Netspar
Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Marie Wahlers

Valuation of Long-Term Liabilities under Solvency II

Extrapolation Methods for the European Interest Rate Market

MSc Thesis 2013-044

Maastricht University

Master Thesis - Financial Economics:

Valuation of Long-term Liabilities under Solvency II - Extrapolation methods for the European interest rate market

Marie Wahlers
i6001744

Supervisor:
Prof. Dr. Peter Schotman

July 30, 2013

Abstract

For the upcoming Solvency II framework, the Smith-Wilson method constitutes the extrapolation technique for valuing long-term liabilities. Next to the Smith-Wilson method this paper considers three other methods to extrapolate the long end of the Euro yield curve. Two of them belong to the popular Nelson-Siegel family and the other represents the equilibrium class of interest rate term structure models. Evaluated by the models' ability to reduce volatility for long-term yields and by their extrapolation errors, the Smith-Wilson method performs best. Unlike the other methods, it does not require parameter estimations. Only since mid-2011, when yields have shifted to unprecedentedly low levels, the Smith-Wilson method deteriorates and the Nelson-Siegel model exhibits substantially better fitted extrapolations.

Contents

1 Introduction **3**

2 Literature Review **8**

2.1 Interest rate concepts 8

2.2 Yield curve models 12

2.2.1 Statistical models 12

2.2.2 Economic models 21

3 Data and Methodology **25**

3.1 Data 26

3.2 First step - Bootstrapping 28

3.3 Second step - Calibration 30

3.4 Third step - Extrapolation 32

3.5 Fourth step - Comparison 33

4 Results **34**

4.1 Calibration 34

4.2 Extrapolation 39

4.3 Comparison 42

4.3.1 Variance test 43

4.3.2 Extrapolation error test 44

5 Discussion **48**

6 Conclusion **54**

References **60**

A Matlab code for the Smith-Wilson calibration and extrapolation **64**

B MCS test results for base models based on the second subsample **64**

1 Introduction

The risky consequences of insurance companies' high asset allocations in the UK and the US have led Europe to move away from traditional actuaries' valuation of liabilities and to turn to a market-consistent approach. In November 2009, the European Union's Council and Parliament approved the European Commission's proposal for the Solvency II Directive (European Insurance and Occupational Pensions Authority, 2013). Despite the potentially promising economic impact of the new directive discounting liabilities at a market interest rate instead of an actuarial and artificial one, at the time of approval, the consequences of the financial crisis were severe. The alarmingly low funding ratios of insurance companies induced them to appeal to the regulators. 'It is clear that a blind application of market consistency can cause volatility and pro-cyclicality [...]. The financial crisis was a textbook case in how in abnormal market conditions market consistency can undermine business. It is good that it happened before we finalised Solvency II' (Carver, 2011). Under the new directive, for the valuation of insurance companies' liabilities a discounting approach is used that requires extrapolation of the market's interest rate term structure. Given the desirable transparency inherent in the market-based approach along with the valuation challenges caused by the financial crisis, the question about an appropriate way of extrapolating the yield curve arises. This research provides insights into potential methods that achieve extrapolations and evaluates their relative performance.

Solvency II is a new supervisory framework that is aimed at unifying the insurance industry in the European Union (EU) by replacing the currently fourteen separate insurance directives. The new coherent regulation is designed to enhance equal protection of European insurance policyholders who face different insurance practices across the EU member countries. The major protective measure entails the reduction of risk for insurance customers in the event of the insurance company's insolvency. When insurance companies cannot meet their obligations, policyholders have to bear the losses of their claims themselves. Limiting insolvencies is only successful when risk is being assessed from the beginning on. Thus, in order to take pre-emptive measures, Solvency II provides signals to warn supervisory authorities such as the European Insurance and Occupational Pension Authority (EIOPA) whenever the risk for policyholders increases. The mechanism ultimately increases the confidence in a stable EU insurance market.

Analogous to the Basel III regulations for the banking industry, the prudential regime of Solvency II is based on three pillars. The first pillar encompasses quantitative requirements regarding solvency capital standards. The second pillar underlies requirements for risk-management as well as supervision, and the third pillar stipulates disclosure procedures (Internal Market DG - Financial Institution Insurance, 2004).

Since capital requirements are central to the health and success of insurance companies, the first pillar is subject to most of the current discussion surrounding the implementation of Solvency II. The extent to which insurance companies' own funds will be affected by the new regime has been tested in field studies under the most recent, fifth Quantitative Impact Study (QIS). It reports that compared to the Solvency I system, the capital requirements under the new directive are estimated to be one to six times higher depending on the type of insurer (European Insurance and Occupational Pensions Authority, 2011). Since capital not only has protective benefits but also comes along with costs, future capital raising may influence the insurer's profitability. Only in the theoretical market of Modigliani and Miller (1958), a firm's value is unaffected by its financing decisions. Relaxing the assumptions about perfect capital markets, gives rise to agency costs and asymmetric information issues, which may outweigh the benefits of holding additional capital to buffer potentially large losses in the future (de Haan & Kakes, 2010).

These capital requirements follow from the valuation of the assets and liabilities of insurance companies under Solvency II. The asset-side of an insurer consists of investment portfolios that ensure sufficient funds to cover the insurance company's obligations. The investment products are usually tradeable and their value is hence constantly determined by the market. On the other hand, the liabilities which take on the form of insurance contracts are not traded and thus not priced by the market. For a comparison to the asset side under the Solvency II's economic balance sheet approach, the liabilities should also be valued market-consistently. This approach is based on a transfer argument as postulated by Article 75 of Solvency II (Directive 2009/138/EC): 'Liabilities shall be valued at the amount for which they could be transferred, or settled, between two knowledgeable willing parties in an arm's length transaction'. This 'amount' is the value of the so-called technical provision that according to Article 77 amounts to the sum of the present value of liability cash flows and a risk margin that ensures the buying party to be willing to take over all obligations. The present value shall be derived by 'using the relevant risk-free interest rate term structure' (Article 75). Since the risk-free interest rate term structure is an unobservable concept, it has to be specified what observed interest rate instruments should be used for its derivation. The EIOPA switched from proposing European investment grade government bonds during QIS 4 to European interest rate swaps in QIS 5.

Furthermore, Article 75's proposed discounting approach creates two main challenges. First, the payment date of insurance companies' projected future cash flows may exceed the term of the market interest rates and hence no present value can be readily computed. Consider, for instance, a life insurance contract bought by a 20-year old woman that may live another 80 years. The pricing of this life insurance is not possible at the observable long-term interest rates since there is no swap contract that matures in 80 years. Second, even for life insurance contracts that are expected to last 'only' 40 years, proper discounting may not be practical because longer-term interest rates are too volatile. The reason for this is that the interest rate term structure completely depends on the market instruments that it is stripped from. Thus, the volatility of the implied rates is solely determined by the liquidity of the market instruments. Less liquid bonds or swaps are more sensitive to trades due to their low trading volume. Highly fluctuating interest rates then directly translate into volatile liability estimates and capital requirements. The resulting necessity to adjust capital on a day-to-day basis is unreasonable and unpractical. Hence, according to EIOPA (2010, p.2): 'Valuation of technical provisions and the solvency position of an insurer or re-insurer shall not be heavily distorted by strong fluctuations in the short-term interest rate'. This suggests the presence of liquidity distortions on the European swap market.

Due to the suggestion that the market of long-term interest rates is either non-existing or illiquid, an extrapolation method is needed to extend the interest rate term structure beyond the last observable or liquid market rate. Mathematically, extrapolation techniques range from simple to advanced methods (Thomas & Maré, 2007). The simplest way to extrapolate is to take the last observed interest rate and assume that rate to hold for all other terms. The result of a constant long-end term structure is not in line with the requirements set out by Solvency II for at least two reasons. First, the simplicity will very likely lead to extrapolated rates that are very different from the true interest rates, which imply inconsistency with the new directive's market valuation approach. Second, all extrapolated rates will, by construction, vary whenever the final rate changes, causing a possibly quite unstable curve and therefore violating EIOPA's guidelines for avoiding fluctuations. Consequently, more advanced extrapolation methods are necessary to suit the Solvency II criteria for extrapolation.

Generally, the methods should be based on plausible economic assumptions about the behavior of interest rate term structures. One of the most important properties is that interest rates approach a long-run level when the term increases indefinitely. As a consequence, the long-run equilibrium rate represents the heart of the method proposed by the Solvency II Directive and is termed the Ultimate Forward Rate (UFR). This technique was developed by Smith and Wilson (2001) and proposes a transition from the last liquid rate to the UFR. Since the UFR reflects market's long-term expectations about interest rates, and expectations rarely change fundamentally, the method is designed to produce stable extrapolations. EIOPA determines the UFR as the sum of expected real interest rates and inflation. With an inflation target of 2% and expected real interest rate proxied by past, generally much higher rates than today, the UFR seems to be too high, given the current very low nominal interest rates. Hence, academics and practitioners from the insurance industry have expressed doubts regarding the market-consistency of the Smith-Wilson method (Academic Community Group, 2012). The main concern is about the discrepancy between the high UFR and the lower real long-term interest rate level. The induced upward bias in extrapolated rates artificially depresses the present value of liabilities. The resulting low capital requirements may not be appropriate in regard to the true interest rate risk. Hereof, if the proposed Smith-Wilson technique turns out not to be best suited for Solvency II, an alternative extrapolation method will be needed. Hence, next to Solvency II's Smith-Wilson method, this research considers two classes of alternative methods that will potentially fulfill Solvency II's criteria.

The first class of methods consists of statistical models which aim at fitting the observed yield curve by imposing a function on it. The Nelson-Siegel method (1987) is the most popular example because it models the yield curve by its three components - level, slope and curvature. In order to further improve the fit of the Nelson-Siegel method, Svensson (1994) extends it with one more yield curve factor - a second curvature component. The resulting family of Nelson-Siegel models is used around the world by all major central banks (Bank for International Settlements, 2005) and has also gained recent attention in the academic world due to the models' reformulation, reinterpretation and theoretical derivation (Diebold, Piazzesi, & Rudebusch, 2005; Diebold & Li, 2006; Christensen, Diebold, & Rudebusch, 2007, 2008).

The second alternative is the class of economic models pioneered by Merton (1974) and Vasicek (1977) who derive a formulation for the pricing of zero coupon bonds from an underlying stochastic differential equation for the instantaneous interest rate. The pricing solution for bonds of different maturities can be transformed into a solution for the zero yields that can then be extrapolated. Since the Vasicek model aims to provide an economical motivation for the yield curve, the model does not perfectly fit the observed yields. Not fitting all yields exactly does not preclude sensible extrapolations, quite the contrary. Extrapolations, in their nature, aim at being different to observed long-term rates in order to avoid undesirable volatility.

Given the alternative models, the paper compares the Smith-Wilson extrapolations to those produced by the Nelson-Siegel, Svensson and Vasicek models. In light of Solvency II's requirement of stable and market-consistent liability valuations, two main evaluation criteria form the basis for comparison. On the one hand, the extrapolations should ensure stability and on the other hand, they should provide an unbiased reflection of the actual, more volatile long-term interest rates. The two qualifications led me to the following two research questions:

Research Question 1. What extrapolation method best fulfills the extrapolation criterion for reducing the volatility at the long end of the yield curve?

Research Question 2. Which of the four models produces extrapolations that reflect the current interest rate environment, thereby providing an unbiased estimate of the actual long-term yields?

I use the European swap market data to construct yield curves. The dataset consists of daily fixed-for-floating swap rates from January 4, 1999 to April 17, 2013 and includes maturities one to 10, 12, 15, 20, 25, 30, 40 and 50 years. I estimate the four models with stripped zero swap rates based on the liquid maturities of one to 20 years. Estimates of the yields for maturities 21 to 50 are derived by extrapolations of each model. These extrapolated rates are compared to the actual stripped zero rates for maturities 25, 30, 40 and 50 that have been excluded from the estimation before.

I find that the Smith-Wilson method performs best. In regard to stability, it is the only model that significantly reduces volatility for the 30-, 40- and 50-year yields, whereas the Vasicek model even increases standard deviations compared to the actual rates. The Nelson-Siegel and Svensson methods, on the other hand, neither increase volatility, nor decrease it. On that account, the more parametrized the yield curve model, the less stable it is. No model manages, however, to significantly stabilize the 25-year yield suggesting the yield to be already very liquid.

In regard to the bias criteria, I find that all models produce an upward bias in their extrapolations. However, the Smith-Wilson method does so at least. One major factor responsible for the common overestimation is that no model is able to capture the convexity effect - a phenomenon that arises due to the non-linear price-yield relation for fixed income securities inducing longer-term instruments to trade at a lower yield.

Not only do the Vasicek and the Svensson models perform least well in terms of stability, but they are inferior to the SW and the Nelson models also in terms of biases. During the last two years of the sample period, the role of the Smith-Wilson and Nelson methods reverses even leading to an underperformance of the Smith-Wilson method compared to all other models. Whereas the large upward bias recently

observed with the Smith-Wilson method is due to the too high and fixed Ultimate Forward Rate, the Nelson model's extrapolations benefit from the lately downward shift of the yield curve.
In general, the findings indicate a trade-off between the two criteria of stability and bias. The lower the standard deviations for the extrapolated yields compared to those of the actual yields, the more strongly the model systematically overestimates long-term yields.

The results are beneficial for the ongoing discussion and implementation of the new Solvency II framework as there is up to now no stand on the best extrapolation method. Extant literature and research has mainly focused on advancing yield curve models to best fit and forecast the current term structure. The aim is to find a model that prices instruments or quantifies interest rate risks as exactly as possible. De Pooter (2007), for example, shows that extending the Nelson-Siegel model by additional factors increases fitting and forecasting accuracy. However, the results cannot be directly extended to extrapolation performance as flexibility may likely increase volatility, thereby opposing the extrapolation aim of stabilizing the long end of the yield curve.
Thomas and Maré (2007) provide a mathematical presentation of potential extrapolating models and evaluate them in terms of each model's hedging accuracy. My focus, however, lies in the usefulness of the extrapolating models to stabilize the long end of the yield curve rather than in their applicability to risk management. Furthermore, the set of extrapolation models by Thomas and Maré (2007) is composed of the Nelson-Siegel class and the Smith-Wilson approach. I add an equilibrium model to the collection of statistical models under consideration. Since Thomas and Maré (2007) apply the models on the South-African interest rate market, the lack of other studies induces the EIOPA to rely on the African case study results for the Smith-Wilson specification instead of the European ones. Thus, this research contributes by shedding light on the performance of the Smith-Wilson model in the European environment.
Annaert, Claes, De Ceuster, and Zhang (2013) study extrapolations in the European interest rate swap market, thereby being closest to what I do. They, however, focus exclusively on the Svensson model by refining its estimation by conditional ridge regressions. They find that their proposed estimation technique outperforms extant econometric procedures. Their evaluation is based on fitting errors of the long-term yield level, volatility and volatility forecast. Instead of focusing on the best estimation technique, I concentrate on finding the best extrapolating model among a set of three other models next to the Svensson model. Therefore, I do not only consider the fitting performance but also the models' usefulness in reducing volatility at the long end.

The paper proceeds as follows. In the first part of section 2, I introduce fixed income concepts that are relevant for the derivation of interest rate term structures, in particular the procedure of bootstrapping the yield curve from market instruments. The second part of section 2 provides a discussion of the two main approaches that I consider for modelling the yield curve. The first approach is statistical and includes in particular the Nelson-Siegel, the Svensson and the Smith-Wilson methods. I summarize the models' features, advantages and disadvantages that have been already studied in the extant literature. The second approach is economic and focuses on the Vasicek model, in particulary describing its mathematical derivation and distributional properties. The literature review ends with two propositions about the expected four models' extrapolation results. Section 3 outlines the necessary steps to judge the models' extrapolation performances. It describes the stripping of zero yields from the dataset of daily Euro swaps, the estimation of the four models, the way of extrapolating, and the tests applied to judge each model's extrapolations. Section 4 first presents the estimation results and second the extrapolation outcomes. The results for the variance and accuracy tests follow. Section 5 discusses the findings in light of the propositions I derive in the course of the literature review. A conclusion is drawn in section 6 giving rise to possible implications and opportunities for future research on the extrapolation topic.

2 Literature Review

This chapter introduces the concepts and notations central to the yield curve models that are used for the extrapolations in this paper. It outlines in general the derivation of an interest rate term structure by considering important characteristics of the fixed income market. Based on the yield curve concept, it motivates the selection of interest rate models by outlining each model's features and performance that have been already extensively studied in the existing literature.

2.1 Interest rate concepts

Interest rates are the groundwork to any term structure model. There are many types of interest rates depending on the credit risk associated with the borrower and the time period of borrowing (Schmidt, 2011). In line with a market-consistent valuation of insurers' liabilities, Solvency II stipulates the use of risk-free interest rates (European Insurance and Occupational Pensions Authority, 2010). In reality, we have to approximate the risk-free interest rate by calculating the interest rate associated with government bonds that are essentially considered to be default-free for most industrialized countries. Along this line, the risk-free yield curve depicts the yields of government bonds as functions of the bonds' maturities (Sundaresan, 2009). Given that bonds' yields do not only differ in their time to maturity but also in other features such as their coupon rate abd their vintage and callability, it is preferable to have a yield concept that can be universally applied. Here, the term structure of interest rates comes in very useful. It is the relation between the yields of zero-coupon bonds as a function of time to maturity (Sundaresan, 2009). Define $P(t,T)$ as the price of a zero coupon bond at time t with maturity T and a face value of unity so that $P(T,T) = 1$. When using continuously compounded interest rates, the price of the zero bond is

$$P(t,T) = 1 * e^{-R(t,T)(T-t)}, \tag{1}$$

where $R(t,T)$ is the bond's yield at time t with maturity T and is mostly referred to as the spot rate of interest. A Spot rate defines the interest rate that you earn when investing an amount today (Hull, 2009). Depending on how long you invest your principal amount, there are infinitely many spot rates, e.g. the two-year spot rate $R(0,2)$ gives the interest earned on an investment from now until in two years. Hence, the term structure of interest rate at time t can be derived from the price of zero bonds by solving equation 1 for the yield as

$$R(t,T) = -\frac{1}{T-t} ln(P(t,T)). \tag{2}$$

Especially for maturities below one year, instead of extracting zero yields from bond prices, interest rates in the money market constitute the beginning part of the yield curve. In particular, the Euro Inter Bank Offered Rate (Euribor) is used to represent spot rates with maturities of one week to one year. According to Hull (2009), the rate at which banks deposit money with each other is regarded as a 'better' risk-free rate since zero-coupon bond rates can be distorted due to tax and other regulatory treatments. Additionally, in contrast to coupon-paying bonds, money market rates give a direct and not an implied or stripped value for the spot rates.

Alternatively, the term structure can be described by a set of forward rates instead of spot rates since the former concept can be derived from the latter and vice versa. A forward rate $F(t,T,S)$ is the interest rate that applies to the future period between T and S but is set at time t by a forward agreement contract (Hull, 2009). By the no-arbitrage condition, the relation between the spot rate and the forward rate can be defined. For instance, investing into a zero bond with maturity S must be equal to buying a zero bond with maturity T and selling forward its unity face value. The proceeds on the two strategies have to be the same to ensure no-arbitrage possibilities,

$$e^{R(t,S)(S-t)} = e^{R(t,T)(T-t)} e^{F(t,T,S)(S-T)}. \tag{3}$$

Solving for the implied forward rate results in

$$F(t,T,S) = \frac{R(t,S)(S-t) - R(t,T)(T-t)}{S-T} = \frac{1}{S-T} * ln\left(\frac{P(t,T)}{P(t,S)}\right). \tag{4}$$

For modeling purposes, instead of taking the spot and forward rates for a specific period, their instantaneous complements are used. The spot rate that applies to an infinitely small investment period is commonly called the short rate and plays an important role in modeling the term structure. One factor interest rate models are entirely determined by the evolution of the short rate which is defined as

$$r(t) = \lim_{T \to t} R(t,T). \tag{5}$$

Similarly, the instantaneous forward rate becomes

$$f(t,T) = \lim_{S \to T} F(t,T,S). \tag{6}$$

Although being a rather artificial interest rate concept, the instantaneous foward rate is more convenient since it only depends on two and not three points in time (Cairns, 2004). Furthermore, in practice, the instantaneous forward rate rather than the spot rate constitutes the yield curve. Svensson (1994, p.14) emphasizes the advantageous interpretation of the forward curve compared to that of the spot rate curve: 'Whereas the yield curve can be interpreted as expected averages of future short rates, the forward curve can be interpreted as indicating the expected time path of future short rates'. Consequently, while the

short rate is of particular interest in the academic world regarding theory-building, the instantaneous forward rate is popular among practitioners such as central banks (Svensson, 1994).
According to equation 2, the term structure depends on prices of bonds. However, the EIOPA proposes swaps and not the commonly used bonds to determine the term structure. The main reason is that the European market for swap contracts has grown large and become very liquid.. Figure 1 exemplary illustrates this development by indicating a sharp rise of the swap market in Germany from 1986 to 2013. Similar developments apply to other European countries.

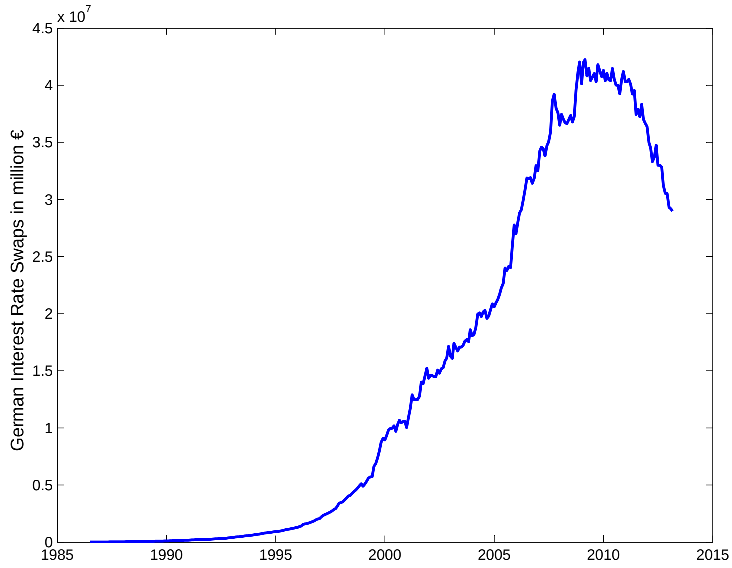


Figure 1: **The rise of the swap market**
End-of-month notional amounts of interest rate swaps contracted by all German banks over the period from 30-Jun-1986 to 31-May-2013. The volumes are reported in millions of current Euro and total 324 observations.

Furthermore, when swap contracts are signed every day in large numbers, it is possible to find each day a set of swaps with constant maturities ranging from one up to 50 years. Hence, the swap market ensures a complete construction of the yield curve every day. Dai and Singleton (2000) argue that this does not apply to the bond market since there are not every day newly issued government bonds with constant maturities. Also, bonds may be distorted as they are specially treated in the repurchase agreement market or are subject to special tax treatments. In contrast, swaps carry little counter party risk so that they may closely approximate a zero term structure(Duffie & Huang, 1996). Another reason for using Euro swaps is that they circumvent a potential disagreement about the selection of a particular European country to serve as benchmark for all EU countries (Beber, Brandt, & Kavajecz, 2009).
Generally, an interest rate swap is a contract to exchange a fixed rate - the swap rate - for a variable rate - the Euribor - at predetermined times in the future for a specific period. Swap contracts, hence, can be interpreted as swapping a variable coupon bond for a fixed bond. Pricing swap contracts involves that, by no-arbitrage, the present value of the fixed cash flows equals that of the floating ones. Since the floating rate is the current interest rate in the money market, the coupon rate is equal to the yield. Thus, newly issued swap contracts can be treated as a set of par yield bonds.
Consequently, the procedure of deriving the yield curve is independent of whether using swaps or bonds

as inputs. Both require the method of bootstrapping that deduces the yields implied by the market instrument prices (Hull, 2009). In matrix notation it holds that $M = C * d$ where M is a vector of k-market instrument prices, C is a matrix of $k * n$ cash flows associated with each instrument, and d is a vector of n-discount factors. The discount factors equal the current prices of the unity zero-coupon bonds $P(0,T)$ that in turn imply the term structure. Thus, solving for d involves inverting C such that $d = C^{-1} * M$.

Bootstrapping the zero yields is relatively fast to compute especially when all k-bonds pay their cash-flows at the same points in time. Nevertheless, matrix inversion is not always possible in practice and due to bonds' varying levels of liquidity, different bonds as inputs result in different discounting functions (Sundaram & Das, 2010). Figure 2 shows the result of bootstrapping the yield curve on March 4, 2013 using Euro vs. Euribor six-months fixed swap rates from Datastream (identifier code: ICEIB). In this example, k equals 16, namely the two- to 10-, 12-, 15-, 20-, 25-, 30-, 40- and 50-year maturity swap contracts. M is a vector of 16 unity prices since the swap rates can be treated as coupon rates of fixed par bonds that pay EUR 1 at maturity. C contains 16x50 entries corresponding to the annual coupon/swap payments. I use Matlab to derive the implied zero swap rates and to specify the spline method for interpolation.

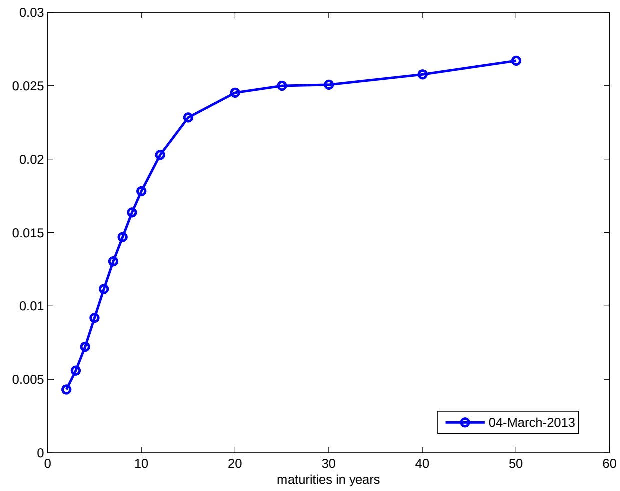


Figure 2: **Exemplary yield curve on 04-March-2013**
The zero yields are bootstrapped from Euro interest rate swap rates of maturities from two to 10, 12, 15, 20, 25, 30, 40 and 50 years. Yields are expressed in percentage points and are continuously compounded. The interpolation is based on the spline method.

As it can be deduced from the bootstrapped yield curve in Figure 2, the most common shape of the yield curve is concave. However, temporarily, the yield curve may be inverted, almost flat or more curved than usual (Diebold & Li, 2006). The yield curve also hints at the necessity to extrapolate since quoted swap rates get more rare as maturity increases. This is indicated by the only few scatters at the long end. Less trading leads the prices of these swaps to be more fluctuating meaning that the resulting yield curve used for discounting liability cash flows may not be sufficiently reliable. Usually, this is difficult to asses for swaps since they are traded over-the-counter (OTC) and not on a public exchange. Table 1 however provides an indication of the extent of trading in different maturity segments of swap contracts. The two trade summary measures are as of April 20, 2012 and obtained by Trioptima, a firm providing post-trade services, such as clearing trades, as third party in the OTC swap contracts.

Maturity in years	0-2	2-5	5-10	10-15	15-20	20-30	30+
Notional values in current billion US$	118,331	80,362	66,631	12,437	7,482	13,155	2,077
Trade Count	824,869	892,622	919,028	169,850	133,363	270,761	41,933

Table 1: **Global interest rate swap trade summary by maturity**
The notional values of interest rate swaps and the trade counts are measured as of the end of 20-April-2012. They are based on the weekly Rates Repository Report published by Trioptima. Submitting traders are the world's 14 largest banks and non-dealers such as asset management firms.

According to Trioptima (2012), the notional amount outstanding of interest rate swaps for maturities higher than 30 years is only 1.76% of the amount for zero- to two-year maturities. Interestingly, the 20- to 30-year maturity buckets seem to be more liquid than the two shorter-maturity buckets of 10- to 15- and 15- to 20-year swaps. Consequently, the swap rate volatility is likely to be lowest for the 5- to 10-year maturity swaps and highest for those with maturities longer than 30 years. The low trade volume for very long maturity swaps stresses the need for stable extrapolations of interest rates since derived interest rates from these observable swaps are likely to be too volatile.

2.2 Yield curve models

There are two general approaches to model the term structure of interest rate. The first approach is purely statistical and results in a static term structure. It involves estimation techniques to fit the current yield curve without considering the underlying economic factors driving it. The second approach, in contrast, sets up a dynamic model that explicitly describes the evolution of the yield curve.

2.2.1 Statistical models

This section outlines the most popular methods of the first approach and describes in more detail the widely-used functional form fitting as well as the more recent Smith-Wilson macroeconomic technique. Generally, the first step towards producing the current term structure is to identify traded market instruments from which to derive the unobservable zero spot rates. The derivation of these implied zero yields involves two main problems that in turn lead to the need for fitting. First, the prices of coupon bonds may be driven by factors other than the spot rates and hence may distort the implied zero-discount prices. For instance, institutional investors' preferences of specific maturity bonds, liquidity effects, tax considerations or embedded options induce bonds of the same maturity to lie clearly outside the price range. By using on-the-run issues that sell close at par, the outlier problem is well restricted (Sundaresan, 2009). However, this leaves a limited set of bonds and results in a lack of bonds maturing at each point in time. Thus, the fitting procedure needs not only to smooth the yield curve but also to interpolate it to cover the missing maturities. The techniques differ in the extent to which they put relatively more focus on either smoothing, or fitting as exactly as possible. Choosing a focus is necessary since there is always a trade-off between a smoother or more flexible yield curve fit (Anderson, 1996). Figure 3 shows a scatter of German federal bonds' yields to maturity that are outstanding as of April 15, 2013 and are provided by the Deutsche Bundesbank. Using Matlab I fit two curves to the yields showing a flexible and a smooth fit, respectively. In contrast to Figure 2 that shows a bootstrapped zero yield curve, Figure 3 plots a curve of quoted yield to maturities in order to illustrate two types of a statistical model's fitting approach.

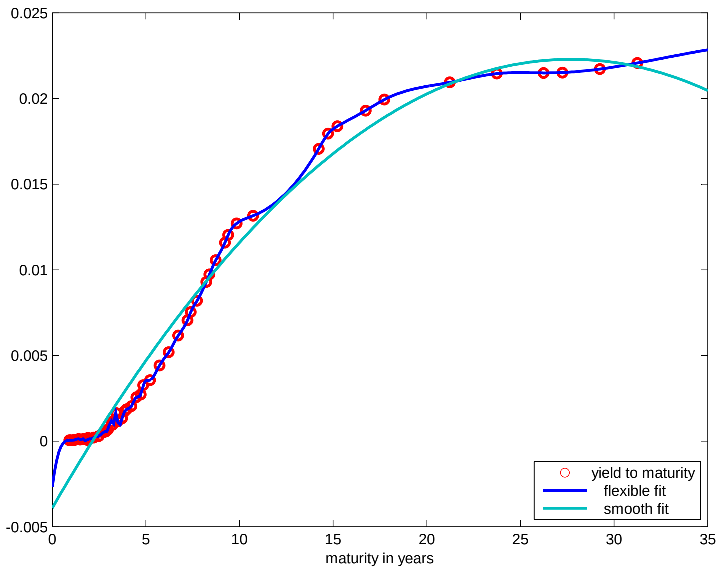


Figure 3: **German yields to maturity with different fits**
Yield to maturities are reported in percentage points and based on German federal bonds and notes outstanding on 15-April-2013. The observations total 51 after excluding inflation-indexed securities.

The most popular type of a flexible yield curve fitting is the spline method used for instance by the Federal Reserve Bank and the Bank of Japan (Sundaresan, 2009). First introduced by McCulloch (1971), the spline approach approximates the zero-coupon price function by specifying polynomial functions between the observed prices. In particular, the discount function is split into N intervals. Each interval exhibits a different function of time $f_N(t)$ which is knotted at each transfer point from one interval to the next one. All N-functions combined build the total discount function (Sundaram & Das, 2010). Most common are cubic splines that define each piecewise function as a third degree polynomial (Sundaresan, 2009). If $f_N(t)$ has a cubic form, then a regression analysis requires $N * 4$ parameters to be estimated (Sundaram & Das, 2010). The estimation of a large number of coefficients may become tedious, especially if the number of intervals N increases. Even though a higher N may mean that the shape of the current term structure can be captured in detail, it also risks overfitting the curve since outliers are not smoothed away. Shea (1984) already pointed out the trade-off between exact fit and overfitting when specifying the number of intervals and the location of division points for intervals. McCulloch (1975) suggests that the number of knot points between intervals is equal to the square root of the number of bonds and that the knots are equally spaced over the spanned maturity. Litzenberger and Rolfo (1984) suggest placing the knots at 1, 5 and 10 years which correspond to short, medium and long maturities respectively, and are usually well represented by liquid benchmark securities in the market. A problem in particular associated with cubic splines is that this method may produce unrealistic forward rate curves. Vasicek and Fong

(1982) introduced an approach that fits piecewise exponential functions. The advantage of exponential representations is that the long end of the forward curve flattens out so that unrealistically increasing rates at long maturities are eliminated. This feature corresponds with the expectation theory of the yield curve in which forward rates reflect expected short rates. Forward rates which constantly increase in the long horizon are based on the assumption that investors can distinguish expectations about interest rates far in the future - which is unrealistic (Anderson, 1996). Shea (1985), however, puts forward that exponential splines do not generally produce better fits than those of cubic splines.

An alternative method to that of splines circumvents the search for optimal knot points and erases the problem of overfitting by imposing a single instead of N parametric functions to fit the yield curve. The most popular method is the Nelson-Siegel approach (1987) extended by Svensson (1994). The latter is used in practice by most of the major central banks in the world including the European Central Bank and the Bank of England (Bank for International Settlements, 2005). Nelson and Siegel (1987) originally introduced the method in order to provide academics and practitioners with a more parsimonious functional form of the yield curve (Anderson, 1996). Thereby, they responded to Friedman (1977, p.411) who stated that 'students of statistical demand functions might find it more productive to examine how the whole term structure of yields can be described more compactly by a few parameters'.
The method specifies a function for the current forward curve so that $f(t)$ is the implied instantaneous forward rate for maturity t given as

$$f(t) = \beta_0 + \beta_1 \exp(-\frac{t}{\tau}) + \beta_2(\frac{t}{\tau} \exp(-\frac{t}{\tau}), \tag{7}$$

where the parameters $\beta_0, \beta_1, \beta_2$ and τ have to be estimated. The single function ensures a smooth curve rather than the spline method's saw-toothed one but is still sufficiently flexible to fit the generally observed shapes of yield curves (Sundaram & Das, 2010). In particular, it can fit different combinations of increasing, decreasing, horizontal, hill and U-shaped curves which makes it very flexible despite its parametrization (Sundaram & Das, 2010). Notwithstanding, the Nelson-Siegel model has been criticized for its non-assurance of being arbitrage-free (Björk & Christensen, 2001). It is a purely statistical model that may produce estimated yield curves that are inconsistent over time. Arbitrage opportunities arise when a model's expected future spot rates over t years are unequal to the yields of today's implied t-year zero bonds with a margin corresponding to Jensen's inequality (Diebold et al., 2005). Diebold and Li (2008) show how to make the dynamic version of the Nelson-Siegel model theoretically consistent by adding a term that adjusts the model to the no-arbitrage constraint. However, the correction term leads to a more complex estimation of the model and thus the model's advantage of being simple to implement disappears. In this line of reasoning, Coroneo et al (2011) study the extent to which the Nelson-Siegel curves accords with the no-arbitrage condition. They find that despite the lack of a theoretical arbitrage-free derivation, the Nelson-Siegel model produces dynamically consistent curves. Still, the result is driven by statistical properties and not by theoretical restrictions.
Nevertheless, the Nelson-Siegel model recently regained its prominence when Christensen et al. (2007)

showed that it underlies a latent three-factor model. Their result is in line with the principal component analysis of yields that identifies three common factors driving the yield curve (Litterman & Scheinkman, 1991). In particular, the three factors affect the level, the slope and the curvature of the yield curve. In order to emphasize the economic interpretation of the Nelson-Siegel function, Diebold and Li (2006) rewrite it as

$$R(t) = \beta_0 + \beta_1(1 - e^{-\frac{t}{\tau}})\frac{\tau}{t} + \beta_2[(1 - e^{-\frac{t}{\tau}})\frac{\tau}{t} - e^{-\frac{t}{\tau}}]. \tag{8}$$

Here, the β-coefficients correspond to the sensitivities of yields to the three driving components. The factor β_0 governs the level of the long-term interest rate. During the estimation, the level factor must be ensured to be positive because negative long-run rates are theoretically and practically unreasonable (Bolder & Stréliski, 1999). Furthermore, the sign of β_1 determines the basic shape of the yield curve. A positive value suggests an inverted yield curve, and a negative factor indicates a positive spread between short- and longer-maturity yields. A high β_2 produces a strong hump or a through in the yield curve determined by a positive or a negative sign, respectively. The sum of the level and the slope coefficients determines the short rate $r(0)$ in the limit (Annaert et al., 2013). The loadings to the three factors depend solely on the parameter τ. It is called the shape parameter since it specifies the position of the hump or the through. Analogously, the level parameter, the shape parameter must be positive (Bolder & Stréliski, 1999). A combination of the three components is able to produce essentially any observed yield curve shape at least in the short end of it.

Thus, compared to the spline method, the Nelson-Siegel approach has on the one hand an economically very intuitive backing but on the other hand produces a smoother curve. It is the former feature that makes it so popular among academics and practitioners; the latter, however, may create problems in fitting the yield curve at the long end (Gürkaynak, Sack, & Wright, 2007). Especially in the UK market, in which bonds are frequently traded up to very long maturities, the yield curve often exhibits a second hump at higher maturities. Problematic for the Nelson-Siegel function is to capture this second bump since the model approaches the asymptotic value too fast (Svensson, 1994). Gürkaynak et al. (2007) associate the additional curvature with the convexity effect on long-term bonds. Bondholders are willing to pay a premium for a higher-duration bond that gains more from the higher convexity when overall yields rise than a shorter-duration bond. Svensson (1994) improves the Nelson-Siegel fit to the entire term structure by adding a fourth term that allows for a second curvature in the forward curve. Written in the specification for the spot rate, it is

$$R(t) = \beta_0 + \beta_1(1 - e^{-\frac{t}{\tau_1}})\frac{\tau_1}{t} + \beta_2[(1 - e^{-\frac{t}{\tau_1}})\frac{\tau_1}{t} - e^{-\frac{t}{\tau_1}}] + \beta_3[(1 - e^{-\frac{t}{\tau_2}})\frac{\tau_2}{t} - e^{-\frac{t}{\tau_2}}]. \tag{9}$$

Compared to the Nelson-Siegel form, two additional parameters β_3 and τ_2 need to be estimated. They are interpreted in the same fashion as β_2 and τ_2, respectively, but for the second and not the first hump or through along the yield curve.

Overall, whether to use the Nelson-Siegel-(Svensson) approach or the spline fitting depends on the type of practitioners and the group of data. In particular, the NSS framework may be rather suited for macroeconomic analysis, whereas practitioners that rely on an exact fit to observed prices may turn to the very flexible spline. However, there is not a unified opinion about what method provides the better fit. For instance, Bliss (1991) shows that, even though less flexible, the parsimonious fitting functions perform better than cubic splines.

Despite the disagreement over the models' fitting performance, the NSS model clearly outperforms in terms of extrapolation of spot rates. Whereas the spline method is unable to extrapolate because it focuses on interpolation, the NSS single functional form is able to give estimated spot rates for any maturity. Given the estimates of the model's parameters, a yield of any maturity can be computed. As t gets larger, the forward and spot rates approach β_0 (see Equations 7, 8 and 9). While β_1 defines the general speed of convergence, the components' decays depend mainly on τ_1 for the Nelson method, and on τ_1 and τ_2 for the Svensson method. A high shape parameter value leads to faster flattening at the long end so that the extrapolated part will be close to the last liquid yield. Nelson and Siegel (1987) themselves put forward that when the shape parameter is estimated to be high, the yield curve is well fitted at short maturities, thereby flattening too quickly at the long end. They conclude that the curve approaches the long-run rate β_0 too rapidly. Their empirical test thus shows the tendency for the model to overestimate long-maturity yields when short yields are high and the curve downward sloping. Reversely, the model tends to underestimate long yields when the curve is upward sloping. This stands in opposition to the Vasicek's tendency to produce errors at the long end but, on average, it also should not produce systematic errors in one direction.

Studies about the Svensson method find it to create a different bias. The model's aim to capture the convexity effect at the long end can lead to unintuitive parameter estimates. Svensson (1994) finds that β_0 may turn out to be around 0 because the functional form assumes that forward rates continue to slope downward indefinitely. The problem arises when there is no perpetuity to use for estimating the infinite yield. This downward bias in the long-run yield, however, does not distort the fit over the observable yields but may create unrealistic extrapolations. Given Svensson's findings, I expect the Svensson's method's extrapolations to tend to underestimate the actual yields when the convexity effect already starts to be present at the last liquid point. Otherwise, the Svensson method cannot know that the effect exists with the consequence that it neglects it all together, and in turn is unlikely to produce unrealistically low extrapolations. Additionally, the estimated parameters may be unstable and may exhibit jumps that are not driven by actual jumps in the yields. According to Svensson (1994, p.24), 'yields were more sensitive to these parameter jumps at maturities beyond 20 years'. Thus, the Svensson model is likely to create more varying extrapolations than the Nelson model whose fast flattening reduces volatility. The erratic parameter estimation reported by Svensson (1994) induces me to expect the Svensson method to be more volatile than the Nelson one but both to be less stable than the Smith-Wilson method and more stable than the cruder Vasicek one.

A rather new method has arisen in the practitioners' world that combines the spline feature of close fitting and the macroeconomic approach of the Nelson-Siegel-(Svensson) method. In particular, the method named after Smith and Wilson (2001) provides an exact fit to the observed prices but at the long end approaches a predetermined constructed macroeconomic rate. In this way, the first part of the curve is the extreme of the splines method, and the second part is the extreme of the Nelson-Siegel-(Svensson) technique.

Smith and Wilson (2001) originally designed a yield curve technique tailored for a good extrapolation. The Smith-Wilson method belongs to the static modeling type and produces the discount function $P(t)$ as its output. Since here I describe the yield curve only at the current point in time, I use $P(t)$ for the earlier defined $P(0,T)$. In the Smith-Wilson method $P(t)$ is the sum of two discounting factors. First, it comprises the long-run discounting factor with the Ultimate Forward Rate (UFR) and second, it consists of a linear combination of so-called Wilson functions that result from the N observed zero-coupon bond prices with maturity m:

$$P(t) = e^{-UFR*t} + \sum_{j=1}^{N} \zeta_j * W(t, m_j). \tag{10}$$

The parameters ζ fit the current price discounting functions, and the Wilson function $W(t, m_j)$ is a symmetric function that requires a predetermined mean reversion α as input so that

$$W(t, m_j) = \begin{cases} e^{-UFR*(t+m_j)} * (\alpha * t - 0.5 * e^{-\alpha*m_j} * (e^{\alpha*t} - e^{-\alpha*t})) & t \leq m_j \\ e^{-UFR*(t+m_j)} * (\alpha * m_j - 0.5 * e^{-\alpha*t} * (e^{\alpha*m_j} - e^{-\alpha*m_j})) & t \geq m_j. \end{cases} \tag{11}$$

As the ζ_j are the only unknown parameters to the method, the Smith-Wilson technique proposes to set up a linear system of N discounting functions according to Equation 10, one for each zero-coupon bond. Using linear algebra, the N-dimensional vector $\zeta = (\zeta_1, \zeta_2, ..., \zeta_N)'$ is determined by solving

$$p = \mu + W\zeta \tag{12}$$

where

$$p = (P(m_1), P(m_2), ..., P(m_N))', \tag{13}$$

$$\mu = (e^{-UFR*m_1}, e^{-UFR*m_2}, ..., e^{-UFR*m_N})' \tag{14}$$

and

$$W = \begin{bmatrix} W(m_1, m_1) & W(m_1, m_2) & \dots & W(m_1, m_N) \\ W(m_2, m_1) & W(m_2, m_2) & \dots & W(m_2, m_N) \\ \vdots & \vdots & \vdots & \\ W(m_N, m_1) & \dots & \dots & W(m_N, m_N) \end{bmatrix}. \tag{15}$$

Having determined the vector ζ, it is possible to compute the price of a zero-coupon bond for all maturities with Equation 10. For instance, the price of a hypothetical zero bond with maturity of $t = 40$ is equal to

$$P(40) = e^{UFR*40} + w * \zeta \tag{16}$$

where $\zeta * w$ is a vector multiplication of the precalculated ζ and $w = (W(40, m_1), W(40, m_2), ..., W(40, m_N))$. When assessing the discount function for all maturities along the line of t=40 above, the construction of the complete continuously compounded yield curve results from solving the discount bond pricing equation for the zero yield as in Equation 2 for maturity t.

The EIOPA lays down that m_N=20, α=0.1 and UFR=4.2%. The first specification implies that the liquidity of market instruments with a maturity of beyond 20 years is insufficient for a reliable curve. Thus, a Smith-Wilson yield curve after Solvency II consists of the observed market yield curve until a maturity of 20 years, and after that of constructed rates. These rates approach a level of 4.2% with a speed of 0.1. For the latter specification - the convergence parameter α - the EIOPA relies upon the results by Thomas and Maré (2007) who find reasonable curves with a speed of 0.1 by using South-African yield curve data. The Ultimate Forward Rate specification results from the sum of an expected long-run inflation rate and a real interest rate estimated by historical data.

In general, the existence of a constant long-term interest rate is consistent with arbitrage-free pricing but cannot be confirmed by market data since coupon bonds and other market instruments used to derive the yield curve have finite maturities (Cairns, 2004). Estimates for the long-term spot rate over time show substantial variation in its level but given the estimates' high standard errors, the finding does not necessarily conflict with the theoretically constant long-run rate (Cairns, 2004). In particular, the Dybvig-Ingersoll-Ross Theorem states that in an arbitrage-free world the long-run rate is never decreasing (Cairns, 2004).

Antonio et al. (2009) summarize three main features inherent in the UFR. First, the UFR is a simple concept that does not require difficult estimation methods. Second, it ensures that changes in interest rates today do not affect the rate at infinite maturity; hence the UFR is stable. Third, the UFR creates consistency across economies since one single rate is applicable to all countries.

The EIOPA identifies four theoretical components of the UFR: The expected real interest rate, the expected long-term inflation rate, a nominal term premium, and a nominal convexity reduction. In principle, using historical data, each determinant can be calculated. However, the European Insurance and Occupational Pensions Authority (2010) chooses to incorporate only the first two components by arguing that the latter two may not produce reliable estimates. The expected real interest rate that reflects an aggregated average of annual real bond returns across various countries is set to 2.2% . Contrarily, the EIOPA decided to completely abstain from historic averages for the expected long-term inflation rate. The reason is that the inflation rate has globally shown a clear downward trend over the last years due to central bank inflation targeting. Hence, the EIOPA sets an annual 2% for all EU countries. The last two components - the convexity effect and the term premium - arise via the discussion of the expectations theory of the yield curve. In its original form, this theory postulates that the instantaneous forward rates are equal to the expected spot interest rates (Anderson, 1996), so that

$$f(t, s) = E[R(s)]. \tag{17}$$

The above relationship rarely holds due to the two effects mentioned above.

First, forward rates derived from observable coupon-bond prices are lower than what investors expect future short rates to be because of the convex relationship between bonds and yields. Bondholders accept a lower yield on coupon-bearing long-term bonds due to the insurance given by the convexity gain when the interest rate level increases. The convexity effect, also referred to as Jensen's Inequality, suggests that implied forward rates are lower than actual expectations of future short rates. As Anderson (1996) points out, this negative premium is usually insignificant and thus the EIOPA chose to neglect it for the determination of the UFR. Otherwise, the EIOPA runs the risk of overestimating the true unconditional forward rate.

Second, risk-averse investors demand a positive term premium for the uncertainty of future interest rates. An investor usually prefers to invest in adjacent short-maturity bonds rather than in a single long-maturity bond since the first strategy ensures a higher liquidity. Whereas Antonio et al. (2009) estimate the term premium for the UFR to be 1.5%, the EIOPA decides to omit it. Still, a consulting group of experts to the EIOPA - the CRO Forum - fears that the UFR may still be too high. Rebel (2012) summarizes the CRO Forum's critique on the Smith-Wilson method by pointing out its three main shortcomings. First, setting a fixed last liquid point drives controversies among experts about liquid instruments in the market. Second, the Smith-Wilson method incorporates a clear distinction between the observed and the extrapolated part of the yield curve by basing the first part completely on market inputs, and the second part solely on the fixed UFR. Third, the distribution of interest rate risk along the Smith-Wilson yield curve is centered around the last liquid point, inducing a counterintuitive need for hedging. Thomas and Maré (2007) show that the distribution of interest rate sensitivity of the Nelson-Siegel method is spread evenly along the curve, whereas with the Smith-Wilson method the entire risk lies in 20- (LLP) and 19-year interest rates. The high exposure to only two maturity rates compared to an even exposure to all rates implies a more frequent rebalancing of hedge portfolios with the Smith-Wilson method than with the Nelson-Siegel method. Any changes in the 20- and 19-year interest rates immediately translate into the extrapolated interest rates requiring altered hedging positions for all maturities beyond 20 years. Even the EIOPA admits that the setting of the convergence speed is rather subjective and arbitrary, thereby calling for future work on finding objective criteria for the determination of α.

Furthermore, the Smith-Wilson method is not constrained to only produce positive values for the discount function in the extrapolated part (European Insurance and Occupational Pensions Authority, 2010). The problem may occur when the convergence is low and the LLP is higher than the sum of α and the UFR. Similar problems do not exist with other parametric models such as the Nelson-Siegel technique which models the forward rate directly and hence, by definition, cannot produce negative discount rates. Although the EIOPA pushes forward to introduce a complete market valuation of assets and liabilities by proposing the Smith-Wilson method, it withdraws from a fair valuation of long-term liabilities by setting the regulatory UFR (Chief Risk Officer Forum, 2010). The fact that the UFR for Japan is distinct from all other countries gives rise to the possibility of changing the UFR in the future if countries happen to experience a sustained low interest rate environment. However, Solvency II does not specify the procedure of deciding and changing the UFR in the future. Notwithstanding, the new directive implies a constant risk of possible regulatory changes in the UFR that are able to trigger immense adjustment of insurance

companies' capital requirements. Whereas insurance companies are able to hedge against the interest rate risk in the market, they cannot hedge against regulatory changes in the UFR. Additionally, the predetermined level of 4.2% for the UFR is inconsistent with the market-value proposition of Solvency II in periods in which the value of long-term assets drops due to disrupted capital markets but the present value of long-dated liabilities stays unrealistically low due to discounting at the much higher UFR. Consequently, the UFR level creates a distorted view on the actual funding ratios. Not only the level of the UFR is controversial but also its composition. Industry experts argue that the UFR should not be based on the central bank's inflation target but be rather specific to each country's price developments driven by wage increases (Chief Risk Officer Forum, 2010).
Next to the concept of the UFR, the consultation group has also heavily criticized the EIOPA's fixing of the LLP at 20 years. The experts argue that the Euro interest rate market is still liquid beyond 20 years by pointing out that the amounts of outstanding European AAA government bonds to be even higher for those with 20- to 30-year maturities than for those with 10- to 20-year maturities. Furthermore, they say that trading volumes of Euro swaps with 20- to 30-year maturities are as high as those with 10 to 20 years. Nonetheless, the amounts of outstanding bonds and swaps with a maturity beyond 30 years shrink to only 10% of those with maturity up to 30 years, so that a LLP at the 30-year maturity is better justified. Furthermore, all institutions will hedge their long-dated liabilities with instruments with 20-year maturity because it will be the last point to be subject to interest rate risk. Consequently, the LLP will ironically turn into one of the most liquid points (Chief Risk Officer Forum, 2010). Additionally, since the liquid part of the curve perfectly fits observed rates, the Smith-Wilson curve may happen to be bumpy which is an issue previously discussed with the spline method. According to the HBS.8.11 regulation, the EIOPA gives the option to choose a different convergence period. Instead of using the original 60-years period, insurers can apply one extended to 90 years after the LLP or one shortened to 10 years after the LLP. The option is contradictory to Solvency' II objective of unifying the approach across European insurers. In case the EIOPA still wants to provide an option, this option should be based on the United States' example introduced in July 2012, so that European and American funding ratios are comparable (Academic Community Group, 2012). Overall, the controversies arise especially due to the fact that the Smith-Wilson method requires subjective decisions regarding its parameters contrarily to all other methods including the Nelson-Siegel method because the latter always relies on estimated parameters that can never be subject to one's own discretion.

In response to the critique put forward by the CRO Forum on the proposed Smith-Wilson method for extrapolating the yield curve under Solvency II, the Dutch financial services provider Cardano has proposed an altered version of it, which led the Dutch government to introduce an update to their regulation for Dutch insurers and pension funds (Rebel, 2012). The alternative Smith-Wilson method differs only in the way the extrapolation works. Whereas the original Smith-Wilson method extrapolates the curve from the LLP to the UFR, the alternative version constructs a weighted average between the market data and the UFR beyond the LLP (De Nederlandsche Bank, 2012). This way, the method does not assume a clear cut-off point between a liquid, market-based curve and a UFR curve but slowly decreases the weight put to the more unreliable market data. Hence, for $1 \leq t \leq 20$, the two methods are the same but differ for $21 \leq t$ in the way that the original model ignores any market rates, whereas the Cardano version employs a mixture of market and artificial rates.

Despite the criticism, the Smith-Wilson approach exerts some favorable features, which led the EIOPA choose it. One such property is that the Smith-Wilson method ensures an exact fit to the current yield curve which neither other static models nor equilibrium interest rate models guarantee. In particular, the Nelson-Siegel method, even though providing a very close fit by minimizing the difference between model and input data, may over- or underestimate observed yields. Furthermore, the least square optimization often results in instable parameters over time. The Smith-Wilson method is solved analytically and is thus never subject to parameter jumps. Whereas least square methods fit the yield curve reasonably well, equilibrium models have been criticized to lack cross-sectional fit (Hull & White, 1990). In this sense, by ensuring no-arbitrage, the Smith-Wilson method is preferable to the Nelson-Siegel method and equilibrium models (Thomas & Maré, 2007).

Furthermore, the Smith-Wilson method performs no smoothing of input data. This way, it circumvents the controversial trade-off between fit and flexibility, associated with regression spline methods, by only focusing on the fitting aspect. Additionally, the macroeconomic approach inherent in the Smith-Wilson method ensures stable long-run forward rates. This may be a desirable feature since pricing errors for long-maturity input market prices become irrelevant and cannot be responsible for spurious volatility at the long end (Antonio et al., 2009). The asymptotic behavior at the long end is in line with the expectations theory of the yield curve, since at points far in the future investors cannot differentiate anymore between, for instance, a forward rate in 30 and 35 years (Anderson, 1996). Lastly, calculations of the yield curve can be performed in Excel making the Smith-Wilson technique employable in the entire industry.

Due to the sharp criticism of practitioners, I expect that the Wilson method to produce the highest rates in the extrapolated curve among all models thereby lying consistently above the actual curve. The upward bias is likely to be particularly strong during the last years of the European Central Bank's aggressive monetary easing and consequently low long-term interest rates. Furthermore, the fixing of the long rate at one level induces the method to be least volatile, thereby giving it its notation of method particularly suited for extrapolations.

2.2.2 Economic models

The second type of modeling of the yield curve focuses on describing its dynamics instead of providing a fit. Every term structure model is a system composed of a set of state variables that evolve over time. The simplest models have one state variable, namely the short rate $r(t)$. Its dynamics consist of a deterministic and a random part (James & Webber, 2000). Based on empirical evidence, the short rate preferably inherits two properties in the model. First, it mean-reverts and second, it is more volatile than longer rates (Sundaresan, 2009). Thus, the deterministic component reflects the feature of the short rate to revert to a long-term level μ at rate α, whereas the random component reflects its volatility as a function of a Brownian motion $W(t)$. The stochastic differential equation for the short rate looks as follows:

$$dr(t) = \alpha(\mu - r(t))dt + \sigma dW(t). \tag{18}$$

Equation 18 describes the evolution of the instantaneous interest rate introduced by Vasicek (1977). The stochastic differential equation is a so-called mean-reverting Ornstein-Uhlenbeck process, and has the following solution:

$$r(t) = \mu + (r(0) - \mu)e^{-\alpha t} + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dW_s. \tag{19}$$

Equation 19 implies that given today's rate, the expected short rate equals the long-run rate μ plus an adjustment term. The conditional expectation is

$$E[r(t)|r(0)] = \mu + (r(0) - \mu)e^{-\alpha t}. \tag{20}$$

The higher the speed of convergence and the smaller the difference between today's rate and the long-run rate, the closer is the expected rate to the long-run mean. The conditional variance is

$$Var[r(t)|r(0)] = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}). \tag{21}$$

The lower the speed of convergence and the higher the variance parameter, the more is the short rate expected to vary in the future.

Despite these appealing properties, one drawback of the model is that the short rate, by following a Gaussian normal distribution, may assume negative values, which is unrealistic. Nevertheless, Anderson (1996) argues that for plausible parameters for the convergence α, the long rate μ, and the volatility σ the drawback is negligible. Others turn to the one-factor equilibrium model of Cox and Ross (1985), which ensures non-negative rates by making the random term of the short rate dynamics dependent on the square root of the rate. Neither of the models, however, fits the current yield curve exactly, since they are both limited in the shapes they can take on (James & Webber, 2000). The reason for this is that the models are both constrained to produce rates that are equal to the short rate at the first maturity on the current yield curve, and to μ at the curve's long end. Even though fitting the models to the current yield curve prevents curve overfitting and thus a stable term structure over time, they may nevertheless turn out to be off the real observed yield curve due to their limited flexibility (Anderson, 1996). Not having an exact fit leads to arbitrage possibilities that are an undesirable feature for a financial economics model. Still, the Vasicek model is popular due to its simplicity and tractability (Sundaram & Das, 2010).

In order to price a zero-coupon bond with the Vasicek model, the dynamics of the short rate under the risk-neutral measure Q have to be specified. Expressing the dynamics of the short/instantaneous rate from Equation 18 under the risk-neutral measure Q yields the rate's risk-adjusted dynamics:

$$dr(t) = \alpha(\tilde{\mu} - r(t))dt + \sigma d\tilde{W}(t), \tag{22}$$

where $\tilde{\mu} = \mu - \frac{\lambda\sigma}{\alpha}$. The parameter λ represents the market price of risk, for which investors want to be compensated when buying interest rate products that depend on the stochastic - hence risky- short rate (James & Webber, 2000). $\tilde{W}(t)$ is the Browninan motion under the risk-adjusted measure. According to arbitrage-free pricing, a zero-coupon bond price is the expected discounted payoff at maturity under the objective measure. It follows that

$$P(t,T) = E_Q[1 * \exp(-\int_t^T r_s ds)]. \tag{23}$$

The zero bond price solution of the Vasicek model to Equation 23 belongs to the so-called affine term structure models. According to Duffie and Kan (1996), affine models have a unique closed form solution for zero bonds in the risk-neutral world, expressed as

$$P(t,T,r(t)) = A(t,T)e^{-B(t,T)X(t)}, \tag{24}$$

where $X(t)$ is a group of state variables that are the drivers of the yield curve process. Duffie and Kan (1996) show that the set of $X(t)$ can be represented by a set of zero yields of different maturities. In the case of the Vasicek model, the single state variable is the zero yield of the shortest possible maturity, or the short rate. Thus, in Equation 24, $r(t)$ replaces $X(t)$. The Vasicek model's affine solution for the term structure follows as

$$R(t,T,r(t)) = -\frac{1}{T-t}[ln(A(t,T)) - B(t,T)r(t)], \tag{25}$$

where

$$B(t,T) = \frac{1-e^{\alpha t}}{\alpha} \tag{26}$$

and

$$A(t,T) = \exp\left[\frac{(B(t,T)-(T-t))(\alpha^2\mu - \frac{\sigma^2}{2})}{\alpha^2} - \frac{\sigma^2 B(t,T)^2}{4\alpha}\right]. \tag{27}$$

In order to fit the Vasicek model to the current term structure, the parameters μ, α, σ, and $r(0)$ have to be determined. Determining the parameters of an interest rate model by using market data is known as calibration. Essentially, there are two types of calibration. First, the model can be calibrated using currently observed yields as inputs. Thus, cross-sectional market data is needed - namely, a set of yields of different maturities, all available at the same point in time. Second, a historical estimation is possible, which requires a time series of the short rate as data input. In contrast to the first type of calibration, the latter involves a constant-maturity yield over different points in time (James & Webber, 2000). Commonly, the model is fitted to current yields in order to deduce an indication of the expected direction of future interest rates, and not to just derive what the previous yield's realizations were (James & Webber, 2000). However, in illiquid markets in which current prices are erroneous, or the cross-section is very thin, historical fitting may be the only option. Additionally, in order to derive an estimate of the unit price of risk, both calibrations are necessary. The reason for this is that current prices are observed under subjective probabilities, whereas historical prices, which do not include expectations, fall under the objective measure (James & Webber, 2000). When comparing Equation 18 and Equation 22, it is clear that time series estimations are needed to find μ, and cross-sectional ones to find $\tilde{\mu}$. Then, one can solve for λ equaling $(\mu - \tilde{\mu})\frac{\alpha}{\sigma}$.

Since this research is concerned with recovering the current yield curve rather than using the Vasicek model to price market instruments, I focus on cross-sectional fitting exclusively. Analogous to fitting the Nelson-Siegel model to current price data, the cross-sectional calibration of the Vasicek model involves finding the parameters $r(0)$, α, μ, and σ that minimize the sum of squared deviations between the market zero yields and those implied by Equation 25.

With four free parameters, the model can exactly fit only four observed bond prices (Sundaram & Das, 2010). Hence, the number of different shapes that the model can produce is limited. The class of models that produce current term structures that do not completely coincide with the observed yield curve are classified as equilibrium models. Due to a growing need to have models, producing yields that are in line with observed ones, a new class has developed - the so-called arbitrage-free models (Sundaram & Das, 2010). In order to provide an exact fit, these models incorporate as many free parameters as there are initial price inputs by making the drift of the short-rate process dependent on time (Sundaram & Das, 2010). In this way, compared to the equilibrium model whose drift is deterministic, the arbitrage-free models are more flexible (Hull, 2009). Through the introduction of time dependence, the Vasicek model can be converted into an arbitrage-free model known as the extended Vasicek model after Hull and White (1990). Other members of the class of one-factor arbitrage-free models are the models developed by Ho and Lee (1986), by Black, Derman, and Toy (1990), and by Black and Karasinski (1991). Nevertheless, the desirable arbitrage-free feature may simultaneously be the class' drawback, since the flexibility implies that the price of risk may change implausibly from one day's calibration to the next one (Sundaram & Das, 2010). Also, the feature of exact fitting trades off with a simpler, and more tractable equilibrium model. Consequently, the choice of a model depends on its use. According to James and Webber (2000), in general, models should be selected based on three features. First, they should fit the current yield curve reasonably well. Second, they should have good dynamics, and third, they should provide computational ease (James & Webber, 2000). In the context of this thesis, the original form of the Vasicek model, and not its extension, is the preferred choice, and there are two reasons for that. First, the model offers a tractable implementation, and is one of the first economic models for the term structure. Thus, Vasicek's resulting current term structure focuses on the most fundamental economic driver, namely the short rate, which is even found to explain 98% of changes in yields (Litterman & Scheinkman, 1991). Hence, it is a suitable basis for evaluating the Smith-Wilson yield curve, which stresses practicalities instead. Second, an arbitrage-free model would not be a good choice for a comparison to other models because, by definition, its calibration to market data results in an exact fit, no matter its underlying drivers.

Contrarily, according to Hull (2009, p.676), the Vasicek yield curve can only be 'upward-sloping, downward-sloping or slightly humped'. In particular, according to Equation 25, the current Vasicek yield curve starts at $r(0)$ and as $T \to \infty$, the long-run yield $R(\infty)$ approaches $\mu - \frac{\sigma^2}{2\alpha^2}$. Consequently, the yield curve is monotonically increasing when $r(0) \leq R(\infty) - \frac{\sigma^2}{4\alpha^2}$, decreasing if the short rate is below the long-run mean μ, and humped if it lies exactly between the first two conditions (Vasicek, 1977).

Hence, if the current yield curve starts out high, I expect the Vasicek curve to monotonically decrease, thereby producing extrapolated yields that underestimate the actual yields. Reversely, low short yields suggest an upward sloping total yield curve and are likely induce the monotonically rising extrapolated Vasicek yields to overshoot the actual yields. If there is no tendency for the actual short yields to systematically start out high or low, the Vasicek model is, on average, unlikely to be biased. Furthermore, the clear-cut conditions regarding the Vasicek yield curve's shape may induce the curve to abruptly change, once the short rate trespasses one of these conditions. Thus, I expect the Vasicek model to be most volatile.

After the analysis of the two statistical models (Nelson and Svensson), the macroeconomic Wilson model, and the economic Vasicek model, the following question arises: Which of the models best fulfills the extrapolation criterion of reducing the volatility at the long end?

By construction, the Smith-Wilson method ensures its extrapolated yields to approach the same UFR every day, thereby giving extrapolations least flexibility. Whereas all other three methods are parametric, the Smith-Wilson one is purely technical and is never subject to parameter instability. In contrast, the Vasicek model's most complex parametrized closed form solution for the yield curve is most exposed to it. Based on that, I make the following proposition:

Proposition 1. The Smith-Wilson method reduces volatility the most, and the Vasicek model the least.

Even though stability is the main objective of using extrapolations, it is not everything. It should not come at the cost of producing biased long ends of the yield curve. Thus, which of the four models produces extrapolations that reflect the current interest rate environment, thereby providing an unbiased estimate of the actual long-term yields?

Given the Nelson model's theoretical interpretation and its popularity, it is likely to fit current yields well. The same applies to the Svensson model, but by adding another factor, the Svensson model is particularly suited to fit observed yields of longer maturities. Given that I fit the models only to yields up to maturity of 20 years and not beyond, the Nelson model is likely to best fit the curve and to produce extrapolations that are closest to the actual yields. Due to the nature of the Vasicek model of being an equilibrium model, it does not perfectly fit yields and only produces limited yield curve shapes that are likely to translate into extrapolations that are further away from the actual yields. Given the predetermined UFR, the Wilson method is least likely to match actual long-term yields, since it is least able to adjust to different interest rate environments. Thus, it is proposed:

Proposition 2. The Nelson-Siegel method produces the least biased extrapolations and the Smith-Wilson method the most biased ones.

3 Data and Methodology

Figure 4 outlines the procedure for evaluating the four different models' extrapolations of the yield curve. The first stage derives the zero yield curve that covers maturities from one to 50 years by bootstrapping swap rates of maturities from one to 10, 12, 15, 20, 25, 30, 40 and 50 years. The resulting yield curve is divided into a liquid part that spans over maturities from one to 20 years, and an illiquid part covering 21- to 50-year yields. The second stage uses the liquid rates as market inputs for the calibration of the Vasicek, the Nelson-Siegel, the Nelson-Siegel-Svensson, and the Smith-Wilson models. Given the estimated parameters, each model extrapolates the yield curve for maturities from 21 to 50 years in the third stage. These extrapolations are compared with the illiquid actual curve derived in stage one on the basis of volatility and fit.

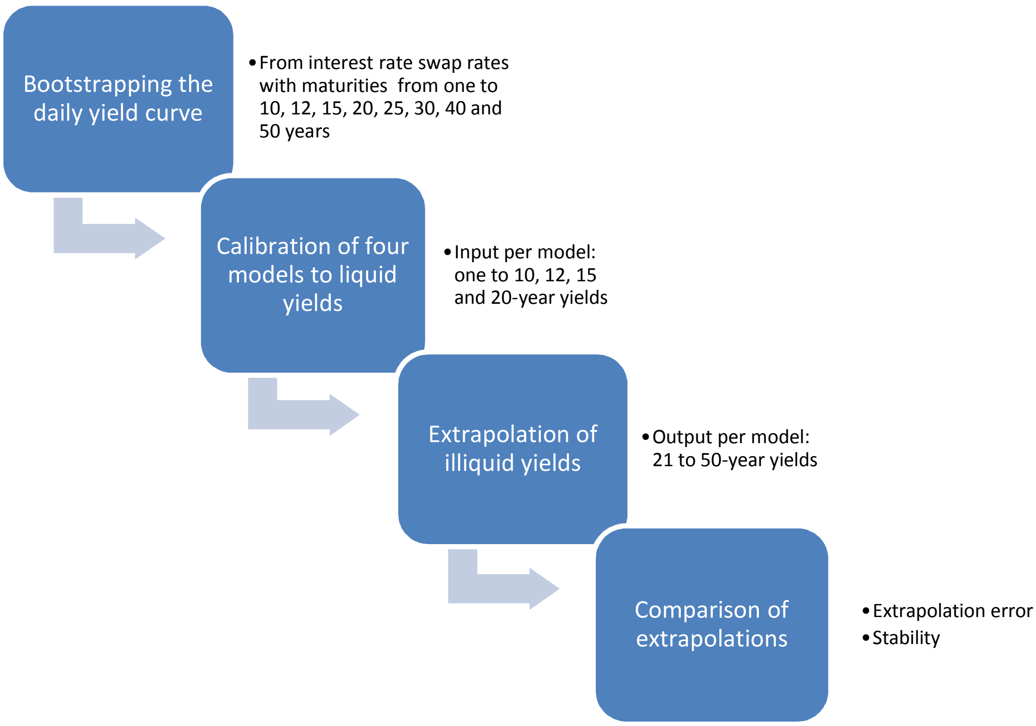


Figure 4: **Methodology in a nutshell**

3.1 Data

I use daily fixed-for-floating Euro swaps from Datastream to bootstrap the yield curve (identifier code: ICEIB). The fixed rates are midpoints and quoted on an European annual 30/360 basis. The swap rates for maturity of one year are annualized percentages and referenced to the three-month Euribor, whereas those for maturities of two to 10, 12, 15, 20, 25, 30, 40 and 50 years have the six-month Euribor as floating leg. The dataset spans from January 4, 1999 to April 17, 2013 and covers all working days excluding holidays. In total, there are 3728 days. Quotes for swap rates of 40- and 50-year maturities are missing for the days before and including February 28, 2005. Only after that did long maturities start to be contracted and quotes became available. Consequently, my dataset consists of daily swap rates of maturities from one to 10, 12, 15, 20, 25 and 30 years for the first 1606 days, and of additional 40- and 50-year swap rates for the subsequently remaining 2122 days. Thus, in total, I have 60164 (15*1606+17*2122) swap rates observations.

In an attempt to quantify the liquidity of swaps which make up my dataset, I first consider bid and offer rates before turning to the above mentioned mid-rates that form the basis for the zero yield curves. The bid-ask rates are retrieved from Datastream along with the mid-rates. All quotes have been gathered by ICAP, a global broker. Column 1 and 2 of Table 2 show the mean and the standard deviation of the daily differences between offer and bid rates over the period of January 4, 1999 to April 17, 2013. Column 3 reports the standard deviation of the spread's change at each maturity.

Maturity in years	Mean ($*10^{-2}$)	StdDev. ($*10^{-2}$)	StdDev. of the change ($*10^{-2}$)
1	0.0286	0.0093	0.0087
2	0.0295	0.0089	0.0072
3	0.0297	0.0092	0.0082
4	0.0293	0.0089	0.0063
5	0.0294	0.0088	0.0062
6	0.0292	0.0090	0.0064
7	0.0291	0.0092	0.0067
8	0.0295	0.0089	0.0066
9	0.0297	0.0085	0.0063
10	0.0295	0.0089	0.0065
12	0.0325	0.0073	0.0072
15	0.0328	0.0073	0.0074
20	0.0325	0.0073	0.0070
25	0.0326	0.0078	0.0081
30	0.0325	0.0072	0.0069
40	0.0335	0.0085	0.0089
50	0.0338	0.0082	0.0086

Table 2: **Measuring the liquidity of the Euro swap rate curve**
Column 1 and 2 report mean and standard deviation of the daily swap rate spreads. The spread is calculated as the difference between the ask and the bid swap rate in percentage points. Column 3 shows the standard deviation of the spread's change in percentage points. It is computed as the spread's first difference. Maturities 25 and 30 are reported for the period from 04-Jan-1999 to 17-April-2013. Maturities 40 and 50 apply to the shorter sample, 01-March-2005 to 17-April-2013.

Overall, the magnitudes of the mean spreads are relatively small reaching from 2.9 basis points for one-year maturity swaps to 3.4 basis points for the 50-year swap. According to the measure at hand, the Euro swap market seems to have sufficient liquidity thereby limiting the extent to which it can be volatile. Even though this might not give a strong case in favor of extrapolations, there is still an apparent tendency of higher maturity swaps to have a higher bid-offer spread.

In line with the results by Trioptima (see Table 1), the 20- and 30-year mean spreads are lower than the 12- and 15-year mean spreads, substantiating a supposed higher liquidity for the first group of swaps. By the same token, the 30-year spread exhibits the lowest standard deviation. Liquidity first increases with maturity but decreases at the very long maturities with the exception of the 30-year maturity. The standard deviation of the spread suggests a pattern that may seem counter-intuitive: The shorter-maturity spreads seem to be more volatile than the longer-maturity ones with the one-year maturity swap being the most illiquid one. That is why the EIOPA's consultation group questions the reliability of the quote dispersion's measure and calls for careful interpretation of it (CFO Forum and CRO Forum, 2010).

In contrast, the standard deviation of the spread's daily change is in line with the expected lower liquidity of the long-maturity swaps. According to this volatility measure, the 40-year swap manifests the lowest stability with a daily standard deviation of 0.89 basis points. The 30-year swap continues to seem especially frequently traded as the standard deviation of its spread's change is relative low compared to the standard deviations of the changes in spreads of other long-maturity swaps.

During stressed financial market conditions, daily variation in the swap rates' spreads are remarkably higher than during normal periods. Figure 5 shows the daily change in the swap rates' spreads over the period from 2005 till 2013 for the short 5-year and the long 50-year maturity. There is almost no variation until the end of 2008 in either spread but with the start of the financial crisis the graph depicts strong fluctuations. Whereas the changes in the spreads are seemingly equal between the two maturities during the pre-crisis period, after 2008 the change in the spread of the 50-year maturity swap is more volatile than that of the 5-year maturity swap. According to the differential reaction during stressed periods, extrapolations suggest to be especially important during financial turmoil. It is not surprising that the call for a new European insurance regulatory framework has been triggered by the financial crisis.

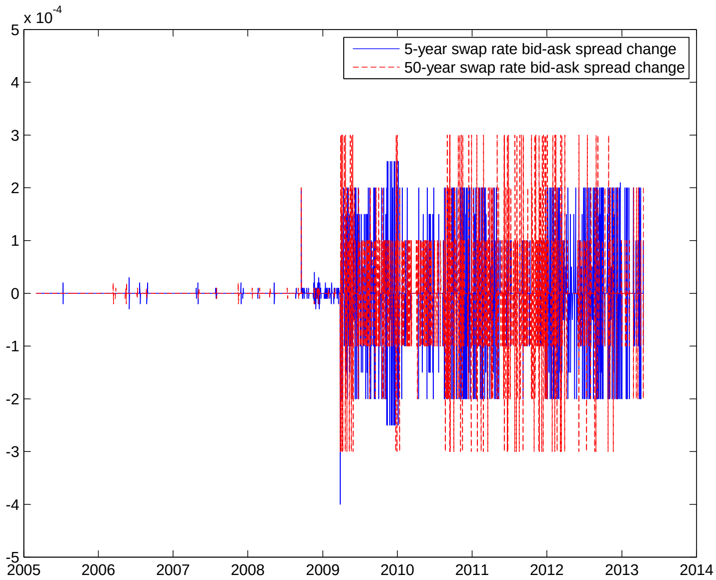


Figure 5: **Liquidity decreases more strongly during stressed market periods for longer maturity swaps** A time series of the swap rate spread's first difference at maturity 5 and 50. The spread is calculated as the difference between the ask and the bid swap rate in percentage points over the period from 01-March-2005 to 17-April-2013.

3.2 First step - Bootstrapping

From the mid-rate swaps, I derive the zero yield curve in Matlab (function: IRDataCurve.bootstrap). The curve consists of continuously compounded spot rates based on 30/360 day-counts as this is the convention for Euro-denominated swaps (Neftci, 2008). In particular, define S_t as the fixed swap rate with maturity t in years, notional value of EUR 1, and $R(t)$ as the zero coupon spot rate with maturity t in years. The relation between zero rates and fixed swap rates is given by

$$1 = \sum_{i=1}^{t-1} \frac{S_t}{e^{R(i)i}} + \frac{1+S_t}{e^{R(t)t}} \tag{28}$$

Equation 28 reminds of the pricing equation for a par yield bond with annual S_t-coupons and face value of EUR 1. Given that $S_1 = R(1)$ and knowing S_2, one can solve for $R(2)$. $R(3)$ is then found recursively by solving Equation 28 for $R(3)$ after inserting S_3, $R(1)$, and the value for $R(2)$, which has been found in the previous step. All other zero rates are calculated similarly. The missing fixed swap rates beyond the 10-year maturity, e.g. between the 12- and 15-year maturities, are derived by linear interpolation [1].

Table 3 summarizes the distributional characteristics of the bootstrapped yields. It reflects the so-called stylized facts of yields in regard to the yield curve's shape, volatility, and dynamics, which have been already studied extensively in the literature, i.e. by Diebold and Li (2006). On average, yield curves have been found to be upward-sloping. This holds true for the average yield curve from maturity one to 25 (see column 1 in Table 3). Furthermore, this average yield curve also offers the common concavity feature since, for instance, the increase in the yield from maturity 1 to maturity 5 is by about 70 basis points higher than that from maturity 20 to 25. However, the curve is downward-sloping at the very long end. The reason for the strong decrease for maturities 40 and 50 is partly due to the shorter and more recent sample period within which yields were generally lower. Still, when calculating the average yields of maturity one to 30 using the restricted sample period, the average yield curve continues to slope downward over the 30- to 50-year maturity range (not reported here). This feature has been extensively studied by Phoa (1997) who assigns the convexity effect to it. Long-term interest rate instruments benefit more from convexity than instruments with short maturities, so that the first outperform the latter when yields rise. Hence, investors are willing to pay more for the 30- to 50-year maturity swaps, thereby pushing their yields down. The gain is more pronounced whenever yields are more volatile because the chance of the convexity gain to materialize increases.

Besides the usual upward slope, the yield curve can take on all kinds of different shapes. Figure 6 shows the development of the one-, 10-, 20-, and 30-year rates over time. Whereas in the first half of the period the one-year yield lies below the higher-maturity rates, in the end of 2008 it increases until it rises above all other rates. Here, for instance, the shape of the yield curve inverts.

Another stylized fact is that 'the short end of the yield curve is more volatile than the long end' (Diebold & Li, 2006, p.343). Table 3 highlights that tendency in terms of the standard deviation of yields. However, the standard deviation of the yields may be a suboptimal measure since it does not take into account the quite strong differences in yield levels across maturities. Consequently, column 6 depicts the volatility of the yield's daily change in percentage points. Here, the picture is much more in line with the liquidity story, namely that the long end of the yield curve is less liquid and hence more volatile. Whereas the standard deviation of the daily change in the one-year yield is 2.8 basis points, it grows to about 5.3 basis points for the 50-year yield's change.

Generally, yields of all maturities feature a high persistence, as can be seen from the autocorrelations. After 60 days, the effect is strongest for the 10-year yield and decreasing towards both the one- and 50-year rates. The least persistent rates are of term 40 and 50. This stands in contrast to the stylized fact documented by Diebold and Li (2006) who argue that short-term yields are less persistent than long-run

[1] Alternatively, I used the spline interpolation method but the difference in derived rates is very small so I stick to the simpler linear interpolation.

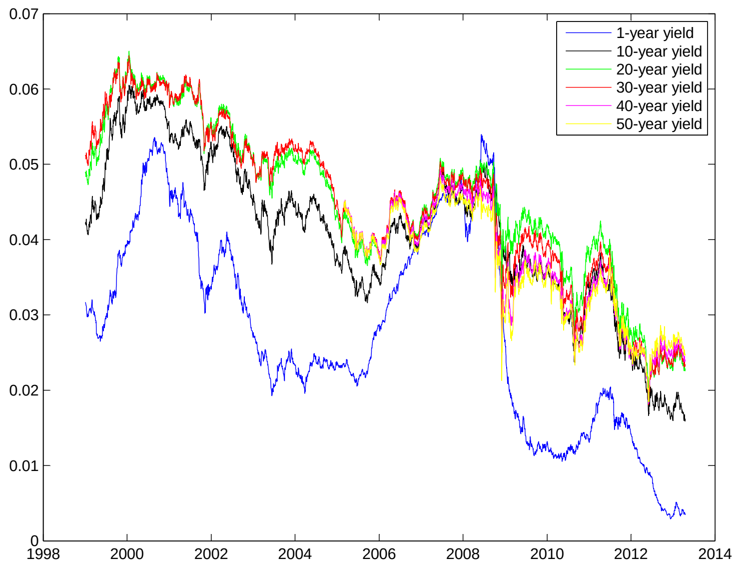


Figure 6: **The yield curve exhibits different shapes over time**
The zero yields are bootstrapped from Euro swap rates, continuously compounded and reported in percentage points. Maturities one to 30 are reported for the period from 04-Jan-1999 to 17-April-2013. Maturities 40 and 50 apply to the shorter sample, 01-March-2005 to 17-April-2013.

rates.

3.3 Second step - Calibration

Using the bootstrapped zero yields with maturity one to 10, 12, 15, and 20 years, I calibrate the four models to fit every day's liquid part of the yield curve. Except for the Smith-Wilson method, I use non-linear least square optimization (Matlab's function 'lsqnonlin') to calibrate the models.

In the case of the Vasicek model, there are four econometric techniques to derive its parameters. Initial approaches to estimate historical parameters proceeded in firstly discretizing the stochastic differential equation of the short rate (see Equation 18), and then performing an Ordinary Least Square regression to find μ, α and σ (James & Webber, 2000). However, the discretization results in an autoregressive model of first order, which is found likely to have a unit root. The inferred parameters are thus invalid (James & Webber, 2000).

Consequently, more sophisticated estimation methods are necessary that do not infer the parameters from the differential equation but rather from the closed yield curve solution (see Equation 25). Alternatives to OLS estimation are the Generalized Method of Moments (GMM) and the Maximum Likelihood Estimation (MLE) (James & Webber, 2000). According to Stanton (Stanton, 1997, p.1976), GMM is employed when MLE is 'too complicated or time-consuming or when we want to specify only certain properties of the distribution'. Since for the Vasicek model the density function for the short rate is known, MLE is likely to be easier. Nonetheless, I am using a numerical optimization procedure for the Vasicek calibration to be consistent with the calibration of the Nelson-Siegel family of models through non-linear least square (NLS). The employed routine involves minimizing the squared deviations of the

Maturity	Mean	Max	Min	StdDev	StdDev of the change ($*10^{-2}$)	Skewness	Autocorr1	Autocorr60
1	0.02817	0.05396	0.00288	0.01342	0.02841	.06070	0.99982	0.98260
2	0.02987	0.05510	0.00321	0.01304	0.03978	-0.09964	0.99979	0.98330
3	0.03182	0.05590	0.00418	0.01261	0.04461	-0.23904	0.99979	0.98452
4	0.03355	0.05641	0.00563	0.01221	0.04385	-0.32405	0.99980	0.98552
5	0.03516	0.05694	0.00749	0.01179	0.04453	-0.37062	0.99980	0.98628
6	0.03664	0.05774	0.00951	0.01143	0.04451	-0.38785	0.99980	0.98678
7	0.03795	0.05839	0.01129	0.01119	0.04274	-0.38789	0.99980	0.98708
8	0.03912	0.05906	0.01301	0.01098	0.04267	-0.38270	0.99980	0.98722
9	0.04013	0.05981	0.01452	0.01078	0.04289	-0.38294	0.99980	0.98727
10	0.04098	0.06050	0.01588	0.01061	0.04297	-0.38262	0.99980	0.98730
12	0.04249	0.06206	0.01786	0.01034	0.04353	-0.37104	0.99979	0.98724
15	0.04415	0.06383	0.01880	0.01021	0.04392	-0.35905	0.99978	0.98694
20	0.04540	0.06503	0.01870	0.01042	0.04521	-0.38734	0.99977	0.98646
25	0.04542	0.06475	0.01838	0.01075	0.04577	-0.37288	0.99976	0.98581
30	0.04490	0.06441	0.01807	0.01098	0.04634	-0.35401	0.99975	0.98511
40	0.03625	0.04937	0.01820	0.00769	0.05153	-0.21709	0.99947	0.96959
50	0.03569	0.04824	0.01843	0.00722	0.05259	-0.14627	0.99944	0.96829

Table 3: **Summary statistics of bootstrapped zero yields**
The zero yields are continuously compounded and stripped from Euro swap rates with maturity one to 10, 12, 15, 20, 25, 30, 40 and 50. Descriptive statistics for maturities 1 to 30 are based on the complete time series from 04-Jan-1999 to 17-April-2013. Maturities 40 and 50 apply to the shorter sample, 01-March-2005 to 17-April-2013. Means, maximums, minimums and standard deviations are reported in percentage points. Autocorrelations correspond to the first- and 60-day lag.

model's yields and the actual yields each day:

$$\underset{\alpha,\mu,\sigma,r(0)}{\text{minimize}} \sum_{t=1}^{10,12,15,20} [R(t) - R(t|\alpha,\mu,\sigma,r(0))]^2. \quad (29)$$

To get more sensible results, I restrict the parameters to a lower bound of zero and an upper bound of [4,0.4,10,0.4], respectively[2]. Consequently, I get 3728 sets of four parameters. I follow the approach by Duffie and Kan (1996) by estimating the short rate, as it is unobservable and should be treated as a limiting, not an actual rate. Instead, some may use, for instance, the three-month Euribor rate as a proxy for $r(0)$ by adding it to the yield inputs. However, this may increase estimation errors in the idiosyncratic short-rate data. The Matlab code for the Vasicek calibration is based on Holborow (2008). Similar to the Vasicek model, the Nelson-Siegel family of models may be fitted in various ways. According to Bolder and Stréliski (1999), there are four different variations for each model of the Nelson-Siegel family. Either model can be fitted by minimizing the sum of squared pricing errors through a NLS optimization or by specifying a log-likelihood equation for the price errors, which are assumed to follow a normal distribution. Furthermore, the NLS or MLE technique can be implemented with a full or a partial estimation. The full estimation derives the parameters simultaneously, whereas the partial estimation fixes either the βs or the τ(s) and estimates the non-fixed group. The partial procedure is similar to the two-step optimization originally used by Nelson and Siegel (1987), in which the non-linear shape parameter is derived in the first step, and in the second step the resulting linear model is estimated by

[2]There is no specific reason to set the boundary values to exact these numbers. However, they give sufficient flexibility but restrict unreasonable jumps.

OLS. Annaert et al. (2013) refine the procedure by using ridge regressions, which aim at reducing the high standard errors due to multicollinearity in the linear explanatory variables.

In my case, analogue to the Vasicek fitting, the Svensson model is fitted by NLS to find its six parameters β_0, β_1, β_2, β_3, β_4, τ_1, and τ_2 that minimize the sum of squared yield errors:

$$\sum_{t=1}^{10,12,15,20} [R(t) - R(t|\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \tau_1, \tau_2)]^2. \tag{30}$$

Furthermore, I apply a full estimation, since Bolder and Stréliski (1999) find the full procedure to outperform the partial estimation or to be at least as good as it. Additionally, the Vasicek's calibration can only be derived by full, not partial optimization.

De Pooter (2007) shows that restricting the Svensson factors during the calibration eliminates once-in-a-while large jumps in estimates. Since β_0, τ_1, and τ_2 should be strictly positive, I specify a lower bound of zero for them in the optimization routine. According to Bolder and Stréliski (1999), the Nelson-Siegel-Svensson model is very sensitive to the starting value. To have a sensible initial guess, I use the European Central Bank's average daily parameter values for their Svensson's calibrations over the period from September 6, 2004 to March 1, 2013. The historical averages are 3.1, -2.6, -3.9, and 0 percentage points for the beta coefficients, and 2.5 and 0.44 for the shape parameters.

Similarly, the Nelson model's four parameters are found via the same optimization problem as in Equation 30, using the same lower bounds and starting values but excluding β_4 and τ_2.

I use the Matlab function 'FitSvensson' to find 3728 sets of six parameters for the Svensson model and Matlab's 'FitNelsonSiegel' to find 3728 sets of four parameters for the Nelson model. The two functions belong to the Matlab's Financial Instrument Toolbox and are designed to derive a term structure using bond market data. As argued before, pricing swaps is equivalent to pricing par coupon-bonds as in Equation 28. Thus, I fit the Nelson-Siegel class to 13 par bonds that have a coupon rate equal to the quoted swap rate, and a unity face value.

The Wilson method is implemented through matrix calculations instead of least square optimizations. The calibration involves the computation of the 13x1 vector ζ for each day by solving Equation 12 (see Appendix A for the Matlab code).

3.4 Third step - Extrapolation

After having determined each model's parameters that fit the liquid curve, the extrapolated yields are obtained by plugging the fitted parameters into the model and calculating the model's implied yields for maturities 21 to 50 years. I end up with a cross-section of 29 extrapolated yields for each day for all four models. For instance, in the case of the Nelson-Siegel model, extrapolated yields $R(t)_e$ are

$$R(t)_e = \hat{\beta}_0 + \hat{\beta}_1(\frac{1 - \exp(-\hat{\tau}t)}{\hat{\tau}t}) + \hat{\beta}_2(\frac{1 - \exp(-\hat{\tau}t)}{\hat{\tau}t} - \exp(-\hat{\tau}t)), \tag{31}$$

for t = 21, 22,...,50. In the same fashion the Svensson and Vasicek extrapolated yields are derived. For the Smith-Wilson method, first the extrapolated discount prices are determined via Equation 10 as

$$P(t)_e = e^{-0.042*t} + \hat{\zeta} * W(t, m_{13x1}), \tag{32}$$

for t = 21, 22,...,50. A loop in Matlab finds the Wilson function's α such that the forward rate at t = 60 does not deviate by more than 3 basis points from 4.2%. α starts at 0.05 and may take on a value up to 0.2 with iterative steps of 0.0001. The resulting 29 discount prices are converted to zero yields with Equation 2.

3.5 Fourth step - Comparison

Besides a comparison of the visual characteristics of the models, tests of two other main criteria were performed. First, I evaluate if the extrapolation's aim of stability is achieved. I conduct the Brown–Forsythe test for equal variances, introduced by Brown and Forsythe (1974), to see whether I can reject the null hypothesis of equal volatilities of the actual yields and the extrapolated yields for each model. Instead of using the yields' level, I use the daily changes in yields:

$$y(t)_i = R(t)_i - R(t)_{i-1}, \tag{33}$$

where i goes from two to the total number of observations of i_{max} which equals 3728 regarding yields of maturities 25 and 30 years and 2122 regarding the 40- and 50-year yields.
For a two-group comparison, the test statistic for the Brown–Forsythe test is

$$F = (2 * (i_{max} - 1) - 2) \frac{(i_{max} - 1)[(\bar{Z}_e - \bar{Z}_{total})^2 + (\bar{Z} - \bar{Z}_{total})^2]}{\sum_{i=1}^{i_{max}-1} (Z_{i,e} - \bar{Z}_e)^2 + \sum_{i=1}^{i_{max}-1} (Z_i - \bar{Z})^2}, \tag{34}$$

where I have i_{max}-1, since relative changes imply one missing observation. $Z_{i,e}$ is the absolute difference between $y(t)_{i,e}$ and the median of $y(t)_e$ and Z is analogously the absolute difference between $y(t)_i$ and the median of $y(t)$. The test is based on the median and not the mean of the yields' changes because in this way it is more robust to non-normality in the data. $\bar{Z}_e$ is the mean of $Z_{i,e}$ and correspondingly, $\bar{Z}$ is the mean of Z_i. $\bar{Z}_{total}$ is the overall mean of $Z_{i,e}$ and Z_i. The Brown–Forsythe test underlies a F-distribution with (2*(i_{max}-1)-2) denominator degrees of freedom and one numerator degree of freedom. I implement the test in Matlab, function 'vartestn'.
Second, I test to what extent the extrapolations are off the actual yield level. Therefore, I calculate the Root Mean Sqaured Error (RMSE) defined as

$$RMSE_{m,t} = \sqrt{\frac{\sum_{i=1}^{i_{max}} (R(t)_{i,e} - R(t)_i)^2}{i_{max}}}, \tag{35}$$

with number of models m = 1, 2, 3, 4, maturity t = 25, 30, 40, 50 and $R(t)_{e,i}$ as the extrapolated yield that corresponds to the actual yield $R(t)_i$ for number of days i = 1, 2, 3...i_{max}. Again, for maturities t = 25, 30, i_{max} equals 3728 but due to the lack of actual yields $R(t)_i$ in the early sample period, maturities

$t = 40$, 50 have i_{max} equal to 2122.

To test what model is the best in terms of small extrapolation errors, I use the Model Confidence Set Test (MCS Test) in Oxmetrics. It belongs to the Multiple Comparison (MulCom) package, which allows to simultaneously analyze the forecasting accuracy of any number of models. Due to the fact that the extrapolations are out-of-sample, they can be considered as predictions. The MCS Test, introduced by Hansen, Lunde, and Nason (2010), finds the best model or a group of best models among a larger set of models. The test starts with the complete collection of models, and tests whether all models are equally good in terms of a pre-specified loss function and depending on a pre-specified confidence level. If it is rejected, an elimination rule decides on which model to be excluded from the set. Then, the equivalence test is repeated and models are eliminated step-by-step until a resulting set of models - the survivors - cannot be rejected anymore. I choose the evaluation criteria to be the model's Mean Squared Error (MSE), which is simply the square of the RMSE defined in Equation 35, and I set the significance level to 0.1. Given the four key maturities of extrapolations, namely 25, 30, 40 and 50 years, I conduct one MCS test per maturity. Each of the four tests includes the four models' time series of extrapolations whereas the first two tests are based on 3728 daily extrapolations per model and the last two on 2122. The model(s) with an MCS p-value higher than 0.1 thus make(s) up my final set of best model(s) per maturity.

4 Results

I present the results in three parts. First, I describe the outcomes for the calibrated parameters that apply to the Vasicek, the Svensson, and the Nelson models. Second, summary statistics describe the extrapolated yields of all models associated with step three. At last, I present the results of the comparison of the extrapolations by the equal variance and the MCS test.

4.1 Calibration

Since the Smith-Wilson method exactly fits the inputted yields, and depends on pre-specified parameters of convergence and the ultimate forward rate, I only present the parameter calibrations for the other three methods. Due to unstable parameters, for each of the three models I also estimate an alternative specification that reduces the free parameters by fixing the convergence parameters.

The first part of Table 4 shows the fitted parameters for the original Vasicek model as outlined before. According to the mean values, the average yield curve is upward sloping, since $r(0)$ is strictly smaller than the long rate μ. Whereas the short rate is relatively stable, the other three parameters change drastically. Especially the conversion parameter α is extremely volatile having a standard deviation, which is more than 38 times higher than that of the short rate. Economically unintuitive, the time series of the long-run average rate μ features large fluctuations. Coupled with a rather stable implied short rate, the strongly changing long-run rate induces unreasonable day-to-day changes in the yield curve. For instance, the fitted yield curve, being upward sloping the one day, may change to be humped-shaped the next day, even though the observed yields did not change noticeably. Figure 7 exemplifies the problem. Although the actual yield curve did not significantly change from one working day, July 2, 2010, to the next working day, July 5, 2010, the calibrated Vasicek yield curve changes from humped-shaped to upward sloping. The underlying issue is attributed to the volatile parameters, namely a short rate that increases by 13%, a long-run rate that grows by 789%, and a mean-reversion that decreases by 92%.

Parameters	$r(0)$	α	μ	σ	Sum of squared residuals in-fit
Original Vasicek					
Mean	0.0245	0.2266	0.1687	0.0241	0.0000041
Std.Dev.	0.0146	0.5645	0.1501	0.1056	0.0000078
Fixed Vasicek					
Mean	0.0245	0.1500	0.0676	0.0197	0.0000060
Std.Dev.	0.0144	0.0000	0.0217	0.0178	0.0000086

Table 4: **Vasicek calibration results**
Rows 1 and 2 present the estimated parameters for the original Vasicek model. Rows 3 and 4 show the estimation results when the convergence parameter α is fixed at 0.15. Both specifications are fitted by non-linear least square each day over the period 04-Jan-1999 to 17-April-2013. Summary statistics for $r(0)$, μ and σ are reported in percentage points. The sum of squared residuals are calculated in sample over maturities 1-20 for each day of the sample period and reported in percentage points.

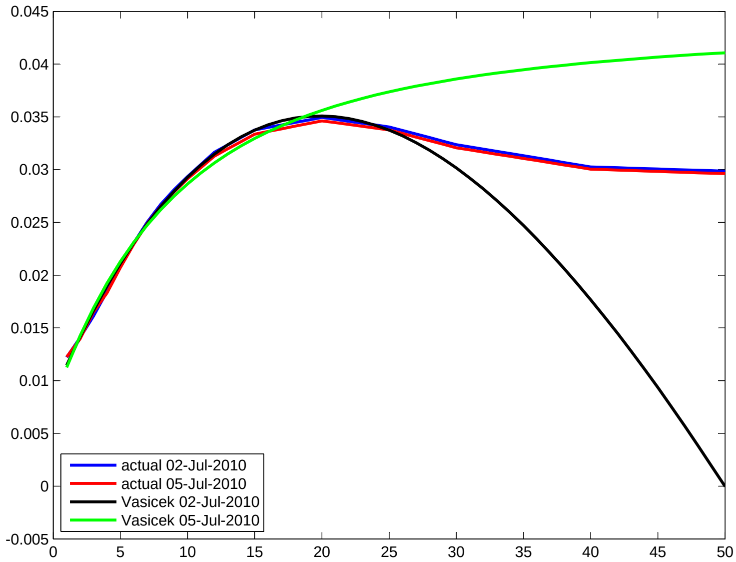
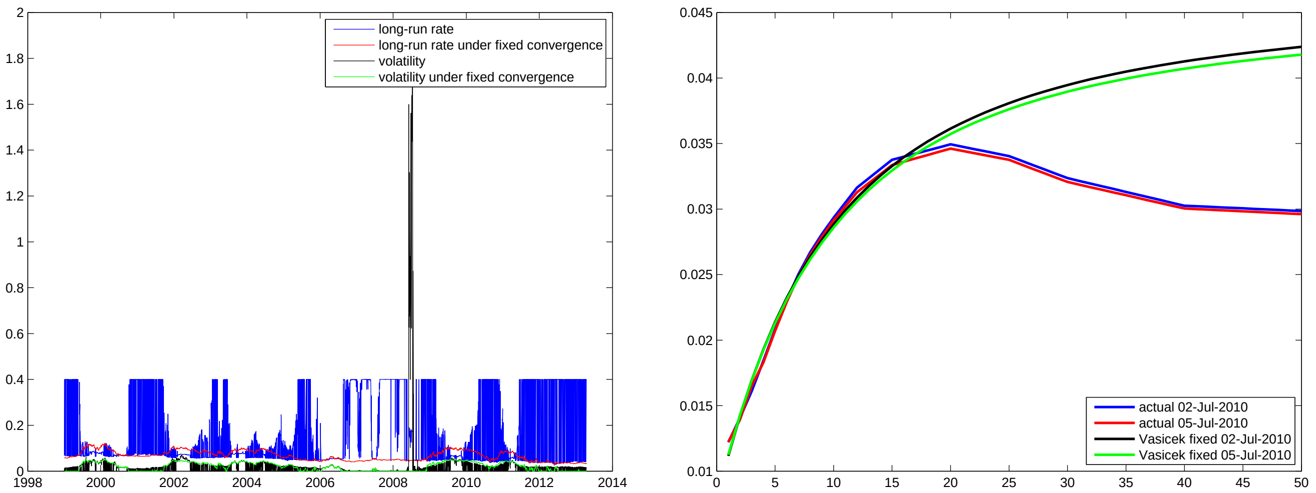


Figure 7: **Example of Vasicek's unstable calibrated parameters that lead to abrupt yield curve changes**
The actual yield curves are bootstrapped from Euro swap rates. The Vasicek yield curves are calibrated to the actual yields of maturity one to 10, 12, 15 and 20 and extrapolated beyond. All yields are continuously compounded and reported in percentage points.

In order to mitigate the variability of the parameters and to derive sensible calibrations in regard to parameter stability, I fix the convergence parameter α, which is the most volatile one. The second part of Table 4 shows the resulting fitted parameters when fixing α at a level of 0.15 percentage points. Indeed, all parameters exhibit stronger stability. Whereas the short rate is only marginally affected, mean and standard deviation of the long-run rate μ and volatility σ decrease substantially. Figure 8(a) emphasizes the effect on the stabilized time series of the two parameters. The associated result is depicted in Figure 8(b), which is the same as Figure 7 but with fixed Vasicek calibrations. The unreasonable change in the shape of the Vasicek yield curve vanishes, and when the actual yield shape remains constant, the

Vasicek shape does as well. The economically favorable effect of reducing the jumps of the long-run rate comes at a small cost in the form of a slight deterioration in the in-sample fit. However, in regard to the purpose of deriving good extrapolations in the next section, the increase in the sum of squared residuals by 0.0000018 is rather negligible compared to the benefit of obtaining more stable curves overall.



(a) Effect on the stability of the parameters

(b) Effect on day to day changes in the shape of the yield curve

Figure 8: **Fixing the conversion-rate parameter α in the Vasicek calibration**

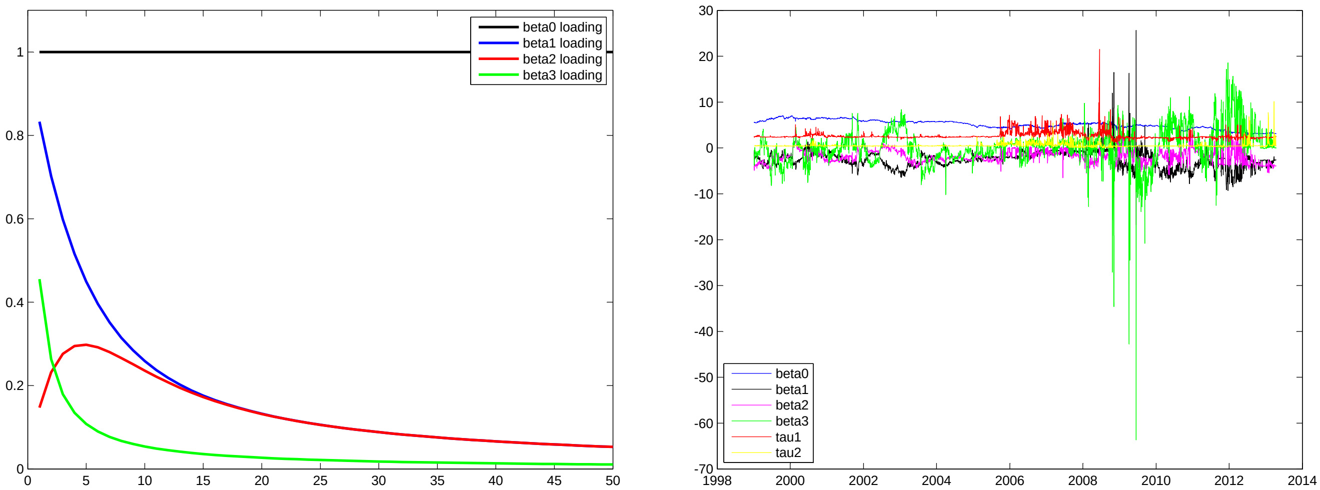
The first part of Table 5 shows the results for the calibration of the 3728 daily short-end yields to the Svensson functional form. Relative to the Vasicek calibration, which fits four parameters, the Svensson model, by freely choosing six parameters, achieves an improved in-sample fit. On average, the Svensson method's sum of squared residuals is 63% lower than that of the Vasicek model. The long-run interest rate β_0 is estimated to be around 5% on average over the whole sample period. This number is comparable to that of 6.7% from the fixed Vasicek model. Also, similar to the Vasicek calibration, the Svensson model's fitted parameters strongly differ in their stability. Figure 9(b) depicts that among all factors, the long-run rate fluctuates the least, and the Svensson model's additional curvature factor the most. Given that I fitted each curve to the daily yields only up to maturity 20, the very volatile β_3 suggests that it may be redundant and may only lead to an overparameterization of the yield curve function. The issue of little precision in β_3 is due to multicollinearity in the explanatory variables, which has been studied for instance by De Pooter (2007). Among the shape parameters, however, the additional shape parameter τ_2 is more stable than the Nelson parameter τ_1. In particular, the latter has twice the standard deviation of the former.

Figure 9(a) shows the average factor loadings corresponding to the mean value of τ_1 and τ_2 from the original Svensson model in the first part of Table 5. The graph strengthens the method's possible decomposition of the yield curve into different factors. Whereas the loading associated with β_0 is constant at 1 and therefore affects all yields similarly, the loading associated with β_1 is steadily decreasing and thus induces an increasing difference between short- and long-term yields. The former component models

Parameters	β_0	β_1	β_2	β_3	τ_1	τ_2	Sum of squared residuals in-fit
Original Svensson							
Mean	0.0511	-0.0252	-0.0187	0.0030	2.6504	0.5417	0.0000026
Std.Dev.	0.0098	0.0187	0.0148	0.0395	0.8482	0.4388	0.0000071
Fixed Svensson							
Mean	0.0508	-0.0290	-0.0129	0.0110	2.600	0.500	0.0000016
Std.Dev.	0.0100	0.0234	0.0139	0.0458	0.000	0.000	0.0000024

Table 5: **Svensson calibration results**
Rows 1 and 2 present the estimated parameters for the original Svensson model fitted by non-linear least square. Rows 3 and 4 show the OLS estimation results when the shape parameters τ_1 and τ_2 are fixed at 2.6 and 0.5, respectively. Both specifications are fitted each day over the period 04-Jan-1999 to 17-April-2013. Summary statistics for β-factors are reported in percentage points. The sum of squared residuals are calculated in sample over maturities 1-20 for each day of the sample and reported in percentage points.

the yield curve's level and the latter its slope. The loading for β_2 causes the yield curve to have a bow part in the short yield segment as the loading's maximum is reached at a maturity of around 5 years. The additional Svensson loading is associated with a second curvature within the 5- to 15-year segment according to Figure 9(a).



(a) Average factor loadings of the Svensson method

(b) Daily parameters of the Svensson method. β-factors in percentages.

Figure 9: **Fitting the yield curve with the Svensson method**

As shown in Figure 9(b), the Svensson method suffers from similar instability problems as the Vasicek model. Also, with the Svensson method, jumps may result without much change in actual yields. According to the original paper by Svensson (1994), the Svensson method is very sensitive to these extreme parameters at maturities beyond 20 years. Since my focus will be exactly on those maturities, I assume that a procedure similar to the one followed for the Vasicek model may solve the problem. Hence, I perform the fitting again while fixing the two convergence parameters τ_1 and τ_2 at their mean values of 2.6 and 0.5 respectively, which come from the original estimation (see first half of Table 5). The non-linear least square procedure reduces to a simple OLS cross-sectional regression, since the loadings are given and do not need to be estimated. I use Matlab's function 'regress' to retrieve the four beta coefficients reported in the second part of Table 5.

In contrast to the favorable results when fixing the convergence parameter in the Vasicek model, the Svensson model's fixed estimation does not produce noticeably less volatile estimated parameters. It only does not, by construction, exhibit any volatility in its shape parameters but the estimated factors β_0, β_1, and β_3 even increase in their variability. In opposition to what happened to the in-sample fit with the Vasicek model, the restricted Svensson model actually improves its mean sum of squared residuals. Hence, it is not only that the fixed loadings very likely stabilize the out-of sample extrapolations but also that the Svensson method fits better in sample. Similar results are found by Annaert et al. (2013), who use a conditional ridge regression to fit the Svensson model. This procedure involves two steps. First, the optimal shape parameters are estimated, and then the model is fitted with OLS.

An alternative solution to the Svensson model overparameterization problem is to simply switch to the Nelson specification by setting β_3 to zero. Table 6 shows in the upper panel the results of calibrating the bootstrapped liquid yields to the original Nelson yield curve form for each trading day between 1999 and 2013. The similarity of the Nelson and the Svensson models is unambiguous. The fitted parameters of the Nelson model are equivalent to the four Svensson parameters up to the third decimal place (compare the first part of Table 6 and Table 5). This is not a result by chance but due to the fact that the Nelson model is nested in the Svensson model. Despite equal parameter calibrations, the in-sample fit is twice as good with the Nelson method than with the Svensson model. This is especially due to the elimination of the very volatile β_3 of the Svensson model. Overall, the finding emphasizes the fact that six free parameters are too many for a calibration to only 13 yields, which end at the 20-year maturity. Restricting the functional form to only four parameters leads to more stable results.

Parameters	β_0	β_1	β_2	τ	Sum of squared residuals in-fit
Original Nelson					
Mean	0.0507	-0.0238	-0.0183	2.6538	0.0000013
Std.Dev.	0.0099	0.0124	0.0145	0.8255	0.0000017
Fixed Nelson					
Mean	0.0484	-0.0168	-0.0348	1.4	0.0000051
Std.Dev.	0.0100	0.0115	0.0156	0	0.0000035

Table 6: **Nelson calibration results**
Rows 1 and 2 present the estimated parameters for the original Nelson model fitted by non-linear least square. Rows 3 and 4 show the OLS estimation results when the shape parameters τ_1 is fixed at 2.6. Both specifications are fitted each day over the period 04-Jan-1999 to 17-April-2013. Summary statistics for β-factors are reported in percentage points. The sum of squared residuals are calculated in sample over maturities 1-20 for each day of the sample and reported in percentage points.

As the standard deviations of the Nelson parameters are similar to those of the Svensson parameters, stability problems may still exist with the Nelson model. Again, a natural choice is to fix the convergence parameter. Regarding the Nelson method, this is a very common practice (see for instance Diebold and Li (2006), Fabozzi, Martellini, and Priaulet (2005) and De Pooter (2007)). The reason for this is that τ determines the maturity at which the loading on the medium term factor achieves its maximum, and that point rarely changes (Annaert et al., 2013). Following Diebold and Li (2006), who set the shape parameter to 1.37, I fix it at 1.4. The new calibrations are derived by OLS regressions just as in the

Svensson case. The resulting fitted parameters all decrease compared to the unrestricted Nelson method (see the lower panel of Table 6). The exception is β_1, which increases from -2.4% to -1.7%. As expected, the in-sample fit deteriorates, since there is one parameter less to be freely chosen during the fitting process. Figure 10 shows that fixing the shape parameter eliminates its strong variability but does not mitigate the other parameters' fluctuations.

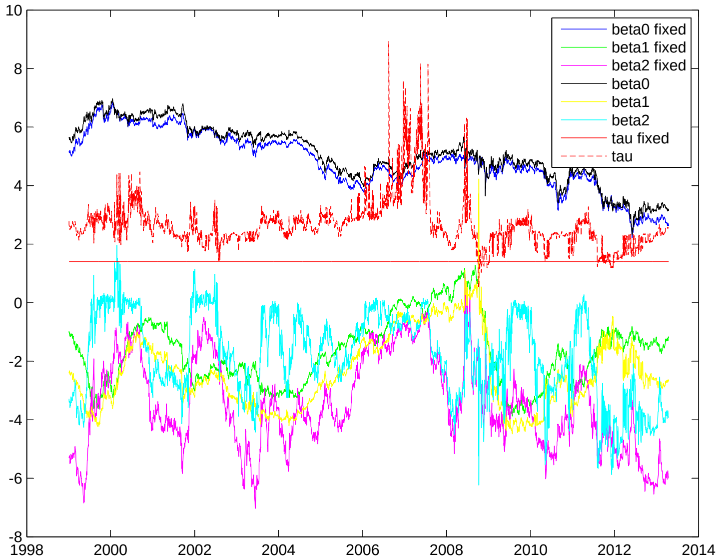
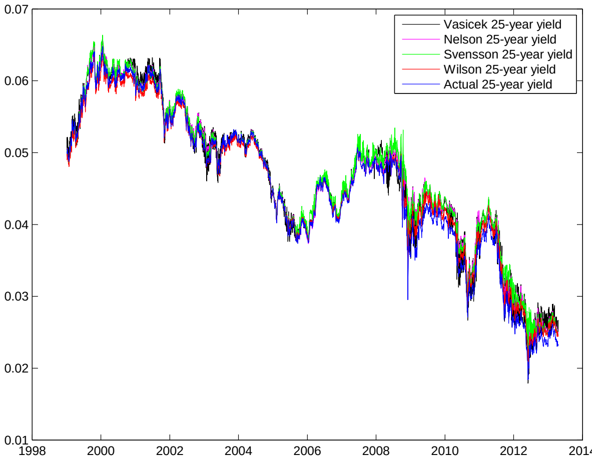


Figure 10: **The daily fitted parameters of the Nelson method and its fixed version**
The β-factors are reported in percentages.

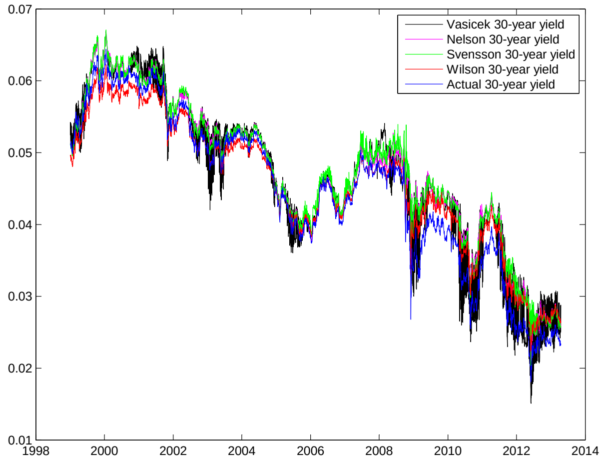
4.2 Extrapolation

The extrapolation results for the original models show the instability issue discussed in the previous section. Only the Wilson model offers a comparably smooth evolution. For the Vasicek, the Svensson and the Nelson models, as the maturity of the extrapolated yields increases, their volatilities increase as well. Whereas Figure 11(a) still depicts seemingly well-behaved 25-year extrapolations and the 30-year yields seem to track the actual yield acceptably well (see Figure 11(b)), the picture strongly changes for very long yields. Figure 11(c) and Figure 11(d) emphasize the extreme volatility of the extrapolated Vasicek yields that even lead to unrealistically temporary negative interest rates. Even though the result obtained by the Vasicek model is most striking, also the Svensson model exhibits extrapolated yields that fluctuate more than the actual yields. Since the main purpose of the extrapolations is to stabilize the long end of the yield curve, the original models are of little use. Hence, I will turn the main focus to the 'fixed' versions of the Vasicek, the Svensson and the Nelson models. Together with the Wilson model, these four specifications make up my base set of models.

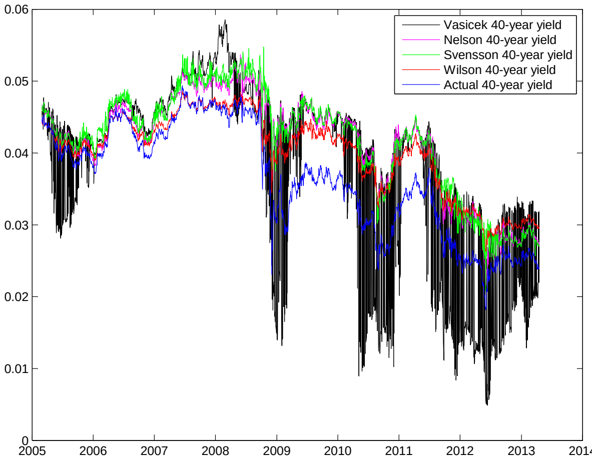
In contrast to the original models, the fixed models seem to achieve more stable extrapolated yields, and are much better comparable to the Wilson method. Figure 12(a) shows that all four models' yields closely follow the overall downward trend of the actual yield with maturity of 25 years. The similarity is reflected in Figure 12(c), which shows all means to be around the actual 4.5%. The similarity between



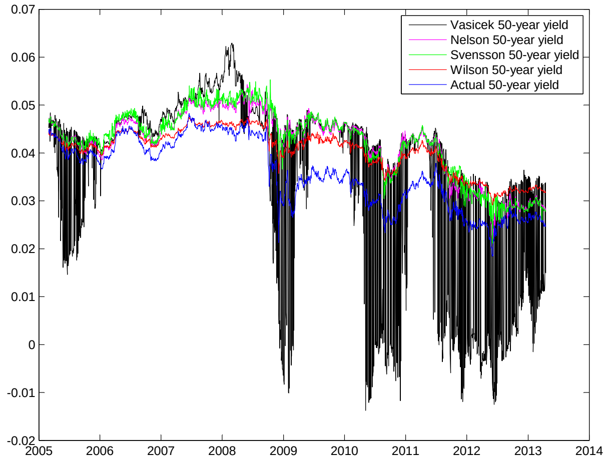
(a) 25-year yield in percentage points



(b) 30-year yield in percentage points



(c) 40-year yield in percentage points

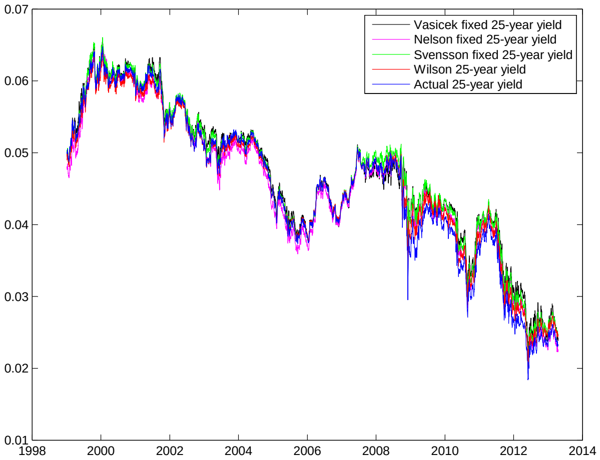


(d) 50-year yield in percentage points

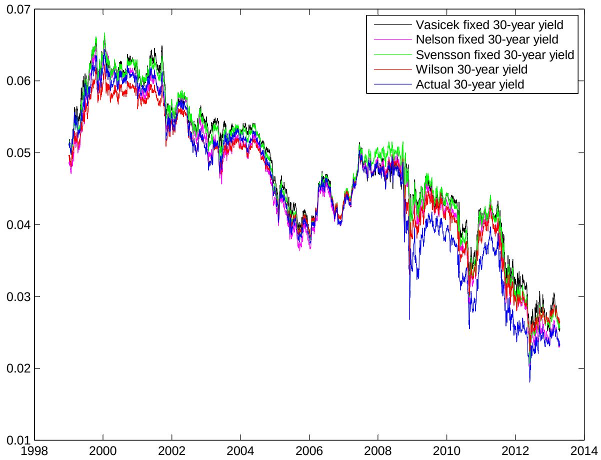
Figure 11: **Original models' extrapolated yields and the actual daily yields**

each base model's 25-year yield and the actual yield also holds for the maximum and the minimum values as well as for the volatility. Distributional differences between extrapolations and actual yields are still quite small at maturity of 30 years (see 12(d)) but become more and more pronounced when moving to higher-maturity yields. Figure 12(b) already illustrates that especially for the post 2008 period, all models evolve above the actual yield.

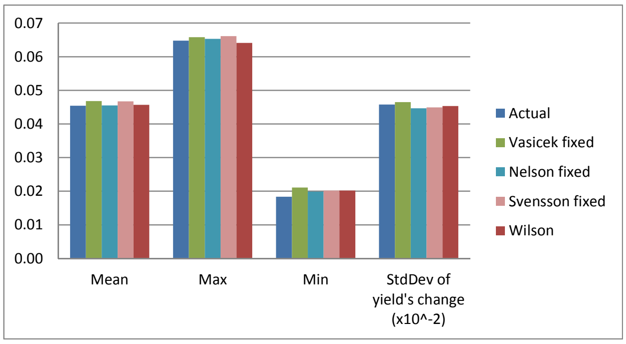
The summary statistics for the 30-year yield reveals the tendency of the maximum and the minimum yields to be highest for the Wilson method, thereby hinting at the possibility for the model to have the strongest upward bias. The minimum value for the Wilson method becomes exceptionally high with the 40-year yield. Figure 13(c) gives a minimum for the Wilson method that is 1.5 times higher than that of the actual yields. However, the upward bias in the level is present in all methods. Figure 13(a) shows a clear-cut discrepancy between the models and the actual long-run yield with maturity of 40 years. It seems that particularly after 2008 the models' yields mirror the actual yield while being shifted up by a constant factor. In the course of year 2012, it appears that the Nelson and the Svensson models are the first to have learned about the general downward shift of the entire yield curve. Whereas the yields produced by the Wilson and the Vasicek models continue to stay substantially above the actual yield, the Nelson and the Svensson models' yields started to get closer.



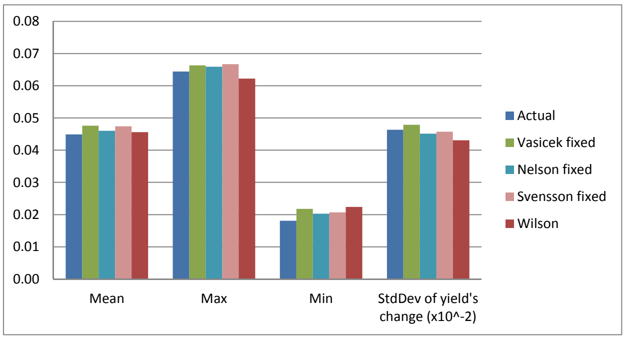
(a) 25-year yield in percentage points



(b) 30-year yield in percentage points



(c) Summary statistics of the 25-year yield

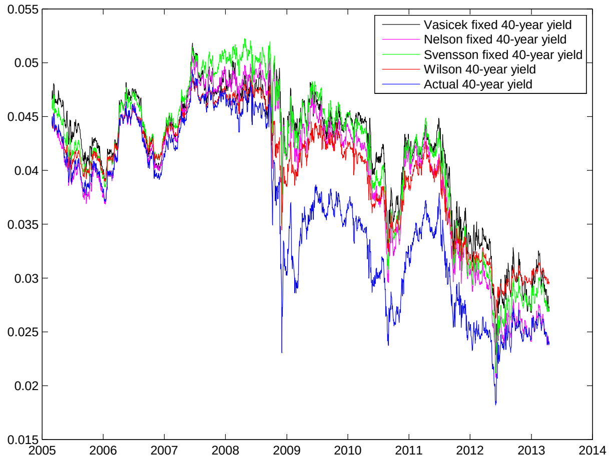


(d) Summary statistics of the 30-year yield

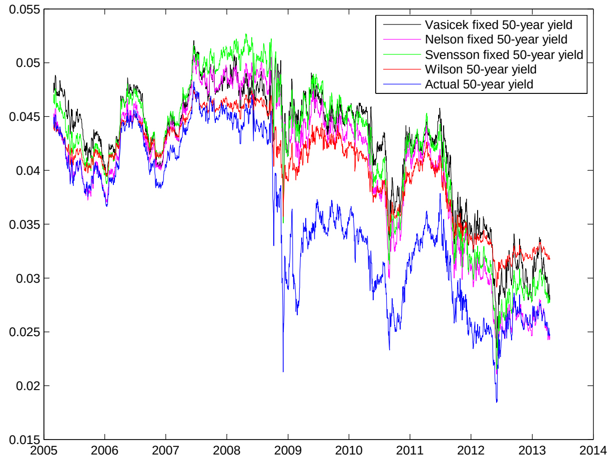
Figure 12: **Base models' extrapolated yields and the actual yield in days at maturities 25 and 30 years**

The Wilson method, by construction, cannot shift down unless changing the predetermined Ultimate Forward Rate. A mitigating solution is to extend the convergence period between the LLP and the UFR, thereby alleviating the drastic rise of the yield reaching the too high UFR level. Figure 14 shows the development of the original Wilson method's extrapolated yield error at maturity of 50 years, which corresponds to the EIOPA's suggested convergence maturity of 60 years, and an alternative Wilson model's 50-year yield error, which is associated with a UFR reached at maturity of 80 years. Extending the transition period by 20 years lowers the deviation from the actual yields particularly after the strong downward trend in the actual yield during 2011. Still, a more radically increased transition is necessary to close the existing discrepancy in levels between the Wilson yield and the actual one in the post financial crisis period, given an unchanged UFR level.

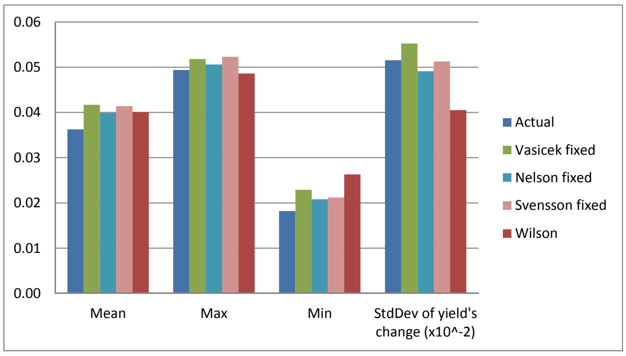
Going from the 40-year yield to the 50-year yield implies shifting up all models' extrapolations by approximately 5 basis points on average. In contrast, the actual yield with maturity of 50 years is six basis points lower than the actual 40-year yield. This phenomenon has already been associated with the convexity effect inherent in the observed yield curves. Consequently, the difference between the minimum value for the actual yield and the Wilson yield widens with maturity of 50 years (compare Figure 13(c) and 13(d)). Despite this seemingly upward bias, the Wilson method is most stable. In terms of its change,



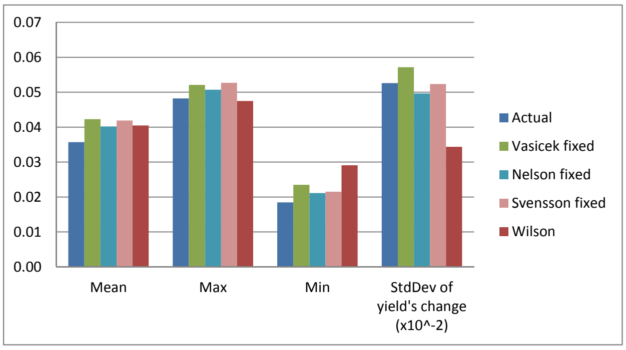
(a) 40-year yield in percentage points



(b) 50-year yield in percentage points



(c) Summary statistics of the 40-year yield



(d) Summary statistics of the 50-year yield

Figure 13: **Base models' extrapolated yields and the actual yield in days at maturities 40 and 50 years**

the Wilson method reduces the volatility of the 50-year yield from 5.3 to 3.4 basis points.

The increased stability of the extrapolated yields induces them to also be more persistent compared to the actual yields. Figure 15(a) shows the daily autocorrelation of order 60 for the group of 25-year, 30-year, 40-year and 50-year yields. Again, there is the tendency of a growing difference in persistence when moving from maturities of 25 to 50 years. Overall, the Nelson method consistently produces the highest autocorrelated yields peaking at 0.987. Still, none of the models' autocorrelation coefficients fall short of 0.97; not even at a lag of 60 days. Statistically, this may suggest unit root processes and infinite second and third moments. This is the reason why in Figure 15(b) I report the skewness of the changes, and not that of the level. There is a consistent picture of skewness across maturities. All yields are negatively skewed with the extrapolated Nelson model yields least skewed and the Svensson model yields most skewed.

4.3 Comparison

Overall, the summary statistics and the visual inspection suggest that, on the one hand, all base models consistently overestimate actual yields but on the other hand they (except for the Vasicek model) seem to reduce the volatility of yields. In the following, I test these two issues formally.

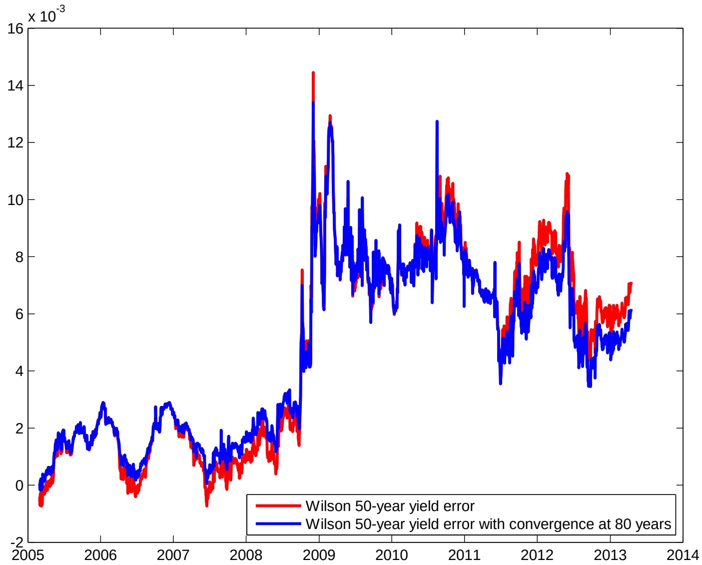
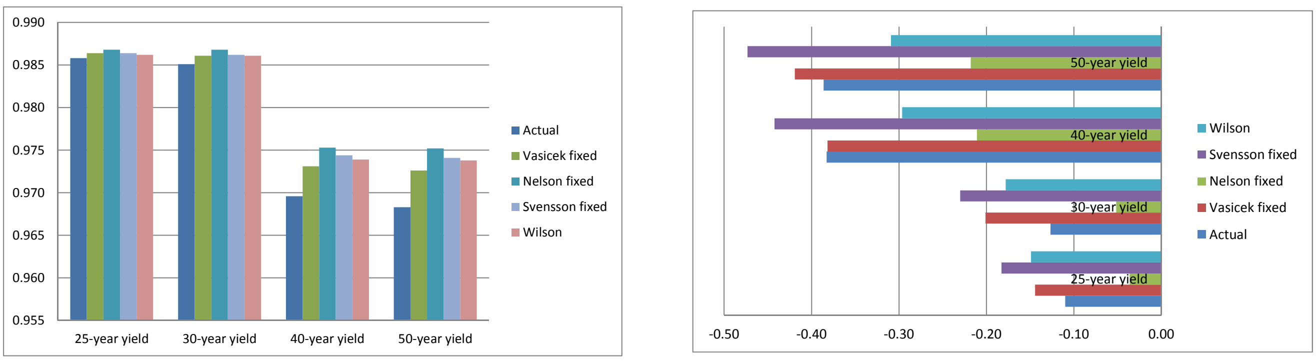


Figure 14: **50-year yield extrapolation error of the Wilson model with convergence at maturity 60 and 80 years**



(a) Autocorrelation coefficients after 60 days

(b) Skewness

Figure 15: **Persistence and skewness of extrapolated yields and the actual yield**

4.3.1 Variance test

Due to the likely non-stationarity in the time series of yields, suggested by the autocorrelations in the previous section, I test the extrapolations on their stability in terms of their changes, not levels. The extrapolated model is successful if I can reject the null hypothesis of equal variance between the actual and the extrapolated series of the yields' changes. Each model is tested separately. Table 7 presents the results for all eight models, namely the four original ones and their variations.

The variance test confirms the instability of Svensson's and Vasicek's extrapolations. For all maturities, both models produce yields that are significantly less stable at the 5% significance level. Similarly, the original Nelson model produces extrapolations that are worse in terms of stability. The exception is the 25-year yields, which are as stable as the actual 25-year yields at the 10% significance level.

The yields obtained by the fixed versions of the three models improve on stability but are still not statistically more stable than the actual yields. Even though in magnitude all standard deviations are substantially lowered when switching from the original to the fixed version, it is not sufficient to have significantly more stable extrapolations relative to the actual yields. The fixed Vasicek model continues to underperform at maturities of 40 and 50 years is performing equally well compared to the rest of the models only for maturities of 25 and 30 years. The fixed Nelson and the fixed Svensson models produce extrapolations which are as stable as the actual yields with high confidence. Still, they render unusable for stabilizing the long end of the yield curve.

In contrast, the Wilson method achieves significantly more stable extrapolations. Figure 16(a) exemplifies the increased stability at the longest-maturity yield. With convergence at 60 years, the Wilson method significantly lowers the volatility at maturities of 30, 40 and 50 years, whereas with convergence at 80 years, it does so only at the two longest maturities. Otherwise, both methods are statistically equally stable to the actual yields. The reason why the original Wilson method outperforms its variation is that, by construction, the longer it takes for the Wilson method with convergence at 80 years to reach the fixed UFR, the more variable the yields are during the transition.

Another point to note is that at the 25-year maturity no model is able to significantly decrease volatility. Even though the Nelson model produces the lowest standard deviation in magnitude, the standard deviation in the Nelson yield's change is still statistically equivalent to that of the actual one for maturity of 25 years as Figure 16(b) visualizes.

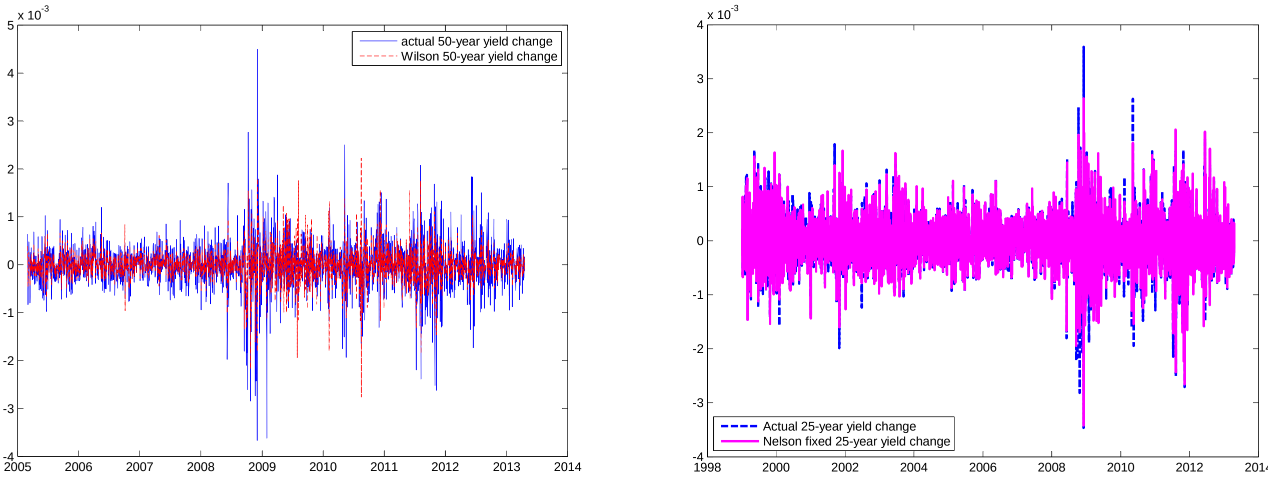
Maturity	**Actual**	**Vasicek**	**Vasicek fixed**	**Nelson**	**Nelson fixed**	**Svensson**	**Svensson fixed**	**Wilson**	**Wilson 80**
25	0.0458	0.1473	0.0465	0.0475	0.0447	0.0641	0.0449	0.0453	0.0466
		(0.00)	(0.30)	(0.14)	(0.74)	(0.00)	(0.44)	(0.57)	(0.57)
30	0.0463	0.2985	0.0479	0.0495	0.0451	0.0695	0.0458	0.0431	0.0460
		(0.00)	(0.08)	(0.01)	(0.96)	(0.00)	(0.77)	(0.00)	(0.45)
40	0.0515	0.8821	0.0552	0.0588	0.0491	0.0944	0.0513	0.0405	0.0459
		(0.00)	(0.00)	(0.00)	(0.64)	(0.00)	(0.96)	(0.00)	(0.00)
50	0.0526	1.5943	0.0572	0.0619	0.0497	0.1017	0.0524	0.0344	0.0407
		(0.00)	(0.00)	(0.00)	(0.68)	(0.00)	(0.83)	(0.00)	(0.00)

Table 7: **Stability test**
The first column reports the standard deviation of the time series of daily actual yield changes. Columns 2-9 show the standard deviation of the time series of daily extrapolated yield changes. All standard deviations are in percentages. P-values are in parenthesis to the null hypothesis that the extrapolated yields have the same variance as the actual yields. The equal variance test is based on the Brown-Forsythe statistic.

Overall, whereas the Vasicek model performs worst in the equal variance test, there is strong evidence that the Wilson method is the most stable one. In particular, the two Wilson models are the only ones that are able to significantly lower the volatility. Since the summary statistics suggested that stability came hand in hand with upward biased yield levels, the next test investigates the models in terms of their extrapolation error.

4.3.2 Extrapolation error test

According to Table 8, the two Wilson method's specifications have the lowest Root Mean Square extrapolation error across maturities, and the Vasicek model has the worst. The overall lowest RMSE of 8.56



(a) The Wilson method significantly reduces volatility

(b) The Nelson fixed method produces significantly equal volatility

Figure 16: **Daily actual yield changes and extrapolated yield changes**

basis points has been produced for the 25-year extrapolations by the Wilson method with convergence at 80 years. The highest RMSE of 144 basis points belongs to the Vasicek model at the 50-year maturity. Within each extrapolation model, the error increases with maturity. A major reason for this is suggested by the exemplary yield curves on September 1, 2006 (see Figure 17). It shows that none of the models is able to account for the convexity effect in their extrapolated yields, which lead to a systematic overestimation of long-term yields.

Maturity	Vasicek	Vasicek fixed	Nelson	Nelson fixed	Svensson	Svensson fixed	Wilson	Wilson 80
25	0.001700	0.002000	0.001800	0.001491	0.002000	0.001800	0.000973	0.000856
30	0.003000	0.003500	0.003100	0.002430	0.003300	0.003100	0.002200	0.001931
40	0.007900	0.006600	0.006100	0.005031	0.006300	0.006100	0.004800	0.004553
50	0.014400	0.007700	0.007200	0.005987	0.007500	0.007200	0.005800	0.005571

Table 8: **Bias criterion**
The out-of-sample Root Mean Square Error is based on the models' extrapolated yields and the actual yield. The RMSE is reported in percentage points. Maturities 25 and 30 apply to the sample period from 04-Jan-1999 to 17-April-2013. Maturities 40 and 50 are based on the period from 01-March-2005 to 17-April-2013.

Comparing the fixed and the original version of the Nelson and the Svensson models shows that for both versions the Nelson model achieves lower extrapolation errors. Even though reducing the estimated parameters usually reduces the fit and increases in-sample errors, in the case of extrapolations, a small number of parameters seems to lower out-of-sample errors. The same applies to the Vasicek 40- and 50-year yields because the fixed version displays lower errors than the original method. However, for maturities 25 and 30, the original model is more accurate. The reason for that is expressed by Figure 18. It depicts that the fixed version's error evolves mainly at the upper outer error of the fluctuating original model. Whereas the original model's volatility ensures that the error is often close to zero, the stable fixed version hits less frequently the zero line.

The formal Model Confidence Set test confirms that the Wilson model with convergence at 80 years significantly outperforms all other models in regard to accuracy. Its superiority over the original Wilson

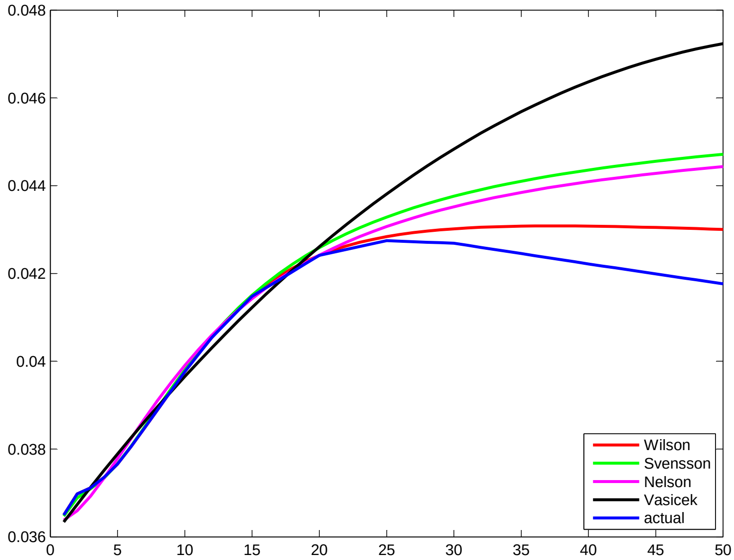


Figure 17: **Original models' yield curves and the actual yield curve on 01-Sep-2006**
The actual yield curve is bootstrapped from Euro swap rates. The models' yield curves are calibrated to the actual yields of maturity from one to 10, 12, 15 and 20 and extrapolated beyond. All yields are continuously compounded and reported in percentage points.

method is especially driven by the last years of the sample period as already discussed in the context of Figure 14. To emphasize this point, I split the sample into two periods. The first period runs from January 4, 1999 to December 31, 2008, as the time series of yields stress a strong shift in overall yields at the end of 2008 after the financial crisis has hit. The second period extends from January 5, 2009 to April 17, 2013. During the first subsample, the original Wilson method exhibits the significantly lowest error at maturities of 40 and 50 years, whereas the longer-transition Wilson method continues to only outperform at maturities of 25 and 30 years. Figure 14 visualizes that the yield produced by the original Wilson model is closer to the actual 50-year yield between 2007 and 2009, a time span that makes up 50% of the total first period for the 50-year yield. During the second period, the Wilson 80-year-convergence method is again superior for the 25- and 30-year yields, yet surprisingly, the Vasicek model produces the significantly lowest 40- and 50-year extrapolations. The latter result is in opposite to the variance test conclusion for which the original Vasicek model performed extraordinary badly. However, the strong Vasicek volatility does not influence the fact that the Vasicek model is still the one with the least bias for the post-financial crisis period. Figure 19 shows that even though the Wilson 80-year-convergence model is extremely stable as compared to the Vasicek model, the former constantly produces a positive error, whereas the latter has errors varying in sign and magnitude from day to day . On average, the Vasicek model's error is closer to zero than the Wilson 80-year-convergence model's error.

Since stability is still a major criteria for an extrapolation model, the original Vasicek model's bias-free performance is negligible. Restricting the collection of models to the base models, the MCS test is 90% confident that for the entire sample period and across extrapolated maturities, the Wilson model is the best one. The fixed Nelson model is, however, equally good at maturity 50 when increasing confidence to 95%. The latter finding is due to the seemingly better performance of the fixed Nelson model in the

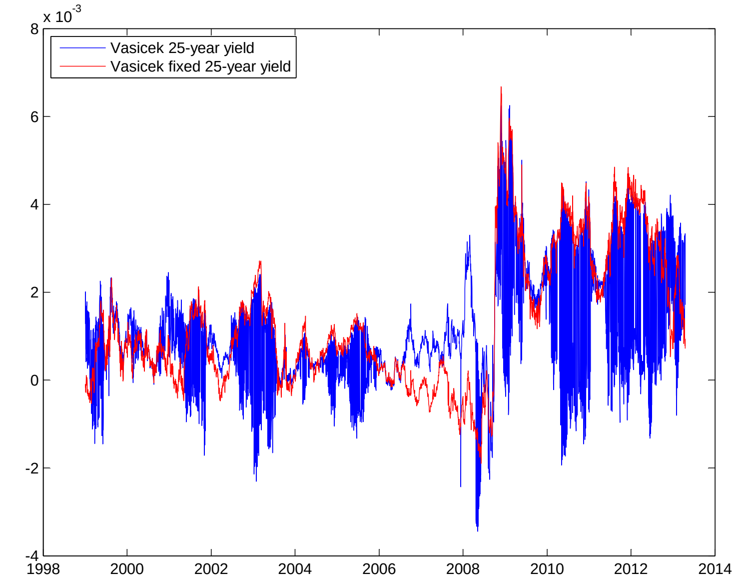


Figure 18: **25-year yield extrapolation error of the original and the fixed Vasicek models**

last years of the sample period (see Figure 13(b)). When dividing the sample again into the pre- and post-crisis periods, the MCS test proves the fixed Nelson 40- and 50-year extrapolations to be superior to the extrapolations produced by all other models during the second sample period. Still, for the first and the second period's 25- and 30-year extrapolations, the Wilson method continues to produce the most accurate extrapolations (see Appendix B for test results).

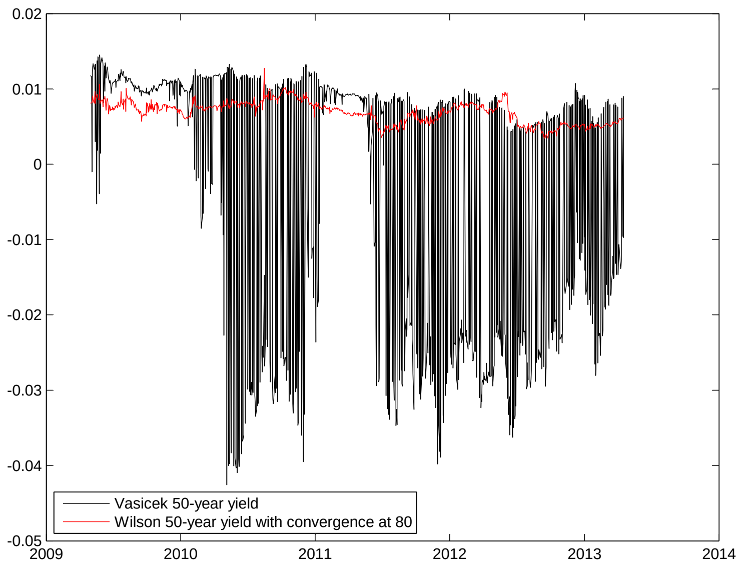


Figure 19: **50-year yield extrapolation error of the Vasicek model and the Wilson model with convergence at 80 for the second subsample from 05-Jan-2009 to 17-April-2013**

5 Discussion

The results of the formal tests regarding the extent to which each model stabilizes long-term yields while minimizing the extrapolation error imply that there is no evidence against Proposition 1 but Proposition 2 turns out to be the opposite to what was initially expected. In accordance with Proposition 1, the Smith-Wilson method creates the most stable extrapolations and the Vasicek model the least ones. The Nelson-Siegel and the Svensson methods turn out to be inbetween. This ranking holds for the original versions as well as for the fixed ones.

The significantly more volatile yield curve extrapolations of the Vasicek model relative to the actual yields renders one-factor equilibrium models futile for extrapolations. The reason for that is, however, not the same reason that dismissed the Vasicek model as a good term structure model for pricing interest rate derivatives. The Vasicek model's feature of having a rough outline of the yield curve instead of an exact fit induces large pricing errors but not necessarily unstable extrapolations. Contrarily, since extrapolations also do not require exact fitting but rather sound and stable yields, the fundamentally motivated model that produces little flexible yield curve shapes, may well have stabilized the yield curve. However, the main problem is the Vasicek's affine term structure solution that is too parametric to create stability.

Fixing the convergence parameter mitigates the volatility but still leaves the model with a non-linear and complex yield curve function (for reference see again Equation 25, 26 and 27). The parametric solution continues to induce high parameters' sensitivity, leading to large yield curve changes without much change in actual yields (see again Figure 7).

After all, the Vasicek model seems to be outdated. In order to price and hedge interest rate products, academics and practitioners have long turned to more advanced models that are, in contrast to the Vasicek model, multi-factor driven, arbitrage-free and have a zero probability of producing negative interest rates. Now, the Vasicek model's chance of reviving by rendering useful as stabilizing extrapolation model turns out to have failed as well. Consequently, the formal variance test gives sound reasons for the EIOPA and its consultants not to have included the Vasicek model in the discussion. Contrarily, the Nelson-Siegel family has been recommended as a potential alternative to be tested in QIS5 (CFO Forum and CRO Forum, 2010). The stability results confirm that the statistical class of models produce substantially smaller variances than the economic model. Even though the Nelson-Siegel and the Nelson-Siegel-Svensson approaches model the yield curve with a parametrized functional form, this form as well is simpler than the Vasicek function. By fixing the shape parameters, the yield curve turns linear, whereas the Vasicek model always continues to be non-linear. Since the linearized Nelson and Svensson models create more stability than the non-linear versions, reducing the estimated parameters again leads to better extrapolations in terms of stability. Along the same lines, within this family, the Nelson-Siegel model performs better than the Svensson method in terms of stability. The finding holds for the original as well as for the fixed version. In both cases, the reason is twofold. First, the Nelson model consists of less parameters than the Svensson model. Second, it also consists of one less explanatory variable, thereby mitigating multicollinearity and consequently noise in the estimates. The latter is in line with Annaert et al. (2013), who find improved extrapolations when conditioning the estimation of the Svensson model parameters on the degree of dependence between the non-linear components of the model. Their results show that when estimating the Nelson and the Svensson models in a more standard, ridge-free way, the standard deviation of the 30-year swap rate increases in comparison to the actual rate. My findings further support the inability of this family to significantly reduce volatility for Euro swaps, since the estimations themselves create too much variability. Thus, previous findings of fit and forecast performance between models cannot be generalized to extrapolation procedures. For instance, De Pooter (2007) finds that a four-factor Nelson-Siegel model produces a better in-sample fit and more accurate out-of-sample forecasts than the original three-factor Nelson-Siegel model. However, my extrapolation results show that more parameters are less favorable because they allow for too much variability at the long end. Consequently, the Smith-Wilson model, a technique that does not require any estimation, performs best on the stability criteria. The fixed UFR ensures, by construction, that the long end does not vary. No matter what interest rate level the last liquid point ends at, the 60-year yield stays the same. By changing the convergence point, the Smith-Wilson method creates the possibility to increase or decrease the stability at one's own discretion. A convergence maturity of, for instance, 50 years instead of 60 years induces all extrapolated yields to reach the UFR-level faster, so that less latitude is given to its different levels. Furthermore, the variance test has shown that a longer convergence, i.e. to a maturity of 80 years, increases volatility. Therefore, from all models considered, the Smith-Wilson one is unique in the discretion it gives regarding the stability of extrapolations. The parameters of the other three models, which govern the stability of the yield curve, are estimated on the basis of the given actual data inputs instead of being fixed by choice. To limit individual insurance companies to employ the Smith-Wilson method as they please, the EIOPA should enforce a unique convergence point instead of having 'no fixed, predefined maturity where the UFR is deemed to be arrived at' (European Insurance and Occupational Pensions Authority, 2010, p.5).

The uniqueness of the Smith-Wilson method has also been stressed by Thomas and Maré (2007), who find that between the Nelson-Siegel family models and the Smith-Wilson model, the Smith-Wilson technique minimizes interest rate risk the most. Also common to my extrapolation results, they find simple extrapolation techniques to even increase risk, and not to lower it. Similarly, in my case, the original Vasicek, Nelson and Svensson models worsen, and not foster the long rates' stability.

Overall, the results of this study and that by Thomas and Maré (2007) about the relative performance of the Smith-Wilson method unambiguously hint at its superiority as extrapolating model. This may relativize the strong criticism put forward by the CRO Forum regarding the appropriateness of the model.

One critique, however, may not sound too unreasonable, as my results underpin the CRO Forum's proposition of moving out the last liquid point by 10 years. At the extrapolated 25-year yield, the variance test presents that no model achieves significantly lower volatility. Instead of indicating failures of the models, the finding rather shows that this yield is already sufficiently liquid and cannot be stabilized any further. The analysis of bid-ask spreads of Euro swaps (see Table 2) as well as the report by TriOptima (see Table 1) corroborate that the 20- to 30-year maturity segment is indeed very liquid. The standard deviation of the spread's change associated with the 30-year swap is substantially lower than that of the 12- to 25-year swaps. In line with that, the same pattern is observed in trade counts and trade volumes.

Extending the liquid part of the curve to 30 years will thus not effect the stability of the extrapolations since they are equally volatile to the actual yield before maturity of 30 years. It will, however, push the Smith-Wilson method closer to Solvency II's objective of market-consistent valuation when the liquid part of the curve is extended and the 'artificial' extrapolated part reduced. On that account, the EIOPA may benefit from considering a last liquid point beyond 20 years for the Euro interest rate term structure. A natural choice may be the British pound yield curve's LLP of 30 years.
Overall, the liquidity of the European interest rate swap market does not give as much reason for extrapolations as other markets. Compared to i.e. the South-African swap market in the study by Thomas and Maré (2007), actual interest rates are much more volatile than their European equivalents. While the 30-year African swap rate has an annual standard deviation of 1.63% between 2000 and 2007, the European swap, based on my data, exhibits 1.07% between 1999 to 2013. This may be the reason why the Nelson-Siegel family models achieve interest rate risk reduction in the African market - even though not as much as the Wilson model - whereas they do not significantly decrease volatility in the European market. Thus, it seems as if the European yield curve is already sufficiently liquid, so that the weaker Nelson-Siegel class cannot improve on it anymore.

Proposition 1 on stability is confirmed. Yet, Proposition 2 turns out to be quite the opposite to the expectation about the models' extrapolation errors. Instead of having the most biased extrapolations, the Wilson method turns out to perform best overall. Even though, the UFR is an artificial construct, the MCS test indicates that the actual yield curve may inhere an UFR similar to the predetermined Wilson model's UFR. Figure 20 stresses the good fit of the Wilson model in contrast to the other models. Not only does it have the best in-sample fit, it also comes closest to the actual yield shape beyond the last liquid point.

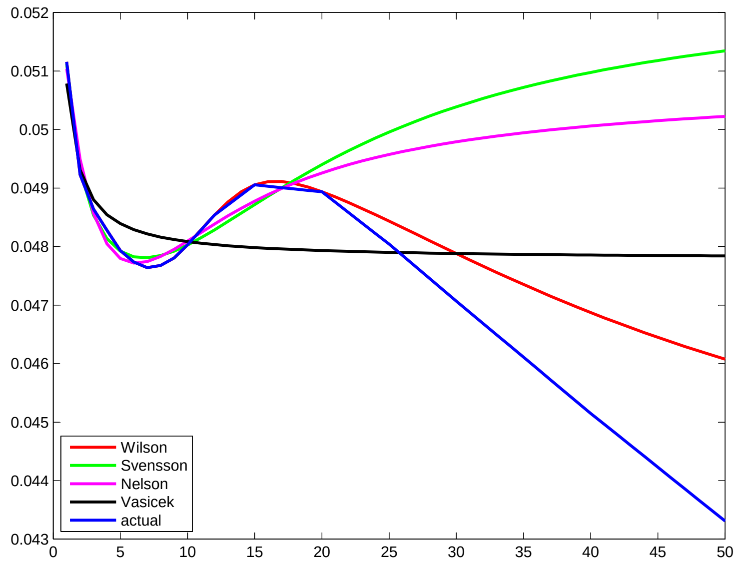


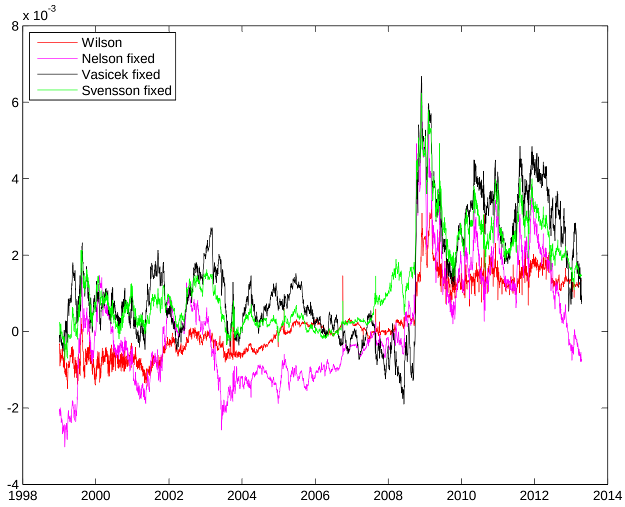
Figure 20: **Original models' yield curves and the actual yield curve on 01-Aug-2008**
The actual yield curve is bootstrapped from Euro swap rates. The models' yield curves are calibrated to the actual yields of maturity from one to 10, 12, 15 and 20 and extrapolated beyond. All yields are continuously compounded and reported in percentage points.

Still, as can be seen in Figure 20, all models including the Wilson model demonstrate an upward bias at maturities 25 to 50. As the series of extrapolation errors in Figures 21(a), 21(c), 21(c) and 21(d) show, this bias is consistent over time across models and is most pronounced for longer-maturity yields and the period after 2008.

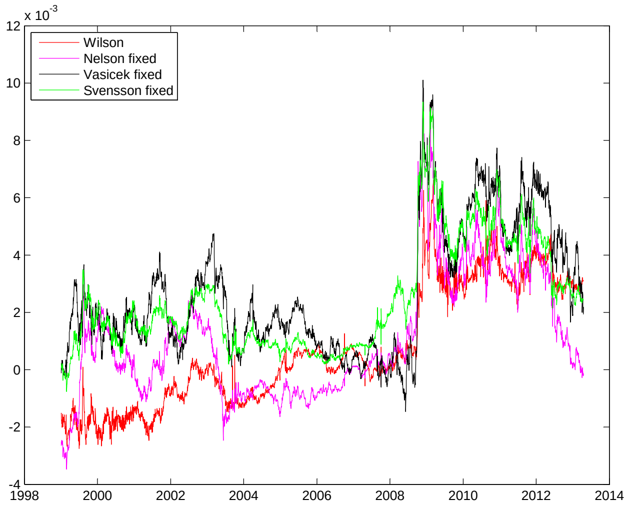
One reason that equally applies to all models is that no model is able to capture the convexity effect in the observed long end of the yield curve. The effect features a downward sloping part of the curve due to higher sensitivity of long-term swaps to interest rate changes. It usually starts beyond a maturity of 20 years yet, since the calibration considers zero rates only up to that maturity, no model is able to take it into account (see again Figure 17).

The Svensson model has been introduced exactly for capturing this curvature at the long end. Accounting for convexity led me expect a possible underestimation of the actual yield curve by the Svensson method as it pursues the downward slope indefinitely. However, given the restricted maturity inputs, instead of producing extrapolations that underestimate the actual yield curve, the Svensson model rather consistently overestimates it. Figure 22 emphasizes that the Svensson method, although theoretically able to fit it, cannot know that convexity exists. It also shows that it does not flatten fast enough or does so to an asymptotic long-run rate that is too high. Respectively, β_0 and the absolute value of β_1 are estimated to be too high for the original as well for the fixed Svensson model to fit the actual level of long end yields.

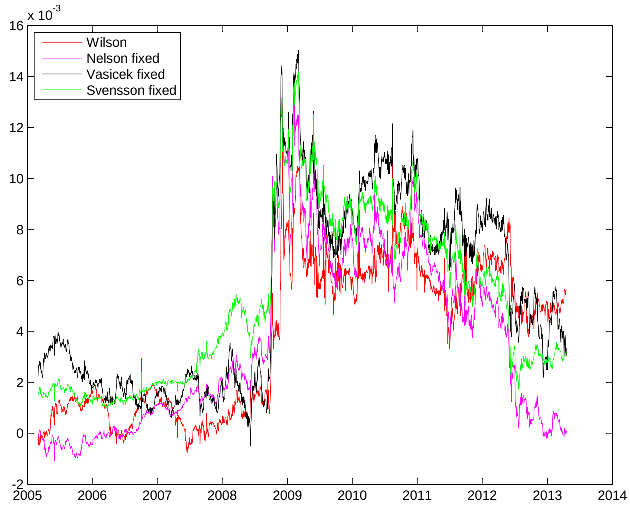
Figure 22 also exemplifies that the issue with the Svensson model's bias is even amplified in the fixed Vasicek model. Again, the estimated parameters are systematically such that an upward bias emerges. In case of the Vasicek model, it is the fixed convergence parameter itself that has led to the bias. Less variability of the fixed Vasicek yields translate into a much lower spot volatility σ from 0.0241 to 0.0197. This way, the Vasicek yield curve is pinned to the monotonically increasing yield curve condition, since the short rate is bound to be below $R(\infty) - \frac{\sigma^2}{4\alpha^2}$. Whereas the original Vasicek yield curve exhibits all three possible shapes, the fixed Vasicek model predominately produces upward sloping curves. Figure 23(a) shows that the original Vasicek model proxies the yield curve on March 8, 2013 with a humped shape but in Figure 23(b) the fixed version fits an upward sloping shape to the very same actual underlying curve. The more frequent switching of shapes with the original instead of the fixed Vasicek model supports my expectation of no bias in the Vasicek model's extrapolations. The extrapolation error test even chooses the original Vasicek model to be superior to all other models for the second subperiod, since it is the only model that does not produce rates which are systematically above the actual rates. However, putting in perspective the fact that the model is highly instable, allows for negative rates and lacks flexibility, it still does not qualify for an extrapolation model. Figure 11(d) shows that the Vasicek's 50-year yield frequently reaches -0.01 percentage points, thereby being completely unfeasible in practice. Furthermore, Figure 24 visualizes the inflexibility when the Vasicek model is unable to fit a yield curve that has two humps. The best the Vasicek model can do here is to fit a straight line. The bad in-sample fit translates into an unfavorable starting point for extrapolations, which consequently turn out to be unboundedly high.



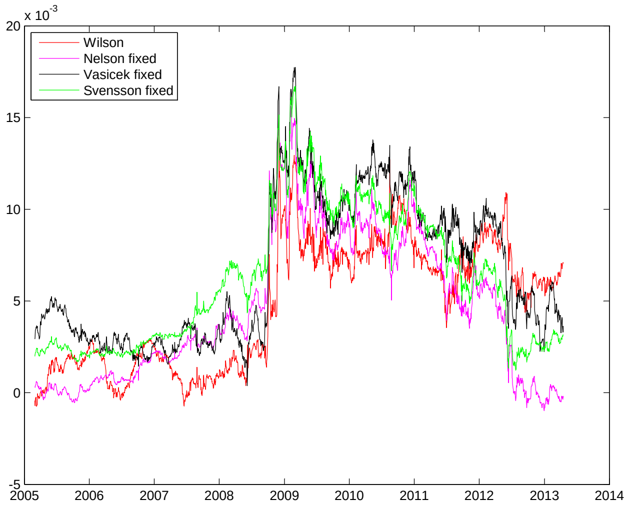
(a) 25-year yield in percentage points



(b) 30-year yield in percentage points



(c) 40-year yield in percentage points



(d) 50-year yield in percentage points

Figure 21: **Daily yield extrapolation errors of the base models**

Generally, the Vasicek model's variations illustrate a trade-off between the two extrapolation criteria - stability and fit. Whereas the original Vasicek model performs badly on the first criteria but is unbiased, the fixed Vasicek model noticeably improves on stability but starts to produce biased results. The same trade-off but in a less pronounced way holds for the Wilson specifications when the original Wilson model outperforms the longer-transition Wilson method on the stability criteria, yet, the reverse is observed during the error test. However, the finding cannot be generalized as the Svensson and the Nelson models improve on both criteria when switching from their original to their fixed versions.

It remains to clarify to what extent my expectation about the Nelson model being the least biased one is fulfilled. As all methods, the Nelson model turned out to be upward biased instead of unbiased in both its original and its fixed versions. It overestimates the actual long-term rates in situations in which the actual yield curve is upward as well as downward sloping. I expected, due to previous findings in the literature,

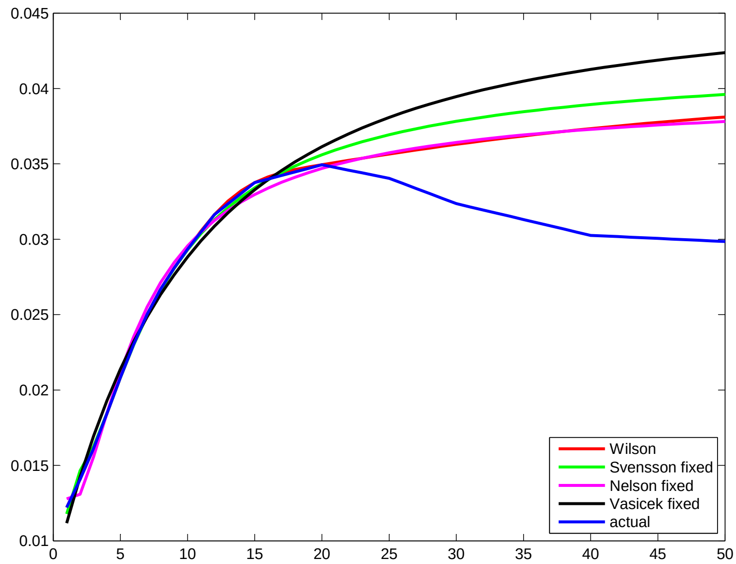
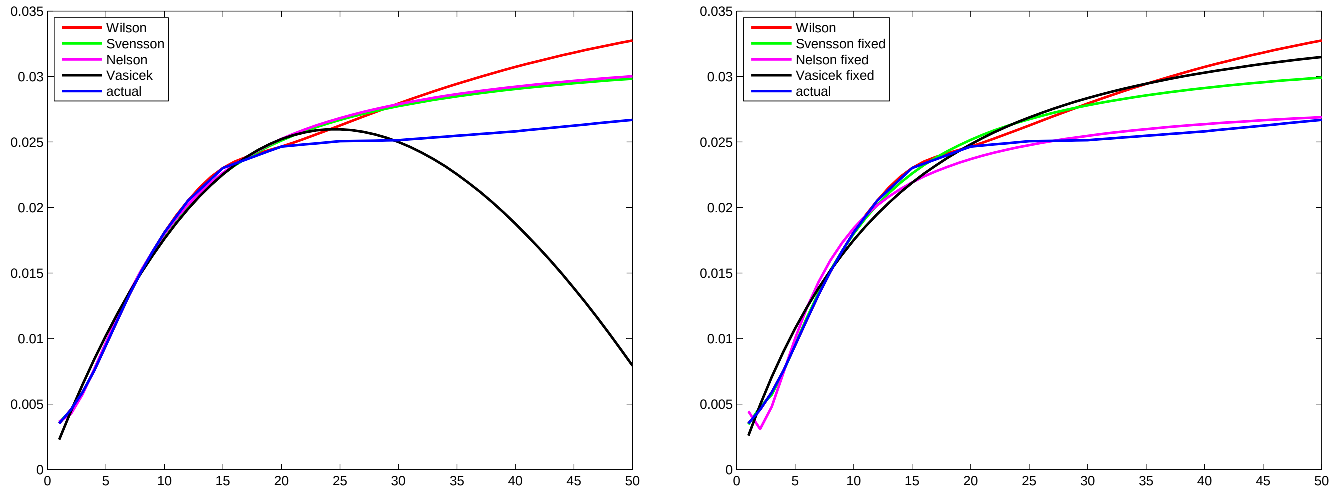


Figure 22: **Base models' yield curves and the actual yield curve on 02-July-2010**
The actual yield curve is bootstrapped from Euro swap rates. The models' yield curves are calibrated to the actual yields of maturity from one to 10, 12, 15 and 20 and extrapolated beyond. All yields are continuously compounded and reported in percentage points.

that the Nelson model would tend to underestimate rates when the yield curve is upward sloping. There are cases that accord with the expectation illustrated in Figure 25. The underestimation is, however, strongest at the longest in-sample yields and diminishes with the longest extrapolated rates. This trend can also be found in Figures 21(a), 21(c), 21(c), and 21(d) when considering the extrapolation errors separately for each maturity. Especially in the first subperiod, the Nelson model frequently produces negative errors at maturities of 25 and 30 years, but at maturities of 40 and 50 years, the underestimation abates. The figures also stand in contrast to the findings of Bolder and Stréliski (1999) who conclude the Svensson model to be less biased than the Nelson-Siegel model. In my case, it is clearly the opposite as the extrapolation errors are smaller and less positively biased for the Nelson than for the Svensson model. Moreover, the Nelson model even performs noticeably well at the end of the sample, as it is chosen by the MCS test to be the best model for the second subperiod. Hence, Proposition 2 indeed holds for the second part of the sample because the Nelson model produces the least biased extrapolations. Also, it holds that the Wilson model is most biased during the post-crisis period. The systematic overestimation of yields by the Wilson method aggravates over time until in the last two years of the sample the Wilson extrapolations perform worst in terms of the bias criteria (see Figure 21(d)). The reason for the radical change of the Wilson model from being the best in the major part of the sample and the worst in the last years is simple. The overall yield levels have heavily come down but by construction, the Wilson model cannot adjust to that beyond the last liquid point. Figure 26(a) exemplifies the perfect fit of the Wilson model until the LLP and how the model then heavily overshoots at the long end in order to reach the predetermined 4.2%. In contrast, the Nelson model almost perfectly fits the long-term rates in the end of the sample. Figure 26(b) illustrates why this is the case. First, there is no convexity effect in the actual curve on April 17, 2013, so that the overestimation due to this issue is absent. Second, the



(a) Original models' and actual yield curves on 08-Mar-2013

(b) Base models' and actual yield curves on 08-Mar-2013

Figure 23: **Upward bias in the Vasicek model when fixing α**

magnitude of the hump is lower for the Nelson than for the Svensson model, so that the Nelson model approaches to a lower asymptote. Whereas the Svensson and the Vasicek models continue to have too little flattening, the Nelson model exactly hits the actual 30- to 50-year yield level. The reason for it is that the actual yields up to maturity of 20 years moved to start out at a very low level allowing the Nelson model to fit a lower β_0 and β_1 because the sum of these two factors determines the intercept. As a consequence, the Nelson curve is less steep and approaches a lower long-run yield level, eliminating the previous overshooting. However, the Nelson model's outperformance out of sample is relativized by its underperformance in sample relative to the Svensson and the Wilson method.

Weighing the results for the stability criteria against those for the bias criteria, it seems that reducing volatility comes at the cost of creating a bias. Since only the Smith-Wilson method improves stability significantly, it is the only method that compensates the bias and, thus, qualifies as an extrapolating model.

6 Conclusion

Responding to insurance companies' call upon regulators to overthink the new European insurance directive, this research investigated the performance of different extrapolation techniques relative to the Smith-Wilson method proposed for Solvency II. The results support the finalized Solvency II framework to include the Smith-Wilson as an extrapolating model. Among the models studied, there were two Nelson-Siegel class models, the equilibrium Vasicek model, and the Solvency II's Smith-Wilson model. The latter model best mitigates the insurers' concern about undesirable fluctuations in their funding ratios - especially during periods of distressed markets such as the recent financial crisis. The reason for selecting the Smith-Wilson method is driven by this research's outcomes regarding the two research questions derived from Solvency II's extrapolation requirements. The investigation of the first question

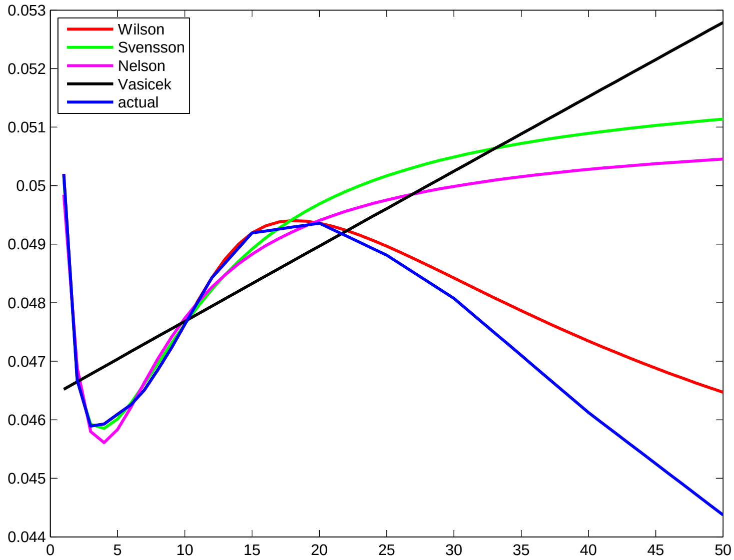


Figure 24: **The Vasicek model cannot fit a yield curve with two humps**
The actual yield curve is bootstrapped from Euro swap rates from 15-Sep-2008. The models' yield curves are calibrated to the actual yields of maturity from one to 10, 12, 15 and 20 and extrapolated beyond. All yields are continuously compounded and reported in percentage points.

aimed at finding the model that possesses the ability to significantly reduce volatility at the long end of the yield curve. A test of equal variance was employed to distinguish between the methods' relative success on that criteria. Neither the popular Nelson-Siegel model nor the Svensson method achieves a reduction of variance. The equilibrium Vasicek model even exacerbates stability. Only the Smith-Wilson model significantly stabilizes long-term rates with a maturity of 30, 40 and 50 years.

The second research question has stemmed from the new directive's approach of calculating liabilities at market instead of traditional book value by asking what method produces unbiased extrapolations relative to the actual market yields. A multiple comparison test of models' out-of sample accuracy was used to evaluate the models' relative extrapolation errors. Surprisingly, no model features bias-free extrapolations, suggesting a trade-off between the first question's stability criterion and the second question's unbiased objective. As extrapolations reduce volatility, they systematically start to overestimate long-term interest rates. However, the extent of the upward bias differs across the models with the Smith-Wilson method best compensating the two opposing criteria. The Wilson model is not only least biased but it is simultaneously the only model that is statistically proved to reduce volatility. Thus, the criticism about the Smith-Wilson model's set parameters turns invalid, as the fixing shows to be even advantageous. The Nelson-Siegel, Svensson and Vasicek methods' major issue is exactly the fact that they consist of estimated instead of fixed parameters. Whereas the Nelson-Siegel class fulfills its task of calibration by well fitting in sample, it tends to neglect reasonability of parameters that govern extrapolated yields. As the Nelson-Siegel method's extrapolations improve their stability with a fixed shape parameter, the question remains whether lowering the upper bound for the level factor during the optimization may alleviate the upward bias. To be kept in mind, however, is that tighter bounds may induce more harm than good when the model loses its flexibility in fitting well in sample. Contrarily, the Smith-Wilson method is not exposed

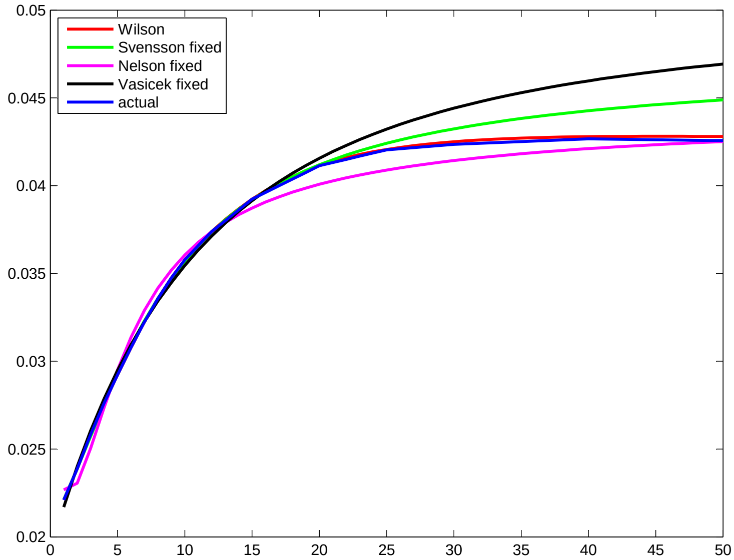
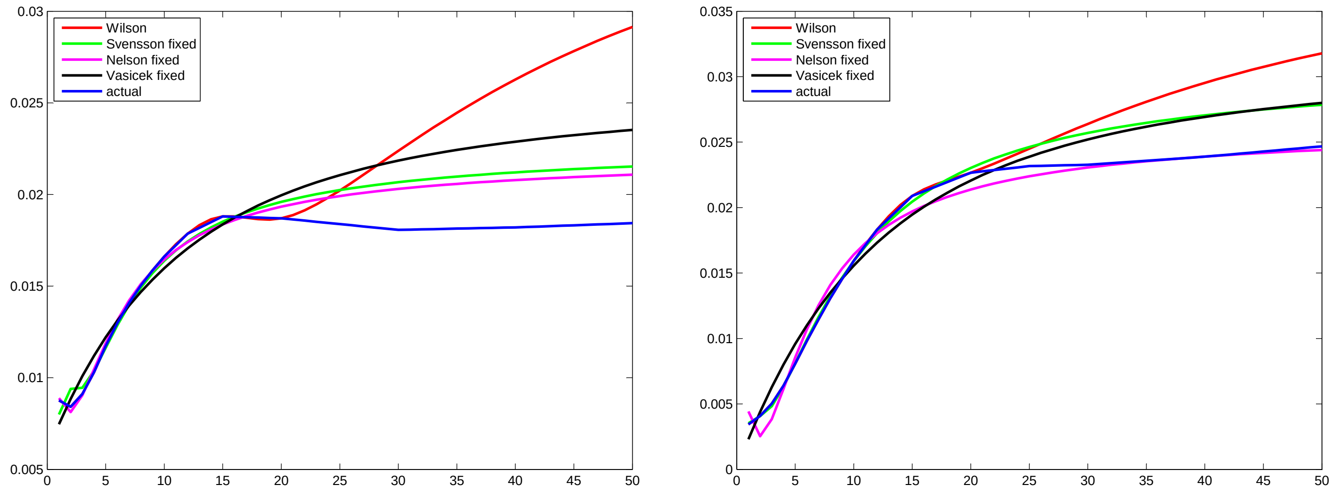


Figure 25: **Base models' yield curves and the actual yield curve on 29-April-2005**
The actual yield curve is bootstrapped from Euro swap rates. The models' yield curves are calibrated to the actual yields of maturity from one to 10, 12, 15 and 20 and extrapolated beyond. All yields are continuously compounded and reported in percentage points.

to fitting considerations. It has been introduced to especially focus on extrapolations, and that with success. Thus, extrapolations improve, the fewer estimated parameters a model encompasses. This finding specifically applies to extrapolating the yield curve, as previous studies about fitting yields found the opposite. They suggest that in-sample performance improves as the number of parameters increases. This research's analyses of extrapolations considered each key maturity separately, leading to distinct outcomes for extrapolating yields with terms of 25, 30, 40 and 50 years. Due to different availabilities of observed swap data, comparing the four classes of results among each other needs caution. Swap rates of maturity 40 and 50 years are only available since March 2005, yet those of 25 and 30 years start already in January 1999. For instance, this is likely to result in mean values of 25- and 30-year extrapolations which are higher when based on the complete and not on the shorter sample. However, no such issue exists when drawing comparisons between each specific extrapolation and its corresponding observed yield as done in both statistical tests. Another weakness arose with the implementation of the calibrations. Although the original Vasicek, Nelson and Svensson models are all fitted with the non-linear least square optimization routine from Matlab, the first is based on a code, whereas the latter two are calculated with a build-in function. For consistency reasons, using the exact same procedure instead of only the same optimization tool would have been preferable. Also, the calibrations of the fixed version of the Nelson and the Svensson models could have been conducted in a statistically more advanced way. Instead of just fixing the shape parameters, other estimation procedures aiming at less rigorously resolving multicollinearity and parameter instability may have further enhanced the models' performances. Especially the ridge or the two-step regression estimation may have led to more stable and less erroneous Nelson-Siegel extrapolations so that this model may have even outperformed the Smith-Wilson method. Despite the data and estimation limitations, the overall results still provide implications



(a) Base models' and actual yield curves on 01-June-2012

(b) Base models' and actual yield curves on 17-April-2013

Figure 26: **Performance reversion between the Wilson and the Nelson models at the end of the sample**

that are considerable for the completion of Solvency II and lead to possibilities for future research.

Regarding the findings about stability, insurance companies in Europe will benefit from stable funding ratios with the implementation of the Smith-Wilson method in Solvency II. Their concern with volatile and pro-cyclical liability values is greatly alleviated. Periods of financial market distress are already difficult per se for insurance companies that face declining asset values. An additional burden arises when they also have to cope with more fluctuating liability values as compared to normal periods. The Smith-Wilson extrapolations manage to mitigate a cyclical pattern in volatility, thereby abating the operational difficulties during crises. Those extrapolations, by being more stable, diminish interest rate risk for liability estimates as the liabilities' sensitivity to interest rate changes is less than based on observed market rates. Especially insurance companies which mainly hold long-term obligations benefit from a reduced duration, which is easier to hedge. Consequently, insurers may consider to boost sales of long-term contracts as these obligations' balance sheet risk is lower under the new directive. Less liability risk may also mean that companies are able to hold less required surplus. Since extra reserves are costly for insurers, it could happen that insurers pass on their cost reduction to policyholders in form of lower insurance premiums or higher benefits.

Insurance companies will not only benefit from stability; they will also gain from higher funding ratios. Even though the Smith-Wilson method exhibits the smallest upward bias, any overestimated extrapolations directly translate into higher discount rates, so that liability estimates decrease and funding ratios increase. The effect grows during crises and shrinks during booms. For instance, the recent financial crisis has led to strong monetary easing that has heavily brought down interest rates. Low actual long-term yields mean that the UFR and the resulting extrapolations are too high. In times of monetary contraction by the central bank, interest rates may rise to a level that exceeds the UFR, which will eventually decrease funding ratios. Thus, instead of having pro-cyclicality, the opposite emerges. The counter-cyclical pattern in funding ratios may be even tagged as an automatic stabilizer for the insurance industry. It may also spur risk-taking, as insurers know about the automatic safety net for their funding ratio. This is the opposite effect of what Solvency II aimed to achieve when postulating enhanced protection against insolvency risk for European policyholders. The systematic overestimation also stands in contrast to the market-consistent approach that targets an actual instead of an artificially reduced reflection of liability values. The adverse consequences need to be closely examined by the EIOPA to ensure that its objectives will be ultimately fulfilled and not reversed.

Several possibilities arise that may help to reduce the observed upward bias for the Smith-Wilson method. An obvious solution is to change the UFR. This can be done either temporarily or permanently. When lowering the UFR level only for limited time, it may be difficult to determine this specific time period. The UFR dynamics have to mirror the actions by central banks. However, the effects of monetary expansions on the yield curve are not clear-cut, which leads to potentially arbitrary changes of the UFR. This danger has already been brought up several times by the CRO Forum which denotes it as a regulatory risk. It is a risk that no insurance company is able to effectively hedge because it is of political and not of financial nature. It gives an extra burden to the risk management team that may even be inclined to partially neglect market risk in order to focus on regulatory risk. When managers' performance depends more heavily on political risk than on interest rate risk, they may be increasingly incentivized to lobby at the EIOPA for higher UFR levels. In order to avoid the consequences of sporadic changes in the UFR, the EIOPA may decide to reduce the UFR before Solvency II is ultimately inaugurated. A permanent reduction would be welcomed in case it becomes clear that the currently low levels of interest and inflation rates are the result of a fundamental economic change of the sort Japan has been undergoing since the 1990s. However, if it turns out not to be an ongoing condition of the economy, the lowered UFR will then be too low and the bias issue will be left unresolved. Alternatively, the different components of the UFR may be adjusted. As this research stresses a strong prevalence of the convexity effect at long-term yields, it may be beneficial to include this component in the calculation of the UFR. In order to be consistent, the value should also contain an estimate for the term premium. It is interesting for future research to see the net effect of the negative convexity component and the positive term premium. Given the strong negative effect in the actual data, the former may even outweigh the latter, leading to a reduction of the UFR without changing the estimates of real interest rate and inflation levels. The fastest and easiest way to slightly mitigate the upward bias is to change the transition to the UFR as shown with this research. An UFR convergence at 80 instead of 60 years has shown to decrease extrapolation errors. Statistically, the longer-transition specification of the Smith-Wilson method has been shown to exert significantly more accurate extrapolations. Another possibility to enhance accuracy is to extend the LLP from 20 to 30 years. According to this research's results, the extension will not adversely affect the stability criteria since the 25-year yield tends to be sufficiently liquid. However, it will bring the Smith-Wilson method closer to the market-consistency objective because the market-based curve makes up a larger proportion of the entire yield curve. Along the same lines, stronger market-based extrapolations can be achieved with the Smith-Wilson specification which is currently used by the Dutch central bank. When long-term yields are a weighted average of observed yields and Smith-Wilson extrapolations, the systematic deviation from actual yields is alleviated.

Another point that Solvency II may copy from the Dutch system is that De Nederlandsche Bank has extended the Smith-Wilson framework, thereby covering not only the insurance industry but pension funds as well. As the EIOPA already hints at implementation tests for European pension funds similar to those conducted under the Quantitative Impact Studies with insurance companies, the new directive is likely to be also introduced to the European pension funds industry in the near future. As more companies, managers, policyholders, and beneficiaries will fall under the new regulation, extrapolation techniques will become increasingly important. This research thus provides an insight into the relative performance of extrapolation methods in the European market as there has not been so far any existing study on comparative evaluations of the Smith-Wilson model with alternative methods.

Since European insurance companies also operate outside Europe, they have obligations that need to be discounted at the foreign currency's market rate. Future research may evaluate to what extent volatility in the English Pound or US Dollar swap curves is reduced with Smith-Wilson extrapolations. As the EIOPA specifies the same level of UFR for all countries except for Japan, Switzerland, and five emerging economies, extrapolations for yield curves denominated in another currency may have less or no bias compared to the Euro market.

Furthermore, this research finds that the Nelson-Siegel model has started to catch up with the Smith-Wilson method and to even outperform it in the last two years. Hence, it may be interesting to keep on observing the relative performance of these two models and to test whether the recent reversion may become permanent. If this is the case, the Nelson-Siegel model may be the preferred choice for Solvency II's extrapolations. Very likely, a continuing superiority of the Nelson-Siegel model will depend on the extent to which the recent interest rate change is fundamental or of a temporary nature only. Beyond the focus on stable and unbiased extrapolations, the Nelson-Siegel and the Smith-Wilson models compete on the amount of hedging that each method induces. Whereas the Smith-Wilson method wins on the extrapolation criteria, the Nelson-Siegel model is less expensive in terms of hedging requirements. Ongoing research in that regard may weigh the extrapolating with the hedging performance of the two models to give a broader judgment about their relative effectiveness for Solvency II.

References

Academic Community Group. (2012). *Summary of comments on eiopa cp 12/003: Draft technical specifications qis of eiopa's. advice on the review of the iorp directive: Consultation paper.* Retrieved from https://eiopa.europa.eu/fileadmin/tx_dam/files/Stakeholder_groups/opinions-feedback/20120801-EIOPA-OPSG-Opinion-CP-003-12-QIS-TS-IORPII.pdf

Anderson, N. (1996). *Estimating and interpreting the yield curve.* Wiley.

Annaert, J., Claes, A. G., De Ceuster, M. J., & Zhang, H. (2013). Estimating the spot rate curve using the nelson–siegel model: A ridge regression approach. *International Review of Economics & Finance*.

Antonio, D., Carlin, S., Hibbert, J., Holmes, C., Liu, Z., Roseburgh, D., & Sorensen, S. (2009). *A framework for extrapolation of long-term interest rates.* Retrieved from http://www.barrhibb.com/documents/downloads/A_Framework_for_Estimating_and_Extrapolating_the_Term_Structure.pdf.

Bank for International Settlements. (2005). Zero-coupon yield curves: technical documentation. *Monetary and Economic Department*. Retrieved from http://www.bis.org/publ/bppdf/bispap25.pdf

Beber, A., Brandt, M. W., & Kavajecz, K. A. (2009). Flight-to-quality or flight-to-liquidity? evidence from the euro-area bond market. *Review of Financial Studies*, *22*(3), 925–957.

Björk, T., & Christensen, B. J. (2001). Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, *9*(4), 323–348.

Black, F., Derman, E., & Toy, W. (1990). A one-factor model of interest rates and its application to treasury bond options. *Financial Analysts Journal*, *46*(1), 33-39.

Black, F., & Karasinski, P. (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, 52–59.

Bliss, R. R. (1991). Testing term structure estimation methods. *Advances in Futures and Options Research*, *9*, 197-231.

Bolder, D., & Stréliski, D. (1999). Yield curve modelling at the bank of canada. *Available at SSRN 1082845*.

Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, *69*(346), 364–367.

Cairns, A. J. (2004). *Interest rate models: an introduction* (Vol. 10). Princeton University Press.

Carver, L. (2011, March). The move away from market consistency. *Insurance Risk*. Retrieved from http://www.risk.net/insurance-risk/feature/2030027/away-market-consistency

CFO Forum and CRO Forum. (2010). *Qis 5 technical specification risk-free interest rates.*

Chief Risk Officer Forum. (2010, August). *Cro forum best practice paper- extrapolation of market data.* Retrieved from http://www.thecroforum.org/cro-forum_extrapolation-of-market-data/

Christensen, J. H. E., Diebold, F. X., & Rudebusch, G. D. (2007). The affine arbitrage-free class of nelson-siegel term structure models.

Christensen, J. H. E., Diebold, F. X., & Rudebusch, G. D. (2008). An arbitrage-free generalized nelson-siegel term structure model.

Coroneo, L., Nyholm, K., & Vidova-Koleva, R. (2011). How arbitrage-free is the nelson–siegel model? *Journal of Empirical Finance*, *18*(3), 393–407.

Cox, I. E., John C, & Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica*, *53*(2), 385–407.

Dai, Q., & Singleton, K. J. (2000). Specification analysis of affine term structure models. *The Journal of Finance*, *55*(5), 1943–1978.

De Nederlandsche Bank. (2012). *Ufr method for calculating the term structure of interest rates.* Retrieved from http://www.toezicht.dnb.nl/en/5/18/51-226790.jsp

de Haan, L., & Kakes, J. (2010). Are non-risk based capital requirements for insurance companies binding? *Journal of Banking & Finance*, *34*(7), 1618–1627.

De Pooter, M. (2007). Examining the nelson-siegel class of term structure models: In-sample fit versus out-of-sample forecasting performance. *Available at SSRN 992748*.

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

Diebold, F. X., Piazzesi, M., & Rudebusch, G. (2005). *Modeling bond yields in finance and macroeconomics* (Tech. Rep.). National Bureau of Economic Research.

Duffie, D., & Huang, M. (1996). Swap rates and credit quality. *The Journal of Finance*, *51*(3), 921–949.

Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical finance*, *6*(4), 379–406.

European Insurance and Occupational Pensions Authority. (2010). *Qis 5 risk-free interest rates - extrapolation method.* Retrieved from http://eiopa.europa.eu/fileadmin/tx_dam/files/consultations/QIS/QIS5/ceiops-paper-extrapolation-risk-free-rates_en-20100802.pdf

European Insurance and Occupational Pensions Authority. (2011, March). *Eiopa report on the fifth quantitative impact study (qis5) for solvency ii.* Retrieved from eiopa.europa.eu/fileadmin/tx_dam/files/.../reports/QIS5_Report_Final.pdf

European Insurance and Occupational Pensions Authority. (2013). *Introducing solvency ii.* Retrieved from https://eiopa.europa.eu/activities/insurance/solvency-ii/index.html

Fabozzi, F. J., Martellini, L., & Priaulet, P. (2005). Predictability in the shape of the term structure of interest rates. *The Journal of Fixed Income*, *15*(1), 40–53.

Friedman, M. (1977). Time perspective in demand for money. *The Scandinavian Journal of Economics*, *79*(4), 397–416.

Gürkaynak, R. S., Sack, B., & Wright, J. H. (2007). The us treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, *54*(8), 2291–2304.

Hansen, P. R., Lunde, A., & Nason, J. M. (2010, March). *The model confidence set* (CREATES Research Papers No. 2010-76). School of Economics and Management, University of Aarhus. Retrieved from http://ideas.repec.org/p/aah/create/2010-76.html

Ho, T. S., & Lee, S.-B. (1986). Term structure movements and pricing interest rate contingent claims. *The Journal of Finance*, *41*(5), 1011–1029.

Holborow, D. (2008). *Matlab code - vasicek yield curve fitting.* Retrieved from https://

www.quantnet.com/threads/matlab-code-vasicek-yield-curve-fitting-various-bond-price-models-available.2177/

Hull, J. (2009). *Option, futures and other derivatives.* Upper Saddle River, N.J., [etc.] : Pearson/Prentice Hall.

Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *The Review of Financial Studies*, *3*(4), 573-592.

Internal Market DG - Financial Institution Insurance. (2004, July). *Framework for consultation on solvency ii* (Tech. Rep.). European Commission. Retrieved from https://eiopa.europa.eu/fileadmin/tx_dam/.../framework_for_consult.pdf

James, J., & Webber, N. (2000). *Interest rate modelling.* Wiley.

Litterman, R. B., & Scheinkman, J. (1991). Common factors affecting bond returns. *The Journal of Fixed Income*, *1*(1), 54-61.

Litzenberger, R. H., & Rolfo, J. (1984). An international study of tax effects on government bonds. *The Journal of Finance*, *39*(1), 1–22. Retrieved from http://dx.doi.org/10.1111/j.1540-6261.1984.tb03857.x doi: 10.1111/j.1540-6261.1984.tb03857.x

McCulloch, J. H. (1971, January). Measuring the term structure of interest rates. *The Journal of Business*, *44*(1), 19-31. Retrieved from http://ideas.repec.org/a/ucp/jnlbus/v44y1971i1p19-31.html

McCulloch, J. H. (1975). The tax-adjusted yield curve. *The Journal of Finance*, *30*(3), 811–830. Retrieved from http://dx.doi.org/10.1111/j.1540-6261.1975.tb01852.x doi: 10.1111/j.1540-6261.1975.tb01852.x

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, *29*(2), 449–470.

Modigliani, F., & Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *The American economic review*, *48*(3), 261–297.

Neftci, S. N. (2008). *Principles of financial engineering.* Academic Press.

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *The Journal of Business*, *60*(4), 473-489.

Phoa, W. (1997). Can you derive market volatility forecasts from the observed yield curve convexity bias? *The Journal of Fixed Income*, *7*(1), 43–54.

Rebel, L. (2012). The ultimate forward rate methodology to valuate pensions liabilities: A review and an alternative methodology. Retrieved from http://www.netspar.nl/files/Evenementen/2012-11-09%20PD/rebel.pdf.

Schmidt, W. M. (2011). Interest rate term structure modelling. *European Journal of Operational Research*, *214*(1), 1-14.

Shea, G. S. (1984). Pitfalls in smoothing interest rate term structure data: Equilibrium models and spline approximations. *Journal of Financial and Quantitative Analysis*, *19*(03), 253–269.

Smith, A., & Wilson, T. (2001). *Fitting yield curves with long term constraints.* (Research Notes,Bacon and Woodrow. Referred to in Thomas, M. and Maré, E. (2007))

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *The Journal of Finance*, *52*(5), pp. 1973-2002. Retrieved from http://www.jstor.org/stable/2329471

Sundaram, R. K., & Das, S. R. (2010). *Derivatives: principles and practice.* New York: McGraw-Hill Irwin.

Sundaresan, S. (2009). *Fixed income markets and their derivatives.* Elsevier Science.

Svensson, L. E. (1994). Estimating and interpreting forward interest rates: Sweden 1992-1994.

Thomas, M., & Maré, E. (2007). Long term forecasting and hedging of the south african yield curve. *2007 ASSA Convention, The Actuarial Society of South Africa*.

TriOptima. (2012). *Interest rate trade repository report* (Tech. Rep.). ICAP Group Company. Retrieved from http://www.trioptima.com/resource-center/historical-reports.html

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of financial economics*, *5*(2), 177–188.

Vasicek, O., & Fong, H. G. (1982). Term structure modeling using exponential splines. *The Journal of Finance*, *37*(2), 339–348.

A Matlab code for the Smith-Wilson calibration and extrapolation

```
function W=Wilson(t,m,a)
%creates the matrix of Wilson functions see Equation 15
% each cell is calculated from Equation 11
%   t stands for the maturity in years
%   m presents the maturity of observed input instruments
%   a is the speed of transition, alpha
for i=1:length(t)
    for k=1:length(m)
W(i,k)=exp(-0.042*(t(i)+m(k)))*(a*min(t(i),m(k))-...
0.5*exp(-a*max(t(i),m(k)))*(exp(a*min(t(i),m(k)))-exp(-a*min(t(i),m(k)))));
    end
end
end
```

```
function P=DiscountPrice(t,m,p)
% Determines the zero discount price of the Smith-Wilson method, see
% equation 10
%   t stands for the maturity in years
%   m presents the maturity of observed input instruments
%   p is the observed zero discount price corresponding to maturity m
for a=0.05:0.0001:0.2
muh=exp(-0.042*m);
zeta=(p-muh)'/Wilson(m,m,a);
P=exp(-0.042*t)+(Wilson(t,m,a)*zeta');
if abs(((log(P(length(t)))-log(P(length(t)-1)))/((length(t)-1)-length(t)))-0.042)<=0.003;
    break;
end
end
```

B MCS test results for base models based on the second sub-sample

```
  ------  MODEL CONFIDENCE SET ESTIMATION  ------
Number of models:     l=4
Sample size:          n=1118
Loss function:        mse
Test Statistic:       MaxT
Resample by:          BlockBootResamp
Bootstrap parameters: B=10000 (resamples), d=2 (block length)

Model Name                      mse(*10^3)   MCS p-val.
Vasicekfixed25                    0.01057    0.0000
Nelsonfixed25                     0.00369    0.0000
Svenssonfixed25                   0.00778    0.0000
Wilson25                          0.00233    1.0000 *

Level 0.1 Model Confidence Set

Model Name                      mse(*10^3)   MCS p-val.
Wilson25                          0.00233    1.0000 *

saving output in C:\MCSout\


This output produced on 13-06-2013, 20:26:16
Time elapsed:  8.64

  ------  MODEL CONFIDENCE SET ESTIMATION  ------
Number of models:     l=4
Sample size:          n=1118
Loss function:        mse
Test Statistic:       MaxT
Resample by:          BlockBootResamp
Bootstrap parameters: B=10000 (resamples), d=2 (block length)

Model Name                      mse(*10^3)   MCS p-val.
Vasicekfixed30                    0.03031    0.0000
Nelsonfixed30                     0.01414    0.0000
Svenssonfixed30                   0.02340    0.0000
Wilson30                          0.01226    1.0000 *

Level 0.1 Model Confidence Set

Model Name                      mse(*10^3)   MCS p-val.
Wilson30                          0.01226    1.0000 *

saving output in C:\MCSout\


This output produced on 13-06-2013, 20:27:15
Time elapsed:  8.45

  ------  MODEL CONFIDENCE SET ESTIMATION  ------
Number of models:     l=4
Sample size:          n=1118
Loss function:        mse
Test Statistic:       MaxT
Resample by:          BlockBootResamp
Bootstrap parameters: B=10000 (resamples), d=2 (block length)

Model Name                      mse(*10^3)   MCS p-val.
Vasicekfixed40                    0.00152    0.0000
Nelsonfixed40                     0.00134    1.0000 *
Svenssonfixed40                   0.00145    0.0000
Wilson40                          0.00136    0.0041

Level 0.1 Model Confidence Set

Model Name                      mse(*10^3)   MCS p-val.
```

```
Nelsonfixed40                       0.00134    1.0000 *

saving output in C:\MCSout\


This output produced on 13-06-2013, 20:28:14
Time elapsed:  8.56

  ------  MODEL CONFIDENCE SET ESTIMATION  ------
Number of models:     l=4
Sample size:          n=1118
Loss function:        mse
Test Statistic:       MaxT
Resample by:          BlockBootResamp
Bootstrap parameters: B=10000 (resamples), d=2 (block length)

Model Name                        mse(*10^3)   MCS p-val.
Vasicekfixed50                      0.00158    0.0000
Nelsonfixed50                       0.00137    1.0000 *
Svenssonfixed50                     0.00149    0.0000
Wilson50                            0.00142    0.0000

Level 0.1 Model Confidence Set

Model Name                        mse(*10^3)   MCS p-val.
Nelsonfixed50                       0.00137    1.0000 *

saving output in C:\MCSout\


This output produced on 13-06-2013, 20:29:02
Time elapsed:  8.42
```