Interest rate models for estimating counterparty credit risk

Dynamic Nelson-Siegel and Displaced Diffusion

Delft University of Technology

Thesis

Menzo Rood (4167341)

Supervisors:
Prof. Dr. Ir. C.W. Oosterlee
Dr. Ir. X. Huang
Dr. Ir. B. Zhang

Committee member:
Dr. ir. J.H.M. Anderluh

August 18, 2020

Abstract

In this study, two interest rate models are analysed in context of counterparty credit risk. The goal of the study is to find a model that performs well on historical simulation for the PFE and EPE. The two models analysed are the Dynamic Nelson-Siegel model and the Displaced Diffusion model. In the Dynamic Nelson-Siegel model, a Nelson-Siegel curve is fitted against the historical yield curves. The fit gives an historical series of the parameter values of the Nelson-Siegel curve, which are modelled via a stochastic process to obtain future yield curve predictions. In historical backtesting, the classic model using AR(1) processes for the parameters performs inadequate. Analysis on the underlying assumptions of the model show that the mean-reverting behaviour that is modeled is the cause. In addition the data is likely to feature heteroskedastic behaviour, which is not incorporated by the model. An adjusted model in which one parameter is modeled with a random walk with drift performs well on longer maturity rates, however shorter maturity rates are not modeled satisfactory. The Displaced Diffusion model uses a lognormal diffusion process that is shifted to model Libor rates. As it is a Libor market model, all libor rates are modelled seperately using correlated Brownian motions. The shift parameter allows negative rates to be modeled, and is initially assumed constant. The backtesting results are mixed; some observed libor rates are modeled inadequately and some cannot be rejected to come from the Displaced Diffusion model and thus are modelled correctly. When backtesting the PFE, the results are good at the short term. At the 2-year window, PFE estimates are not always conservative but the number of excesses are of medium severity when compared to the probabilities used in the green-orange-red system dictated by the Basel committee for VaR backtesting.

Keywords: Counterparty Credit Risk, Interest Rate Models, Dynamic Nelson-Siegel, Displaced Diffusion, Historical Backtesting, Model Validation

Acknowledgements

This report is written as part of the MSc program Applied Mathematics at Delft University of Technology. I'm very grateful for the useful comments and help received by my ING supervisors and TU Delft supervisor. Xinzheng added many useful comments and also asked for realism in my ideas, Bowen explained me many things and provided many ideas and papers. I found the combination of guidance very good. Kees provided many helpful comments and support over the whole project. A special thanks goes out to the ING Model Validation Trading Risk team for being a welcoming and incredibly talented team. I would like to thank Alejandro in particular, who despite being a very experience professional makes time to help the intern with technical problems. More than once. I would also like to thank my family and friends for the many laughs and small moments of happiness during a pandemic.

Master's Thesis

This study is performed as a partial fulfillment of the requirements for the MSc degree in Applied Mathematics at Delft University of Technology

Contents

Summary

The goal of this study is to find an interest rate model that can be used as benchmark in calculating the Potential Future Exposure (PFE) and Expected Positive Exposure (EPE), both Counterparty Credit Risk measures. The reason for this study is that current models sometimes give unsatisfactory results when the models are used in backtesting.

The available data in the study are euro zero rate curves from 01-01-2007 to 27-02-2019 and thus spanning well over 12 years. Data quality is high, in the delivered dataset no points are missing. Preliminary analysis on the data show multiple phenomena:

- Negative rates are observed
- Interest rates are tested to be heteroskedastic on the short to medium term
- Interest rates are tested to have constant variance on the longer term
- Mean reversion couldn't be statistically shown
- Correlations between rates are generally high when the maturities are close in the tenor structure, and near zero for rates further away in the tenor structure

The second and third point together may indicate that different maturities possibly are generated by different stochastic processes. In addition to the modeling criteria, a benchmark-model is preferably easy to use and computationally attractive. The two interest rate models that are selected based on those criteria are the Dynamic Nelson-Siegel model and the Displaced Diffusion model. Dynamic Nelson-Siegel is analysed because it is computationally attractive and easy to use. It does however feature mean reversion and doesn't incorporate heteroskedastic behaviour. Displaced Diffusion on the other hand satisfies all modeling criteria, save for the constant variance found to be present in the long maturity rates. As it is a libor-market model it is computationally demanding, as multiple maturity libor rates have to be modeled and calibrated, including the correlation structure.

The performance of the models is tested by means of backtesting. In the backtesting procedure, partial series of the historical data are used as backtesting scenarios. These scenarios may overlap. The model is calibrated to data up to a point in time and subsequently used to make prediction over a time horizon, using the start of the backtesting scenario as inital condition. These predictions are then compared with the realised values at the end of the scenario.

In case of testing the empirical distribution, the realised values are compared by tracking the percentage of paths lower than the realised values. Via arguments of the probability integral transform, all these percentages should follow a uniform distribution. In the case of the PFE, the historical performance is tested by using the simulated rates to price a certain product. For these generated prices, an empirical quantile is calculated. The performance of the model is then tested by comparing the number of exceedences of this quantile. A large number of exceedences of a quantile indicates the model underestimates risk.

In addition to the backtesting, the models are validated as well. The underlying assumptions of the models are analysed to find reasons for possible bad performance on backtesting. The validation also points out possible flaws in the model that could lead to model risk.

Backtesting

The Dynamic Nelson-Siegel model does not perform well on both backtesting the empirical distribution and the PFE. This hold even for zero rate maturities for which the model seemed to be adequate based on the initial data analysis. The model is almost always too conservative, except when the model is used to calculate a Forward Rate Agreement in which ING receives a fixed lag over a long time horizon, then the model is not conservative enough.

The Displaced Diffusion model performs well on only part of the libor rates. The short to medium rates have been found to feature heteroskedastic behaviour, however these rates are not found to be statistically matching with the Displaced Diffusion paths. The longer maturity rates do seem to fit the model well. When backtesting the PFE the model performs adequatly, but on the conservative side.

The PFE results of both model using a green-orange-red light system is shown in table 1. A green light is given when the number of exceedences is lower than the corresponding 95% quantile of the binomial distribution, with n the number of observations in the backtesting experiment and p the PFE quantile. An orange light when that number is between the 95% and 99.9% quantile and a red light for a number of exceedences that falls in the 99.9% quantile. It is clear that Displaced Diffusion performs well overall, as it is not too conservative that no PFE exceedences are measures, and the number of orange results is not too much. It is directly clear that Dynamic Nelson-Siegel does not perform well.

Table 1: Results of PFE backtesting

Fixed rate	FRA type	Maturity (start x end in months)	Test window (in months)	$E_{95\%}^{DNS}$	$E_{99\%}^{DNS}$	$E_{95\%}^{DD}$	$E_{99\%}^{DD}$
K=0.01	payer	6x9	1	0	0	6	3
	payer	6x9	3	0	0	8	4
	payer	24x36	24	0	0	2	0
	receiver	6x9	1	1	0	0	0
	receiver	6x9	3	0	0	6	3
	receiver	24x36	24	20	0	2	1
K variable	payer	6x9	1	0	0	5	2
	payer	6x9	3	0	0	5	2
	payer	24x36	24	0	0	0	0
	receiver	6x9	1	0	0	1	1
	receiver	6x9	3	0	0	0	0
	receiver	24x36	24	20	0	3	0
K=-0.01	payer	6x9	1	0	0	5	2
	payer	6x9	3	0	0	6	1
	payer	24x36	24	0	0	3	0
	receiver	6x9	1	0	0	6	3
	receiver	6x9	3	0	0	0	0
	receiver	24x36	24	0	0	0	0
K=0.05	payer	6x9	1	0	0	0	0
	payer	6x9	3	2	0	0	0
	payer	24x36	24	0	0	0	0
	receiver	6x9	1	0	0	1	1
	receiver	6x9	3	0	0	0	0
	receiver	24x36	24	20	13	0	0

Model Validation

After the backtesting, the underlying assumptions of both models are analysed. The purpose for both is different since the difference in performance. Dynamic Nelson-Siegel is analysed to find the reasons for it's lacking performance in backtesting. Displaced Diffusion seems to be performing adequatly in CCR measures, and thus the assumptions need to be analysed to make further improvements and to see whether model risk is present that may prevent usage of the model in practice.

For Dynamic Nelson-Siegel, the following is found

- The yield curve fit is adequate. The beta parameters represent level, slope and curvature well. However in some cases the observed yield curves feature a second slope. In these cases the Svensson model could be a better fit.

- The assumption that the beta processes are independent is not warranted. The processes show low correlation over the whole process, however when taking parts of the data, correlations persistently takes values close to 1 and -1. An analysis is done where the correlation is incorporated in simulation, estimated based on the daily data in the month prior to the initial value. Results did not improve statistically, however visually it clearly shows the model becomes less conservative.

- The assumption that the beta parameters follow an AR(1) process is not warranted. The first beta parameter, corresponding to level, is found to be represented better with a random walk with drift. The AR(1) processes are sufficient for the other two beta parameters, however the residuals still do not have constant variance. Conditional Heteroskedasticity could not be shown in the residuals.

- The λ-parameter that determines the time on which the curvature takes it's maximum has been analysed by comparing the backtesting results using small changes in it's value. The original model estimated λ as the value for which the curvature takes it's maximum on 2.5 years. The comparative backtesting shows that maximization on 2 years performs slightly better.

Using the knowledge from validating the DNS model, and adjusted model is proposed. Three changes are proposed: define λ on 2 years, model β_1 as a random walk with drift and model the correlations between the parameters using the daily parameter process as a proxy. The adjustments improve the model drastically, and all higher maturity rates fit the model statistically, as well as the 1-month zero rate. Further improvements can be made by correcting the kurtosis of the model, which is found to be too low, as well as some inaccuracies in modeling the β_3 parameter. Caution should be taken with the model as the adjustments may be data-specific.

For Displaced Diffusion, the following is found in the validation process

- The shift parameter is fitted using maximum likelihood on all the seperate libor rate models to test if the assumption of a single constant shift parameter is warranted. Large differences are found; a small shift fits the shorter rates better, while a larger shift is found to be better for longer maturity rates.

- The shift parameter presents a hard lower bound and thus the risk that an interest rate attends values lower than this bound is not in the model. A backtesting experiment in which the shift parameter was estimated using MLE shows that it happens frequently that the interest rates on later times are lower than the estimated shift. A solution is to monitor the shift parameter closely and use conservative expert judgement.

- The residuals when calibrating the model to the entire dataset can't all be shown to be normal. The cause are extreme values, which are not captured by the Gaussian residuals.

- On very long time horizons the model can explode when no drift is fitted, and tend to the lower bound when a negative drift is fitted. On time horizons up to 30 years, this behaviour is not as pronounced.

It is concluded that, given adjustments, both models have advantages. The Displaced Diffusion model is perceived as a good model for benchmarking CCR measures. The reason is that the model performs well in distribution on half of the rates, and is conservative in the cases where the distribution does not match. Especially in PFE estimation the model is predominantly conservative but not too much, since PFE exceedences do take place. In addition the model does not need adjustments that are possibly data-specific, in contrast to the DNS model. A downside of the model is that calibration and simulation is slow, given the use of a correlation matrix based on all rates.

1 Introduction

Counterparty credit risk (CCR) is one of the most important and complex fields in risk management. Gregory (2011) even mentions most market participants see counterparty risk the key financial risk, as a result of the 2007 crisis. The complexity of the subject is due to the uncertainty that is involved in CCR. It concerns the risk of a counterparty failing to meet its obligation. Since contracts can change in value due to changes in the underlying risk factor, this means a default of a counterparty can induce costs to replace the position that were not there at the inception of a contract. Since the value of a contract can change in both directions, this risk is bidirectional, which adds to the complexity. Especially economic downtime makes that a solid CCR policy is necessary to prevent a "domino" effect in case of default.

Crucial in the quantification of CCR is the use of interest rate models. Given a certain financial contract or portfolio, movements in interest rate determine whether or not the holder is exposed to this type of risk. Current popular interest rate models sometimes give unsatisfactory results in estimating certain counterpart risk measures. This study will examine multiple interest rate models in view of this.

The two risk measures central in this research are the estimation of the Potential Future Exposure (PFE) and the Expected Positive Exposure (EPE). The PFE can be seen as a CCR-equivalent of the Value-at-Risk, used in other areas of risk management. It represents a quantile, and broadly speaking says: "If our counterparty defaults, how bad does it get?". The EPE is the average of the positive exposures over the time horizon of a portfolio. Both risk measures are calculated using the assumption that the recovery rate is zero and thus after a default, all profitable payments will defer [1].

The main research question is

> Which interest rate model can used to benchmark estimations of the Expected Positive Exposure as well as the Potential Future Exposure?

To be able to answer this question, multiple sub questions are formulated.

- What characterizes the data to be modeled?
- On which criteria is the choice of interest rate models based?
- How do selected models perform in terms of these criteria?
- Why do the models perform as they do?
- What risk is not captured by the models?

These subquestions offer a stepwise approach to give a meaningful answer to the research question. The first subquestion gives rise to the desirable behaviour potential benchmark models. The second subquestion adds practical criteria to the model selection process, on which two models can be selected. The third research question presents a large part of the study, in that it encompasses backtesting of the models and thus offer insight in the PFE and EPE estimations of the model. The fourth subquestion has the goal of understanding why the models perform well or not. It allows to specify in what cases the models can be used as benchmark, and in what cases it is not usable. The fifth subquestion focusses on the validation of the models. Aspects not included in the model can lead to model risk and need to be analysed in order to be used in a financial environment.

1.1 Scope

Many aspects are relevant in the world of Counterparty Credit Risk. In this section an overview is given of the topics that are researched and the aspects that are not.

[1] Note that non-profitable payments will in almost all cases will not defer, as the counterparty-in-default will be under a curator, processing payments to be done.

This study is confined to interest rate models. Other types of financial instruments could be considered to, such as Forex, Stock and Commodities. However, interest rate derivatives are currently traded in much higher volumes than other products. In addition, only one currency is used, and any cross-currency interest rate products are ignored for the purpose of this study.

The focus of this study is on the implementation of two interest rate models, and analyse to analyse them in a CCR context. As will be reasoned later, the estimation and calibration of parameters is an important criterion for the models to be used in a risk management environment. Existing methods for parameter estimation will be used, and no further effort is done to optimize them.

1.2 Structure

The report starts with an overview of the basic mathematical concepts used in the study. This encompasses two parts, interest rates including interest rate products and counterparty credit risk. chapter 3 is devoted to model selection. Literature and analysis on the data lead to criteria for models to be studied. Baed on these criteria, Diplaced Diffusion and Dynamic Nelson-Siegel are explained to be good contenders and are described shortly. This is then followed by chapter 4 in which the methodology is presented. The main focus of the chapter is introducing the backtesting framework that is used for testing the empirical distribution, the product distribution and the PFE performance. The following two chapters, 5 and 6, contain the full analysis and validation of Dynamic Nelson-Siegel and Displaced Diffusion respectively. The results are then summarised and concluded in chapter 7.

2 Background on Interest Rates and Counterparty Credit Risk

In this section an introduction is given to the basic interest rate products and derivatives used as well as to counterparty risk. It aims to define them mathematically. This section borrows from Gregory (2011), Gregory (2012) and chapters 11, 12 and 14 of Oosterlee and Grzelak (2019).

2.1 Interest Rate Products and Concepts

The most basic form of interest rate is the (continuously compounded) money-market account. The idea is that continuously, at an infinitesimal time, deterministic interest is gained over the initial value. This initial value is standardised at 1. This gives the following ODE

$$\begin{aligned} dM(t) &= r(t)M(t)dt \\ M(t_0) &= 1 \end{aligned} \tag{1}$$

With $r(t)$ deterministic, we can simply integrate to obtain $M(t) = e^{\int_{t_0}^{t} r(s)ds}$.

A well known example of an interest rate product is the bond security. It pays interest on a predetermined amount, called the notional. This percentage of interest paid based on the notional is called coupon. At the end of the life-cycle of the contract, called maturity, a final payment is done of the notional in addition to the final coupon. The types of coupon can be either predetermined, or determined on various times during its lifespan. The first is called fixed-income security, the latter a float-rate note. A special, very basic type is called the Zero-Coupon Bond

Definition 1 (Zero-Coupon Bond) . *The zero-coupon bond, denoted as $P(t, T)$, pays one unit of the currency at time T, with no intermediate payments. We have $P(T, T) = 1$ at maturity T.*

Given the Zero-Coupon Bond, the zero rate can be defined.

Definition 2 (Zero rate) . *Denoting maturity T_i and t_i the date of observation, the zero rate is the constant rate for which the bond can be priced under continuously compounding as*

$$P(t_i, T) = e^{r_{t_i}^{T_i} T_i}. \tag{2}$$

Then $r_{t_i}^{T_i}$ is the zero rate for maturity T_i at time t_i

The instantaneous short-rate is defined in the same way as the money-market account described earlier, as the rate on a riskless investment over an infinitesimal time period from now. The instantaneous forward rate is instead the instantaneous rate at a future time. Suppose we enter a contract, in which at time T_1 we provide a zero-coupon bond with maturity T_2, in which we have $t < T_1 < T_2$. We denote the price of this contract at time t as $P_f(t, T_1, T_2)$. By assuming no-arbitrage and market completeness, if we would buy two zero-coupon bonds, one with maturity T_1 and one with T_2, it would follow that

$$P_f(t, T_1, T_2) = \frac{P(t, T_2)}{P(t, T_1)}.$$

The forward rate is defined as the rate in period $[T_1, T_2]$ such that the value of the contract equals

$$P_f(t, T_1, T_2) = e^{-(T_2 - T_2) r_F(t, T_1, T_2)}.$$

Now, by the above two equalities, the forward rate can be specified in terms of the zero-coupon bonds of the two maturities. We simply equate them to find an expression for $r_F(t, T_1, T_2)$

Definition 3 (Forward rate) *Forward Rate $r_F(t, T_1, T_2)$ is defined as*

$$r_F(t, T_1, T_2) = -\frac{logP(t, T_2) - logP(t, T_1)}{T_2 - T_1}.$$

Then by letting $T_1 \to T_2$ we define the ***instantaneous forward rate*** *$f^r(t, T_2)$ as*

$$f^r(t, T_2) = -\frac{\partial}{\partial t} logP(t, T_2).$$

Strongly related to the notion of forward rates, is the yield curve. A yield of a bond is the interest rate implied by a bond based on the price it is traded on. So if a zero-coupon bond with a maturity of exactly one year costs 99 cents right now, the (implied) interest rate is slightly over 1% on a yearly basis. Bonds of different maturities feature different yields, which warrants the idea of a yield curve. By the way the forward rates are defined, the yields are an important tool to assess forward rates.

Based on the current prices for a certain interest rate product at multiple maturities, discount factors $p(t_i)$ can be calculated. Since these maturities are at discrete time points, the entire yield curve is then interpolated between the discount factors.

The final concept introduced in this section is the (forward) Libor rate.

Definition 4 (Forward Libor Rate) *Using a period $\tau_i = T_i - T_{i-1}$, often called tenor, we define the forward Libor rate $l_t^i = l(t; T_{i-1}, T_i)$ using the zero-coupon bond price $P(t, T_{i-1})$*

$$l_t^i = \frac{1}{\tau_i} \frac{P(t, T_{i-1}) - P(t, T_i)}{P(t, T_i)}. \tag{3}$$

By noting that $P(t, T_0) = P(t, t) = 1$, it is easy to see that for a given tenor structure we can go back and forth between zero rates and Libor rates, which will be convenient when comparing the different models.

The bonds as defined before can be seen as one-way transactions. After an initial payment equal to the present value of the bond, the buyer gains a certain amount of cash over time. A product that differs in this, is the Forward Rate Agreement. With an FRA, there is a receiver and a payer in the contract. The payer will have to pay a fixed rate at specific time with respect to a certain notional. The receiver will receive this fixed rate, and in return will pay a floating rate in return, often based on the Libor rate.

Definition 5 (Forward Rate Agreement) *A Forward Rate Agreement (FRA) is an agreement between two parties in which a fixed rate is exchanged against a floating rate at a specified future time. At a future instant T_i relative to the inception of the contract, the payer will pay a fixed rate K over time period $\tau_i = [T_{i-1} - T_i]$. In return, the receiver pays the Libor rate over the same time period. Forward rate agreements are noted in the text as FRA T_{i-1} x T_i.*

The payoff of the FRA contract is

$$V(T_{i-1}) = \frac{\tau_i (l_{T_{i-1}}^i - K)}{1 + \tau_i l_{T_{i-1}}^i}.$$

The FRA can be defined assuming a notional on which the interest rates are calculated. In this definition, the notional is equal to 1, however any notional can be made by summing various FRAs with the same specifications. Under risk-neutrality, the price of a forward rate agreement (see Oosterlee and Grzelak (2019)) is given by

$$V(t_0) = P(t_0, T_{i-1]}) - P(t_0, T_i) - \tau_i K P(t_0, T_i).$$

2.2 Counterparty Credit Risk

When engaging in a trade of some contract with a counterparty, there is always a risk that the counterparty cannot meet its obligations. Traditional credit risk concerns losses that follow from this. If one would pay for a zero-coupon bond now from partner Q, and Q defaults [2], a loss has been taken that is not foreseen. Note that this type of risk is one-sided, in that the issuer of the bond faces no risk when the counterparty defaults. Counterparty credit risk differs in this, in that it is concerned with profits that are lost when a counterparty defaults. In addition, derivative contracts in which counterparty risk is important are often two-sided. This can be best explained by an example of an interest rate swap.

Suppose company X enters in a swap with company Y. Company X pays a fixed rate of 1% every year for 10 years to Y based on a notional. Company Y pays a yearly rate equal to the Libor rate, which

[2] assuming a 0% recovery

is determined two days before the transaction. We assume that at inception, both parties value the exchange of rates equally, i.e. the value of the contract is zero. However, during the contract's lifespan, the value of the contract will change with the Libor rate. In case the Libor rate will be relatively much larger than the fixed rate, the contract gains in value since it receives a larger rate than the 1% it has to pay. If in this case company Y defaults, it misses out on its expected profits of the deal. On the other hand, if the Libor rate is well below 1% and company X defaults, it is company Y who misses out on a contract with a positive value. This is exactly what counterparty risk is concerned with.

Many risk measures exist to gauge or deal with this type of risk. Before dealing with them, basics such as Mark-to-Market (MM) and (Expected) Exposures will be clarified. In the previous example, it was indicated the deal was valued zero at the start. This is often the case with MtM at the start of the contract. MtM is an intrinsic value of a derivative contract and it is the net present value of all (expected) payments determined by the contract with respect to the counterparty. An exposure is then defined as the cost of replacing the derivative contract in case the counterparty defaults with zero recovery. In the case the intrinsic value is positive, a default of the counterparty means that a loss is incurred, and as such it is called a positive exposure. If on the other hand the intrinsic value is negative, money is owed to the counterparty. In the case of a default of the counterparty, there is however nothing gained, as in most cases a curator takes care of remaining business.

When trying to gain insight in one's exposures, both the current exposure and the exposure in the future are relevant. In this research, the focus is on the exposure in the future. Two measures are in particular important. The first is the Potential Future Exposure. It represents a quantile of the disptribution of exposures. More specifically, it gives the future exposure against a confidence interval of certain magnitude. The second is the Expected Positive Exposure, which is an average of the expected exposure over time. In this, the expected exposure is defined as the expectation of the non-negative values of the MtM.

The Potential Future Exposure is defined in the same way as the Value at Risk, meaning it is just a quantile. The main difference is that it is a quantile of the distribution of potential gains instead of the loss distribution. In addition the PFE is often taken over longer time horizons.

Definition 6 (Potential Future Exposure) *The Potential Future Exposure is defined as the $\alpha-$quantile for which we have*

$$PFE_{\alpha} = min\{\epsilon \in \mathbb{R} : P(E > \epsilon) \leq 1 - \alpha\}. \tag{4}$$

The EPE can be generally defined as in Ghamami and Zhang (2014).

Definition 7 (Expected Positive Exposure)

$$EPE = \int_0^T EE_t dt. \tag{5}$$

In which EE_t is the expected value of the exposure at time $t \geq 0$ and $T > 0$ is the time to maturity for the longest position.

This general definition is not entirely the same as the definition from the Basel framework. As Bonollo et al. (2015) remarked, the official definition of Expected Exposure and EPE already incorporates Monte Carlo simulation in its definition, leading to the following

Definition 8 (EPE Basel) *In the Basel III framework, EPE is defined as*

$$EPE = \frac{\sum_{k=1}^{K} EE_k \cdot \triangle_k}{T}. \tag{6}$$

Here EE_k, the Expected Exposure, is defined as

$$EE_k = \frac{1}{N} \sum_{n=1}^{N} MtM(t_k, S_{k,n})^+,$$

with N the number of paths, MtM the mark-to-market value based on the simulated underlying $S_{k,n}$ at time t_k.

Since the introduction of the EE and EPE measures by the Basel Committee on Banking Supervision, it has been found that these measures underestimate the risk. The problem is that these measure can decrease over time, while in general risk does not decrease over time. This has been counteracted by introducing the Effective EE and Effective EPE. Effective EE is simply a non-decreasing version of the EE, and the Effective EPE is defined as the mean of this Effective EE. The effective measures are not used in the further study, as good estimation of the EE and EPE in principle also lead to good estimates for the effective versions of these risk measures.

3 Model Selection

This chapter starts with a description of the literature review performed. The goal is to gain an understanding of which type of models perform well, and which kind of phenomena are generally modelled in interest rate models. Then some of these phenomena are tested in the data that is provided for this study. Based on this, a list of criteria is made, on which the models are selected that are central in the study. These models are defined and described at the end of the chapter.

3.1 Previous reviews of interest rate models

A comparison of various short-term interest rate models is done in Chan et al. (1992). In that study, various SDE's of the general form

$$dr = (\alpha + \beta r)dt + \sigma r^y dW,$$

were analysed. With certain values for the parameters, this collapses to various well-known interest rate models. For example, with $y = 0$ this is the Vasicek model and with $y = \frac{1}{2}$ this is the CIR square root model. In total there are 8 models compared, including an unrestricted model, in which all parameters are estimated. The study finds that models with $y < 1$ are rejected. The reason is that models that feature a relation between the interest rate volatility and the level of the interest rate score better, leading to the conclusion that this is one of the most important features of interest rate models. An additional conclusion is that the parameter β was found to be insignificant, suggesting the mean-reversion is less important in the short term.

A review more focused on term-structure pricing models is done in Rebonato (2004) and Rebonato (2003). One of the conclusions from both reviews, is the necessity for fast and easy computations in the models used. Especially for pricing this is important, since traders need to be able to hedge in matter

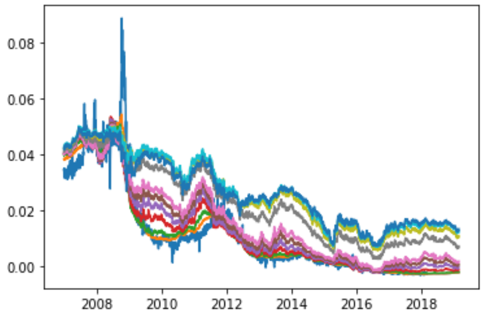


Figure 1: Euro zero rates

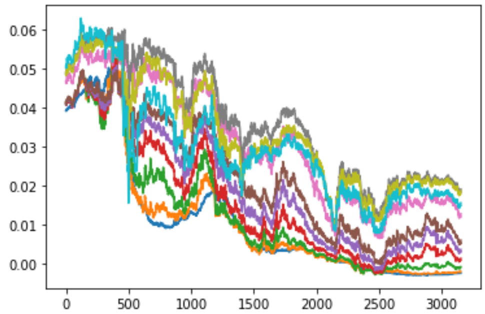


Figure 2: Libor rates

of seconds. However this is also important in backtesting, as often many rates in different currencies and different portfolios are analysed. Rebonato continues that the calibration of models is seen as very important, as well as the ability to model volatility smiles and skews.

In an internal report of ING (see FI/FM-Quants (2019)), the focus is on five criteria. These are the ability to allow for negative interest rates, fitting of the initial curve, computational complexity, achievable and stable calibration, and facilitation of real-world scenarios. The later means that the market price of risk is used to adjust market expectations, and this can be achieved by specifying the dynamics under the real-world measure.

In Aas et al. (2018) a comparison is done on multiple interest rates under Solvency II. The study compares the CIR++ model, G++ model and Libor Market model and uses them to find a best estimate of the liabilities as under the Solvency II framework. The study concludes that, even though the models are quite different in the simulated process and the mean and volatility of the theoretical distribution, the resulting best estimates are very close. Because the underlying distribution are different but certain estimates are very comparable, this can also be interpreted that calibration is of high importance.

This study takes a different view of the research discussed in this section, but also has some similarities. None of the studies compare and describe the models in a counterparty credit risk setting. In this study, the models are used to find historical values for the PFE and EPE, and are subsequently compared with realised values by means of back-testing. In addition, as will be described later in this chapter, the models analysed will not be nested, such as in Chan et al. (1992). In addition, the choice of models in this study will enable to compare the usefulness of having multiple factors as opposed to only having one-factor. Aas et al. (2018) has some similarities with this study, in that it compares quite different models, both in the distribution of the analysed models as well as the modeled rates. There two main differences are that in that study is in a Counterparty Credit Risk setting, as well as the strong focus on historical estimation and back-testing. The aforementioned study is strictly in the risk-neutral world, whereas the efforts in this study are mainly under the real-world measure.

3.2 The data

The data for this research is the pre-processed zero rates for euro, provided by ING. The zero rates are effectively the on-date yield on a Zero Coupon Bond. The data encompasses various maturities $T \in [30/365, 0.5, 1, 2, 3, 4, 5, 10, 15, 20, 30]$ in years. The data is in daily intervals. A visualisation of the zero rates is given in figure 1. The interest rates have been steadily following a downward movement over the whole time period studied. From the zero rates, the Libor rates can be constructed as described in chapter 2. They are shown in figure 2.

It is unconventional to use daily data in counterparty credit risk, given the longer time horizons that are present. In addition there is a strong weekly cycle present in the data, as can be seen in the autocorrelation function shown figure 3. The remainder of the study uses monthly data, generated by

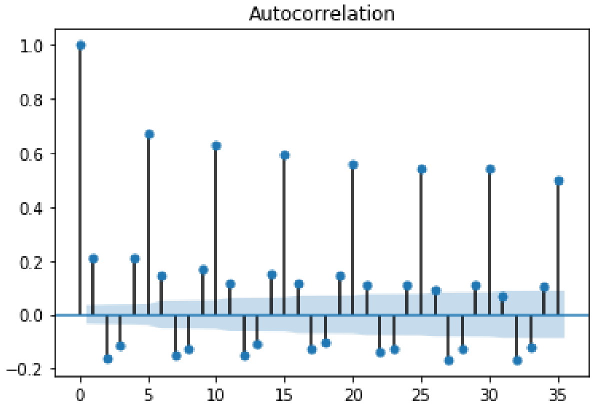


Figure 3: ACF of a short-maturity zero rate

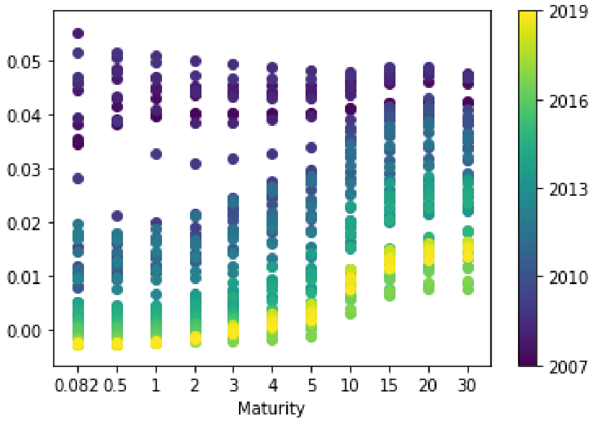


Figure 4: Yield curve in ten-week intervals

taking every 20th datapoint.

Another way of visualizing the data is by scatter-plotting the yield curves over time. This is shown in figure 4. It uses 10 week intervals, and a (sequential) coloring pattern to give insight in the yield curve movement over time. Darker blue colors show the zero rates at earlier times in the dataset and yellow colors corespond to zero rates at the end of the dataset. An initial thought from the visual is that higher maturities seem less volatile than the shorter maturities.. In addition it shows that the interest rates of shorter maturities - between one month and five years - decreased rapidly in the years after the 2008 financial crisis.

Some of the properties mentioned in the literature as given before can be roughly checked on the data to see if they make sense and if they can be used to justify the selection of the models. From the figures explained just before, it is clear that the models need to facilitate negative interest rates. It is however not directly clear if the rates are stationary - there seems to be a downward trend in the interest rates. To gain an impression if this is indeed true in the data, an augmented Dickey-Fuller test (the augmented version of the test introduced in Dickey and Fuller (1979)) has been done for all zero rates, as well as the differences of the zero rates. The null-hypothesis of this test is that the rate has a unit root, meaning the process is not converging to a (constant) long term mean

Based on the first column of table 2 it cannot be rejected that there is a unit root in the time series.

Though this is somewhat unexpected for interest rates, in this case it is perhaps understandable. The interest rates in this study are from 2007 until 2019. This time-horizon is on the short side to really see mean-reverting behaviour. In addition, across this time-period both the effects of the 2007 crisis and the low-rate policy of the ECB afterwards may have resulted in a forced downward pattern [3]. In addition to the ADF test on the levels, the p-values of the ADF test on the differences of the interest rates is shown in the second column of table 2. It is convincing that the monthly changes in interest rates are stationary in the sense that they do not contain a unit root.

In addition to mean-reverting behaviour, one of the important conclusions from the literature review in the previous section is the dependence of variance on the level of the interest rates. Two statistical tests are used, respectively a White test (White (1980)) and a Breusch-Pagan test (Breusch and Pagan (1979)).

The tests for heteroskedasticity need to be applied to some regression model. In this case an AR(1) process is taken. An AR(1) process can be seen as a discretized Ornstein-Uhlenbeck process, using Euler-Maruyama as following

$$r_{t_{i+1}} = r_{t_i} + (\theta - \alpha r_{t_i})\triangle t + \sigma\sqrt{\triangle t} Z_{t_i} = \theta\triangle t + (1-\alpha\triangle t)r_{t_i} + \epsilon_{t_i}$$

The White-test has the null hypothesis that the variance of the error term ϵ_{t_i} of a regression model is constant. From column 4 of table 2 it is statistically significant that shorter rates do seem to be heteroskedastic. However, for the longest maturities, it can't be rejected at the 5% confidence level that the variance of the errors is constant. The Breusch-Pagan test is different in that it only tests for linear heteroskedasticity. In the case of the Breusch-Pagan test, the null hypothesis ís rejected at the the 5% level.

The p-values of the test are shown in columns 3 and 4 of table 2. The difference between shorter and longer maturity rates is striking, and may warrant the idea that different maturity zero rates follow a different type of process. From these tests, the models that will be tested should in principle feature heteroskedasticity for shorter maturity rates. Models which have variance dependent on the interest rate level are expected to perform well on the short and medium range. However at longer maturities this heteroskedastic behaviour is not observed. A normally distributed process may suffice in this case or even outperforms heteroskedastic models. For this reason it is important to test both types of models.

Various statistical tests				
Zero rate maturity	ADF on interest rate	ADF on differences	White-test	BP-test
1-month	0.452	0.000	0.000	0.000
6-month	0.337	0.000	0.000	0.000
1-year	0.286	0.000	0.000	0.000
2-year	0.225	0.000	0.000	0.000
3-year	0.453	0.000	0.000	0.000
4-year	0.555	0.000	0.000	0.000
5-year	0.609	0.000	0.000	0.000
10-year	0.741	0.000	0.109	0.037
15-year	0.809	0.000	0.960	0.776
20-year	0.746	0.000	0.952	0.944
30-year	0.678	0.000	0.623	0.585

Table 2: Statistics and Heuristics

[3] Note however that detrending the zero rates, by fitting a straight line using OLS, also does not lead to a rejection of the null hypothesis that the series contains a unit root

Finally, the distribution of the daily changes in interest rates are analysed to gain an initial insight in the interest rate distribution. Figure 5 shows the monthly changes of zero rates of various maturities. It is clear that for especially the shorter maturity rates, the distribution is not bell shaped and has more extreme observations than a normal distribution. The zero rates with a very high maturity seem to follow the traditional bell shape of the normal distribution better. This re-enforces the statement made earlier that the short maturity and long maturity interest rates may follow different types of processes.

3.2.1 Data quality

The data used in the study are the zero rates from 2007-01-01 till 2019-02-28. The data features daily rates and do not contain any missing values. The data is also tested for stale series, indicating that missing values were already filled the data was obtained for this study. Around 50 datapoints in the daily data did not change over the day. All these points were found in the last few years of the data, in the zero rates with 1 month and 6 month maturity, and are non-consecutive. These rates have shown very small variance in the mentioned time period, which can explain the stale data points. Moreover, when taking transforming the data to monthly, only two stale observations remains present. Overall, the data quality is thus high.

3.3 Model criteria

Using the insights gained fromthe literature and initial data analysis, multiple criteria for the models can be specified. The criteria can be roughly cast into three categories. The first is the modeling category. This states which properties of the observed values should be incorporated in the model. Common properties, as established in the previous section, are mean-reversion in the long term, the ability to facilitate negative interest rates and/or jumps, and the ability to model the yield curve/term structure.

The second category is computational. As mentioned in Rebonato (2003) in practice some models are slightly outdated, but still in use because newer models are often computationally infeasible. Models that have analytical solution are much faster, however this is often not the case for more complex products. In addition, Rebonato states the importance of calibration.

The third category is the practical use in risk management. In estimating the PFE and EPE, overestimation of risk results in unnecessary capital that needs to be held. On the other hand underestimation of these risk factors exposes the firm to preventable risk. The accuracy of the model to use it as benchmark for estimates is part of this category as well.

See table 3 for an overview of the model criteria.

Modeling	Mean-reversion
	Negative rates
	Volatility smiles
	Exact fit of current yield curve
	Correlations in the yield curve
	Features Heteroskedastic behaviour
Computational	Computational complexity
	Ease of calibration
	Adaptable under the real-world measure
Practical	Ease of use
	Effects on capital requirements
	Accuracy in EPE and PFE estimation

Table 3: Criteria for interest rate models

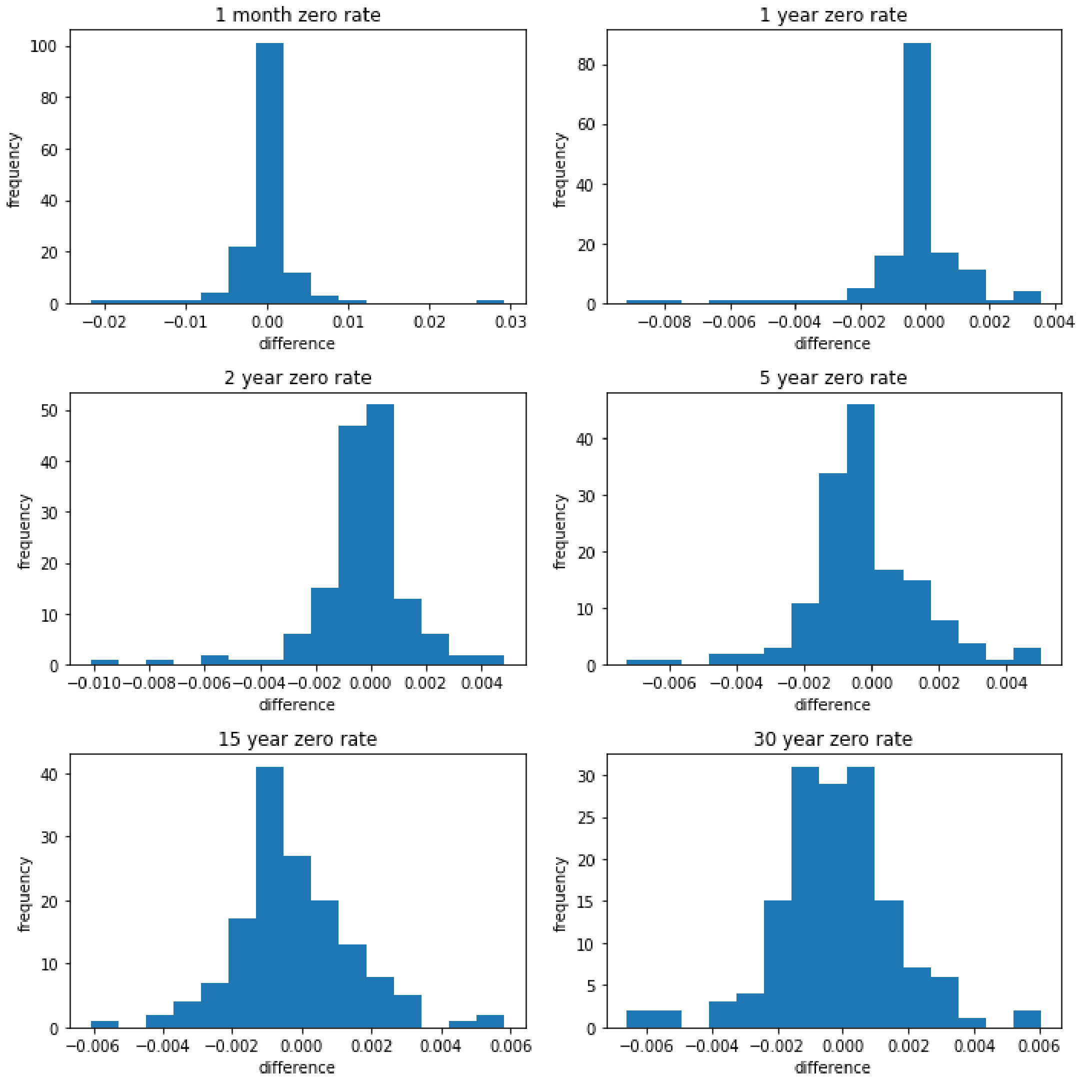


Figure 5: Histograms of the monthly changes for various zero rates

3.4 Models of attention

The models need to facilitate negative rates. In addition, only models that focus on modeling the yield curve and its correlation structure are considered. Given that the rates for different maturities show different types of behaviour, both a normal and a lognormal will be analysed.

The models researched in this study are the Displaced Diffusion model as introduced by Rubinstein (1983) and the dynamic Nelson-Siegel model, suggested by Diebold and Li (2006). The latter is based on the Nelson-Siegel curve introduced in Nelson and Siegel (1987). Both models feature a correlation structure on the yield curve, and facilitate negative rates. Dynamic Nelson-Siegel offers in addition flexibility and relatively easy estimation of parameters and simulation. Displaced Diffusion is used as a Libor market model and thus is harder to calibrate, but models the correlation structure more precisely.

3.4.1 Displaced Diffusion

The Displaced-Diffusion (DD) model as conceived originally by Rubinstein (1983) is used to model a company in both a risk-free and a risky part, as an opposing model for the classic Geometric Brownian Motion. The model can also be used for interest rates. As described in Oosterlee and Grzelak (2019) the DD model can be seen as a shifted log-normal process, which can be considered as first-order approximation to the Constant Elasticity of Variance model (CEV). According to Rebonato (2003), the CEV model is popular because on one hand, it features the empirically observed log-normal nature of the forward rate process, and on the other hand can imply a volatility smile. However, it is computationally demanding. With the DD process being computationally simpler, but still very similar to the CEV process, the DD process is very attractive.

From a modeling perspective, the DD process can feature negative rates because of the displacement. In addition, by modelling the process as log-normal it facilitates that the variance is dependent on the level of the interest rate, which is a desirable property, as described in section 3.2.

A drawback is that not all parameters are easy to estimate, especially the displacement parameter. In addition, since the DD process is used as replacement process for the CEV, it is a Libor market model and needs a lot of high quality data to gain insight in the correlation structure. Because DD models Libor rates, the zero rates need to be converted to Libor rates first. After simulation, the rates need to be converted back to zero rates in order to compare results with other models.

Definition 9 (Displaced Diffusion) *Displaced Diffusion is a model that models the Libor rate. It is characterized by the following SDE*

$$dl_t^i = \sigma_i\big(\beta l_t^i + (1-\beta)l_t^i\big)dW_t^i, \tag{7}$$

with σ_i the volatility and β the displacement parameter. By setting $\theta = \frac{1}{\beta}(1-\beta)l_{t_0}^i$ and $\hat{\sigma}_i = \beta\sigma_i$ This can be seen as a shifted log-normal distribution

$$\frac{dl_t^i + \theta}{l_t^i + \theta} = \hat{\sigma}_i dW_t^i. \tag{8}$$

The model models the Libor rates as a shifted lognormal processes, where the shift is used to account for negative interest rates. Given the lognormal nature of the model, it can explode over longer time periods as well as reach the lower bound and never return. Over short to medium periods this should however not matter. The variance of l_t given l_0 is dependent on t, as following (using Z as a standard normal random variable)

$$
\begin{aligned}
\text{Var}(l_t) &= \text{Var}\Big(l_0\text{exp}(-\frac{1}{2}\sigma^2 t + \sigma\triangle tZ)\Big) \\
&= l_0^2\text{exp}(-\sigma^2 t)\text{Var}\Big(\text{exp}(\sigma\triangle tZ)\Big) \\
&= l_0^2\text{exp}(-\sigma^2 t)\left(\text{exp}(\sigma^2 t) - 1\right)\text{exp}(\sigma^2 t) \\
&= l_0^2\left(\text{exp}(\sigma^2 t) - 1\right) \qquad (9)
\end{aligned}
$$

,
using the variance for a log-normal distribution in the second step. And thus for predictions that are further in the future, rather high forward rates are possible in this model, as the variance increases with time. Given that some of the CCR measures are susceptible to large outliers, this is something to look out for when testing this model.

3.4.2 Dynamic Nelson-Siegel

The dynamic Nelson-Siegel method is slightly different, in that it is not an equilibrium model or no-arbitrage model. Rather, it models the current yield curve to a Nelson-Siegel curve. One then dynamically changes the parameters. The model is interesting in all three categories identified in the previous section. From a modeling perspective, it has strong emphasis on the yield curve, and can incorporate negative interest rates. It is mean-reverting because the parameters are modelled as AR(1) processes. From a computational perspective the model can be attractive, however this depends on the ways of estimating the parameters. Many methods exist, for example using ridge regression as described in Annaert et al. (2012). In addition, more practical methods are described in Ibáñez (2016). Finally, the model is useful in a risk management setting, due to the strong focus on forecasting. Diebold and Li (2006) found that for 1 year predictions and longer, the model shows promising results. Since this is the range in which counterparty risk measures are estimated, it can be interesting to see how the model performs in EPE and PFE estimations.

In the model, the Nelson-Siegel curve is parametrised as

$$
y_{\tau_i}(t) = \beta_{1,t} + \beta_{2,t}\frac{1 - e^{\lambda_t \tau_i}}{\lambda_t \tau_i} + \beta + 3, t\big(\frac{1 - e^{-\lambda_t \tau_i}}{\lambda_t \tau_i} - e^{-\lambda_t \tau_i}\big), \qquad (10)
$$

in which τ_i is the maturity. Diebold and Li (2006) use the approach to fix parameter λ_t to simplify calculations. Then, using a certain process, the parameters $(\beta_t^1, \beta_t^2, \beta_t^3)$ are dynamically varied to obtain a forecast for the future yield curve. This yield curve can then be used to obtain future values for the respective products that depend on the yield curve. In the research by Diebold and Li (2006) the model was found to perform best in out-of-sample forecast when modeling the parameters as AR(1) processes.

The rationale behind the dynamic Nelson-Siegel model, is that it tries to capture the high dimensional yield curve dynamics with lower dimensional state dynamics, see Diebold and Rudebusch (2013). It is tractable via regression when one only considers the three latent variables, and assumes the lambda constant. The dimensionality reduction is done by using the factor loadings 1, $\frac{1-e^{\lambda_t \tau_i}}{\lambda_t \tau_i}$ and $\big(\frac{1-e^{-\lambda_t \tau_i}}{\lambda_t \tau_i} - e^{-\lambda_t \tau_i}\big)$.

The three latent variables can be explained as following. The $\beta_{2,t}$ is a short term factor, given that its loading is a function that starts at one and decays fast to zero. Its effect on the yield curve shape is that it determines the yield curve slope. The loading on $\beta_{3,t}$ on the other hand starts at zero, increases, and then decreases to zero, and thus can be seen as a medium term factor. This represents the curvature of the yield curve. Finally, $\beta_{1,t}$ has a constant factor loading and is the long term factor, given that the other two loadings decay to zero. The parameter $\beta_{1,t}$ can be understood as the yield curve level.

4 Methodology

In this chapter the setup for model testing is explained. First, the procedures for parameter estimation are selected and reflected upon. Then, the general approach is defined. Since the models under scrutiny are modeling different rates - forward rates, Libor rates and yield curves - a unifying "numeraire" is chosen to compare results.

Many of the criteria from chapter 2 are easy to verify - the model allows for negative interest rates or not, the same holds for mean-reversion. The criteria that are not easily verified, and thus the main focus of the research forthcoming, are predictive power, the ability to model real-world scenarios and ease of calibration. These criteria are all connected; assuming a certain parameter as constant (e.g. as done in Diebold and Li (2006)) may lead to easy estimation of parameters. But they in turn may over-simplify certain aspects and result in wrong estimates, perhaps underestimating risk.

The models as used are calibrated using historical data, and thus are modeled under the real-world measure $\mathbb{P}$. The implication is that the forecasts are not risk-neutral.

In addition, Monte Carlo simulation is used in all cases. Binomial-tree methods could be used as well, however this is more difficult in high dimensional situations. Given that dynamic Nelson-Siegel is a three-factor model and Displaced-Diffusion is used as a Libor market model with a factor for each seperate maturity, Monte Carlo is preferred.

In this section it is described how the models are tested. The performance of the models is tested in three ways. The first is performance of the model on historical backtesting with regards to the empirical distributions generated by the models. The aim is to test whether the realised interest rates could be generated by the theoretical models. The second is performance of the models on historical backtesting in a Counterpart Credit Risk setting. The aim is to test how the models perform historically on the expected exposures, potential future exposure and indirectly on the expected positive exposure. Finally, the models key assumptions are validated, including analysis on the normality of residuals and model fit.

Real-World Measure

In this study, the main goal is to find a model that can serve as a benchmark in estimating the PFE and EPE. The main reason is that the current models in some cases are over-conservative when in historical backtesting. Since the historical backtesting takes place under the Real-World meaure, the analysis described in the coming sections takes place under the Real-World measure as well. In most cases, simulation under the Real-World measure comes naturally, as models will be calibrated using historical data. However in some cases, a drift needs to be added to the model to be able to capture the "Market price of risk".

4.1 Backtesting the empirical distribution

Then, following most of the internal works of ING from FI/FM-Quants (2019), backtesting is done. The methodology encompasses multiple tests. First of all, the goodness of the empirical distribution of the simulated zero rates is tested. This is done by employing the goodness of fit tests of Kolmogorov-Smirnov, Anderson-Darling and Cramer-von Mises. Second, the goodness of fit of the Expected Exposures is tested. This uses the same tests as just mentioned, however this time by using the simulated and realised rates in the pricing functions for a certain portfolio (as will be described shortly after). The third tests are the actual performance on the various CCR-measures: the PFE, EPE and EEPE.

The general procedure of backtesting is as follows: The models are calibrated and used to generate possible future paths for a certain time point. Then, at the next time point this is repeated, and so forth. Doing this allows to test how the model performs if it were employed against historical data. A key concept is the backtesting window, with an example shown in figure 6. The first value in the superscript time points refer to the instance of the time point, he second determines the calibration period. The initial value of a calibrated model is equal to the observed market value at the time point at the start of the backtesting window. For example at $t_{0,0}$ a model is calibrated, and is used to simulate paths. In

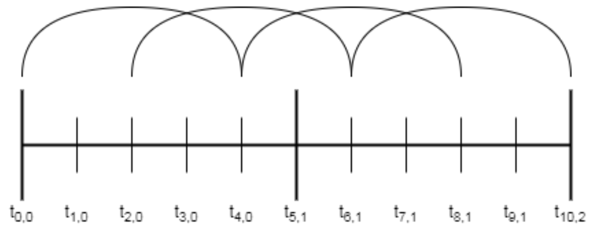


Figure 6: Backtesting example with five-monthly recalibration and four-month overlapping intervals

the case of $t_{0,0}$, simulated paths are checked against the result at time $t_{4,0}$. The next window then uses $t_{2,0}$ as initial value, however uses the same calibrated set of parameters as used in $t_{0,0}$. The simulated results are checked against the results on $t_{6,1}$. At $t_{5,1}$, the models are recalibrated. Note however that the parameters do not change within one backtesting period

Three aspects need to be specified for a backtesting experiment. These are

- Initial recalibration
- Periodical recalibration
- Overlapping intervals

First, models are calibrated initially. Models require a certain amount of data to make meaningful predictions, and thus a predefined number of daapoints can't be used as backtesting scenarios, but only as model input. This period is set at three years for all experiments to follow.

The second is that the models are recalibrated periodically. In the figure this is indicated with the second subscript. Only the starting point is important for determining whether a recalibration takes place - it does not change in one window.

The third is that intervals can overlap. This helps in generating more backtesting scenarios, especially for longer time horizons. If for example a yearly window is chosen, non-overlapping intervals result in only 9 scenarios, given the burn in period. Scenarios that overlap can result in up to 100 scenarios. This does however lead to the problem that the various backtesting windows have a high correlation. The solution to this taken is that the pseudo random numbers used for the respecting windows are also overlapping.

The empirical distribution is tested by using that if the theoretical model is equal to the data generating distribution, the quantiles should match. For every backtesting window, paths are generated. Then the ratio of paths that are smaller than the observed rate should follow a Uniform(0,1) distribution.

4.1.1 Uniformity tests

The test the uniformity of the quantiles, three tests are used: Anderson-Darling (Anderson and Darling (1954)), Cramer-von Mises (Cramér (1928),von Mises (1928)) and Kolmogorov-Smirnov (Kolmogorov (1928),Smirnov (1939),Smirnov (1944)). With these the empirical density function is tested against a predefined probability distribution in this study.

Definition 10 (Goodness-of-Fit test statistics) ***Anderson-Darling*** *uses the test statistic*

$$AD_n = n \int_{-\infty}^{\infty} \frac{[\tilde{F}(s;n) - F(s)]^2}{F(s)(1-F(s))} dF(s)$$

Cramér-von Mises *uses the test statistic*

$$CvM_n = n \int_{-\infty}^{\infty} [\tilde{F}(s;n) - F(s)]^2 dF(s)$$

Kolmogorov-Smirnov *uses the test statistic*

$$KS_n = sup_s |\tilde{F}(s;n) - F(s)|$$

All definitions use $\tilde{F}(s;n)$ *to denoted the empirical distribution and* $F(s)$ *the theoretical CDF*

The goodness-of-fit tests all focus on different areas. Anderson darling uses the scaling $\frac{1}{F(s)(1-F(s))}$ which leads to a stronger focus on the tail. Cramér-von Mises does not feature such scaling and thus tests the whole distribution equally. The Kolmogorov-Smirnov test uses that largest difference between the theoretical and empirical distributions. This can be at all places in the distribution, so there is no specific focus. The test is strict in the sense that one relatively large deviation in an otherwise good fit will be rejected. Together, these tests allow to verify the goodness-of-fit from various angles.

When backtesting the empirical distribution as described in the previous section, the theoretical distribution used is U(0,1) and thus

$$F(s) = s\mathbb{1}_{[0,1]}(s)$$

and the empirical distribution is

$$\tilde{F}(s;n) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{[0,s]}(x_j)$$

in which x_j are the fraction of generated paths that are lower than the realised values, from the previous section. Calculating the test statistics is then straightforward.

4.1.2 Set up of the empirical distribution experiment

Backtesting of the empirical distribution is done over various windows, both overlapping and non-overlapping. The windows in this study are 1-month, 3-months, 1-year and 2-year. This gives a thorough overview of the performance of the models an the short to medium range. Ideally, even longer backtesting windows are taken into account. CCR-measures in practice are calculated over the maturity of the whole portfolio, which can be as long as 30 years. However, given that the data stretches just over 12 years and a burn in of 3 years is used, windows that are larger than a few years are not feasible. This is because at that point all windows overlap to a certain extent, and the number of observations become very low.

The steps in backtesting the empirical distribution are described as following, using $\triangle$ the length of the backtesting scenarios and δ the time between he start of two consecutive backtesting scenarios.

- Take the initial values from starting point t_i of the backtesting scenario and if indicated recalibrate parameters
- Generate MC paths of the interest rate over the backtesting window time horizon until time $t_i + \triangle$, resulting in an empirical distribution
- Calculate the emperical quantile of the realised interest rate at $t_i + \triangle$.
- Repeat above for $t_{i+1} = t_i + \delta$, until no realised values remain for $t_{i+1} = t_i + \delta$.

In addition, various maturity rates are tested. Since the data consists of 11 different maturities, for readability not all rates are shown in the report. The only window which is non-overlapping is the 1 month backtesting window. The other windows are all overlapping. The tested maturities are the 1-month rate, the 2-year rate, the 5-year rate and the 30-year rate. With these, the short, medium and long maturity rates are all represented. In addition, these rates all feature different types of behaviour as analysed in chapter 3. The backtesting set-up is summarized in table 4.

Table 4: Setup of the backtesting experiment for the empirical distribution

Window	Overlapping (y/n)	Tested Maturities
1-month	n	1m, 2y, 5y, 30y
3-months	y	1m, 2y, 5y, 30y
1-year	y	1m, 2y, 5y, 30y
2-year	y	1m, 2y, 5y, 30y
5-year	y	1m, 2y, 5y, 30y

4.2 Backtesting performance of CCR-measures

This section decribes the backtesting procedure for Counterpart Credit Risk measures. The core of the backtesting as described in the previous section remains the same - a backtesting window is used in which sample paths are generated. These paths are then compared against the historically realised values. However, in the case of CCR-measures, the compared values are not interest rates but rather the prices of products depending on the interest rates. A model performing well on the distribution level does not always perform well on the product level. Backtesting of products that are far out-the-money and far in-the-money often depends more on the tails of the distribution. In addition, by using the simulated rates in a pricing function, different behaviour is possible than what is modeled in the original model.

The contract used to test the models is a Forward Rate Agreement as described in chapter 2. Since the price of an FRA depends on the price of specific zero coupon bonds, the Libor rates generated by the DD model need to be transformed to zero rates. This is done recursively using the definition of Libor rate in equation . Note that the pricing function can be written such that it also includes the underlying Libor rate. However it still needs a ZCB price in order to be calculated. In addition, interpolation needs to be done in some cases. To make interpolation unifying between both models, this is done on the zero rates when testing both models.

Interpolation is necessary in situations where FRA product is based on rates with a maturity not included in the generated paths. In these cases, linear interpolation is used. Interpolation is always done on the zero rates as mentioned before. Note that in the dynamic Nelson-Siegel model, the entire yield curve is generated. Thus estimates using the dynamic Nelson-Siegel model never need interpolation.

4.2.1 Mark-to-Market testing

The Mark-to-Market value of the products are backtested by incorporating the respective pricing function. The main goal is to test the performance of the models on the whole distribution of the products. By doing this, insight is gained in the performance of the model on estimating the EPE. Estimating the EPE from the market is not possible, since it would in principle require multiple values of the same portfolio or product on the same date. Both FI/FM-Quants (2019) and Ruiz (2014) mention this, and the solution suggested in both cases testing the historical distribution of the product following from the model against historical realised values of the product. When this is modeled correctly, calculation of the EE and EPE using the theoretical model is reliable.

To test the performance of the model on the MtM values, a backtesting experiment is done. The backtesting is done in similar fashion as the empirical distribution, such that the MtM distribution is tested in it's entirety. The backtesting procedure can be summarised as following, noting δ the time between backtesting scenarios and $\triangle$ the length of the backtesting scenarios.

- Take the initial values from starting point t_i of the backtesting scenario and if indicated recalibrate parameters
- Generate MC paths of the interest rate over the backtesting window time horizon until time $t_i + \triangle$
- Use a pricing function on both the realised rate and the simulated rates at time $t_i + \triangle$
- Calculate the ratio of paths that are lower than the observed prices, and safe this value.
- Repeat above for $t_{i+1} = t_i + \delta$, until no realised values remain for $t_{i+1} = t_i + \delta$.

Since the empirical distribution of the risk factors are already extensively calculated, only a few simulation experiments are done for the MtM model. The products used in this study are FRAs, and their price is a deterministic function of the zero rates at the end of the backtesting scenario. In principle, the MtM backtesting should yield similar results. To experiment setup is shown in table 5.

Table 5: Setup of the MtM backtesting experiment

Contract	Window
FRA6x9 payer	1 month
FRA6x9 receiver	1 month
FRA24x36 payer	24 months
FRA24x36 receiver	24 months
FRA120x180 payer	12 months
FRA120x180 receiver	12 months

The FRA contracts used in the MtM backtesting all are contracts that are values 0 at inception. This happens when the value for $K = \frac{P(t_0,T_{i-1})-P(t_0,T_i)}{(T_{i-1}-T_i)P(t_0,T_i)}$. It is straightforward to show this by substituting this K in the price of an FRA, which is given in equation 2.1.

4.2.2 PFE exceedence testing

The PFE is tested by counting the number of times a realized value exceeds the calculated value. This should not happen too often nor too little. In case the realized values cross the PFE significantly more than 5% of the times, the model underestimates the risk. On the other hand if it happens significantly less than 5% of the time, the model is over-conservative and can lead to unnecessary reserve capital.

Whether the number of exceptions differs significantly can be tested using the exception counting test also used for VaR backtesting. The methods described in Kupiec (1995) are used in this study. The underlying principle is that the number of exceedences in case the realised values come from the theoretical distribution, the number of observed exceedences can be seen as a binomial distribution $B(n, p)$. The parameter n is the number of observations and p the probability of exceeding the PFE and thus equal to the quantile specified for the PFE. This is then compared with the distribution $B(n, \frac{x}{n})$ in which x is the number of exceedences. Kupiec (1995) uses the likelihood ratio test statistic to determine if $p = \frac{x}{n}$, defined as

$$LR = -2\ln[(1-p)^{n-x}p^x] + 2\ln[(1-\frac{x}{n})^{n-x}(\frac{x}{n})^x]$$

Under the null-hypothesis that $p = \frac{x}{n}$ this statistic is chi-square distributed with 1 degree of freedom. And thus the value of the LR statistic found can be compared with the value for the corresponding significance level in a chi-square table. If the LR statistic is larger, the null hypothesis that $p = \frac{x}{n}$ is rejected and thus the model is incorrect.

Ruiz (2014) suggests a similar green-orange-red light approach as is used in backtesting the VaR for the PFE. The principle is based on much the same as above, but simplifies it by setting thresholds for which the number of exceedences are suspiciously high (orange) or very unlikely high (red). In the paper, orange is specified when the likelihood ratio defined earlier exceeds the 95% confidence level of the cumulative distribution function of the binomial distribution, and red the 99.9 % confidence level. When using 250 observation, 4 exceedences is the limit for a green rating, meaning the model is conservative in the respective case. Between 4 and 9 is orange, which shouldn't happen to often, and 9 or higher is red, which is problematic. In the backtesting in this study, the number of observations is generally smaller. To account for this, the Likelihood ratio as defined before is simply compared to the chi-squared distribution with 1 degree of freedom critical values at both the 95% level and the 99.9% level. The red-orange-green rating are then given accordingly. When the model is (over)conservative, a green light is given. This is of importance for the $PFE_{95\%}$ in cases where zero exceedences take place - in principle this is unlikely to happen when the observed values indeed are generated by the theoretical model. However for the purpose of risk management this is still a green light as it is conservative.

4.2.3 Set up of the PFE backtesting experiment

Backtesting of the CCR measures varies more parameters than the empirical distribution experiment, which only varies the backtesting horizon. In the case of backtesting the PFE, the type of contract, the maturity and the window over which the contract is backtested are taken into account. In addition, the fixed leg is varied in order to gain insight in how the model performs when the contract that is researched is far in-, at-, or out-the-money.

The type of FRA contract is either payer or receiver, and allows to see if differences exist in these type of contracts. Given that the interest rates that are used as input in this study have a noticeable downward trend, it should be the case that when a party benefits from decreasing rates, the PFE behaves accordingly. That is, when the contract value is going up, the PFE should go up as well. In the case of a payer contract, the issuer pays a fixed leg and receives a floating leg and thus the contracts value decreases as the interest rate is is based on decreases. PFE profiles in this case should generally be low.

The maturity of the contract and window of the backtesting determine what rates are used and the window over which the rates are analysed. An example is the FRA6x9 backtested over 1 month intervals. The price of an FRA6x9 contract is determined at inception using the 6 month zero rate and the 9 month zero rate. Then after one month, the value of the contract is determined using the 5 month rate and the 8 month rate.

The choice for which maturity contracts and backtesting windows depend on three thing. What rates are needed to be tested, what window is tested and the need for interpolation. In principle a cross-section of all rates need to be examined, each over a short and medium backtesting window [4], whilst minimizing the need and effects of interpolation. But also some care has to be taken to not let the number of backtesting experiments blow up. For this reason two contracts are analysed, the FRA6x9 and FRA24x36 contract

The backtesting set-up for the PFE in its entirety is summarized in table 6.

4.3 Model Validation

In the validation of the models, various assumptions are verified. As the underlying assumptions and parameters are different between the models, different approaches are taking. However, some overlap can be found between the assumptions. The assumptions verified are generally driven by earlier found remarks. In both models the residuals of the models are tested using normality tests. In addition, the residuals are tested for heteroskedasticity, as this was found for some maturities in the data at earlier times. Testing the residuals of a model over the whole dataset allows to obtain insight in the fit of

[4] As with the backtesting of the empirical distribution, longer backtesting windows longer than 5 years are not feasible given the time horizon of the dataset

Table 6: Experiment setup for testing PFE

Fixed rate	FRA type	Maturity (start x end in months)	time since inception
K=0.01	payer	6x9	1 month
	payer	6x9	3 months
	payer	24x36	24 months
	receiver	6x9	1 month
	receiver	6x9	3 months
	receiver	24x36	24 months
K variable	payer	6x9	1 month
	payer	6x9	3 months
	payer	24x36	24 months
	receiver	6x9	1 month
	receiver	6x9	3 months
	receiver	24x36	24 months
K=-0.01	payer	6x9	1 month
	payer	6x9	3 months
	payer	24x36	24 months
	receiver	6x9	1 month
	receiver	6x9	3 months
	receiver	24x36	24 months
K=0.05	payer	6x9	1 month
	payer	6x9	3 months
	payer	24x36	24 months
	receiver	6x9	1 month
	receiver	6x9	3 months
	receiver	24x36	24 months

the model. It also can show if the assumption of a driving normal distribution, as both models use, is warranted. Frequent extreme observations for example will result in rejection of normality of the residuals. In both models parameters are tested for sensitivity. In the case of the dynamic Nelson-Siegel model, this is done via the λ parameter which determines where the curvature takes it's maximum. In the case of Displaced Diffusion, this is the shift parameter which determines the lower bound of the model, but also the size of the volatility parameter. In addition both parameters are assumed constant, which is an assumption that needs to be verified whether it is warranted based on the data. The models are then tested on other types of behaviour. In the case of DNS, this is the inherent mean reverting behaviour of the underlying AR(1) processes for the parameters. In the case of DD, this is the long term behaviour given the lognormal nature of the rates. The correlations of both models are analysed as well. For DNS this is the correlation between the parameter processes. For DD this is the correlation structure inherent to the model. Finally, the fit to the original yield curve is analysed. This is only relevant for DNS, as the DD model uses the current yield curve as starting point. In DNS the yield curve is approached using a Nelson-Siegel curve. Some types of yield curves may not be possible to be modeled as a Nelson-Siegel curve leading to possible mistakes in the estimation.

5 Dynamic Nelson-Siegel

In this chapter, all steps described in chapter 4 on methodology are performed on the dynamic Nelson-Siegel model. Before doing that, the calibration and simulation procedures are described.

5.1 Calibration

The dynamic Nelson-Siegel model is a three factor model, and for all three risk factors three parameters need to be estimated. In addition there is a "price of risk" parameter λ_t. The essence of the model is relatively simple. In the first step, one needs to fit the Nelson-Siegel curves to all historical yield curves. Then in the second step, one estimates the processes underlying the parameter processes obtained in the first step.

As mentioned before there are many ways of doing the first step. In this study the approach of Diebold and Li (2006) is taken, in which the parameter λ_t is taken as a constant. It is chosen to be the value for which the factor loading for curvature takes it's maximum on the medium term, i.e.,

$$\text{argmax}_{\lambda_t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}-e^{-\lambda\tau}\right)$$

The industry practice is to take either 2 or 3 years for τ. In the aforementioned research, 2.5 years is taken as compromise. Using this constant value, obtaining the historical parameter processes can be done by using Ordinary Least Squares.

The parameters $\beta_1, \beta_2, \beta_3$ are assumed to follow independent AR(1) processes. Note that the specification of the AR(1) models used in this study is

$$X_{t+1} = \mu(1-\phi) + \phi X_t + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N(0, \sigma^2)$$

The reason is that this way, the process is easily written in the form

$$X_{t+1} - \mu = \phi(X_t - \mu) + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N(0, \sigma^2) \tag{11}$$

Which helps in both simulating, by instead simulating $Y_t = X_t - \mu$, as well as in estimating parameters. The parameters of interest are μ, ϕ, σ for each of the betas. The parameters μ and ϕ are simply estimated by performing linear regression on the function $x_{i+1} = x_i + 1$. Then σ is obtained by the square root of the sample variance of the residuals of the model.

5.2 Mathematical properties of the yield curve

In this section, mathematical properties of the yield curve are described. By noting the factorloadings of curvature and slope as $L_{1,\tau_i}, L_{2,\tau_i}$, the notation of the DNS yield curve can be simplified as

$$y_{\tau_i}(t+n) = \beta_{1,t+n} + \beta_{2,t+n}L_{1,\tau_i} + \beta_{3,t+n}L_{2,\tau_i}. \tag{12}$$

in which τ_i is the maturity. The mean of the zero rate maturity processes, given the current Nelson-Siegel curve (i.e. with the β's known) is

$$\begin{aligned}
\mathbb{E}[y(t+n)|\mathcal{F}_t] &= \mathbb{E}[\beta_1(t+n) + \beta_2(t+n)L_{1,\tau_i} + \beta_3(t+n)L_{2,\tau_i}|\mathcal{F}_t] \\
&= \mathbb{E}[\beta_1(t+n)|\mathcal{F}_t] + \mathbb{E}[\beta_2(t+n)|\mathcal{F}_t]L_{1,\tau_i} + L_{2,\tau_i}\,\mathbb{E}[\beta_3(t+n)|\mathcal{F}_t] \\
&= \mu_1(1-\phi_1^n) + \beta_1(t)\phi_1^n + L_{1,\tau_i}(\mu_2(1-\phi_2^n) + \beta_2(t)\phi_2^n) + L_{2,\tau_i}(\mu_3(1-\phi_3^n) + \beta_3(t)\phi_3^n) \\
&:= \hat{\mu}_n.
\end{aligned}$$

It is now clear that, given the auto-regressive coefficients are in $(-1, 1)$, for n large the yield curve process is mean reverting to $\mu_1 + L_1\mu_2 + L_2\mu_3$. Following similar steps, using that the individual AR(1) processes are independent, the variance is

$$
\begin{aligned}
\text{Var}(y(t+n)|\mathcal{F}_t) &= \text{Var}(\beta_1(t+n)|\mathcal{F}_t) + (L_{1,\tau_i})^2\text{Var}(\beta_2(t+n)|\mathcal{F}_t) + (L_{2,\tau_i})^2\text{Var}(\beta_3(t+n)|\mathcal{F}_t) \\
&= \sigma_1^2\frac{1-\phi_1^{2n}}{1-\phi_1^2} + (L_{1,\tau_i})^2\sigma_2^2\frac{1-\phi_2^{2n}}{1-\phi_2^2} + (L_{2,\tau_i})^2\sigma_3^2\frac{1-\phi_3^{2n}}{1-\phi_3^2} := \hat{\sigma}_n^2,
\end{aligned}
$$

in which σ_i^2 is the variance of the noise term ϵ_i. Since the AR(1) processes are independent Gaussian processes, their linear combination is also a Gaussian process.Using this, it is straightforward to construct a confidence interval, using $\alpha = 0.05$ for illustrational purposes, of the form

$$
[\hat{\mu_n} - 1.96\hat{\sigma_n}, \hat{\mu_n} + 1.96\hat{\sigma_n}].
$$

Furthermore, the correlations in the yield curve can be calculated exactly. Any two zero rates are a linear combination of three independent normal random variables, and thus the correlation is as following

$$
\begin{aligned}
&Corr(Z_1 + L_{1,1}Z_2 + L_{2,1}Z_3; Z_1 + L_{1,2}Z_2 + L_{2,2}Z_3) \\
&\qquad = \frac{Cov(Z_1 + L_{1,1}Z_2 + L_{2,1}Z_3, Z_1 + L_{1,2}Z_2 + L_{2,2}Z_3)}{\sqrt{Var(Z_1 + L_{1,1}Z_2 + L_{2,1}Z_3)Var(Z_1 + L_{1,2}Z_2 + L_{2,2}Z_3)}} \\
&\qquad = \frac{Cov(Z_1, Z_1) + L_{1,1}L_{1,2}Cov(Z_2, Z_2) + L_{2,1}L_{2,2}Cov(Z_3, Z_3)}{\sqrt{(\sigma_1^2 + L_{1,1}^2\sigma_2^2 + L_{2,1}\sigma_3^2)(\sigma_1^2 + L_{1,2}^2\sigma_2^2 + L_{2,2}^2\sigma_3^2)}} \\
&\qquad = \frac{\sigma_1 + L_{1,1}L_{1,2}\sigma_2 + L_{2,1}L_{2,2}\sigma_3}{\sqrt{(\sigma_1^2 + L_{1,1}^2\sigma_2^2 + L_{2,1}\sigma_3^2)(\sigma_1^2 + L_{1,2}^2\sigma_2^2 + L_{2,2}^2\sigma_3^2)}},
\end{aligned}
$$

in which it is used multiple times that for $X, Y \sim N(\mu, \sigma^2)$ with X, Y independent, $X + Y \sim N(\mu + \mu, \sigma^2 + \sigma^2)$. This expression for the correlation structure implied by DNS can later be used to test the theoretical correlation structure of the yield curve against the historically observed correlations.

5.3 Simulation

Though the confidence interval derived before is very convenient for PFE estimation, since it is defined as a quantile, it is still useful to perform simulation. On one hand, this is because definitions in Basel are based on simulations, on the other hand because the effective EPE estimation is path dependent. Given the underlying AR(1) process it is convenient to substitute and model a process without a constant (see equation (11) and thereafter).

Then, the process for n steps ahead can be simulated by using

$$
X_{t+n} = \phi^n X_t + \sum_{i=0}^{n-1} \phi^i \epsilon_{t+n-i}.
$$

For one path, this can be expressed using the following matrix form.

$$
\begin{bmatrix} X_{t+1} \\ X_{t+2} \\ \vdots \\ X_{t+n-1} \\ X_{t+n} \end{bmatrix} = \begin{bmatrix} \phi^1 \\ \phi^2 \\ \vdots \\ \phi^{n-1} \\ \phi^n \end{bmatrix} X_t + \begin{bmatrix} \phi^0 & 0 & \cdots & \cdots & 0 \\ \phi^1 & \phi^0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \phi^{n-2} & & & \phi^0 & 0 \\ \phi^{n-1} & \phi^{n-2} & \cdots & & \phi^0 \end{bmatrix} \begin{bmatrix} \epsilon_{t+1} \\ \epsilon_{t+2} \\ \vdots \\ \epsilon_{t+n-1} \\ \epsilon_{t+n} \end{bmatrix}.
$$

This is easily and efficiently simulated for n steps and multiple paths p. The paths are expressed as the columns in the following matrix form

$$\begin{bmatrix} X_{t+1,1} & \dots & X_{t+1,p} \\ \vdots & \ddots & \vdots \\ X_{t+n,1} & \dots & X_{t+n,p} \end{bmatrix} = \begin{bmatrix} \phi^1 & \dots & \phi^1 \\ \vdots & \ddots & \vdots \\ \phi^n & \dots & \phi^n \end{bmatrix} X_t + \begin{bmatrix} \phi^0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ \phi^{n-1} & \dots & \phi^0 \end{bmatrix} \begin{bmatrix} \epsilon_{t+1,1} & \dots & \epsilon_{t+1,p} \\ \vdots & \ddots & \vdots \\ \epsilon_{t+n,1} & \dots & \epsilon_{t+n,p} \end{bmatrix}.$$

Using this, the entire simulation of the process of one dynamic Nelson-Siegel parameter for multiple paths is simply a matrix-multiplication and addition. Within the current format, no extra variance-reducing techniques are employed.

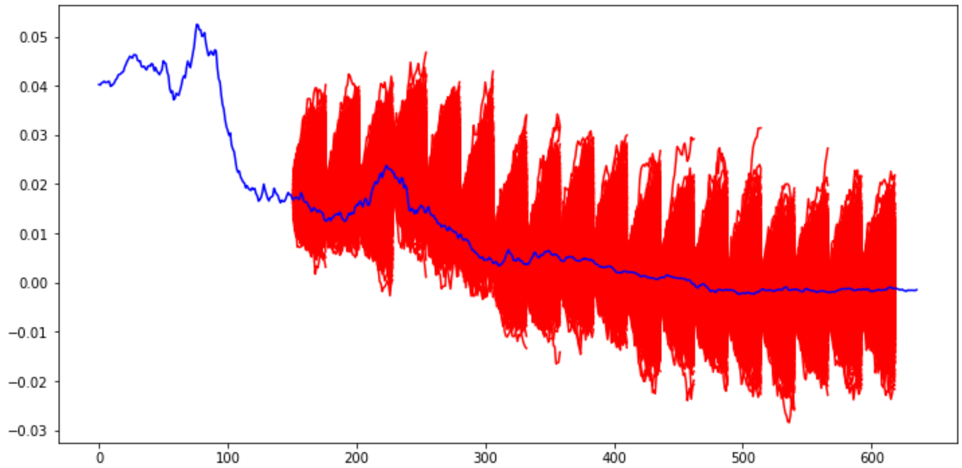
5.4 Performance Backtesting

In this section, the empirical distribution and PFE are tested against historical data. Some parameters are equal throughout this section, to be able to test the relevant aspects of the model. Recalibration is done on a half-year basis. The FRA used in PFE calculations varies in run time and duration, and in addition is tested for various values of the fixed leg, as explained in the setup of the backtesting experiment. By doing this it is possible to get an indication of how the model performs when backtesting the PFE of an out-of-the-money, at-the-money and in-the-money Forward Rate Agreement.

5.4.1 Empirical distribution

First the empirical distribution is tested. The dynamic Nelson-Siegel model is used to generate paths and these are subsequently checked against the realized values, as described in chapter 4. Figure 7 shows an overal example of the backtesting, with non-overlapping intervals. The 2 year zero rate is shown, and similar results hold for other maturities. From the figure it is clear that the variance remains stable over time, even when interest rates are lower. In addition it is visible that the realized interest rates tend to the middle or lower part of the generated paths. This is a strong indication that the simulated distribution may not fit the data particularly well.

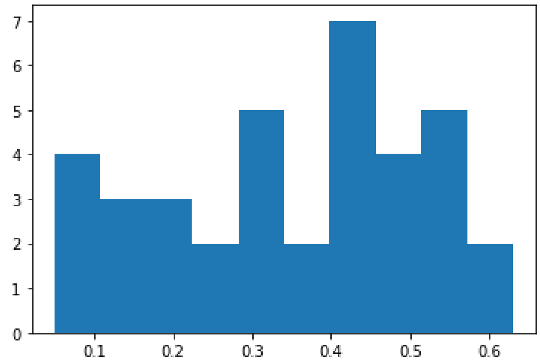
Figure 7: Example backtesting paths against the 2-year rate



Various backtesting horizons with overlapping and non-overlapping intervals have been chosen. The goodness of fit of one month intervals all return a p-value of zero. The hypothesis that the real world rate follows a dynamic Nelson-Siegel model with recalibration as described earlier is strongly rejected. The same result also applies to longer time horizons, using 3-month non overlapping intervals, 1 year overlapping intervals, 2 year overlapping intervals and 5-year overlapping intervals. This is in contrast with Diebold and Li (2006), who perceived particular success on the medium term, using 1 year prediction intervals.

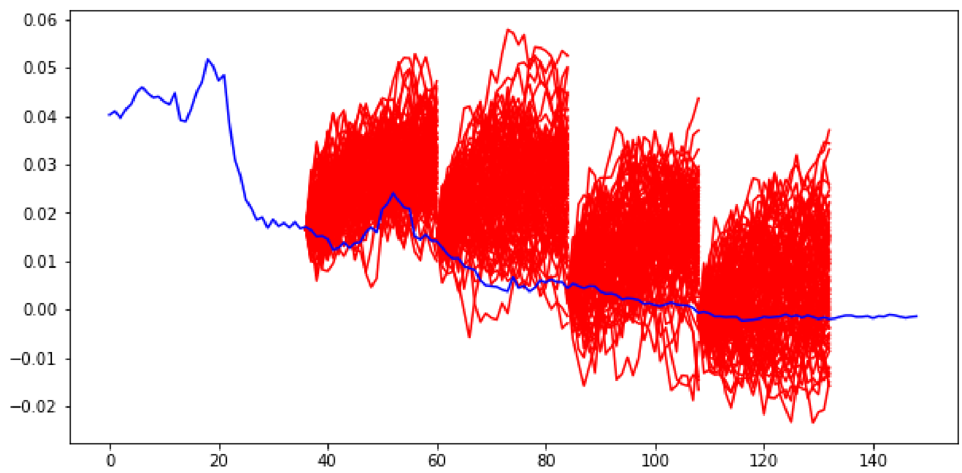
Before continuing to the PFE testing, some evidence is shown on why the predictions do not perform well on the historical data. Figure 8 shows the quantiles of the realized values against the generated empirical distribution. The realized rates never belong to the upper quantiles, as the x-axis stops at 0.6. The dynamic Nelson-Siegel model assigns thus a too high probability of the interest going up by a large amount.

Figure 8: Histogram of the quantiles of realized values against the empirical distribution



As an example for this behaviour, non-overlapping paths are generated over a 2-year horizon. The result is shown in figure 9. The issue with the model is that is tends to some long term mean because of the mean reverting behaviour of the model, which is problematic.

Figure 9: Backtesting using 2 year non-overlapping intervals



5.4.2 MtM backtesting

In this section the results of the MtM backtesting experiment are presented. Table 7 contains the results of the goodness-of-fit tests for the dynamic Nelson-Siegel model. Most of the tests fail, which is as expected based on the results of the backtesting experiment for the empirical distribution.

Table 7: Results of of the MtM backtesting experiment

Contract	Window	KS	CvM	AD
FRA6x9 payer	1 month	0.000	0.000	0.000
FRA6x9 receiver	1 month	0.000	0.000	0.000
FRA24x36 payer	24 months	0.000	0.000	0.000
FRA24x36 receiver	24 months	0.000	0.000	0.000
FRA120x180 payer	12 months	0.000	0.000	0.000
FRA120x180 receiver	12 months	0.000	0.000	0.000

5.4.3 Backtesting the PFE

In this section the backtesting results are presented on estimating the PFE. Based on the previous results it is expected that for payer FRAs, the PFE will overestimate risk as the model overestimated the probablity of rates going up. Conversely for receiver FRAs, it is expected that the PFE underestimates risk. Because of readability the full PFE profiles are shown in Appendix A, section 8.1. Only some PFE profiles are used as an example. The full results are shown in table 8.

The experiments using a fixed rate of $K = 0.01$ is used as an example, as the other results are similar. The full results are show in appendix A in figures 27 to 32. For most contracts, the PFE as estimated by the DNS model is overconservative, as can be seen from both the exceedences in table 8.1 and the mentioned figures. The exception is the case of a RA24x36 receiver contract, in which far too many exceedences are observed. All these are observed at the start of the backtesting. The cause is that the interest rates of the relevant maturity show a very sharp downward movement at the places where the model underestimates the PFE. Conversely, the PFE estimates for the FRA24x36 payer contract are far more conservative, being an indication that the DNS models probability mass lies above the realised distribution. This is a direct consequence of what is shown in figure 9 - the mean reversion property of the model calibrated on data up to the start of the backtesting window is too strong on a 2 year time horizon. The results for the other values of K are similar.

5.5 Model Validation

In this section the key assumptions of the dynamic Nelson-Siegel model as identified throughout the report. It starts with analysing the yield curve fit, in which various aspects of the fit of a Nelson-Siegel curve are analysed. Then the underlying assumptions of the beta processes are analysed, first the independence of the modeled processes and then the betas as AR(1) processes. Finally, the assumption of λ being a constant, maximizing the curvature on the 2.5 year time horizon, is analysed.

5.5.1 Yield curve fit

Estimation of the values is done by regression using ordinary least squares, in which β_{t-1} is used as independent variable, to predict dependent variable β_t. The values for μ, ϕ are easily obtained, and the variance of ϵ is subsequently obtained from the residuals. Some tests and heuristics about the fit, some also done in Diebold and Li (2006), are

- the Nelson-Siegel curves need to fit the yield curves well
- the correlations between the estimated β's need to be small, to warrant the assumption that the parameters are independent processes
- the parameters need to resemble level, curvature and slope of the yield curve
- the range of interest rates needs to be realistic

Table 8: Exceedences of the PFE for DNS

K	contract	time	#Obs.	95%	99 %	$LR_{95\%}$	$LR_{99\%}$
K=0.01	6x9 payer	1 month in	112	0	0	11.489	2.251
	6x9 payer	3 months in	110	0	0	11.285	2.211
	24x36 payer	24 months in	89	0	0	9.130	1.789
	6x9 receiver	1 month in	112	1	0	5.951	2.251
	6x9 receiver	3 months in	110	0	0	11.285	2.211
	24x36 receiver	24 month in	89	20	0	24.162	1.789
K variable	6x9 payer	1 month in	112	0	0	11.489	2.251
	6x9 payer	3 months in	110	0	0	11.285	2.211
	24x36 payer	24 months in	89	0	0	9.130	1.789
	6x9 receiver	1 month in	112	0	0	11.489	2.251
	6x9 receiver	3 months in	110	0	0	11.285	2.211
	24x36 receiver	24 month in	89	20	0	24.162	1.789
K= -0.01	6x9 payer	1 month in	112	0	0	11.489	2.251
	6x9 payer	3 months in	110	0	0	11.285	2.211
	24x36 payer	24 months in	89	0	0	9.130	1.789
	6x9 receiver	1 month in	112	0	0	11.489	2.251
	6x9 receiver	3 months in	110	0	0	11.285	2.211
	24x36 receiver	24 month in	89	0	0	9.130	1.789
K=0.05	6x9 payer	1 month in	112	0	0	11.489	2.251
	6x9 payer	3 months in	110	2	0	3.070	2.211
	24x36 payer	24 months in	89	0	0	9.130	1.789
	6x9 receiver	1 month in	112	0	0	11.489	2.251
	6x9 receiver	3 months in	110	0	0	11.285	2.211
	24x36 receiver	24 month in	89	20	13	24.162	41.761

- The correlations in the theoretical yield curve match the observed correlations

The last item is not tested in previously mentioned study. This is done by comparing the long term correlations dictated by the DNS model (as described in section 5.2) with the observed correlations in the yield curve.

The first item is can be verified by looking at the fit of the curves and the residuals. In figure 10 the yield curve at the various times is shown. The shape seems to be good for most of the dates, but it is by no means a perfect fit. Especially outliers are skewing the fit. For pricing this is an undesired aspect of the model, since it is not able to replicate the prices of the initial contracts. However, for out-of-sample forecasts, this is not necessarily an issue.

The second item is easily verified, by simply checking the correlations between the processes. These are all below 0.4, which makes it reasonable to assume that the processes are independent. Diebold and Li (2006) stop here, however next section of this report will give a more critical look to this assumption.

The third item concerns the modeling of level, slope and curvature of the processes. Diebold and Li (2006) define the level as the limit in maturity of the yield curve, however since the data has 30 years as longest maturity, 30 years is taken instead. The slope of the yield curve is defined as the 30-year yield minus the 1-month yield and the curvature as two times the 2-year maturity zero rate, minus the sum of the 10-year and 3-month zero rates. Figure 11 shows plots of level, slope and curvature against the respective β-parameters. The fit is good, even though some periods are shifted to some extent. The correlation between β_1 and the level is 0.99, between β_2 and slope -0.96 and between β_3 and curvature is 0.99, and thus resemble the stylized features well.

The estimated parameters are shown in table 9. Then using the formulation of section 4 the long term mean and variance of the process can be determined.

Figure 10: The yield curve versus the fitted Nelson-Siegel curve

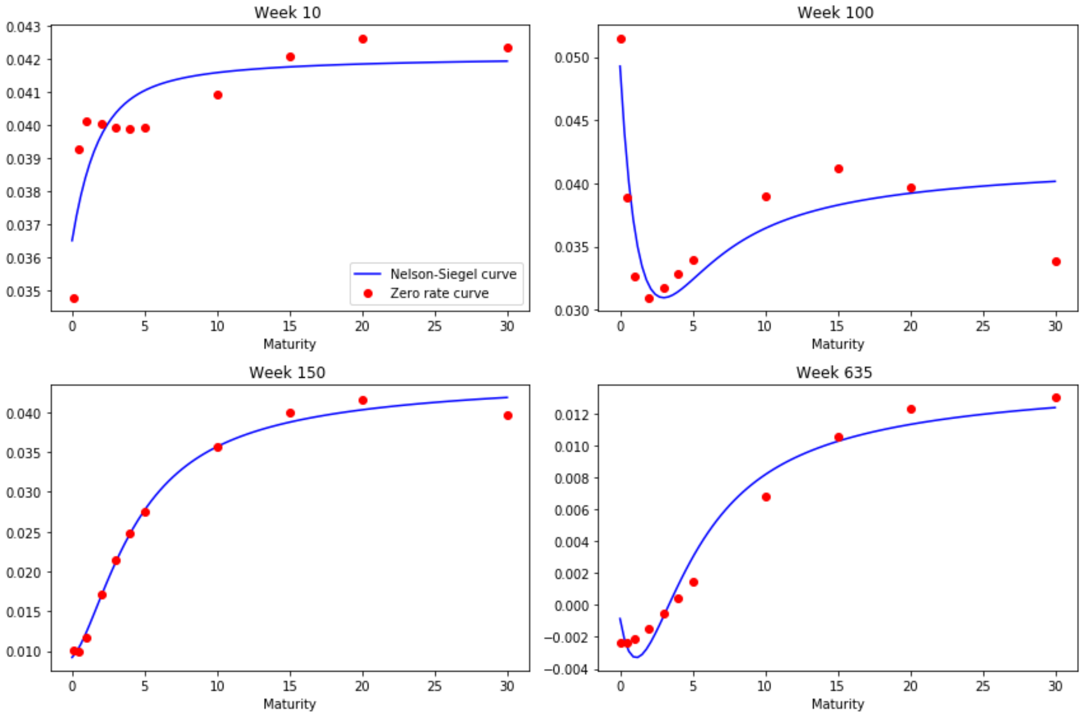


Table 9: Dynamic Nelson-Siegel parameters

AR(1) parameters			
Parameter Process	AR-coefficient γ_i	Constant μ_i	Noise volatility σ_i
β_1	0.9879	0.0288	0.00228
β_2	0.9295	-0.0161	0.00357
β_3	0.8140	-0.0290	0.00905

Using these parameters, it is straightforward to derive exact confidence intervals, as described in section 4. In figure 12, the 95% confidence interval is shown for a very long time horizon (1000 weeks, thus about 20 years). From the figure it is clear that, because of the relatively high auto-regressive parameters, quite some changes are expected even in the very short term. In addition the model seems to attach a low probability that the rates go beyond 5%, even over a 20 year time horizon. This reflects the calibration dataset well. Figure 13 then shows 50 paths forecasted using the methods described in section 4. The modeled interest rates are within realistic bounds of what is observed in the dataset.

The fifth item can be verified by using the exact expression for the yield curve correlation in section 5.2. Figure 14 shows the yield correlations for a fitted DNS model after 10 years [5]. Although the scales are slightly different, the correlation structure resembles the observed correlation structure of the yield curve very well. The good fit suggests the use of level, slope and curvature as principal components for the yield curve are well chosen.

[5]This is because the volatilities in the correlation are those of an AR(1) process, and thus a long time horizon is used to be able to use the long term variance

Figure 11: The parameters of DNS against their stylized counterparts

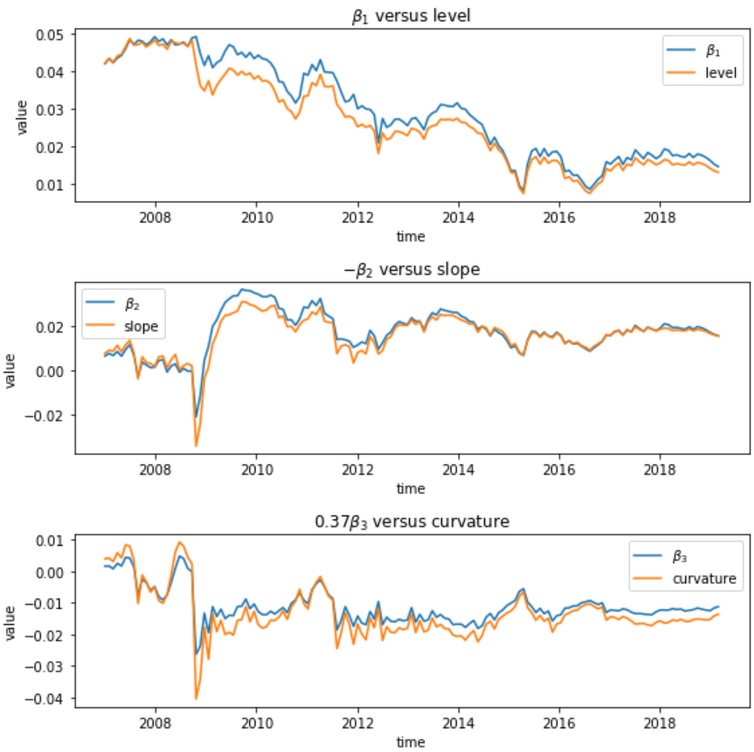


Figure 12: 95% CI for the 2-year zero rate under DNS

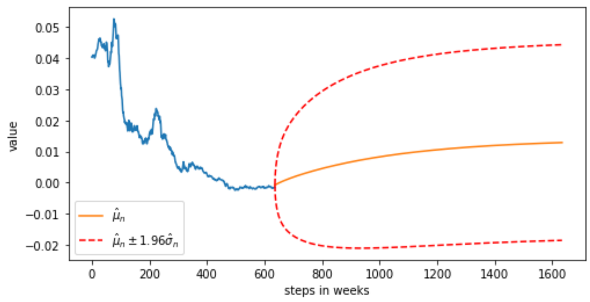


Figure 13: 50 paths for the 2-year zero rate under DNS

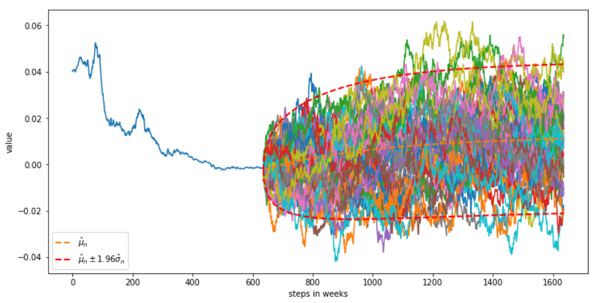
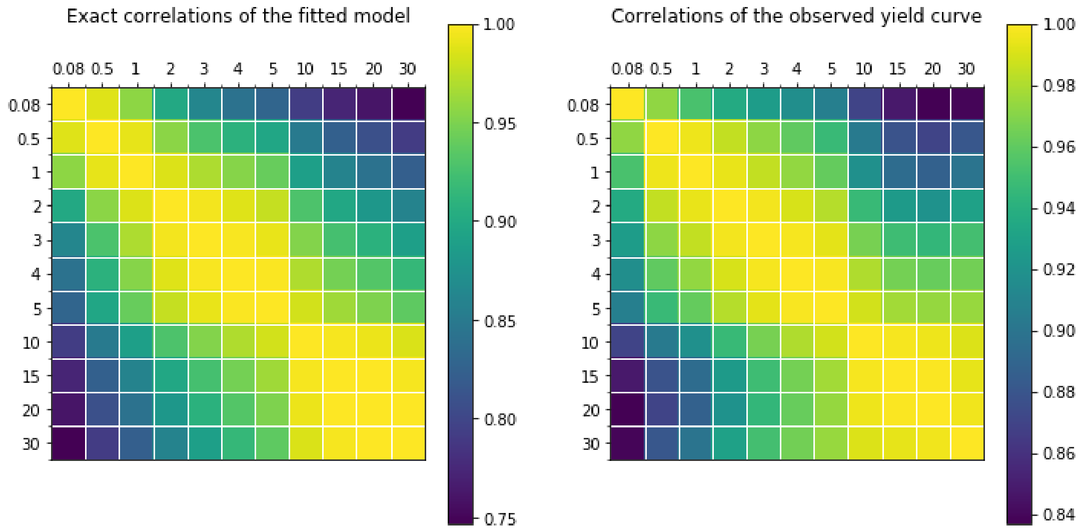


Figure 14: Theoretical versus Observed yield correlation



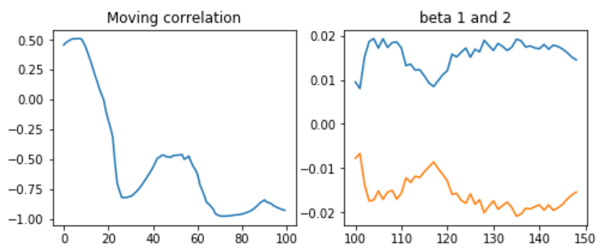


Figure 15: Moving correlation and sample paths of β_1, β_2

5.5.2 Modeling β's as independent processes

In chapter 5 the fit of the β's was tested using much of the same results as Diebold and Li (2006). One of the main assumptions in the model is that the β's are modelled independently, and correlations over the whole process are low. There is however reason to believe the parameters vary in how much they are correlated. Figure 15 shows the correlation between β_1 and β_2 over time, using a window of 50 observations. It is evident that the assumption of no correlations does not hold. Especially the later periods, as shown on the right side of figure 15 show that the respective β's move in opposite direction.

This observation can also explain the difference in behaviour between shorter maturity rates and longer maturity rates. Longer maturity rates predominantly depend on β_1, as the factor loadings of the other β's are close to zero. And thus longer maturity rates are driven by normal increments. However for the shortest maturities, where the factor loading of β_3 is still very low, the process is modeled as a linear combination of β_1 and β_2. Given their strong negative correlation in the more recent part of the dataset, this results in very low variation, as is observed in reality.

Similar results are found for the correlation structure between β_1 and β_3 as well as between β_2 and β_3. It is thus recommended to model the correlation structure in order to better match the interest rates across all maturities. There are many possibilities for this to research in future studies. Suggested ways of modeling the correlation structure are using time dependent or stochastic models. In addition given the abrupt changes in correlation, potential models can use regime switching.

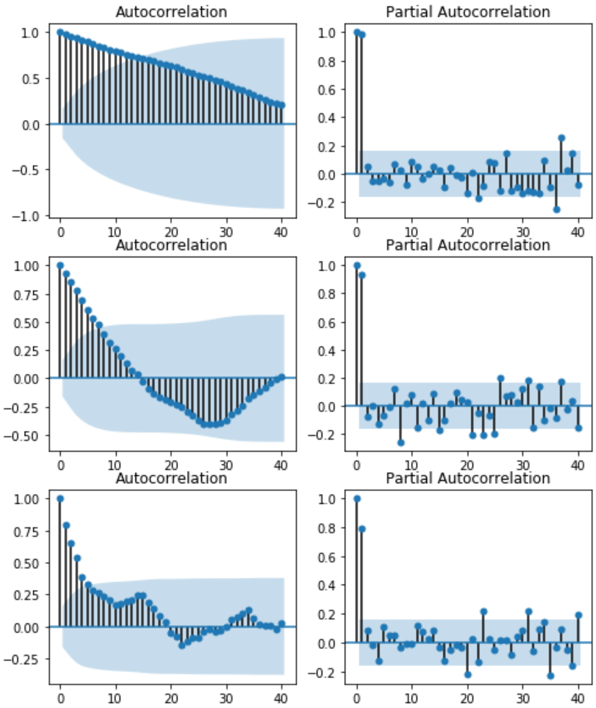


Figure 16: ACF and PACF plots of the β's

5.5.3 Modeling β's as AR(1) processes

The work of Diebold and Li (2006) concludes that the model performs best on long horizons when the β-parameters are modeled as AR(1) processes. However using the same approach, results are not as optimistic in this study. In this section, various tests are done to analyse whether the choice of AR(1) processes is justified. In addition, different models are proposed based on this analysis and are tested on a small scale.

Tests on the β series

As a heuristic to determine the order and type of an ARMA(p,q) process, Shumway and Stoffer (2005) suggest to examine the autocorrelation function (ACF) and partial autocorrelation function (PACF). For AR(p) processes, the ACF tails off and the PACF cuts off after lag p. As can be seen in figure 16, this is the case for all β parameters and thus suggest that an AR(1) process might be viable.

However AR(1) processes are stationary. And thus an Augmented Dickey-Fuller test is done on the beta processes to check whether they are stationary. The results are shown in the first column of table 10. Based on this,the null-hypothesis that the time series of the β_1 parameter is non-stationary cannot be rejected. This is puts a serious doubt on whether the use of an AR(1) process to model β_1 is justified.

Next, the differences of the β processes are tested for normality, as the increments of AR(1) processes are normal. The normality tests used are Anderson-Darling, Cramer-von Mises and Kolmogorov-Smirnov. All tests are done with respect to a normal distribution with mean and variance determined by the sample mean and sample variance over the whole time series. The respective p-values are shown at the right-hand side of table 10. The conclusion can be made that the increments of β_1 are normally distributed, but that is not the case for β_2 and β_3.

Two main causes have been found in the study that can explain why the tests fail for β_2 and β_3. The first is that the tails of the observed interest rates are fatter than the theoretical distribution. The

Table 10: ADF test on the β processes and Normality tests on the β differences

differences	ADF	AD	CvM	KS
β_1	0.799	0.430	0.531	0.567
β_2	0.041	0.005	0.009	0.002
β_3	0.000	0.000	0.002	0.003

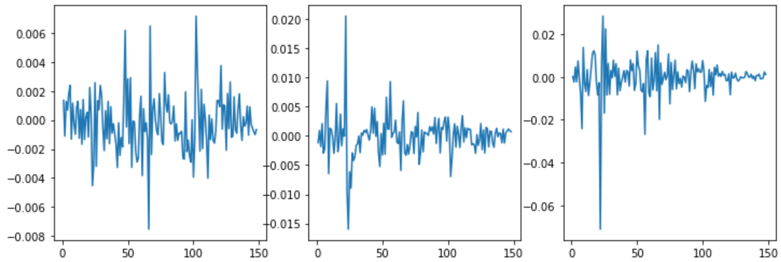


Figure 17: Difference plots of $\beta_1, \beta_2, \beta_3$ (from left to right)

kurtosis for the difference processes of $\beta_1, \beta_2, \beta_3$ are 2.22, 9.48 and 23.39 respectively, and thus only the tails for the first process are close to a normal distribution. The second explanation found is that the variance of the differences do not seem to be constant over time. Both explanations can be seen in figure 17. The middle figure shows the β_2 differences and the right-hand figure the β_3 differences, both exhibit extreme values and tapering behaviour, indication non-constant variance.

Based on the observed diminishing variance for β_2 and β_3 and the earlier formulated hypothesis that interest rate variance may be dependent on the interest rate level, something interesting has been found. When rescaling the aforementioned differences by dividing them by the values for β_1, most of the normality tests cannot be rejected. In other words, the variance of β_2 and β_3 may be dependent on the yield-curve level. [6]

Tests on the fitted AR(1) models

This section analysed the fitted AR(1) models. The previous section indicates in some cases such as the acf plots, that AR(1) may be good for β_2 and β_3. This section presents deeper analysis of the fit of the AR(1) models.

The residuals of the fitted AR(1) processes are shown in figure 18. The first thing to note is that the residuals are very similar to the differences shown in figure 17. This can be explained by the fact that the autoregressive parameters are relatively high, and relatively small constant. When subtracting β_i from the AR(1) process equation, the following is obtained

$$\begin{aligned}\beta_i &= \mu + \phi\beta_{i-1} + \epsilon_i \\ \beta_i - \beta_{i_1} &= \mu + \phi\beta_{i-1} + \epsilon_i\beta_{i_1} \\ \beta_i - \beta_{i_1} &= \mu + (\phi - 1)\beta_{i-1} + \epsilon_i.\end{aligned}$$

Since the left-hand side is equal to the difference, and the right-hand side tends to a normal distribution for ϕ close to one. Because of the similarities between the residuals and the differences, the same observations hold - extremes are not captured well in the model, as is the tapering variance over time.

[6]Note that dependence on the yield-curve level is different from the dependence of variance on the interest rate level as is present in a log-normal model, as the dependence does not come from the level of the process itself.

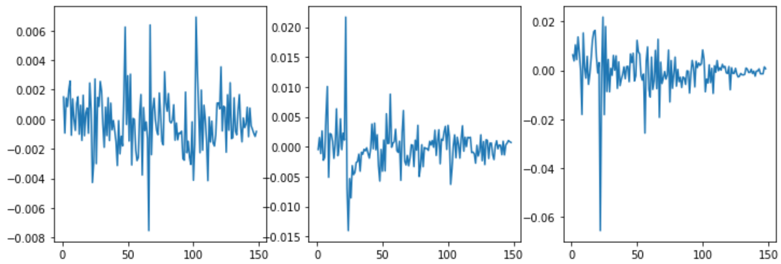


Figure 18: Residuals of the fitted AR(1) processes

Table 11 shows the results of the normality tests, as well as a Ljung-Box test and White test on the residuals. The tests confirm the observations made earlier: The normality tests indicate the residuals of the β_1 model are normal, but not for β_2 and β_3. In addition the White-test confirms (at the 5% significance level) that the variance of the residuals of the latter processes are not constant. Finally, the Ljung-Box test cannot reject the null-hypothesis that the residuals contain no serial correlation. This indicates that modeling the variance of the processes with conditional heteroskedasticity (such as ARCH and GARCH models) may not be the correct choice.

Table 11: Normality, Ljung-Box and White tests on the AR(1) residuals

process	AD	CvM	KS	LB	White
β_1	0.435	0.512	0.567	0.299	0.666
β_2	0.013	0.020	0.013	0.363	0.000
β_3	0.001	0.002	0.006	0.272	0.045

5.5.4 Assuming λ as a constant

λ has been assumed to be constant in similar ways as described in Diebold and Li (2006) and in Nelson and Siegel (1987). The result is that the factor loading of the curvature is always maximized at 2.5 years. However is is very well possible that the curvature has its maximum at different times, see for example week 100 in figure 10.

The parameter λ_t is fixed, and some variations can be applied to analyse whether the model is sensitive to changes in λ. Table 12 shows the backtested distribution with λ being calibrated to 2 and 3 years respectively, as opposed to the 2.5 years mentioned earlier. The choice of λ has an impact on the outcomes for the backtesting of the 30 year maturity. With the 2 year version, it is no longer possible to reject that the zero rate follows from the DNS model at the 5% confidence level.

The observed improvement in backtesting results for the 30 year rate for $\lambda = 0.8966$ can be explained by the fact that it represents the level of the yield curve. As shown in section 5.5.1, β_1 and the level are highly correlated. When the curvature's maximum is shifted to the left, the value of the factor loading of the curvature becomes lower and thus is included to a smaller extent in the 30-year rate process.

5.6 Adjusting the DNS model

Given the poor results of the base DNS as described in Diebold and Li (2006), the results from the model validation is used to test an adjusted version of the model. Three main findings from the model validation are used in this adjusted model:

- Use λ such that is optimizes the curvature on 2 years
- Model β_1 as an random walk with drift

Table 12: Goodness of fit with different values of λ

t	maturity	KS	CvM	AD
t=2 & $\lambda = 0.8966$	1 month	0.000	0.000	0.000
	2 years	0.000	0.000	0.000
	30 years	0.117	0.057	0.000
t=3 & $\lambda = 0.5877$	1 month	0.000	0.000	0.000
	2 years	0.000	0.000	0.000
	30 years	0.000	0.000	0.000

- Model the correlation structure using the most recent daily data

The first two have been treated extensively in the model validation. The third one is a practical solution to the correlation modeling. Since correlation is not observed directly, the approach uses the data points between the monthly data as proxy. This way a correlation structure can be obtained that is close to the instantaneous correlation.

Using the same backtesting settings as previously, the results improve drastically in comparison with the original model. The p-values for the uniformity tests are shown in table 13. The original model resulted in 0 p-values for all zero rate maturites. The improved model can't be rejected to be the data generating model for the shortest and longest few maturities at the 5% significance level. The model is still far from perfect, even with the improvements. Many rates are rejected to be generated by the theoretical model. As the p-values for the Anderson-Darling test are lower, this is likely because of the extreme values modeled incorrectly.

Table 13: Goodness of fit for the adjusted DNS model

maturity	KS	CvM	AD
1 month	0.460	0.445	0.096
6 month	0.375	0.299	0.016
1 years	0.073	0.016	0.000
2 years	0.026	0.029	0.000
3 years	0.015	0.064	0.000
4 years	0.084	0.047	0.001
5 years	0.051	0.042	0.003
10 years	0.099	0.141	0.083
15 years	0.103	0.206	0.142
20 years	0.336	0.355	0.226
30 years	0.304	0.322	0.199

The adjusted model is by no means suitable to be employed in practice exactly as described in this section, but rather serves as a "proof of concept". First of all the model has some problems with modeling the kurtosis. This can be seen in figure 19. The characteristic "flying-bird" shaped empirical distribution indicates that the model features lower kurtosis than the realised values, as also is mentioned in Ruiz (2014). Additionally, all rates that are statistically a bad fit coincide with rates that depend on the curvature, modeled by β_3. Closer inspection of the backtested β_3 process shows that during one interval the auto-regressive parameter is estimated to be close to zero. This happens between 2012 and 2015, as can be seen in figure 11 in the plot of β_3 against curvature. The model calibrated on this period a low probability of an upward period, however an upward period is in fact observed afterwards. The simulated paths for zero rates with maturities between 1 year and 5 years all have multiple entries in the empirical distribution equal or close to one because of this period. One solution to this problem, as also adopted in another way in one of INGs models (FI/FM-Quants (2019)), is to constraint changes in parameter values

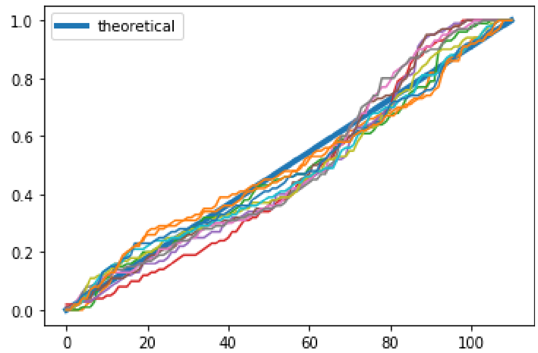


Figure 19: Improved DNS - empirical versus the theoretical distribution

in such a way that they can't change by more than a percentage of the previous value.

In addition, the calibration of the correlations uses the daily rates of the previous month in the backtesting window as proxy for the monthly instantaneous correlation. This assumption has to be verified in more detail.

The first result of this section is the way the model improvements were found. Rather than viewing the dynamic Nelson-Siegel approach as a fixed model, it is recommended to use it as a framework. The parameters may be modelled via different processes in that framework, possibly including stochastic volatility or correlation. It is recommended to analyse the parameter processes carefully, and model them accordingly.

The second principle that may have a wider application is the use of the intra-scenario data in estimating certain parameters. In this study, the default timestep between backtesting scenarios is one month. But since daily data is available during these monthly windows, it can serve as natural proxy for instantaneous parameters.

6 Displaced Diffusion

This chapter describes the full analysis of the Displaced Diffusion model. In section 6.1 the calibration of the model is explained, followed by the simulation equations in section 6.2. After the model is specified, the backtesting results are shown and explained in section 6.3. The model is validated in section 6.4 to gain insight in the risks of the model and the situations in which the model performs inadequate.

6.1 Calibration

The DD model is seen in this context as an approximation to the CEV-LLM. The instantaneous correlations matrix need to be estimated by historical data. The volatilities for the various tenors can be easily estimated via historical data by using $\hat{l}^i_{t_i} + \theta = l^i_{t_i}$ with the latter being log-normal. It is then easy to see, using Ito's lemma, that $\ln(\hat{l}^i_{t_i} + \theta) = \ln(l^i_{t_0} + \theta) - \frac{1}{2}\sigma_i^2(t_i - t_0) + \sigma_i(W^i_{t_i} - W^i_{t_0})$.

Volatility

Then the historical volatility can be obtained by using

$$
\begin{aligned}
\mathrm{Var}\Big(\ln\frac{\hat{l}^i_{t_i}+\theta}{\hat{l}^i_{t_{i-1}}+\theta}\Big) &= \mathrm{Var}\Big(\ln\frac{l^i_{t_i}}{l^i_{t_{i-1}}}\Big)\\
&= \mathrm{Var}\Big(-\frac{1}{2}\sigma_i^2(t_i-t_{i-1})+\sigma_i(W^i_{t_i}-W^i_{t_{i-1}})\Big)\\
&= \sigma_i^2\mathrm{Var}\Big(W^i_{t_i}-W^i_{t_{i-1}}\Big)\\
&= \sigma_i^2(t_i-t_{i-1}). \qquad (13)
\end{aligned}
$$

And thus by adding θ to the observed Libor rates and then taking the log-differences, historical estimates for σ_i can be obtained as

$$
\sigma_i = \sqrt{\frac{\mathrm{Var}\Big(\ln\frac{\hat{l}^i_{t_i}+\theta}{\hat{l}^i_{t_{i-1}}+\theta}\Big)}{\triangle t}}, \qquad (14)
$$

in which $t_i - t_{i-1} = \triangle t$ is fixed. Estimating the volatility in this way does depend on the choice of θ, since the added term doesn't cancel out in calculating the sample variance. This is however as expected, as shown in Oosterlee and Grzelak (2019) a displaced diffusion model with shift parameter β can be seen as a shifted log-normal parametrised as

$$
\frac{\mathrm{d}(l_t+\theta)}{l_t+\theta} = \hat{\sigma_i}\beta\mathrm{d}W_t,
$$

with $\theta = \frac{1}{\beta}(1-\beta)l_{t_0}$ and the estimate for σ_i equal to $\hat{\sigma_i}\beta$. Intuitively this behaviour is expected. Since the entire distribution is shifted and volatility is dependent on the level of the interest rate, the estimated variance will naturally change.

Displacement parameter

Then, a way to estimate the displacement parameter has to be decided. One approach is to use expert opinion. Another is to use a maximum likelihood method, as for example described in Fries et al. (2017). Since the estimation of the parameters σ_i as described before in fact depends on the shift θ by means of the log-differences, these have to be taken into account in the maximum likelihood estimation. Fries et al. (2017) derived that the log-likelihood function is equal to

$$
\theta = \arg\max_\theta\Big(\sum_{i=1}^{n}\log(f(l_{t_{i+1}};l_{t_i},\theta)),\Big)
$$

in which

$$
f(l_{t_{i+1}};l_{t_i},\theta) = \frac{1}{\sqrt{2\pi}\sigma(\theta)(l_{t_{i+1}}+\theta)}\exp\Big(-\frac{1}{2}\frac{\log(l_{t_{i+1}}+\theta)-\log(l_{t_i}+\theta)}{2\sigma(\theta)^2}.\Big).
$$

An important remark has to be made here, in that this method generally results in different estimates of displacement parameters for every maturity (subscript i, left out in the formulations just shown). The representation of the Displaced Diffusion model from Oosterlee and Grzelak (2019) allows for differences in shifts depending on the initial values l_{t_0}. However as discussed in a later chapter this is not enough to account for the rather large difference in estimated shift parameters between shorter and longer Libor forward rates. One solution is to fix the displacement parameter and optimize the above function for all maturities at once.

Correlation structure

The last parameter that has to be estimated is the correlation. Using the definition of correlation, the sample correlation is equal to

$$\mathrm{Corr}(\ln\frac{l^i_{t_i}}{l^i_{t_{i-1}}}, \ln\frac{l^j_{t_i}}{l^j_{t_{i-1}}}) = \frac{\mathrm{Cov}(\ln\frac{l^i_{t_i}}{l^i_{t_{i-1}}}, \ln\frac{l^j_{t_i}}{l^j_{t_{i-1}}})}{\sigma_i\sigma_j\triangle t}.$$

The covariance of the log-differences can be determined as following

$$\begin{aligned}
\mathrm{Cov}\Big(\ln\frac{l^i_{t_i}}{l^i_{t_{i-1}}}, \ln\frac{l^j_{t_i}}{l^j_{t_{i-1}}}\Big) &= \mathrm{Cov}\Big(-\frac{1}{2}\sigma_i^2\triangle t + \sigma_i(W^i_{t_i} - W^i_{t_{i-1}}), -\frac{1}{2}\sigma_j^2\triangle t + \sigma_j(W^j_{t_i} - W^j_{t_{i-1}})\Big) \\
&= \sigma_i\sigma_j\mathrm{Cov}(W^i_{t_i} - W^i_{t_{i-1}}, W^j_{t_i} - W^j_{t_{i-1}}) \\
&= \sigma_i\sigma_j\Big(\mathrm{Cov}(W^i_{t_i}W^j_{t_i}) - \mathrm{Cov}(W^i_{t_i}W^j_{t_{i-1}}) - \mathrm{Cov}(W^i_{t_{i-1}}W^j_{t_i}) + \mathrm{Cov}(W^i_{t_i}W^j_{t_i})\Big) \\
&= \sigma_i\sigma_j\Big(\rho_{i,j}t_{i-1} - 2\rho_{i,j}t_i + \rho_{i,j}t_{i-1}\Big) \\
&= \sigma_i\sigma_j\rho_{i,j}\triangle t. \qquad (15)
\end{aligned}$$

The various covariances in the third step are easy to calculate using $\mathrm{d}W^i_t\mathrm{d}W^j_t = \rho_{i,j}\mathrm{d}t$ and thus constructing a correlated increment using $W^j_t = \rho_{i,j}\hat{W}^i_t + \sqrt{1-\rho^2_{i,j}}\hat{W}^i_t$, with the hat implying independence. In addition it is used that $\mathrm{Cov}(W_t, W_s) = \min(t, s)$. But then using simply the correlation between the observed log-differences can be used to estimate $\rho_{i,j}$ for all i, j.

The DD models every Libor rate in the term structure as a shifted lognormal process correlation to all other rates. Thus the estimation uses the correlation matrix of the differences on which the above derivation is applied.

Drift

Because the model is estimated in a risk-neutral setting, a drift is added to the model and estimated. Assuming σ given, it holds that

$$\begin{aligned}
\mathbb{E}\left[\ln\frac{l^i_{t_i}}{l^i_{t_{i-1}}}\right] &= (\mu - \frac{1}{2}\sigma^2)\triangle t + \sigma^2\triangle t\,\mathbb{E}[Z] \\
&= (\mu - \frac{1}{2}\sigma^2)\triangle t.
\end{aligned}$$

And thus the drift is estimated as

$$\mu = \frac{\mathbb{E}\left[\ln\frac{l^i_{t_i}}{l^i_{t_{i-1}}}\right]}{\triangle t} + \frac{1}{2}\sigma^2.$$

Note that $l^i_{t_i}$ denotes the DD rate with the shift added, and thus is positive.

6.2 Simulation

The basis for the simulation is an Euler-Maruyama scheme. The processes are discretised as

$$\hat{l}^i_{t+\triangle t} + \theta = \hat{l}^i_t + \theta + \mu(\hat{l}^i_t + \theta)\triangle t + \sigma_i(\hat{l}^i_t + \theta)\triangle W^i_t,$$

in which $\triangle W^i_{t+\triangle t} = W^i_{t+\triangle t} - W^i_t \sim N(0, \triangle t)$. Then, by substituting $\hat{l}^i_t + \theta = l^i_t$, one step can be calculated as

$$l^i_{t+\triangle t} = l^i_t(1 + \mu_i \triangle t + \sigma_i \sqrt{\triangle t} Z^i).$$

Since the modeled increments are independent, this can be substituted recursively to simulate k steps as following

$$l^i_{t+k\triangle t} = l^i_t \prod_{j=1}^{k} \left(1 + \mu_i \triangle t + \sigma_i \sqrt{\triangle t} Z^i_j\right).$$

In this, Z_i a standard normal random variable. Given θ, σ_i this is straightforward to simulate for every separate process. However, since multiple correlated paths are simulated, the corresponding correlated Brownian motions need to be build first. Given the correlation matrix as estimated in the previous section, this is easily done by using a Cholesky decomposition. Now, for $i \in \{1, ..., 11\}$ Libor-maturities and k steps, one entire MC simulation of all maturities can be calculated using

$$\begin{bmatrix} l^1_{t+k\triangle t} \\ \vdots \\ l^{11}_{t+k\triangle t} \end{bmatrix} = \begin{bmatrix} l^1_t \\ \vdots \\ l^{11}_t \end{bmatrix} \prod_{j=1}^{k} \left(1 + \begin{bmatrix} \mu_1 \triangle t \\ \vdots \\ \mu_{11} \triangle t \end{bmatrix} + \begin{bmatrix} \sigma_1 \sqrt{\triangle t} Z^1_j \\ \vdots \\ \sigma_{11} \sqrt{\triangle t} Z^{11}_j \end{bmatrix} \right)$$

When simulating this, it is very convenient to do this as a cumulative product. Then, for every path one has to simulate one standard normal correlated matrix. This can be done by using the lower Cholesky decomposition $\mathbf{L}$ and an uncorrelated standard normal matrix $\hat{\mathbf{Z}}$ as

$$\mathbf{Z} = \begin{bmatrix} Z^1_1 & \dots & Z^{11}_1 \\ \vdots & \ddots & \vdots \\ Z^1_k & \dots & Z^{11}_k \end{bmatrix} = \mathbf{L}\hat{\mathbf{Z}}.$$

6.3 Performance backtesting

In this section, the results of the backtesting of DD is presented. First, the empirical distribution is analysed in section 6.3.1. Then section 6.3.1. continues with the backtesting results of the PFE.

6.3.1 Empirical distribution

In this section the empirical distribution is tested. The Displaced Diffusion model is used to generate paths and these are subsequently checked against the realized values, as described in chapter 4. Figure 20 shows an overal example of the backtesting, with non-overlapping intervals. Qualitatively, the empirical distribution appears to be of lower variance at the lower interest rates. In contrast with the results of the dynamic Nelson-Siegel model, this results in smaller quantiles as well, enabling better results on the PFE estimation.

The result should be an uniform distribution. The results for a non-overlapping backtesting experiment with one month windows are shown in table 14.

Table 14: Goodness of fit on one month intervals

maturity	KS	CvM	AD
1 month	0.001	0.000	0.000
2 years	0.079	0.110	0.000
5 years	0.000	0.000	0.000
30 years	0.123	0.054	0.041

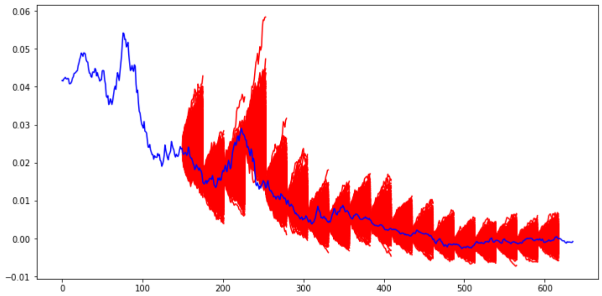


Figure 20: 6-month non-overlapping backtesting sample paths against the 2-year zero rate

It is clear that the model does not perform very well on the short to medium term, and it is rejected at the 5% confidence level that the realized values come originate the theoretical model. However for the longest maturity, it cannot be rejected that the data could come from the theoretical distribution.

For 3-month overlapping intervals the results remain the same. The results are shown in table 15.

Table 15: Goodness of fit on three month intervals

maturity	KS	CvM	AD
1 month	0.001	0.000	0.000
2 years	0.003	0.007	0.000
5 years	0.008	0.011	0.001
30 years	0.038	0.02	0.018

As a final experiment the model is tested on a one year and two year overlapping intervals. The results are shown in table 16 and table 17 respectively. It is clear that the model does not perform well on longer backtesting windows in terms of empirical distribution.

Table 16: Goodness of fit on one year overlapping intervals

maturity	KS	CvM	AD
1 month	0.001	0.000	0.000
2 years	0.000	0.000	0.000
5 years	0.006	0.001	0.000
30 years	0.058	0.011	0.005

To gain insight in why exactly the distributions cannot be verified to be the sample, figure 21 shows the empirical distribution for the 2-year overlapping backtesting window against the theoretical $U(0,1)$ distribution. In the 5 year rate predictions, lower quantiles are over represented. For the 2 and 30 year rate forecasts this is the opposite and the realised values are more often found in the higher quantiles of the generated distributions.

6.3.2 MtM backtesting

In this section the results of the MtM backtesting experiment are presented. Table 18 contains the results of the goodness-of-fit tests. Most of the tests fail, which is as expected based on the results of the backtesting experiment for the empirical distribution. However the model does perform well on the FRA120x180 contract.

Table 17: Goodness of fit on two year overlapping intervals

maturity	KS	CvM	AD
1 month	0.000	0.000	0.000
2 years	0.001	0.003	0.003
5 years	0.000	0.000	0.000
30 years	0.063	0.031	0.013

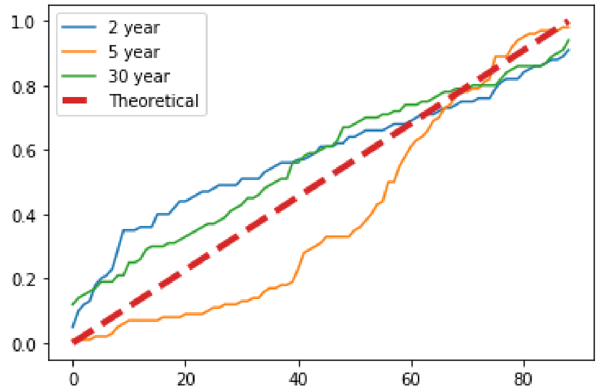


Figure 21: Empirical plot for the 2,5 and 30 year maturities versus the theoretical

6.3.3 PFE backtesting

The number of exceedences of the PFE for various FRA contracts is shown in table 19. The results are mixed. In some cases the model seems to perform well, but for some contracts the model is on the conservative side. There is one case where the number of exceedences is too high. However it is perfectly possible in this many tests that one of the values exceeds the 95% quantile of the chi-squared distribution.

Figure 22 shows the situations in which the model isn't conservative at the 95% confidence level. There are two possible explanations for the large number of exceedences even at the $PFE_{99\%}$.

- Initial calibration period is too short, as all the exceedences take place at the start
- half-yearly recalibration is too slow to catch the increased volatility present at the start. The PFE's can be seen to be increasing even after the higher volatility period.

Overall the model performs well and is on the conservative side. It falls under the green light for all but one value, which is orange, meaning the excess is manageable. As Ruiz (2014) mentions, the ideal result is lots of green lights, a few orange and no red light. This is the case in the backtesting as performed in the study.

6.4 Model Validation

In this section, the underlying assumptions of the DD model are analysed. The goal is to research if there are critical errors following from these assumptions, that can lead to model risk in certain cases. Three assumptions are researched:

- The constant shift incorporated in the model. The assumption of using a constant shift is research, as well the possibility of including the shift parameter in the calibration process.
- The use of a normal distribution for generating increments.
- The long term behaviour of the model.

Table 18: Results of of the MtM backtesting experiment

Contract	Window	KS	CvM	AD
FRA6x9 payer	1 month	0.000	0.000	0.000
FRA6x9 receiver	1 month	0.000	0.000	0.000
FRA24x36 payer	24 months	0.000	0.001	0.001
FRA24x36 receiver	24 months	0.000	0.000	0.000
FRA120x180 payer	12 months	0.065	0.113	0.009
FRA120x180 receiver	12 months	0.000	0.000	0.000

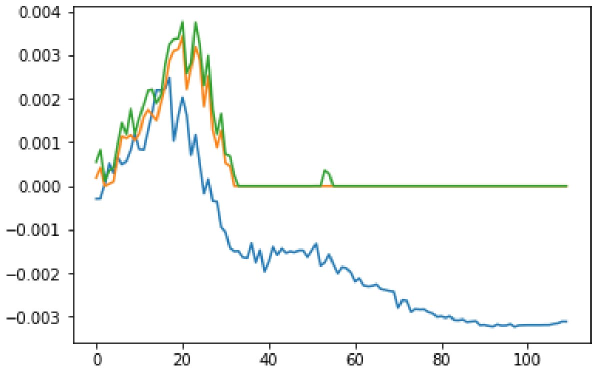


Figure 22: PFE's of the FRA6x9 payer with K=0.01 after 3 months

6.4.1 The shift parameter

The model is bounded below by the shift parameter, and thus any scenario involving even lower interest rates is deemed impossible. In principle, this is somewhat undone by re-calibrating the model frequently, including a review of the used shift parameter. This way, sharp movements towards the initial lower bound may make that a lower shift parameter should be used. There is still a problem when estimating the CCR measures over longer time horizons. As mentioned before (in Actuarial Association of Europe (2016)), European policy on for example a cashless society could mean that interest rates go even lower. In such cases, even if improbable, risk will be severely underestimated.

In addition to the shift parameter being a hard lower bound, one constant shift parameter is assumed for the model for all of the Libor rates. To verify whether this assumption is feasible, the shift parameter has been estimated using the maximum likelihood method as described in section 6.1. The shift is estimated separately for every process using a maximum shift of 0.4, or 40%, as an arbitrary maximum for optimisation purposes. The results are shown in table 20.

Longer maturities are estimated to have a larger shift (about -7% is estimated to be the lower bound for the 5-year) than shorter maturities (about -0.5 % seems to be the lower bound). This result can be explained as following. Looking at the Libor rate between one month and half year, it is quite clear the variance of the process is close to zero. Since the DD model is dependent on the level, and the estimated volatility is not close to zero, it must be concluded that *if* the process is indeed a shifted lognormal process, *then* the process is close to its minimum.

On the other hand for processes with maturities further in the future, large variance is still present even at values close to 0%. Since there is reason to believe the interest rate variance is dependent on the level (see chapter 2), the conclusion that these rates can become quite more negative is not far-fetched. An economic interpretation of the bounds on interest rates is given in Actuarial Association of Europe (2016). It is mentioned that in the current economy the interest rates are mainly bounded because of the cost of holding cash, and the risk of the crowd making a mass retrieval in case of too low interest rates.

Table 19: Exceedences of PFE for DD

K	contract	time	#Obs	95%	99 %	$LR_{95\%}$	$LR_{99\%}$
K=0.01	6x9 payer	1 month in	112	6	3	0.029	2.184
	6x9 payer	3 months in	110	8	4	1.055	4.606
	24x36 payer	24 months in	89	2	0	1.771	1.789
	6x9 receiver	1 month in	112	0	0	11.490	2.251
	6x9 receiver	3 months in	110	6	3	0.047	2.253
	24x36 receiver	24 month in	89	2	1	1.771	0.013
K variable	6x9 payer	1 month in	112	5	2	0.070	0.566
	6x9 payer	3 months in	110	5	2	0.049	0.599
	24x36 payer	24 months in	89	0	0	9.130	1.789
	6x9 receiver	1 month in	112	1	1	5.951	0.013
	6x9 receiver	3 months in	110	0	0	5.782	0.009
	24x36 receiver	24 month in	89	3	0	0.559	1.789
K= -0.01	6x9 payer	1 month in	112	5	2	0.070	0.566
	6x9 payer	3 months in	110	6	1	0.047	0.009
	24x36 payer	24 months in	89	3	0	0.559	1.789
	6x9 receiver	1 month in	112	6	3	0.029	2.184
	6x9 receiver	3 months in	110	0	0	5.782	0.009
	24x36 receiver	24 month in	89	0	0	9.130	1.789
K=0.05	6x9 payer	1 month in	112	0	0	11.490	2.251
	6x9 payer	3 months in	110	0	0	11.285	2.211
	24x36 payer	24 months in	89	0	0	9.130	1.789
	6x9 receiver	1 month in	112	1	1	5.951	0.013
	6x9 receiver	3 months in	110	0	0	11.285	2.211
	24x36 receiver	24 month in	89	0	0	9.130	1.789

However in the case of a cashless society, such a lower bound ceases to exist. While a cashless society is not the case in the very near future, it is certainly possible in the medium or far future.

Finally, for maturities of 10, 15 and 20 years, the maximum likelihood does not seems to converge to a realistic value. Note that a bound of 0.4 is set for finding the optimal value. This might indicate that the shifted lognormal process may not be the correct process for these Libor rates.

Including the shift parameter in calibration

Given the availability of a calibration method for the shift parameter, it is interesting to see if it is viable to include in the calibration of the backtesting. However, one of the main model risks of the DD model is when interest rates go lower than the lower bound. It is tested if this happens when calibrating the data historically and checking the obtained values for theta against the minima observed over the two years following.

It turns out that it happens frequently that the shift parameter, when estimated using data up to a point, is exceeded within two years by the realised rates. This not only shows that this way of estimating the shift can lead to model risk, it also stresses the importance of having a conservative value for theta.

Table 20: Maximum-Likelihood of the shift

maturity	Shift
1 month	0.0042
6 months	0.0036
1 year	0.0044
2 years	0.0066
3 years	0.0128
4 years	0.0291
5 years	0.0685
10 years	0.3999
15 years	0.3999
20 years	0.3999
30 years	0.0318

Backtesting using different shifts

From section 6.3 it is clear that the DD model does not perform on all maturities. Based on the finding that the shift might differ for the various zero rates, the backtesting is repeated on small scale for $\theta = 0.005$ and $\theta = 0.05$. In addition this can provide some insight in the sensitivity of the backtesting to changes is the shift paramater

It is expected for the large theta to see better results for high maturity rates, and the small theta for shorter maturity rates. The backtesting is only done on the empirical distribution, using monthly non-overlapping intervals. The results are shown in table 21 and table 22. The model perform better for the shortest rate when $\theta = 0.005$, and for the 5 year rate when $\theta = 0.05$, compared to the original case when using 0.02 as shift. The statistical tests however still predominantly reject that the observed values come from the respective distribution at the 5% confidence level.

Table 21: Goodness of fit for $\theta = 0.005$

maturity	KS	CvM	AD
1 month	0.041	0.053	0.044
2 years	0.010	0.023	0.008
5 years	0.0	0.003	0.001
30 years	0.196	0.246	0.169

Table 22: Goodness of fit for $\theta = 0.05$

maturity	KS	CvM	AD
1 month	0.000	0.000	0.000
2 years	0.000	0.000	0.000
5 years	0.005	0.018	0.025
30 years	0.109	0.054	0.042

6.4.2 Normality of residuals

In this section, the residuals over the whole dataset are analysed. This is provides a general overview of the fit of the model. In addition some conclusions can be derived on the correctness of modeling randomness via (standard) normal random variables.

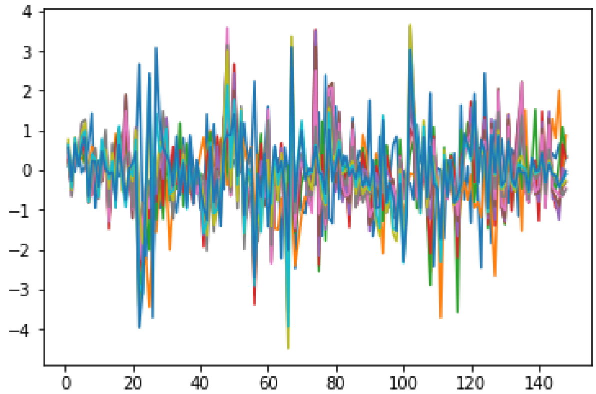


Figure 23: Residuals of the DD model

The principle of obtaining the standardized residuals is straightforward. Given that the discretized log returns including shift are modeled as

$$\log(\frac{l_{t_{i+1}} + \theta}{l_{t_i} + \theta}) = (\mu - 0.5\sigma^2)\triangle t + \sigma\sqrt{\triangle t}Z$$

with $Z \sim N(0,1)$, the residuals are simply obtained by isolating Z. The result is:

$$Z = \frac{\log(\frac{l_{t_{i+1}}+\theta}{l_{t_i}+\theta}) - (\mu - 0.5\sigma^2)\triangle t}{\sigma\sqrt{\triangle t}}$$

The parameters μ, σ are estimated as described in section 6.1, and a shift of 2% is used. The residuals are shown in figure 24 for all Libor rates. The residuals seem stationary, however some larger deviation are seen. This is rather uncommon for a normal distribution.

The normality tests are performed on the residuals to verify the normality assumption statistically. The Anderson-Darling test, Cramer-von Mises test and Kolmogorov-Smirnov are used, with reference to a standard normal distribution. The results are shown in table 23.

Table 23: Goodness of fit for $\theta = 0.05$

maturity	KS	CvM	AD
1 month	0.004	0.003	0.003
6 month	0.000	0.000	0.000
1 years	0.014	0.007	0.005
2 years	0.199	0.323	0.210
3 years	0.136	0.213	0.260
4 years	0.006	0.035	0.037
5 years	0.013	0.031	0.028
10 years	0.457	0.200	0.154
15 years	0.085	0.064	0.035
20 years	0.101	0.060	0.046
30 years	0.169	0.065	0.034

The results are mixed. As was analysed in chapter 3, heteroskedastic behaviour could not be rejected for shorter maturity rates. Since this is one of the main modeling features of the Displaced Diffusion model, it is expected to see a good fit for the shorter maturity rates. However for the shortest three

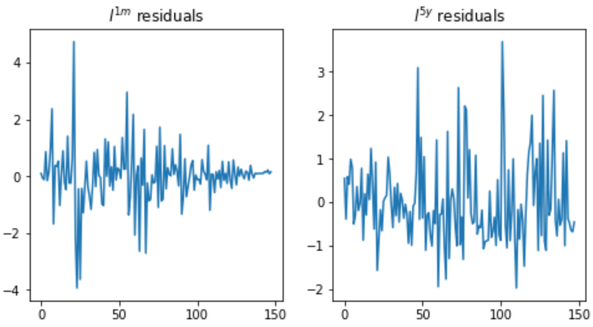


Figure 24: Residuals of the 1 month and 5 year libor rate

maturities, all normality tests are rejected at the 5% confidence level. Then the early mid maturities are convincingly normally distributed, as expected. The late mid maturities are again rejected at the 5% confidence interval. The longer maturities seem to have normally distributed residuals, only for the tails of the distribution this is rejected at the 5% level.

The plots from figure 24 seem to have more extremes than one would expect from a standard normal distribution. In addition the statistical tests fail mostly on Anderson-Darling, which has a heavier weight on the tails of the distribution. Thus the assumption of an underlying normal distribution is not entirely correct and the reason is likely to be in the tails of the distribution.

Finally, the residuals still seem to feature heteroskedastic behaviour in the cases where the goodness-of-fit tests fail. Figure 24 show the residuals of the 1 month and 5 year Libor rate. It is apparent that the 1 month residual variance still depends on the level of the rate. This may indicate that the approximation of the CEV model by the DD model is incorrect in this case. It is recommended to study the use of the CEV, as it can model the y parameter in the SDE Chan et al. (1992) researched (see section 3.1).

6.4.3 Long term behaviour

The Displaced Diffusion model with a drift under the real world measure can behave in two different ways on the very long time horizon. It mainly depends on the sign of the drift parameter. If it is negative, the process on the long term will tend to the lower bound - the shift. When it is positive, the process is not bounded above. In earlier chapters, some example paths were generated to show the overall behaviour of the the theoretical model.

There are 100 paths are generated over a 83 year horizon using both the model fitted with a drift and without to see the long term behaviour. The 83 year horizon is based on 1000 steps using monthly intervals. All paths for the 3 year Libor rate with drift are shown in figure 25 and with drift in figure 26. It is clear that the long term behaviour can explode over very long time horizons when not including a drift. In contrast, when including a drift the model tends to the lower bound over time. In principle this is as expected. This behaviour also holds for the other maturity Libor rates.

However it is worth to mention that products in a counterparty credit risk setting are tested over the time they run. It is not common for products to run over 30 years. In that time frame (around the 500th observation in the long term plots), behaviour is still somewhat reasonable, given that Libor forward rates are modeled. Thus it can be concluded that the long term behaviour is realistic for reasonable long term behaviour, but caution has to be taken in the case of extremely long time horizons.

7 Conclusion

In this report, analysis is done on the dynamic Nelson-Siegel and Displaced Diffusion models when used in a counterparty credit risk setting. The models are used to calculate the PFE. They have then been

Figure 25: Long term behaviour without drift

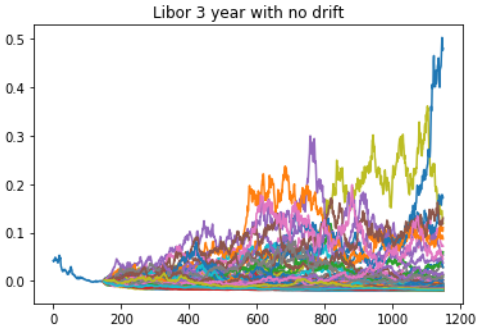
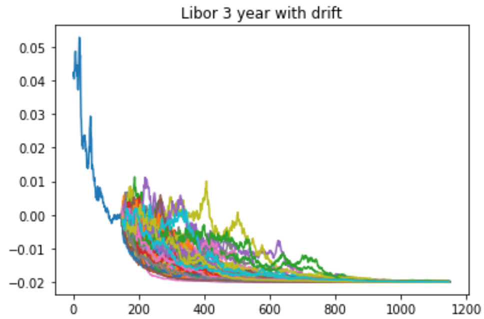


Figure 26: Long term behaviour with drift



backtested against historical data and performance on PFE estimation has been done. At the beginning, some analysis has been done on the specific time series to gain a better understanding of when models perform stronger than at other times.

In this report, analysis is done on the dynamic Nelson-Siegel and Displaced Diffusion models when used in a counterparty credit risk setting. Based on literature and preliminary data analysis,

Some evidence has been found that the interest rate processes for shorter maturities are heteroskedastic. Another observation done, based on the statement just made, is that shorter and longer maturity interest rates could follow different processes. In addition the interest rates do not seem to feature mean reverting behaviour. However, the timespan used in this study is just over 12 years of data - often not enough to show true mean reverting behaviour.

Given that Displaced Diffusion features heteroskedastic behaviour but no mean reversion, one would expect that it should outperform the dynamic Nelson-Siegel model. This is the case in our study. The results show the following

- Dynamic Nelson-Siegel is in most cases overconservative. The reason being that an AR(1) process does not capture the behaviour of β_1 well and because of that the auto-regressive parameters are too large.
- Dynamic Nelson-Siegel is underestimating risk when using longer backtesting windows and on contracts that are out of the money. The reason is that the model tends to increase to a mean, while in reality the interest rates predominantly have been going down.
- Displaced Diffusion performs better, both in backtesting the empirical distribution and the PFE. In most cases, the model is on the conservative side, the PFE exceedences are tested to be statistically plausible. The model seems to be too conservative with receiver contracts, given the downward trend captured in the model. In addition, the model is overconservative when the fixed leg is much higher than the actual rates.

Even though the results for backtesting of the PFE are useable in some cases for the Displaced Diffusion model, the model does not always perform well at the empirical distribution. This is especially the case for longer time horizons and for shorter maturity rates. It is therefor concluded that the model can be used as benchmark for shorter maturity contracts that are at- or in-the-money.

Recommendations for future study

It is recommended to challenge the models over longer time horizons using a larger data set. In the current study, with a burn in time of 3 years, one can only use backtesting windows of a couple of years at best, without compromising the number of scenario's. In addition, overlapping windows need to be used for all time horizons longer than a couple of months, for the same reason.

The use of daily rates as proxy for the instantaneous correlation matrix used for the adjusted DNS model is recommended to study more closely. Especially for parameters that are not observed in the market, the daily subset of data can represent a useful way of estimation. Given that CCR analysis is often done over longer time horizons with relatively large step-sizes, it may be relevant for other models as well, not just for the DNS model.

Model recommendations

It is recommended use the dynamic Nelson-Siegel model as a framework rather than a fixed model. Diebold and Li (2006) analyse various different types of models for the beta parameters (from random walks to vector-autoregressive models) and come to the conclusion that AR(1) models are best. However as reasoned in the conclusions, that result can be put to doubt on the data in this study, especially for the β_1 parameter representing the level of the yield curve. When analysing the parameters, different processes were found to be more fitting. In addition, the correlation may be modelled. It is suggested to perform multiple analyses on the processes and their dependencies to find a suitable model.

The assumption that all zero rates (or Libor rates) follow the same type of process can be challenged, as written in the conclusions. One solution is given by the previous recommendation - since the beta parameters in the dynamic Nelson-Siegel model affect different maturities in different ways, using different processes for these respective parameters can in principle model this. Other possibilities are regime-switching processes such as for example described in Hamilton (2008), but in which one lets the different rates switch from regime independently. Regime switching models are also a good option because of their macro-economic interpretation. Many of the interest rates are forced to be low by ECB policy. Once the lower rate policy is lifted, the regime of interest rates are likely to change.

A single constant as shift parameter might be too restrictive. It is recommended to check if the shift is still realistic every time the model is recalibrated. One way to do this is as done in chapter 6 - by finding the various shifts by maximum likelihood estimation. However, that method does not work equally well on all maturities, so some care is required. In addition it was found that in calibrating the shift on part of the data, and backtesting on the two years after, it happens often that the realised rates attend values lower than the estimated lower bound. This is in important model risk. It is suggested to take the expert based estimation, taking into account the most negative observed zero rates till date. This should include other currency/national rates that are low.

References

Aas, K., L. R. Neef, L. Williams, and D. Raabe (2018). Interest rate model comparisons for participating products under solvency ii. *Scandinavian Actuarial Journal*, 203–224.

Actuarial Association of Europe (2016, December). Negative interest rates and their technical consequences. Technical report. Obtained on 5 March 2020 from https://actuary.eu/memos/negative-interest-rates/.

Anderson, T. W. and D. A. Darling (1954). A test of goodness of fit. *Journal of the American Statistical Association 49*(268), 765–769.

Annaert, J., A. G. P. Claes, M. J. D. Ceuster, and H. Zhang (2012). Estimating the yield curve using the nelson-siegel model: A ridge regression approach. *International Review of Economics and Finance*, 482–496.

Bonollo, M., L. Di Persio, I. Oliva, and A. Semmoloni (2015). A quantization approach to the counterparty credit exposure estimation. *SSRN Electronic Journal*.

Breusch, T. S. and A. R. Pagan (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica 47*(5), 1287–1294.

Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders (1992). An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance 47*(3), 1209–1227.

Cramér, H. (1928). On the composition of elementary errors. *Scandinavian Actuarial Journal 1928*(1), 13–74.

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*(366), 427–431.

Diebold, F. X. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics 130*(2), 337 – 364.

Diebold, F. X. and G. D. Rudebusch (2013). *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach.* Princeton University Press.

FI/FM-Quants (2019, June). Technical model documentation for trading risk models - potential future exposure, v1.4.

Fries, C., T. Nigbur, and N. Seeger (2017). Displaced relative changes in historical simulation: Application to risk measures of interest rates with phases of negative rates. *Journal of Empirical Finance 42*, 175–198.

Ghamami, S. and B. Zhang (2014). Efficient monte carlo counterparty credit risk pricing and measurement. *FEDS Working Paper* (2014-114).

Gregory, J. (2011). *Counterparty Credit Risk: The new challenge for global financial markets.* Wiley.

Gregory, J. (2012). *Counterparty Credit Risk and Credit Value Adjustment: A continuing challenge for glocal financial markets.* Wiley.

Hamilton, J. (2008, 01). *Regime-Switching Models.*

Ibáñez, F. (2016). Calibrating the dynamic nelson-siegel model: A practitioner approach. Working Papers Central Bank of Chile 774, Central Bank of Chile.

Kolmogorov, A. (1928). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari 4*, 83–91.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives 3*(2), 73–84.

Nelson, C. and A. F. Siegel (1987). Parsimonious modeling of yield curves. *The Journal of Business 60*(4), 473–89.

Oosterlee, C. and L. Grzelak (2019). *Mathematical Modeling and Computation in Finance.* World Scientific Publishing Europe Ltd.

Rebonato, R. (2003, 03). Term-structure models: a review.

Rebonato, R. (2004). Review paper. interest-rate term-structure pricing models: a review. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 460*(2043), 667–728.

Rubinstein, M. (1983). Displaced diffusion option pricing. *The Journal of Finance 38*(1), 213–217.

Ruiz, I. (2014). Backtesting counterparty risk: How good is your model? *Journal of Credit Risk 10*(1), 87–120.

Shumway, R. H. and D. S. Stoffer (2005). *Time Series Analysis and Its Applications (Springer Texts in Statistics)*. Berlin, Heidelberg: Springer-Verlag.

Smirnov, N. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bull. Maths. Univ. Moscou 2*, 3–14.

Smirnov, N. (1944). Approximate laws of distribution of random variables from empirical data. *Uspekhi Mat. Nauk 10*, 179 – 206.

von Mises, R. (1928). *Wahrscheinlichkeit, Statistik und Wahrheit*. Julius Springer.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica 48*(4), 817–838.

8 Appendix A: PFE profiles

8.1 Dynamic Nelson-Siegel

In this section, the PFE's that are calculated in the backtesting of the Dynamic Nelson-Siegel model are show against the realised values of the FRAs that are tested. The plots represent all the backtesting experiments described in table 6.

8.2 Displaced Diffusion

In this section, the PFE's that are calculated in the backtesting of the Displaced Diffusion model are show against the realised values of the FRAs that are tested. The plots represent all the backtesting experiments described in table 6.

FRA using a 1% fixed leg

Figure 27: FRA6x9 payer after 1 month

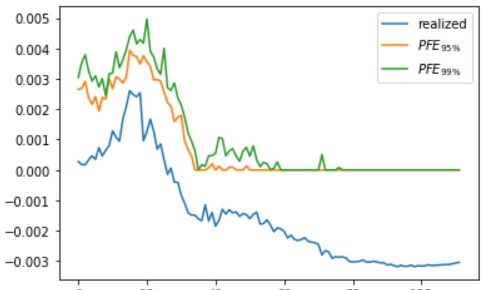


Figure 28: FRA6x9 receiver after 1 month

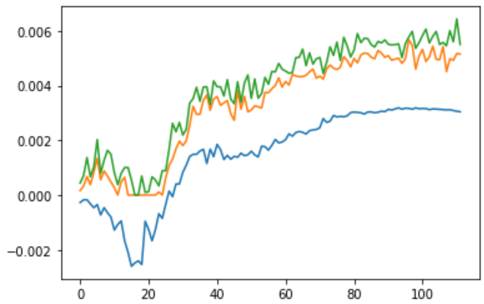


Figure 29: FRA24x36 payer after 24 months

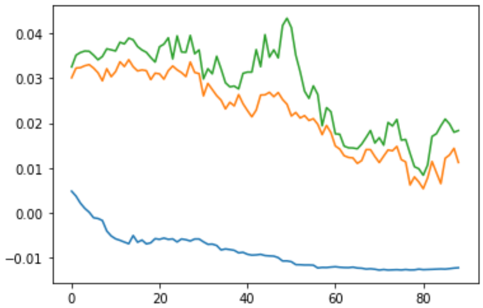


Figure 30: FRA24x36 receiver after 24 months

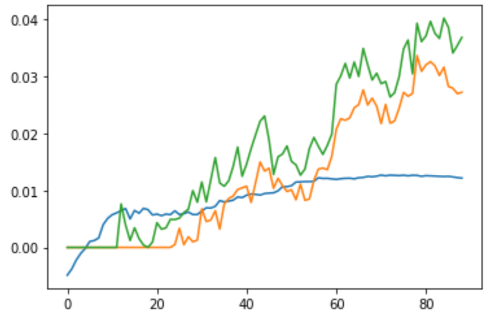


Figure 31: FRA6x9 payer after 3 months

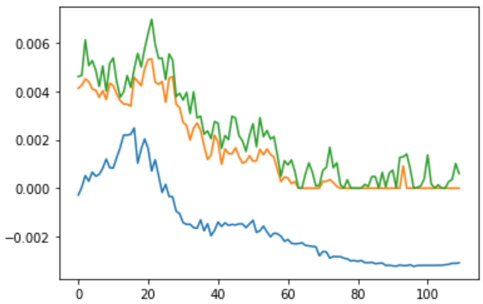
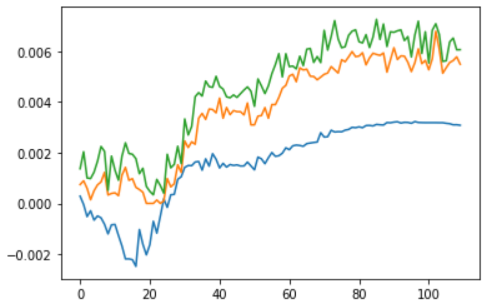


Figure 32: FRA6x9 receiver after 3 months



FRA using a variable fixed leg

Figure 33: FRA6x9 payer after 1 month

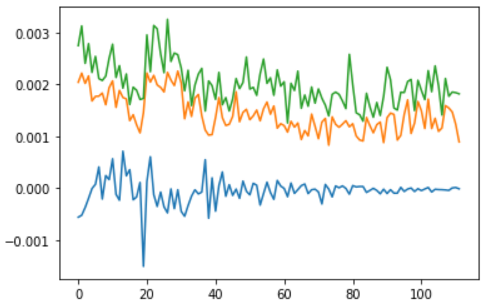


Figure 34: FRA6x9 receiver after 1 month

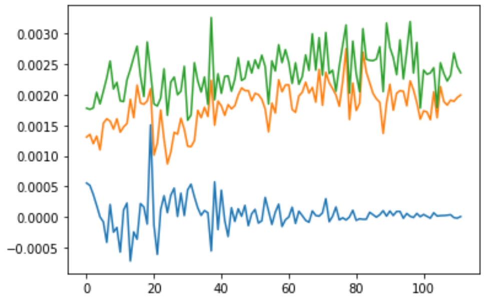


Figure 35: FRA24x36 payer after 24 months

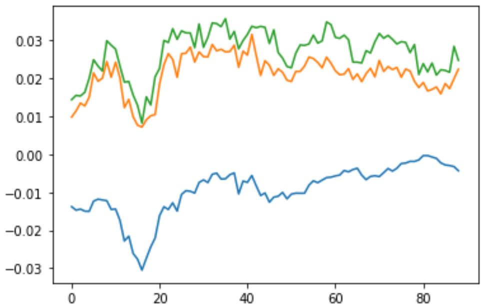


Figure 36: FRA24x36 receiver after 24 months

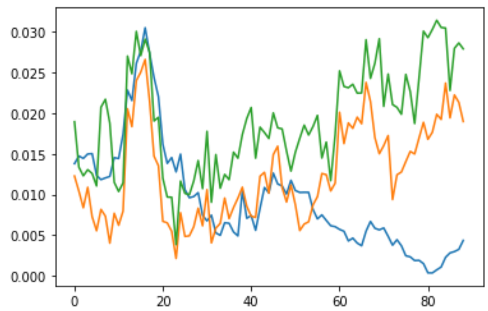


Figure 37: FRA6x9 payer after 3 months

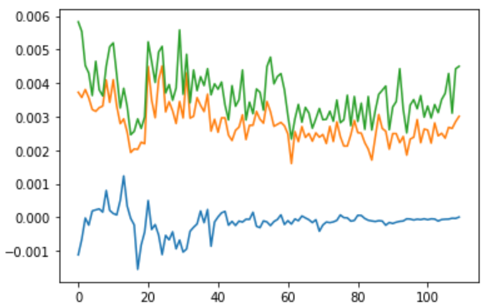
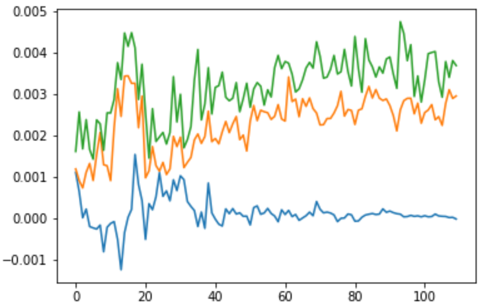


Figure 38: FRA6x9 receiver after 3 months



FRA using a 5% fixed leg

Figure 39: FRA6x9 payer after 1 month

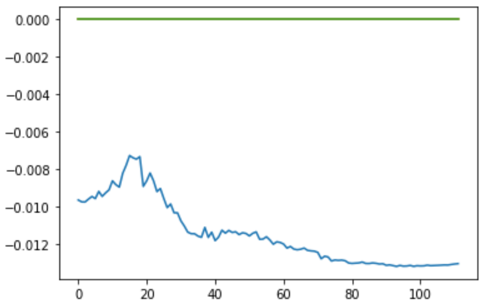


Figure 40: FRA6x9 receiver after 1 month

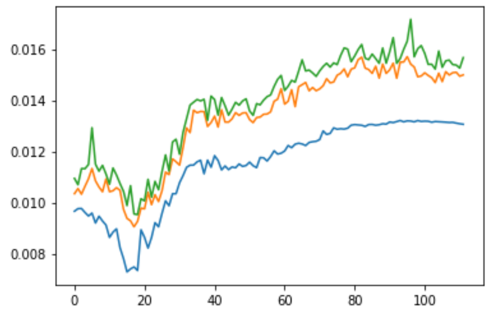


Figure 41: FRA24x36 payer after 24 months

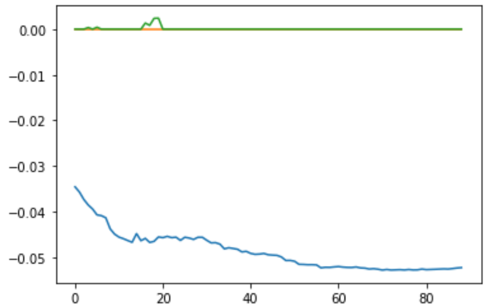


Figure 42: FRA24x36 receiver after 24 months

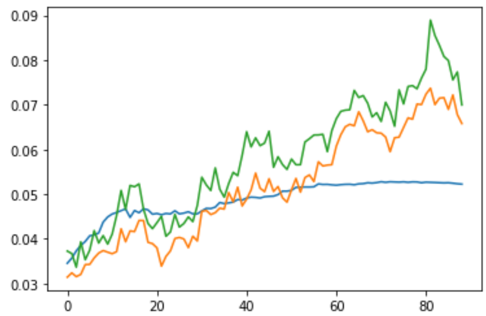


Figure 43: FRA6x9 payer after 3 months

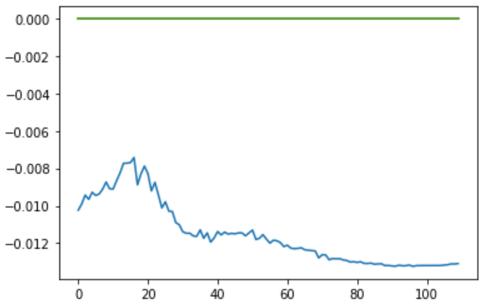
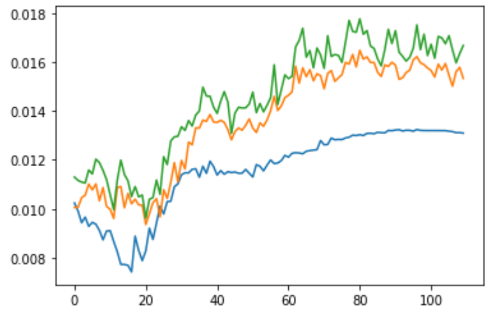


Figure 44: FRA6x9 receiver after 3 months



FRA using a -1% fixed leg

Figure 45: FRA6x9 payer after 1 month

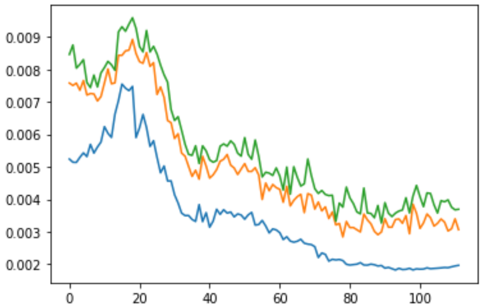


Figure 46: FRA6x9 receiver after 1 month

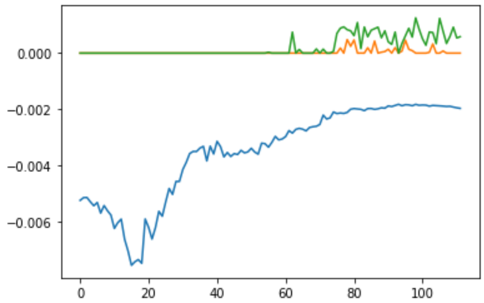


Figure 47: FRA24x36 payer after 24 months

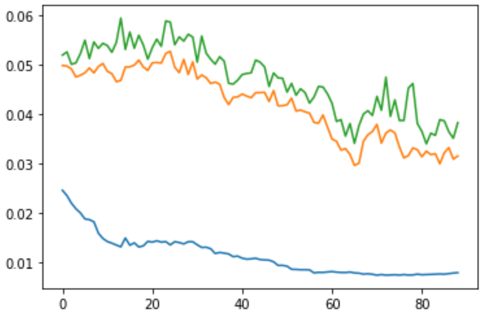


Figure 48: FRA24x36 receiver after 24 months

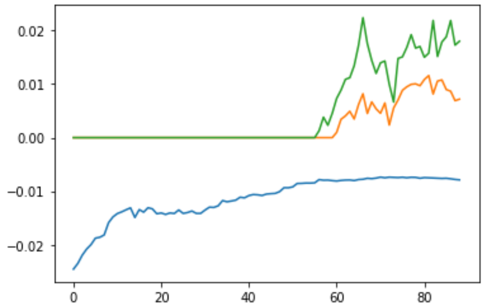


Figure 49: FRA6x9 payer after 3 months

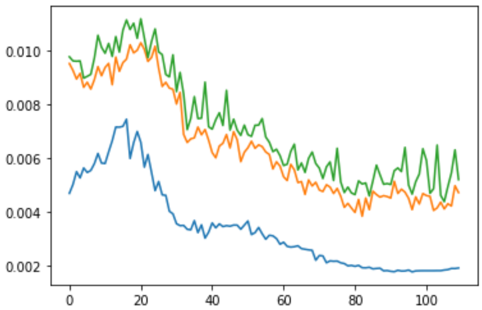
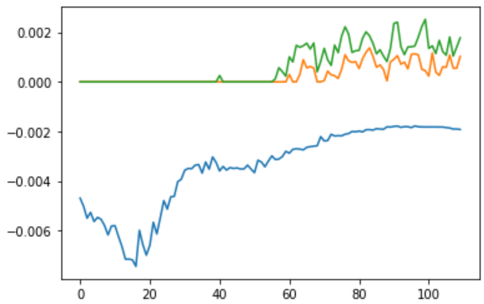


Figure 50: FRA6x9 receiver after 3 months



FRA using a 1% fixed leg

Figure 51: FRA6x9 payer after 1 month

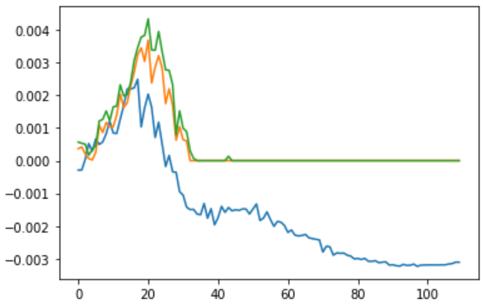


Figure 52: FRA6x9 receiver after 1 month

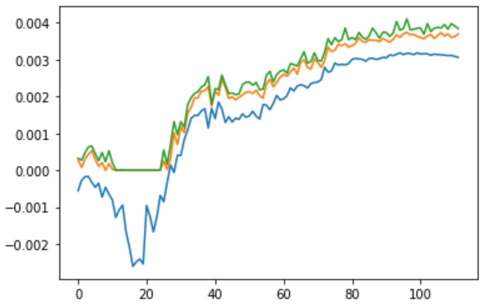


Figure 53: FRA24x36 payer after 24 months

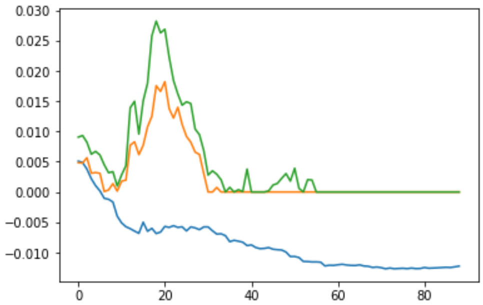


Figure 54: FRA24x36 receiver after 24 months

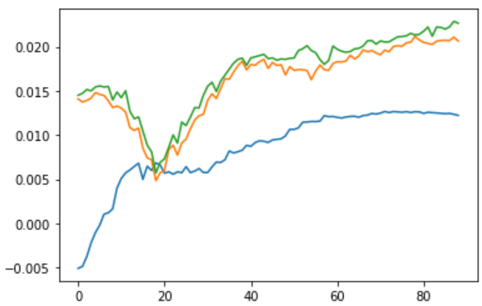


Figure 55: FRA6x9 payer after 3 months

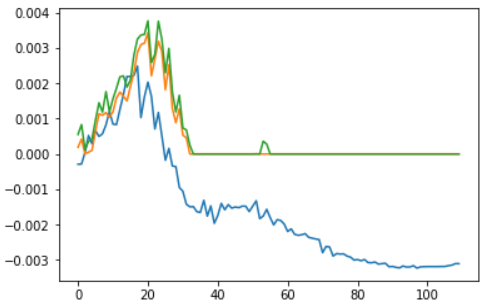
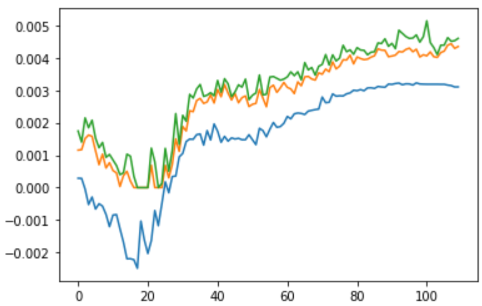


Figure 56: FRA6x9 receiver after 3 months



FRA using a variable fixed leg

Figure 57: FRA6x9 payer after 1 month

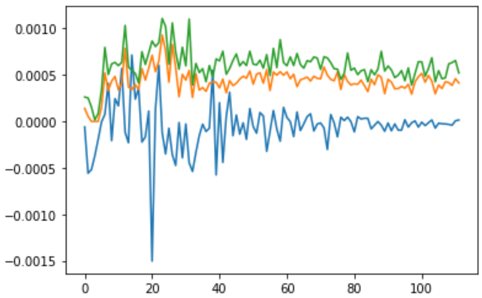


Figure 58: FRA6x9 receiver after 1 month

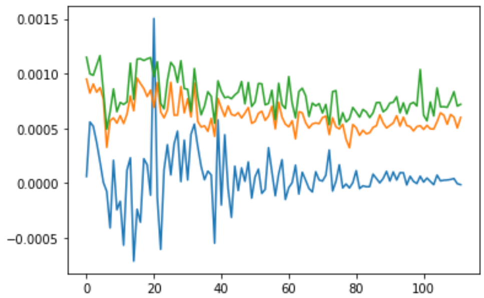


Figure 59: FRA24x36 payer after 24 months

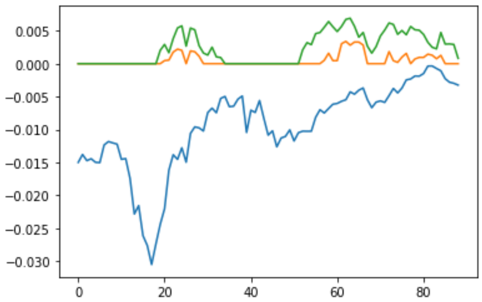


Figure 60: FRA24x36 receiver after 24 months

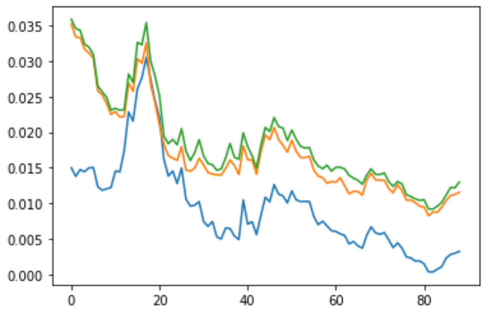


Figure 61: FRA6x9 payer after 3 months

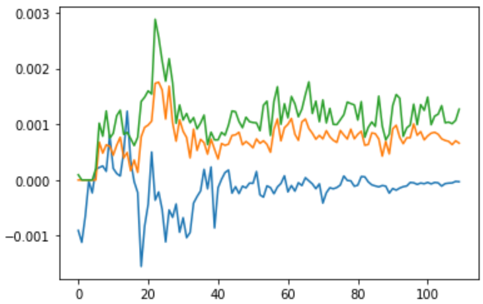
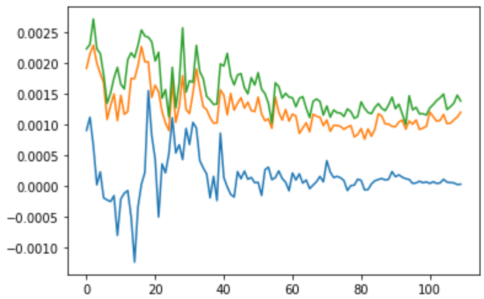


Figure 62: FRA6x9 receiver after 3 months



FRA using a 5% fixed leg

Figure 63: FRA6x9 payer after 1 month

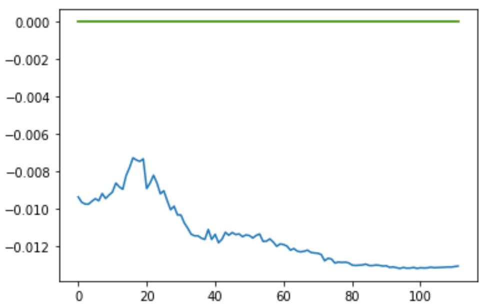


Figure 64: FRA6x9 receiver after 1 month

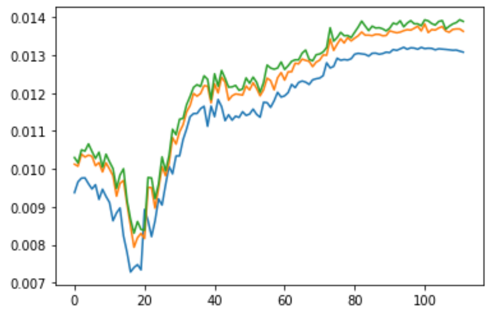


Figure 65: FRA24x36 payer after 24 months

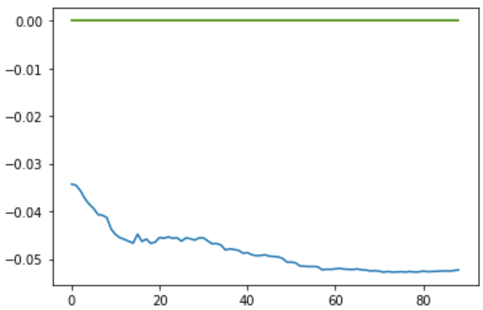


Figure 66: FRA24x36 receiver after 24 months

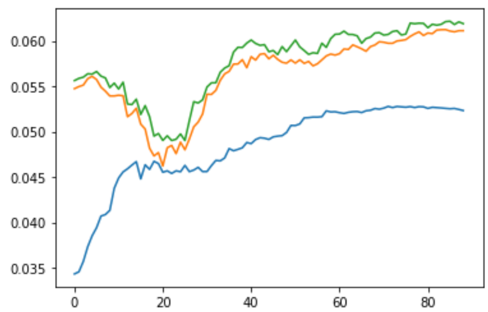


Figure 67: FRA6x9 payer after 3 months

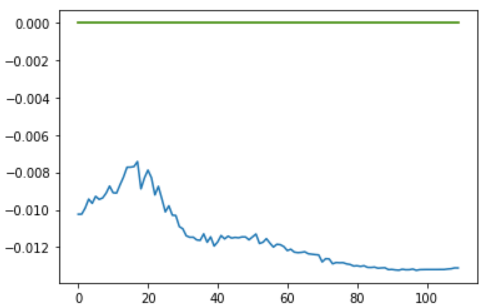
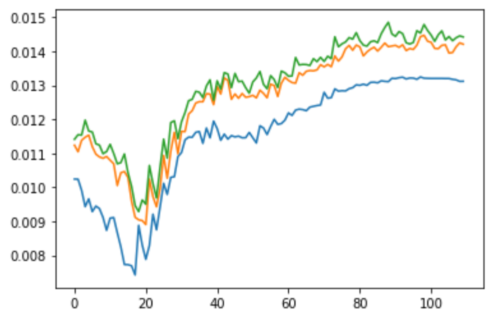


Figure 68: FRA6x9 receiver after 3 months



FRA using a -1% fixed leg

Figure 69: FRA6x9 payer after 1 month

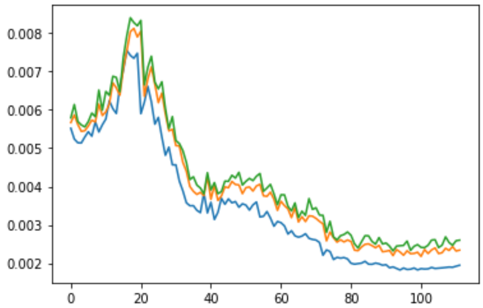


Figure 70: FRA6x9 receiver after 1 month

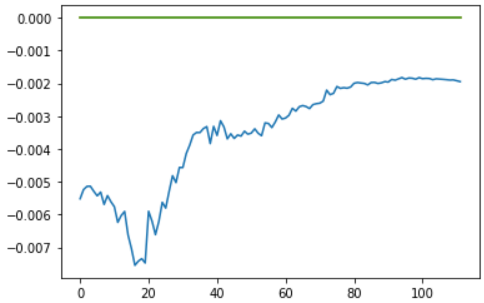


Figure 71: FRA24x36 payer after 24 months

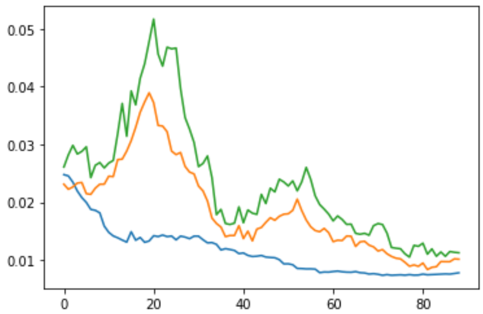


Figure 72: FRA24x36 receiver after 24 months

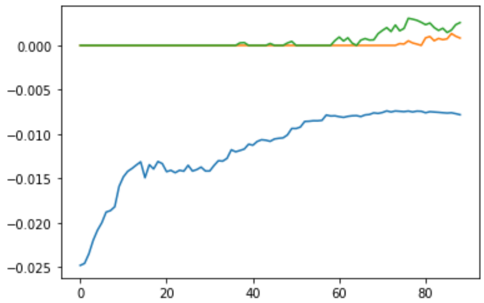


Figure 73: FRA6x9 payer after 3 months

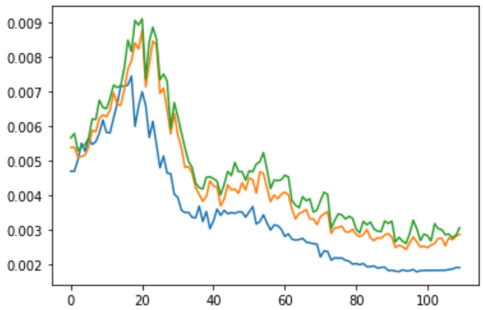


Figure 74: FRA6x9 receiver after 3 months