DANMARKS
NATIONALBANK

WORKING PAPERS

February 2016 | No. 106

MODELLING DANISH GOVERNMENT BOND YIELDS IN A LOW-RATE ENVIRONMENT

Nicolaj H. Christensen, Anders Nysteen and Niklas B. D. Pedersen.

Danmarks Nationalbank

DANMARKS
NATIONALBANK

Please direct any enquiries to
Danmarks Nationalbank,
Communications,
Havnegade 5,
DK-1093 Copenhagen K
Denmark
E-mail: kommunikation@nationalbanken.dk

Henvendelser kan rettes til:
Danmarks Nationalbank,
Kommunikation,
Havnegade 5,
1093 København K.
E-mail: kommunikation@nationalbanken.dk

ISSN (online) 1602-1193

DANMARKS NATIONALBANK WORKING PAPERS

MODELLING DANISH GOVERNMENT BOND YIELDS IN A LOW-RATE ENVIRONMENT

Contact for this working paper:

Nicolaj H. Christensen
Danmarks Nationalbank
nhc@nationalbanken.dk

Anders Nysteen
Danmarks Nationalbank
any@nationalbanken.dk

Niklas B. D. Pedersen
Danmarks Nationalbank
nped@nationalbanken.dk

RESUME

Modellering af danske statsrenter i et lavrentemiljø.
Vi undersøger fordelene ved at modellere udviklingen i danske statsrenter i et lavrentemiljø ved brug af en dynamisk rentemodel med en nedre grænse, en såkaldt skyggerente-model. Vi benytter en skyggerente-udvidelse af den velkendte arbitrage-fri Nelson-Siegel model. I litteraturen har skyggerentemodellen vist at kunne forbedre cross-sectional fit og renteforecasts, når renterne er tæt på nul. Vi finder imidlertid ikke de samme forbedringer ved at benytte en skyggerentemodel på dansk data. Skyggerentemodellen udfordres af, at de danske statsrenter over de seneste år er blevet stadig mere negative. Trods udfordringerne foretrækker vi stadig en skyggerentemodel frem for Gaussiske modeller, da førstnævnte fanger den asymmetriske fordeling af fremtidige renter, som karakteriserer lavrentemiljøet. Set fra et risikostyringsperspektiv finder vi, at skyggerente-udvidelsen kun påvirker løbetidspræmierne i en kort periode i begyndelsen af 2015, hvor renterne var på deres laveste. Til sidst viser vi, at en rentemodel med to frem for tre faktorer forbedrer renteforecasts i lavrentemiljøet.

ABSTRACT

Modelling Danish Government Bond Yields in a Low-Rate Environment.
We model the dynamics of Danish government bond yields in a low-rate environment using a term structure model with a lower bound, a so-called shadow rate model. Specifically, we use a shadow rate extension of the well-known Arbitrage-Free Nelson--Siegel model. In the literature, shadow rate models have been shown to improve the cross-sectional fit and the forecast performance when interest rates are close to zero. We do not, however, identify such improvements when using a shadow rate model on Danish yield data. The reason is that the shadow rate model is challenged when rates continue to decline into negative territory as has been the case in Denmark in recent years. Despite the challenges, we still prefer a shadow rate model over Gaussian models as the former captures the asymmetric distribution of future rates that characterises the low-yield environment. From a risk management perspective, we find that the shadow rate extension influences the term premia estimates notably only for a brief period in the beginning of 2015, where rates were at their lowest. Finally, as an aside we find that a term structure model with two rather than three factors improves the forecast performance in the low-rate period.

KEYWORDS

Term structure model. Shadow-rate model. Arbitrage-free Nelson-Siegel. Low-rate environment. Government debt management. Risk management.

JEL CLASSIFICATION

C32; C53; G17.

ACKNOWLEDGEMENTS

The authors wish to thank Lars Mayland Nielsen, Jacob Stæhr Wellendorph Ejsing, Kristian Loft Kristiansen, Oliver Juhler Grinderslev, Søren Lejsgaard Autrup, Kim Padkjær Abildgren, and other colleagues at Danmarks Nationalbank for valuable comments and suggestions on preliminary versions of this paper. The authors alone are responsible for any remaining errors.

Modelling Danish Government Bond Yields in a Low-Rate Environment

Nicolaj H. Christensen, Anders Nysteen and Niklas B. D. Pedersen

Government Debt Management, Danmarks Nationalbank *

Abstract

We model the dynamics of Danish government bond yields in a low-rate environment using a term structure model with a lower bound, a so-called shadow rate model. Specifically, we use a shadow rate extension of the well-known Arbitrage-Free Nelson–Siegel model. In the literature, shadow rate models have been shown to improve the cross-sectional fit and the forecast performance when interest rates are close to zero. We do not, however, identify such improvements when using a shadow rate model on Danish yield data. The reason is that the shadow rate model is challenged when rates continue to decline into negative territory as has been the case in Denmark in recent years. Despite the challenges, we still prefer a shadow rate model over Gaussian models as the former captures the asymmetric distribution of future rates that characterises the low-yield environment. From a risk management perspective, we find that the shadow rate extension influences the term premia estimates notably only for a brief period in the beginning of 2015, where rates were at their lowest. Finally, as an aside we find that a term structure model with two rather than three factors improves the forecast performance in the low-rate period.

*Electronic address: `governmentdebt@nationalbanken.dk`; URL: `http://www.nationalbanken.dk/en/governmentdebt`

I. INTRODUCTION

The Danish Government Debt Management Office is responsible for the management of the central government debt. The overall objective of the debt policy is to cover the financing requirement at the lowest possible long-term borrowing costs, while taking the degree of risk into account. The expected costs and risk of a given debt structure is evaluated using various measures. One of such measures is a Cost-at-Risk (CaR) model, which can be interpreted as the issuer-side equivalence of the popular Value-at-Risk framework used by many asset managers.

Forecasting the dynamics of future interest rates is a central input into the CaR model. The interest rate projections are generated in a term structure model that aims at capturing the yield dynamics across maturities and throughout time. Many of the term structure models proposed in the literature rely on a Gaussian framework as this eases the model implementation.[1] A popular model within the Gaussian framework is the Arbitrage-Free Nelson–Siegel (AFNS) model proposed by Christensen et al. [7].

The low-rate environment presents some challenges for the Gaussian models. Most notably, Gaussian models, including the AFNS-model, assume that interest rate projections are symmetrically distributed. This is most likely not the case when rates are close to zero as rate increases are unbounded whereas the scope for large rate declines is restricted to some extent, resulting in an asymmetric distribution.[2]

Another challenge is that the AFNS-model assumes that interest rate volatility is constant over time. This has been a particularly strict assumption in the period with low rates since the short end of the curve has experienced a large decline in volatility, cf. Figure 1. One explanation of the decline in rate fluctuations is that market participants had expectations of short rates staying low for long. As the AFNS-model fails to account for the decline in volatility, it risks projecting too wide a distribution of future interest rates, especially in the

[1] Another popular model has been the Cox-Ingersoll-Ross model that allows for time-varying volatility, see Cox et al. [8].

[2] Breeden and Litzenberger [5] show that when rates have been close to zero, the distribution of expected future rates becomes asymmetric as market participants attach greater likelihood of rate increases over further rate declines. One reflection of the asymmetric interest rate distribution is that central banks to a large extent abstained from further reducing policy rates after they had hit near-zero levels and instead focused on unconventional monetary measures as forward guidance and quantitative easing.

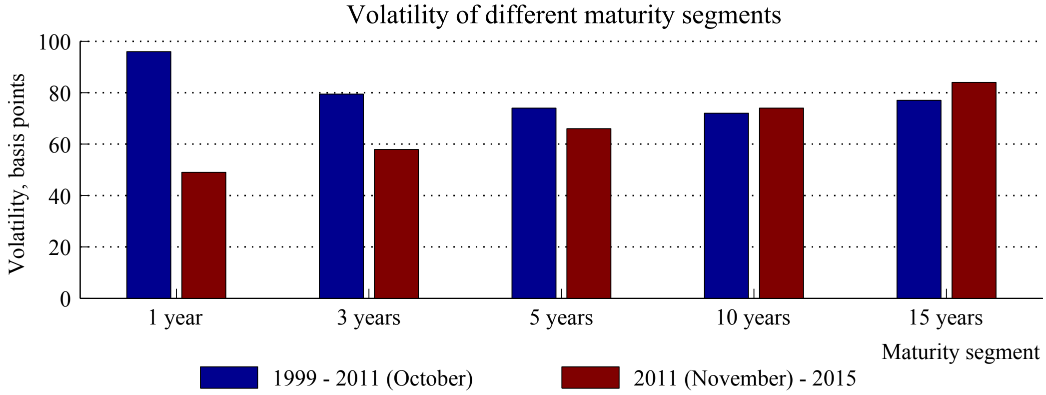

Figure 1: Volatility of Danish government bond yields with different maturities, before and during low-rate period. The volatility is calculated as the (annualised) standard deviation of daily zero-coupon bond yield changes.

short end of the yield curve. These challenges are well-known in the literature and have been addressed by introducing a lower bound in the models. This results in a so-called Shadow Rate (SR) model, which was originally proposed by Black [4], see Section III for details on the implementation. Choosing a lower bound for the Danish case is described in Section IV.

Motivated by the low-yield environment, there has been an increasing amount of work on SR-models in recent years. For example, Christensen and Rudebusch [6] extend the popular AFNS-model with a lower bound and find that this improves the model forecast of Japanese government yields. Also, Bauer and Rudebusch [2] and Lemke and Vladu [16] find improved forecast performance by using a SR-model on U.S. government bond yields and euro swap rates, respectively. Since Danish government bond yields have been near zero or negative for a number of years, we find it interesting to investigate whether the benefits of using a SR-model also apply to the Danish case.

A SR-model has the most relevance when the short end of the yield curve has been stable around a certain lower interest rate level, e.g. zero per cent. To a large extent this has been the case in the U.S. and Japan but not in Denmark where short rates, at times, have moved deeply into negative territory.[3] The Danish interest rate development gives rise to some challenges for a SR-model, and as a result we do not achieve the usual improvements of

[3] The monetary policy in Denmark is aimed at keeping the exchange rate of the Danish Krone stable against the euro. Due to a large inflow of currency in early 2015 the leading monetary policy rate was reduced to -75 basis points. The policy rate is currently at -65 basis points.

cross-sectional fit and forecast performance, as outlined in Section V. Despite the challenges, we still prefer a SR-model over a Gaussian model as the former captures the asymmetric distribution of future rates that characterises the low-yield environment.

II. DATA

Choice of input data

We use zero-coupon yields of Danish government bonds for the period January 1999 to December 2015.[4] A sufficient span of maturities should be included in order to capture the full dynamics of the interest rate development. We therefore use nine different maturities (3 and 6 months, 1, 2, 3, 5, 7, 10 and 15 years). As we use end-of-month data we end up with 204 data points for each maturity segment.

Yield curve data

Two characteristics of the Danish government yield development can be highlighted. First, yields have been on a clear downward trend during the period considered, cf. Figure 2. Second, yields of different maturities have tended to move in tandem. The high degree of correlation across maturities motivates that the variation in yields can be captured by a few common factors.[5] Based on a principal component analysis, it turns out that most of the variation can be captured by two common factors, although a third factor has useful contribution in some periods, cf. Figure 3. Against this background, we consider term structure models with both two and three factors in the following.

[4] Zero-coupon yields are extracted from Scanrate's RIO using an extended Nelson–Siegel approach.

[5] Other financial assets also exhibit a factor structure. For example, in the Capital Asset Pricing Model the expected return of a stock is assumed to be driven by one factor (market risk).

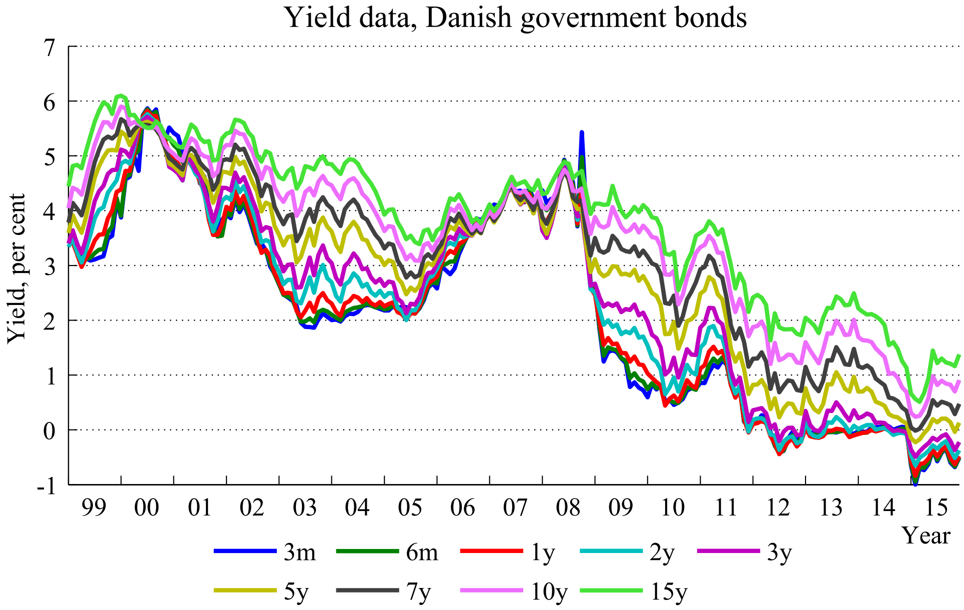

Figure 2: Zero-coupon yields of Danish government bonds in the period from January 1999 to December 2015, end-of-month observations. The considered maturities are 3 and 6 months, 1, 2, 3, 5, 7, 10 and 15 years.

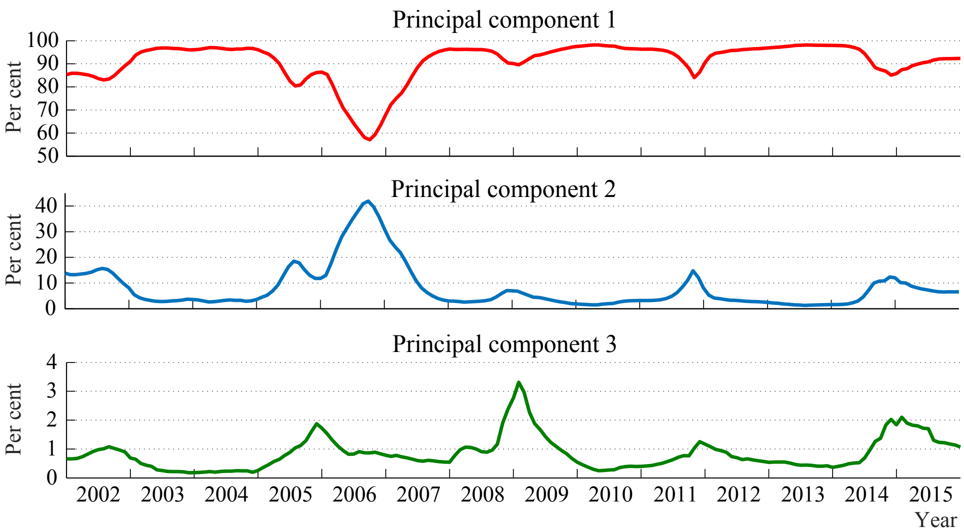

Figure 3: Principal components are linearly independent variables that are obtained by an orthogonal transformation of the yield data. The fraction of variance described by each principal component is illustrated over time. The estimation is performed on a rolling window of 36 months. In sum, the three factors explain 99.7 per cent of the variation across the entire period.

III. THE MODEL

In this section we outline the well-known framework of the AFNS-model. We use this to introduce a lower bound, which results in a SR-model. The notation is based on a three-factor model but can be directly applied to a two-factor model by leaving out the third factor.

The AFNS-model: A factor model with no lower bound

Nelson and Siegel [17] proposed a method to fit a smooth curve to a set of yields based on four parameters. Diebold and Li [9] made the parameters time-varying, thereby making it possible to estimate the evolution of yield curves over time. Although the Diebold–Li model has some clear benefits, it has one drawback; it does not rule out opportunities for arbitrage. This issue was addressed by Christensen et al. [7], who introduced an Arbitrage-Free Nelson–Siegel model, the AFNS-model, by combining the Diebold–Li model with the theoretical framework proposed in Duffie and Kan [10]. Specifically, they identify a closed-form expression of the zero-coupon yield, which contains an adjustment term that makes the model free of arbitrage by imposing a specific structure on the factor dynamics under the risk neutral measure, $\mathbb{Q}$. Appendix A lays out the details for our implementation of the AFNS-model.

At a given time, t, the AFNS-model defines the zero coupon yield for a given maturity measured in years, τ, by three factors, X_t^i, and a yield-adjustment term, $A(\tau)/\tau$,

$$y_t(\tau) = X_t^1 + X_t^2 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + X_t^3 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) - \frac{A(\tau)}{\tau}. \tag{1}$$

Each factor is weighted by a factor loading, given as the coefficients of each X_t^i. In the AFNS framework, the factor loadings are time-independent as they only depend on the maturity and a model parameter, λ. The three factors are usually interpreted as the level, slope, and curvature of the yield curve, cf. Figure 4.[6] The yield-adjustment term is given by a closed-form solution, cf. Christensen et al. [7], see Appendix A for details.

[6] An alternative interpretation is that the factors control the long, short and medium end of the yield curve, respectively. Factor one loads equally on all maturities and can therefore be interpreted as determining the the long end of the curve. The loading on factor two is declining in τ and therefore impacts the shorter end of the curve the most. Factor three loads the most on medium-term yields.

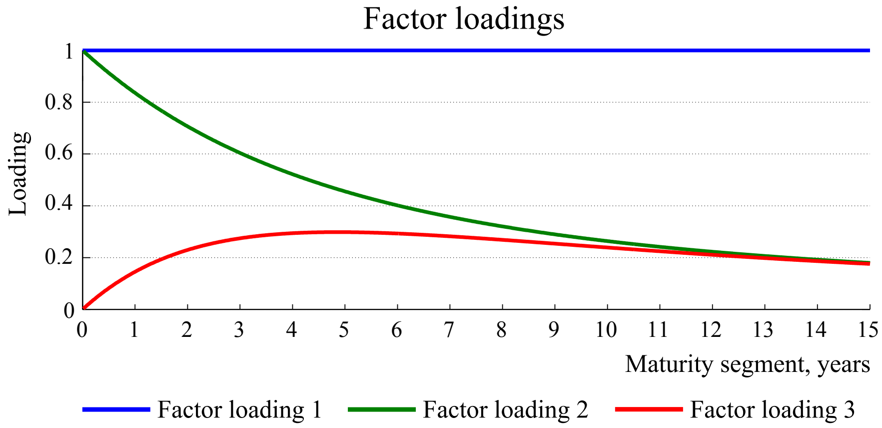

Figure 4: Factor loadings in a three-factor AFNS-model as a function of the maturity, shown for $\lambda = 0.37$. The factor loadings are given as the coefficients of each X_t^i in equation (1).

The instantaneous short rate, $s_t = y_t(0)$, is given as the sum of factor one and two,

$$s_t = X_t^1 + X_t^2. \tag{2}$$

In the AFNS framework we choose to consider the case of independent factor dynamics under the physical measure $\mathbb{P}$,[7]

$$d\bar{X}_t = \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_{22}^{\mathbb{P}} & 0 \\ 0 & 0 & \kappa_{33}^{\mathbb{P}} \end{pmatrix} \left[\begin{pmatrix} \theta_1^{\mathbb{P}} \\ \theta_2^{\mathbb{P}} \\ \theta_3^{\mathbb{P}} \end{pmatrix} - \bar{X}_t \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} d\bar{W}_t^{\mathbb{P}}, \tag{3}$$

where the factors are given in vectorised form $\bar{X}_t = (X_t^1, X_t^2, X_t^3)^T$. For the independent specification (diagonal κ- and σ-matrices), each factor is reverting toward a long-run mean level, $\theta_i^{\mathbb{P}}$, with a speed of $\kappa_{ii}^{\mathbb{P}}$, and $\bar{W}_t^{\mathbb{P}}$ is a Wiener process with standard deviation σ_{ii}, $i \in \{1, 2, 3\}$. An independent factor structure is implemented here, as it eases the numerical optimisation of the model since only ten parameters must be estimated, $\{\kappa_{11}^{\mathbb{P}}, \kappa_{22}^{\mathbb{P}}, \kappa_{33}^{\mathbb{P}}, \theta_1^{\mathbb{P}}, \theta_2^{\mathbb{P}}, \theta_3^{\mathbb{P}}, \sigma_{11}, \sigma_{22}, \sigma_{33}, \lambda\}$. More general model specifications with correlated factor dynamics (i.e. non-zero off-diagonal elements in the κ- and σ-matrices) are not considered, as they tend to improve the in-sample fit at the cost of out-of-sample performance due to over-fitting, cf. Christensen et al. [7]. The parameters are estimated using a Kalman filter, see Appendix A for details.

[7] In order to obtain the factor dynamics under the physical measure the Girsanov Theorem is used under the assumption of an affine specification of the risk premium.

Extending the AFNS model with a lower bound

We now introduce a lower bound for the short rate development using the framework presented in Black [4]. In the SR-model, the introduction of a lower bound implies that yields are restricted from moving below a certain level, $r_{\min}$. The actual short rate, r_t, is thus given as the maximum of the lower bound, $r_{\min}$, and the underlying shadow rate, s_t,

$$r_t = \max(r_{\min}, s_t). \tag{4}$$

Following Christensen and Rudebusch [6] we assume that the shadow rate has the same dynamics as the short rate in the AFNS-model.

A stylised example can help illustrating the potential benefit of introducing a lower bound. When the short end of the yield curve is flat, the AFNS-model can have difficulties in fitting the observed yields, cf. Figure 5 (left). In such a yield environment, a lower bound can provide additional flexibility so that the long end of the curve can be fitted without compromising the fit of the short end of the curve, cf. Figure 5 (right).[8] A side effect of truncating the short end of the yield curve is that the distribution of future rates becomes narrower, which results in lower model-implied yield volatility.[9]

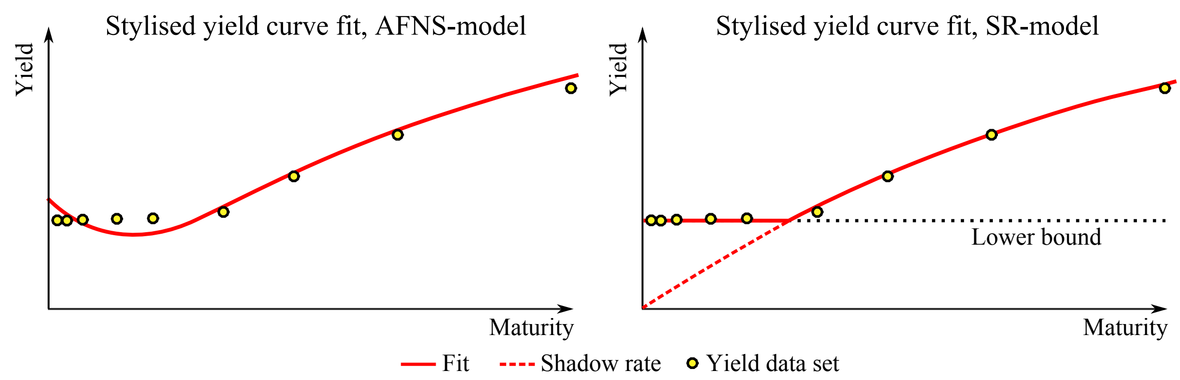

Figure 5: Stylised yield curve fit of the AFNS-model (left) and the SR-model (right) when the short end of the yield curve is flat.

[8] The benefit of the lower bound becomes smaller if the short end is above the lower bound and if the short end of the curve is not flat.

[9] The shadow rate is used as the starting point when simulating future interest rate paths. If the shadow rate is close to or below the lower bound, several of the simulated paths will move below the lower bound and will therefore be truncated due to the max-function in equation (4). Truncation of the interest rate

Estimation of the parameters in a SR-model is more complicated because yields are no longer given on a closed-form as in the AFNS-model. Therefore, approximations have been proposed to directly estimate bond prices in the SR-model.[10] We use the approximation proposed by Krippner [12], who use an option based approach, see Appendix A for details on the implementation. Using the Krippner approximation is quick but not completely accurate. Priebsch [18] provides a more accurate approximation, but since it involves solving a doublet integral, it is considerably more time-consuming to estimate the model parameters. We have implemented both approximations and we find that the estimation error in the Krippner approximation is small and we therefore use this method throughout. A comparison of the two approaches is given in Appendix C.

paths will result in less variation and therefore a decline in volatility.

[10] Yields can still be estimated using standard Monte Carlo simulations, but since the parameter estimation is done through an iterative process, a simulation approach is very time-consuming.

IV. INTRODUCING A LOWER BOUND

The introduction of a lower bound is motivated by the fact that investors, in principle, have the option to hold cash rather than investing in bonds with a negative yield. The existence of an asset that yields a return of zero (i.e. cash), however, does not rule out negative interest rates. This is because holding large amounts of cash is costly in terms of storage and security as well as inconvenient in terms of carrying out transactions over large geographical distances. The costs and inconvenience of holding large amounts of cash can help explaining why demand for cash has not increased materially in Denmark even though interest rates have been well below zero for a while.[11]

In recent years there has been much debate about the extent to which a lower bound exists. One point of view is that the lower bound can be eliminated by reducing the value of holding cash, for example by introducing some form of taxation on paper currency.[12] This potential policy tool has not been used so far. Instead, central banks have turned to unconventional monetary measures such as quantitative easing (QE) after policy rates hit near-zero levels. The preference of using policy measures such as QE over conventional rate reductions indicates that there may be some form of lower bound with respect to how low rates can go.

In the implementation of a SR-model a specific lower bound level must be defined. This is of course a challenging task since the lower bound is unknown and cannot as such be defined as a definite value. If the chosen level is too low, the model will not be able to use the lower bound effectively. On the other hand, if the lower bound is too high, several of the observations will fall outside the distribution of forecasted yields, thereby systematically deteriorating the forecast performance.

In practice, one approach is to use the lowest observed yield as the lower bound, which for our data set is a 3-month rate of -100 basis points. Another approach is to use the level that optimises the one-month forecast ability of the model. For the three-factor model, this value is around -70 basis points, cf. Figure 6. Since yields have been below -70 basis points, this lower bound level implies that several observations will fall outside the distribution of future

[11] See Jensen and Spange [11] for a further discussion of the demand for cash at negative interest rates.

[12] For example, Agarwal and Kimball [1] propose that the lower bound can be eliminated by taxing paper currency by using a time-varying paper currency deposit fee between private banks and the central bank.

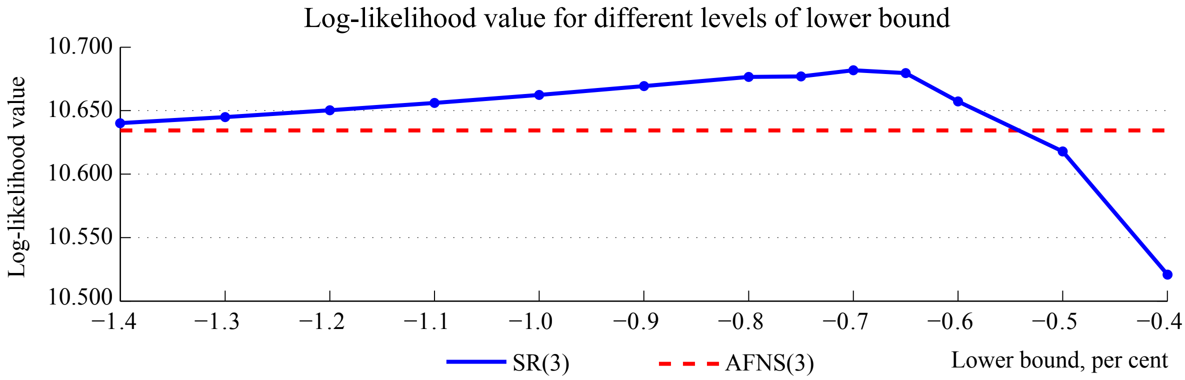

Figure 6: Log-likelihood value for various lower bound levels in the three-factor SR-model, compared to the log-likelihood value obtained with the three-factor AFNS-model. The log-likelihood value measures the one-month in-sample forecast performance of the model. The estimation period is January 1999 to December 2015.

rates. A third approach could be to use a lower bound level that varies over time. A floating lower bound can potentially provide more flexibility, but systematic forecast errors will still arise as Danish yields have moved to new low levels more than once. We prioritise not being continuously surprised when forecasting, and therefore choose to implement a constant lower bound equal to the lowest yield observed throughout the data period, i.e. a lower bound of -100 basis points.

V. PERFORMANCE OF THE MODELS

In this section we describe the performance of the SR-model with a focus on risk management implications. In the Danish Debt Management Office, interest rate models are primarily used for projecting paths of future interest rates, which are used as an input to evaluating the expected costs and risk of a given issuance and swap strategy. The expected cost of a particular issuance and swap strategy is closely linked to estimations of term premia. At time t, the term premium (TP) for a given maturity τ can be defined as the difference in expected costs from issuing a bond with τ years to maturity compared to rolling over a very short bond,

$$\mathrm{TP}_t(\tau) = y_t(\tau) - \frac{1}{\tau}\int_0^{\tau} r_{t+T}\,\mathrm{d}T. \tag{5}$$

At a given point in time the only unknown in calculating the term premium is the average path of the future short rate, r_t. To estimate the future short rate, we use the different term structure models proposed in Section III. The resulting term premia are shown in Figure 7, from which two main features can be highlighted. First, the shadow rate extension of the three-factor model only has substantial impact on term premia estimates during the first half of 2015, where yields were close to the selected lower bound. Second, the term premia estimates vary considerably between the two- and three-factor models. In the following, we split the analysis into two separate parts: First, we compare the results of a SR-model relative to an AFNS-model. Second, we examine the difference between using a two- and a three-factor model.

Impact from the shadow rate model

As shown in Figure 7, the introduction of a lower bound results in term premia estimates that are higher as well as lower compared to the estimates in an AFNS-model. This is because the introduction of a lower bound influences the projected short rate path (and with that the estimated term premium) in two ways.[13] First, the SR-model implies a slower reversion toward the long term interest rate level. The direct effect of a slower mean reversion is to

[13] When discussing the impact of the SR-model we only consider the three-factor models as the conclusions are qualitatively equivalent for the two-factor models.

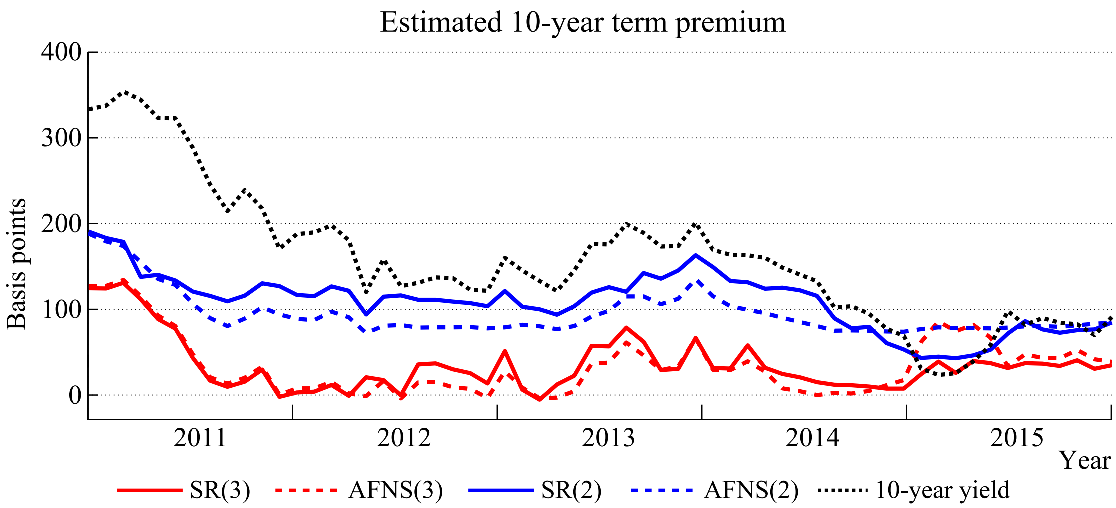

Figure 7: Estimated 10-year term premium for the two- and three-factor models. The actual 10-year yield is included for comparison. The numbers in the parentheses indicate the used number of factors.

drag down each individual simulated short rate path. This is reflected in a more conservative short rate *median* path in the SR-model compared to the AFNS-model, cf. Figure 8. When considering the *mean* path, this reversion is countered by the impact of truncating a large share of the negative yield paths with the SR-model, which will increase the *mean* path of the short rate.

Over the low-rate period the two counter-acting effects have neutralised each other to a large extent. As a result, the term premia estimates of the SR-model have been close to the estimates of the AFNS-model. The largest divergence between the term premia estimates is observed in early 2015, where the SR-model estimates are well below those of the AFNS-model. The primary reason is that the AFNS-model in some cases projected very negative rates in this period, which resulted in a low mean path of future rates, thereby increasing the AFNS term premia estimates.

For the term premia estimates to be considered reliable, the models must, among other things, be able to perform reasonable forecasts of future rates. We therefore analyse the out-of-sample forecast performance of the two models, see Figure 9. Due to the inherent mean reversion of the models, they over-shoot the continued downward trend of the realised

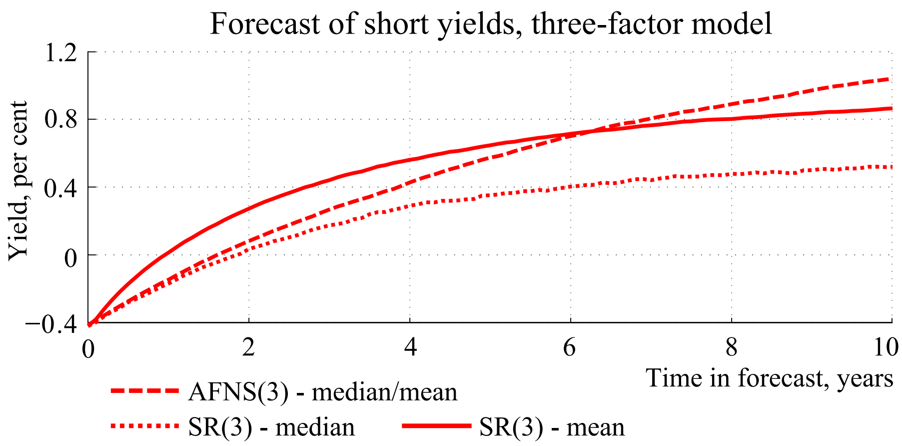

Figure 8: Mean and median paths of the short rate in the three-factor AFNS- and SR-model. Based on 50,000 simulations using December 2015 as the starting date. The mean path is above the median path of the SR-model due to the truncation of the lowest rates.

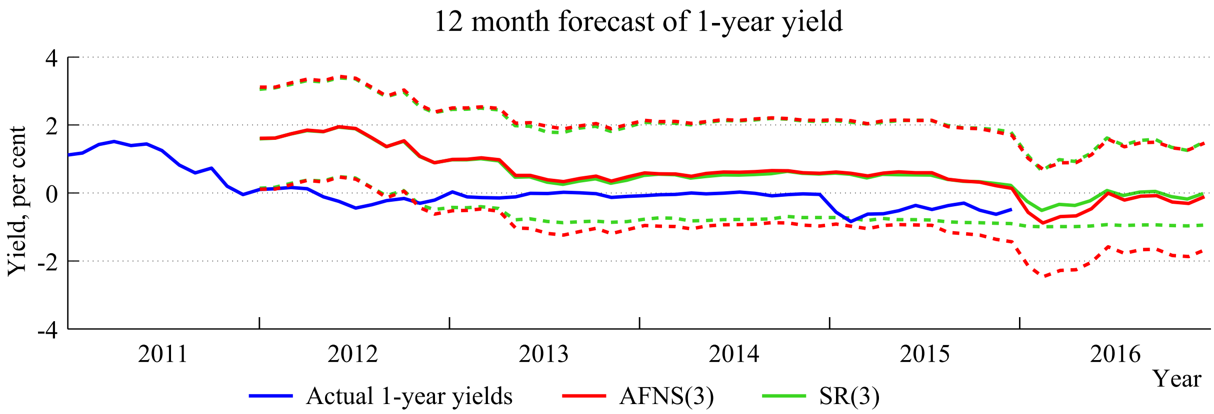

Figure 9: 12 month forecasts of the 1-year yield compared to the actual yield (blue line). For the forecasts, the mean (solid line) and the 5th and 95th percentiles (dashed lines) are indicated at the time where the forecast is realised, i.e. 12 months after the forecast was made. The analysis is done for the last 60 months of the data set using real-time parameter estimates based on a fixed starting date of January 1999.

yield development.[14] Based on the average forecast error, we find that the introduction of a lower bound does not improve the forecasts significantly, cf. Table I. This contrasts with the forecast improvements from using a SR-model on U.S. and Japanese data by Bauer

[14] The issue of too aggressive mean reversion can potentially be addressed by using a small-sample bias correction as proposed by Bauer et al. [3]. We leave this for future reasearch.

and Rudebusch [2] and Christensen and Rudebusch [6], respectively. The lack of forecast improvements in the Danish case can be explained by the fact that Danish yields have been at their lowest only briefly, whereas U.S. and Japanese short rates have remained around the same lower bound for a longer time. As a consequence, the lower bound plays a small role in the Danish case although rates have been very low.

Model / Maturity		3m	6m	1y	2y	3y	5y	7y	10y	15y	Mean RMSE
AFNS(3)	3 mo.	43	42	42	45	48	52	51	48	44	46
	6 mo.	65	66	70	77	83	88	87	82	75	77
	12 mo.	102	105	112	123	131	137	137	130	119	122
SR(3)	3 mo.	45	43	41	43	46	49	49	46	43	45
	6 mo.	67	67	69	75	80	84	83	78	73	75
	12 mo.	104	106	111	121	127	133	132	126	116	120

Table I: Root mean square of the deviations between the forecasted yields and the realised yields, given in basis points. The forecasts are made for time horizons of 3, 6, and 12 months (rows), and are made across all maturities (columns). The forecast period is 60 months. Thus we compare 57, 54 and 48 forecasts for each maturity segment of the 3, 6 and 12 month forecasts, respectively.

Benefits of using a SR-model

Although term premia estimates and forecast performance are only affected to a small degree by the introduction of the lower bound, there are still benefits from using a SR-model compared to an AFNS-model in the Danish low-yield environment. First, the SR-model distributes the future rates asymmetrically and in a reasonable way above the lower bound, cf. Figure 10.[15] Second, the SR-model is better than the AFNS-model in capturing the decline in interest volatility in the low rate period, although the divergence between realised and model-implied conditional volatility is still significant, cf. Figure 11.

In sum, we prefer the SR-model to the AFNS-model as it contains the expected asymmetry of the future rate distribution, and as it, to some extent, replicates the decline in volatility

[15] As noted in the introduction, Breeden and Litzenberger [5] show that when rates are low, the distribution of expected future rates becomes asymmetric as rate increases are more likely than further rate declines.

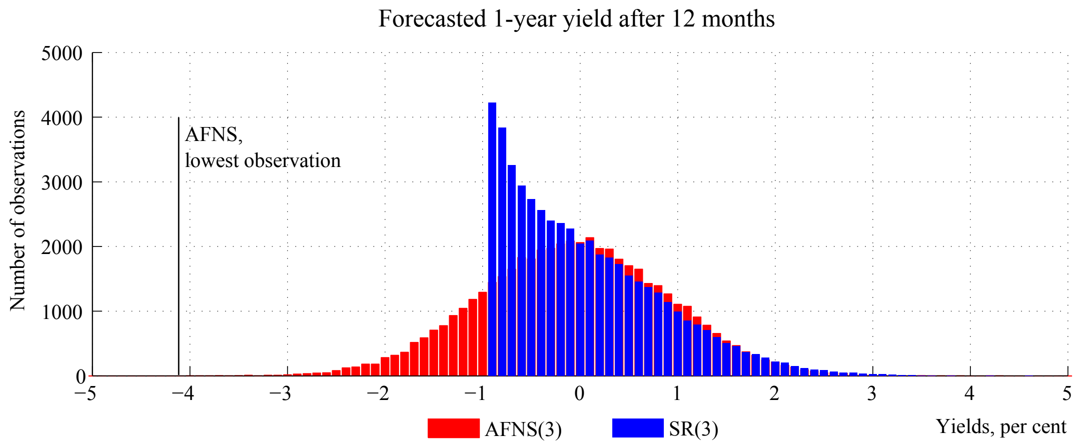

Figure 10: Histogram of the 12 month forecasts of the 1-year yield, starting from December 2015 and using 50,000 simulated paths.

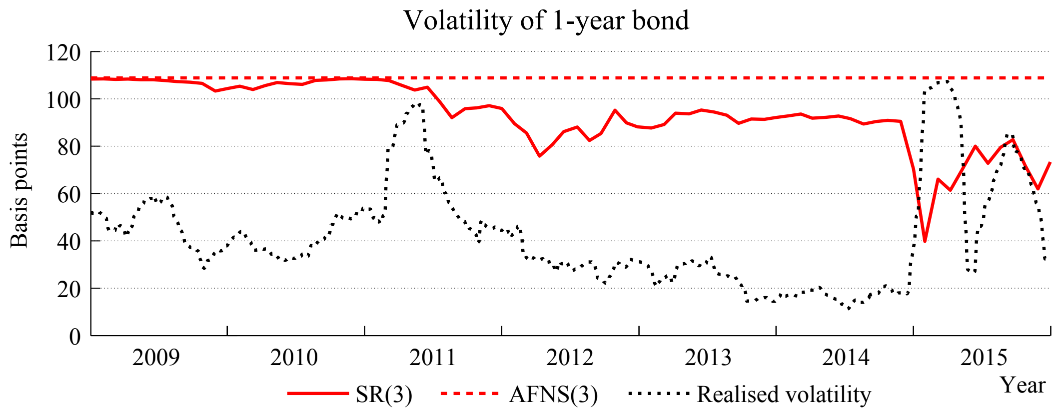

Figure 11: Model-implied conditional volatility for the three-factor SR-model (solid line) and the AFNS-model (dashed line) compared to the realised volatility of the changes in 1-year rates based on a 3-month rolling estimation window. Volatility is expressed as annualised standard deviation.

of the short end of the curve. However, we do not find that the SR-model is better at forecasting.

Impact from number of model factors

In most periods, the term premia estimates from the two-factor models are well above the estimates based on the three-factor models. This is because the two-factor models project that the short rates will stay low for a longer period. For example, at the end of 2015 the two-factor SR-model forecasts that the short rate on average stays negative for around

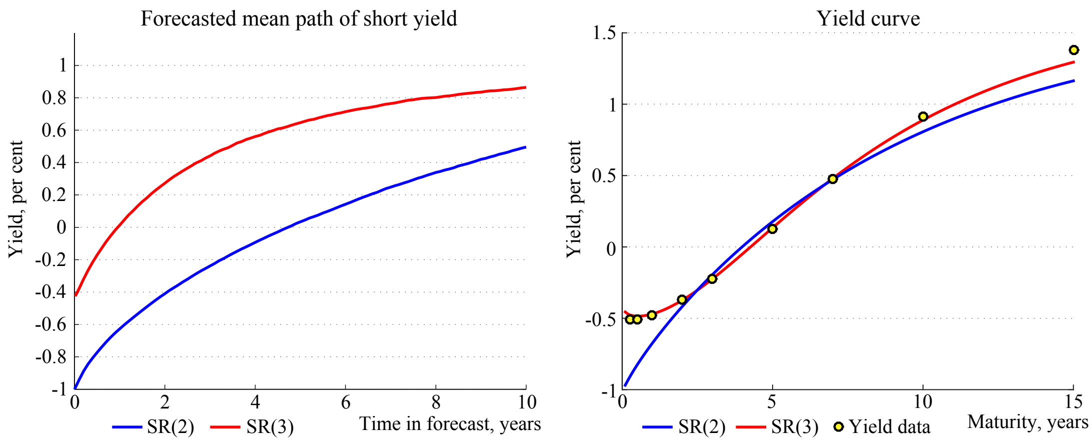

Figure 12: Left: Mean path based on 50,000 forecasts of the short yield, forecasted from December 2015. Right: Yield data (yellow bullets) and estimated yield curves at December 30, 2015.

five years, cf. Figure 12 (left). This is much more conservative than the projections of the three-factor SR-model, which explains the lower term premia estimates of the three-factor models.

From a model point of view, the difference in term premia estimates can mainly be explained by how the models fit the very short end of the yield curve. As an example, a fitted yield curve from end December 2015 is shown, cf. Figure 12 (right). Due to the smaller number of factors, the two-factor model considerably underestimates the short end of the curve. A lower starting point will, all else equal, result in a lower average mean path and thereby a higher term premia estimate. Figure 13 shows that the under-shooting of the short rate is a general issue of the two-factor model. In contrast, the three-factor model fits the short end of the yield curve better, although there may be signs of an over-fitting issue, which implies that the model will tend to overshoot the short rate.[16]

Next, we compare the forecast performance of the two- and three-factor models, cf. Table II. The two-factor model captures the downtrend in the Danish yields better than the three-factor model during forecast periods of 3, 6, and 12 months. The primary reason is a slower reversion for the mean path of the short rate in the two-factor model during the first years of forecasting as illustrated in Figure 14.

[16] Krippner [13] finds the same on U.S. data and designates the problem as a sign of over-fitting. It should be noted that it is important to include short maturities in a three-factor model to limit the over-shooting issue in the short end.

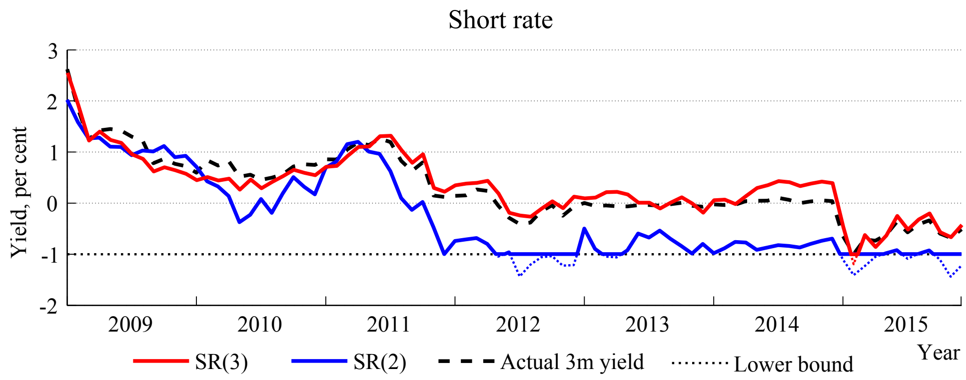

Figure 13: Short rate in a SR-model with two and three factors, respectively. The short rate is defined in equation (4). The underlying shadow rate of each model is illustrated by a dotted line. The actual three-month zero-coupon yield is included to illustrate the divergence to the modelled short rates.

Model / Maturity		3m	6m	1y	2y	3y	5y	7y	10y	15y	Mean RMSE
SR(3)	3 mo.	45	43	41	43	46	49	49	46	43	45
	6 mo.	67	67	69	75	80	84	83	78	73	75
	12 mo.	104	106	111	121	127	133	132	126	116	120
SR(2)	3 mo.	49	40	31	32	37	41	39	37	37	38
	6 mo.	47	43	44	53	61	65	64	61	61	55
	12 mo.	58	61	69	84	93	99	98	95	93	83
Random walk	3 mo.	32	31	31	32	34	36	37	38	39	34
	6 mo.	43	44	46	50	54	58	60	61	61	53
	12 mo.	61	62	66	73	79	87	90	92	90	78

Table II: Root mean square of the deviations between the forecasted yields and the realised yields, given in basis points. The forecasts are made for time horizons of 3, 6, and 12 months (rows), and are made across all maturities (columns). The forecast period is 60 months. Thus we compare 57, 54 and 48 forecasts for each maturity segment of the 3, 6 and 12 month forecasts, respectively. The SR(3) results from table I are included for comparison.

Although forecast performance is important, one should not solely rely on this measure when evaluating interest rate models. To illustrate this point, it can be seen from Table II

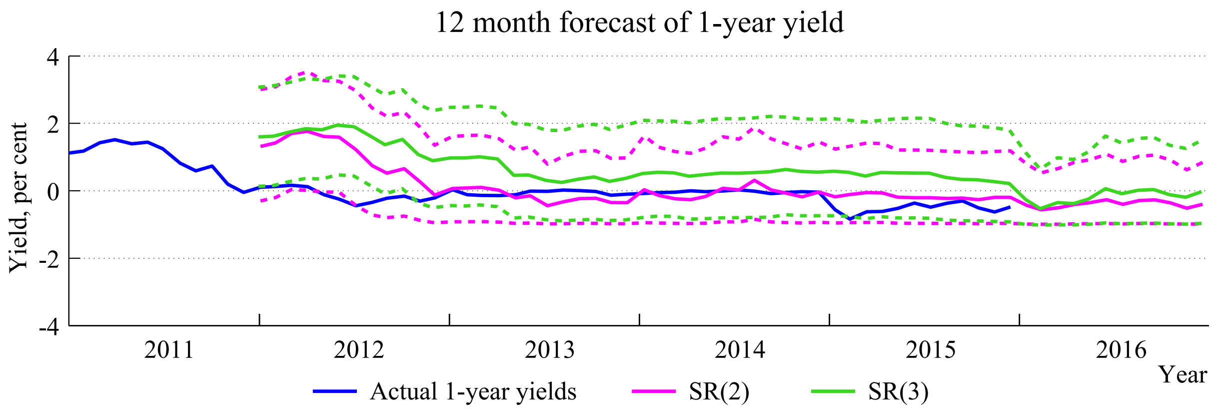

Figure 14: 12 month forecasts of 1-year yield compared to the actual yield (blue line), illustrated for the two- and three-factor SR-model. For the forecasts, the mean (solid line) and 5th and 95th percentiles (dotted lines) are indicated at the time, where the forecast should be realised, i.e. 12 months after the forecast was made.

that the forecast performance can be improved by simply assuming a random walk process of the factors. A random walk implies that yields on average stay at their current (negative) level for the entire future. This implies a negative equilibrium interest rate level, which we consider unrealistic from an economic point of view. Also, superior forecast performance in a low-yield environment does not imply that a two-factor model will provide better forecasts if rates start to lift off. In a risk management setting, we therefore consider two- as well as three-factor models going forward in a CaR analysis. However, for the purpose of estimating term premia, we use a three-factor model due to its superior short-end fit.

VI. FINAL REMARKS

We have implemented a SR-model on Danish government bond yields to capture the asymmetric distribution of future rates and the decline in interest rate volatility, which have characterised the low-yield environment. A model able to address these issues should potentially generate more credible interest rate projections, which is important from a risk management perspective. We find, however, that the SR-model has not been better at forecasting the interest rate development in recent years. The reason is that the SR-model is challenged when rates continue to decline into negative territory as has been the case in Denmark. Given that short rates in other countries recently have moved well below zero as well, we see a research potential for developing the traditional SR framework to better reflect that the lower bound cannot be considered a fixed value. A possible extension of our model implementation would be to include macroeconomic variables in order to improve the estimation of the long term interest rate level and the speed of mean reversion toward this, see e.g. Bauer and Rudebusch [2].

VII. REFERENCES

[1] Ruchir Agarwal and Miles Kimball. Breaking through the zero lower bound. *International Monetary Fund, Working paper*, (15/224), 2015.

[2] Michael D. Bauer and Glenn D. Rudebusch. Monetary policy expectations at the zero lower bound. *Federal Reserve Bank of San Francisco, Working Paper Series*, (2013-18), 2013.

[3] Michael D. Bauer, Glenn D. Rudebusch, and Jing Cynthia Wu. Correcting estimation bias in dynamic term structure models. *Journal of Business & Economic Statistics*, 30(3):454–467, 2012.

[4] Fischer Black. Interest rates as options. *The Journal of Finance*, 50(5):1371–1376, 1995.

[5] Douglas T. Breeden and Robert H. Litzenberger. Central bank policy impacts on the distribution of future interest rates. *Working paper*, 2015.

[6] Jens H.E. Christensen and Glenn D. Rudebusch. Estimating shadow-rate term structure models with near-zero yields. *Journal of Financial Econometrics*, 2014.

[7] Jens H.E. Christensen, Francis X. Diebold, and Glenn D. Rudebusch. The affine arbitrage-free class of Nelson–Siegel term structure models. *Journal of Econometrics*, 164(1):4–20, 2011.

[8] John C. Cox, Jonathan E. Ingersoll Jr, and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407, 1985.

[9] Francis X. Diebold and Canlin Li. Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337–364, 2006.

[10] Darrell Duffie and Rui Kan. A yield-factor model of interest rates. *Mathematical finance*, 6: 379–406, 1996.

[11] Carina M. Jensen and Morten Spange. Interest rate pass-through and the demand for cash at negative interest rates. *Danmarks Nationalbank Monetary Review 2nd Quarter*, pages 55–66, 2015.

[12] Leo Krippner. A tractable framework for zero-lower-bound Gaussian term structure models. *Reserve Bank of New Zealand Discussion Paper Series*, (DP2013/02), 2013.

[13] Leo Krippner. A comment on Wu and Xia (2015), and the case for two-factor shadow short rates. *Centre for Applied Macroeconomic Analysis Working Paper*, (48/2015), 2015.

[14] Leo Krippner. *Zero Lower Bound Term Structure Modeling: A Practitioners Guide*. Palgrave Macmillan, 2015.

[15] Jeffrey C. Lagarias, James A. Reeds, Margaret H. Wright, and Paul E. Wright. Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1):112–147, 1998.

[16] Wolfgang Lemke and Andreea Liliana Vladu. A shadow-rate term structure model for the euro area. *Conference paper, Annual Conference 2015 (Muenster): Economic Development - Theory and Policy*, First preliminary version, 2014.

[17] Charles R. Nelson and Andrew F. Siegel. Parsimonious modeling of yield curves. *Journal of Business*, pages 473–489, 1987.

[18] Marcel A. Priebsch. Computing arbitrage-free yields in multi-factor Gaussian shadow-rate term structure models. *Finance and Economics Discussion Series*, page 63, 2013.

Appendix A: MODEL IMPLEMENTATION

The mathematical basis for our implementation of the Arbitrage-Free Nelson–Siegel (AFNS) model and the shadow rate (SR) extension with the Krippner-approximation are reviewed below.

Arbitrage-free Nelson–Siegel model

The AFNS-model is implemented using the assumptions from Christensen et al. [7] for the risk neutral $\mathbb{Q}$-dynamics of the factors. Combining the assumptions with a diagonal covariance matrix we obtain the factor dynamics,

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} d\bar{W}_t^{\mathbb{Q}}, \tag{A1}$$

where λ is a model parameter, and $\bar{W}_t^{\mathbb{Q}}$ is a vector of Wiener processes under the risk neutral measure $\mathbb{Q}$ with standard deviation σ_{ii}. Next let the short rate be related to the factors as follows

$$r_t = X_t^1 + X_t^2. \tag{A2}$$

Based on equation (A1) and (A2) Christensen et al. [7] find a closed-form solution for the yield on a zero-coupon bond,

$$y_t(\tau) \ = B(\tau)\bar{X}_t - \tfrac{A(\tau)}{\tau}, \tag{A3}$$

where τ is the time to maturity, and $\bar{X}_t = (X_t^1, X_t^2, X_t^3)^T$. The first term in equation (A3) is given by

$$B(\tau)\bar{X}_t = X_t^1 + X_t^2 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + X_t^3 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right).$$

The second term in equation (A3) is the yield adjustment term, $A(\tau)/\tau$. This ensures that the model becomes free of arbitrage. For the diagonal specification of the σ-matrix in equation (A1) the term is given by

$$\begin{aligned}\frac{A(\tau)}{\tau} &= \sigma_{11}^2\frac{\tau^2}{6} + \sigma_{22}^2\left[\frac{1}{2\lambda^2} - \frac{1-\exp(-\lambda\tau)}{\lambda^3\tau} + \frac{1-\exp(-2\lambda\tau)}{4\lambda^3\tau}\right] \\ &\quad + \sigma_{33}^2\left[\frac{1}{2\lambda^2} + \frac{\exp(-\lambda\tau)}{\lambda^2} - \frac{\tau\exp(-2\lambda\tau)}{4\lambda} \right. \\ &\quad \left. - \frac{3\exp(-2\lambda\tau)}{4\lambda^2} - \frac{2(1-\exp(-\lambda\tau))}{\lambda^3\tau} + \frac{5(1-\exp(-2\lambda\tau))}{8\lambda^3\tau}\right]. \end{aligned} \tag{A4}$$

Next, we assume an independent specification of the factor dynamics under the physical measure[17], $\mathbb{P}$,

$$d\bar{X}_t = \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_{22}^{\mathbb{P}} & 0 \\ 0 & 0 & \kappa_{33}^{\mathbb{P}} \end{pmatrix} \left[\begin{pmatrix} \theta_1^{\mathbb{P}} \\ \theta_2^{\mathbb{P}} \\ \theta_3^{\mathbb{P}} \end{pmatrix} - \bar{X}_t \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} d\bar{W}_t^P. \tag{A5}$$

For this independent specification (diagonal κ- and σ-matrices), each factor is reverting toward a long-run mean level, $\theta_i^{\mathbb{P}}$, with a speed of $\kappa_{ii}^{\mathbb{P}}$, and $\bar{W}_t^{\mathbb{P}}$ is a Wiener process with standard deviation σ_{ii}.

Estimation of the model

To estimate the parameters in the model we use the Kalman filter which is described in detail in Appendix B of Christensen and Rudebusch [6]. For a given parameter set the Kalman filter is an iterative procedure filtering out the errors between one-period forecasts of the yield curve and the observed yields. We use the log-likelihood function to transform the information about how well our model fit the data down to a single number. The optimal parameters are the one maximizing the log-likelihood function. The Kalman filter use equation (A3) as the measurement equation and equation (A5) as the transition equation. The optimisation is performed according to Appendix B.

[17] In order to obtain the factor dynamics under the physical measure, $\mathbb{P}$, we use the Girsanov Theorem under the assumption of an affine specification of the risk premium.

We would like to add a note on using the Kalman filter from Appendix B in Christensen and Rudebusch [6]: The implementation involves a method for solving the conditional variance for the process, Q_t. For a diagonal covariance matrix, $\Sigma = \mathrm{diag}(\sigma_{11}, \sigma_{22}, \sigma_{33})$, the expression of Q_t may be reduced as[18]

$$
\begin{aligned}
Q_t &= \int_0^{\Delta t} \exp\left(- \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_{22}^{\mathbb{P}} & 0 \\ 0 & 0 & \kappa_{33}^{\mathbb{P}} \end{pmatrix} s \right) \Sigma\Sigma' \exp\left(- \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_{22}^{\mathbb{P}} & 0 \\ 0 & 0 & \kappa_{33}^{\mathbb{P}} \end{pmatrix}' s \right) \mathrm{d}s \\
&= \mathrm{diag}\left(\sigma_{11}^2 \frac{1-\exp(-2\kappa_{11}^{\mathbb{P}}\Delta_t)}{2\kappa_{11}^{\mathbb{P}}}, \sigma_{22}^2 \frac{1-\exp(-2\kappa_{22}^{\mathbb{P}}\Delta_t)}{2\kappa_{22}^{\mathbb{P}}}, \sigma_{33}^2 \frac{1-\exp(-2\kappa_{33}^{\mathbb{P}}\Delta_t)}{2\kappa_{33}^{\mathbb{P}}} \right). \qquad \text{(A6)}
\end{aligned}
$$

Shadow rate model

Introducing a lower bound to the model implies that (A3) no longer holds and yields must therefore be approximated. We choose to approximate the yield curve in the SR-model with the approximation proposed by Krippner [12], see Appendix C for details. To implement the Krippner approximation we start by recalling the relation between the forward rate and the yield,

$$y(\tau) = \frac{1}{\tau}\int_0^{\tau} f(s)ds. \qquad \text{(A7)}$$

Within the AFNS model using diagonal σ- and κ-matrix as seen in equation (A5) Christensen and Rudebusch [6] describes the forward rate as

$$f(\tau) = X_t^1 + \exp(-\lambda\tau)X_t^2 + \lambda\tau\exp(-\lambda\tau)X_t^3 - \frac{\partial A(\tau)}{\partial\tau}, \qquad \text{(A8)}$$

where

$$-\frac{\partial A(\tau)}{\partial\tau} = -\tfrac{1}{2}\sigma_{11}^2\tau^2 - \tfrac{1}{2}\sigma_{22}^2\left(\tfrac{1-\exp(-\lambda\tau)}{\lambda}\right)^2 - \tfrac{1}{2}\sigma_{33}^2\tau^2\left[\tfrac{1-\exp(-\lambda\tau)}{\lambda\tau} - \exp(-\lambda\tau)\right]^2. \qquad \text{(A9)}$$

[18] For a general specification of either the κ- or σ-matrix, Q_t must be evaluated using a basis transformation involving the eigenvalues and eigenvectors of the κ-matrix, see Krippner [14].

Krippner [12] shows how to approximate the forward rate in a SR-model. First, Krippner considers two worlds; one with cash and one without. Based on this approach he argues that the last marginal increment of the forward rate will always be none-negative due to the availability of holding cash. This, combined with the assumption that an European option approximately has the same properties as an American, Krippner derives the following result for any SR-model with a Gaussian shadow rate,

$$\bar{f}(\tau) = f(\tau)\Phi\left(\frac{f(\tau)}{\omega(\tau)}\right) + \omega(\tau)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{f(\tau)}{\omega(\tau)}\right]^2\right), \tag{A10}$$

where Φ is the Gaussian CDF function and $\omega(\tau)^2$ is the variance under the forward measure related to the forward rate. Using a three-factor AFNS-model for the shadow rate (which is a Gaussian model) Christensen and Rudebusch [6] find a closed form solution for $\omega(\tau)$. We reduce their result only to consider the independent specification of the $\mathbb{P}$-dynamics as in equation (A5),

$$\begin{aligned}\omega^2(\tau) &= \sigma_{11}^2\tau + \sigma_{22}^2\frac{1-\exp(-2\lambda\tau)}{2\lambda} \\ &\quad + \sigma_{33}^2\left[\frac{1-\exp(-2\lambda\tau)}{4\lambda} - \frac{\tau\exp(-2\lambda\tau)}{2} - \frac{\lambda\tau^2\exp(-2\lambda\tau)}{2}\right].\end{aligned} \tag{A11}$$

Christensen and Rudebusch [6] argue that the size and properties of the error attached to this approximation is unknown. We find that the approximation error for Danish government bond yield data is small, see Appendix C.

To estimate the parameters in the SR-model we also need to modify the Kalman filter because our yield function no longer is a linear function of the factors.[19] We use the extended Kalman filter (EKF) as proposed in Appendix B in Christensen and Rudebusch [6]. By using the EKF the yield function is again linearized through a first order Taylor expansion, which involves the derivatives of the shadow yield curve with respect to the factors. In the Krippner approximation, the derivatives are given as

[19] A linear function is written as $y = A + BX$, as in equation (A3).

$$\frac{d\bar{y}(\tau)}{X_t^1} = \frac{\Phi\left(\frac{f(\tau)}{\omega(\tau)}\right)}{\tau}, \tag{A12}$$

$$\frac{d\bar{y}(\tau)}{X_t^2} = \frac{\exp(-\lambda\tau)\Phi\left(\frac{f(\tau)}{\omega(\tau)}\right)}{\tau}, \tag{A13}$$

$$\frac{d\bar{y}(\tau)}{X_t^3} = \frac{\lambda\tau\exp(-\lambda\tau)\Phi\left(\frac{f(\tau)}{\omega(\tau)}\right)}{\tau}. \tag{A14}$$

Appendix B: DETERMINING THE OPTIMUM

In this section a convergence analysis is carried out. The ability of converging toward the same parameter optimum reflects the robustness of the model and is important for determining the number of runs with different starting values that should be carried out in order to determine the optimum.

In order to determine the optimal set of parameters in the Kalman filter, two search algorithms are implemented; a global search algorithm followed by a local search algorithm. Simulated annealing is used as the global search algorithm, followed by a local Nelder–Mead simplex method from Lagarias et al. [15]. In the simulated annealing algorithm we use a sufficiently high initial temperature that makes the span of the search field wide.

Next, to perform the convergence test, the algorithm setup is passed with different random starting values. The starting values for the parameters are randomly picked within a reasonable range: $\kappa_{ii} \in [0.1; 1.0]$, $\theta_i \in [-0.05; 0.05]$ and $\sigma_{ii} \in [0.001; 0.05]$ and $\lambda \in [0.1; 0.9]$ where $i \in \{1, 2, 3\}$.

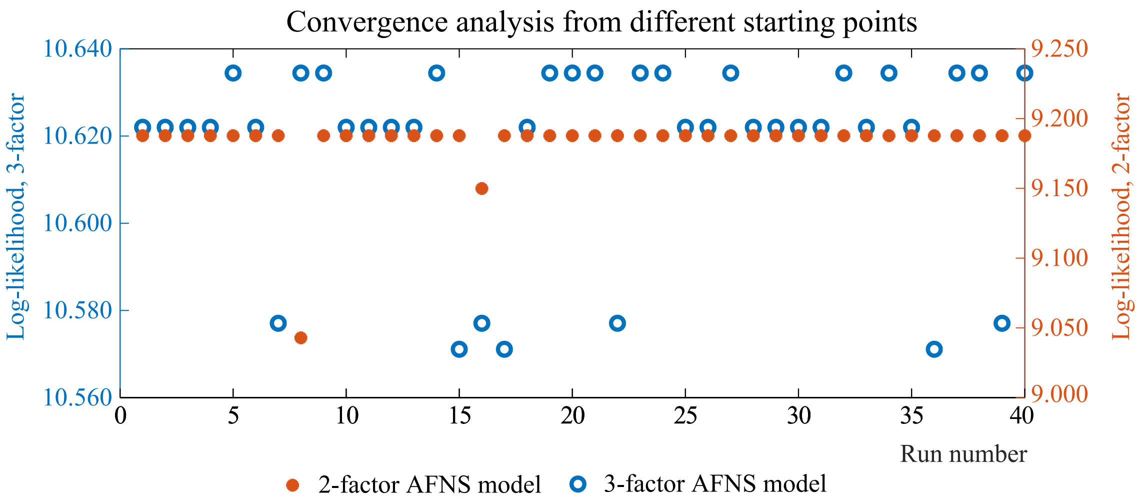

Figure 15: The log-likelihood value of the parameter optimum in the two- and three-factor AFNS model, using different starting points as stated above. The test is performed with data from January 1999 to December 2015.

The analysis shows that it is easier to identify a global optimum in a two-factor model rather than a three-factor model cf. Figure 15 and hence to find the same optimum. Specifically, the three-factor model convergences to the global optimum 15 out of 40 times compared to the two-factor model which finds the global optimum 38 out of 40 times. The higher degree of robustness of the two-factor model is due to the fewer number of parameters which have to be determined. When using the three-factor model, it is thus necessary to perform a couple of searches in order to confirm the optimum.

Appendix C: COMPARING WITH OTHER ESTIMATION APPROACHES

In order to obtain the yield curve in the shadow rate model the most exact approach would be to simulate the yield curve, as no closed form solution is available. This may e.g. be carried out by Monte Carlo simulations. The simulation approach is considered as the true solution when the number of simulations tends to infinity. This procedure is very time consuming, and therefore we implement the Krippner [12] approximation outlined in Appendix A. In this section we investigate the error attached to the Krippner approximation compared to the simulated yield curve. We also compare to another approximation introduced by Priebsch [18], who argues that his approximation is more accurate than Krippner's. Priebsch approximates the yield on a zero-coupon bond using a moment generating function. He argues that the approximation is quite accurate when only using the first two moments, see Priebsch [18]. According to Figure 13, the shadow short rate is lowest in February 2015, which is thus where the approximation error is expected to be large.

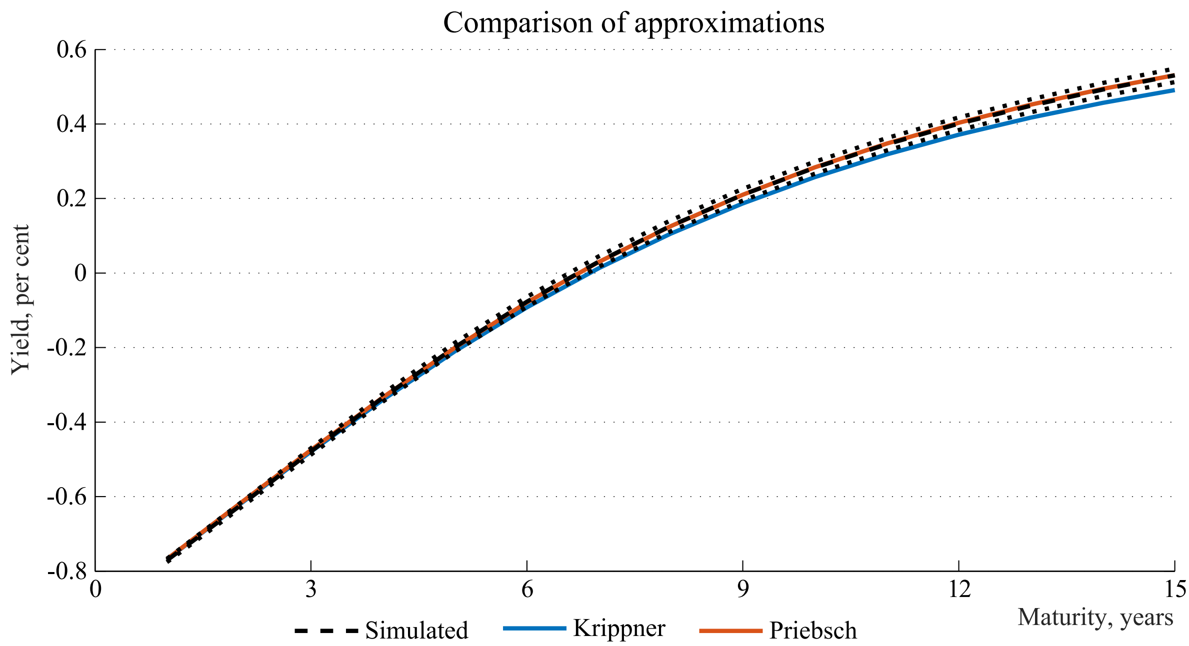

Figure 16: Approximate yield curve with the three-factor SR-model. Estimated using data from January 1999 to February 2015 with lower bound at -1.00 %. The dashed lines indicate the 5th and 95th percentiles of the simulated solution with 50,000 simulations.

As illustrated in Figure 16, the error for the Krippner approximation is increasing in maturity. The maximal error for the three-factor SR-model is observed at 15 years and is

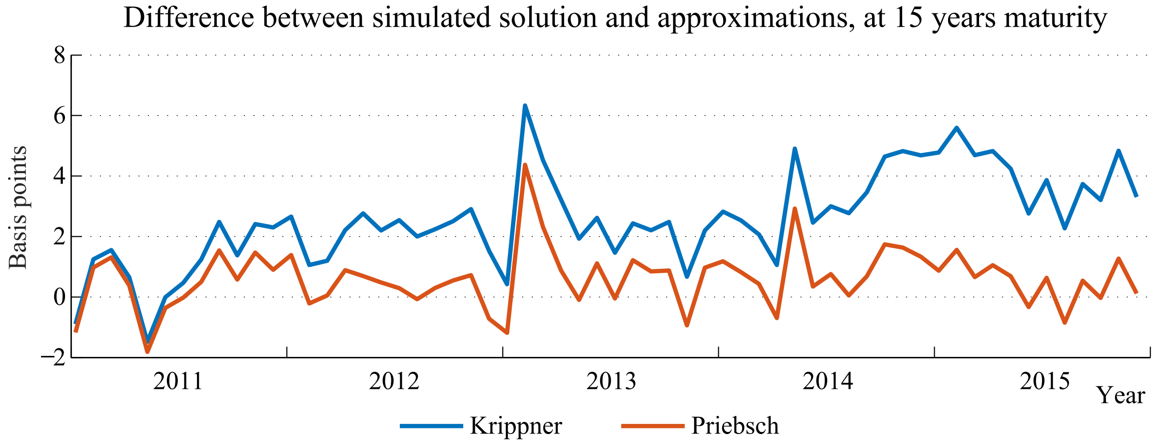

Figure 17: Difference at the 15 year maturity between simulated version (with 50.000 simulations) and approximations using data from January 1999 with lower bound at -1.00 %.

6 basis points for the Krippner approximation compared to less than 1 basis point from the Priebsch approximation. After six years the Krippner approximation breaches the 5th percentile for the simulated solution.

Next, we investigate how the maximal approximation error evolves over the 60 month forecast period, cf. Figure 17. The error in the Krippner approximation is slightly larger than the Priebsch approximation over the whole period, but does never exceed 7 basis points. On the other hand the Krippner approximation is almost a 100 times faster than the Priebsch approximation with our implementation. Therefore we do not find the benefit from using the Priebsch approximation big enough to compensate for the additional time consumption.