended to be more varying and were highly dependent on the dataset used. When using shorter term to maturities the variance explained by the level diminished a little and the slope and curvature increased slightly. The importance of the three components for bonds was also shown to be similar for all large countries e.g. the U.S., Germany, France, UK, Australia and Japan. He also pointed out that some risk factors are ignored with PCA, for example if there would be a shift in the curve for maturities from 30 to 100 years relative to shorter yields. This would not show up in the PCA since the range only goes to 30 years. Another yield curve risk that does not show up using Constant Maturity Treasuries is the the yield spread between liquid and illiquid assets. This was a large factor in the US Treasury market in 1998 that was greater than the curvature factor.

Rebonato, Mahal, Joshi, Buchholz and Nyholm [17] evolved a method to construct a future yield curve over a period of years using a simple semi-parametric approach. They used PCA to analyze percentage changes of USD LIBOR and Swap yields from 29th of September 1993 to 4th of December 2001. They found that the first component explains more than 90% of the variation. To model future yields they tried different models. The first one, referred to as the *naive* simply used historical simulation to evolve future yields. The drawing of historical yields, was done by randomly choosing a starting date and then randomly choosing a window of consecutive yields that should be included. In this manner a yield curve was evolved. This procedure preserved the co-dependence structure of the original data across maturities. But they found that the assumption that the increments were IID did not hold. The method fails to capture the serial autocorrelation structure and the cross-sectional co-dependence. The difference from IID behavior was so strong that the procedure was unsatisfactory. To cope with this they introduced mean-reversion for the shortest and longest maturities and also introduced "springs" for maturities in between. The springs were introduced to match the curvatures of the empirical data. The last refinement consisted of a jump frequency in the sampling windows, that is a small probability of jumping out of the sampling window and start a new sampling. By these enhancements they managed to replicate the statistics of the empirical data very well. They claimed that one possible application of the model is off-balance sheet transaction, e.g. swaps, which exposes the counter-party to potential future credit exposure, which is dependent of the deal will be in the money at the time of a possible default.

Rebonato and Nyholm [18] extends their paper from 2005 by comparing their semi-parametric approach with a parametric approach for long-horizon yield curve forecasts. The parametric approach is a model by Bernadell, Coche and Nyholm [3] that is a three-factor representation of observed yields built upon the Nelson and Siegel methodology [14]. They found that their semi-parametric model discovers the same features as the parametric model without the use of Markov chains. The data focused on was US Treasury yields with maturities; 3 months, 6 months, 1 years, 2 years, 5 years, 10 years, 20 years and 30 years from 1st of January 1986 to 7th of May 2004 at a daily frequency.

Fiori and Iannotti [7] used Principal Component Analysis and Monte Carlo Simulation to assess Italian banks interest rate exposure. They looked at the Euro Area par yield curve with maturities from 1 month to 30 years from 4th of January 1999 to 30th of September 2003. The first three components explained 95% of the total variation. Instead of simulating from the normal distribution, which have been proved to be empirical incorrect, they focused on heavy-tailed distributions. In particular they used an Gaussian kernel estimator with optimal bandwidth. The distribution functions of these were obtained by integrating the kernel densities. From these Value-a-Risk calculations were made. They compared the parametric simulation (Normal) with the non-parametric (Gaussian kernel). The results showed that the parametric approach has limitations when the interest rates are increasing, especially when interest rates were low as in December 2001.

Duffee [6] investigated the idiosyncratic variation of Treasury Bill Yields and found that Treasury bills of three month or less term to maturity exhibits price movements that are idiosyncratic, i.e. they are not related to changes in other interest rates. He believed that this is caused by an increased market segmentation.

Nelson and Siegel [14] introduced a simple and perspicuous model for the shapes of the yield curve. They applied it to US Treasury bill and Treasury bonds yields. The parametric model was successful in representing typical yield curve shapes: monotonic, humped and S-shaped. They found that the model explained 96 % of the variation in bill yields across maturities from 1981 to 1983. The model is explained more thoroughly in Section 3.17.

Bank of Canada investigated the Nelson Siegel and Svensson parametric yield curve models [5]. They investigated the optimization problem, the robustness and data filtering applied to Canadian Government Securities. They found that the Svensson model was the best model.

Chapter 3

Theoretical Background

3.1 Inflation

Inflation is the rate of change in general prices over time [12]. It can be expressed in terms of an inflation rate i. The prices 1 year in the future can on average be expressed as today's prices multiplied by a factor $(1 + i)$. The inflation compounds in a similar way to interest rates, that is after k years at the same inflation rate i, the future prices will be the original prices times $(1 + i)^k$. In reality the inflation rate changes over time.

Inflation can also be seen as it reduces the purchasing power of money. A classic example is that a dollar today does not yield as much bread and milk as a dollar did 10 years ago. Therefore we can think of the prices as increasing or the value of money as decreasing. If we assume that the inflation rate today is i, then the value of a dollar next year expressed in today's purchasing power will be $1/(1 + i)$.

When conducting a study over a certain time period, it can be very convenient to express prices in the same kind of dollars. Therefore we consider constant (real) dollars, which are defined relative to a reference year (typically the starting year). This yields that the dollars will maintain the purchasing power of the dollars for the reference year over time. These dollars should be distinguished from the actual (nominal) dollars that are used in daily transactions.

From this we are able to define a new interest rate, namely the real interest rate, defined as the rate that real dollars increase if they were put into a bank account that pays nominal interest rate. To illustrate this, think of putting money into a bank account at time zero and then withdrawing them after a year. The purchasing power of the bank balance can be expressed as

$$1 + r_0 = \frac{1 + r}{1 + i}, \tag{3.1}$$

where r is the nominal interest rate, i is the inflation rate and r_0 is the real interest rate. This equation tells us that the money in the bank grows with $(1 + r)$ nominally but it is at the same time deflated by $\frac{1}{1+i}$. Rearranging the expression we obtain

$$r_0 = \frac{r - i}{1 + i}. \tag{3.2}$$

For small inflation levels the real rate is approximately equal to the nominal rate minus the inflation rate.

3.2 Present Value

The idea behind the concept of Present Value (PV) [12] is to describe the time value of money. If you invest 1 USD at the bank today at a rate r_f (consider as the risk-free rate) you money will be worth $1 \cdot e^{r_f \cdot 1}$ after one year using continuous compounding. On the other hand 1 USD recieved one year from now is worth $1/\left(e^{r_f \cdot 1}\right)$ USD today which is also referred to as the present value. The procedure of transforming future cash-flows into today's value is called discounting.

3.2.1 Compounding of rates

There are two ways of compounding rates; frequent and continous compounding. If we assume that the risk-free is r_f and the interest is compounded at m equally spaced times per year. We do also assume that the cash flows start out at time zero and at the end of each period, and that the total number of periods is n. Then we obtain cashflows as follows $(x_0, x_1, ..., x_n)$ and this yields the following formula for the present value

$$PV = \sum_{k=0}^{n} \frac{x_k}{[1 + r/m]^k} \tag{3.3}$$

If we instead assume that the interest is compounded continuously and that the cash flows occur at times $t_0, t_1, ..., t_n$. Then the formula boils down to

$$PV = \sum_{k=0}^{n} x\left(t_k\right) e^{-rt_k} \tag{3.4}$$

3.3 Forward Rates

Forward rates [12] are interest rates for money to be borrowed between two future time points, but the terms are agreed upon today. If we for example consider a 2-year horizon. Assume that the spot rates s_1 and s_2 are known and given on a yearly basis. After two years 1 USD would have grown to $(1+s_2)^2$ USD in a 2-year account. One other option is to first put the money in a 1-year account for one year, which will yield $(1+s_1)$ after one year. If we then already have agreed to borrow the money for 1 more year at a rate f, then we will recieve $(1+s_1)(1+f)$ at the end of year 2. f is called the forward rate. These two ways of borrowing money should be equal if there is no arbitrage and therefore we get the following equation

$$(1+s_2)^2 = (1+s_1)(1+f) \tag{3.5}$$

which can be solved for f

$$f = \frac{(1+s_2)^2}{(1+s_1)} - 1 \tag{3.6}$$

as can be seen the forward rate is determined by the two spot rates s_1 and s_2.

3.4 Yield

The yield [12] of a bond is the rate of interest that is implied by the payment structure. More precise it is the interest rate for which the present value of the bond's cash flow streams (coupons and face value) is exactly equal to its current price. The more formal term of yield is **yield to maturity** (YTM) to separate it from other yield measures. Yields are always quoted on annual basis. The yield is exactly the internal rate of return of a bond at its current price, but yield is the term used in the market. The calculation for the yield, y, of a bond maturing at time T using continuous compounding is

$$PV = F \cdot e^{-y(T-t)} + \sum_{i=1}^{N} C_i \cdot e^{-y(T-t_i)}, \tag{3.7}$$

where F is the face value of the bond, N is the number of coupons and C_i is the coupon paid at time t_i. Equation (3.7) must be solved for y to determine its yield. It can only be solved by hand for very simple cases, otherwise it has to be done by an iterative process, which can easily be implemented on a computer.

3.4.1 Yield Curve

The general conditions in the fixed-income securities is closely related to the yield of a bond [12]. In general all yields move together in this market, but all yields are not exactly the same. The reason for the variation in yield over bonds is that they have different credit quality ratings. An AAA-bond (highest rating) has in general a higher price and thus has a lower yield than a B-bond with the same maturity. There are however other explanations to the differences in bond yields. Time to maturity is another factor that affects the yield of a bond. Intuitively bonds with long time to maturity tend to offer higher yields than bonds with shorter time to maturity with the same rating. Therefore a "normally" shaped yield curve has higher yield with longer time to maturity. It does in fact happen that bonds with longer time to maturity have lower yields than bonds with shorter time to maturity, this is termed *inverted yield curve*. This shape appears when the short-term rates increases rapidly, but the investors believe that the increase is only temporary, and therefore the long-term rates (which is an expectation about the future) remain close to their former level.

3.5 Interest Rate Swap

A swap [10] is a an agreement between two counter-parties to exchange cash flows according to a prearranged formula in the future. The first swaps were settled in 1981 and since then the market has grown fast. Many billions of dollars of swap contracts are negotiated each year.

The "plain vanilla" interest rate swap is the most common one. The setup is that one party, A, agrees to pay another party, B, cash flows of a predetermined fixed rate on a given notional principal at specific points in time during an interval. A the same time party B pays cash flows to party A, based on a floating rate on the given notional amount over the same time period. The currency is the same for the two interest rate cash flows. The swaps have duration for 2 to 15 years.

Swaps appear because companies have comparative advantages. One company may have an advantage in the fixed rate market and the other in the floating market. The companies tend to borrow money in the market they have an comparative advantage. However they might want to match a cash flow in the other market, this is where the swap is used to transform the rate. The floating rate used in these agreements is typically the LIBOR rate, see Section 3.7 for further details.

Financial institutions have a natural advantage funding themselves in the floating rate market. This implies that a interest rate swap occurs when they want to match for example a deal with an airline who wants to borrow money with a fixed rate. Then PK Air turns to their lender and matches a floating rate with a fixed rate that is offered to the airline. All of PK Air's interest rate swaps agreements so far has been in US Dollars.

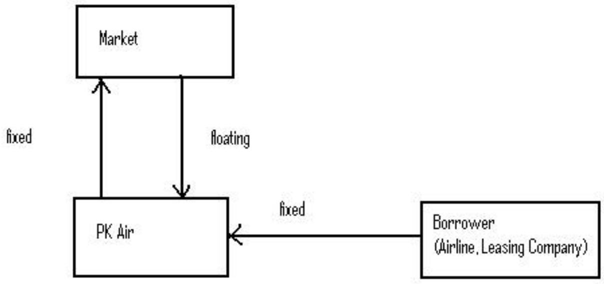

Figure 3.1: Interest Rate Swap

3.6 Swap Breakage

Swap breakage occurs when a swap agreement is terminated early, for PK Air there could be two causes of this; one is an early repayment and the other is a default of the counter party. The fixed rates are of course at the same level throughout the agreement. The floating rate will naturally change over time. The fixed rate was calculated from the yield curve at time $t_0 = 0$. On the time of default $t_{default}$, the yield curve will look different and the floating rates would imply another fixed rate for the remaining term. In the present value calculation there is a high probability that the risk-free rate that is used for discounting has also changed. The present value can be both positive and negative and of course also zero, resulting in an income, cost or zero result breakage.

3.6.1 Example of Swap Breakage

Say that we want to swap a floating rate for a fixed rate for a time interval of 5 years and that the swaps are made on a yearly basis. The swap is on a notional amount, N, of 50 000 USD. Assume that today's yield curve has the following setting:

Table 3.1: Fictional Yield Curve

Term to maturity	1 yr	2 yr	3 yr	4 yr	5yr
rate	0.0058	0.0124	0.0186	0.0233	0.0270

where the rates are compounded on a yearly basis. From this we can calculate the implied forward rate f_{ij} between year i and j. For the second year this is

$$(1 + 0.0124)^2 = (1 + 0.0058)(1 + f_{12})$$

if solved for f_{12}

$$f_{12} = \frac{(1 + 0.0124)^2}{(1 + 0.0058)} - 1 = 0.0190$$

in the same manner we can calculate all forward rates.

Table 3.2: Implied Forward Rates

interval	0-1 yr	1-2 yr	2-3 yr	3-4 yr	4-5 yr
forward rate	0.0058	0.0190	0.0312	0.0375	0.0419

To calculate the fixed rate, we calculate the Present Value (PV) *in thousands USD* of the cashflow streams and we set the PV to zero and solve for r_{fixed}

$$50\left((0.0058 - r_{fixed})\, e^{-0.0058 \cdot 1} + ... + (0.0419 - r_{fixed})\, e^{-0.0270 \cdot 5}\right) = 0$$

$$\Rightarrow r_{fixed} = \frac{0.0058 \cdot e^{-0.0058 \cdot 1} + ... + 0.0419 \cdot e^{-0.0270 \cdot 5}}{e^{-0.0058 \cdot 1} + ... + e^{-0.0270 \cdot 5}} = 0.0265$$

If we now assume that the agreement is terminated after 1 year and we have the following yield curve at that time

Table 3.3: New Fictional Yield Curve

Term to maturity	1 yr	2 yr	3 yr	4 yr
rate	0.0122	0.0152	0.0223	0.035

As before we back out the implied forward rates

Table 3.4: New implied Forward Rates

interval	0-1 yr	1-2 yr	2-3 yr	3-4 yr
forward rate	0.0122	0.0182	0.0367	0.0740

From this we obtain the following Present Value

$$PV = \big((0.0122 - 0.0265) \cdot e^{-0.0122 \cdot 1} + ... + (0.0740 - 0.0265) \cdot e^{-0.035 \cdot 4}\big) \cdot 50$$
$$= +1.7058$$

If the counter-party defaults we will have a Swap Breakage Gain of 1705.8 USD.

3.7 LIBOR rate

The information in this section was found at one of the British Bankers Association (BBA) webpages [2].

3.7.1 Historical Background

In the early 1980's an increasing number of banks in the London market were actively trading new instruments as currency options and interest rate swaps. Most of the banks found these new instruments attractive but they were at the same time bothered about that the underlying rates that had to be agreed on before entering the contract. The BBA were therefore asked by the banks they represented to create a uniform measure for the market and produce a benchmark index. The concept was that banks could now reference their contracts against a standard rate which removed the negotiation of the underlying rate. This improved the generation of instruments. In 1984 BBA cooperated with for example the Bank of England and others which resulted in the BBAIRS, that is the BBA standard for Interest Swap rates.

In this standard the fixing of BBA Interest Settlement rates, the predecessor of BBA-LIBOR, was included. BBAIRS became standard market practice the 2nd September 1985. The first BBA-LIBOR rates were given in 1986 and in three currencies; US Dollars, Japanese Yen and Sterling. Today it is calculated in 10 currencies for 15 maturities.

Since the BBA-LIBOR was introduced there has been two significant changes. Firstly in 1998 the question was changed from "At what rate do you think interbank term deposits will be offered by one prime bank to another prime bank for a reasonable market size today at 11 am?". The new question, which is still used today is "At what rate could you borrow funds, were you to do so by asking for and then accepting inter-bank offers in a reasonable market size just prior to 11 am?".

Secondly in 1999 with introduction of the Euro on the 1st of January 1999 the number of currencies BBA-LIBOR was calculated for obviously diminished.

3.7.2 BBA-LIBOR

BBA-LIBOR stands for London InterBank Offered Rate. The rate is derived for 10 currencies with 15 maturities for each, the shortest being overnight and the longest being 12 months. The BBA-LIBOR should be interpreted as a benchmark, it provides you with an idea of what the average rate of a prominent bank, for a specific currency, can get unsecured funding for a certain term in the specific currency. In other words, it yields the lowest real-world cost of unsecured funding in the London market.

The rates are derived by a calculation based on submissions from a panel, consisting of the most active banks in the specific currency.

3.7.3 Definition

It is important to know that BBA-LIBOR is based upon the offered rate, and **not** the bid rate. Each contributing bank is ask to base their BBA-LIBOR submission on the the question mentioned above: "At what rate could you borrow funds, were you to do so by asking for and then accepting inter-bank offers in a reasonable market size just prior to 11 am?". In this way the submissions are based upon the lowest perceived rate that a bank that participates in a certain currency panel could go into the inter-bank money market and get sizable funding for a certain maturity.

The BBA-LIBOR is not derived from actual transactions, it would be hard to have this as an requirement since not all banks require funds in marketable size each day for the currencies and maturities they quote. This should not be interpreted as that the rates does not reflect the true cost of interbank funding. The banks are aware of their credit and liquidity risk profiles for the rates it has dealt with, and are therefore able to build a curve that accurately predicts the correct rates for currencies or maturities which it has not quoted at the moment.

All BBA-LIBOR rates are quoted as annualized interest rates, which is a market convention. For an example say that the overnight rate is given as 2 %, this does not imply that that the overnight loan interest rate is 2 %. Instead it means that the annual loan interest rate is 2 %.

3.7.4 Applications

BBA-LIBOR is the primary benchmark for short term interest rates in the world. The rate is also used as an indicator of the strain in money markets and it is also often used as a measure of the market's expectations of future central banks interest rates. Different studies show that approximately 350 trillion USD of swaps and 10 trillion USD of loans are indexed to BBA-LIBOR. It is the foundation for settling interest rate contracts on most of the world's major futures and options exchanges.

3.7.5 Selection of panels

The contributing banks are chosen for currency panels with the goal of reflecting the market balance for a specific currency based upon these three guidelines

1. Scale of market activity
2. Credit rating
3. Perceived expertise in the currency concerned

Each of the 10 panels, with 8 to 16 contributors, is chosen by the independent Foreign Exchange and Money Markets Committee (FX & MM Committee) to obtain the best way of representing the activity in the London money market for a certain currency. Therefore BBA-LIBOR submissions from panel members will on average be the lowest interbank unsecured loan offers from the ones available on the money market.

The FX & MM Committee evaluates each panel every year, based upon a review from BBA of the contributors. The review evaluates each bank according to the criteria listed above. The review is not limited to only present contributors as any Bank can submit itself to the evaluation process for any currency.

3.7.6 Calculation

The calculations are made by Thomson Reuters. They analyze the data from the banks and calculate the rates according to the definitions given by FX & MM Committee. The whole procedure is supervised by BBA. In the contributing banks every cash desk has an application from Thomson Reuters installed. Each morning the currency dealer takes their own rates of the day and puts them in to the application between 11.00 and 11.20. The banks are not able to see the other banks rates until they are submitted by Thomson Reuters. The data is also submitted by nine other data vendors.

All the BBA-LIBOR rates from Thomson Reuters are calculated in the same way using an trimmed arithmetic mean. The contributors are ranked by Thomson Reuters and sorted in descending order. The trimming is done by dropping the top and bottom quartiles. The remaining 50 % of the data are then averaged to get the BBA-LIBOR quote. The reason to drop the top and bottom quartiles is to improve the accuracy of the quotes by removing outliers. By dropping these the possibility for any individual contributor to influence the quote is removed.

3.8 U.S. Interest Rate Swaps

These rates can be used as a proxy for the continuation of the LIBOR rates. These are quoted as the fixed rates swapped for 3 month LIBOR semiannually.

3.9 US Government Debt

3.9.1 Treasury Bills

The U.S. government offers debt obligations called Treasury bills [19] , also known as T-bills. Of the total U.S. government marketable debt the T-bills occupy approximately one-fourth. In the middle of 1995, there was around three-quarters of a trillion dollars worth of bills outstanding. Since the bills are liabilities of the government, these obligations are considered as default free. The secondary market for these are one of the most active, characterized by its low bid-ask spread and extremely high liquidity.

The T-bills are issued at discount from the face value and have no coupons or stated interest. The earnings for the holder is simply the difference between the discount issue price and the face value, which is paid out by the Treasury at time of maturity. The discount is determined in an auction, where new bills are offered to dealers and other investors. The auction is conducted by the Federal Reserve Bank of New York acting on behalf of the U.S. Treasury. The bills are sold to those who offer the highest price, and this auction ensures that the resulting interest costs of the Treasury is minimized. The discount is related to the term of the bond and what investors believe about the market future.

T-bills are offered in the following maturities: 91-day (3-month), 182-day(6-month) and 364-day(1-year). The three- and six-month bills are auctioned every week, and the one year every fourth week to meet the huge demand for funds from the U.S. government of refinancing an outstanding debt. Since 1993, the bills that have been issued in multiples of 1000 USD, with 10000 USD as the smallest amount. The amount which has the lowest commission rates is the one known as a round lot and is 5 million USD. These are normally traded by large market participants. The commission rates is generally in between 12.50 to 25 USD per 1 million USD and is affected by the maturity of the bill, three-months have the lowest commission.

No certificates or papers of the debt is issued when the T-bills are issued. The claim is only registered in the computer system of the Federal Reserve.

Due to their low risk and short maturity, T-bills are very popular instruments for market participants. The range of holders goes from individuals to governments and everything in between. For individuals and commercial investors the T-bills is interesting since they are relieved from state and local taxes. Foreign banks are large holder of bills, their interest is mostly based on the safe return of the bills. For the U.S. government the T-bills are an important way to raise funds to finance the U.S. debt outstanding. The Federal Reserve does also hold a lot of bills to be able to use its monetary instruments to influence the economy.

3.9.2 Treasury Bonds and Treasury Notes

Contrary to the T-bills, the Treasury bonds (T-bonds) and Treasury notes (T-notes) [19] are interest-bearing securities with maturities greater than one year. The difference between notes and bonds is that they are issued in different maturities. Notes are issued with maturities of 2, 3, 4, 5, 7 and 10 years and bonds with maturities greater then 10 years, e.g. 30 years. The primary market for these securities is a bit complicated. Almost all Treasury debt that is negotiable is offered at auctions on a yield basis. The speculators submit sealed bids with the lowest yield to maturity that they would accept, instead of regular bidding auction. The bids are conducted at local Federal Reserve Banks and they are opened at a prespecified time in the future. The bids specify the volume and the yield (in percent in two decimal places). There is also a possibility for small investors to place a noncompetitive bid, which means that they will pay a price equal to the mean of the accepted bids. The Treasury itself does intervene the process by setting "stop-out" levels, which is the highest level that they are willing to issue debt. At the prespecified times the average level of the accepted bids are calculated. Then the Treasury sets a coupon on the debt closest to 1/8 of 1 percent, where the level will yield an average price of the successful bids equal to 100 USD (per 100 USD of face value) or lower.

3.10 Smoothing Spline

If the available data set is very noisy, e.g. daily yields of a certain maturity, it might be a good idea to fit your data with a smoothing spline [13]. A smoothing spline, s, is constructed for a given smoothing level p and weights w_i. The objective of the spline is to minimize

$$p \sum_i w_i \left(y_i - s\left(x_i\right)\right)^2 + (1-p) \int \left(\frac{d^2 s}{dx^2}\right)^2 dx$$

if the weights, w_i, are unspecified all points are assigned a weight of 1. The smoothing level, p, is defined between 0 and 1. The lower limit produces a least-square straight line fit to the data. If you set it to the upper boundary it produces a cubic-spline interpolant. An application of this procedure is given in Section 5.1.3 later on.

3.11 Correlation

To get an idea of what correlation is, consider the joint density function of two random variables [1]. This density could be said to describe a mountain-like shape. If the mountain is symmetric around the two axes of the random variables, little information about one variable is obtained by knowing the value of the other variable, i.e. the correlation is low between the variables. On the other hand, if the shape of the joint density function is ridge-like between the axes, the correlation will be high.

To get a feeling for the correlation, scatter plots is useful. A scatter plot is a plot of two synchronous returns of financial time series against each other, e.g. gold returns against crude oil. A low correlation corresponds to symmetrical dispersed scatter plot, i.e. a high value of one variable does not imply a high/low value of the other. A high correlation is described by a ridge between the axes, if it has negative the slope the correlation is negative, and vice versa.

Correlation measures the co-movements between two return series. Strong positive correlation says that an upward movement in one return series is accompanied by an upward movement in the other series. For strong negative correlation an upward movement in one series corresponds to a downward movement in the other series.

The covariance measure can be used to measure the co-movements of two random variables X and Y, it is defined as

$$Cov\,(X,Y) = E\left[(X-\mu_X)\,(Y-\mu_Y)\right] \tag{3.8}$$

where μ_X and μ_Y are the means of the random variables. The drawback with this measure is that it is not only determined by the co-movements of the returns but also by the size of them, that is in general monthly returns will have higher covariance than daily returns. Since covariance is not independent of the units of measurement, it is not suitable to make comparisons. To cope with this the correlation measure is introduced. It is a standardized form of the covariance that it is independent of the units of measurement. For two random variables it is standardized by dividing the covariance by the product of their standard deviations, that is

$$Corr\,(X,Y) = \frac{Cov\,(X,Y)}{\sqrt{V\,(X)\cdot V\,(Y)}} \tag{3.9}$$

where $V\,(X)$ and $V\,(Y)$ are the variances. This standardization procedure will yields a correlation value of -1 to +1, where the first refers to perfect negative correlation and the latter to perfect positive correlation. If the two random variables are statistically independent their correlation coefficient should be insignificantly different from zero, and the variables are referred to as orthogonal. Notice that the converse is not necessarily true, if two random variables are orthogonal does not imply independence (they could have zero covariance but still be related by the higher moments of their joint density function).

3.11.1 Cross-Correlation

The autocorrelation function could be used to determine a sequence's structure in the time domain. Cross-correlation uses the same concept but instead of comparing a time shifted version of the signal with itself, it compares to different sequences with each other [9]. The cross-correlation function of two sequences $\{X_t\}$ and $\{Y_t\}$ is defined as

$$\phi_{X,Y}\,(h) = E\,(X_t Y_{t+h}) = \lim_{T\to\infty}\frac{1}{2T+1}\sum_{t=-T}^{T} X_t Y_{t+h} \tag{3.10}$$

The function is a statistical comparison of two sequences as a function of the time shift between them. It is very useful as a practical tool for determining timing differences between the sequences.

3.11.2 Cross-Correlation Coefficient

The Cross-Correlation Coefficient is closely linked to the usual correlation coefficient and is defined as

$$\ell_{X,Y}(h) = \frac{\phi_{X,Y}(h)}{\sqrt{\phi_{X,X}(h=0)\,\phi_{Y,Y}(h=0)}} \tag{3.11}$$

as with the traditional correlation coefficient, the value of the cross-correlation coefficient lies between -1 and 1. It can be used to determine for which lag h to sequences obtain the highest correlation.

3.12 Least Squares Method

Let x_1,...,x_n be outcomes of independent stochastic variables X_1,...X_n. The means of the stochastic variables are known except for an unknown parameter. Therefore it is assumed that $E(X_i) = \mu_i(\theta)$ for $i = 1, 2, ..., n$ where $\mu_1, \mu_2, ..., \mu_n$ are known and θ is an unknown parameter with parameter space Ω_Θ. This implies that $X_i = \mu_i(\theta) + \epsilon_i$, i.e. $X_i =$ "known function of θ" + "error". These errors are assumed to have zero means. In general their variances are assumed to be identical. Let

$$Q(\theta) = \sum_{i=1}^{n} [x_i - \mu_i(\theta)]^2 \tag{3.12}$$

be the sum of the observations deviations from the $\mu_i(\theta)$'s. The expression between the hard brackets are recognized as the error of observation i *if θ is the correct parameter value.* The method is to use the θ-value that minimizes the sum of squares as θ-estimate. To find this *Least Squares Estimate of θ* [4] the derivative $dQ(\theta)/dQ$ is calculated and set equal to zero to find the minimum.

A more general method can be applied if the distribution contains k unknown parameters $\theta_1, \theta_2, ..., \theta_k$. The expression for Q is now a function of these parameters

$$Q(\theta_1, \theta_2, ..., \theta_k) = \sum_{i=1}^{n} w_i [x_i - \mu_i(\theta_1, \theta_2, ..., \theta_k)]^2 \tag{3.13}$$

where $w_1, ..., w_n$ are weights that could be assigned to the observations. In general these are set to 1, which corresponds to equal uncertainty for all data. The Least Square Estimate is as before determined by minimization, e.g. by setting the partial derivatives of Q to zero and solving for θ_i.

3.13 Maximum Likelihood Method

The idea behind the Maximum Likelihood Method is to as an estimate of θ use the value that makes our sample data as likely as possible [4]. The value of θ will therefore naturally depend on our sample data and will be the Maximum Likelihood Estimate, θ^*_{obs} of θ.
Let x_1,...,x_n be outcomes of independent stochastic variables X_1,...X_n that has a distribution that depends on an unknown parameter θ with parameter space Ω_θ. In most cases X_1,...X_n are assumed to be independent and that they have the same distribution, but the method also works without these assumptions.

If X is a continuous stochastic variable it has density function $f(x;\theta)$, and in the discrete case it has probability mass function $p(x;\theta)$. If the sample is from a discrete distribution and all the $X_i's$ are independent, will the probability of obtaining $P(X_1 = x_1, ..., X_n = x_n)$ be

$$P(X_1 = x_1, ..., X_n = x_n) = P(X_1 = x_1) \cdots P(X_1 = x_1) \tag{3.14}$$

$$= p_{X_1}(x_1;\theta)\, p_{X_2}(x_2;\theta) \cdots p_{X_n}(x_n;\theta) \tag{3.15}$$

The Likelihood Function is defined as

$$L(\theta) = \begin{cases} P(X_1 = x_1, ..., X_n = x_n;\theta) & \text{(discrete)} \\ f_{X_1,...f_{X_n}}(x_1, ..., X_n = x_n;\theta) & \text{continuous} \end{cases}$$

The idea behind the Maximum Likelihood method is to let the argument θ take all values in Ω_θ and find the value of θ that maximizes the function. This value is called θ^*_{obs} and is the estimate. In the discrete case the idea behind the method is to maximize the probability of obtaining the sample data.

When maximizing $L(\theta)$ it is often convenient to maximize $lnL(\theta)$. Since the logarithm is a monotonic increasing function will both $L(\theta)$ and $lnL(\theta)$ have maximums at the same point. The Likelihood function is normally a product and therefore the logarithm yields a sum which facilitates the maximization which is done by taking the derivative and solving for it equal to zero.

3.14 Linear Regression

Linear regression models are formulated in this way

$$Y = \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k \tag{3.16}$$

On the left side is the *dependent variable*, Y, and on the right there are the *independent variables* $X_1, X2, ..., X_k$ [1]. The independent variables can also be called *explanatory variables*. The coefficients $\beta_1, \beta_2, ..., \beta_k$ are model parameters and they measure the influence of its corresponding independent variable on Y. X_1 is in general assumed to be equal to 1 so the model includes a constant β_1. The goal of the regression is to find estimates of the true parameter values and predictions of the dependent variable by using historical data on the dependent and independent variables.

3.14.1 The Simple Linear Model

The simplest case is when $k = 2$ and $X_1 = 1$, i.e. we have a constant for all t. For notational purposes the constant is set to α (interpreted as the intercept with the vertical axis), and β_2 is denote by β (the slope of the line) and X_2 is denoted X. This results in the following setting

$$Y_t = \alpha + \beta X_t \tag{3.17}$$

To see the relationship between the data all pairs of (X_t, Y_t) is plotted against each other in a scatter plot. All the points will not lie along a straight line so an error process is introduced to the equation

$$Y_t = \alpha + \beta X_t + \epsilon_t \tag{3.18}$$

An estimated straight line through the scatter plot yield a predicted or fitted value of Y_t for each X_t from

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t \tag{3.19}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the estimates of the intercept and the slope. The difference between the real value of Y and the fitted value Y at time t is denoted ϵ_t and is called the *residual* at time t, i.e. $\epsilon_t = Y_t - \hat{Y}_t$. So each real data point Y_t is described by

$$\hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t + \epsilon_t \tag{3.20}$$

The estimates are obtained by minimizing the sum of the squares of the residuals, which is known as *ordinary least squares (OLS) criterion.* OLS is an unbiased estimation.

The OLS estimates for the Simple Linear Model are given by

$$\hat{\beta} = \sum_t \frac{\left(X_t - \bar{X}\right)\left(Y_t - \bar{Y}\right)}{\sum_t \left(X_t - \bar{X}\right)^2} \tag{3.21}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \tag{3.22}$$

where $\bar{Y}$ and $\bar{X}$ denotes the sample means of X and Y.

3.14.2 Multivariate Models

In the multivariate case the model takes the form

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + ... + \beta_k X_{kt} + \epsilon_t \tag{3.23}$$

and assuming that the model has a constant term yields $X_{1t} = 1$ for all $t = 1, ..., T$. There will be T equations with k unknown parameters. To simplify things the model can be written in matrix notation. Let the dependent variable column vector be $\mathbf{y} = (Y_1, Y_2, ..., Y_T)'$ and put the independent variables into a matrix X where the jth column of X corresponds to the data on X_j. The first column of X will be a column of 1s if there is a constant in the model. Denote the vector of true parameters $\beta = (\beta_1, \beta_2, ..., \beta_k)'$ and let $\epsilon = (\epsilon_1, \epsilon_2, ..., \epsilon_T)'$ be the vector of error terms. The representation of the model then boils down to

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon \tag{3.24}$$

The matrix form of the OLS estimators of β is given by

$$\hat{\beta} = \left(\mathbf{X}'\mathbf{X}\right)^{\mathbf{-1}} \mathbf{X}'\mathbf{y} \tag{3.25}$$

3.14.3 Goodness of Fit

There are numerous ways of measuring how well the model describes the data. Some measures are presented below.

3.14.3.1 R^2 and adjusted R^2

R^2 is the coefficient of determination of a model. It is a product of the concept of analysis of variance (ANOVA). It has three key metrics;the total sum of squares, $TSS = (\mathbf{y} - \bar{\mathbf{y}})^{\mathbf{T}} (\mathbf{y} - \bar{\mathbf{y}})$, the explained sum of squares, $ESS = (\hat{\mathbf{y}} - \bar{\mathbf{y}})^{\mathbf{T}} (\hat{\mathbf{y}} - \bar{\mathbf{y}})$ and the residual sum of squares, $RSS = \epsilon^{\mathbf{T}}\epsilon$. R^2 is the proportion of the total sum of squares that is explained by the model

$$R^2 = \frac{TSS}{ESS} = 1 - \frac{RSS}{TSS} \tag{3.26}$$

From this measure it is possible to tell how much of the variation in $\mathbf{y}$ that is explained by $\beta\mathbf{X}$.

Adjusted R^2 takes into account the number of explanatory variables in the model. Different from R^2, R^2_{adj} increases only if the new term improves the model more than expected by chance. The measure can be negative and it is less or equal to R^2. The definition is

$$R^2_{adj} = 1 - \frac{RSS/\left(T - k - 2\right)}{TSS/\left(T - 1\right)} \tag{3.27}$$

where T is the sample size and k is the number of explanatory variables. R^2_{adj} is in general used for determining if an explanatory variable improves the model

3.15 Principal Component Analysis

This section follows the derivation in the book by Alexander [1]. Before the analysis can be done, the data must be stationary. This is done by taking the absolute difference from day to day in the $T \times K$ rate data matrix $\mathbf{X}$. Then you normalize your data matrix $\mathbf{X}$ by subtracting each column with its mean μ and standard deviation σ to obtain a new matrix $\mathbf{X}^*$ with mean 0 and standard deviation 1.
Principal Components Analysis (PCA) starts out from the symmetric correlations matrix of the variables in $\mathbf{X}^*$

$$\mathbf{V} = \frac{\mathbf{X}^{*\mathbf{T}} \cdot \mathbf{X}^*}{T} \tag{3.28}$$

the components are found by calculating the eigenvalues and eigenvectors of the correlation matrix $\mathbf{V}$. It will be shown that the principal components can be described as a linear combination of these columns. The components will be orthogonal to each other since they are eigenvectors. The first principal component will describe the most of the variance in $\mathbf{X}^*$, the second the second most and so forth. This is achieved by choosing the weights from the $k \times k$ eigenvector matrix of $\mathbf{V}$ in the following way

$$\mathbf{VW} = \mathbf{W\Lambda} \tag{3.29}$$

where $\mathbf{\Lambda}$ is the $k \times k$ diagonal matrix of eigenvalues from $\mathbf{V}$. Next the columns of $\mathbf{W}$ is ordered according to the size of their corresponding eigenvalue. Therefore if $\mathbf{w_{ij}}$ for $i, j = 1, ...k$, then the mth column of $\mathbf{W}$, $\mathbf{w_m} = (w_{1m}, ..., w_{1k})$ is the $k \times 1$ eigenvector that corresponds to eigenvalue λ_m. The columns of $\mathbf{W}$ is ordered so that $\lambda_1 > \lambda_2 > ... > \lambda_k$.
The mth principal component is defined in the following way

$$P_m = w_{1m} \cdot X_1^* + w_{12} \cdot X_2^* + ...w_{1k} \cdot X_k^* \tag{3.30}$$

where X_i^* is the ith column of the normalized matrix $\mathbf{X}^*$. This can be written as

$$P_m = \mathbf{X}^* \cdot \mathbf{w_m} \tag{3.31}$$

The complete $T \times m$ matrix of principal components can thus be expressed as

$$\mathbf{P} = \mathbf{X}^* \cdot \mathbf{W} \tag{3.32}$$

To understand that this leads to orthogonal (uncorrelated) components, notice that

$$\begin{aligned}
\mathbf{PP}' &= \mathbf{W^T X^{*T} X^* W} = \{\mathbf{X}^{*\mathbf{T}} \cdot \mathbf{X}^* = T\mathbf{V}\} \\
&= T\mathbf{W^T W \Lambda} = \{\mathbf{W} \quad \text{is orthogonal} \quad \mathbf{W^T} = \mathbf{W^{-1}}\} \\
&= T\mathbf{\Lambda} \quad \text{,where } \mathbf{\Lambda} \text{ is the diagonal matrix of eigenvalues}
\end{aligned}$$

Therefore the vectors of $\mathbf{P^T P}$ are uncorrelated.
Each principal components variance is determined by its corresponding eigenvalues proportion of the total variation. For example the variation in $\mathbf{X}^*$ explained by P_m is given by

$$\frac{\lambda_m}{\Sigma\lambda_{ii}} = \frac{\lambda_m}{k} \quad \text{where } k \text{ is the number of variables in the system.} \tag{3.33}$$

If the components in the original system is highly correlated the first eigenvalue will be much greater than the others, i.e. the first principal component explains the most of the variation. Equation (3.32) can be expressed as $\mathbf{X}^* = \mathbf{P} \cdot \mathbf{W}'$ because $\mathbf{W}' = \mathbf{W^{-1}}$, in vector form

$$X_i = w_{i1}P_1 + w_{i2}P_2 + ... + w_{ik}P_k \tag{3.34}$$

It is therefore possible to represent each input data vector as a linear combination of the principal components. This is known as the *Principal Component Representation* of the original variables. It is often enough to only include the first few principal components, since these explain almost all the variation.

3.16 Ornstein-Uhlenbeck process

An Ornstein-Uhlenbeck process is a mean-reverting process given by the stochastic differential equation

$$dX_t = \kappa\left(\theta - X_t\right)dt + \sigma dW_t \tag{3.35}$$

where κ, θ and σ are parameters and W_t is a standard Brownian Motion. κ is the speed of mean-reversion and θ is the mean and σ is the volatility.

To solve the equation we introduce a change of variable, letting $Y_t = X_t - \theta$. Then Y_t satisfies the following Stochastic Differential Equation

$$dY_t = dX_t = -\kappa Y_t dt + \sigma dW_t \tag{3.36}$$

In equation 3.36 the process dY_t has a drift towards zero at an exponential rate κ. Therefore a change of variable is motivated to remove the drift. This is done by

$$Y_t = e^{-\kappa t}Z_t \Leftrightarrow Z_t = e^{\kappa t}Y_t, \tag{3.37}$$

Applying Itô's formula to equation 3.37 yields

$$\begin{aligned} dZ_t &= \kappa e^{\kappa t}Y_t dt + e^{\kappa t}dY_t &(3.38)\\ &= \kappa e^{\kappa t}Y_t dt + e^{\kappa t}\left(-\kappa Y_t dt + \sigma dW_t\right) &(3.39)\\ &= 0dt + \sigma e^{\kappa t}dW_t &(3.40)\end{aligned}$$

The solution is obtained by integrating from s to t

$$Z_t = Z_s + \sigma \int_s^t e^{\kappa u} dW_u \tag{3.41}$$

Retransforming the change of variables yields

$$Y_t = e^{-\kappa t} Z_t = e^{-\kappa(t-s)} Y_s + \sigma e^{-\kappa t} \int_s^t e^{\kappa u} dW_u, \tag{3.42}$$

and finally

$$X_t = Y_t + \theta = \theta + e^{-\kappa(t-s)} \left(X_s - \theta\right) + \sigma \int_s^t e^{-\kappa(t-u)} dW_u. \tag{3.43}$$

3.17 Nelson-Siegel's Yield Curve Model

Nelson and Siegel formulated their model to fit the shape of the yield curve in the following way

$$y\left(m\right) = \beta_0 + \beta_1 \frac{\left(1 - e^{-\frac{m}{\tau}}\right)}{m/\tau} + \beta_2 \left(\frac{\left(1 - e^{-\frac{m}{\tau}}\right)}{m/\tau} - e^{-\frac{m}{\tau}}\right) \tag{3.44}$$

where m is term to maturity, β_0 is the long term component, β_1 is the short term component, β_2 is the medium term component. The long term component is a constant throughout the maturity, the medium component starts out at zero and decays to zero and is therefore only affecting the medium term. The short term component has the fastest decay monotonically to zero. τ determines the if the fit will be best at short or long maturities and the location of the maximum of the medium component β_2. A graphical illustration is shown in Figure 3.2.

Some constrains has to be imposed on the model to obtain plausible results when fitting it to empirical yields. First β_0, the long term component has to be greater than zero. Secondly, $\beta_0 + \beta_1$ has to be greater than zero to guarantee non-negative yields at the short end. Thirdly, the decay parameter, τ has to lie in the maturity spectrum, i.e. be greater than zero.

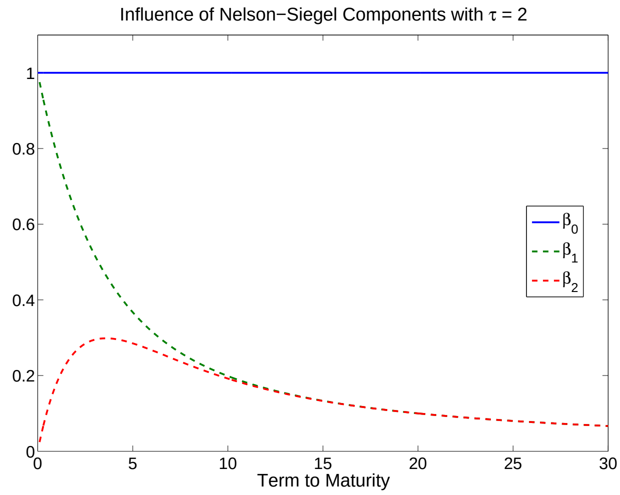

Figure 3.2: Nelson-Siegel components influence

Chapter 4

Data Selection

The floating rate that PK Air uses as a reference is the LIBOR rate, this data was obtained from Bloomberg. Since the LIBOR rates longest maturity is 12 months, US Swap rates have been used as a proxy for longer time to maturity. The idea was to look at the most homogeneous data sets. The time series of data were available in the following setting for daily quotes:

Table 4.1: LIBOR and US Swap Time Series 1988-2009

LIBOR	1 m	2 m	3 m	6 m	12 m	
USSWAPS	2 yr	3 yr	4 yr	5 yr	7 yr	10 yr

Since the time series only go back to 1988, there was a need for time series who stretched further back in time. Therefore US Government Securities yields were also studied, specifically Treasuries with constant time to maturity [15]. The longest homogeneous time series were available from 2nd of January 1962 and were composed in this way and were available in daily, weekly and monthly quotes.

Table 4.2: US Treasuries Constant Time to Maturity Series 1962-2009

Treasury	1 yr	3 yr	5 yr	10 yr

Another setting with more maturities were available from 1982

Table 4.3: US Treasuries Constant Time to Maturity Series 1982-2009

Treasury	3 m	6 m	1 yr	3 yr	5 yr	10 yr

Chapter 5

Results

5.1 Principal Component Analysis

5.1.1 LIBOR and Swaps

Starting out with the LIBOR and Swaps rates some problems appeared with the raw data. It turned out that the rates had sometimes been quoted on different days, typically one rate was missing out on a certain day. To deal with this problem interpolation was applied to obtain homogeneous data sets. It was now possible to construct a yield surface

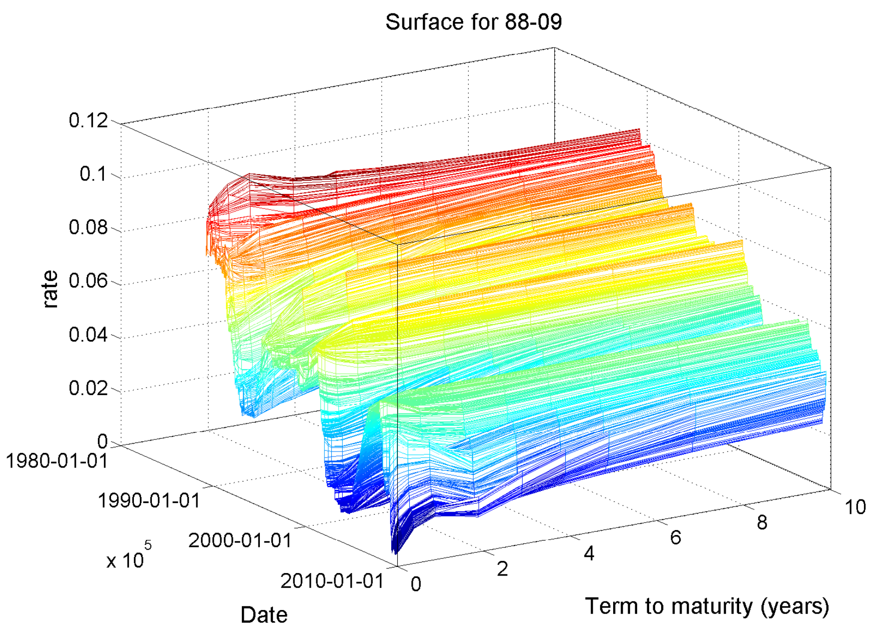

Figure 5.1: Nominal Yield Surface LIBOR and Swaps 1988 to 2009

Secondly the rates were transformed to real rates using the relation (3.1). The US inflation data was given on a monthly basis [8] and was applied to the rates that corresponded to that month. This yielded a real yield surface

as can bee seen the yield surface shape changes.

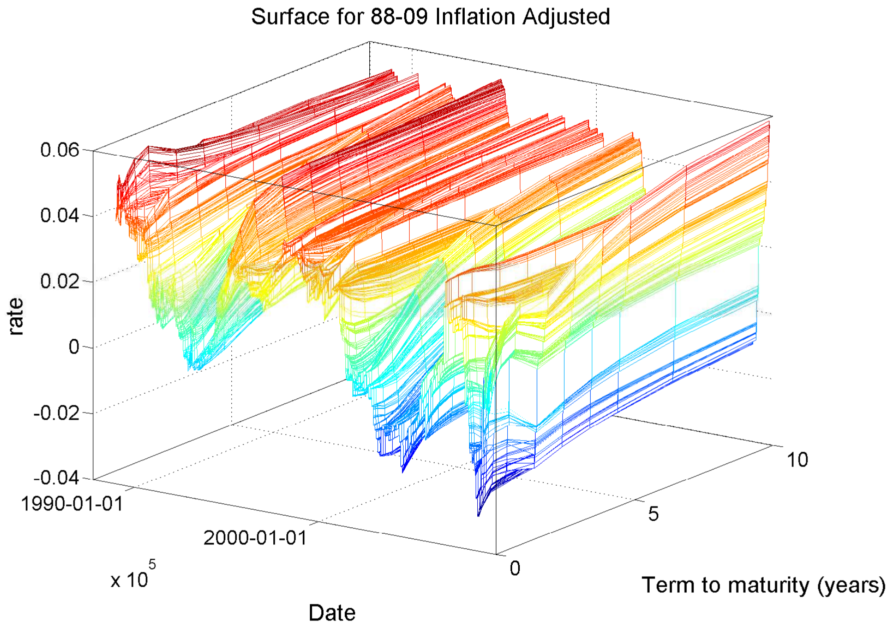

Figure 5.2: Real Yield Surface LIBOR and Swaps 1988 to 2009

Principal Component Analysis has an application to term structures [1]. Term structures are special because they yield an ordering of the system that gives an intuitive interpretation of the principal components as shall be seen. Before performing the PCA it is important that the time series are stationary, the yields are in general not stationary , therefore absolute differences of the yield data were considered, $\mathbf{X}$, as shown in Figure 5.3. As a last step the differences are normalized, that is subtracting the mean and dividing by the standard deviation for each of the columns to obtain a new matrix $\mathbf{X}^*$, a graphical illustration of this for 1 month LIBOR can be seen in Figure 5.4.

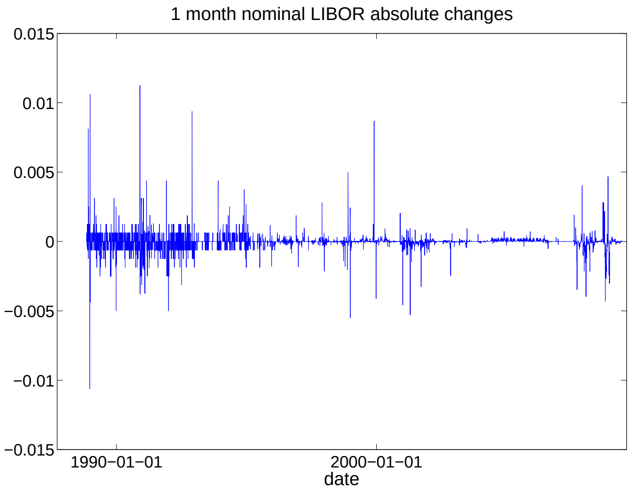

Figure 5.3: 1 month nominal LIBOR absolute changes

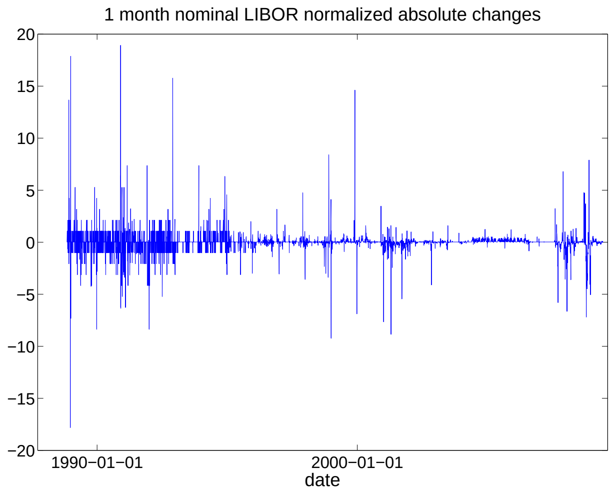

Figure 5.4: 1 month nominal LIBOR normalized absolute changes

From the normalized absolute changes the correlation matrix was obtained

Table 5.1: Correlation Matrix LIBOR and Swaps

	1 m	2 m	3 m	6 m	1 yr	2 yr	3 yr	4 yr	5 yr	7 yr	10 yr
1 m	1										
2 m	0.620	1									
3 m	0.555	0.648	1								
6 m	0.463	0.556	0.741	1							
1 yr	0.377	0.439	0.624	0.757	1						
2 yr	0.112	0.129	0.196	0.229	0.273	1					
3 yr	0.108	0.105	0.163	0.205	0.249	0.934	1				
4 yr	0.105	0.096	0.148	0.186	0.228	0.899	0.935	1			
5 yr	0.089	0.091	0.143	0.179	0.222	0.887	0.928	0.936	1		
7 yr	0.086	0.089	0.139	0.171	0.212	0.858	0.906	0.918	0.957	1	
10 yr	0.087	0.083	0.125	0.160	0.202	0.7924	0.842	0.860	0.916	0.936	1

The correlations shows typical yield curve behavior. The correlation diminishes with the spread between the maturities. Another observation is that the longer term to maturity, the higher the correlation with the neighboring maturities than for shorter maturities.

Table 5.2: Eigenvalues of Correlation Matrix LIBOR and Swaps

Principal Component	Eigenvalue	Cumulative R^2
P_1	5.8070	52.8
P_2	3.0414	80.5
P_3	0.7745	87.5
P_4	0.3842	91.0

The first three components, which are referred to as *the level, the slope* and *the curvature* explains more than 87% of the variation. The table does also show that the first component corresponds to the largest eigenvalue. Figure 5.5 shows a graphical illustration of the first three principal components (PCs).

A hump occurs for the shorter maturities, this could be explained as shall be seen later on by the inclusion of the shorter maturities. The first PC, the level, is approximately constant over the maturities. The second PC, the slope, approximately raises the yield for maturities up to 2 years and lowers it for longer. The third PC, the curvature, approximately raises the yield for all maturities except the interval 3 months to 3 years.

5.1.1.1 Different Valuation Basis

What happens if the valuation basis is changed from absolute returns to log returns? Log returns is defined as

$$X_n = logS_n - logS_{n-1} \tag{5.1}$$

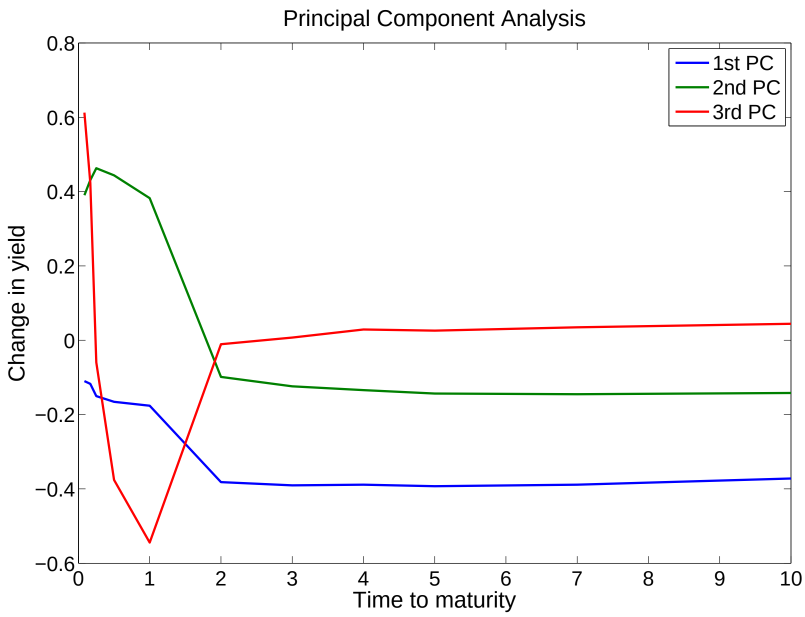

Figure 5.5: Principal Components nominal LIBOR and Swaps

where S_i is the yield at day i. The answer is that the results are almost the same as shown in Figure 5.6 and Table 5.3

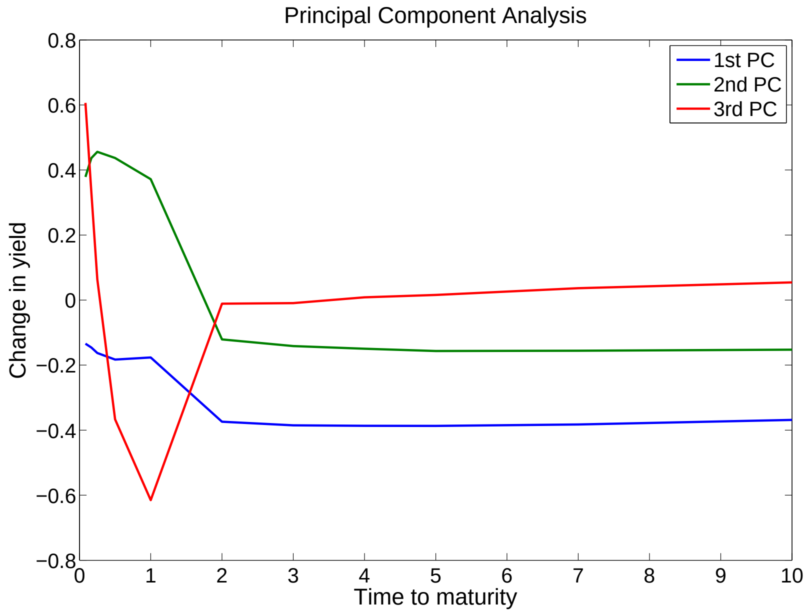

Figure 5.6: Principal Components nominal log LIBOR and Swaps

Table 5.3: Variance Explained log LIBOR and Swaps 1988-2009

Component	Cumulative R^2
P_1	54.03
P_2	84.11
P_3	90.79
P_4	93.68

The results are therefore indifferent of the valuation basis, the small differences can be explained by numerical errors in MATLAB.

5.1.1.2 Weekly data

The same procedure was done for weekly data on the yields of LIBOR and US Swaps. The Correlation Matrix changed as can be seen in Table 5.4.

Table 5.4: Correlation Matrix LIBOR and Swaps Weekly

	1 m	2 m	3 m	6 m	1 yr	2 yr	3 yr	4 yr	5 yr	7 yr	10 yr
1 m	1										
2 m	0.679	1									
3 m	0.734	0.839	1								
6 m	0.637	0.753	0.857	1							
1 yr	0.469	0.577	0.726	0.824	1						
2 yr	0.310	0.346	0.442	0.591	0.641	1					
3 yr	0.274	0.303	0.394	0.537	0.598	0.979	1				
4 yr	0.257	0.284	0.365	0.502	0.562	0.941	0.968	1			
5 yr	0.229	0.271	0.343	0.482	0.547	0.926	0.963	0.964	1		
7 yr	0.209	0.247	0.323	0.447	0.515	0.890	0.938	0.946	0.976	1	
10 yr	0.191	0.222	0.311	0.423	0.494	0.873	0.914	0.922	0.951	0.962	1

5.1.2 US Treasuries Constant Maturity

5.1.2.1 Daily Data

The Treasury yield curve go back from 1962, but from then data are only available for maturities from 1 to 10 years. This longest yield surface is shown in Figure 5.7.

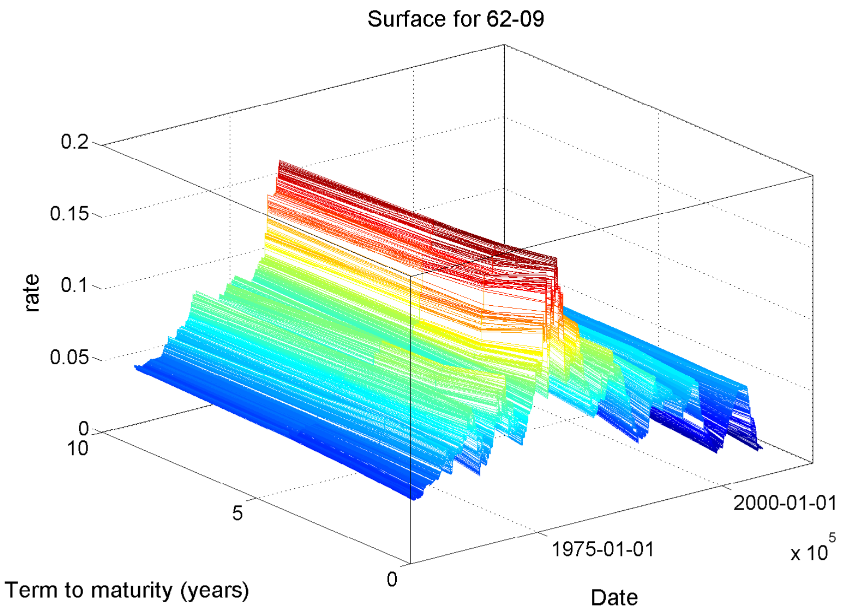

Figure 5.7: T-notes 62-09 Nominal Yield Surface

If the whole period is analyzed at once, the following results are obtained

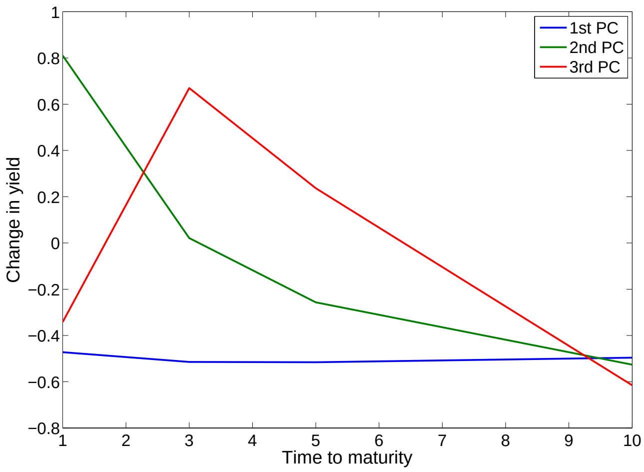

Figure 5.8: Principal Components Nominal T-notes 62-09

Table 5.5: Variance Explained T-Notes 1962-2009

Component	Cumulative R^2
P_1	89.05
P_2	96.43
P_3	98.89

The *level* explains almost all the variation.
The time interval between 1962 to 2009 was also split into smaller windows of 5 years. The visual results of how the principal components propagates over time can be seen in Appendix A. The total variance explained is shown in Table 5.6.

Table 5.6: Variance Explained T-notes 1962-2009 5 year windows

Cum. R^2	62-67	67-72	72-77	77-82	82-87	87-92	92-97	97-02	02-07	07-09
P_1	80.66	86.94	81.28	91.26	91.46	91.37	92.66	89.21	89.81	84.48
P_2	91.99	94.81	91.95	96.61	96.90	97.76	98.02	97.15	97.52	96.73
P_3	96.91	98.42	97.50	98.82	98.93	99.25	99.40	99.20	99.48	99.43

From the table is clear that the *level* explains more than 80 % of the total variation and all three components explains more than 96 %. This implies that when leaving out the shorter term to maturities the explanatory power increases. From the graphical illustrations of the components the "hump" that occurred with the LIBOR and US Swaps has now disappeared. The explanation is that the shorter term to maturities are not included in the T-notes and that they do not covary with longer maturities as stated by Duffee [6]. Another observation is that the *level* is approximately constant over the different time windows. The same relation holds for the *slope* and the *curvature* has the same shape except for the intervals 1967 to 1977.

The next dataset is from 1982 to 2009 and the nominal yield surface is shown in Figure 5.9 and compared to the previous set, maturities of 3 months and 6 months were included.

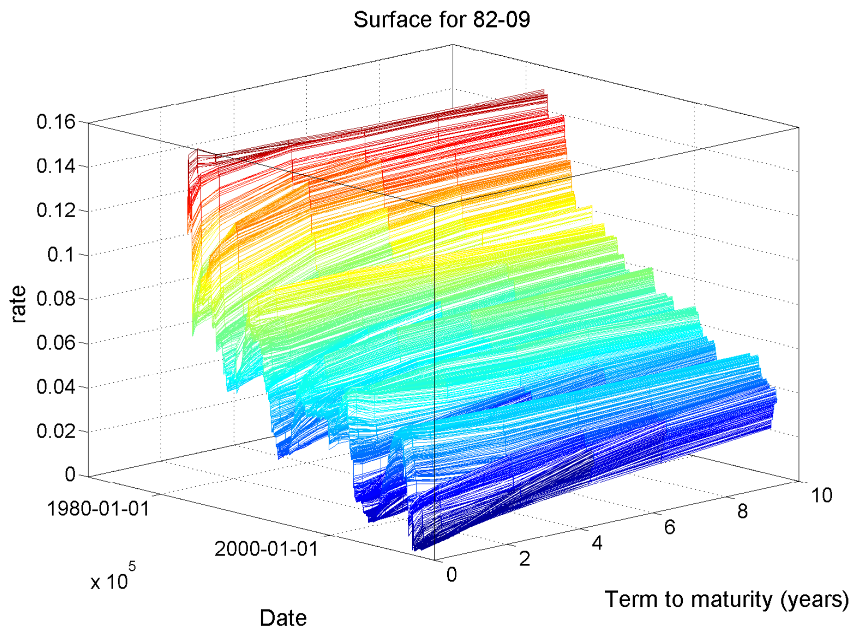

Figure 5.9: Nominal T-bond Yield Surface 1982-2009

The principal components are shown in Figure 5.10

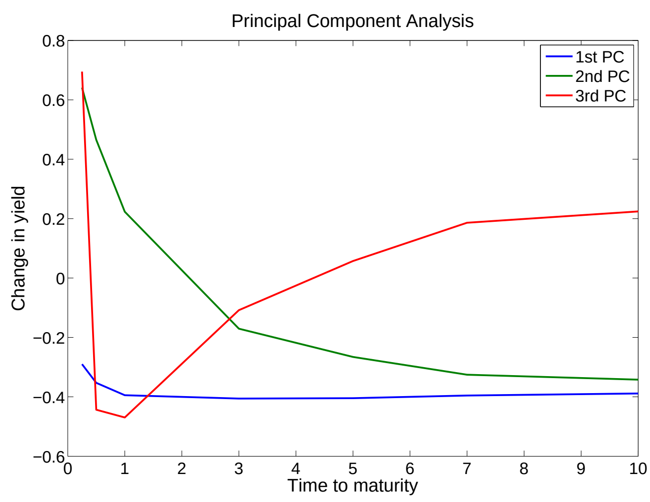

Figure 5.10: T-bond 1982-2009 Principal Components

Here the "hump" resurrects again and could as before be originated to the shorter maturities. According to Table 5.7 the same pattern continues, the *level* explains the most of the variation, although it has diminished compared to the window-study above, one possible explanation for this is the inclusion of shorter maturities.

Table 5.7: Variance Explained T-bonds 1982-2009

Component	Cumulative R^2
P_1	78.47
P_2	93.54
P_3	96.60

5.1.2.2 Weekly data

For the weekly data I started out with the T-notes (1 to 10 year maturities) from 1962-2009. The principal components are shown in Figure 5.11

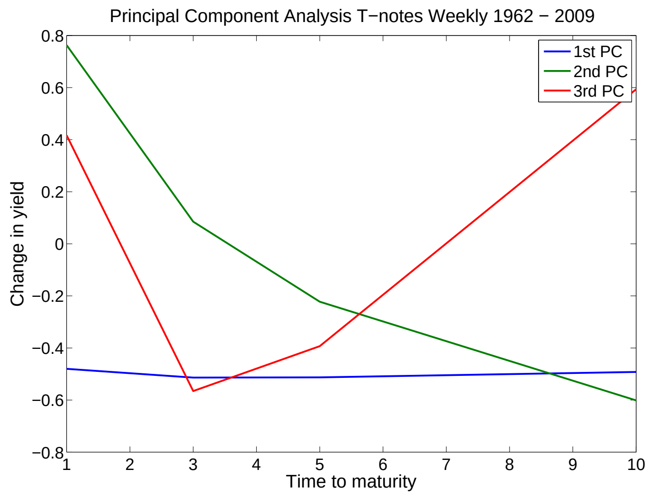

Figure 5.11: Principal Components absolute changes Nominal Weekly T-notes 1962-2009

Table 5.8: Variance Explained T-Notes Weekly 1962-2009

Component	Cumulative R^2
P_1	91.83
P_2	97.95
P_3	99.42

From Table 5.8 we see that the *level* itself explains almost all variation. Another observation is that explanatory power is higher than for the daily data.

5.1.2.3 Monthly Data

The principal components of the monthly data for the period 1962-2009 is shown in Figure 5.12

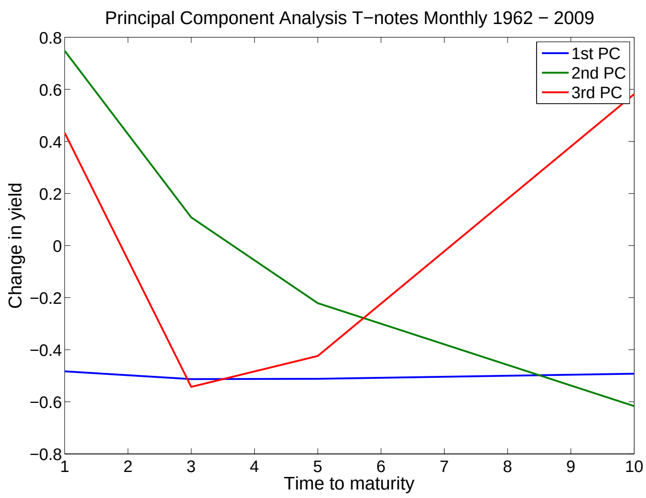

Figure 5.12: Principal Components absolute changes Nominal Monthly T-notes 1962-2009

Table 5.9: Variance Explained T-Notes Monthly 1962-2009

Component	Cumulative R^2
P_1	93.43
P_2	98.84
P_3	99.77

Table 5.9 shows that the variation explained increases when the frequency of the data decreases, e.g. from daily to monthly data.

5.1.2.4 Conclusion PCA

The results are indifferent of valuation basis, i.e. with relative or absolute changes. The explanatory power increases when the data frequency decreases. When shorter maturities were added to the analysis, the explanatory power decreased. This can be observed by comparing the result of LIBOR and Swaps with T-notes. It was also confirmed that the first three Principal Components explains almost all the variation in the system. These results are analogue with previous research.

5.1.3 Correlation with the PURCS Cycle

Since the PURCS Cycle is the driver of all different subsimulations in SAFE, PK Air would also want it to affect the evolution of the yield curve in some sense. Therefore the correlation between PURCS and the yields was studied. First PURCS was compared against 1 month Nominal LIBOR, the series are shown in Figure 5.13 and a scatter point of their absolute changes is shown in Figure 5.14.

As can be seen in the scatter plot they do not seem to co-vary at all and the correlation coefficient is calculated to 0.0592. So there is no correlation between these two time series? It turned out that there actually is! The problem lies in the scope of Signal Theory. Since the PURCS Cycle is given on a yearly basis and the 1 month LIBOR on a daily basis, the latter will be very noisy compared to former. Therefore the 1 month LIBOR should be smoothed to give a fair description of the correlation. For this purpose Smoothing Splines were used in MATLABs *Curve Fitting Toolbox (cftool)*, see Section 3.10. An illustration of the different smoothing levels is shown in Figure 5.15. The next step was to analyze how the correlation was affected by the smoothing level and the result is given in Figure 5.16. As can bee seen the correlations is about 0.6 and to retain the most information in the original data the smoothing factor was chosen to 0.995.

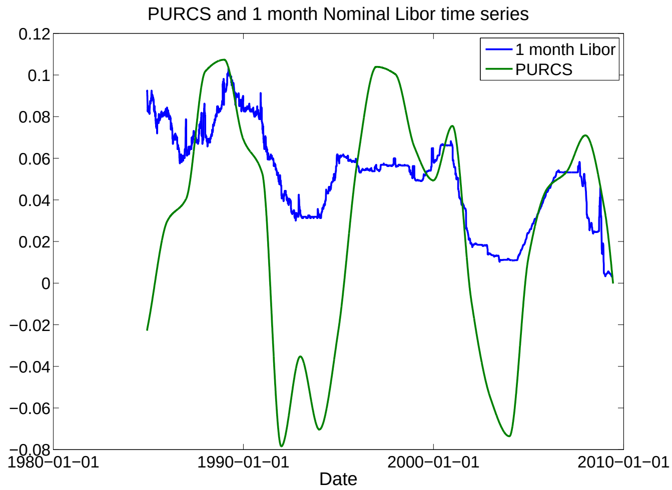

Figure 5.13: PURCS and 1 month Nominal LIBOR Time Series

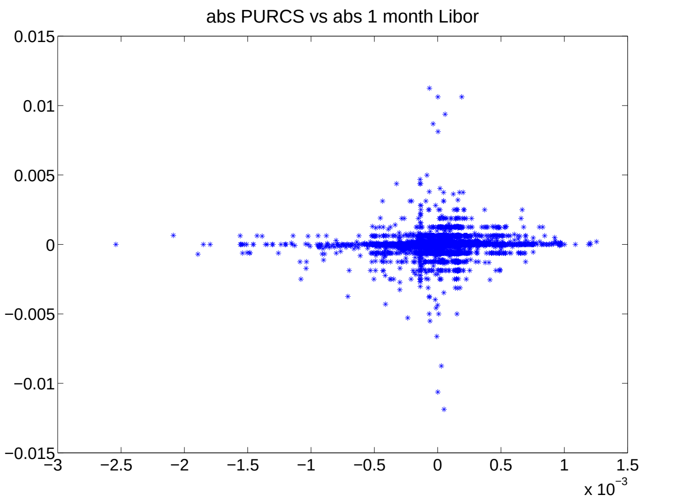

Figure 5.14: absolute changes of PURCS vs 1 month Nominal LIBOR

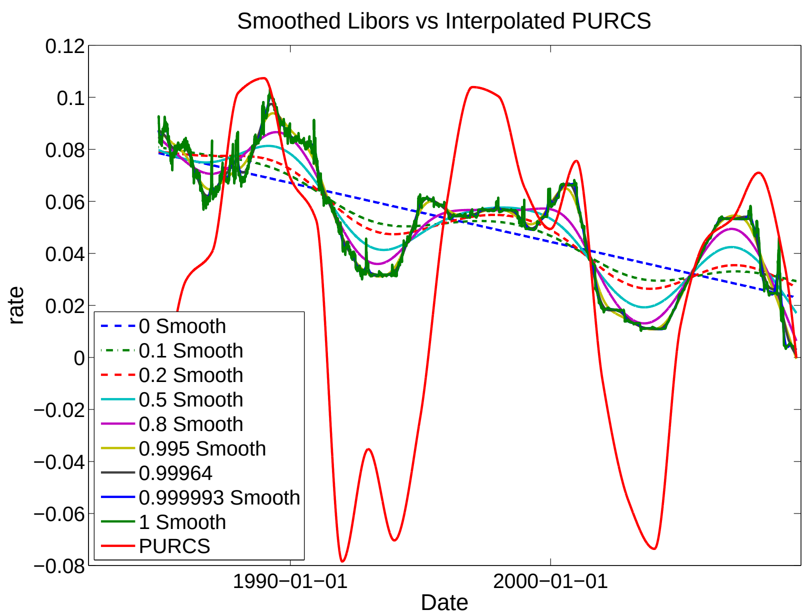

Figure 5.15: PURCS and 1 month Nominal LIBOR with different smoothing

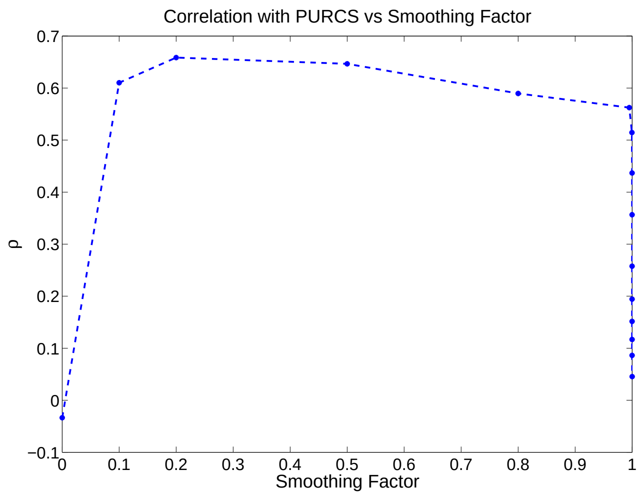

Figure 5.16: Correlation between PURCS and 1 month Nominal LIBOR with different smoothing

5.1.3.1 Crosscorrelation with PURCS

To determine if the correlation could be higher if the two time series were shifted crosscorrelation was also studied for the different absolute changes of the smoothing factors. It could be determined that the highest correlation was obtained for lag $h = 0$, i.e. for the original times series not shifted. A graphical illustration is shown in Figure 5.17.

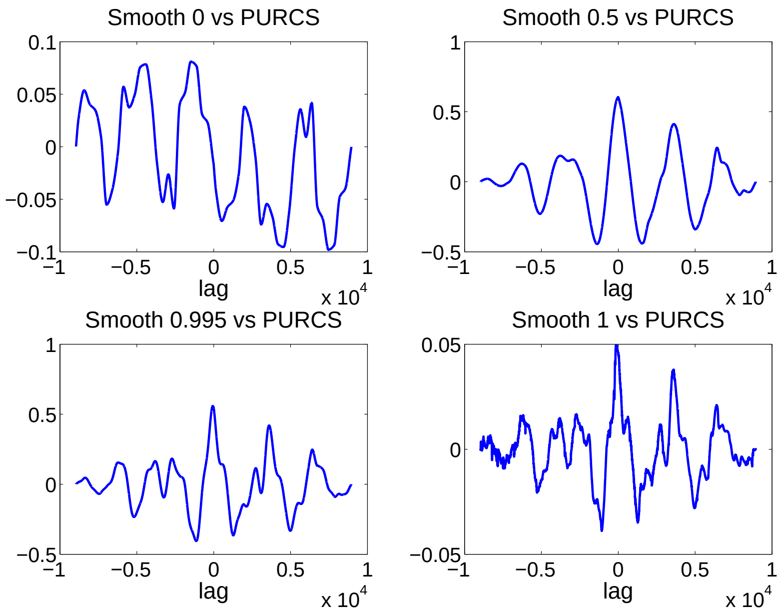

Figure 5.17: Crosscorrelation Coefficient between PURCS and 1 month Nominal LIBOR with different smoothing factors

5.2 Simulation: PCA

The idea was to simulate future yield curves using the *The Principal Component Representation*

$$X_i = w_{i1}P_1 + w_{i2}P_2 + ... + w_{ik}P_k. \tag{5.2}$$

Since the first three principal components explained 93.68 % of the total variation, focus was only on these and (5.2) boils down to

$$X_i = w_{i1}P_1 + w_{i2}P_2 + w_{i3}P_3 + \epsilon, \tag{5.3}$$

$$where \quad \epsilon \sim N\left(0, Var\left(w_{i4}P_4 + ... + w_{ik}P_k\right)\right) \tag{5.4}$$

To do the simulation we keep the eigenvectors, w_{i1}, w_{i2}, w_{i3} fixed and study the distributions of the first three principal components, P_1, P_2, P_3.

5.2.1 LIBOR and US Swaps

5.2.1.1 Nominal

The distribution of the first PC is shown in Figure 5.18, the best fits for the distributions are obtained with a $t_{4.57}$-location-scale[1] distribution with parameters as shown in Table 5.10. The estimation was done in MATLAB using *dfittool*.

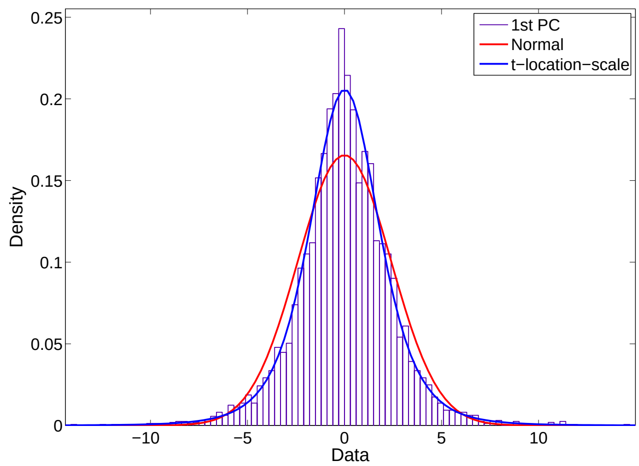

Figure 5.18: Distribution of Principal Component 1 LIBOR and Swaps

[1]If the random variable X is t_p distributed and μ and σ are estimated then the corresponding location-scale variable is $Y = \mu + \sigma X$.

Table 5.10: Estimate of t-location-scale PC 1 LIBOR and Swaps

Parameter	Estimate	Standard Error
μ	0.0089	0.029
σ	1.837	0.032
ν	4.571	0.300

The distribution of the second PC is shown in Figure 5.19 and the parameter estimation in Table 5.11

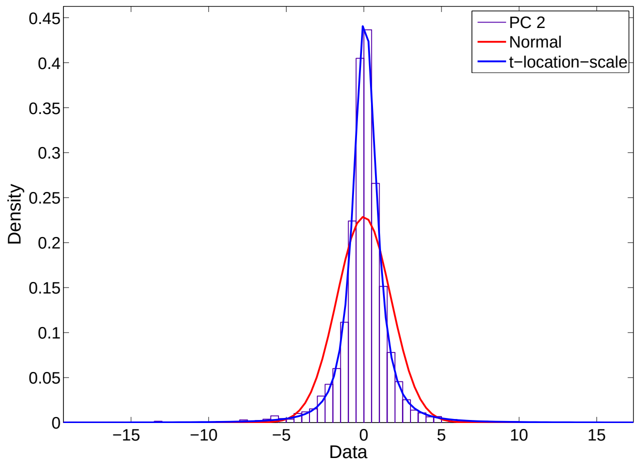

Figure 5.19: Distribution of Principal Component 2 LIBOR and Swaps

Table 5.11: Estimate of t-location-scale PC 2 LIBOR and Swaps

Parameter	Estimate	Standard Error
μ	0.0680	0.014
σ	0.7841	0.016
ν	2.014	0.073

For the third component the following t-location-scale estimation was obtained

Table 5.12: Estimate of t-location-scale PC 3 LIBOR and Swaps

Parameter	Estimate	Standard Error
μ	-0.003	0.0059
σ	0.33	0.0074
ν	1.89284	0.071

From these findings the simulation was done by Monte Carlo simulation from the corresponding t-location scale distributions. The number of days simulated was 8000 days (about 16 years) since this is the average length of PK Air's deals. To make the simulations more stable the simulations are done 1000 times and then averaged. The positive thing with averaging is that the simulations become more stable, but the negative thing is that the yields flatten out and lose some of their original shape.

The simulation is started from the following yield curve from the simula-

Table 5.13: Fictional Starting Yield Curve

TTM	1m	2m	3m	6m	1yr	2yr	3yr	4yr	5yr	7yr	10yr
yield	0.01	0.015	0.0175	0.02	0.023	0.025	0.027	0.028	0.029	0.0295	0.0296

tion according to (5.3) the normalized changes X^*_{sim} were obtained. These were tranformed back to X_{sim} using

$$X_{sim} = X^*_{sim} \cdot V(X_i) + E(X_i) \tag{5.5}$$

where X_i is the original changes column i, e.g. the column for changes of 3 month LIBOR over time. The next step was to transform back the yield data by

$$S_{i+1} = \Delta_{i,i+1} + S_i \tag{5.6}$$

where S_i are the yields at time i and $\Delta_{i,j}$ is the absolute yield change from time i to j. A simulation for 8000 days conducted 1000 times and then averaged resulted in the following yield surface

The result is not that pleasing since the yields go negative. The surface

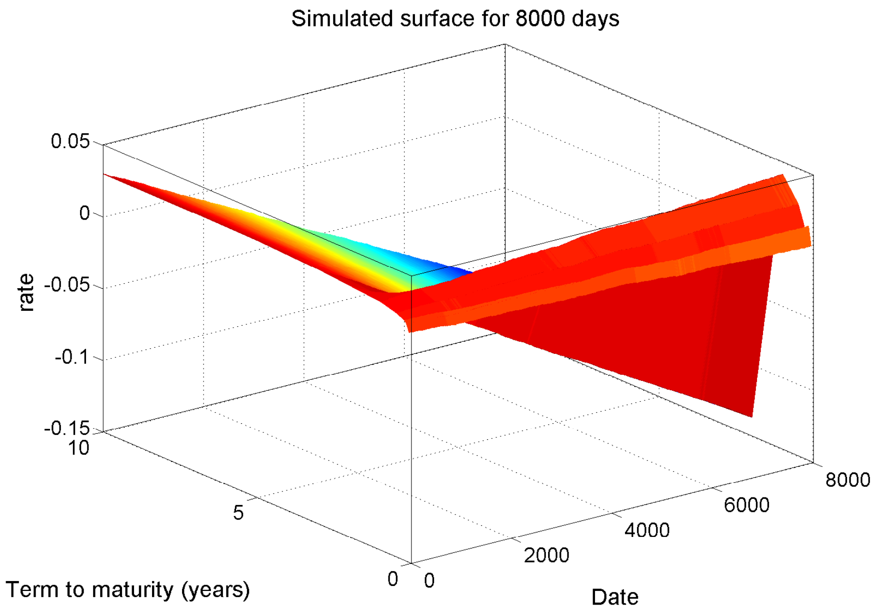

Figure 5.20: Simulated Surface for 8000 days Nominal LIBOR and Swaps

also shows a hump for shorter maturities. Lots of the movements have been flattened out because of the averaging. A Principal Component Analysis on the simulated changes correlation matrix is visible in Figure 5.21.

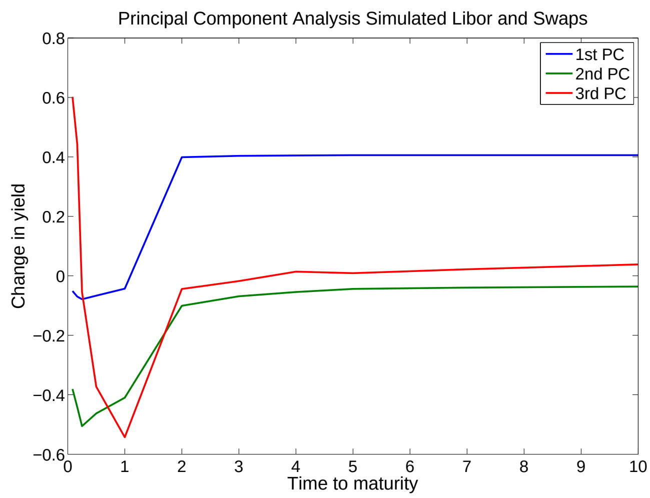

Figure 5.21: Principal Components Simulated Nominal LIBOR and Swaps

If comparing with Figure 5.5 the level component have shifted from approximately -0.4 to 0.4, the slope has also been inverted, but the curvature remains in the same setting. The correlation matrix of the simulated data is shown in Table 5.14

Table 5.14: Correlation Matrix Simulated Nominal LIBOR and Swaps

	1 m	2 m	3 m	6 m	1 yr	2 yr	3 yr	4 yr	5 yr	7 yr	10 yr
1 m	1										
2 m	0.975	1									
3 m	0.662	0.811	1								
6 m	0.324	0.525	0.923	1							
1 yr	0.104	0.320	0.813	0.974	1						
2 yr	-0.015	-0.023	0.025	0.064	0.113	1					
3 yr	-0.036	-0.057	-0.035	-0.002	0.049	0.998	1				
4 yr	-0.030	-0.060	-0.063	-0.041	0.006	0.994	0.999	1			
5 yr	-0.047	-0.079	-0.081	-0.055	-0.004	0.993	0.998	0.999	1		
7 yr	-0.042	-0.078	-0.090	-0.069	-0.020	0.991	0.998	0.999	0.999	1	
10 yr	-0.034	-0.074	-0.098	-0.082	-0.036	0.989	0.996	0.998	0.999	0.999	1

5.2.1.2 Real

The normalized original change matrix X^* had the setting shown in Table 5.15 The simulation was conducted in the same manner as earlier and it

Table 5.15: Correlation Matrix Real LIBOR and Swaps

	1 m	2 m	3 m	6 m	1 yr	2 yr	3 yr	4 yr	5 yr	7 yr	10 yr
1 m	1										
2 m	0.876	1									
3 m	0.869	0.901	1								
6 m	0.839	0.873	0.938	1							
1 yr	0.794	0.822	0.893	0.928	1						
2 yr	0.677	0.691	0.744	0.746	0.738	1					
3 yr	0.667	0.677	0.728	0.733	0.725	0.974	1				
4 yr	0.662	0.670	0.720	0.723	0.714	0.959	0.974	1			
5 yr	0.662	0.674	0.725	0.727	0.717	0.955	0.971	0.974	1		
7 yr	0.666	0.678	0.728	0.729	0.718	0.944	0.962	0.967	0.983	1	
10 yr	0.673	0.683	0.732	0.732	0.720	0.920	0.938	0.944	0.967	0.975	1

yielded the yield curve in Figure 5.22. It showed the same pattern as for the nominal yields, although not as neagtive, but still inplausible.

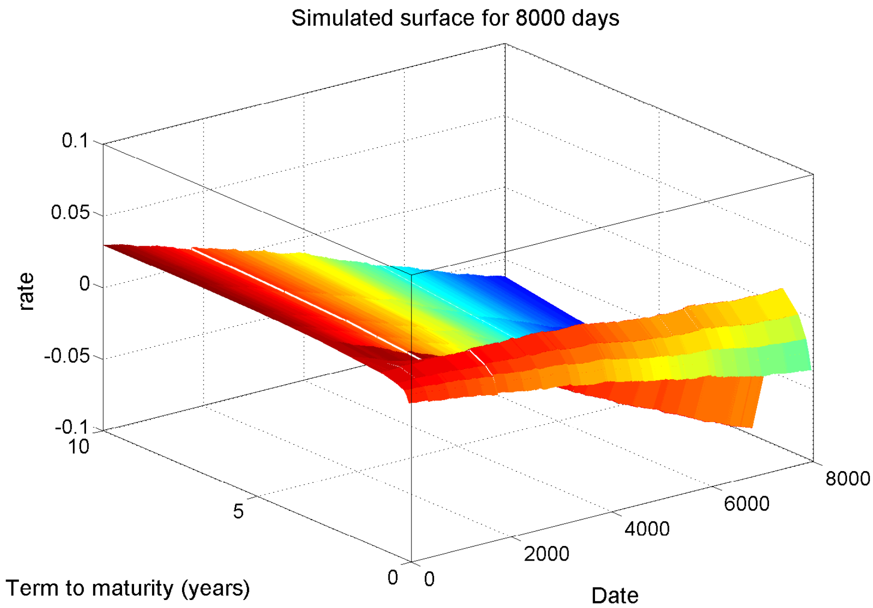

Figure 5.22: Simulated Surface for 8000 days 1000 times Real LIBOR and Swaps

A PCA on the simulated changes yielded the following results

The variance explained is displayed in Table 5.16 The simulated correlation matrix is given in Table 5.17

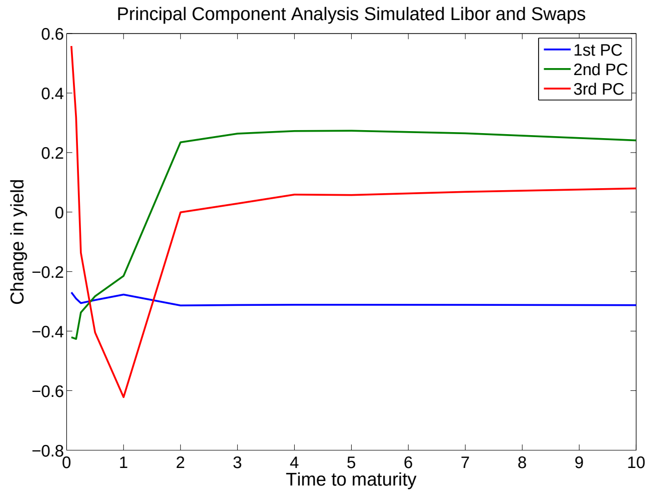

Figure 5.23: Principal Components Simulated Real LIBOR and Swaps

Table 5.16: Variance Explained Simulated Real LIBOR and Swaps 2008 - 2009

Component	Cumulative R^2
P_1	89.17
P_2	94.88
P_3	100

5.2.2 US Treasuries Constant Time to Maturity

For comparison purposes the US Treasuries were studied. First the longest set of data is analyzed, that is from 1962 to 2009. For this interval T-notes were only available with maturities of 1, 3, 5 and 10 years. The correlation matrix had the following setting The simulated surface can be seen in Figure 5.24

The resulting surface for this setting is more pleasing than for the LIBOR and Swaps rates. An upward trend is apparent. PCA of the simulated data are shown in Figure 5.25 and the correlation matrix in Table 5.19

From these results the conclusion was that the simulations were highly dependent on the propagation of the historical datas used as a foundation for the simulations. For the LIBOR and Swaps there has been a downward trend since 1988 and the simulations therefore continue in the same pattern. For the T-notes the historical data were Λ-shaped but with more upward trends than downward and the simulations exhibit an upward trend.

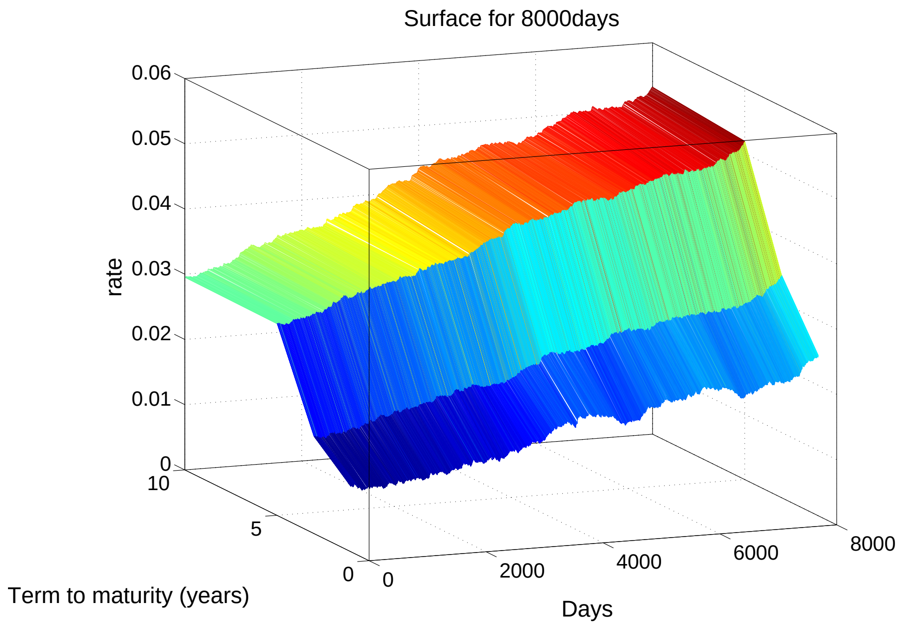

Figure 5.24: Simulated Surface for Nominal T-notes 1962-2009

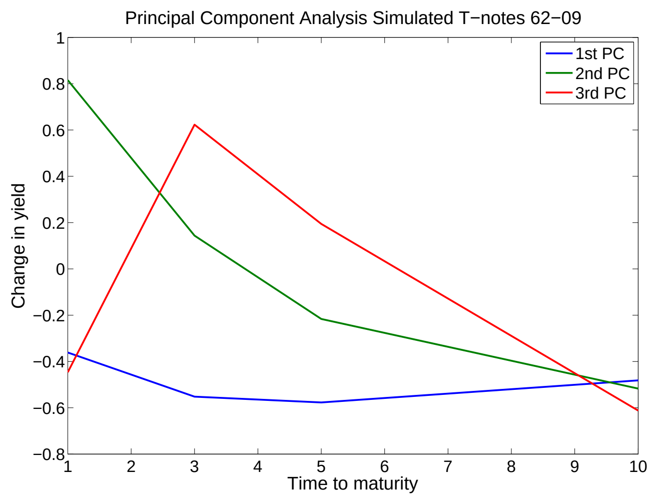

Figure 5.25: Principal Components Simulated Yields

Table 5.17: Correlation Matrix Simulated Real LIBOR and Swaps 1988 - 2009

	1 m	2 m	3 m	6 m	1 yr	2 yr	3 yr	4 yr	5 yr	7 yr	10 yr
1 m	1										
2 m	0.982	1									
3 m	0.86	0.938	1								
6 m	0.730	0.846	0.978	1							
1 yr	0.595	0.736	0.925	0.984	1						
2 yr	0.768	0.832	0.892	0.868	0.822	1					
3 yr	0.765	0.825	0.879	0.852	0.803	0.999	1				
4 yr	0.770	0.826	0.872	0.842	0.789	0.998	0.999	1			
5 yr	0.770	0.825	0.872	0.842	0.790	0.998	0.999	0.999	1		
7 yr	0.776	0.830	0.874	0.842	0.788	0.998	0.999	0.999	0.999	1	
10 yr	0.789	0.842	0.882	0.846	0.790	0.998	0.999	0.999	0.999	0.999	1

Table 5.18: Correlation Matrix Nominal T-notes 1988 - 2009

	1 y	3 y	5 y	10 y
1 y	1			
3 y	0.847	1		
5 y	0.800	0.941	1	
10 y	0.729	0.873	0.926	1

5.3 Simulation: Rebonato

Since the simulation from the *Principal Component Representation* was mostly unplausible a test with the simulation conducted by Rebonato, Mahal, Joshi, Buchholz and Nyholm [17] was also implemented. I started out with their first approach; *the naive*. As described earlier it is an historical simulation method, where the future yields are simulated from the historical yield changes. The method is as follows:

1. Draw a random starting yield change curve from the historical yield curve changes.
2. Draw a random number between 5 and 50 for the window length of the subsimulation.
3. In each extraction of a yield curve change there is a 5 % probability of jumping out of the window and go back to step 1.
4. Continue with the subsimulations until the desired simulation length is obtained.

5.3.1 T-notes

US Treasury yields with maturities of 3 months to 10 years between 2nd of January 1986 to 7th of May 2004 were studied to be able to relate to the study by Rebonato and Nyholm [18]. The difference from the other approach was that no averaging was done and percentage changes were used to keep the yields positive. One yield curve outcome is shown in Figure 5.26.

The Principal Components before and after are shown in Figures 5.27 and

Table 5.19: Correlation Matrix Simulated Nominal T-notes 1988 - 2009

	1 y	3 y	5 y	10 y
1 y	1			
3 y	0.586	1		
5 y	0.417	0.916	1	
10 y	0.215	0.581	0.852	1

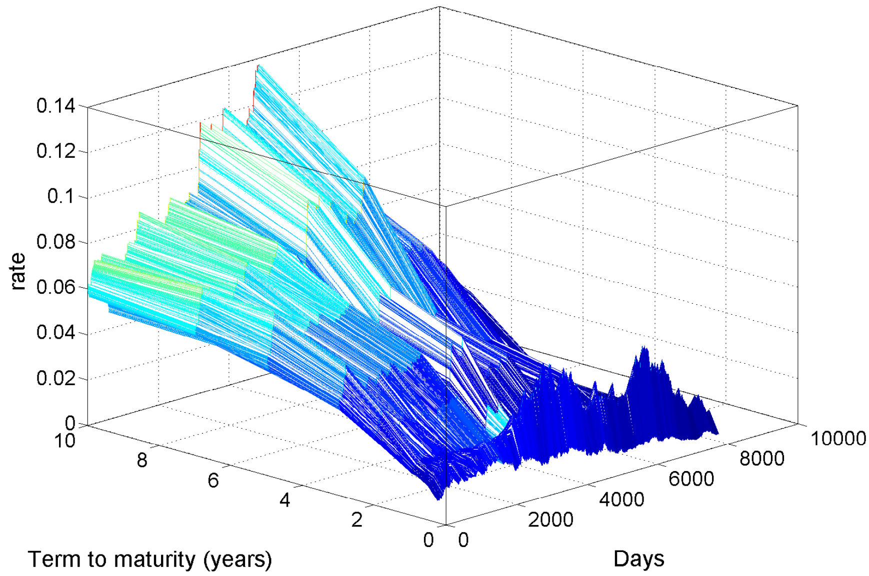

Figure 5.26: Simulated Nominal Surface T-yields 1962 - 2009

5.28. As can be seen the components are preserved. To clarify this the correlation matrices are shown in Table 5.20 and 5.21. As stated in Rebonato, Mahal, Joshi, Buchholz and Nyholm [17] this procedure asymptotically recovers the correlation matrix and the same holds for the eigenvalues and eigenvectors.

The final result with this *naive* approach is not pleasing although that the correlation structure and the principal components are recovered. Therefore the author's *spring and mean reversion* model was tried. It uses spring constants and mean-reversion levels to match the variance of the curvature of the simulation with the empirical variance of the curvature. It is not clear

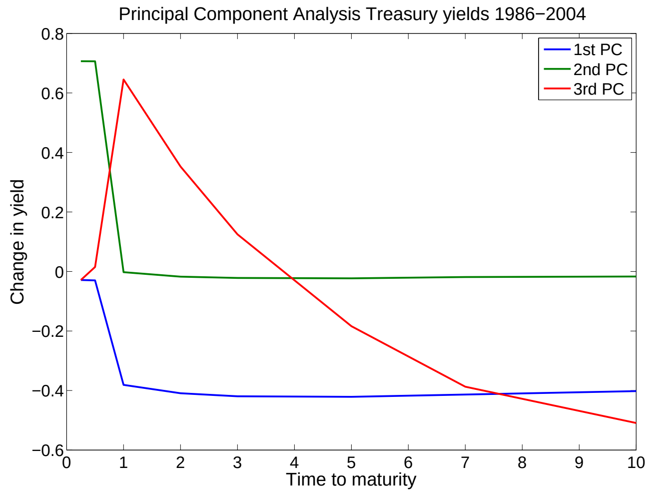

Figure 5.27: Principal Components Treasury Yields 1982 - 2004

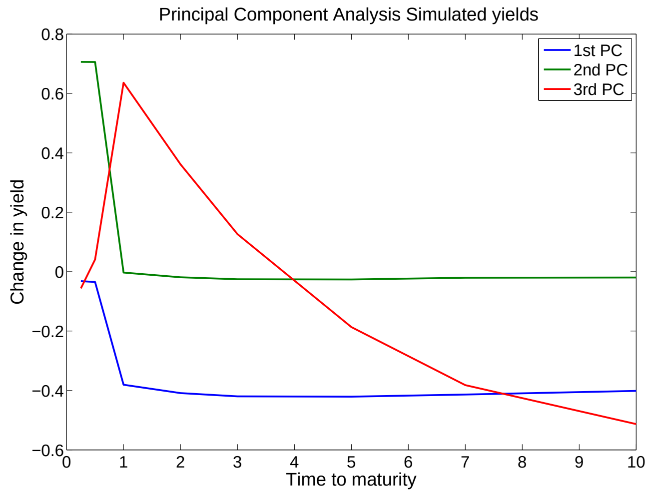

Figure 5.28: Principal Components Simulated Yields

Table 5.20: Correlation Matrix Nominal Treasury Yields 1982 - 2004

	3 m	6 m	1 y	2 y	3 y	5 y	7 y	10 y
3 m	1							
6 m	0.786	1						
1 y	0.053	0.057	1					
2 y	0.038	0.047	0.889	1				
3 y	0.037	0.042	0.860	0.964	1			
5 y	0.038	0.039	0.816	0.908	0.953	1		
7 y	0.043	0.043	0.769	0.859	0.912	0.964	1	
10 y	0.045	0.042	0.730	0.810	0.871	0.940	0.974	1

Table 5.21: Correlation Matrix Simulated Nominal Treasury Yields

	3 m	6 m	1 y	2 y	3 y	5 y	7 y	10 y
3 m	1							
6 m	0.790	1						
1 y	0.054	0.070	1					
2 y	0.041	0.057	0.888	1				
3 y	0.040	0.047	0.858	0.962	1			
5 y	0.043	0.044	0.811	0.899	0.951	1		
7 y	0.049	0.050	0.766	0.853	0.909	0.961	1	
10 y	0.051	0.047	0.721	0.802	0.866	0.936	0.973	1

from their paper on how to do this matching. In this study a non-linear least squares routine was tried on the data. Unfortunately, the results were still implausible, the final curvatures obtained translated distributions, either greater or less than zero. This yielded even more implausible results than the *naive* approach. Therefore this model was abandoned.

5.4 Simulation: Brute Force Model

Since the two former models did not result in plausible future yield curves, a need for a new model arose. Another criteria was that the yield curves should be affected by the PURCS cycle. As stated in the article by Rebonato [17] - *Brute-force simulation and analysis of the results is therefore often the only possible investigative route.* Due to the failure of the former models, the focused shifted to brute force modeling. The idea was to add the cycle as a factor to the yield curve evolution. To keep the correlation structure smoothing splines were applied to the LIBOR and Swaps yields with a smoothing factor of 0.995, see Section 5.1.3. The relationship between the PURCS cycle and the smoothed yields was set up in the following way

$$a(j) + log(y(i,j)) = log(1 + c(i) \cdot d(j)) \tag{5.7}$$

where $a(j)$ is a translation factor, $y(i,j)$ is the yield at date i and maturity j, $c(i)$ is the PURCS cycle at date i and $d(j)$ is a scaling factor. 1 is added to the cycle since it has negative values. The reason that there is no scaling factor on the yields is that this is covered by the scaling on the PURCS cycle. If a scaling factor is introduced the optimization routine fails to find plausible solutions, it will just keep on rescaling.

This relationship was studied for historical data for the LIBOR and Swaps and the PURCS cycle. Expression (5.7) was re-expressed as

$$F(i,j) = a(j) + log(y(i,j)) - log(1 + c(i) \cdot d(j)) \tag{5.8}$$

To determine the functions a and d, for each maturity j a non-linear least squares optimization routine, *lsqnonlin*, in MATLAB was used. The results are shown in Figure 5.29 and 5.30

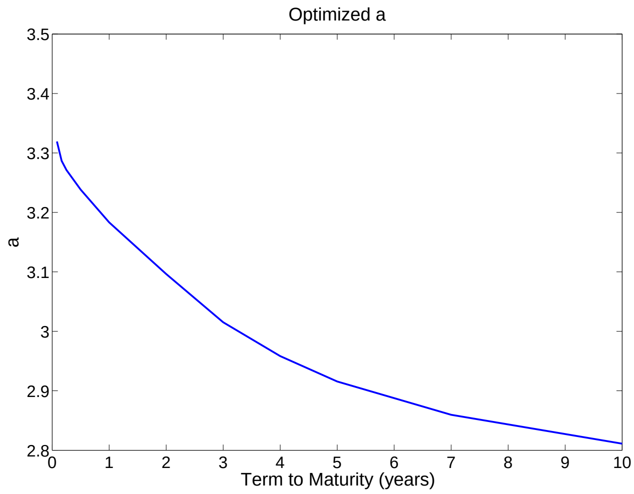

Figure 5.29: Least Squares optimized a

As can be seen, a do not vary that much and is therefore set to 3. For d an exponential function

$$d(t_j) = m + pe^{-rt_j} \tag{5.9}$$

was fitted with least squares to the function to describe the factor variation across maturities. The fitted function is shown in Figure 5.31 and its values in Table 5.22 and Goodness of Fit values in Table 5.23.

Table 5.22: Estimated parameters for function d with 95 % Confidence bounds

parameter	value	Lower Bound	Upper Bound
m	0.4152	-0.211	1.041
p	5.835	5.245	6.426
r	0.2812	0.2067	0.3557

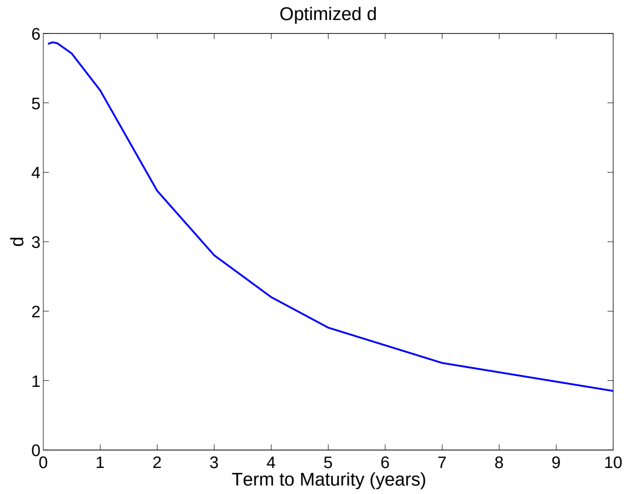

Figure 5.30: Least Squares optimized d

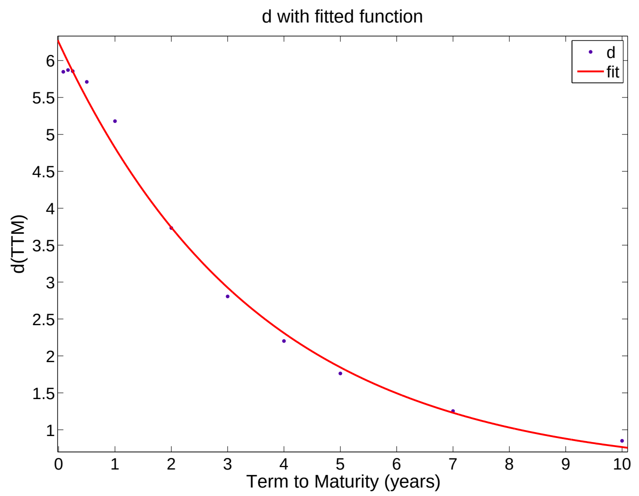

Figure 5.31: Least Squares fitted function to d

Table 5.23: Goodness of Fit for function d

Measure	Value
Sum of Squared Errors	0.3051
R-Squared	0.9926
Adjusted R-Squared	0.9907

Having determined a and d, equation (5.8) could be rewritten as

$$F(i,j) - a = log(y(i,j)) - log(1 + c(i) \cdot d) \tag{5.10}$$

$$e^{F(i,j)-a} = e^{log(y(i,j))-log(1+c(i)\cdot d)} \tag{5.11}$$

$$= \frac{y(i,j)}{(1 + c(i) \cdot d)} \tag{5.12}$$

in other words, $e^{F(i,j)-a}$ is the yields without PURCS cycle influence, and that is what should be simulated. From this it is clear that we only need the parameter d, since the value of a does not affect the expression.

The last equation could now be re-expressed as

$$z(i,j) = \frac{y(i,j)}{(1 + c(i) \cdot d)} \tag{5.13}$$

where z is referred to as *the Cycle Neutral Yield* at date i and maturity j. By using this "logarithmic approach" the PURCS cycle is affecting the yields as a factor. The Cycle Neutral surface is shown in Figure 5.32.

The next step was to apply Nelson-Siegel's curve fitting technique to this surface. This was conducted by running an nonlinear least squares routine, *lsqnonlin*, in MATLAB. It was done with the same conditions as stated in Section 3.17. The time series of the four parameters are shown in Figure 5.33, and the error of the fit in Figure 5.34. The downward trend since 1988 to 2009 is evident in β_0.

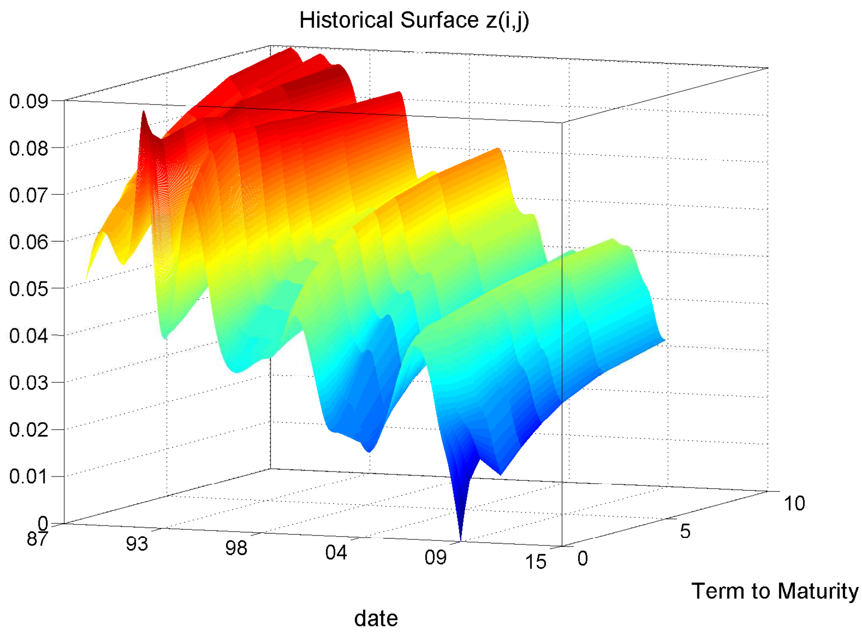

Figure 5.32: Historical PURCS Cycle Neutral Yield Surface

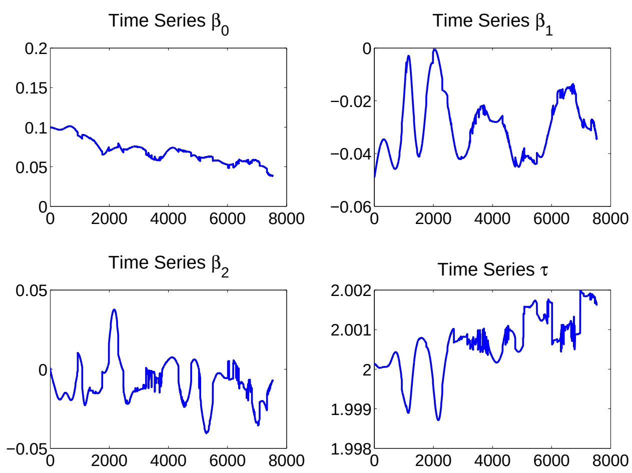

Figure 5.33: Nelson-Siegel estimated parameters from z

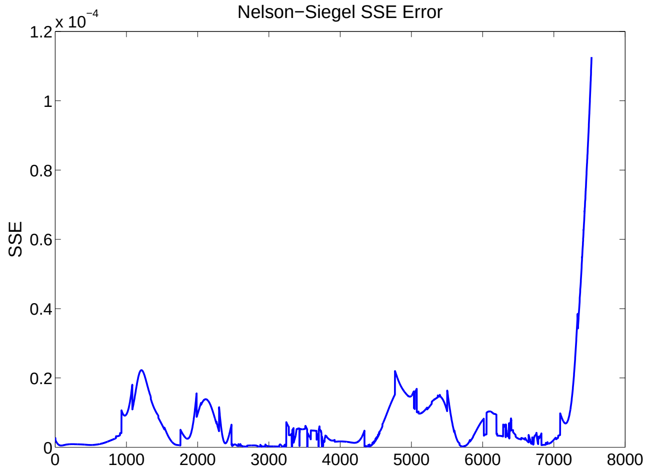

Figure 5.34: Nelson-Siegel Sum of Squares fit error

How to model these series? Since we are still in a world of low yields post the financial crisis, it is fairly reasonable to expect that the yields will rise with time. PK Air had also expressed a desire to have mean-reversion in the model and therefore Ornstein-Uhlenbeck processes are applied, see Section 3.16.

By looking at the time series in Figure 5.33, it seemed as τ was a combination of β_0 and β_1. If this is possible simulation time can be reduced. Therefore the following relationship was tried on historical data with a multiple linear regression. The obtained parameters are shown in Table 5.24 and the Goodness of Fit values in Table 5.25.

$$\tau = v + q\beta_0 + s\beta_1 \tag{5.14}$$

The following parameter estimates were obtained

Table 5.24: Estimated parameters for fitting β_0 and β_1 to τ with 95 % Confidence bounds

parameter	value	Lower Bound	Upper Bound
v	2.0021	2.0021	2.0022
q	-0.0365	-0.0370	-0.0360
s	-0.0367	-0.0374	-0.0360

Table 5.25: Goodness of Fit for regression

Measure	Value
Sum of Squared Errors	8.7159e-004
R-Squared	0.7712
Adjusted R-Squared	0.7711

5.4.1 Calibration of Ornstein-Uhlenbeck processes

To Calibrate the processes Maximum-Likelihood Estimation was applied, see Section 3.13 for a general description and Appendix B for a derivation of the estimates for the Ornstein-Uhlenbeck process. Since PK Air wants the simulation to be as fast as possible, the data set of estimated Nelson-Siegel parameters were chosen to be represented by 30 points instead, with equal distance in between themselves. Later when the simulation is conducted more frequent data sets can be obtained by linear interpolation. The Maximum Likelihood parameter estimation results are shown in Tables 5.26, 5.27, and 5.28.

Table 5.26: Estimated Ornstein-Uhlenbeck parameters β_0

parameter	value
μ	0.0526
λ	0.0994
σ	0.0053

Table 5.27: Estimated Ornstein-Uhlenbeck parameters β_1

parameter	value
μ	-0.0279
λ	0.7090
σ	0.0133

The condition in Nelson Siegel's model that $\beta_0 + \beta_1 > 0$ had still to be governed. To cope with these negativity problems, a help function was introduced as soon as their sum was smaller than 0.01. If this study has been done 2 years ago, it would probably have been unnecessary with this "pillow function", but since the financial crisis and its aftermath is considered as an extreme event, these kind of helping functions is justified. If the sum the

Table 5.28: Estimated Ornstein-Uhlenbeck parameters β_2

parameter	value
μ	-0.0077
λ	1.7382
σ	0.0274

two parameters is defined as

$$f(i) = \beta_0(i) + \beta_1(i) \tag{5.15}$$

Then the help function $g(i)$ is

$$g(i) = \begin{cases} ae^{\frac{f(i)}{a}-1}, & if \quad f(i) < 0.01 \\ f(i), & if \quad f(i) \geq 0.01 \end{cases}$$

a graphical illustration of the function is given in Figure 5.35. As can bee seen it is linear above the threshold of 0.01.

The "help" from the function is then portioned out in the following way

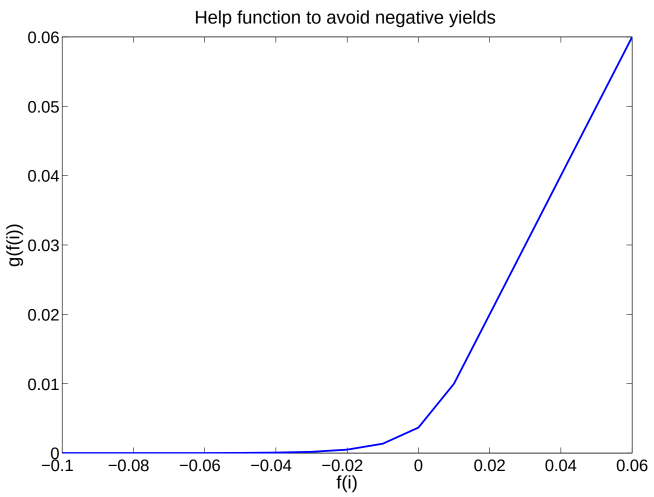

Figure 5.35: Help function with threshold 0.01

$$\beta_0(i) = \frac{\beta_0(i)}{(\beta_0(i) + abs(\beta_1(i)))} \cdot \left(ae^{\frac{f(i)}{a}-1} - f(i)\right) + \beta_0(i) \tag{5.16}$$

$$\beta_1(i) = \frac{\beta_1(i)}{(\beta_0(i) + abs(\beta_1(i)))} \cdot \left(ae^{\frac{f(i)}{a}-1} - f(i)\right) + \beta_1(i) \tag{5.17}$$

as can bee seen from the equations above the help is related to their respective proportions. After non-negativity is assured, the next step is to generate the yields with the Nelson-Siegel function, see equation 3.44, the results are shown in Figures 5.36 and 5.37.

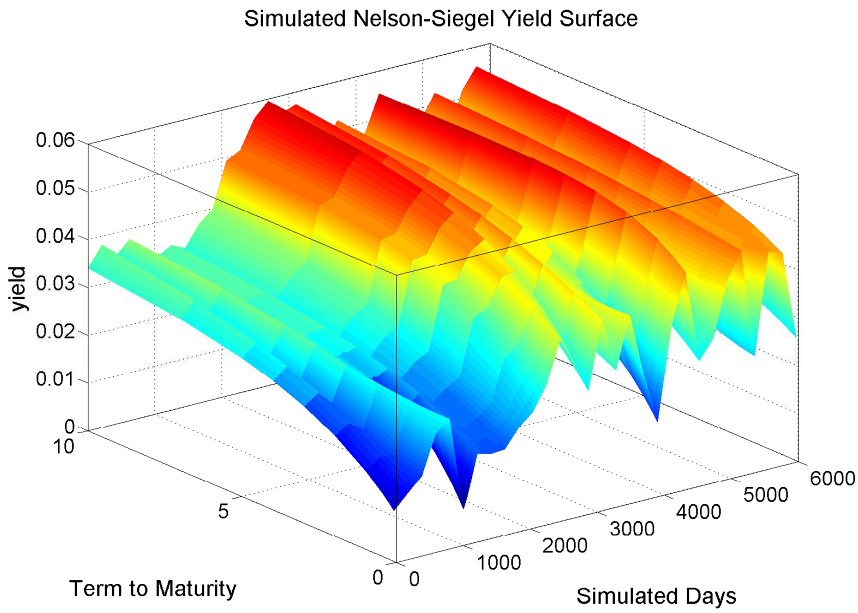

Figure 5.36: Simulated PURCS Cycle Neutral Yield Surface

The final step of the simulation is to add a simulated future PURCS Cycle influence, by rearranging equation (5.13) to

$$y\left(i,j\right)=z\left(i,j\right)\cdot\left(1+c\left(i\right)\cdot d\right) \tag{5.18}$$

the final result is shown in Figures 5.38 and 5.39

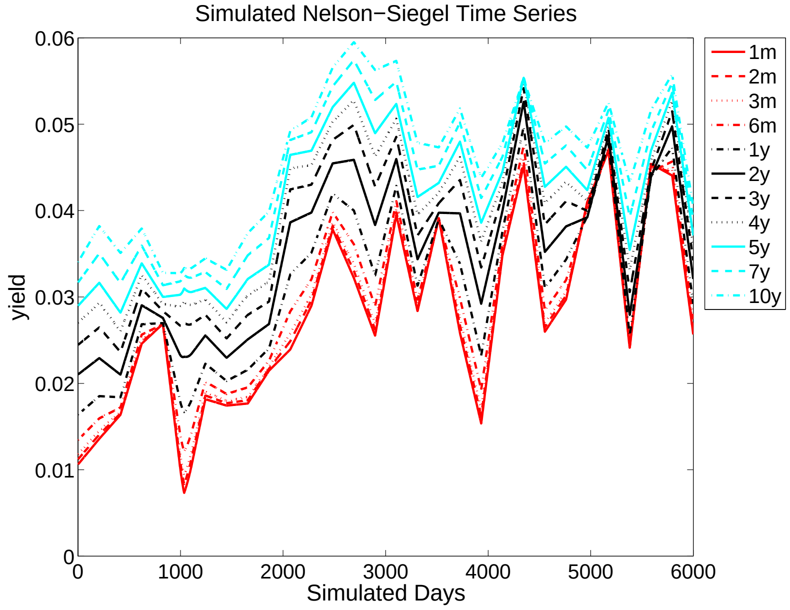

Figure 5.37: Simulated PURCS Cycle Neutral Yield Time Series

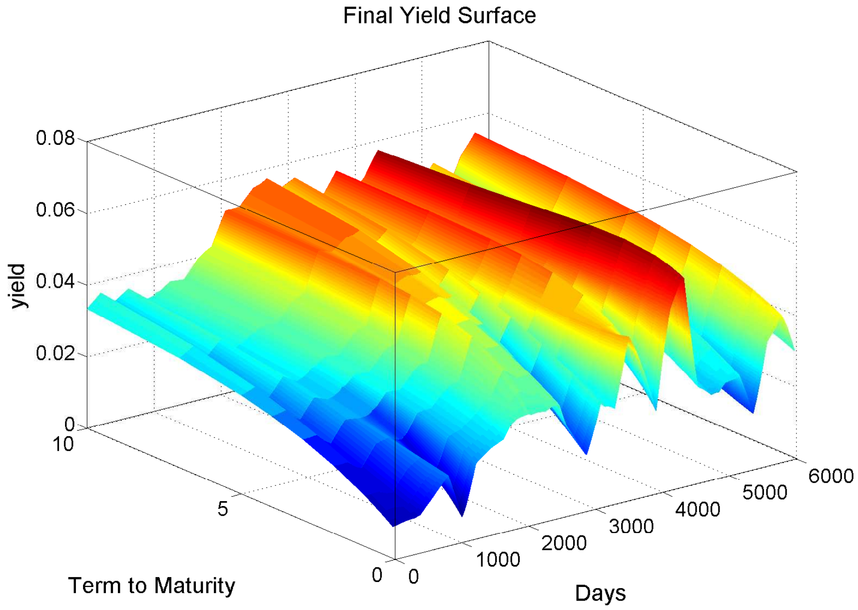

Figure 5.38: Final Simulated Yield Surface

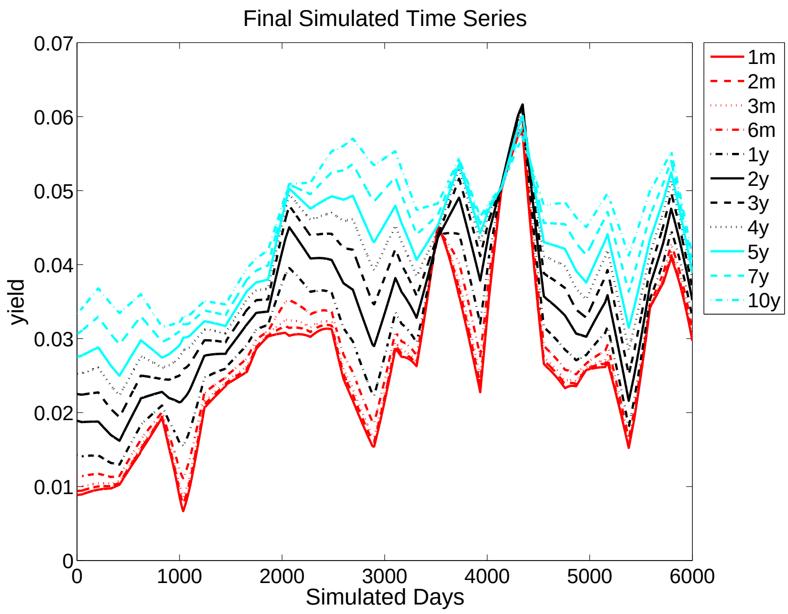

Figure 5.39: Final Simulated Yield Time Series

To produce statistics over the simulations, 10000 yield surfaces were generated and percentiles were calculated to visualize the results. The results for the 1 month, 1 year and 10 year term to maturities over the 30 time steps are shown in Figures 5.40, 5.41 and 5.42. The Cyclical behavior is as expected more evident in the shorter maturities.

The same results are available for all term to maturities and can in turn be used for PK Air's Swap Breakage Scenario calculation.

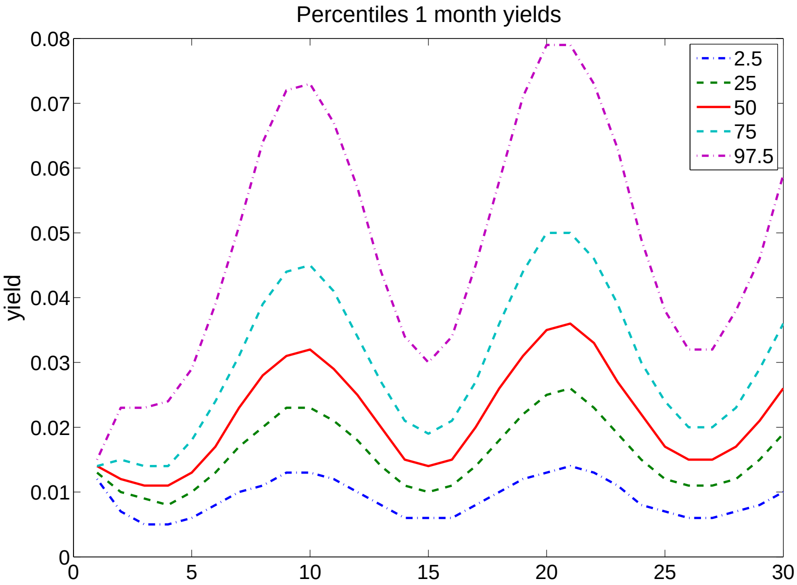

Figure 5.40: Percentiles for 10 000 simulations of 1 month yield

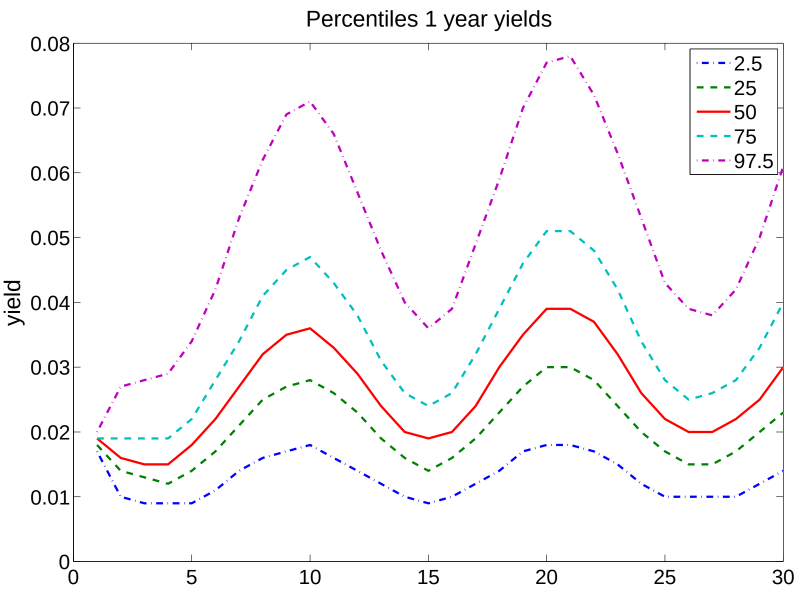

Figure 5.41: Percentiles for 10 000 simulations of 1 year yield

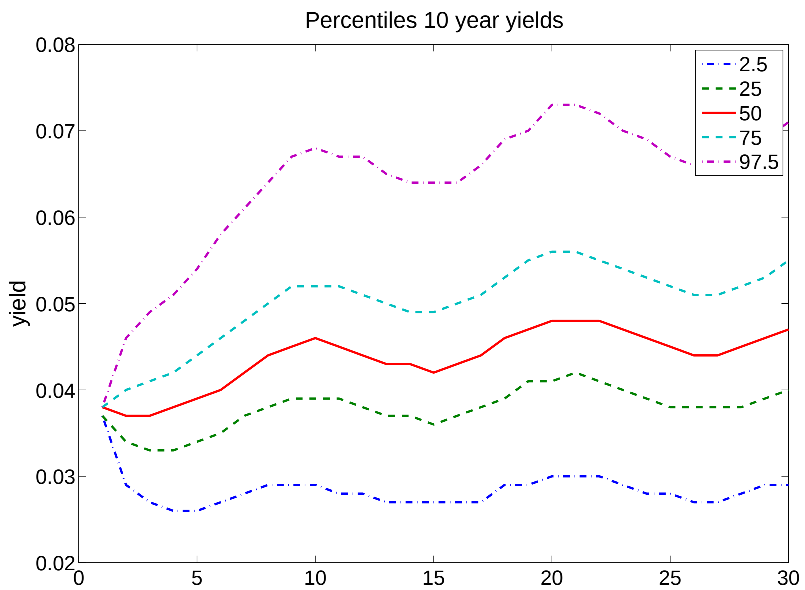

Figure 5.42: Percentiles for 10 000 simulations of 10 year yield

Chapter 6

Discussion

6.1 Conclusions

With Principal Component Analysis, the results were analogue with precedent studies. The first three components are enough to describe the variation of the yield curve and the inclusion of shorter maturities decreases the explanatory power.

Long horizon simulations of the yield curve from the *Principal Component Representation* did not yield plausible results. The term structure loses its shape and this historical simulation procedure exhibits the same trend as the historical data. This trend results in either negative yields or a "zero carpet" depending on the valuation basis.

The "Rebonato approach" also did not yield plausible results. The problem with this approach was the implementation of the spring constants by matching the simulated variance of the curvature to the historical ones. It is not stated in their paper on how to do this, and efforts in this thesis failed.

Because of the lack of long-term models in the Real World Measure the Brute Force Model was developed. The model manages to produce plausible future yield curve scenarios *with cyclical influence*, which was demanded by PK Air since it influences all other parts of their model. The model also follows the other criteria specified by PK Air: tractability, few parameters and fast to simulate. For instance, it is easy to edit the input parameters of the Ornstein-Uhlenbeck processes with PK Air's view of the future. The help function to assure non-negativity is justified since the financial crisis is an extreme event - and must be treated thereafter. This author believes that this function is better than ordinary methods such as the *absorption* and *reflection* methods.

The model could be even better with some refinement and development. Maybe some people argue that the PURCS Cycle effect is to large and that it could be wise to investigate another way of linking the yields to the cycle than what has been done here. Some information is also lost by smoothing the yields, representing frequent data with fewer points, fitting functions and manipulating the data back and forth. This could all be added to a "noise part" in the simulation that could be added in the end, but this needs further investigation.

The Ornstein-Uhlenbeck process could probably be replaced by an even more sophisticated process, e.g. with varying volatility, mean-reversion level and speed of mean-reversion.

The fitting techniques with least squares, regression and maximum-likelihood could also be inspected and developed. The same holds for the Nelson-Siegel model, there is an extension of the model - the Nelson-Siegel and Svensson model, which introduces another parameter and allows for a second "hump" on the curve. But as always one is limited by time so this lies outside of the scope of this thesis.

Correlation is often criticized of being a limited measure of dependence. It could be interesting to look at other dependence measures, e.g. cointegration and copulas.

The data set used for the Brute Force Model is the US LIBOR and Swaps rates, it would be interesting to estimate the model from a longer set of data, e.g. the Treasury Yields, but they have the drawback of not having the shorter maturities for the longer period. Another topic for discussion is how valid the data from 1962 are? Has the financial markets changed since then in terms of globalization and size? This author believes that the US LIBOR and Swaps are more relevant, although the longer period Treasury Yields could serve as a reference.

6.2 Future Studies

During the work with this thesis many ideas and questions arose, some of them have been stated in the previous section. Unfortunately these lie out of the scope of this thesis. Some of these are:

- Investigate if the use of the Nelson-Siegel and Svensson model improves the model and also consider other curve fitting techniques.

- Study the estimation techniques used and see if these could be replaced or refined by even better ones, e.g. Maximum Likelihood, Least

Squares and Regression.

- Evaluate the link between the yield curve and the business cycle - are there other methods?
- Develop or replace the Ornstein-Uhlenbeck processes.
- Evaluate the forecasting ability of the model.
- Estimate the model from other data sets, e.g. other lengths and countries.
- Investigate other dependence measures than correlation, for instance cointegration and copulas.

To conclude this model presents one approach for simulating yield curves under the influence of a business cycle. From the research by this author this is one of the first in this field. Therefore this author hopes it could serve as a foundation for further developments and refinements.

Appendix A

PCA Results

Here are the results of PCA on T-notes between 1962 and 2009 with 5 year windows.

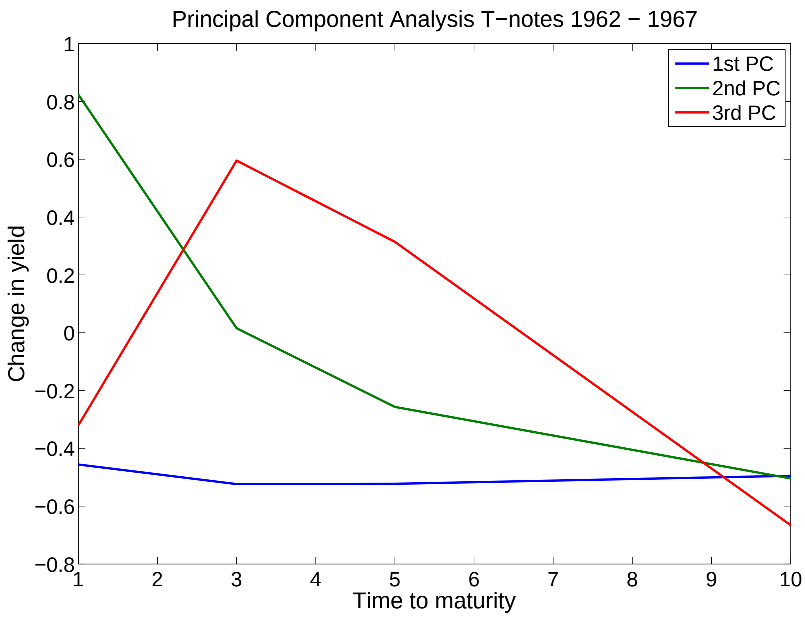

Figure A.1: Principal Components T-notes 1962-1967

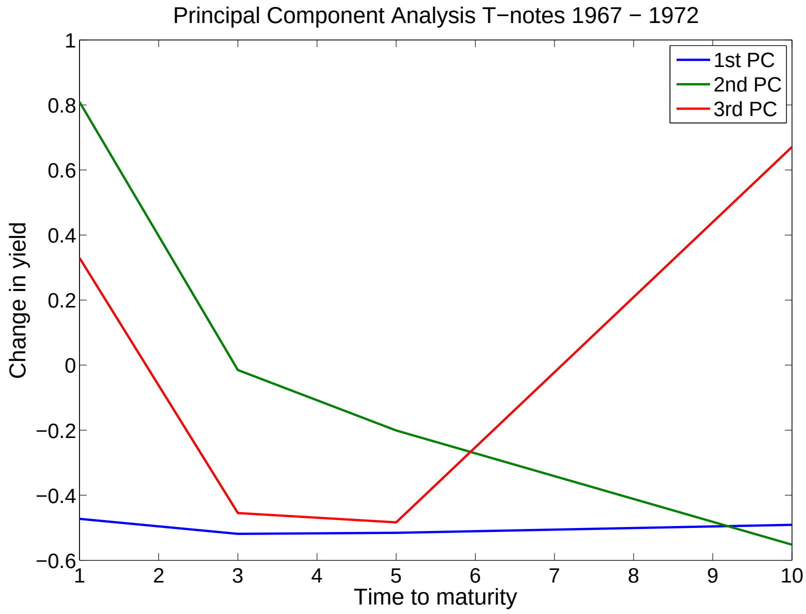

Figure A.2: Principal Components T-notes 1967-1972

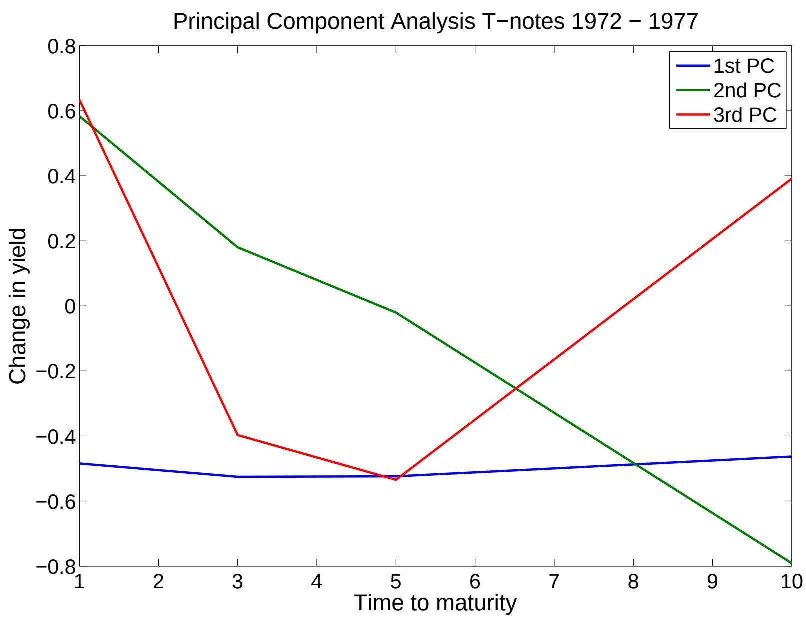

Figure A.3: Principal Components T-notes 1972-1977

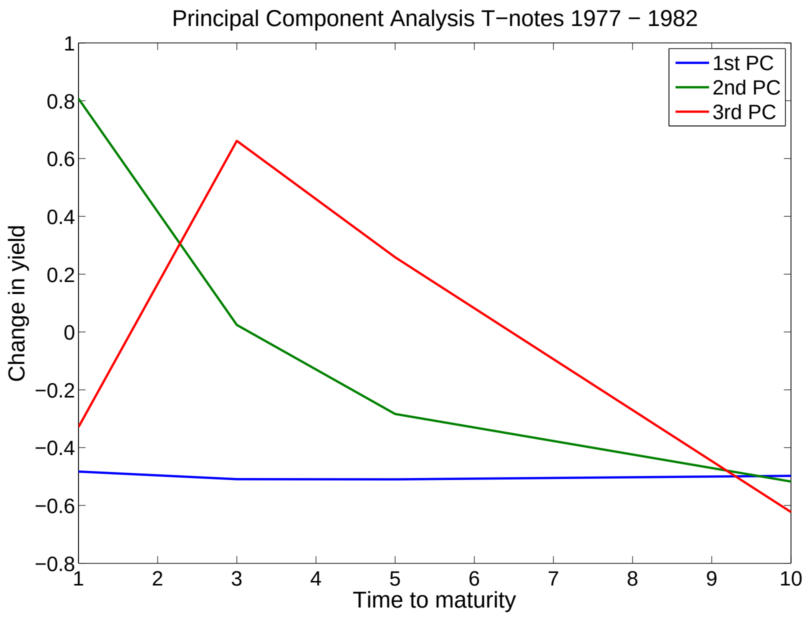

Figure A.4: Principal Components T-notes 1977-1982

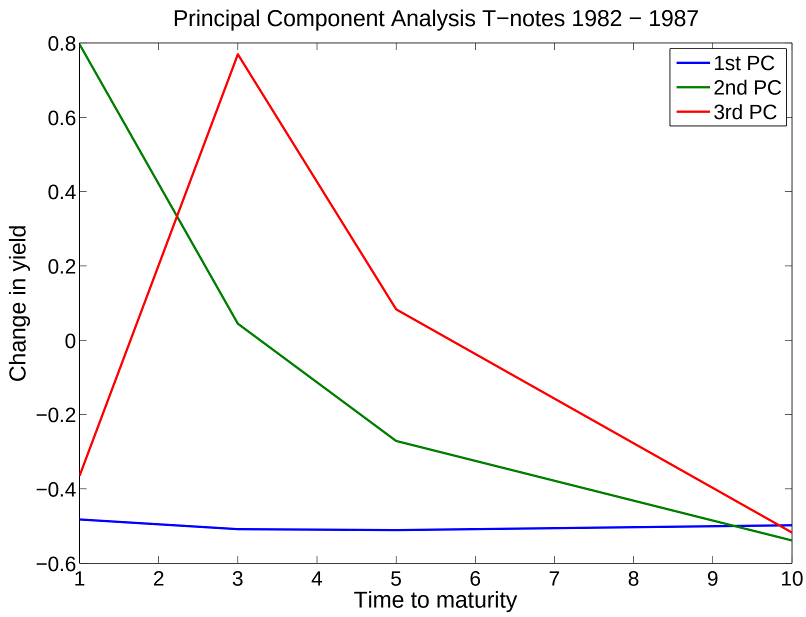

Figure A.5: Principal Components T-notes 1982-1987

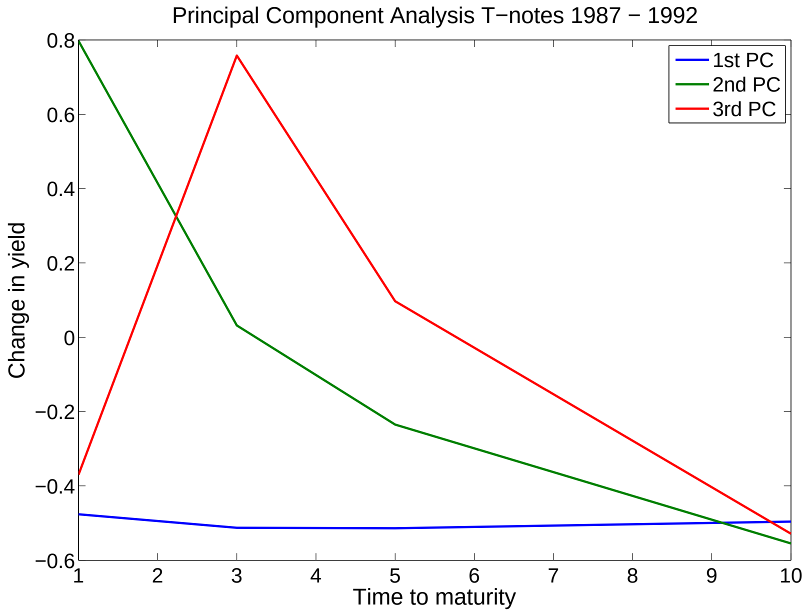

Figure A.6: Principal Components T-notes 1987-1992

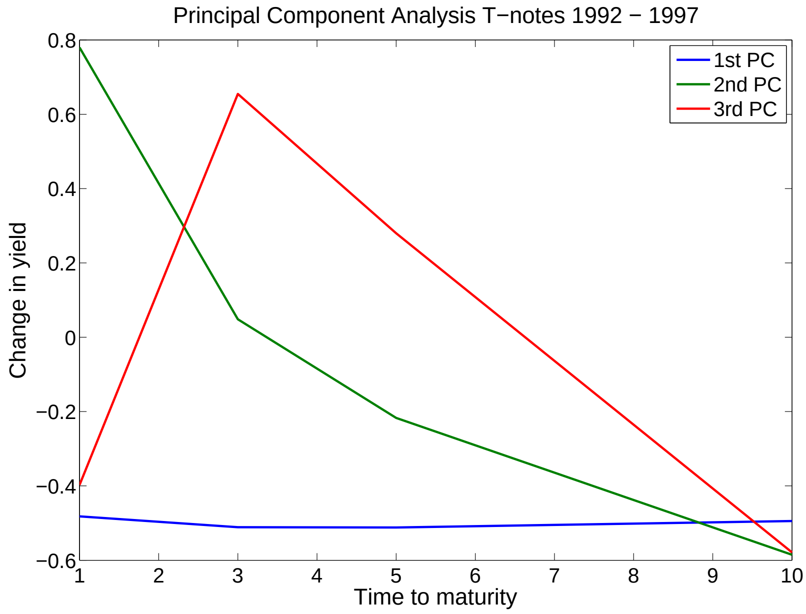

Figure A.7: Principal Components T-notes 1992-1997

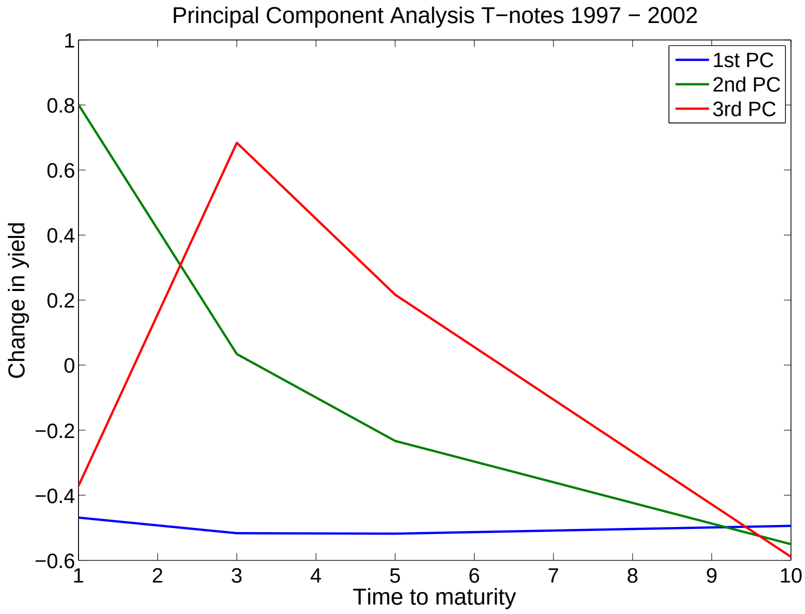

Figure A.8: Principal Components T-notes 1997-2002

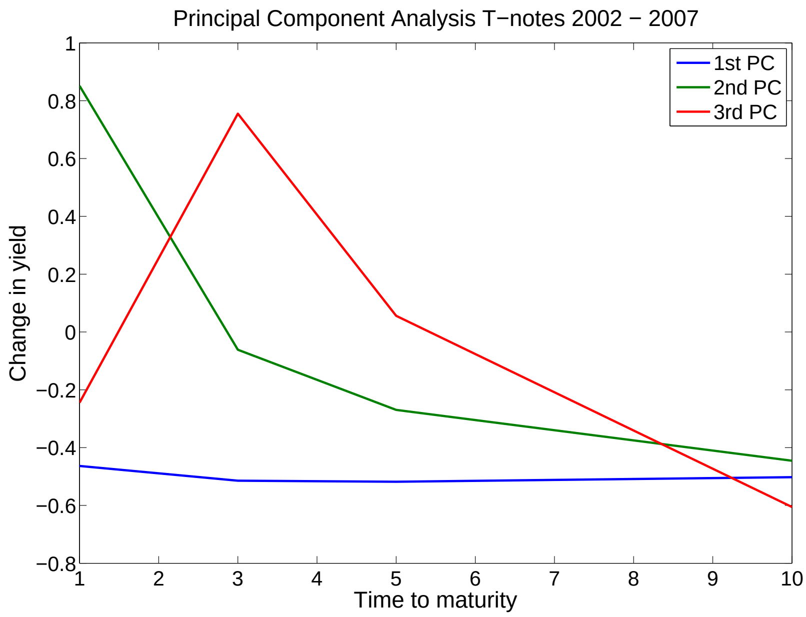

Figure A.9: Principal Components T-notes 2002-2007

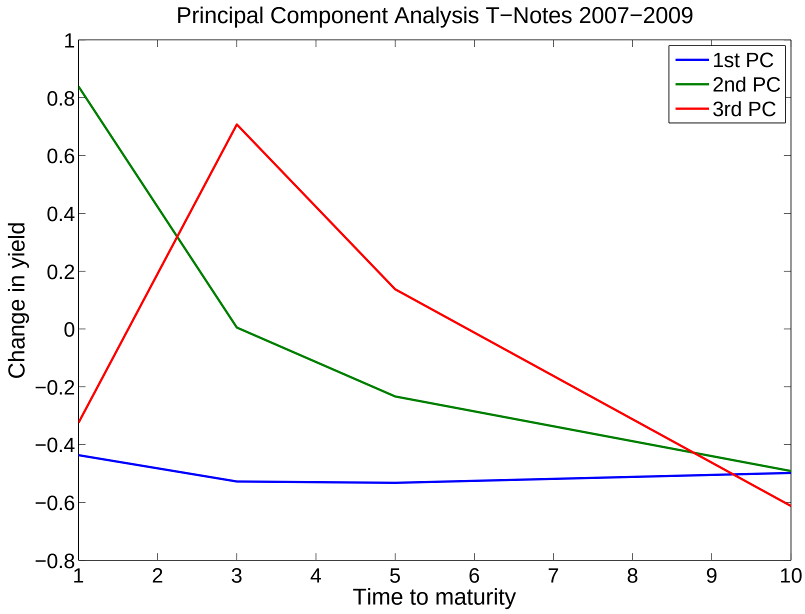

Figure A.10: Principal Components T-notes 2007-2009

Appendix B

Maximum Likelihood Estimation Ornstein-Uhlenbeck

From the solution from Section 3.16

$$X_t = \theta + e^{-\kappa(t-s)} (X_s - \theta) + \sigma \int_s^t e^{-\kappa(t-u)} dW_u$$

the conditional mean and variance can be derived

$$E[X_t \mid X_s] = \theta + e^{-\kappa\delta} (X_s - \theta)$$

where δ has been introduced for the time step $(t - s)$.

$$\begin{aligned} Var[X_t \mid X_s] &= E\left[\left(\sigma \int_s^t e^{-\kappa(t-u)} dW_u\right)\right] \\ &= \text{(by Itô Isometry)} \\ &= E\left[\sigma^2 \int_s^t e^{-2\kappa(t-u)} d_u\right] \\ &= \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\delta}\right) \end{aligned}$$

Therefore X_t is normally distributed with $E[X_t \mid X_s] = \theta + e^{-\kappa\delta} (X_s - \theta)$ and $Var[X_t \mid X_s] = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\delta}\right)$.

From this the conditional probability density function for X_{i+1} given X_i with time step δ as

$$f(X_{i+1} \mid X_i; \kappa, \theta, \sigma) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} exp\left[-\frac{\left(X_i - X_{i-1}e^{-\kappa\delta} - \theta\left(1 - e^{-\kappa\delta}\right)\right)^2}{2\hat{\sigma}^2}\right]$$

where

$$\hat{\sigma}^2 = \sigma^2 \frac{1 - e^{-2\kappa\delta}}{2\kappa}$$

The log-likelihood function for a sample $X_0, X_1, ..., X_n$ is derived from the conditional density function [22]

$$\begin{aligned}
\ell(\kappa, \theta, \hat{\sigma}) &= \sum_{i=1}^{n} log f(X_i \mid X_{i-1}; \kappa, \theta, \hat{\sigma}) \\
&= -\frac{n}{2} log(2\pi) - n \cdot log(\hat{\sigma}) - \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^{n} \left[X_i - X_{i-1} e^{-\kappa\delta} - \theta \left(1 - e^{-\kappa\delta} \right) \right]^2
\end{aligned}$$

To find the maximum of the log-likelihood surface the partial derivatives is set to zero and solved for

$$\begin{aligned}
\frac{d\ell(\kappa, \theta, \hat{\sigma})}{d\theta} =& 0 = \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^{n} \left[X_i - X_{i-1} e^{-\kappa\delta} - \theta \left(1 - e^{-\kappa\delta} \right) \right] \\
&\Rightarrow \theta = \frac{\sum_{i=1}^{n} \left[X_i - X_{i-1} e^{-\kappa\delta} \right]}{n \left(1 - e^{-\kappa\delta} \right)}
\end{aligned}$$

$$\begin{aligned}
\frac{d\ell(\kappa, \theta, \hat{\sigma})}{d\kappa} =& 0 = -\frac{\delta e^{-\kappa\delta}}{2\hat{\sigma}^2} \sum_{i=1}^{n} \left[(X_i - \theta)(X_{i-1} - \theta) - e^{-\kappa\delta} (X_{i-1} - \theta)^2 \right] \\
&\Rightarrow \kappa = -\frac{1}{\delta} log \frac{\sum_{i=1}^{n} (X_i - \theta)(X_{i-1} - \theta)}{\sum_{i=1}^{n} (X_{i-1} - \theta)^2}
\end{aligned}$$

$$\begin{aligned}
\frac{d\ell(\kappa, \theta, \hat{\sigma})}{d\hat{\sigma}} =& 0 = \frac{n}{\hat{\sigma}} - \frac{1}{\hat{\sigma}^3} \sum_{i=1}^{n} \left[X_i - X_{i-1} e^{-\kappa\delta} - \theta \left(1 - e^{-\kappa\delta} \right) \right]^2 \\
&\Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left[X_i - X_{i-1} e^{-\kappa\delta} - \theta \left(1 - e^{-\kappa\delta} \right) \right]^2
\end{aligned}$$

The conditions depend on each other. Although θ and κ are independent of $\hat{\sigma}$. Therefore θ can be found be substituting κ into the expression for θ.

To faciliate some notations are introduced

$$S_x = \sum_{i=1}^{n} S_{i-1}$$
$$S_y = \sum_{i=1}^{n} S_i$$
$$S_{xx} = \sum_{i=1}^{n} S_{i-1}^2$$
$$S_{xy} = \sum_{i=1}^{n} S_{i-1} S_i$$
$$S_{yy} = \sum_{i=1}^{n} S_i^2$$

results in

$$\theta = \frac{S_y - e^{-\kappa\delta} S_X}{n\left(1 - e^{-\kappa\delta}\right)}$$
$$\kappa = -\frac{1}{\delta} ln \frac{S_{xy} - \theta S_x - \theta S_y + n\theta^2}{S_{xx} - 2\theta S_x + n\theta^2}$$

substituting κ into θ yields

$$n\theta = \frac{S_y - \left(\frac{S_{xy} - \theta S_x - \theta S_y + n\theta^2}{S_{xx} - 2\theta S_x + n\theta^2}\right) S_X}{1 - \left(\frac{S_{xy} - \theta S_x - \theta S_y + n\theta^2}{S_{xx} - 2\theta S_x + n\theta^2}\right)}$$

remove denominators

$$n\theta = \frac{S_y \left(S_{xx} - 2\theta S_x + n\theta^2\right) - \left(S_{xy} - \theta S_x - \theta S_y + n\theta^2\right) S_x}{\left(S_{xx} - 2\theta S_x + n\theta^2\right) - \left(S_{xy} - \theta S_x - \theta S_y + n\theta^2\right)}$$

collecting terms

$$n\theta = \frac{\left(S_y S_{xx} - S_x S_{xy}\right) + \theta\left(S_x^2 - S_x S_y\right) + \theta^2 n\left(S_y - S_x\right)}{\left(S_{xx} - S_{xy}\right) + \theta\left(S_y - S_x\right)}$$

and this yields

$$n\theta\left(S_{xx} - S_{xy}\right) - \theta\left(S_x^2 - S_x S_y\right) = \left(S_y S_{xx} - S_x S_{xy}\right)$$

and the final expressions are obtained as

$$\theta = \frac{\left(S_y S_{xx} - S_x S_{xy}\right)}{n\left(S_{xx} - S_{xy}\right) - \left(S_x^2 - S_x S_y\right)}$$

$$\kappa = -\frac{1}{\delta} ln \frac{S_{xy} - \theta S_x - \theta S_y + n\theta^2}{S_{xx} - 2\theta S_x + n\theta^2}$$

$$\hat{\sigma}^2 = \frac{1}{n}\left[S_{yy} - 2\alpha S_{xy} + \alpha^2 S_{xx} - 2\theta\left(1-\alpha\right)\left(S_y - \alpha S_x\right) + n\theta^2\left(1-\alpha\right)^2\right]$$

with

$$\sigma^2 = \hat{\sigma}^2 \frac{2\lambda}{1-\alpha^2}$$

and

$$\alpha = e^{-\kappa\delta}$$

Bibliography

[1] C. Alexander. *Market Models - A Guide to Financial Data Analysis*. John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2001.

[2] British Bankers Association. Libor rate. `http://www.bbalibor.com/bba/jsp/polopoly.jsp?d=1625`.

[3] C. Bernadell, J. Coche, and K. Nyholm. Yield curve prediction for the strategic investor. Technical report, Risk Management, European Central Bank, 2005.

[4] G. Blom, J. Enger, G. Englund, J. Grandell, and L. Holst. *Sannolikhetsteori och statistikteori med tillämpningar*. Studentlitteratur, Studentlitteratur, Lund, 2005.

[5] D. Bolder and D. Stréliski. Yield curve modelling at the bank of canada. Technical report, Bank of Canada, 1999.

[6] G. R. Duffee. Idiosyncratic variation of treasury bill yields. *The Journal of Finance*, 2:527–550, 1996.

[7] R. Fiori and S. Iannotti. Scenario based principal component value-at-risk: An application ot italian banks' interest rate risk exposure. Technical report, Banca d'Italia, 2006.

[8] Financial Trend Forecaster. Us historical inflation data. `http://inflationdata.com/Inflation/Inflation_Rate/HistoricalInflation.aspx`.

[9] O. Hinton. Describing random sequences. `http://www.staff.ncl.ac.uk/oliver.hinton/eee305/Chapter6.pdf`.

[10] J. Hull. *Options, Futures and other Derivative Securities, 2nd Ed.* Prentice Hall, Englewood Cliffs, New Jersey 07632, 1993.

[11] R. Litterman and J. Scheinkman. Common factors affecting bond returns. *The Journal of Fixed Income*, pages 54–61, June 1991.

[12] D. G. Luenberger. *Investment Science*. Oxford University Press, Inc., 198 Madison Avenue, New York, NY 10016, USA, 1998.

[13] Mathworks. Curve fitting toolbox user's guide. Technical report, Mathworks, 2008.

[14] C. R. Nelson and A.F. Siegel. Parsimonious modeling of the yield curves. *Journal of Business*, 60:473–489, 1987.

[15] Federal Reserve Bank of St. Louis. Economic research. `http://research.stlouisfed.org/fred2/categories/115`.

[16] W. Phoa. Yield curve risk factors: Domestic and global contexts. Technical report, The Capital Group of Companies.

[17] R. Rebonato, S. Mahal, M. Joshi, L-D. Buchholz, and K. Nyholm. Evolving yield curves in the real-world measures: a semi-parametric approach. *The Journal of Risk*, 7:29–61, 2005.

[18] R. Rebonato and K. Nyholm. Long horizon yield curve forecasts: Comparison off semi-parametric and parametric approaches. *Applied Financial Economics*, 18:1597–1611, 2008.

[19] A. Santomero and D. Babbel. *Financial Markets, Instruments and Institutions, Irwin McGraw-Hill.* Irwin McGraw-Hill, Boston, Massachusetts, 1997.

[20] K. P. Scherer and M. Avellaneda. All for one... one for all?: A principal component analysis of latin american brady bond debt from 1994 to 2000. Technical report, Courant Institute of Mathematical Sciences, New York University, United States, 2000.

[21] G. M. Soto. Using principal component analysis to explain term structure movements: Performance and stability. *Progress in Economics Research*, 8:203–225, 2004.

[22] M.A. van den Berg. Ornstein-uhlenbeck process. `http://www.sitmo.com/doc/Ornstein-Uhlenbeck_process`.