ERASMUS UNIVERSITEIT ROTTERDAM

Master Thesis Quantitative Finance - Erasmus School of Economics

Affine Term Structure Models Approaching The Zero-Lower Bound

Tim Everaert[1]
499050

Supervised by:
Dr. Michel van der Wel

Co-read by:
Dr. Maria Grith

June 27, 2019

Abstract

This thesis reviews the affine term structure model class as proposed by Duffie and Kan (1996). Since prior research is scarce on evaluating data including the zero-lower bound, I contribute by investigating a recent U.S. Treasury data set. I estimate one-, two-, and three-factor Vasicek (1977) and Cox et al. (1985) models using the Kalman filter approach of De Jong (2000) in an empirical study while including the zero-lower bound. I also perform a simulation study for the three-factor models under a zero-lower bound environment by lowering the short rate and the volatility. I provide evidence that the three-factor Vasicek (1977) model obtains the best fit and captures the stylized facts whereas the Cox et al. (1985) only captures the yield curve's level and slope. This evidence is less apparent on data including the zero-lower bound. In the simulation study, the Cox et al. (1985) model is more accurate than the Vasicek (1977) model on CIR data whereas the performance is close on Vasicek data.

[1]The content of this thesis is the sole responsibility of the author and does not reflect the view of either Erasmus School of Economics or Erasmus University.

Contents

1 Introduction

Affine Term Structure Models (ATSMs) are valuable to policy makers, practitioners and the academic community for investigating bond prices, monetary policy, and the macroeconomic elements determining discount rates (Piazzesi (2010), Gurkaynak and Wright (2012), Duffee (2013), Diebold and Rudebusch (2013)). Today, these ATSMs might be even more valuable in times of historically low yields and an emerging zero-lower bound (Hamilton and Wu (2012a), Krippner (2013a), Wu and Xia (2016), Bauer and Rudebusch (2016)). Friedman (1977) originally noticed that researchers of statistical functions should investigate the modelling of the whole term structure of yields by using a few parameters, resulting into two traditional approaches in the literature. The first deals with Gaussian factors such as the Vasicek (1977) approach, whereas the second deals with non-Gaussian factors as in the Cox, Ingersoll and Ross (1985) approach. Some years later, Duffie and Kan (1996) generalized both in the class of ATSMs.

More recently, the literature has focused primarily on Gaussian term structure models (Christensen et al. (2011), Joslin et al. (2011), Hamilton and Wu (2012b), Adrian et al. (2013), de Los Rios (2015)) whereas the developments of non-Gaussian estimation methods are limited. However, there are some novel non-Gaussian developments (De Jong (2000), Ait-Sahalia and Kimmel (2010), Creal and Wu (2015)). For instance, De Jong (2000) applies a state-space framework and estimates ATSMs by using the Kalman filter and quasi-maximum likelihood (QML) of which the estimation result is known to be inconsistent which can be eliminated by several procedures as illustrated by Gallant and Tauchen (1996), Gourieroux et al. (1993), Frachot et al. (1995), Dai and Singleton (2000), and Lamoureux and Witte (2002). However, the bias in the QML estimates is very small as shown by De Jong (2000) who also argues that these procedures are numerically exhaustive. Therefore, I implement the QML-Kalman filter approach of De Jong (2000).

In this thesis, I present a new perspective on the yield curve in two manners. First, I investigate the impact of the zero-lower bound on the Gaussian model of Vasicek (1977) in comparison with the non-Gaussian model of Cox et al. (1985) since non-Gaussian models are unable to handle negative interest rates. Therefore, I assess the zero-lower bound's impact on the estimation performance of the n-factor models of both Vasicek (1977) and Cox et al. (1985) by applying the Kalman filter approach of De Jong (2000) in an empirical study, for $n = 1, 2, 3$. Second, I investigate the accuracy of both three-factor models in an extensive simulation study while replicating a zero-lower bound environment.

To examine the impact of the zero-lower bound, defined by historically low yields and low volatility, I equally divide the data set into a period excluding and a period including the zero-lower bound respectively spanning from January 1982 until May 2000 and from June 2000 until October 2018. The data set contains monthly observations on Treasury yields with maturities of (3,12,60,120) months. I assess which factor model obtains the best fit and whether the factors represent the level, slope and curvature factor in each period. In addition, using the obtained empirical parameter values, I simulate data from the three-factor models to replicate a zero-lower bound (regular) environment by decreasing (increasing) the long-term short rate and variance parameters. Subsequently, I perform an ordinary simulation where I estimate the Vasicek (1977) (Cox et al. (1985)) model on Vasicek (CIR) data, and a cross-simulation study where I estimate the Vasicek (1977) model on CIR data and reversed.

To the best of my knowledge, an extensive empirical study by applying the Kalman filter approach of De Jong (2000) has not been performed on the n-factor model of Vasicek (1977) and of Cox et al. (1985), while incorporating the zero-lower bound, for $n = 1, 2, 3$. Since De Jong (2000) examined the time-period from 1970 until 1991, his data does not include the zero-lower bound in the U.S. which spans from 2008 until 2017. Hence, by examining a recent data set, I add an interesting dimension which encompasses the main contribution of this thesis: the impact of the zero-lower bound. In addition, I contribute to prior research by examining the estimation performance of the three-factor models in both an ordinary and a cross-simulation study where I replicate a controlled zero-lower bound and a regular environment.

In the ordinary simulation study, I find that the three-factor Cox et al. (1985) model appears to be more accurate and less variable for most of the parameters than the Vasicek (1977) model which has problems estimating its mean-reversion and variance parameters. Nevertheless, the bias on the average short rate is somewhat smaller for the Vasicek (1977) model. The parameter standard deviations in the Vasicek (1977) case appear somewhat larger under the zero-lower bound period, which is reversed in the Cox et al. (1985) case. In the cross-simulation study, I find that the Cox et al. (1985) model appears to be more accurate and less variable for all parameters than the Vasicek (1977) model on CIR data whereas the estimation estimation results are close when estimating on Vasicek data.

In the empirical study, I find that the three-factor models are better specified than the one- and two-factor models in both the period excluding and including the zero-lower bound. The third factor enhances the fit of the average term structure considerably for the excluding zero-lower bound period where the steepness in the middle appears to be only completely seized by the Vasicek (1977) model which has a relatively tiny RMSE that is considerably larger in the other period. The serial correlation is close to zero for both models in the period excluding, while it is considerably higher in the period including the zero-lower bound. Consequently, the Vasicek (1977) three-factor model appears to be the better specified in both periods, although substantially less in the including zero-lower bound period.

Moreover, I find that the first factor describes the level of the yield curve for each model in both periods. The second factor enjoys substantial support as the yield curve's slope for all models in both periods except for the two-factor Cox et al. (1985) model in the including zero-lower bound period. In the Vasicek (1977) case, the third factor has the interpretation as the (reversed) curvature factor. In the Cox et al. (1985) case, the third factor's interpretation as the curvature factor appears relatively little given the tiny correlation between the third factor and its proxy. For both models, the correlation is substantially lower in the period including the zero-lower bound for the first and third factor.

The outline of the remaining parts of this thesis is as follows: section 2 provides an overview of the developments on Gaussian, non-Gaussian, and zero-lower bound estimation methods and models. Section 3 presents a discussion on the theory of ATSMs, the empirical implementation of ATSMs using a state-space model along with the Kalman filter, and a brief description of the Vasicek (1977) as well as the Cox et al. (1985) specification. Section 4 provides the ordinary and cross-simulation study including a detailed description of the setup. Section 5 presents an extensive empirical study on the n-factor Vasicek (1977) and the Cox et al. (1985) models' performance when subject to the zero-lower bound, for $n = 1, 2, 3$. Section 6 provides a conclusion and discussion.

2 Related Literature

As described shortly in the introduction, Friedman (1977) suggested the innovative notion that researchers of statistical demand functions should investigate the modelling of the whole term structure of yields by using just a few parameters. With this notion, Friedman (1977) lay the groundwork for the extensive literature on term structure models. In the classical version of these models characterizing the yield curve, one or more factors move the short rate and the whole yield curve through no-arbitrage. Three factors are often incorporated in order to seize the level, slope and curvature, i.e. the stylized facts, of the yield curve. Consequently, two traditional approaches emerged in the literature.

The first approach handles Gaussian factors as in the Vasicek (1977) approach. The Vasicek model follows an Ornstein-Uhlenbeck stochastic process, i.e. a stationary Gauss-Markov process, which implies it is a Gaussian process where the random variables follow a multivariate normal distribution. Furthermore, the volatility is assumed to be constant in the Vasicek approach. The second approach handles non-Gaussian factors as in the Cox, Ingersoll and Ross (1985) approach, often denoted by the CIR approach. The mutually independent factors in the CIR approach do not follow a normal distribution and are therefore considered to be non-Gaussian. Moreover, the CIR approach relies on a square root volatility framework instead of a constant volatility framework as in the Vasicek approach. This difference between the short rate processes of the Vasicek and CIR approach is explained in more detail in section 3.

Duffie and Kan (1996) made a generalization of the Vasicek (1977) and Cox et al. (1985) approach by developing the affine term structure model (ATSM) approach. In this affine class approach, the factors contain an affine volatility framework that consists of a generalized version of the square-root volatility framework of the CIR approach. Duffie and Kan (1996) allow on top of that for correlation of the factors in their affine class. The affine term-structure model is highly docile since yields are affine factor alterations. Apart from the Vasicek and CIR case, Duffie and Kan (1996) present a brief overview on other special cases of their affine model which includes the model of Langetieg (1980), El Karoui and Rochet (1989), Jamshidian (1989), Jamshidian (1991), Pennacchi (1991), Chen and Scott (1992), Heston (1991), Jamshidian (1992), Longstaff and Schwartz (1992) and Chen (1993).

This generalization of Duffie and Kan (1996) lay the basis for the estimation of ATSMs on which there is a consensus in the literature that it can be challenging (Duffee (2002), Ang and Piazzesi (2003), Kim and Orphanides (2005)). For instance, a standard ATSM generates poor predictions of future yields according to Duffee (2002), in comparison with assuming that yields reflect random walks. Duffee argues that the lack of success of the ATSM is due to the feature that the risk premium is defined as the multiple of the risk-variance. Furthermore, Ang and Piazzesi (2003) incorporate macroeconomic elements as well as traditional unobserved factors in their ATSMs. They find that models which incorporate macroeconomic elements obtain a better forecasting performance than traditional ATSMs which contain solely unobserved factors. Moreover, Kim and Orphanides (2005) argue that estimating dynamic no-arbitrage ATSMs on a small sample including an adaptable market price of risk results into difficulty because of the substantially resolute character of interest rates. To surmount the problem, they apply survey predictions of a yield with a short maturity as a supplementary variable for estimating ATSMs.

Although estimating ATSMs appears to be challenging, the scientific community made substantial improvements in estimating Gaussian ATSMs (Christensen et al. (2011), Joslin et al. (2011), Hamilton and Wu (2012b), Adrian et al. (2013), de Los Rios (2015)). For example, the novel canonical Gaussian dynamic term structure model (DTSM) results from the research of Joslin et al. (2011) which includes observable portfolios of interest rates in the role of pricing factors. They find that predictions of the pricing factors remain invariant to imposing arbitrage-free conditions even when there are several constraints imposed on the factor process of bond yields. They also claim that regular maximum likelihood algorithms reach the optimal global solution immediately for their normalization.

Instead of constructing DTSMs that include observable portfolios of interest rates, Christensen et al. (2011) devise an approach based on no-arbitrage affine DTSMs which comes close to the popular Siegel and Nelson (1988) term structure representation. These models are associated with a canonical no-arbitrage ATSM specification with three factors. They find that enforcing the Siegel and Nelson (1988) framework on this canonical specification substantially enhances its empirical tractability. Besides, they find that imposing no-arbitrage conditions substantially improves predictive performance.

With respect to estimating ATSMs, Hamilton and Wu (2012b) suggest to reconsider maximum likelihood estimation and instead use minimum-chi-square estimation. Since maximum likelihood estimation is asymptotically identical to minimum-chi-square estimation, they claim that it is interesting to use minimum-chi-square estimation by providing evidence that it is easier to compute than maximum likelihood estimation. Under some conditions, Hamilton and Wu (2012b) state that researchers are able to compute small-sample standard errors with minimum-chi-square estimation and to infer with certainty whether a certain estimate illustrates a global optimal solution of the likelihood function.

Adrian et al. (2013) suggest a different approach and use ordinary least squares regressions to price yield curve elements in a cross-sectional as well as a time series manner. They claim that their approach grants substantially rapid estimation of ATSMs accompanying a vast amount of pricing elements. Furthermore, they argue that their approach produces a term structure of interest rates with limited pricing errors in comparison with ordinarily disclosed specifications without enforcing cross-equation constraints during estimation. Adrian et al. (2013) find this result for their approach both in- and out-of-sample.

Similar to Adrian et al. (2013), de Los Rios (2015) introduces a novel procedure to estimate Gaussian DTSMs based on regressions. However, his novel estimation procedure relies instead on asymptotic least-squares characterized by the arbitrage-free prerequisites inherent to Gaussian DTSMs and he presents an empirical analysis on Canadian zero-coupon bonds. de Los Rios (2015) claims that his estimator maintains its easy computation and remains asymptotically efficient under varying conditions where several newly introduced procedures might be deprived of their tractability.

The above mentioned improvements of Christensen et al. (2011), Joslin et al. (2011), Hamilton and Wu (2012b), Adrian et al. (2013), and de Los Rios (2015) allow for easier estimation of Gaussian term structure models which resulted in an increased popularity of the implementation and research with respect to estimation of Gaussian ATSMs. By contrast, the developments of estimation methods for the non-Gaussian term structure models remain limited. However, there are some novel non-Gaussian developments (Ait-Sahalia and Kimmel (2010), Creal and Wu (2015), De Jong (2000)).

For instance, one of the earliest non-Gaussian estimation methods was proposed by De Jong (2000) who presents an empirical study on the yield curve by applying the ATSM approach of Duffie and Kan (1996). He computes the ATSMs in a theoretically correct manner through integrating the cross-section and time series dimension appropriately. De Jong (2000) applies a state space framework and estimates ATSMs by implementing a Kalman filter and quasi-maximum likelihood (QML) for which he discretizes the continuous-time mechanism of the factors. In addition, he applies a broad framework for measurement errors. De Jong (2000) consciously examines the empirical behavior of the n-factor ATSMs where he evaluates the added value of every supplementary factor, for $n = 1, 2, 3$. Furthermore, De Jong (2000) argues that incorporating a greater number of maturities than factors generally classifies every parameter in the ATSM. He finds that a three-factor ATSM, where he approves for correlation among the factors, captures a sufficiently well fit of the yield curve's time series in addition to its cross-section where the three factors deserve the interpretation of the level, slope, and curvature.

Instead of using QML like De Jong (2000), Ait-Sahalia and Kimmel (2010) develop another non-Gaussian estimation method. They design and implement a procedure for closed-form maximum likelihood estimation of multifactor ATSMs. They perform their procedure on nine Dai and Singleton (2000) ATSMs which they estimate with distinct market price of risk specifications on U.S. treasury data and show by performing a simulation study that their procedure precisely approximates true but improbable maximum likelihood estimation. Ait-Sahalia and Kimmel (2010) conclude that estimation with real and simulated data implies that their estimation procedure is substantially closer to the true maximum likelihood estimate than Euler and QML.

Apart from the aforementioned methods, Creal and Wu (2015) devise a new non-Gaussian estimation method to investigate spanned along with unspanned stochastic volatility models with three or four factors. They optimize a concentrated likelihood constructed from linear regressions which increases the speed of optimization significantly by lowering the dimension of the numerical optimization problem. Nevertheless, it produces the equivalent estimator as maximizing the regular likelihood. Creal and Wu (2015) state that their approach enhances the numerical behavior of estimation by removing parameters from the objective function that create problems for traditional methods. As suggested by Collin-Dufresne and Goldstein (2002), the closed-form log-likelihood function is not established for unspanned stochastic volatility models. Therefore, Creal and Wu (2015) implement the expectation maximization (EM) algorithm of Dempster et al. (1977). They observe that spanned stochastic volatility models produce a sufficient fit of the cross-section of yields while causing a loss in capturing the volatility, whereas unspanned stochastic volatility models capture volatility while causing a loss in capturing the cross-section.

In recent years, the academic literature has its focus on evaluating the effect of fiscal stimulus programs to increase spending and inflation during times in which the economy is subject to an interest rate climate with a prominent existence of the zero-lower bound (Hamilton and Wu (2012a), Krippner (2013a), Bauer and Rudebusch (2016), Wu and Xia (2016)). This research on evaluating fiscal stimulus programs in times of historically low interest rates culminates in developments of so-called zero-lower bound models to examine the effect and implications of the occurrence of the exceptional zero-lower bound phenomenon (Krippner (2013b), Christensen and Rudebusch (2016)).

For instance, Krippner (2013b) argues that if nominal yields come close to the zero-lower bound, implementing the common approach of Gaussian ATSMs would be theoretically unsound due to their implicit substantial probabilities of negative yields. Therefore, he introduces a docile alteration of Gaussian ATSMs that imposes the zero-lower bound which comes close to the completely no-arbitrage, although barely docile, structure as suggested by Black (1995). Krippner (2013b) implements his framework to U.S. term structure data by using the iterated extended Kalman filter with robust estimation. He shows that the results for a two-factor model are almost equivalent to the results from a similar Black (1995) model. Moreover, using data sets containing longer maturities Krippner (2013b) illustrates that his zero-lower bound framework allows for direct implementation under conditions where it would be computationally troublesome or impossible to implement the Black (1995) framework by estimating Gaussian ATSMs with two and with three factors.

Another development with respect to zero-lower bound models is the work of Christensen and Rudebusch (2016) who implement a shadow-rate DTSM that considers the zero-lower bound. They compare the shadow-rate DTSM's performance with the regular affine Gaussian DTSM's performance that does not consider the zero-lower bound. Close to the zero-lower bound, they detect an evident decrease in prediction accuracy of the regular Gaussian DTSM, whereas the shadow-rate DTSM predicts sufficiently. Nevertheless, predictions on the premiums for ten year maturity yields are largely equal among the two term structure models. At last, Christensen and Rudebusch (2016) do not discover an improvement in the estimation on U.S. yields when incorporating a marginally positive lower bound.

All together, although zero-lower bound models are available in the literature, it is interesting to investigate the performance of one of the first non-Gaussian estimation methods, i.e. the Kalman filter approach of De Jong (2000), while subject to the zero-lower bound because research is quite scarce involving non-Gaussian estimation methods in general, and especially when zooming in on the effect of the zero-lower bound. To the best of my knowledge, an empirical study by applying the Kalman filter approach of De Jong (2000) has not been performed on the n-factor models of Vasicek (1977) and Cox et al. (1985), for $n = 1, 2, 3$ while incorporating the zero-lower bound.

3 Models

This section provides the models and algorithms employed in this thesis. First, it provides an overview on the approach of Affine Term Structure Models (ATSMs), in the spirit of Duffie and Kan (1996). Second, it discusses the empirical application of the ATSM based on a state-space model. Then, it provides a brief description on the Vasicek (1977) as well as the Cox et al. (1985) model. In addition, it presents the Kalman filter, theory on quasi-maximum likelihood and the estimation methodology.

3.1 Affine Term Structure Models

In this subsection, I closely follow the terminology of Duffie and Kan (1996) who suggest that in term structure models characterizing the yield curve, factor(s) move the short rate and the whole yield curve by the no-arbitrage assumption. The zero-coupon bond price $P_t(\tau)$ for time t until maturity τ is obtained by discounting its expected payment governed through "risk-neutral" measure $\mathbb{Q}$ by the short rate r_s,

$$P_t(\tau) = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{\tau} r_s ds\right)\right]. \tag{1}$$

Duffie and Kan (1996) introduce their ATSM approach where the short rate r_t contains a constant A_0 as well as various latent factors F_t multiplied by a scale vector B_0,

$$r_t = A_0 + B_0' F_t, \tag{2}$$

with $F_t \in R^n$. Duffie and Kan (1996) assume that the latent factors F_t pursue a diffusion mechanism accompanying a framework for the affine volatility under the "real world" measure $\mathbb{P}$,

$$dF_t = \Lambda(F_t - \mu)dt + \Sigma \begin{pmatrix} \sqrt{\alpha_1 + \beta_1' F_t} dW_{1t} \\ \vdots \\ \sqrt{\alpha_n + \beta_n' F_t} dW_{nt} \end{pmatrix}, \tag{3}$$

with W_{it} autonomous Wiener processes, Λ the mean-reversion matrix, μ the mean of the factors, Σ the correlation matrix and α and β the variance parameters.

To price bonds Duffie and Kan (1996) require a stochastic factor mechanism under $\mathbb{Q}$. Therefore, they assume that factor j's market price of risk ψ_j is reliant on a multiple of the volatility, $\psi_j\sqrt{\alpha_j + \beta_j' F_t}$. The converted mechanism $dW_{jt}^* \equiv dW_{jt} + \psi_j\sqrt{\alpha_j + \beta_j' F_t}dt$ is due to this assumption a Wiener mechanism governed by the identical martingale $\mathbb{Q}$. Under $\mathbb{Q}$, the stochastic factor mechanism reads,

$$dF_t = \Lambda^*(F_t - \mu^*)dt + \Sigma \begin{pmatrix} \sqrt{\alpha_1 + \beta_1' F_t} dW_{1t}^* \\ \vdots \\ \sqrt{\alpha_n + \beta_n' F_t} dW_{nt}^* \end{pmatrix}, \tag{4}$$

with the same definition for the parameters as in (3). Both the mean-reversion parameters and the risk-neutral intercept are connected with the real-world parameters via

$$\Lambda^* = \Lambda - \Sigma\Psi\mathbb{B}' \tag{5}$$

and

$$\Lambda^*\mu^* = \Lambda\mu + \Sigma\Psi\alpha', \tag{6}$$

where $\alpha = (\alpha_1, \dots, \alpha_n)'$, $\mathbb{B} = (\beta_1, \dots, \beta_n)$, and $\Psi = \text{diag}(\psi_1, \dots, \psi_n)$.

Duffie and Kan (1996) exhibit in their ATSM approach that the zero-coupon bond price denoted by $P_t(\tau)$ for time to maturity τ, consists of an exponential combination of the factors F_t,

$$P_t(\tau) = \exp[-A(\tau) - B(\tau)'F_t], \tag{7}$$

with $A(\tau)$ the factor intercept and $B(\tau)$ the factor loadings. Thus, due to this relation the zero-coupon bond yields $Y_t(\tau)$ to maturity consist of a combination of the factors F_t. The time to maturity τ specifies the factor loadings $B'(\tau)/\tau$ and intercept $A(\tau)/\tau$ in,

$$Y_t(\tau) \equiv -\ln P_t(\tau)/\tau = A(\tau)/\tau + B(\tau)'/\tau \cdot F_t. \tag{8}$$

These loadings $A(\tau)$ and $B(\tau)$ are in line with the regular differential equation structures,

$$\frac{dA(\tau)}{d\tau} = A_0 - (\Lambda^*\mu^*)'B(\tau) - \frac{1}{2}\sum_i\sum_j B_i(\tau)B_j(\tau)a_{ij} \tag{9}$$

and

$$\frac{dB(\tau)}{d\tau} = B_0 - (\Lambda^*)'B(\tau) - \frac{1}{2}\sum_i\sum_j B_i(\tau)B_j(\tau)b_{ij}, \tag{10}$$

where the vectors b_{ij} and the scalars a_{ij} are constructed as $a_{ij} + b'_{ij}x \equiv [\Sigma\text{diag}(\alpha + \mathbb{B}'x)\Sigma']_{ij}$.

The class of ATSMs encompasses various popular models as special cases. For instance, the generalization of the Vasicek (1977) model to a multivariate version, i.e. the model of Langetieg (1980), arises when $\mathbb{B} = 0$. The generalization of the Cox et al. (1985) model arises when $\mathbb{B}$ is diagonal and $\alpha = 0$. This generalization ensures that every yield is positive. In an empirical setting, De Jong (2000) argues that not every parameter of the ATSM can be classified and therefore some normalizations are inevitable. The first classification problem involves the "intercepts" of the model. When A_0 is a free parameter, the variance intercept, α, and the average of the factors, μ, are not independently classified. As a result, in line with De Jong (2000), I normalize $\mu = 0$. Due to the normalization, A_0 equals the average short rate which is governed by the $\mathbb{P}$ measure, whereas α consists of the mean volatility for each factor.

For the next normalization, De Jong (2000) relies on Pang and Hodges (1996) who demonstrate that the price of a bond remains constant when applying invertible factor alterations. This means that the price of a bond remains constant when performing scale alterations of the factors. Therefore, I follow De Jong (2000) and normalize $B_0 = \iota$ which equates the instantaneous short rate with the total value of the factor(s) and a constant ($r_t = A_0 + \iota' F_t$), while maintaining generality. Another development of the work of Pang and Hodges (1996) is that solely the product matrix $\mathbb{K} \equiv \Sigma^{-1}\Lambda\Sigma$ can be classified. Subsequently, De Jong (2000) assumes that Σ equals the identity matrix while maintaining generality. However, he claims that in an empirical setting it is more convenient to normalize the diagonal elements of Σ to be one and he parameterizes $\Lambda = \text{diag}(-\kappa_1, \ldots, -\kappa_n)'$. Consequently, I follow this convention. For every parametrization, the vector $\kappa = (\kappa_1, \ldots, \kappa_n)'$ comprises (minus) the mean-reversion eigenvalue matrix which does not depend on the normalization and is continually classified.

At last, Duffie and Kan (1996)'s presence prerequisites in their ATSM approach, as reviewed by Dai and Singleton (2000), enforce various complementary constraints concerning Σ and $\mathbb{B}$. Following De Jong (2000), I parametrize $\mathbb{B}'\Sigma = \tilde{\beta}$, where $\tilde{\beta}$ is diagonal and consider $(\tilde{\beta}_{11}, \ldots, \tilde{\beta}_{nn})$ to be free parameters. The process below recaps the parametrization of the model which is similar to equations (2) and (3),

$$r_t = A_0 + \iota' F_t \tag{11}$$

and

$$dF_t = \Lambda F_t dt + \Sigma \begin{pmatrix} \sqrt{\alpha_1 + \tilde{\beta}_{11}(\Sigma^{-1}F_t)_1} dW_{1t} \\ \vdots \\ \sqrt{\alpha_1 + \tilde{\beta}_{nn}(\Sigma^{-1}F_t)_n} dW_{nt} \end{pmatrix}, \tag{12}$$

where $\Lambda = \text{diag}(-\kappa_1, \ldots, -\kappa_n)$. Every parameter in this process is classifiable, other than the diagonal elements of Σ which are set to one. The market price of risk (ψ) classification, is reviewed below in 3.2.

3.2 Empirical Application Of The Affine Term Structure Model

De Jong (2000) argues that the most straightforward method of estimating a factor model is by picking an equal amount of yields as the amount of distinct maturities. Then by inversion of the model, the factors can be extracted. For a few models, the discrete transition density of the factors is recognized such as for the multifactor Vasicek (1977) models. Multiplication of the Jacobian of the transformation and the aforementioned discrete transition density results in the explicit likelihood function. Chen and Scott (1992) and Pearson and Sun (1994) pursue the aforementioned procedure to analyze two-factor Cox et al. (1985) models. However, the preference of maturities to compose the factors is quite subjective while the results of the model are dependent on the preference. Alternatively, I closely follow De Jong (2000) and my estimation depends on the state-space representation of the ATSM.

The non-Gaussian estimation method of De Jong (2000) that I examine is based on the Kalman Filter. De Jong (2000) uses Quasi-Maximum Likelihood (QML) for estimating parameters via Kalman-filter equations, modified from Hamilton (1995). He employs a state-space framework where he gathers both the zero-coupon bond yields observed with error for time t with maturities τ_1 to τ_k and the coefficients, in matrix B and in the vectors y_t and A, characterized as

$$y_t \equiv \begin{pmatrix} Y_t(\tau_1) \\ \vdots \\ Y_t(\tau_k) \end{pmatrix}, \quad A \equiv \begin{pmatrix} A(\tau_1)/\tau_1 \\ \vdots \\ A(\tau_k)/\tau_k \end{pmatrix}, \quad B \equiv \begin{pmatrix} B'(\tau_1)/\tau_1 \\ \vdots \\ B'(\tau_k)/\tau_k \end{pmatrix}. \tag{13}$$

The time period between two observations is denoted by h in the state-space representation of the model,

$$y_t = A + BF_t + e_t, \quad \text{var}\,(e_t) = H, \tag{14}$$

$$F_{t+h} = \Phi F_t + v_{t+h}, \quad \text{var}\,(v_{t+h}) = q\,(F_t). \tag{15}$$

where (y_t, A, B, F_t) are defined as before and Φ, e_t, v_t are respectively the transition matrix and the innovations of the equations. Equation (14) is called the measurement equation of which the coefficients are dependent on the parameters under the risk-neutral distribution, $(A_0, \mu^*, \Lambda^*, \alpha, \tilde{\beta}, \Sigma)$. These parameters rely on the underlying parameters $(A_0, \kappa, \alpha, \tilde{\beta}, \Sigma, \phi)$ via the two constraints stated in (5) and (6) because of the representation of the model. Equation (15) renowned as the transition equation which equates to the discretization of (3) where the normalization $\mu = 0$ is enforced. In the transition equation, the parameters adhere to the conditional factor mean and variance which are respectively denoted by $\mathbb{E}[F_{t+h}|F_t] = \Phi F_t$, and $\text{var}(v_{t+h}) = \text{var}(F_{t+h}|F_t) \equiv q(F_t)$.

The state-space model provides an understandable observation on the market prices of risk classification. Governed by $\mathbb{Q}$, the factor average (μ^*) obtains classification as the mean of the actual yield vector, y_t, granted there are for n factors at a minimum $n+1$ yields to maturity. For classification of the short rate intercept, i.e. A_0, an extra yield is required. Since I follow De Jong (2000) by restricting the average of the factors governed by the measure $\mathbb{P}$ (μ) as being equal to zero, the parameters ψ are accurately classified in constraint (6). This is in contradiction with Pearson and Sun (1994) and Dai and Singleton (2000) who acquire classification for ψ via constraint (5) since they estimate n factor models based on n yields to maturity. However, De Jong (2000) argues that this method solely provides classification when the volatility relies upon the level of the factors.

The ATSM forecasts the accurate relation $y_t = A + BF_t$ between the yields and the factors. However, this exact relation is not gratified when incorporating a greater number of maturities than factors. Hence, a type of estimation error must be implemented on which the theoretical ATSM unfortunately appears quiet but various suggestions emerged from prior research. Chen and Scott (1992) use four maturities to examine a two-factor ATSM where they assume that two interest rates are measured with no error to perform inversion to extract the factors, whereas the estimation of the other interest rates involves a normally distributed estimation error. Various papers assume an estimation error for all yields that has no correlation on the cross-sectional and serial level (Jegadeesh and Pennacchi (1996), Geyer and Pichler (1999), Duan and Simonato (1999)). Frachot et al. (1995) and Lund (1997) argue that the diagonal error covariance matrix will be inappropriate when effected by linear alterations of the data. Therefore, both papers suggest an ordinary, non-diagonal, cross-sectional error covariance matrix.

De Jong (2000) takes these notions into account and assumes the estimation errors to be on average equal to zero and not serially correlated. Moreover, he allows for cross-sectional correlation through constant covariance matrix H which he defines as LDL' to assure that H is positive definite, with L lower triangular where the diagonal components are set equal to one and D a diagonal eigenvalue matrix. The definitions applied in prior research are less general than the definition of De Jong (2000) which ensures that the estimated parameters remain constant despite alterations of the yield vector y_t.

3.3 The Gaussian Case: The Vasicek Model

The first approach deals with Gaussian factors such as the Vasicek (1977) approach where the model follows the stochastic differential equation for the short rate r_t with k mutually independent factors

$$dr_{i,t} = \kappa_i(\theta_i - r_{i,t})dt + \sqrt{\tilde{\alpha}_i}dW_{i,t}, \quad \text{for } i = 1, \dots, k \tag{16}$$

where $W_{i,t}$ is a Wiener process governed by $\mathbb{Q}$, $\tilde{\alpha}_i$ the constant variance parameter, θ_i the long term average level of the yield, and κ_i the mean-reversion speed. This model follows an Ornstein-Uhlenbeck stochastic process, i.e. a stationary Gauss-Markov process, which implies it is a Gaussian process where the random variables follow a multivariate normal distribution.

In this subsection, I present the three-factor ($k = 3$) Vasicek (1977) model which can be easily simplified to the one-dimensional and the two-dimensional case to obtain respectively the one-factor ($k = 1$) and the two-factor ($k = 2$) ATSM. The term structure equals $P_t(\tau) = \exp[-\tilde{A}(\tau) - B(\tau)r_t]$, with the definition of Duan and Simonato (1999) for intercept $\tilde{A}(\tau)$ and factor loading matrix $B(\tau)$, that is,

$$\tilde{A}(\tau) = \tilde{A}_1(\tau) + \tilde{A}_2(\tau) + \tilde{A}_3(\tau), \tag{17}$$

$$B(\tau) = [B_1(\tau), B_2(\tau), B_3(\tau)], \tag{18}$$

where the functions $\tilde{A}_i(\tau)$ and $B_i(\tau)$ equal,

$$\tilde{A}_i(\tau) = R_{i,\infty}(\tau - B_i(\tau)) + \frac{\tilde{\alpha}_i}{4\kappa_i}B_i(\tau)^2, \tag{19}$$

$$B_i(\tau) = \frac{(1 - e^{-\kappa_i\tau})}{\kappa_i}, \tag{20}$$

with $R_{i,\infty} \equiv \theta_i - (\psi_i\tilde{\alpha}_i/\kappa_i) - (\tilde{\alpha}_i/2\kappa_i^2)$ defined as the interest rate for bonds with infinite maturity.

The functional forms for transition matrix Φ and the variance of the measurement errors of the transition equation Q_t in the three-dimensional case are given by Duan and Simonato (1999), that is,

$$\Phi = \begin{bmatrix} e^{-\kappa_1 h} & 0 & 0 \\ 0 & e^{-\kappa_2 h} & 0 \\ 0 & 0 & e^{-\kappa_3 h} \end{bmatrix} \tag{21}$$

and

$$Q_t = \begin{bmatrix} Q_{1,t} & 0 & 0 \\ 0 & Q_{2,t} & 0 \\ 0 & 0 & Q_{3,t}, \end{bmatrix} \tag{22}$$

where

$$Q_{i,t} = \frac{\tilde{\alpha}_i}{2\kappa_i}(1 - e^{-2\kappa_i h}). \tag{23}$$

The parameters that I estimate are the long-term average short rate A_0 $(= \theta)$, mean-reversion κ, market price of risk ψ and variance parameter $\tilde{\alpha}$. I assume that the factors are uncorrelated, i.e. Σ is the identity matrix, to allow for a tractable analytic expression for the models.

3.4 The non-Gaussian Case: The CIR Model

The second approach deals is the Cox et al. (1985) approach which deals with non-Gaussian factors. Chen and Scott (1992) made a multifactor Cox et al. (1985) representation where the short rate r_t follows the stochastic differential equation with k mutually independent factors,

$$dr_{i,t} = \kappa_i(\theta_i - r_{i,t})dt + \sqrt{\beta_i r_{i,t}}dW_{i,t}, \quad \text{for } i = 1, \dots, k \tag{24}$$

where $dW_{i,t}$ are independent Wiener processes. The drift factor $\kappa_i(\theta_i - r_{i,t})$ is equal to the drift factor in the Vasicek (1977) model which guarantees mean reversion of the yield by mean-reversion κ_i to the long-term average level θ_i. The fluctuating variance parameter is β_i. The factors in the CIR model do not follow a normal distribution and are therefore considered to be non-Gaussian.

As in the previous subsection, I present the three-factor ($k = 3$) model, in this case the model of Cox et al. (1985), which can be easily simplified to the one-dimensional and the two-dimensional case. The term structure equals $P_t(\tau) = \exp[-\tilde{A}(\tau) - B(\tau)r_t]$, with the same definition for $\tilde{A}(\tau)$ and $B(\tau)$ as in equations (17) and (18), respectively.

However, in the Cox et al. (1985) case the volatility fluctuates ($\tilde{\alpha} = 0$), whereas in the Vasicek (1977) case the volatility is constant ($\beta = 0$). Therefore, the specification of the functions $\tilde{A}_i(\tau)$ and $B_i(\tau)$ differs in the Cox et al. (1985) case where $\tilde{A}_i(\tau)$ as well as $B_i(\tau)$ are derived from a direct generalization of the regular Cox et al. (1985) equations,

$$\tilde{A}_i(\tau) = -\frac{2\tilde{\phi}_i}{\beta_i}\ln\left(\frac{2\gamma_i e^{[(\kappa_i^*+\gamma_i)/2]\tau}}{(\kappa_i^* + \gamma_i)(e^{\gamma_i\tau} - 1) + 2\gamma_i}\right), \tag{25}$$

$$B_i(\tau) = \frac{2(e^{\gamma_i\tau} - 1)}{(\kappa_i^* + \gamma_i)(e^{\gamma_i\tau} - 1) + 2\gamma_i}, \tag{26}$$

$$\kappa_i^* = \kappa_i + \psi_i\beta_i, \quad \tilde{\phi}_i \equiv \kappa_i \cdot \theta_i \quad \gamma_i \equiv \sqrt{(\kappa_i^*)^2 + 2\beta_i}, \tag{27}$$

as provided by Hull and White (1993) with κ^* as the speed of mean-reversion governed by $\mathbb{Q}$.

The specification for transition matrix Φ and the estimation error variance of the transition equation Q_t are the same as in equations (21) and (22). However, the specification of $Q_{i,t}$ is different and is defined as,

$$Q_{i,t} = r_{i,t}\frac{\beta_i}{\kappa_i}(e^{-\kappa_i h} - e^{-2\kappa_i h}) + \theta_i\frac{\beta_i}{2\kappa_i}(1 - e^{-\kappa_i h})^2. \tag{28}$$

The parameters that are estimated are the same as in the Vasicek (1977) case, except for the variance parameter $\tilde{\alpha}$ which is zero in the Cox et al. (1985) case. Instead, I estimate the fluctuating variance parameter β. Similar to the Vasicek case, I assume that the factors are uncorrelated.

3.5 The Kalman Filter And Quasi-Maximum Likelihood

Based on the state-space representation as presented in equations (14) and (15), the most suitable approach for estimating the parameters is quasi-maximum likelihood (QML) through the Kalman filter as devised by Hamilton (1995). I closely follow the terminology of De Jong (2000) for the appropriate Kalman filter equations in the ATSM. The primary prerequisites are the unconditional mean and unconditional variance, $\hat{F}_0 = \mathbb{E}(F_t)$ and $\hat{P}_0 = \mathrm{var}(F_t)$, respectively. Equation (16) shows the prediction step where via the transition equation, as presented in (15), the factors $F_{t|t-h}$ are predicted on the left and the covariance matrix of the transition equation $P_{t|t-h}$ is estimated in the equation on the right,

$$F_{t|t-h} = \Phi \hat{F}_{t-h} \quad \text{and} \quad P_{t|t-h} = \Phi \hat{P}_{t-h} \Phi' + Q_t, \tag{29}$$

with Φ the transition matrix, Q_t the measurement-error matrix, and h the step size between observations.

By performing QML, the likelihood contributions per observation are summed up and maximized. These (components of the) likelihood contributions are displayed in equation (30) where u_t, V_t, and $-2\ln L_t$, can be interpreted as the prediction error of the observation equation, the variance of the observation equation and an alteration of the standard likelihood, respectively,

$$u_t = y_t - A - BF_{t|t-h}, \quad V_t = BP_{t|t-h}B' + H \quad \text{and} \quad -2\ln L_t = \ln|V_t| + u_t' V_t^{-1} u_t, \tag{30}$$

where $(y_t, A, B, F_{t|t-h}, P_{t|t-h}, H)$ are respectively the yield vector, factor intercept, factor loadings, predicted factors, predicted transition equation variance and the observation equation measurement-error matrix. The Kalman filter as devised by Hamilton (1995) is a repetitive algorithm that updates the factor predictions and the transition equation variance. The components of the updating step are,

$$K_t = P_{t|t-h} B' V_t^{-1}, \quad \hat{F}_t = F_{t|t-h} + K_t u_t \quad \text{and} \quad \hat{P}_t = (I - K_t B) P_{t|t-h}, \tag{31}$$

where $K_t, \hat{F}_t, \hat{P}_t$ are respectively known as the Kalman gain, the updated factor and the updated variance estimate. The identity matrix is denoted by I and the other parameters are the same as in (30).

The steps outlined above comprise the Kalman filter as used by De Jong (2000), which I apply to acquire parameter estimates and maximize the likelihood function. However, to examine the contribution of each additional factor of which more detailed explanations are provided in section 4.3, the Kalman smoother is required. The Kalman smoother is just like the Kalman filter an iterative algorithm, but it iterates backwards and provides smoothed predictions of the factors along with smoothed predictions of the variance of the transition equation. The Kalman smoothing steps are,

$$\hat{F}_{t-h|T} = \hat{F}_{t-h} + \hat{P}_{t-h} \Phi' P_{t|t-h}^{-1} (\hat{F}_{t|T} - \hat{F}_{t|t-h}), \tag{32}$$

$$\hat{P}_{t-h|T} = \hat{P}_{t-h} - \hat{P}_{t-h} \Phi' P_{t|t-h}^{-1} (P_{t|t-h} - P_{t|T}) P_{t|t-h}^{-1} \Phi \hat{P}_{t-h}. \tag{33}$$

where $\hat{F}_{t-h|T}$ and $\hat{P}_{t-h|T}$ are the smoothed factor estimates and the smoothed transition equation variance, respectively. The other parameters are the same as in (29) and (31).

I obtain asymptotic standard errors from the Fisher information matrix when assessed at the QML estimates for parameter vector ν of the model. Because I minimize the negative log-likelihood -lnL, the hessian matrix $H(\nu)$ at the maximum likelihood estimates

$$H(\nu) = \frac{\delta^2}{\delta\nu_i \delta\nu_j}(\text{-lnL}), \tag{34}$$

is identical to the Fisher information matrix. Hence, I determine for the optimal parameter vector $\hat{\nu}_{ML}$ the covariance matrix by $\text{Var}(\hat{\nu}_{ML}) = [H(\hat{\nu}_{ML})]^{-1}$, and the estimated parameters are consequently asymptotically normally distributed

$$\hat{\nu}_{ML} \overset{a}{\sim} N(\nu_0, \text{Var}(\hat{\nu}_{ML})), \tag{35}$$

with the square root of the diagonal components of $\text{Var}(\hat{\nu}_{ML})$ as standard errors.

If the factors and measurement errors adhere to normal distributions, De Jong (2000) argues that the Kalman filter QML estimates are efficient as well as consistent. However, the distribution of the factors in the case of many ATSMs does not adhere to the normal distribution. Nevertheless, De Jong (2000) claims that the estimated parameter vector retrieved by the Kalman filter is probably consistent because of the QML assumption, when the factor's mean and variance are appropriately constructed. However, there is a slight issue with the argument of De Jong (2000) that appears since the factor's variance is determined by prevailing values of the factors of which exact measurement is impossible since they are underlying factors. Hence, the conditional variance employed in the likelihood contributions is imprecise. Another issue is the undermining of the rules for the Kalman filter updating step. This issue arises since the distribution of the factors in the case of many ATSMs does not adhere to the normal distribution. Consequently, the QML parameter vector retrieved by the Kalman filter is inconsistent.

Therefore, to evaluate the bias in the Kalman filter QML estimator, De Jong (2000) implements a modest Monte Carlo simulation analysis on the basis of which he concludes that the Kalman filter QML estimator does not contain a systematic bias. He solely finds significant overestimation involving speed of mean-reversion variable κ, although the bias appears quite limited. Hence, De Jong (2000) provides evidence for the findings of Lund (1997) who indicates that the bias is quite limited for variables obtained from the Kalman filter QML estimator during estimation of ATSMs. As a result, in line with De Jong (2000), I abstain from applying numerically exhaustive simulation-based estimation procedures as used by Gallant and Tauchen (1996), Gourieroux et al. (1993), Frachot et al. (1995), Dai and Singleton (2000), and Lamoureux and Witte (2002). Instead, I present the Kalman filter QML estimates.

3.6 Estimation Methodology And Constraints

As stated before, the most straightforward method of estimating a factor model is by picking an equal amount of yields as the amount of distinct maturities. Then by inversion of the model, the factors can be extracted. However, the preference of maturities to compose the factors is quite subjective while the results of the model are dependent on the preference. Furthermore, solely using the same number of maturities as factors overlooks possible valuable information in the alternative maturities.

Therefore, in line with De Jong (2000), I incorporate more yields than maturities and assume that all yields, i.e. yields with a maturity of (3,12,60,120) months, are observed with error. With respect to optimization, I apply the "*fminunc*" minimization algorithm in "*MATLAB*" to minimize the negative log-likelihood function. After performing the optimization, I multiply the negative log-likelihood by minus one to obtain the regular log-likelihood for each model. To ensure that the long-term average short rate A_0, the speed of mean-reversion κ, as well as the variance parameters $\tilde{\alpha}_i$ and β_i remain positive during the optimization, I perform a logarithmic transformation.

With respect to the Cox et al. (1985) factor models, I need to make an additional constraint due to the square-root diffusion process as stated in (24). Since β_i is positive, the short rate r_t also needs to be positive to ensure that no complex numbers arise from taking the square-root of a negative number. As stated before, the short rate depends on the factors, i.e. $r_t = A_0 + \iota' F_t$, which must stay positive. During my research I experimented with conditions, such as if statements, where some of the factors are allowed to become negative while the short rate remained positive. However, there is a problem with this approach when generating data in the simulation study. Since I generate the factors by implementing an Euler discretization from the diffusion process in (12) with α equal to zero imposed, the factors must be constrained to be positive to prevent complex numbers in the simulated data. Hence, to resolve this problem I take the absolute value of the factors to ensure positive factors after each updating step in the Kalman filter as stated in (31).

4 Simulation Study

I use a simulation study to examine the Vasicek (1977) and Cox et al. (1985) three-factor models since this version outperforms the versions with less factors in the research of De Jong (2000). I examine these models while subject to a "*regular period*" and a "*zero-lower bound period*". First, it introduces the setup of my simulation study to enable the reader to replicate the results. Then, it presents the results of the ordinary simulation to investigate the performance under the two different periods of the three-factor models for both specifications. I also perform a cross-simulation where I estimate the Vasicek (1977) model on data generated by the Cox et al. (1985) model and reversed.

4.1 Setup

To simulate data for both models, I follow De Jong (2000) who performs a small Monte Carlo analysis to examine a bias in the QML estimates of an affine one-factor ATSM. First, I simulate the latent factors from the diffusion process (3) for which I implement an Euler discretization scheme where each of the simulated months contains 25 in-between steps. The diffusion process in (3) simplifies for both models since β is zero for the Vasicek (1977) model whereas $\tilde{\alpha}$ is zero for the Cox et al. (1985) model. In the latter model, I restrict the factors to be positive since $\tilde{\alpha}$ is zero and β is positive to obtain feasible numbers from the square-root part of (3). For each month, I store the values of the factors and compute the yields for maturities of (3,12,60,120) months by implementing the model in (8). Moreover, I simulate measurement errors from the normal distribution with expectation zero and variance H for every observation. In line with De Jong (2000), I assume that the measurement error covariance matrix is constructed as $H = h^2 I$ where h is the step size per latent factor observation. This allows the measurement errors for the distinct maturities to have an identical variance and to be serially and cross-sectionally uncorrelated. Every simulation for the four distinct maturities results in a sample of 221 monthly observations, equal to the length of a period in the empirical study, and involves four distinct maturities. I estimate the model parameters $(A_0, \kappa_i, \tilde{\alpha}_i, \psi_i)$ and $(A_0, \kappa_i, \beta_i, \psi_i)$ on a 1000 simulations by implementing the Kalman-filter QML estimator in respectively the Vasicek (1977) and Cox et al. (1985) case for $i = 3$ factors.

Regarding the population parameters underlying the simulated data for both the regular and cross-simulation, I vary the long-term average short rate A_0 as well as the variance parameters $\tilde{\alpha}$ or β to replicate the regular and the zero-lower bound period. Based on prior research, I set A_0 equal to 0.06 under the regular and to 0.01 in the zero-lower bound period. From the empirical study in section 5, I obtain the population parameters for the market price of risk and the mean-reversion by taking the average of both ψ_i and κ_i from the including and excluding zero-lower bound period to obtain legitimate base case values. For the Vasicek (1977) model, there appears to be a ratio of 1/3 between the $\tilde{\alpha}$ estimates for the two different periods. Therefore, I set the population parameters $\tilde{\alpha}$ to (1,2,6) which are approximately the estimates for the period excluding and obtain the population parameters for the zero-lower bound period by multiplying them by 1/3. For the Cox et al. (1985) model, this ratio is not that clear. Therefore, I take the average of the estimates on the real data for β for both periods and double the average variance under the regular and decrease the average variance by a half under the zero-lower bound period.

4.2 Results

As explained in section 4.1, I obtain the population true values from the empirical study and subsequently simulate data sets from the Vasicek (1977) as well as the Cox et al. (1985) three-factor specification under the regular and zero-lower bound period. Afterwards, I estimate the three-factor models a 1000 times under the regular and zero-lower bound period.

Table 1: Simulation results Vasicek model

This table presents the population values underlying the data generating process for the 1000 simulations as well as the proceeding mean, median and standard deviation for the three-factor Vasicek model under the regular and the zero-lower bound (ZLB) period. As before, the parameters A_0 and $\tilde{\alpha}_i$ are respectively multiplied by $10^2, 10^4, 10^2$. The superscript * means an extra multiplication by 10 per superscript * to show the numbers behind the fourth decimal.

period	Statistic	A_0	κ_1^*	κ_2	κ_3	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	ψ_1	ψ_2	ψ_3
Regular	True Value	6.0000	0.0045	0.3370	3.2410	1.0000	2.0000	6.0000	1.3765	-21.1550	-17.9250
	Mean	6.0170	0.0249	0.4085	3.9685	0.1857	0.5017	2.7590	1.9545	-20.8791	-18.1328
	Median	5.9667	0.0048	0.4164	3.6628	0.2098	0.5882	3.1544	1.3760	-21.1562	-17.9259
	Std. Dev.	0.8218	0.0747	0.0754	1.1761	0.0908	0.2933	1.4558	34.9331	7.6184	5.9770
ZLB	True Value	1.0000	0.0045	0.3370	3.2410	0.3333	0.6667	2.0000	1.3765	-21.1550	-17.9250
	Mean	0.9895	0.0601	0.4979	4.0083	0.0735	0.1429	0.6014	-1.1663	-22.7084	14.7562
	Median	0.9646	0.0049	0.4596	3.9560	0.0636	0.1241	0.5259	1.3894	-21.1559	-17.9244
	Std. Dev.	0.3558	0.1763	0.2342	1.1596	0.0531	0.1216	0.4152	166.9726	15.1508	303.2229

Table 1 presents the simulation results of the three-factor Vasicek (1977) model which has problems with estimating some of the parameters, especially the mean-reversion κ_i and variance $\tilde{\alpha}_i$ are severely biased under both periods. However, the median values for the other parameters are quite acceptable. Interestingly, the market price of risk parameters' standard deviation for the Vasicek (1977) specification under the zero-lower bound period are substantially larger than for the regular period. The mean values for the market price of risk parameters are severely biased due to extreme outliers that for example cause the average ψ_1 to be negative. Hence, the Vasicek (1977) model performs somewhat better under the regular period although there are some severe biases.

Table 2: Simulation results CIR model

This table presents the population values underlying the data generating process for the 1000 simulations as well as the proceeding mean, median and standard deviation for the three-factor CIR model under the regular and the zero-lower bound period. As before, the parameters A_0 and β_i are respectively multiplied by $10^2, 10^4, 10^2$. The superscript * means an extra multiplication by 10 per superscript * to show the numbers behind the fourth decimal.

period	Statistic	A_0	κ_1	κ_2	$\kappa_3{}^*$	β_1^*	β_2^{**}	β_3^{**}	ψ_1	ψ_2	ψ_3
Regular	True Value	6.000	0.0035	0.3903	3.2165	4.4780	100.9570	0.2500	4.5450	-8.0200	-17.2800
	Mean	6.0392	0.0036	0.4147	3.1016	4.0223	91.8780	0.2507	4.5283	-8.0089	-17.2771
	Median	6.0739	0.0036	0.4059	3.0759	4.0754	92.1945	0.2500	4.5346	-8.0119	-17.2792
	Std. Dev.	0.1874	0.0001	0.0496	0.2437	0.2362	3.3534	0.0029	0.0275	0.0143	0.0073
ZLB	True Value	1.000	0.0035	0.3903	3.2165	1.1195	27.3925	0.0625	4.5450	-8.0200	-17.2800
	Mean	0.9127	0.0036	0.4289	3.3324	1.0737	26.5069	0.0624	4.5421	-8.0178	-17.2795
	Median	0.9173	0.0036	0.4265	3.3148	1.0776	26.4680	0.0624	4.5448	-8.0195	-17.2800
	Std. Dev.	0.0429	0.0001	0.0247	0.1559	0.0156	0.3590	0.0004	0.0191	0.0106	0.0042

Table 2 presents the simulation results of the three-factor Cox et al. (1985) model which performs quite well under both periods. It does not have the severe biases in estimating the mean-reversion κ_i as well as its variance of the factors β_i. The mean and median values for the parameters are quite similar under both periods although the bias for the average short rate A_0 is larger under the zero-lower bound than under the regular period. Besides, the standard deviation for all parameters appears somewhat smaller under the zero-lower bound period for the three-factor Cox et al. (1985) model.

Figure 1: Histograms for 1000 simulations of Vasicek model

This figure contains histograms on the parameters estimated on a 1000 simulations for the three-factor Vasicek model under the regular and zero-lower bound (ZLB) period. The red vertical line represents the population values as specified in table 1. The x-axis is scaled to a percentage bandwidth around the true value and hence the histograms might exclude outliers. This percentage for the other histograms is respectively $50, 50, 50, 50, 30, 30, 10, 10, 0.05, 0.01$ for each period. The top ten histograms refer to the regular period, whereas the other ten histograms refer to the ZLB period.

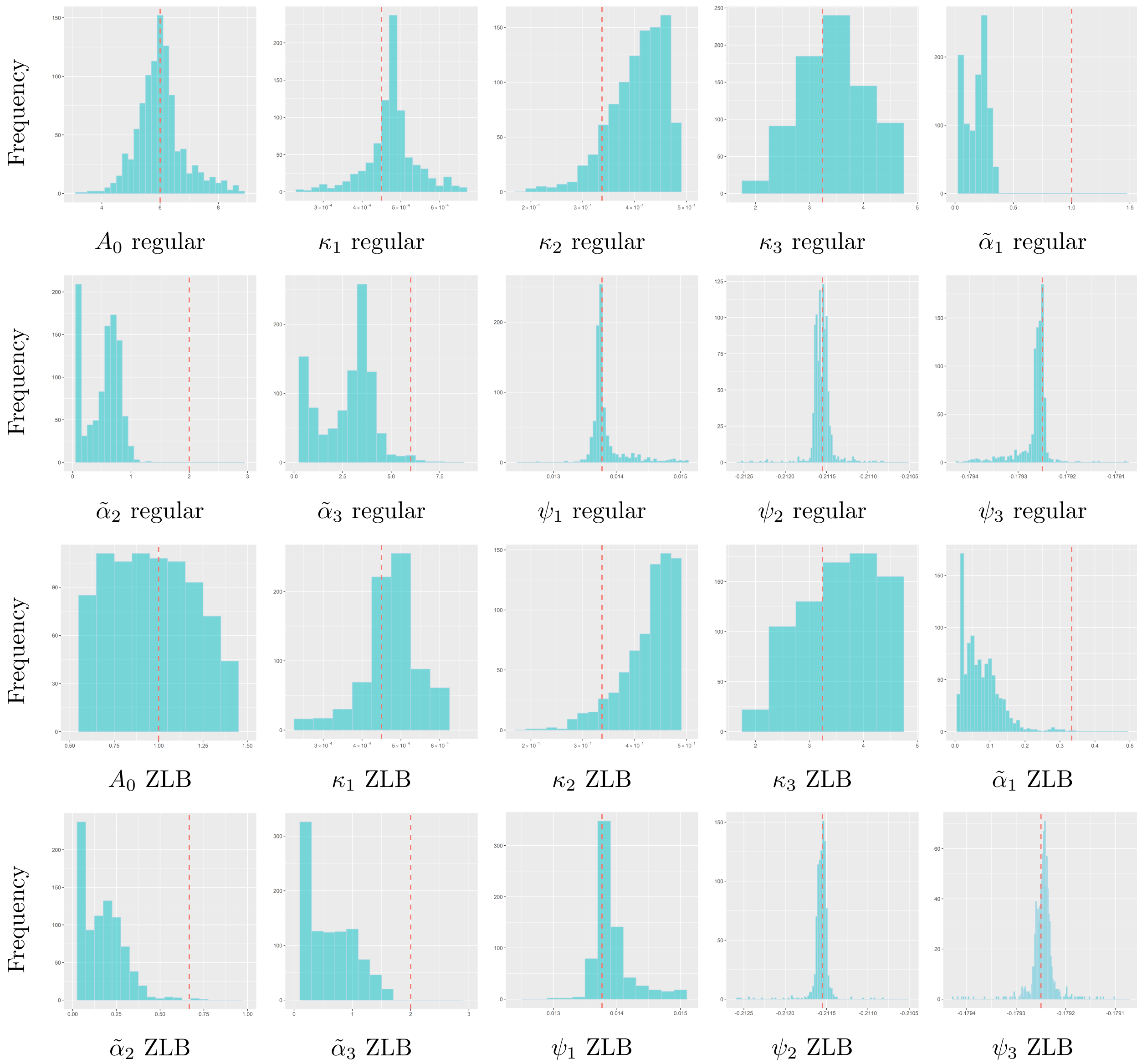


Figure 1 presents the histograms for the three-factor Vasicek (1977) model. Under the regular period, there is a substantial mass of estimates around the true value except for the variance parameters $\tilde{\alpha}$. Furthermore, we can see that the mass is somewhat more widespread under the zero-lower bound period.

Figure 2: Histograms for 1000 simulations of CIR model

This figure contains histograms on the parameters estimated on a 1000 simulations for the three-factor CIR model under the regular and zero-lower bound (ZLB) period. The red vertical line represents the population values as specified in table 2. The x-axis is scaled to a percentage bandwidth around the true value, except for β_i which has as domain $\in (0, 1.5 * \beta_i^p)$. Hence, the histograms might exclude outliers. This percentage for the other histograms is respectively $50, 50, 50, 50, 10, 0.05, 0.01$ for each period. The top ten histograms refer to the regular period, whereas the other histograms refer to the ZLB period.

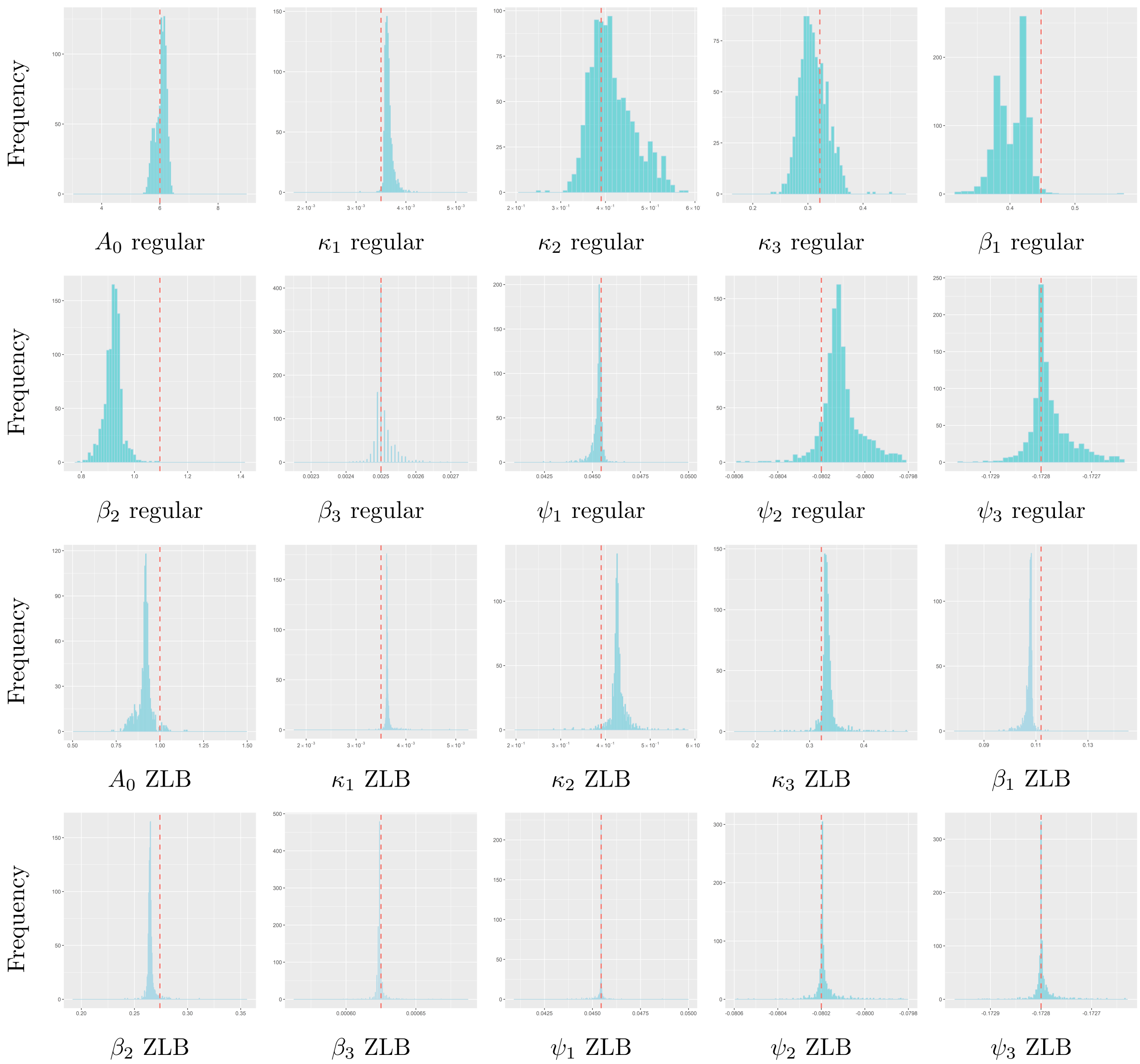


Figure 2 presents the histograms on the estimation results in the three-factor Cox et al. (1985) case. Under the regular period, there is a substantial mass of estimates around the true value except for the variance parameters β_1 and β_2. Furthermore, we can see that the mass is somewhat more widespread under the regular period. However, for almost all parameter estimates under the zero-lower bound period the mass does not incorporate the true value except for the market of price parameters ψ. Hence, the estimates under the zero-lower bound period for the three-factor Cox et al. (1985) model are heavily concentrated around a certain value that is not necessarily the true value of the parameter.

Overall, the Cox et al. (1985) model appears to be more accurate and less variable for most of the parameters than the Vasicek (1977) model which has problems estimating its variance parameters. Nevertheless, the bias on the average short rate A_0 is somewhat smaller for the Vasicek (1977) model.

For the cross-simulation, I also obtain the population true values from the empirical study and subsequently simulate data sets from the Vasicek (1977) as well as the Cox et al. (1985) three-factor specification under the regular and zero-lower bound period. However, I estimate the Vasicek (1977) model a 1,000 times on data generated by the Cox et al. (1985) model and reversed.

Table 3: Simulation results Vasicek model on CIR data

This table presents the population values underlying the data generating process for the 1000 simulations as well as the proceeding mean, median and standard deviation for the three-factor CIR model under the regular and the zero-lower bound period. As before, the parameters A_0 and $\tilde{\alpha}_i$ are respectively multiplied by 10^2 and 10^4. The superscript * means an extra multiplication by 10 per superscript * to show the numbers behind the fourth decimal.

period	Statistic	A_0	κ_1^{**}	κ_2	$\kappa_3{}^*$	$\tilde{\alpha}_1^{**}$	$\tilde{\alpha}_2^{**}$	$\tilde{\alpha}_3$	ψ_1	ψ_2	ψ_3
Regular	True Value	6.0000	0.3500	0.3903	0.3217	-	-	-	4.5400	-8.0200	-17.2800
	Mean	6.8420	0.0084	0.3304	0.2990	0.0115	0.0025	0.0033	3.3424	-7.9236	-17.4540
	Median	6.7891	0.0074	0.3193	0.3146	0.0098	0.0004	0.0012	3.3748	-7.9283	-17.3814
	Std. Dev.	0.1522	0.0058	0.0264	0.0372	0.0111	0.0062	0.0066	0.1693	0.0180	0.2083
ZLB	True Value	1.0000	0.3500	0.3903	0.3217	-	-	-	4.5400	-8.0200	-17.2800
	Mean	1.0431	0.0707	0.3169	0.3672	0.0016	0.0175	0.0145	-8.2847	-22.7445	-13.7108
	Median	1.0392	0.0286	0.3064	0.3799	0.0004	0.0103	0.0000	3.4287	-7.8793	-17.2868
	Std. Dev.	0.0813	0.0966	0.0413	0.0277	0.0030	0.0338	0.1461	50.4612	79.4864	24.7483

Table 3 presents the simulation results of the three-factor Vasicek (1977) model on CIR data which has severe problems with estimating the average short rate A_0, the mean-reversion of the first factor κ_1 and the first factor's market price of risk ψ_1 which are very biased in both periods. Interestingly, the market price of risk parameters' standard deviation are for the Vasicek (1977) specification in the zero-lower bound period substantially larger than in the regular period. The mean values for the market price of risk parameters are severely biased due to extreme outliers that for example cause the average ψ_1 to be negative in the former period. Actually, every parameter estimate is biased and are considerably more biased compared to the estimation results of the Cox et al. (1985) model on CIR data in Table 2. Hence, the Vasicek (1977) model is not able to outperform the Cox et al. (1985) model on CIR data.

Table 4: Simulation results CIR model on Vasicek data

This table presents the population values underlying the data generating process for the 1000 simulations as well as the proceeding mean, median and standard deviation for the three-factor Vasicek model under the regular and the zero-lower bound (ZLB) period. As before, the parameters A_0 and β_i are multiplied by 10^2. The superscript * means an extra multiplication by 10 per superscript * to show the numbers behind the fourth decimal.

period	Statistic	A_0	κ_1^*	κ_2	κ_3	β_1	β_2	β_3	ψ_1	ψ_2	ψ_3
Regular	True Value	6.0000	0.0045	0.3370	3.2410	-	-	-	1.3765	-21.1550	-17.9250
	Mean	6.5737	0.0062	0.6872	1.8382	0.0151	0.0339	0.0729	1.3237	-21.1799	-17.8889
	Median	6.4337	0.0047	0.3779	2.3251	0.0138	0.0238	0.0643	1.3700	-21.1546	-17.9236
	Std. Dev.	1.0343	0.0374	0.6063	1.0910	0.0058	0.0303	0.0413	0.2413	0.9183	0.8767
ZLB	True Value	1.0000	0.0045	0.3370	3.2410	-	-	-	1.3765	-21.1550	-17.9250
	Mean	0.8299	0.0052	0.6244	1.9160	0.0062	0.0085	0.0223	1.3258	-21.1402	-17.9078
	Median	0.9171	0.0047	0.4136	1.8805	0.0053	0.0067	0.0202	1.3631	-21.1540	-17.9237
	Std. Dev.	0.3418	0.0063	1.1671	1.1249	0.0035	0.0107	0.0108	0.2699	0.3322	0.3633

Table 4 presents the simulation results of the Cox et al. (1985) model on Vasicek data which performs better than the Vasicek (1977) model on CIR data although the average short rate A_0 and mean-reversion parameter κ_3 are still very biased. The mean and median values for the parameters are quite similar in both periods. The Cox et al. (1985) model obtains a better fit of the first two mean-reversion parameters on Vasicek data than the Vasicek (1977) model on the same data as visible in Table 1. However, the latter model has a considerably smaller bias for the average short rate A_0 although its market price of risk parameters' standard deviation is considerably larger. Overall, the results for estimating both models on Vasicek data is quite similar with biases for different parameters.

Figure 3: Histograms for 1000 cross simulations of Vasicek model on CIR data

This figure contains histograms on the parameters estimated on a 1000 simulations of CIR data for the three-factor Vasicek model under the regular and zero-lower bound (ZLB) period. The red vertical line represents the population values as specified in table 4. The purple vertical line represents the median value for the $\tilde{\alpha}$ parameters since the true value is unknown. The x-axis is scaled to a percentage bandwidth around the true value, except for $\tilde{\alpha}_i$ which has as domain $\in (0, d * \tilde{\alpha}_i^p)$ with $d = 4$ for all histograms except for $\tilde{\alpha}_3$ ZLB which has $d = 1000$. The domain of κ_1 is $\in (0, \kappa_1^p)$. Hence, the histograms might exclude outliers. This bandwidth percentage for the other histograms is respectively 50, 50, 50, 50, 50, 10 for each period. The top ten histograms refer to the regular period, whereas the other histograms refer to the ZLB period.

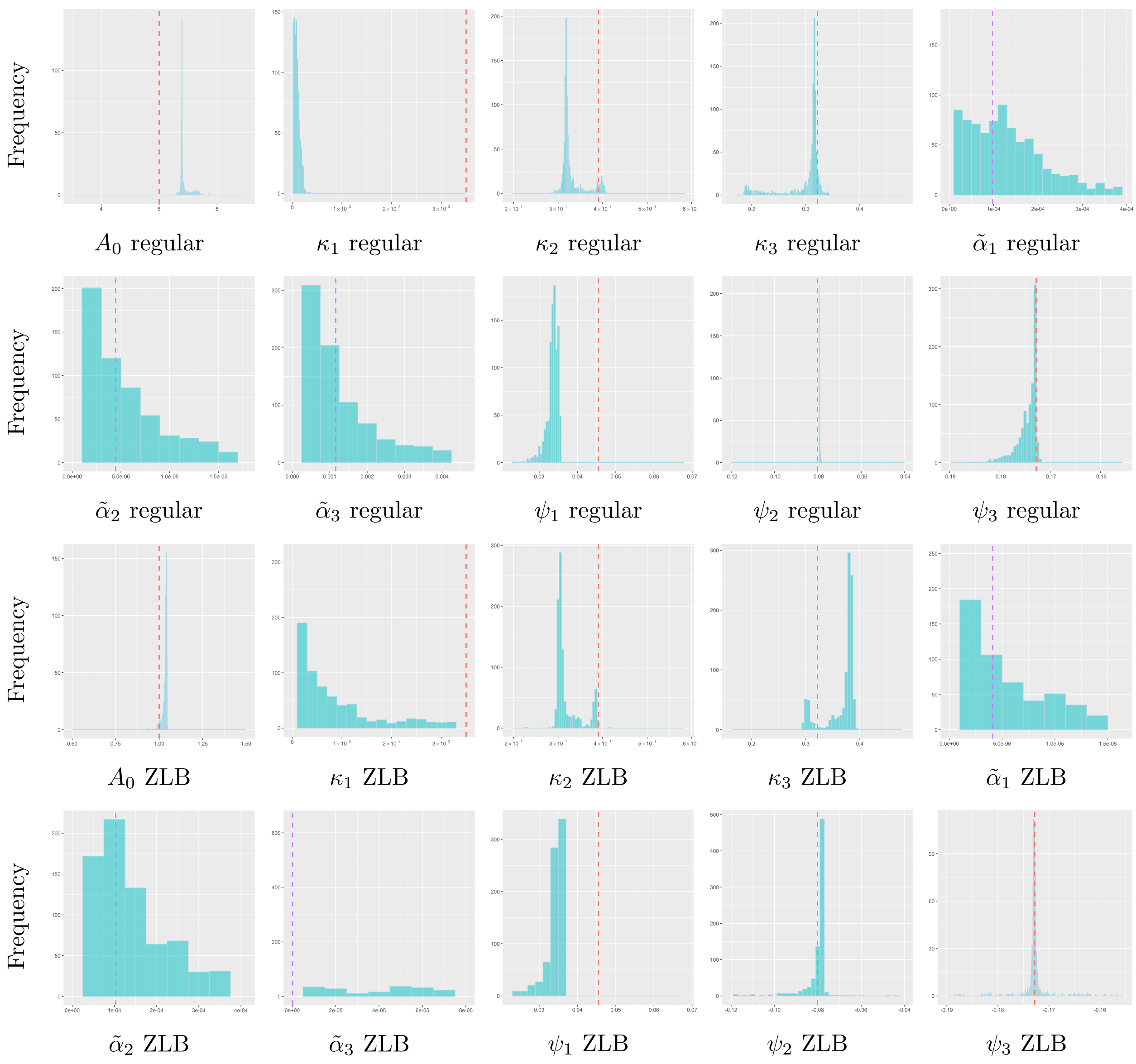


Figure 3 presents the histograms for the three-factor Vasicek (1977) model on CIR data. The red vertical line represents the true values as specified in table 3 and the purple vertical line represents the median value for the $\tilde{\alpha}$ parameters since the true value is unknown. In the regular period, there is only a substantial mass of estimates around the true value of mean-reversion κ_3 and the market price of risk parameters ψ_1 and ψ_2. In the zero-lower bound period, the mass is relatively somewhat closer to its true value for the average short rate A_0 and mean-reversion κ_1 while the estimate of κ_3 is worse. The estimate for $\tilde{\alpha}_3$ is very widespread considering its domain. Looking at Figure 2 clearly shows that the Cox et al. (1985) outperforms the Vasicek (1977) model on CIR data.

Figure 4: Histograms for 1000 cross simulations of CIR model on Vasicek data

This figure contains histograms on the parameters estimated on a 1000 simulations of Vasicek data for the three-factor CIR model under the regular and zero-lower bound (ZLB) period. The red vertical line represents the population values as specified in table 4. The purple vertical line represents the median value for the β parameters since the true value is unknown. The x-axis is scaled to a percentage bandwidth around the true value, except for β_i which has a bandwidth around its median value. Hence, the histograms might exclude outliers. This percentage for the histograms is respectively $50, 50, 100, 100, 50, 100, 50, 50, 0.05, 0.05$ for each period. The top ten histograms refer to the regular period, whereas the other histograms refer to the ZLB period.

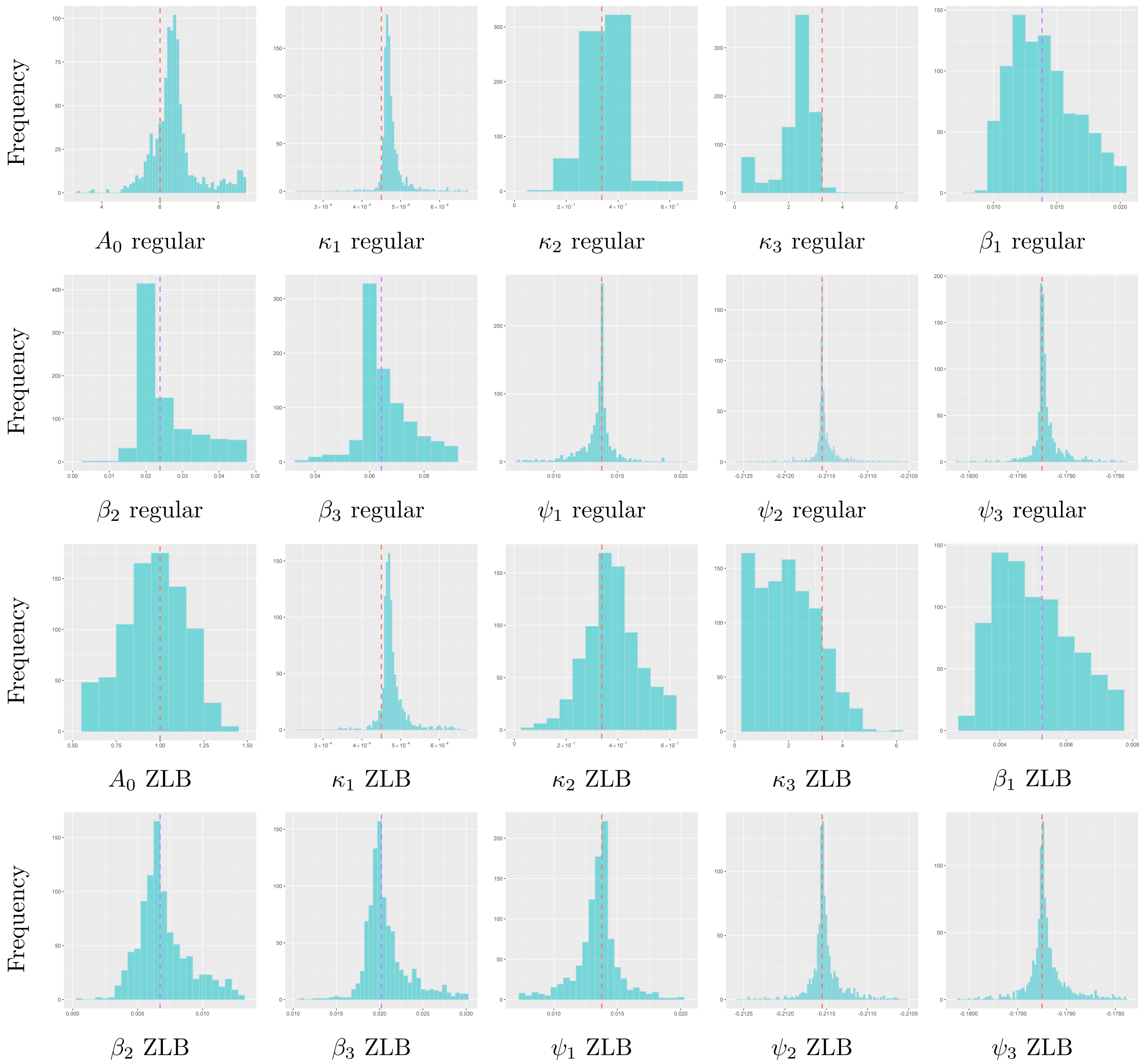


Figure 4 presents the histograms on the estimation results for the three-factor Cox et al. (1985) model when simulating data from the Vasicek (1977) model. The red vertical line represents the true values as specified in table 4 and the purple vertical line represents the median value for the β parameters since the true value is unknown. Under the regular period, there is a substantial mass of estimates around the true value for the parameters which is somewhat more widespread under the zero-lower bound period. Looking at Figure 4 shows that the estimation results for the Cox et al. (1985) model on Vasicek data are comparable in both periods with the Vasicek (1977) model on Vasicek data.

Overall, the Cox et al. (1985) model appears to be more accurate and less variable for all parameters than the Vasicek (1977) model on CIR data when comparing Table 2 and Figure 2 with Table 3 and Figure 3. Hence, the Vasicek (1977) model is not able to outperform the Cox et al. (1985) model on CIR data. However, the estimation performance of the Cox et al. (1985) and Vasicek (1977) are close when estimating on Vasicek data. The first model obtains a better fit of the first two mean-reversion parameters whereas the latter model has a considerably smaller bias for the average short rate although its market price of risk parameters' standard deviation is considerably larger. Consequently, performing an empirical study using real Treasury yield data might be interesting to examine the performance of the Vasicek (1977) and Cox et al. (1985) models while subject to the zero-lower bound.

5 Empirical Study

This section provides an extensive empirical study on the performance of the Vasicek (1977) and the Cox et al. (1985) one-, two- and three-factor models while subject to the zero-lower bound. First, it presents a detailed description regarding the setup for my empirical study to enable the reader to replicate the results. Second, it reviews the data and presents descriptive statistics used in this empirical study. Third, it presents the results of estimating the n-factor models for both specifications to assess the impact for $n = 1, 2, 3$ caused by the zero-lower bound.

5.1 Setup

To contribute to prior research, I assess the impact caused by the zero-lower bound on the Vasicek (1977) and Cox et al. (1985) model specifications through employing a non-Guassian estimation method, i.e. the Kalman filter approach of De Jong (2000). I assess this impact on the Gaussian model of Vasicek (1977) in comparison with the non-Gaussian model of Cox et al. (1985) because non-Gaussian models are unable to handle negative interest rates. Therefore, I estimate Vasicek (1977) and Cox et al. (1985) n-factor models for $n = 1, 2, 3$ on data excluding as well as including the zero-lower bound to investigate which factor model obtains the best fit and whether the existence of the zero-lower bound in the data set impacts the estimation performance.

However, a common problem with regard to estimating ATSMs is the number of factors to incorporate. Therefore, in the spirit of De Jong (2000), I perform an empirical study on the one-, two- and three-factor models of both Vasicek (1977) and Cox et al. (1985). I evaluate each factor model's fit containing n state variables graphically and by Root Mean Squared Error (RMSE) to assess every additional factor's contribution and whether the factors deserve the interpretation of respectively the level, slope and curvature factor, for n = 1, 2, 3.

I present a thorough specification study for every model by evaluating the residuals of each factor model of Vasicek (1977) and Cox et al. (1985). In addition, I compare the average actual and average fitted term structure for each factor model and I provide a comparison between the factors and their proxies as suggested in the literature. I also compare the actual yields with the fitted version of these yields for each factor model of both Vasicek (1977) and Cox et al. (1985) in the Appendix. Furthermore, I ensure that the short rate, $r_t = A_0 + \iota' F_t$, does not become negative by taking the absolute value of the factors in the Cox et al. (1985) model after the updating step in the Kalman filter.

I apply the "*fminunc*" minimization algorithm in "*MATLAB*" to minimize the negative log-likelihood function. Via the inverted Fisher information matrix, I take the square root of its diagonal to obtain the estimation error for the parameter vector. However, obtaining a proper inverse is known to be problematic for Vasicek (1977) and Cox et al. (1985) ATSMs due to a lot of free parameters and the computation by *fminunc* using finite differences. The resulting inverse often contains negative numbers on its diagonal resulting in complex standard errors for the estimated parameters. Therefore, I set the seed to "*rng(1)*" and run the models a 100 times from random initial values to obtain maximum likelihood estimates with feasible standard errors that do not differ much from similar estimates with complex standard errors.

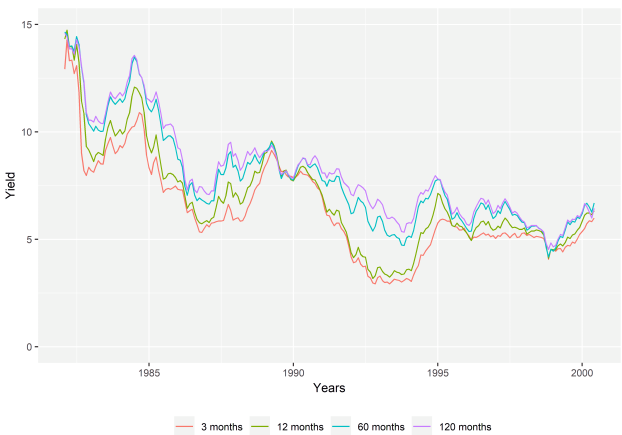
5.2 Data

I use the Federal Reserve Bank data available via the Wharton Research Data Services (WRDS). The data set encompasses U.S. Treasury interest rates with a constant maturity on a monthly basis and spans from January 1982 until October 2018 leading to 422 observations with maturities of (3,12,60,120) months. The lack of data for the shortest maturity prior to January 1982 determined the starting point of this data set. The choice for the maturities is in line with De Jong (2000) who also examines yields with maturities of (3,12,60,120) months. However, his data set spans from January 1970 until February 1991 and thus does not contain the recent occurrence of the zero-lower bound. He uses the McCulloch and Kwon (1993) zero-coupon bond data set consisting of zero-coupon bond yields which are computed by applying McCulloch's interpolation procedure from prices of coupon bonds.

As can be seen in Figure (A.1) in the Appendix, this time period includes the exceptional zero-lower bound phenomenon which spans approximately from August 2008 until March 2017. To examine the impact on the estimation methods of the zero-lower bound, I divided the data set into two equal time periods resulting in a time period excluding the zero-lower bound and a time period including the zero-lower bound. The former period spans from January 1982 until May 2000, the latter spans from June 2000 until October 2018. This division results in 221 observations for each period to ensure that both periods contain ample observations for adequate estimation.

Figure 5: U.S. Treasury yields with a constant maturity per period

On the left, this figure shows the U.S. Treasury yields with a constant maturity for the period excluding the zero-lower bound (ZLB). On the right, it shows the period including the ZLB. The excluding ZLB period spans from January 1982 until May 2000, where the including ZLB period spans from June 2000 until October 2018. The lines represent the yield to maturity for a maturity of 3, 12, 60, and 120 months.



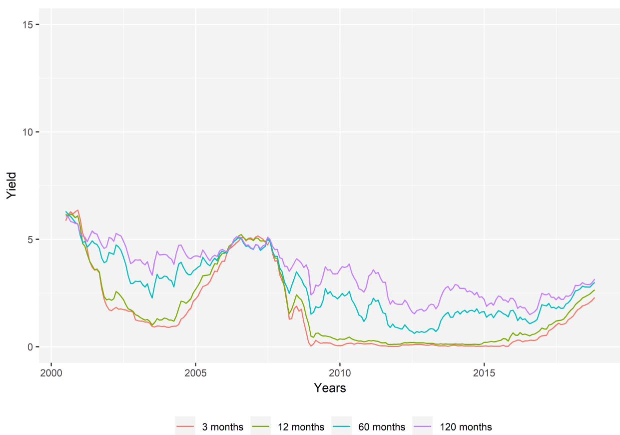


Figure 5 visualizes this division and depicts the first difference between the period excluding the zero-lower bound and the period including the zero-lower bound. It shows the U.S. Treasury yields with a constant maturity for the period excluding the zero-lower bound on the left and including the zero-lower bound on the right. The lines represent the zero-coupon bond yield for the maturities of (3,12,60,120) months. In the including zero-lower bound period, we observe as expected a clear shift in the level of the yields starting from the second half of the time period.

Table 5: Descriptive statistics

This table displays descriptive statistics on U.S. Treasury yields with a constant maturity for the period excluding the zero-lower bound (ZLB) and the period including the ZLB. The former period spans from January 1982 until May 2000, the latter spans from June 2000 until October 2018. The variables 3M, 12M, 60M and 120M represent the yield to maturity for a maturity of 3, 12, 60, and 120 months, respectively. This results in 221 observations for each yield to maturity within each period.

		Yield To Maturity (%)			
Period	**Statistic**	3M	12M	60M	120M
Excluding ZLB	Minimum	2.930	3.180	4.180	4.530
	1st Quartile	5.090	5.390	6.040	6.300
	Median	5.820	6.270	7.370	7.650
	Mean	6.374	6.883	7.839	8.138
	3rd Quartile	7.970	8.110	8.910	9.110
	Maximum	14.280	14.730	14.650	14.590
	Std. Dev.	2.234	2.395	2.360	2.300
	Observations	221	221	221	221
Including ZLB	Minimum	0.010	0.100	0.620	1.500
	1st Quartile	0.100	0.260	1.590	2.360
	Median	0.950	1.240	2.520	3.460
	Mean	1.545	1.762	2.741	3.439
	3rd Quartile	2.200	2.570	3.910	4.420
	Maximum	6.360	6.180	6.300	6.100
	Std. Dev.	1.789	1.739	1.394	1.167
	Observations	221	221	221	221

Table 5 presents descriptive statistics on both periods, i.e. the period including the zero-lower bound and the period excluding the zero-lower bound. The table shows two interesting differences between these two periods. First, as expected, the yields to maturity are substantially higher in the period excluding the zero-lower bound in comparison with the period including the zero-lower bound. Second, all the yields to maturity have a smaller standard deviation in the period including the zero-lower bound compared to the period excluding the zero-lower bound.

As already deduced from Figure 5, we can see a clear shift in the level of the yields in the including zero-lower bound period. We observe in Table 5 that the maximum yield to maturity for all maturities in the period excluding the zero-lower bound ranges from 14.280 to 14.730, whereas in the period including the zero-lower bound it ranges from 6.100 to 6.360. The minimum yield to maturity for all maturities in the period excluding the zero-lower bound ranges from 2.930 to 4.530, whereas in the period including the zero-lower bound it ranges from 0.010 to 1.5. Moreover, in Table 5 we observe the second difference between the two periods, i.e. the standard deviation of the yield to maturity for all maturities is lower for the period including the zero-lower bound. Specifically, for the period excluding the zero-lower bound, the yield to maturity standard deviation for all maturities ranges from 2.234 to 2.395 whereas for the period including the zero-lower bound it ranges from 1.167 to 1.789.

Hence, I examine the implications of the aforementioned two differences between the period including and the period excluding the zero-lower bound on estimating the one-, two-, and three-factor Vasicek (1977) as well as the one-, two-, and three-factor Cox et al. (1985) ATSMs by employing the non-Gaussian estimation method of De Jong (2000).

5.3 Results

This subsection presents the results of the empirical study by applying the Kalman filter approach of De Jong (2000) on the Vasicek (1977) and Cox et al. (1985) factor ATSMs. I demonstrate an extensive analysis for the one-, two-, and three-factor Vasicek (1977) and Cox et al. (1985) models including a comparison between the models where I examine whether there is an effect of the occurrence of the zero-lower bound on the estimation performance of these models. Hence, I first discuss the results for the period excluding the zero-lower bound, after which I present the results for the period including the zero-lower bound containing a comparison of the results for the two periods.

5.3.1 Results For Period Excluding The Zero-Lower Bound

The parameters that I estimate for the i-factor models are the long-term average short rate A_0, the speed of mean-reversion κ_i, market prices of risk ψ_i and variance parameters $\tilde{\alpha}_i$ and β_i, for $i = 1, 2, 3$. I provide estimates based on the excluding zero-lower bound period in table 6, for the i-factor models of Vasicek (1977) where the fluctuating variance parameter β is zero and the i-factor models of Cox et al. (1985) where the constant variance parameter $\tilde{\alpha}$ is zero. For both models, I assume that the factors are uncorrelated to allow for a tractable analytic expression for the models.

Table 6: Estimation results for factor models

This table presents the n-factor Vasicek and the n-factor CIR model QML estimation results, for $n = 1, 2, 3$. The table further provides the mean-reversion variable κ^* along with the half-life of the factors, $[\ln(2)/\kappa^*]$, governed by measure $\mathbb{Q}$. These results are based on the period excluding the zero-lower bound.

Model	Model	$A_0(\times 100)$	κ_i	$\tilde{\alpha}_i(\times 10^4)$	$\beta_i(\times 100)$	$\psi_i(\times 10^{-2})$	κ^*	2 ln L
One-factor	Vasicek	7.7423	0.0090	0.9820		-0.0998	0.0090	9169.55
		(5.0427)	(0.0129)	(52.0789)		(0.0036)	[77.40]	
	CIR	8.1673	0.0157		0.1160	-0.0917	0.0051	9208.00
		(1.2393)	(0.0131)		(0.5631)	(0.0017)	[44.15]	
Two-factor	Vasicek	1.6565	0.0098	0.9994		-0.1011	0.0098	10016.56
		(3.3478)	(0.0118)	(56.2993)		(0.0008)	[70.73]	
			0.6064	1.3571		-0.1558	0.6064	
			(0.0027)	(55.0532)		(0.0229)	[1.14]	
	CIR	4.3256	0.0022		0.2428	0.0460	0.0133	9904.65
		(2.5242)	(0.4387)		(9.6555)	(0.0075)	[51.97]	
			0.6102		0.5539	-0.1421	0.5315	
			(0.0366)		(8.5378)	(0.0251)	[1.30]	
Three-factor	Vasicek	0.0000	0.0007	0.9880		-0.0137	0.0007	10150.58
		(2.0518)	(0.0954)	(278.5723)		(0.0071)	[990.21]	
			0.4600	1.8167		-0.2064	0.4600	
			(0.0257)	(108.0129)		(0.0074)	[1.51]	
			3.6973	6.4548		-0.1362	3.6973	
			(0.0051)	(89.3922)		(0.0046)	[0.19]	
	CIR	2.6500	0.0038		0.1544	0.0282	0.0082	9848.22
		(1.1110)	(0.4243)		(8.0099)	(0.0045)	[84.69]	
			0.5108		0.6432	-0.0136	0.5020	
			(0.0717)		(8.0989)	(0.0061)	[1.38]	
			0.4838		0.0023	-0.1037	0.4836	
			(0.0887)		(157.9264)	(0.0616)	[1.43]	

For the one-factor Vasicek (1977) model the long-term average short rate is 7.7423, whereas for the Cox et al. (1985) model it is 8.1673. The estimated mean-reversion coefficients are quite small, 0.0090 in the Vasicek (1977) case and 0.0157 in the Cox et al. (1985) case. The implicit mean-reversion coefficient governed by risk-neutral measure $\mathbb{Q}$ is yet smaller in the Cox et al. (1985) model with 0.0051, indicating a 44.15 years half-life ($\ln(2)/\kappa^*$), whereas for the Vasicek (1977) model the half-life is 77.40 years. The moderate speed of mean-reversion indicates a substantially flat-fitted term structure. The constant variance parameter $\tilde{\alpha}_1$ in the Vasicek (1977) model is highly insignificant due to its large standard error, where the fluctuating variance parameter β_1 in the Cox et al. (1985) model is also insignificant although its standard error is substantially smaller. This might be evidence that a affine volatility structure which combines $\tilde{\alpha}_1$ and β_1 might be more appropriate. The market price of risk is significantly negative for both models which indicates that owning a large maturity bond requires a positive risk premium. The likelihood is slightly higher for the Cox et al. (1985) model indicating a better fit on the data.

With respect to the two-factor Vasicek (1977) and Cox et al. (1985) models, their long-term average short rate decreases in both models to 1.6565 and 4.3256, respectively. The obtained likelihood is in the two-factor case slightly higher for the Vasicek (1977) model. Moreover, the two factors in both models show clear distinct characteristics. The mean reversion is comparable for the first factor in both models, whereas the second factor has a much higher mean-reversion coefficient under $\mathbb{Q}$, i.e. 0.6064 and 0.5315 with half-lives 1.14 and 1.30 years, in the Vasicek (1977) and the Cox et al. (1985) case, respectively. Moreover, each models' variance parameters are insignificant where the variance for the second factor is larger in comparison with the first. Again, the standard error for the estimates of both variance parameters in the Vasicek (1977) are relatively large. Moreover, the Vasicek (1977) two-factor model measures for the first factor a significantly negative market price of risk of -0.1011 whereas the Cox et al. (1985) model measures a significantly positive market price of risk of 0.0460.

In the three-factor case, dissimilarities arise between the Vasicek (1977) and Cox et al. (1985) model. For instance, their long-term average short rate decreases even further in comparison with the two-factor case although for the latter model it declines less. It appears that A_0 becomes less relevant in the latter case since the first factor incorporates its level as visible in Figure 8. Moreover, for the Vasicek (1977) model, the first factor has a very small mean-reversion coefficient of 0.0007 with implied half-life of 990.21 whereas the third factor has a relatively large mean-reversion coefficient of 3.6973 with half-life 0.19 years. In the Cox et al. (1985) model, for the first and second factor the mean-reversion is comparable with the two-factor case and the third factor resembles the characteristics of the second factor. Furthermore, the instant variance of the third factor is substantially higher in the Vasicek (1977) case relative to the other factors, whereas it is lower in the Cox et al. (1985) case. The first factors' market price of risk display another difference, since ψ_1 is significantly positive in the Cox et al. (1985) case, whereas negative in the Vasicek (1977) case. Like the two-factor case, the Vasicek (1977) model obtains the highest likelihood.

Provided with the estimation results, I compose the models' residuals which are interchangeable with the Kalman filter estimation errors. I obtain these by calculating the discrepancy between the actual and the predicted interest rates. For an adequate model specification, the mean of the residuals for every maturity ought to be near zero. Moreover, there ought to be no serial correlation in the residuals.

Table 7: Residuals of the factor models

This table presents the residuals' summary statistics based on the n-factor Vasicek and the n-factor CIR model where ρ_k represent the serial correlation for order k. The residuals are scaled to percentage points. These residuals are based on the period excluding the zero-lower bound.

		Vasicek				CIR			
Model	**Statistic**	0.25	1	5	10	0.25	1	5	10
	Mean	-1.0685	-0.5963	0.1996	0.3720	-0.7475	-0.2723	0.5338	0.7084
	Stand. Dev.	1.0760	1.0240	0.6582	0.5714	1.1890	1.2030	1.0474	1.0296
	ρ_1	0.8235	0.7671	0.4740	0.3587	0.4250	0.3665	0.1926	0.2164
	ρ_{12}	0.3923	0.3641	0.0823	0.0236	0.1642	0.1273	0.0095	0.0444
One-factor									
	Corr. matrix	1.0000	0.9269	0.5543	0.2605	1.0000	0.9433	0.6900	0.5439
		0.9269	1.0000	0.7798	0.5100	0.9433	1.0000	0.8411	0.7036
		0.5543	0.7798	1.0000	0.9220	0.6900	0.8411	1.0000	0.9704
		0.2605	0.5100	0.9220	1.0000	0.5439	0.7036	0.9704	1.0000
	Mean	-0.0124	0.1194	0.1344	0.1185	0.1592	0.2809	0.2135	0.1506
	Stand. Dev.	0.7380	0.8566	0.8738	0.8211	0.9142	1.0050	0.9044	0.8154
	ρ_1	0.1836	0.2721	0.2815	0.2191	0.1678	0.2688	0.2824	0.2301
	ρ_{12}	-0.0039	0.0337	0.0418	0.0399	0.0055	0.0389	0.0319	0.0271
Two-factor									
	Corr. matrix	1.0000	0.8798	0.8181	0.8385	1.0000	0.9256	0.8542	0.8686
		0.8798	1.0000	0.9780	0.9752	0.9256	1.0000	0.9647	0.9583
		0.8181	0.9780	1.0000	0.9928	0.8542	0.9647	1.0000	0.9940
		0.8385	0.9752	0.9928	1.0000	0.8686	0.9583	0.9940	1.0000
	Mean	-0.0639	-0.0267	-0.0061	0.0145	0.1502	0.2758	0.2051	0.1602
	Stand. Dev.	0.9570	1.0184	1.0009	0.9849	0.9168	1.0099	0.9145	0.8288
	ρ_1	0.1543	0.0827	0.0384	0.0323	0.1726	0.2776	0.3207	0.3073
	ρ_{12}	-0.0064	-0.0184	-0.0162	-0.0113	0.0044	0.0392	0.0232	0.0275
Three-factor									
	Corr. matrix	1.0000	0.9777	0.9396	0.9421	1.0000	0.9237	0.8428	0.8445
		0.9777	1.0000	0.9795	0.9804	0.9237	1.0000	0.9668	0.9616
		0.9396	0.9795	1.0000	0.9957	0.8428	0.9668	1.0000	0.9953
		0.9421	0.9804	0.9957	1.0000	0.8445	0.9616	0.9953	1.0000

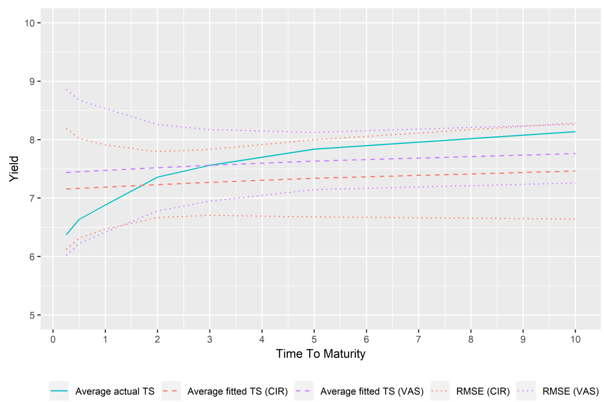
Table 7 presents summary statistics on the residuals for the n-factor models of both specifications, for $n = 1, 2, 3$. The one-factor Vasicek (1977) model on average overestimates the yields of shorter maturities, whereas it underestimates yields for longer maturities. The Cox et al. (1985) one-factor model shows the same pattern, although its over- and underestimation is more balanced. The standard deviation for the Vasicek (1977) model tends to decrease for longer maturities, whereas the Cox et al. (1985) case it remains high. The first- and twelfth-order autocorrelation also tend to decrease to a lower level for longer maturities for the Vasicek (1977) model, when in fact for the Cox et al. (1985) model they are lower for almost all maturities and remain constant at this lower level.

The two-factor models obtain on average a substantially better fit than the one-factor models. The Vasicek (1977) model slightly underestimates every maturity except the three-month maturity. The Cox et al. (1985) model underestimates on average a bit more but still there is quite an enhancement visible. Moreover, the standard deviation of the residuals is somewhat smaller for both models and the serial correlation, which is of greater importance, is substantially lower for both models.

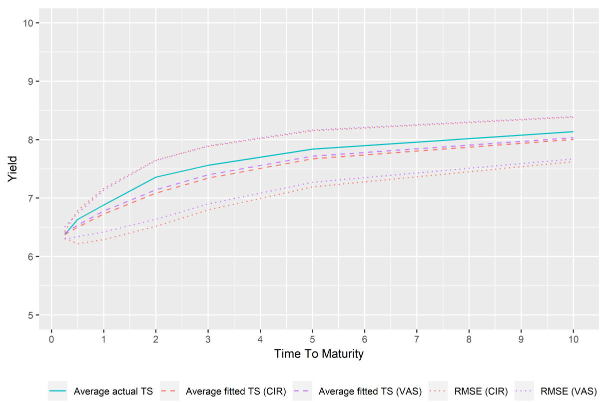
In the three-factor specification, the fit is on average substantially better in the Vasicek (1977) case as well as somewhat better in the Cox et al. (1985) case. However, the residuals' standard deviation increased for the former whereas it remained constant for the latter model. The first-order autocorrelation decreases for the Vasicek (1977) specification, whereas it remains constant on average for the Cox et al. (1985) specification when compared to the two-factor case. Importantly, the twelfth-order autocorrelation for both models is near zero. Judging by the average mean and serial correlation of the residuals, the Vasicek (1977) three-factor model appears to the best specified.

Figure 6: Fit of the factor models

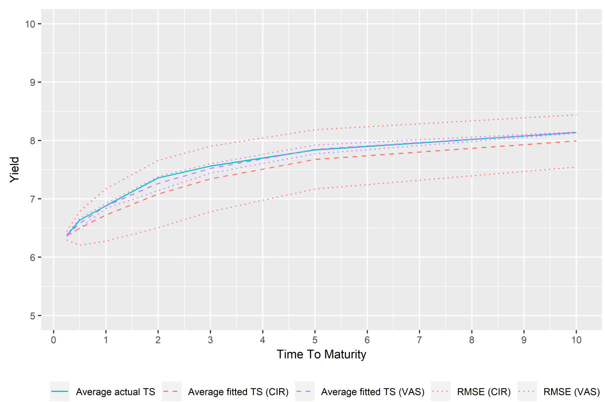
These graphs present, regarding both the n-factor models of Vasicek and CIR, the average fitted as well as the average actual term structure (TS), for $n = 1, 2, 3$. In addition, the root mean squared error (RMSE) is displayed for each model. The graphs are based on the period excluding the zero-lower bound.



One-factor models



Two-factor models

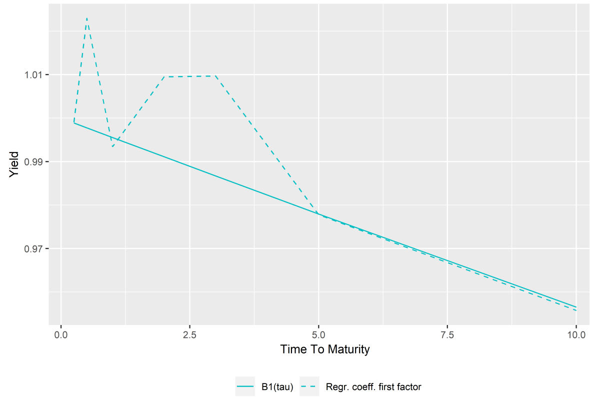


Three-factor models

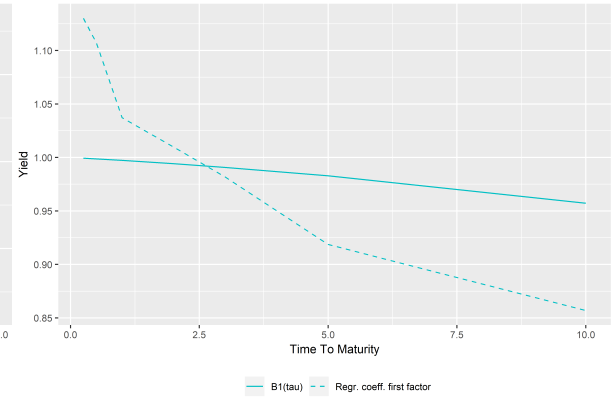
Figure 6 illustrates the average actual and the average fitted term structure for the maturities of (3,6,12,24,36,60,120) months for each model. I compute the fitted term structure by $y_{t|T}(\tau) = A(\tau) + B(\tau)F_{t|T}$ for all maturities τ. By implementing the Kalman smoother of Hamilton (1995), I acquire the smoothed factor estimates, $F_{t|T}$. The graphs also contain the RMSE of the discrepancy between the actual and the predicted interest rates. The one-factor models obtain a poor fit which is visible in the upper-left graph where the Vasicek (1977) model fixates on the term structure's longer end, whereas the Cox et al. (1985) on the shorter end. Adding a second factor enhances the fit considerably although the steepness of the actual term structure in the middle is not completely seized. Of the two models, the three-factor Vasicek (1977) model appears to only capture this steepness with a relatively tiny RMSE.

Figure 7: Regression of observed yields on the fitted factor(s)

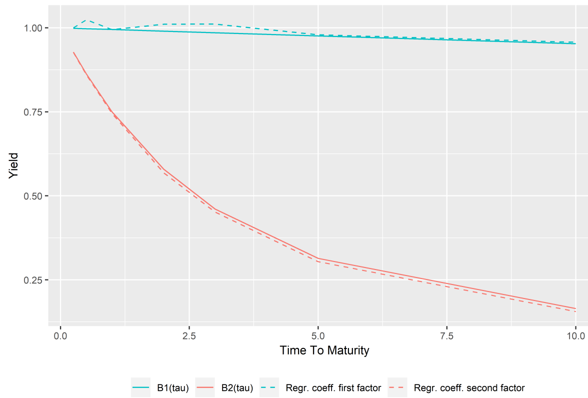
These graphs present the coefficients of a regression for the n-factor Vasicek and the n-factor CIR model in first differences of the actual interest rates on the fitted factor(s) and a constant, for $n = 1, 2, 3$. The graphs are based on the period excluding the zero-lower bound.



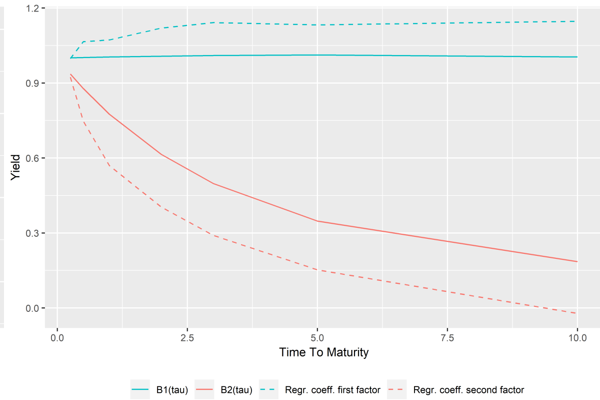
One-factor Vasicek model



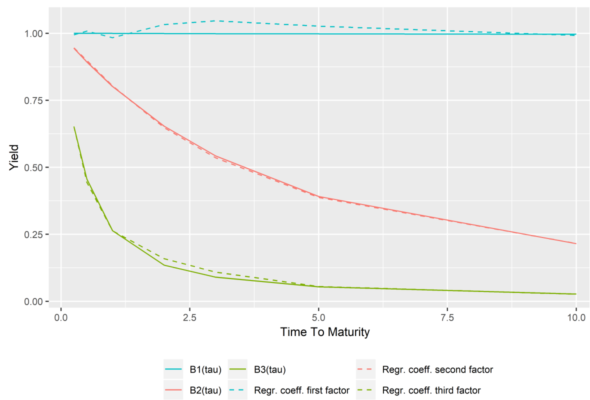
One-factor CIR model



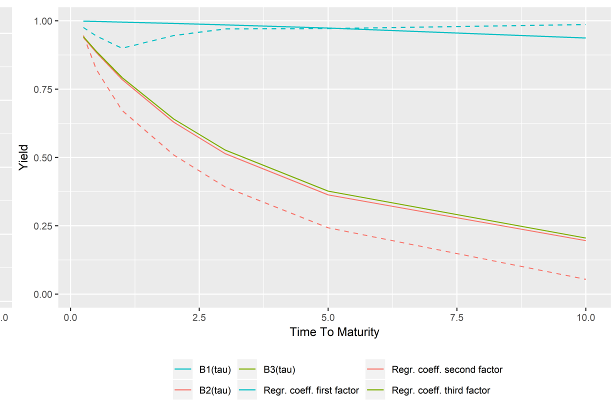
Two-factor Vasicek model



Two-factor CIR model



Three-factor Vasicek model



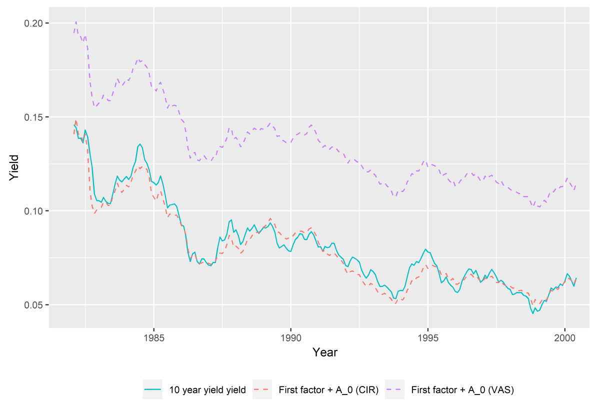
Three-factor CIR model

To further assess the model specification, I perform a regression on the estimated smoothed factors and a constant where the actual yields are the dependent variables. I carry out the regression in first differences since the data is practically non-stationary. The factor loadings $B(\tau)$ are illustrated by the solid lines in Figure 7 which ought to converge to zero for large maturities. However, the first factor loading $B_1(\tau)$ is substantially flat for all models. Although $B_2(\tau)$ decreases much faster, it is still significantly above zero. In the three-factor Vasicek model $B_3(\tau)$ appears to be negligible for longer maturities.

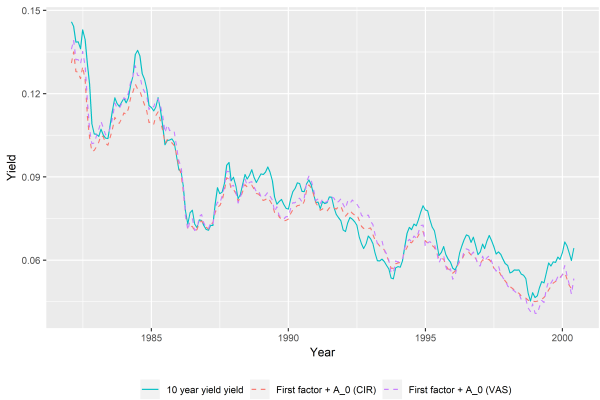
Moreover, the regression coefficients ought to be close to the factor loadings $B(\tau)$. In the upper-left corner in Figure 7, we see that this holds for longer maturities in the one-factor Vasicek (1977) model whereas it is not for the one-factor Cox et al. (1985) model displayed in the upper-right corner. Looking at the middle- and lower-left graphs, we see that the Vasicek (1977) two- and three-factor models' regression coefficients are comparable with their respective factor loadings. The two- and three-factor Cox et al. (1985) model are displayed in the middle- and lower-right graphs, which does not illustrate the third smoothed factor since the fit is substantially bad, as visible in Figure (A.2) in the Appendix.

Figure 8: Fit of the factors

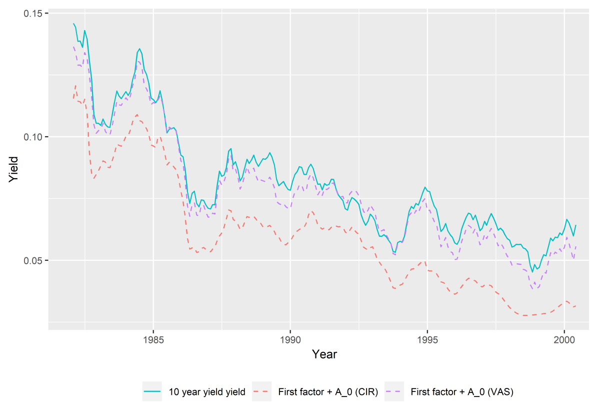
These graphs present the fit of the n-factor Vasicek and the n-factor CIR model, for $n = 1, 2, 3$. The graphs are based on the period excluding the zero-lower bound.



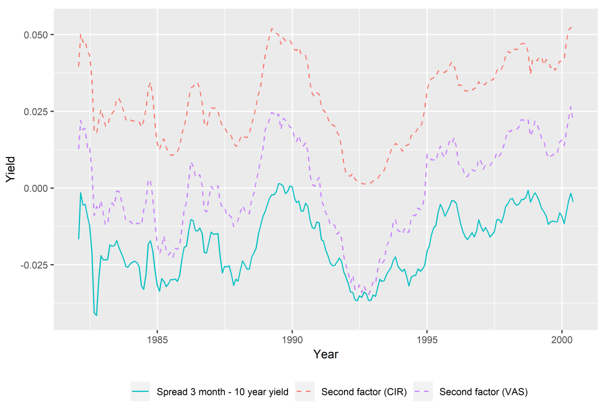
First factor one-factor models



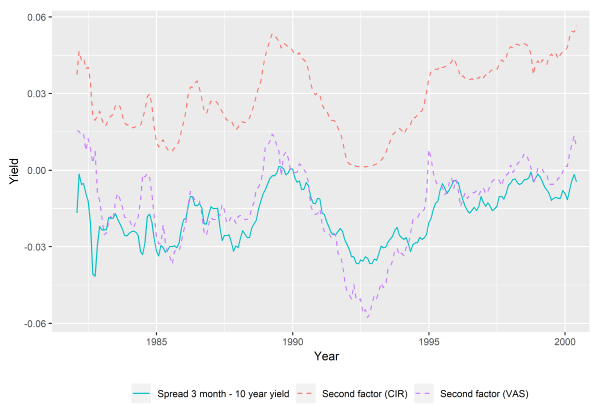
First factor two-factor models



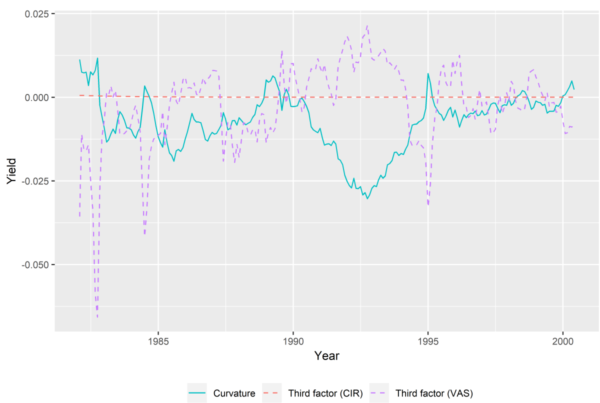
First factor three-factor models



Second factor two-factor models



Second factor three-factor models



Third factor three-factor models

Figure 8 illustrates the fitted factors along with parts of the data to interpret the factors. The upper-left, upper-right and middle-left graphs illustrate the fit of the first factor for the n-factor models, for respectively $n = 1, 2, 3$. The fitted first factor is defined as the first factor plus the estimated intercept A_0 which is compared to the ten-year yield. The correlation between these two lines is very high for each model: 0.9919 versus 0.9817, 0.9768 versus 0.9796, 0.9889 versus 0.9707, for the one-, two- and three-factor Vasicek (1977) model versus the one-, two- and three-factor Cox et al. (1985) model, respectively. Therefore, the first factor represents for each model the level of the yield curve.

In the middle-right and lower-left graph of Figure 8 we see the three-month and ten-year interest rate spread in comparison with the second factor for the n-factor model of Vasicek (1977) and of Cox et al. (1985), for respectively $n = 2, 3$. The correlation between these two lines is 0.9430 versus 0.9424, 0.7801 versus 0.9308, for the two- and three-factor Vasicek (1977) model versus the two- and three-factor Cox et al. (1985) model, respectively. Hence, the interpretation of the second factor enjoys substantial support as the yield curve's slope in all models although there is somewhat less support for this interpretation for the three-factor Vasicek (1977) model. The reason that the second factor in the two- and three-factor Cox et al. (1985) model appears to be increased by a certain level compared to the proxy for the second factor is because in the Cox et al. (1985) the factors are restricted to be positive.

To interpret the third factor, I graph the third factor and compare it to a proxy for the curvature in the lower-right panel. I construct the curvature proxy by subtracting the total of the ten-year plus the three-month yield from twice the two-year yield. The correlation is -0.6389 in the Vasicek (1977) case which can be interpreted as a reversed curvature effect, whereas the correlation is 0.2541 in the Cox et al. (1985) case which is more smooth than the curvature proxy. The support for third factor's interpretation as the curvature factor appears relatively little given the correlation between the third factor of the Cox et al. (1985) model and its proxy.

5.3.2 Results For Period Including The Zero-Lower Bound

Similar to the period excluding the zero-lower bound, the parameters that I estimate for the i-factor models in the period including the zero-lower bound are the long-term average short rate A_0, the speed of mean-reversion κ_i, market prices of risk ψ_i and variance parameters $\tilde{\alpha}_i$ and β_i, for $i = 1, 2, 3$. However, in this section I explicitly focus on the best performing models in the excluding zero-lower bound period, i.e. the three-factor models, to examine potential differences due to the zero-lower bound in the data set. The results for the one- and two-factor models are stated in the tables (A.1) and (A.2) and in the figures (A.3), (A.4) and (A.5) in the Appendix. I have also incorporated the results for the three-factor models in the Appendix to enable the reader to easily compare the results between all factor models.

I provide estimates based on the including as well as the excluding zero-lower bound period in table 8 on the next page, for the three-factor model of Vasicek (1977) where the fluctuating variance parameter β is zero and the three-factor model of Cox et al. (1985) where the constant variance parameter $\tilde{\alpha}$ is zero. The results for both periods are displayed to allow the reader to easily compare the results for the three-factor model of both Vasicek (1977) and Cox et al. (1985) in the two periods. For both models, I assume that the factors are uncorrelated to allow for a tractable analytic expression for the models.

Table 8: Estimation results three-factor models in both periods

This table presents the three-factor Vasicek and the three-factor CIR model QML estimation results. The table further provides the mean-reversion variable κ^* along with the half-life of the factors, $[\ln(2)/\kappa^*]$, governed by measure $\mathbb{Q}$. These results are based on the period excluding and including the zero-lower bound (ZLB).

Period	Model	$A_0(\times 100)$	κ_i	$\tilde{\alpha}_i(\times 10^4)$	$\beta_i(\times 100)$	$\psi_i(\times 10^{-2})$	κ^*	2 ln L
		0.0000	0.0007	0.9880		-0.0137	0.0007	10150.58
		(2.0518)	(0.0954)	(278.5723)		(0.0071)	[990.21]	
			0.4600	1.8167		-0.2064	0.4600	
	Vasicek		(0.0257)	(108.0129)		(0.0074)	[1.51]	
			3.6973	6.4548		-0.1362	3.6973	
Excl. ZLB			(0.0051)	(89.3922)		(0.0046)	[0.19]	
		2.6500	0.0038		0.1544	0.0282	0.0082	9848.22
		(1.1110)	(0.4243)		(8.0099)	(0.0045)	[84.69]	
	CIR		0.5108		0.6432	-0.0136	0.5020	
			(0.0717)		(8.0989)	(0.0061)	[1.38]	
			0.4838		0.0023	-0.1037	0.4836	
			(0.0887)		(157.9264)	(0.0616)	[1.43]	
		0.0011	0.0002	0.3520		0.0379	0.0002	10415.59
		(3.4894)	(0.0454)	(124.9400)		(0.0079)	[3465.74]	
			0.2140	0.9750		-0.2167	0.2140	
	Vasicek		(0.0048)	(61.9792)		(0.0105)	[3.24]	
			2.7848	1.1581		-0.2223	2.7848	
Incl. ZLB			(0.0053)	(46.2474)		(0.8235)	[0.25]	
		1.9275	0.0032		0.2934	0.0627	0.0216	9769.90
		(1.2903)	(0.1245)		(4.4439)	(0.0025)	[32.14]	
	CIR		0.2698		0.4525	-0.1468	0.2034	
			(0.0138)		(1.3897)	(0.0026)	[3.41]	
			0.1595		0.0002	-0.2419	0.1595	
			(0.0628)		(41.5707)	(0.0338)	[4.35]	

For the including zero-lower bound period, some dissimilarities arise between the models. For instance, their long-term average short rate decreases even further in comparison with the two-factor case although for the Cox et al. (1985) model it declines less. Moreover, for the Vasicek (1977) model, the first factor has a very small mean-reversion coefficient of 0.0002 with implied half-life of 3465.74 whereas the third factor has a relatively large mean-reversion coefficient of 2.7848 with half-life 0.25 years. In the Cox et al. (1985) model, for the first and second factor the mean-reversion is comparable with the two-factor case and the third factor resembles the characteristics of the second factor. Furthermore, the third factor's instant variance is substantially higher in the Vasicek (1977) case relative to the other factors, whereas it is lower in the Cox et al. (1985) case. In both models, the first factor's market price of risk is positive. In addition, the Cox et al. (1985) model obtains a lower log-likelihood.

When comparing the results of the two periods, we observe that the characteristics of the parameters of the three-factor models for the including zero-lower bound period appear to be quite similar to the results for the period excluding the zero-lower bound. Overall, both the three-factor model of Vasicek (1977) and of Cox et al. (1985) find that there is a lower mean-reversion for all the factors in the including zero-lower bound period. In addition, they both find in according with the descriptive statistics a lower variance for their factors in the including zero-lower bound period as depicted by the lower $\tilde{\alpha}$ and β parameters except for the first factor in the Cox et al. (1985) case.

Table 9: Residuals of the three-factor models in both periods

This table presents the residuals' summary statistics based on the three-factor Vasicek and the three-factor CIR model where ρ_k represent the serial correlation for order k. The residuals are scaled to percentage points. These residuals are based on the period excluding and including the zero-lower bound (ZLB).

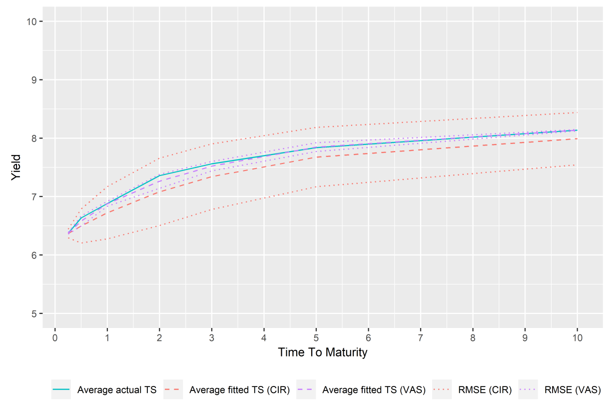
		Vasicek				**CIR**			
Period	**Statistic**	0.25	1	5	10	0.25	1	5	10
	Mean	-0.0639	-0.0267	-0.0061	0.0145	0.1502	0.2758	0.2051	0.1602
	Error %	-1.0025	-0.3879	-0.0778	0.1782	2.3564	4.0070	2.6164	1.9685
	Stand. Dev.	0.9570	1.0184	1.0009	0.9849	0.9168	1.0099	0.9145	0.8288
	ρ_1	0.1543	0.0827	0.0384	0.0323	0.1726	0.2776	0.3207	0.3073
Excl. ZLB	ρ_{12}	-0.0064	-0.0184	-0.0162	-0.0113	0.0044	0.0392	0.0232	0.0275
	Corr. matrix	1.0000	0.9777	0.9396	0.9421	1.0000	0.9237	0.8428	0.8445
		0.9777	1.0000	0.9795	0.9804	0.9237	1.0000	0.9668	0.9616
		0.9396	0.9795	1.0000	0.9957	0.8428	0.9668	1.0000	0.9953
		0.9421	0.9804	0.9957	1.0000	0.8445	0.9616	0.9953	1.0000
	Mean	-0.0116	-0.0216	-0.0514	-0.0348	-0.0156	-0.0378	-0.0108	0.0036
	Error %	-0.7508	-1.2259	-1.8752	-1.0119	-1.0097	-2.1453	-0.3940	0.1047
	Stand. Dev.	0.4322	0.4453	0.5769	0.5289	0.4406	0.4818	0.6660	0.7404
	ρ_1	0.1205	0.0641	0.4224	0.3813	0.1883	0.3012	0.7070	0.8346
Incl. ZLB	ρ_{12}	0.0186	-0.0048	0.1654	0.1430	0.0377	0.0567	0.2469	0.3647
	Corr. matrix	1.0000	0.9722	0.6484	0.6449	1.0000	0.9266	0.4406	0.2071
		0.9722	1.0000	0.7409	0.7349	0.9266	1.0000	0.5763	0.2940
		0.6484	0.7409	1.0000	0.9868	0.4406	0.5763	1.0000	0.9153
		0.6449	0.7349	0.9868	1.0000	0.2071	0.2940	0.9153	1.0000

Table 9 presents summary statistics on the three-factor models' residuals for both periods. In the including zero-lower bound period, the fit for both models is quite sufficient where the Cox et al. (1985) model's fit seems to be a little bit better since its mean residuals are closer to zero on average. The first-order autocorrelations decrease in the Vasicek (1977) compared to the two-factor case as visible in Table (A.2) and are considerably lower than in the Cox et al. (1985) case. Importantly, the twelfth-order autocorrelations for the Vasicek (1977) model are on average substantially closer to zero than the Cox et al. (1985) model. Judging by the average mean and serial correlation of the residuals, the Vasicek (1977) three-factor model appears to the better specified in the including zero-lower bound period.

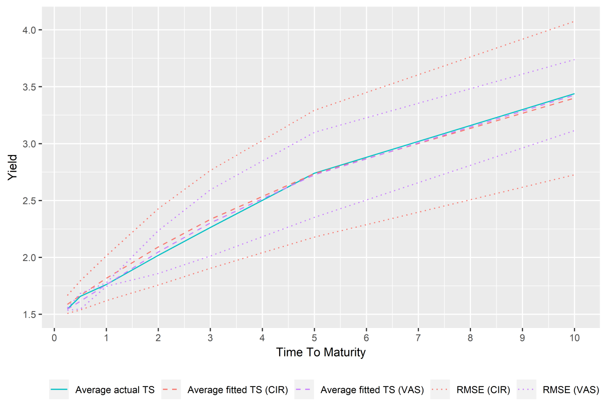
To compare the residuals of the periods, we look at the mean error percentage constructed as the mean error for a particular yield divided by its mean value as stated in Table 5. The Vasicek (1977) model obtains on average a lower error percentage in the excluding, whereas the Cox et al. (1985) obtains it in the including zero-lower bound period although the difference with the Vasicek (1977) model is smaller than in the excluding zero-lower bound period. The first- and twelfth-order autocorrelations are closer to zero in the period excluding than in the period including the zero-lower bound. Especially, the twelfth-order autocorrelations for the Vasicek (1977) model are on average closer to zero. Although the error percentage is somewhat smaller in the Cox et al. (1985) case for the including zero-lower bound period, there is substantially more serial correlation and therefore the Vasicek (1977) three-factor model appears to the better specified albeit substantially less than for the period excluding the zero-lower bound.

Figure 9: Fit of the three-factor models in both periods

These graphs present, regarding both the three-factor models of Vasicek and CIR, the average fitted as well as the average actual term structure (TS). In addition, the root mean squared error (RMSE) is displayed for each model. The graphs are based on the period excluding and including the zero-lower bound (ZLB).



Excluding ZLB



Including ZLB

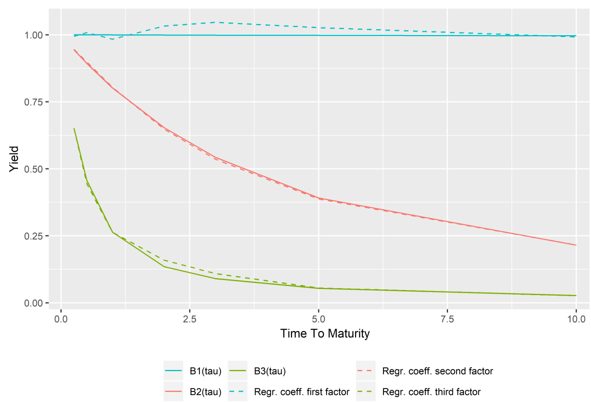
Figure 9 illustrates the average actual and the average fitted term structure for the maturities of (3,6,12,24,36,60,120) months for each model. The third factor provides a considerably better fit in both periods as visible in Figures 6 and (A.3). In both periods, the three-factor models appear to capture the yield curve quite well where the Vasicek (1977) model has a relatively smaller RMSE than the Cox et al. (1985) model which is the smallest in the period excluding the zero-lower bound.

As before, I perform a regression on the estimated smoothed factors and a constant where the actual yields for every maturity from three months to ten years are the dependent variables. The factor loadings $B(\tau)$ for both periods are illustrated by the solid lines in Figure 10 on the next page which ought to converge to zero for large maturities. For both models, we see that the first factor loading $B_1(\tau)$ is substantially flat in both periods. For the Vasicek (1977) model, the third factor loading $B_3(\tau)$ in the including zero-lower bound period appears to take over the pattern of the second factor loading $B_2(\tau)$ in the excluding zero-lower bound period whereas the second factor loading $B_2(\tau)$ in the three-factor model converges to zero for longer maturities in line with theory. The third factor loading $B_3(\tau)$ in the Cox et al. (1985) model has an unconventional form in both periods not in line with theory as visible in Figure (A.2) in the Appendix. Moreover, the latter model appears better specified in the excluding compared to the including zero-lower bound period since its factor loadings converge faster to zero. Overall, it appears the the Vasicek (1977) model is better specified in both periods.

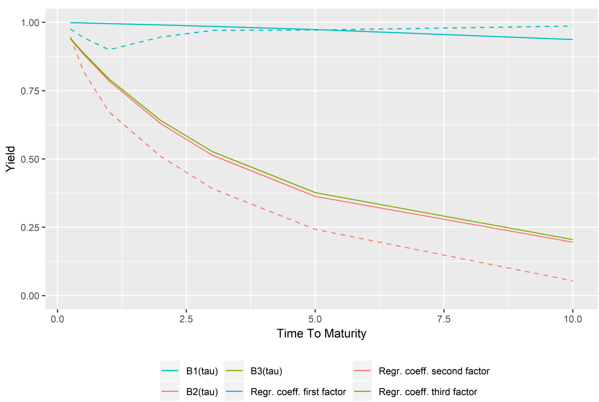
In addition to the theoretical convergence to zero of the factor loadings $B(\tau)$ for large maturities, the regression coefficients ought to be close to the factor loadings $B(\tau)$. In Figure 10, we see that the Vasicek (1977) model's regression coefficients are comparable with their respective factor loadings in both periods. However, the fit for the first factor in the including zero-lower bound period is substantially worse compared to the period excluding the zero-lower bound. The three-factor Cox et al. (1985) model illustrates in both periods a somewhat bad fit for the second factor as well as a substantially bad fit for the third smoothed factor as visible in Figure 10 and (A.2), respectively. Based on the condition that the regression coefficients ought to be close to the factor loadings $B(\tau)$, it seems that again the Vasicek (1977) model appears to be better specified in both periods.

Figure 10: Regression of observed yields on the fitted factors

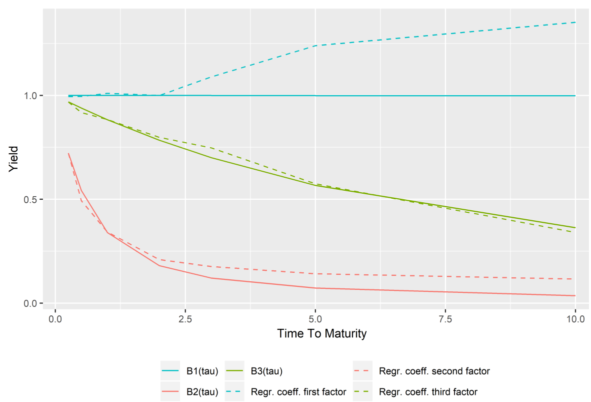
These graphs present the coefficients of a regression for the three-factor Vasicek and the three-factor CIR model in first differences of the actual interest rates on the fitted factor(s) and a constant for both the period including and excluding the zero-lower bound (ZLB).



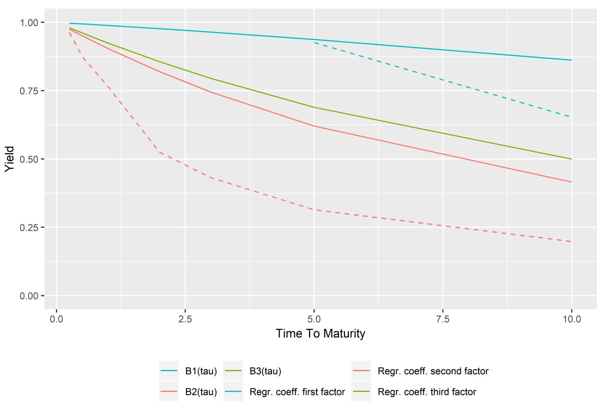
Three-factor Vasicek model in excl. ZLB period



Three-factor CIR model in excl. ZLB period



Three-factor Vasicek model in incl. ZLB period



Three-factor CIR model in incl. ZLB period

Figure 11 situated on the next page, illustrates the fitted factors along with parts of the data to interpret the factors. The fitted first factor is defined as the first factor plus the estimated intercept A_0 which is compared to the ten-year yield. The first factor represents the level of the yield curve in both periods for both three-factor models due to the high correlation between the first factor and its proxy. However, the evidence for this representation is stronger for the period excluding than for the period including the zero-lower bound since the first factor and its proxy's correlation is 0.9889 versus 0.9707 and 0.6848 versus 0.8040, respectively.

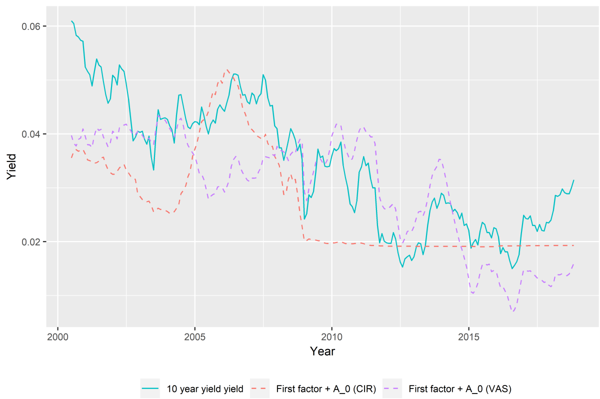
To interpret the second factor, I compare the second factor with the three-month and ten-year interest rate spread. The correlation for respectively the Vasicek (1977) and Cox et al. (1985) model between the second factor and its proxy is 0.7801 versus 0.9308 and 0.8771 versus 0.7893 for the period excluding and including the zero-lower bound, respectively. Hence, the interpretation of the second factor enjoys substantial support in both periods as the yield curve's slope for both models. The reason that the second factor in the three-factor Cox et al. (1985) model appears to be increased by a certain level compared to the proxy for the second factor is because in the Cox et al. (1985) case the factors are restricted to be positive. Furthermore, it appears that the fitted first and second factor remain flat during the zero-lower bound time frame in the Cox et al. (1985) case.

I graph the third factor and compare it to a proxy for the curvature defined as before. The correlation for respectively the Vasicek (1977) and Cox et al. (1985) model between the third factor and its proxy is -0.6389 versus 0.2541 and -0.5621 versus 0.0643 for respectively the period excluding and including the zero-lower bound. Hence, in the Vasicek (1977) case the third factor can be interpreted as a reversed curvature effect in both periods, whereas for the Cox et al. (1985) the evidence for the interpretation of the third factor as the curvature factor appears relatively little for both periods.

Figure 11: Fit of the factors for the three-factor models

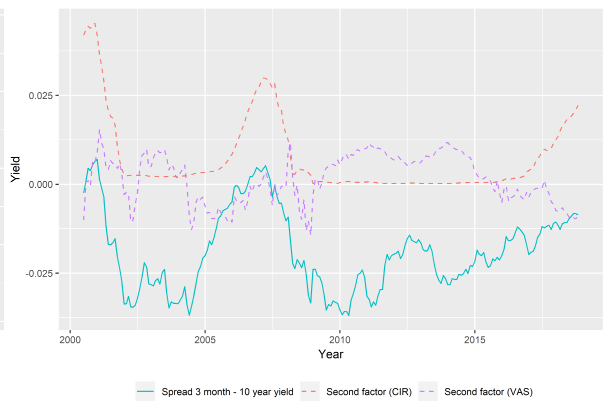
These graphs present the fit of the three-factor Vasicek and the three-factor CIR model. The graphs are based on the period excluding and including the zero-lower bound (ZLB).

First factor in excl. ZLB period



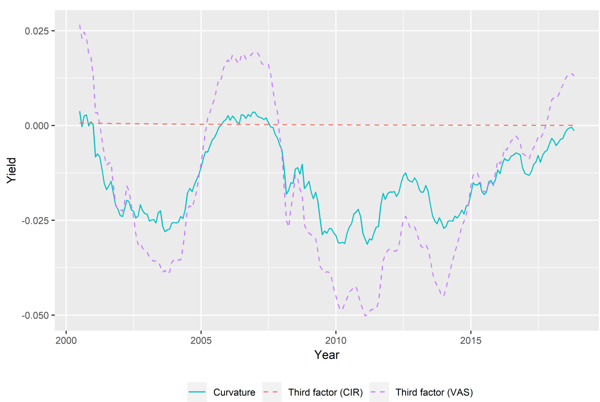
First factor in incl. ZLB period

Second factor in excl. ZLB period



Second factor in incl. ZLB period

Third factor in excl. ZLB period



Third factor in incl. ZLB period

The impact of the zero-lower bound on the three-factor models is also visible in in Figure (A.8) in the Appendix which contains the fitted yield versus the actual yield. In the period excluding the zero-lower bound, the Vasicek (1977) obtains a somewhat better fit whereas in the including zero-lower bound period the fit is substantially better than for the Cox et al. (1985) model which has substantial problems with capturing the curvature of the five- and ten-year yields as well as the yields near the zero-lower bound. The appendix also contains the impact of the zero-lower bound on the one-factor and two-factor models in respectively Figure (A.6) and (A.7). In the period excluding the zero-lower bound, the one-factor models obtain a similar fit but looking at the period including the zero-lower bound we observe that the Cox et al. (1985) model has more problems with capturing the yields near the zero-lower bound. With respect to the two-factor models, they obtain a similar fit in the period excluding the zero-lower bound, whereas in the including the zero-lower bound period I observe that the Cox et al. (1985) model has substantial problems to capture the curvature of the five- and ten-year yields.

6 Conclusion and Discussion

In this thesis, I evaluated the term structure of yields while subject to the zero-lower bound by performing the Kalman filter approach of De Jong (2000). The main contribution of this thesis lies in evaluating the impact of the zero-lower bound on the estimation performance of the Vasicek (1977) and Cox et al. (1985) model specifications. In particular, I contributed to prior research by examining the performance of the three-factor specification for both models by simulating from a controlled zero-lower bound environment in both an ordinary and a cross-simulation setup. Furthermore, I expanded on prior research by performing an extensive empirical study on the performance of the one-, two- and three-factor models of both specifications where I zoomed in on the effect of the zero-lower bound.

In the ordinary simulation study, I find that the three-factor Cox et al. (1985) model appears to be more accurate and less variable for most of the parameters than the Vasicek (1977) model which has problems estimating its mean-reversion and variance parameters. Nevertheless, the bias on the average short rate is somewhat smaller for the Vasicek (1977) model. The parameter standard deviation in the Vasicek (1977) case appears to be somewhat larger under the zero-lower bound period, which is reversed in the Cox et al. (1985) case. In the cross-simulation, I find that the Cox et al. (1985) model appears to be more accurate and less variable for all parameters than the Vasicek (1977) model on CIR data whereas the estimation estimation results are close when estimating on Vasicek data.

In the empirical study, I find that the three-factor models are better specified than the one- and two-factor models in both the period excluding and including the zero-lower bound. The third factor enhances the fit of the average term structure considerably for the excluding zero-lower bound period where the steepness in the middle appears to be only completely seized by the Vasicek (1977) model which has a relatively tiny RMSE that is considerably larger in the other period. The serial correlation is close to zero for both models in the period excluding, while it is considerably higher in the period including the zero-lower bound. Consequently, the Vasicek (1977) three-factor model appears to be the better specified in both periods, although substantially less in the including zero-lower bound period.

Moreover, I find that the first factor describes the level of the yield curve for each model in both periods. The second factor enjoys substantial support as the yield curve's slope for all models in both periods except for the two-factor Cox et al. (1985) model in the including zero-lower bound period. In the Vasicek (1977) case, the third factor has the interpretation as the (reversed) curvature factor. In the Cox et al. (1985) case, the third factor's interpretation as the curvature factor appears relatively little given the tiny correlation between the third factor and its proxy. For both models, the correlation is lower in the period including the zero-lower bound for the first and third factor.

Imperatively, I contemplated the limitations of this thesis which lead to a suggestion for further research. Since I used data on U.S. Treasury yields where the minimum yield is 0.01% for the three-month maturity bond, the data set does not contain negative yields while the yields in several European countries are negative even for longer maturities at the time of writing. Hence, it might be interesting to assess the performance of the Vasicek (1977) and Cox et al. (1985) models in a (purely) negative interest rate environment. However, this additional research expands beyond the scope of this thesis.

References

Adrian, T., Crump, R. K., and Moench, E. (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics*, 110(1):110–138.

Ait-Sahalia, Y. and Kimmel, R. L. (2010). Estimating affine multifactor term structure models using closed-form likelihood expansions. *Journal of Financial Economics*, 98(1):113–144.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics*, 50(4):745–787.

Bauer, M. D. and Rudebusch, G. D. (2016). Monetary policy expectations at the zero lower bound. *Journal of Money, Credit and Banking*, 48(7):1439–1465.

Black, F. (1995). Interest rates as options. *the Journal of Finance*, 50(5):1371–1376.

Chen, L. (1993). Stochastic mean and stochastic volatility: a three-factor model of the term structure and its application in pricing of interest rate derivatives,". *Federal Reserve Board*.

Chen, R.-R. and Scott, L. (1992). Pricing interest rate options in a two-factor cox–ingersoll–ross model of the term structure. *The review of financial studies*, 5(4):613–636.

Christensen, J. H., Diebold, F. X., and Rudebusch, G. D. (2011). The affine arbitrage-free class of nelson–siegel term structure models. *Journal of Econometrics*, 164(1):4–20.

Christensen, J. H. and Rudebusch, G. D. (2016). Modeling yields at the zero lower bound: Are shadow rates the solution? In *Dynamic Factor Models*, pages 75–125. Emerald Group Publishing Limited.

Collin-Dufresne, P. and Goldstein, R. S. (2002). Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility. *The Journal of Finance*, 57(4):1685–1730.

Cox, J. C., Ingersoll Jr, J. E., and Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica: Journal of the Econometric Society*, pages 363–384.

Creal, D. D. and Wu, J. C. (2015). Estimation of affine term structure models with spanned or unspanned stochastic volatility. *Journal of Econometrics*, 185(1):60–81.

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

De Jong, F. (2000). Time series and cross-section information in affine term-structure models. *Journal of Business & Economic Statistics*, 18(3):300–314.

de Los Rios, A. D. (2015). A new linear estimator for gaussian dynamic term structure models. *Journal of Business & Economic Statistics*, 33(2):282–295.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.

Diebold, F. X. and Rudebusch, G. D. (2013). *Yield curve modeling and forecasting: the dynamic Nelson-Siegel approach*. Princeton University Press.

Duan, J.-C. and Simonato, J.-G. (1999). Estimating and testing exponential-affine term structure models by kalman filter. *Review of Quantitative Finance and Accounting*, 13(2):111–135.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, 57(1):405–443.

Duffee, G. R. (2013). Bond pricing and the macroeconomy. In *Handbook of the Economics of Finance*, volume 2, pages 907–967. Elsevier.

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

El Karoui, N. and Rochet, J.-C. (1989). *A pricing formula for options on coupon-bonds*. Asociación Sudeuropea de Economía Teórica.

Frachot, A., Lesne, J., and Renault, E. (1995). Indirect inference estimation of factor models of the yield curve. Technical report, working paper, CREST.

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

Gallant, A. R. and Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 12(4):657–681.

Geyer, A. L. and Pichler, S. (1999). A state-space approach to estimate and test multifactor cox-ingersoll-ross models of the term structure. *Journal of Financial Research*, 22(1):107–130.

Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of applied econometrics*, 8(S1):S85–S118.

Gurkaynak, R. S. and Wright, J. H. (2012). Macroeconomics and the term structure. *Journal of Economic Literature*, 50(2):331–67.

Hamilton, J. D. (1995). Time series analysis. *Economic Theory. II, Princeton University Press, USA*, pages 625–630.

Hamilton, J. D. and Wu, J. C. (2012a). The effectiveness of alternative monetary policy tools in a zero lower bound environment. *Journal of Money, Credit and Banking*, 44:3–46.

Hamilton, J. D. and Wu, J. C. (2012b). Identification and estimation of gaussian affine term structure models. *Journal of Econometrics*, 168(2):315–331.

Heston, S. L. (1991). *Testing continuous time models of the term structure of interest rates*. School of Organization and Management, Yale University.

Hull, J. and White, A. (1993). One-factor interest-rate models and the valuation of interest-rate derivative securities. *Journal of financial and quantitative analysis*, 28(2):235–254.

Jamshidian, F. (1989). An exact bond option formula. *The journal of Finance*, 44(1):205–209.

Jamshidian, F. (1991). Bond option evaluation in the gaussian interest rate model. *Research in Finance*, 9:131–170.

Jamshidian, F. (1992). A simple class of square-root interest-rate models. *Applied Mathematical Finance*, 2(1):61–72.

Jegadeesh, N. and Pennacchi, G. G. (1996). The behavior of interest rates implied by the term structure of eurodollar futures. *Journal of Money, Credit and Banking*, 28(3):426–446.

Joslin, S., Singleton, K. J., and Zhu, H. (2011). A new perspective on gaussian dynamic term structure models. *The Review of Financial Studies*, 24(3):926–970.

Kim, D. H. and Orphanides, A. (2005). Term structure estimation with survey data on interest rate forecasts. *Finance and Economics Discussion Series*, 2005:48.

Krippner, L. (2013a). Measuring the stance of monetary policy in zero lower bound environments. *Economics Letters*, 118(1):135–138.

Krippner, L. (2013b). A tractable framework for zero-lower-bound gaussian term structure models.

Lamoureux, C. G. and Witte, H. D. (2002). Empirical analysis of the yield curve: The information in the data viewed through the window of cox, ingersoll, and ross. *The Journal of Finance*, 57(3):1479–1520.

Langetieg, T. C. (1980). A multivariate model of the term structure. *The Journal of Finance*, 35(1):71–97.

Longstaff, F. A. and Schwartz, E. S. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *The Journal of Finance*, 47(4):1259–1282.

Lund, J. (1997). *Econometric analysis of continuous-time arbitrage-free models of the term structure of interest rates*. Department of Finance, Faculty of Business Administration, Aarhus School of ....

McCulloch, J. H. and Kwon, H.-C. (1993). *US term structure data, 1947-1991*. Department of Economics, Ohio State University.

Pang, K. and Hodges, S. (1996). Non-negative affine yield models of the term structure. In *Financial Options Research Center, Warwick Business School, University of Warwick Working Paper*.

Pearson, N. D. and Sun, T.-S. (1994). Exploiting the conditional density in estimating the term structure: An application to the cox, ingersoll, and ross model. *The Journal of Finance*, 49(4):1279–1304.

Pennacchi, G. G. (1991). Identifying the dynamics of real interest rates and inflation: Evidence using survey data. *The review of financial studies*, 4(1):53–86.

Piazzesi, M. (2010). Affine term structure models. In *Handbook of financial econometrics: Tools and Techniques*, pages 691–766. Elsevier.

Siegel, A. F. and Nelson, C. R. (1988). Long-term behavior of yield curves. *Journal of financial and quantitative analysis*, 23(1):105–110.

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

Wu, J. C. and Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291.

Appendix

Figure A.1: U.S. Treasury yields with a constant maturity

This figure shows the U.S. Treasury yields with a constant maturity for the period from January 1982 until October 2018. In the lower right corner of this figure, we observe the exceptional zero-lower bound (ZLB) period which spans approximately form August 2008 until March 2017. The lines represent the yield to maturity for a maturity of 3, 12, 60, and 120 months.

Figure A.2: Regression of observed yields on the fitted factor(s)

These graphs present the coefficients of a regression for the three-factor CIR model in first differences of the actual interest rates on the fitted factor(s) and a constant. The result for the excluding zero-lower bound period is on the left, whereas for the period including the zero-lower bound on the right.

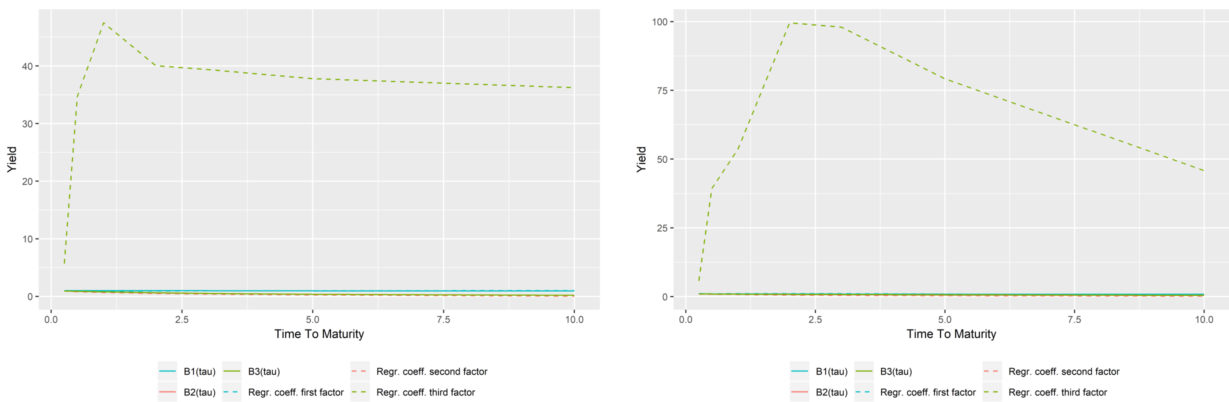


Table A.1: Estimation results factor models

This table presents the n-factor Vasicek and the n-factor CIR model QML estimation results, for $n = 1, 2, 3$. The table further provides the mean-reversion variable κ^* along with the half-life of the factors, $[\ln(2)/\kappa^*]$, governed by measure $\mathbb{Q}$. These results are based on the period including the zero-lower bound.

Model	Model	$A_0(\times 100)$	κ_i	$\tilde{\alpha}_i(\times 10^4)$	$\beta_i(\times 100)$	$\psi_i(\times 10^{-2})$	κ^*	2 ln L
One-factor	Vasicek	6.0460	0.1556	0.0381		0.3142	0.1556	9739.04
		(0.9459)	(0.0037)	(45.6791)		(0.0238)	[4.46]	
	CIR	2.4966	0.1858		0.2442	-0.4954	0.0648	9812.66
		(1.0633)	(0.0096)		(0.5493)	(0.0076)	[10.70]	
Two-factor	Vasicek	4.1868	0.0002	0.2531		-0.02380	0.0002	10118.84
		(1.5379)	(0.0873)	(115.5277)		(0.0057)	[3465.74]	
			0.2169	0.6524		-0.0953	0.2169	
			(0.0045)	(74.2825)		(0.0267)	[3.20]	
	CIR	2.5968	0.0812		0.2769	-0.0115	0.0780	9867.64
		(1.9926)	(0.4318)		(150.3961)	(0.0045)	[8.89]	
			0.1916		0.2097	-0.1015	0.1703	
			(0.0504)		(96.6670)	(0.0016)	[4.07]	
Three-factor	Vasicek	0.0011	0.0002	0.3520		0.0379	0.0002	10415.59
		(3.4894)	(0.0454)	(124.9400)		(0.0079)	[3465.74]	
			0.2140	0.9750		-0.2167	0.2140	
			(0.0048)	(61.9792)		(0.0105)	[3.24]	
			2.7848	1.1581		-0.2223	2.7848	
			(0.0053)	(46.2474)		(0.8235)	[0.25]	
	CIR	1.9275	0.0032		0.2934	0.0627	0.0216	9769.90
		(1.2903)	(0.1245)		(4.4439)	(0.0025)	[32.14]	
			0.2698		0.4525	-0.1468	0.2034	
			(0.0138)		(1.3897)	(0.0026)	[3.41]	
			0.1595		0.0002	-0.2419	0.1595	
			(0.0628)		(41.5707)	(0.0338)	[4.35]	

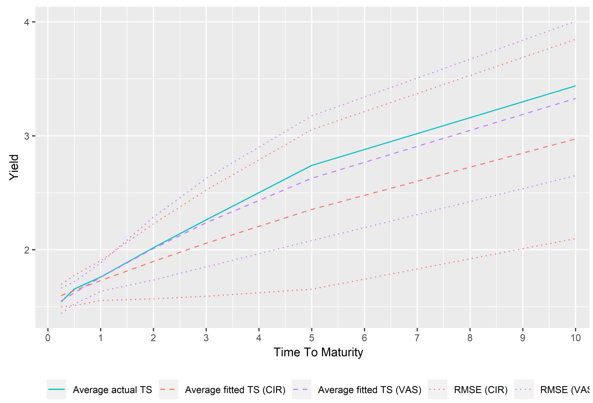
Table A.2: Residuals of the factor models

This table presents the residuals' summary statistics based on the n-factor Vasicek and the n-factor CIR model where ρ_k represent the serial correlation for order k. The residuals are scaled to percentage points. These residuals are based on the period including the zero-lower bound.

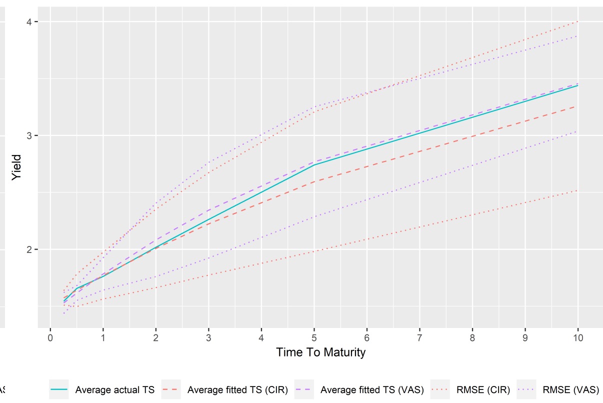
		Vasicek				CIR			
Model	**Statistic**	0.25	1	5	10	0.25	1	5	10
	Mean	-0.0061	0.0038	0.1133	0.1120	0.0229	0.0462	0.1850	0.1742
	Stand. Dev.	0.1984	0.2358	0.5519	0.6655	0.4352	0.4997	0.6965	0.7527
	ρ_1	0.6422	0.7681	0.9124	0.9434	0.1758	0.3631	0.7516	0.8628
	ρ_{12}	0.2823	0.1995	0.3364	0.4810	0.0206	0.0552	0.2699	0.4211
One-factor									
	Corr. matrix	1.0000	0.5808	-0.2018	-0.3879	1.0000	0.9248	0.5077	0.3094
		0.5808	1.0000	0.2991	-0.0033	0.9248	1.0000	0.6539	0.4239
		-0.2018	0.2991	1.0000	0.9138	0.5077	0.6539	1.0000	0.9272
		-0.3879	-0.0033	0.9138	1.0000	0.3094	0.4239	0.9272	1.0000
	Mean	0.0044	-0.0348	-0.0455	-0.0378	-0.0058	0.0128	0.1570	0.1834
	Stand. Dev.	0.1810	0.2520	0.5187	0.4573	0.4363	0.4880	0.6949	0.7674
	ρ_1	0.4904	0.7752	0.8898	0.8531	0.1335	0.2840	0.7146	0.8396
	ρ_{12}	0.1424	0.2340	0.3102	0.2680	0.0218	0.0351	0.2472	0.3867
Two-factor									
	Corr. matrix	1.0000	0.5365	0.2074	0.1479	1.0000	0.9223	0.4211	0.2032
		0.5365	1.0000	0.6583	0.5383	0.9223	1.0000	0.5857	0.3307
		0.2074	0.6583	1.0000	0.9734	0.4211	0.5857	1.0000	0.9262
		0.1479	0.5383	0.9734	1.0000	0.2032	0.3307	0.9262	1.0000
	Mean	-0.0116	-0.0216	-0.0514	-0.0348	-0.0156	-0.0378	-0.0108	0.0036
	Stand. Dev.	0.4322	0.4453	0.5769	0.5289	0.4406	0.4818	0.6660	0.7404
	ρ_1	0.1205	0.0641	0.4224	0.3813	0.1883	0.3012	0.7070	0.8346
	ρ_{12}	0.0186	-0.0048	0.1654	0.1430	0.0377	0.0567	0.2469	0.3647
Three-factor									
	Corr. matrix	1.0000	0.9722	0.6484	0.6449	1.0000	0.9266	0.4406	0.2071
		0.9722	1.0000	0.7409	0.7349	0.9266	1.0000	0.5763	0.2940
		0.6484	0.7409	1.0000	0.9868	0.4406	0.5763	1.0000	0.9153
		0.6449	0.7349	0.9868	1.0000	0.2071	0.2940	0.9153	1.0000

Figure A.3: Fit of the factor models

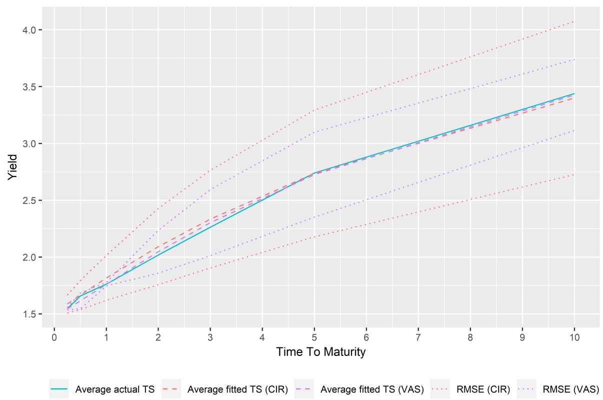
These graphs present, regarding both the n-factor models of Vasicek and CIR, the average fitted as well as the average actual term structure (TS), for $n = 1, 2, 3$. In addition, the root mean squared error (RMSE) is displayed for each model. The graphs are based on the period including the zero-lower bound.



One-factor models



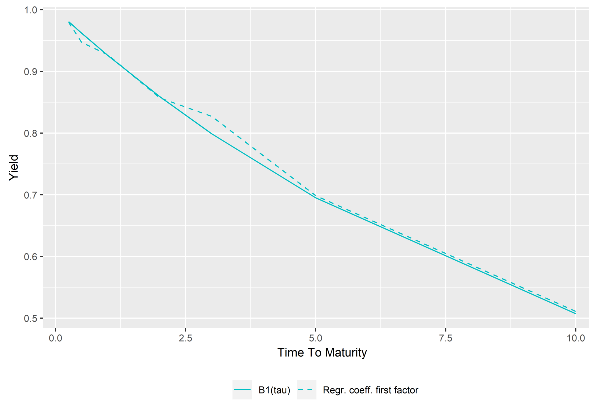
Two-factor models



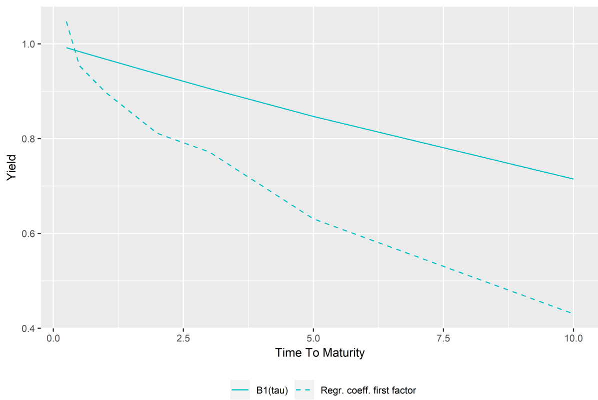
Three-factor models

Figure A.4: Regression of observed yields on the fitted factor(s)

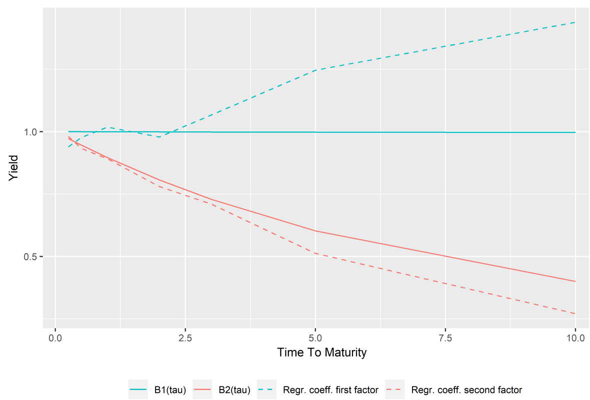
These graphs present the coefficients of a regression for the n-factor Vasicek and the n-factor CIR model in first differences of the actual interest rates on the fitted factor(s) and a constant, for $n = 1, 2, 3$. The graphs are based on the period including the zero-lower bound.



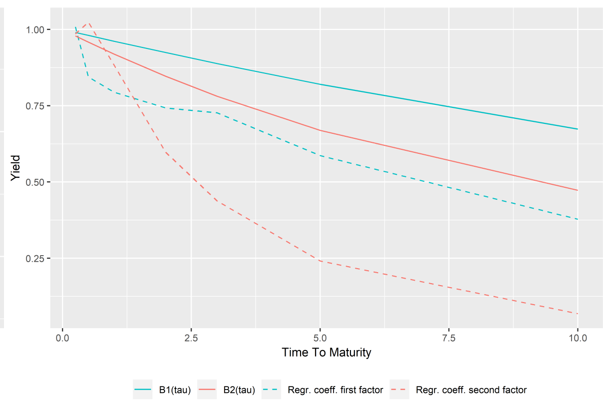
One-factor Vasicek model



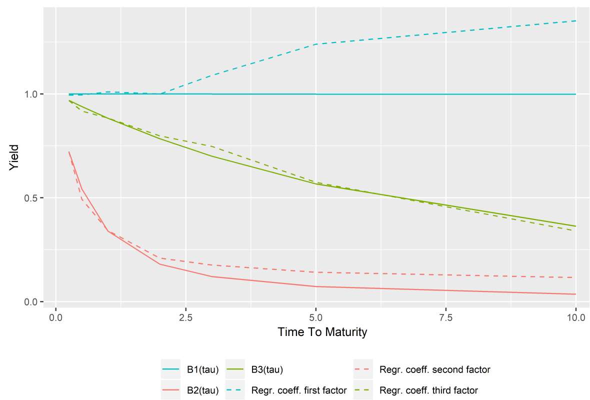
One-factor CIR model



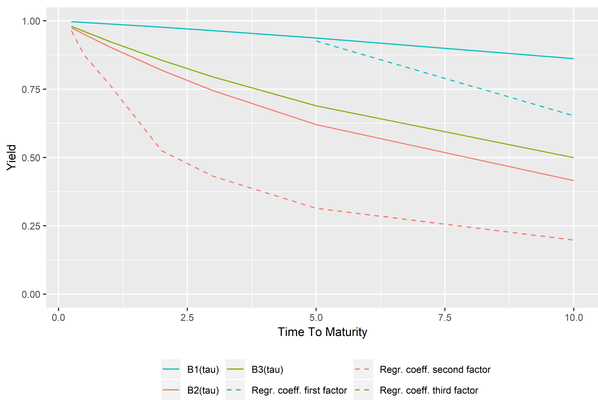
Two-factor Vasicek model



Two-factor CIR model



Three-factor Vasicek model



Three-factor CIR model

Figure A.5: Fit of the factors

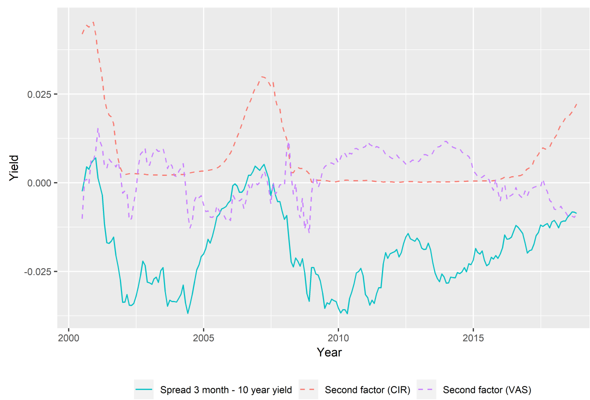
These graphs present the fit of the n-factor Vasicek and the n-factor CIR model, for $n = 1, 2, 3$. The graphs are based on the period including the zero-lower bound.

First factor one-factor models

First factor two-factor models

First factor three-factor models

Second factor two-factor models



Second factor three-factor models

Third factor three-factor models

Figure A.6: Fit of the one-factor model

These graphs present the fitted yields and the actual yields for (3,12,60,120) months with the fit of the one-factor models on the left for the period including the zero-lower bound and the fit of the one-factor models on the right for the period including the zero-lower bound.

Figure A.7: Fit of the two-factor model

These graphs present the fitted yields and the actual yields for (3,12,60,120) months with the fit of the two-factor models on the left for the period excluding the zero-lower bound and the fit of the two-factor models on the right for the period including the zero-lower bound.

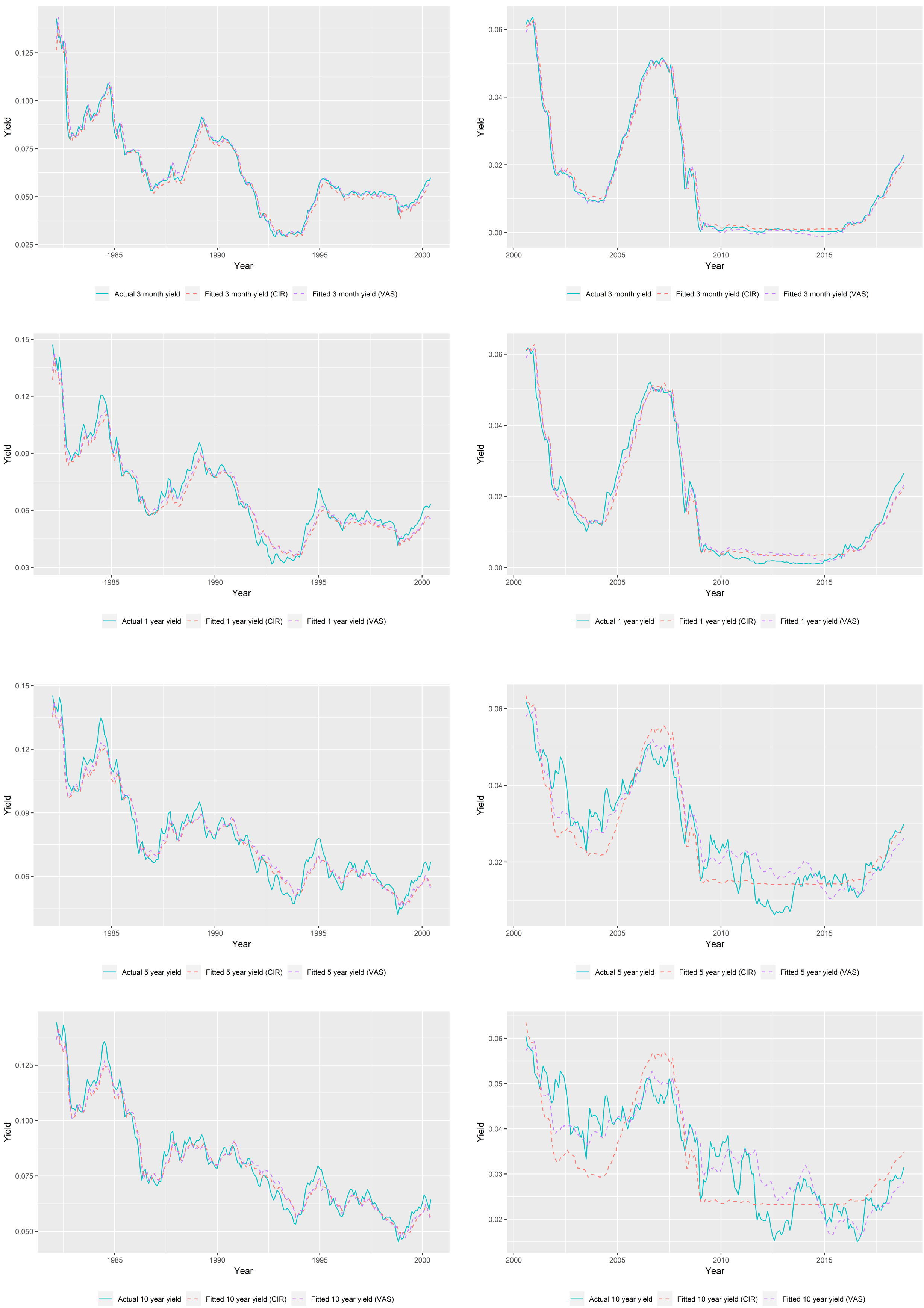


Figure A.8: Fit of the three-factor model

These graphs present the fitted yields and the actual yields for (3,12,60,120) months with the fit of the three-factor models in the period excluding the zero-lower bound on the left and the fit of the three-factor models in the period including the zero-lower bound on the right.