Erasmus University Rotterdam

Master Thesis

Calibrating Bayesian risk scenarios of interest rates with the help of Macro-Economic Factors

Kasper Wijshoff
376235

supervised by
M. van der Wel
R. Damen (Zanders)
second reader:
J.A. Oorschot

January 31, 2019

Abstract

Over recent years, patterns in yield data have proven to raise difficulties in modelling them. Levels of yields are currently at a lower bound and are very persistent. However, many developments have taken place into the modelling techniques of these yields in the past few years. This paper investigates if a Bayesian estimation of a term structure model can be improved by adding Macro-Economic Factors to the model. Both a simulation and empirical study were used to access the performance of the model. The simulation study was not able to generate stable results for the model with economic variables. For the empirical study, a dataset on Euribor yield data from 2008 to 2017 were used. It shows that the extension of the parameter space leads to stable results when some tuning parameters are used. Nonetheless, the economic variables add an improvement to the forecasting of yields and are able to predict changing patterns in the yield curve when extreme scenarios are considered at a one year horizon.

Contents

1 Introduction

Measurement and management of interest rate risk in the banking book is one the main objectives for risk departments of banks. This is mostly due to the fact that all interest-bearing objects on balance-sheets are affected by this risk. The focus of this paper lies mainly on the risk that occurs when the entire yield curve for different tenors changes, the so-called Gap Risk. This risk type is parted in two different sub-types, namely parallel Gap Risk and non-parallel Gap Risk. With parallel Gap Risk the entire yield curve is affected by the same change and therefore just one quantity of shock needs to be determined that will be applied to the entire yield curve. On the other hand, with non-parallel Gap Risk also the slope and shape of the yield curve change. Subsequently, for each level of maturity a different shock is applied to this yield curve. Essential to both approaches, regardless of their differences, is that an underlying model is used to correctly predict those shocks. Given old sets of data on interest rates, existing models have been able to adequately predict shocks. However, with the most recent data these kind of models are not able to do this anymore, because of a persistent and downwards trend in the data. So in essence, the question remains as to how to quantify shocks on the yield curve realistically, given the recent data on interest rates.

A first step in answering the question of how to measure Gap Risk is to model interest rates to be able to obtain shocks in the yield curve. Arguably if a model is developed that is able to correctly describe the structure of interest rates, then it could also be used to derive reasonable shocks in the yield curve. On the other hand, it needs to be examined if a model that perhaps could deliver the best forecasts of the interest rate also could model the tails of the distribution correctly. These tails ultimately determine the shocks that will be applied in a scenario and not in the forecast of the interest rate for the next period. Therefore, in this thesis I do not only focus on the out-of-sample forecasting performance of the models, but also on the model as a whole. Can it be concluded from the distribution that follows from the estimated model that this model can be used to adequately derive shocks for the application of calibrating Gap Risk?

The contribution of this thesis to the already existing literature is explained in two ways. First of all, Macro-Economic Factors are incorporated into the model of Bauer (2017), which uses a Bayesian approach in estimating a term structure model. The usage of these factors is already proven by numerous papers, but, to my knowledge, has not yet been used alongside a Bayesian approach in estimating a term structure model. Second, this research uses the European market for the examination of the Gap Risk. This could possibly lead to new insights in the literature, because most papers on this topic only investigate U.S. data. Next to that, a recent sample in this research will be considered namely starting from 2009 till 2018 which will give new insights to the existing literature for the proposed methods.

Ang and Piazzesi (2003) were the first to include Macro-Economic Factors such as inflation and economic growth rates in a term structure model. Others such as Koopman and van der Wel (2013) and De Pooter et al. (2010) follow this approach in their models, but used more variables. To reduce the dimensionality of their panel they used Principal Component Analysis. Bauer and Rudebusch (2016b) stresses the importance of adding Macro-Economic Factors to term structure by resolving criticism that arose in the literature over the previous years. Literature on the term structure of interest rates has empirically proven to deliver inconsistent results over time (De Pooter et al., 2010). When the structure in interest rates changes, the performance of term structure models also changes. Also, where most models proposed in the literature tend to focus on one aspect, such as using Macro-Economic Factors or imposing a lower bound for decreasing interest rates, there are not many models that try to combine aspects from each other. Mostly these combinations are understandably left out for future research. This paper, however, will further expand the model posed by Bauer (2017) by also including Macro-Economic Factors following a similar approach as to Ang and Piazzesi (2003).

As mentioned before, the model of Bauer (2017) uses a Bayesian framework to estimate a term structure model. He chooses for this technique to address the problems of model uncertainty and a large parameter space in term structure models. Therefore, in this thesis I also follow the Bayesian methodology because the same problems could arise with a different dataset. Moreover, by also adding the Macro-Economic Factors the problem of parameter space will enlarge. Within this framework a decision needs to be made about the priors for the different parameters of the model. The priors used in Bauer (2017) tend to lead to optimal outcomes, so in this setting the same priors will be used. However, the parameters for the Macro-Economic Factors have a more economic interpretation and a different set of priors could be used. Wright (2013) shows that when estimating autoregressive models of macro economic variables with a Bayesian approach, the use of democratic priors improves the forecast performance. These priors are constructed by using a Normal distribution with as mean the average of long-term Blue Chip forecasts. In this paper, due to the scoping of the research for the Macro-Economic Factors the same priors as for the risk factors are used.

As previously mentioned, a major problem with existing models in estimating yield data comes from recent structures in the data. A Bayesian approach could lead to a model that is able to cope with European data over the past few years by considering model uncertainty. Kim and Orphanides (2005) show that a sufficient long data sample is needed to obtain multiple mean-reversions. If European data is used, the sample will presumably be too short to avoid this problem. In addition, after the housing crisis of 2008 the term structure has drastically changed, so data from before that year is arguably not that useful anymore. If, just like other researchers, data from before 2008 is considered, the practical application of the proposed method is less relevant. For the proposed method to be practical it needs to be able to cope with recent historical data of interest rates.

There is abundant literature available on the topic of the proposed method in this paper that discusses various alternative approaches of modelling term structure models. One of the first that made significant breakthroughs on this topic are Vasicek (1977) and Cox et al. (1985), who propose models for estimating interest rates that are still applicable, but tend to be outperformed by newer methods.Nelson and Siegel (1987) also propose a method for modeling the yield curve, but its application was developed by Diebold and Li (2006). This adaptation is used as the well-known Nelson-Siegel model and uses the level, slope and curvature of interest rates as explaining factors to describe the yield curve. A paper that is worth mentioning as an alternative in this paper is that of Laurini and Hotta (2010), because they use a Bayesian approach to estimate a Nelson-Siegel curve.

Finally, Joslin et al. (2011) proves that the estimation of the affine term structure models could be done in a more simpler way. They argue that some risk factors could be numerically solved and that this leads to a lower dimension of the parameters that are to be estimated. Moreover, they showed that the way of solving as done by previous papers (Duffie and Kan, 1996; Dai and Singleton, 2000), only leads to local optima and not to estimated coefficients that are globally optimal. In recent years, the model derived from Joslin et al. (2011) remains one that is widely used in the literature as a starting point for affine term structure models. An illustration of this is that Bauer (2017) also estimates this model but with a different approach.

In recent years, further adaptations of the general affine term structure model were developed that focus on the structure of interest rates from most recent years. In these years the short rate in Europe and America has remained at a constant small negative level. This has led to poor performance of modeling this structure in the previously described conventional affine term structure models (Bauer and Rudebusch, 2016a). Because of this, an adaptation of these models has been derived by some that introduce a Zero Lower Bound (ZLB). This ZLB is used as a minimum for forecasts of the short rate and lets these forecasts stay for a period at this level. Examples of this approach include Hamilton and Wu (2012); Kim and Singleton (2012); Monfort et al. (2017).

Others have taken a different road in improving the forecasting performance, not by developing a new method that is more sophisticated, but by using a set of models that all have their own advantages. Bates and Granger (1969) already show that this is indeed the case when making forecasts for airline passenger data. This finding is further proved by Diebold and Pauly (1987) and Aiolfi et al. (2010) for making combinations of more general economic applications. Moreover, De Pooter et al. (2010) also show that combinations can significantly improve the performance for the forecasting of the term structure of interest rates.

Bauer (2017) is not the first to use a Bayesian way of estimating a term structure model, others are for example: Jones (1998, 2003); Elerian et al. (2001); Eraker (2001); Sanford and Martin (2005). However, these researchers try to estimate more general term structure models, such as one without the formulation of Joslin et al. (2011). This leads to more complex models in which the estimates are harder to compute analytically, which is due to the fact that when using a Bayesian approach, a Markov Chain Monte Carlo (MCMC) algorithm is needed and the dimension of the parameters is not reduced as in Bauer (2017). He also proposes a variant of the MCMC algorithm by dividing the parameters in different blocks so that some blocks can be computed in a more efficient way.

In this paper I show that the proposed Bayesian approach of estimating a term structure model of European yields is suitable for the purpose of generating extreme scenarios of the yield curve. For the same specification as in Joslin et al. (2011) the Bayesian estimation results are reliable and fit the dynamics of the yield curve rather well. Because this setup only uses historic data of the yield curve itself, it is only possible to further predict a continuation of the current patterns in the data. Therefore, I also investigate the extension of adding Macro-Economic Factors to the term structure model. This enlargement of the parameter space does cause the estimation results to be less stable, because tuning parameters are needed to tailor the Bayesian drawings of the parameters. Still, the additional factors help in explaining the shape of the yield curve, which becomes visible when extreme scenarios of this curve are generated. Concluding, the Macro-Economic Factors do not improve the reliability of the estimates but do help in anticipating a change in the shape of the yield curve, which is not possible when only historic data of the yields are used.

2 Data

In this section, information is given regarding the choices that I make for using data in the research. I use two different parts of data in this research, specifically one set of data on interest rate yields and another set consisting data on Macro-Economic variables.

Additionally, descriptive statistics on both sets are given and for the yield data stylized facts are also discussed.

2.1 Yield Data

As the research in this paper is mostly relevant for European banks, the data that I use is from the European markets. The reference yield curve that is most suitable for the measurement of Gap Risk is the 6 month Euribor rate. Therefore, Euribor based yields from the fixed-for-variable rate swap market will be considered in this paper. The data on these yields are available on a Bloomberg terminal[1]. Data are available for all different types of maturities, but due to the high consistency of yields with similar maturities a set has been chosen to overcome this problem[2]. Therefore, maturities have been chosen that are not that similar to each other. The resulting set of maturities represents a sufficient set of maturities that both exhibit short term and long term maturities. However, I am aware that this set has not been widely used by other researchers. Most of them used a set of data for the U.S., where, in comparison with European data, the problem is less present. The following maturities are used in this paper: 6 months, 1 year, 3 years, 5 years, 7 years, 9 years and 10 years. These yields are plotted in figure 1, where a clear trend across maturities can be noted. The yields are steeply increasing in the beginning of the year 2008 and then all yields decreased sharply in the years thereafter. Eventually, the different yields reached a lower bound which stabilized them, and with the short rates with a maturity of 6 months or 1 year even became negative.

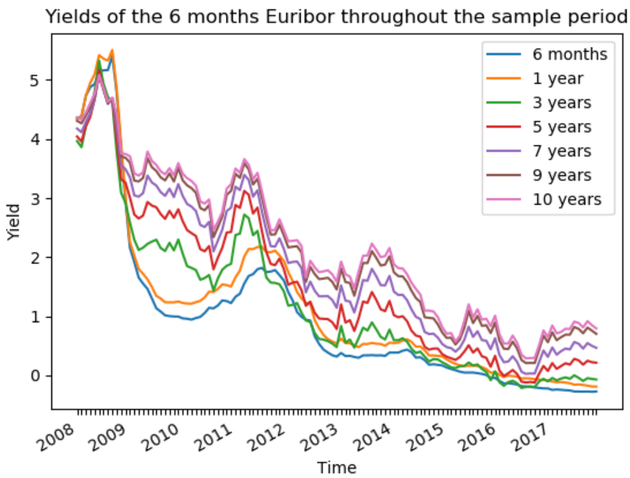

Figure 1: Overview of monthly yields throughout the period that is examined in this paper. The period starts at 01-01-2008 and ends at 01-12-2017. The amount of years in the legend represent the different maturities of the different time series.

[1] The names for these time series are the following: "EUR006M INDEX", "EUR012M INDEX" and "EUSAx Curncy" where x stands for the amount of years till maturity.

[2] To further illustrate this problem I ran standard regressions for each pair of two maturities and this resulted in R^2 of more than 0.9999 for yields with maturities that were adjacent in the set

In table 1 some summary statistics are given for this dataset. From this table, problems of the current historic yield data can already be noted, if the stylized facts of the yield curve are kept in mind. First of all, in figure 2 the average yield curve over time is given which does not satisfy the first stylized fact, implying that the average yield curve should be increasing and concave. Furthermore, the shape of the yield curve is very persistent and therefore it is not able to generate different forms of the curve. Next to that, the short end of the yield curve is not more volatile than the long end, which can also be noted from table 1. However, the other stylized facts are mostly satisfied. All different time-series are rather persistent with the smallest autocorrelation equal to 0.9302 for 3 lags. Because all autocorrelations are rather high, the difference between shorter tenors and larger tenors is not that high.

Table 1: This table contains the different central moments for data on 6 months Euribor with different maturities. Also, their first three autocorrelations are given in the three most right columns.

	Central Moments				Autocorrelations		
	Mean	Std. Dev.	Skewness	Kurtosis	AutoCorr. 1	AutoCorr. 2	AutoCorr. 3
6 Months	0.9880	1.4020	1.8377	2.8644	0.9928	0.9746	0.9487
1 Year	1.1813	1.4069	1.6731	2.4034	0.9926	0.9736	0.9464
3 Years	1.2272	1.3347	1.1645	0.8520	0.9904	0.9754	0.956
5 Years	1.5397	1.3371	0.7731	-0.2375	0.9901	0.978	0.963
7 Years	1.8286	1.3223	0.5491	-0.7162	0.9902	0.9793	0.9663
9 Years	2.0670	1.2951	0.4322	-0.9079	0.9895	0.9787	0.9664
10 Years	2.1658	1.2842	0.3904	-0.9595	0.9894	0.9782	0.9661

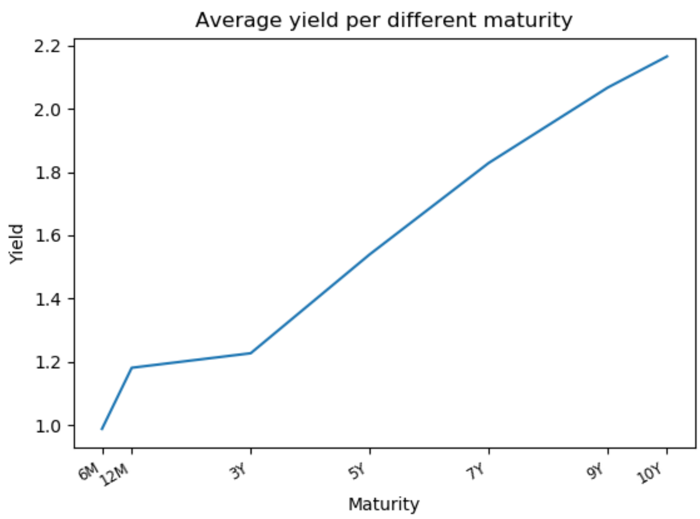

Figure 2: In this figure the average yield curve over time is plotted. It is clearly increasing over the tenors, but presumably not concave.

For the last stylized fact the cross correlations of the yields are given in table 2 and they show that these yields are highly correlated with each other. Therefore, the data supports the last stylized fact. The historic data are not able to satisfy all stylized facts on yield data, which further illustrates the kind of problems that could arise when this data is used for modelling a term structure.

Table 2: Table containing the cross correlations of the different tenors. Only for selected tenors are the correlations given.

	6 Months	1 Year	3 Years	5 Years	7 Years	9 Years	10 Years
6 Months	1.0000	0.9978	0.9531	0.9172	0.8905	0.8730	0.8665
1 Year	0.9978	1.0000	0.9612	0.9294	0.9052	0.8887	0.8825
3 Years	0.9531	0.9612	1.0000	0.9919	0.9791	0.9690	0.9649
5 Years	0.9172	0.9294	0.9919	1.0000	0.9968	0.9920	0.9896
7 Years	0.8905	0.9052	0.9791	0.9968	1.0000	0.9988	0.9978
9 Years	0.8730	0.8887	0.9690	0.9920	0.9988	1.0000	0.9998
10 Years	0.8665	0.8825	0.9649	0.9896	0.9978	0.9998	1.0000

To model the data, Principal Components are used in this paper. More specifically, the first 3 factors on the yield data are used. This factorization ensures that the dimension of the model is not too large, but is still able to use data on several maturities. The loadings of the Principal Components are rescaled in the same manner as in Joslin et al. (2014) in order to work with the same setup.

2.2 Macro-Economic Variables

Data on Macro-Economic Factors are also needed for this research, but first I have to determine which kind of variables to include. Most papers, on the topic of Macro-Economic variables, apply their research to U.S. interest rates and, unsurprisingly, use U.S. data on macroeconomic time series. Following the approach of Ang and Piazzesi (2003) data on two groups of variables will be used and these can be extracted from Eurostat. The data from Eurostat ensures that for the same region data is used as for the interest rates. The first group of variables represents European inflation and consists of the Consumer Price Index growth rate and the Producer Price Index growth rate. The second group is based on the real European economic growth and is based on the Unemployment Rate, Employment growth rate and the Industrial Production growth rate. All growth rates are calculated by taking the logarithm of the current index divided by the index of a year ago: $\log \frac{I_t}{I_{t-12}}$.

Some descriptive statistics are given in table 3. Similar to the data on yields the Macro-Economic variables are rather persistent. For example, the variable with the highest persistence is the unemployment rate with 0.9843. The two variables that represent inflation are quite similar to each other, but the Producer Price Index is more volatile and its higher moments are also more negative. The other variables are more distinct to each other and seem to therefore cover different aspects of how European economic growth behaves. Similar to the paper of Ang and Piazzesi (2003) the individual variables are normalized to have a mean of zero and unit variance. The two groups of variables are also linear decomposed to construct the Macro-Economic Factors. For both groups only one Principal Component is used so this results in two extra variables that will help to model the term structure of yields.

Table 3: This table contains the values of the first four central moments and autocorrelations of the macro-economic variables in the Euro-zone that are used in this research to obtain the factors. CPI stands for the growth rate of Consumer Price Index, PPI for the growth rate of Producer Price Index, UE for the unemployment rate in percentages of the working population, Employ for the growth rate of the percentage of employment, and IP for the growth rate Industrial Production. All growth rates are calculated by taking the logarithm of the yearly change of one variable.

	Central Moments				Autocorrelations		
	Mean	Std Dev	Skewness	Kurtosis	AutoCorr1	AutoCorr2	AutoCorr3
CPI	0.7846	0.4324	-0.2119	-0.0951	0.9747	0.9356	0.8872
PPI	0.7634	1.5040	-0.3250	-0.3294	0.9765	0.9285	0.8627
UE (%)	9.0245	1.1253	-0.2253	-0.4406	0.9833	0.9532	0.9164
EMPLOY	0.0300	0.3828	-0.8261	0.3011	0.9843	0.9642	0.9357
IP	0.2686	2.1880	-2.3584	6.7276	0.9590	0.9140	0.8461

3 Methodology

In the following section information is given on the models that are used in this paper. In addition, derivations and other formulas that are needed to generate the results are given. The first part of the section elaborates on the modeling of a Gaussian Dynamic Term Structure model. The second part elaborates how the Bayesian approach is applied to estimate the model in sections 3.1. Thereafter, it is explained how the adding of the Macro-Economic factors to the model will change the dynamics of the model and the estimation procedure. The last part of this section explains how the estimated model is used to generate the extreme scenarios of the yield curve.

3.1 Gaussian Dynamic Term Structure Model

Following the works of Joslin et al. (2011) and Bauer (2017) I estimate a Gaussian Dynamic Term Structure Model to the data described in section 2. The model of Joslin et al.

(2011) is called JSZ in the remainder of this paper. The canonical form of JSZ is used to start the Bayesian approach of the model, by taking the Maximum Likelihood (ML) estimates as starting values for the parameters. The generic representation of the model by Joslin et al. (2014) is defined by the following set of equations:

$$X_t = K_{0X}^{\mathbb{P}} + K_{1X}^{\mathbb{P}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{P}}, \tag{1a}$$

$$X_t = K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}} \tag{1b}$$

and

$$r_t = \rho_{0X} + \rho_{1X} X_t, \tag{2}$$

where X_t is the vector of latent risk factors, $\Sigma_X \Sigma_X'$ is the conditional covariance matrix of X_t, $\epsilon_t^{\mathbb{P}}, \epsilon_t^{\mathbb{Q}} \sim \mathcal{N}(0, I_{\mathcal{N}})$, $\mathbb{P}$ and $\mathbb{Q}$ stand for respectively the historical and the risk-neutral measure, K_{0X} is the level of the risk factor and K_{1X} is the autoregressive component of the risk factors under the two different measures. Furthermore, r_t stands for the one-period interest rate, and because monthly data is considered in this paper r_t represents the one-month interest rate. According to Joslin et al. (2011), $K_{0X}^{\mathbb{P}}, K_{1X}^{\mathbb{P}}$ can be simply estimated by the Ordinary Least Squares method which is used in obtaining the ML parameters for the model.

3.1.1 Stochastic discount factor

When a world of no-arbitrage is assumed for the interest rates, there must exist a probability measure, $\mathbb{Q}$, that prices all financial assets. This measure is related to the historical measure, $\mathbb{P}$, by the Stochastic Discount Factor (SDF). Following Bauer (2017) I use the same exponentially affine function for the SDF

$$-\log S_{t+1} = r_t + \frac{1}{2}\lambda_t'\lambda_t + \lambda_t'\epsilon_{t+1}, \tag{3}$$

where S_{t+1} is the change of measure at time $t+1$ and λ_t represents the market prices of risk at time t. The vector of λ_t is of dimension $(N \times 1)$ and is priced under the $\mathbb{P}$ measure. This pricing, first posed in Duffee (2002), results in the affine risk-price specification

$$\lambda_t = \Sigma_X^{-1}(\lambda_0 + \lambda_1 X_t), \tag{4}$$

where λ_0, also a $(N \times 1)$ vector, represents the difference in level of the two measures and λ_1, a $(N \times N)$ matrix, represents the difference in the autoregressive component of the measures.

Additional to the affine function of equation (3), any one-period pricing kernel is implied by

$$S_{t+1} = exp(-r_t)\frac{f^{\mathbb{Q}}(X_{t+1}|X_t)}{f^{\mathbb{P}}(X_{t+1}|X_t)}, \tag{5}$$

where $\frac{f^{\mathbb{Q}}(X_{t+1}|X_t)}{f^{\mathbb{P}}(X_{t+1}|X_t)}$ is a Radon-Nikodym derivative and translates to the following together with equation (3):

$$\frac{f^{\mathbb{Q}}(X_{t+1}|X_t)}{f^{\mathbb{P}}(X_{t+1}|X_t)} = \frac{d\mathbb{P}}{d\mathbb{Q}}(X_{t+1};\lambda_t) = exp(\frac{1}{2}\lambda_t'\lambda_t + \lambda_t'\epsilon_{t+1}). \tag{6}$$

Further pursuing the same derivation as in Bauer (2017), the innovations $\epsilon_t^{\mathbb{Q}}$ from equation (1b) are related to the pricing innovations by

$$\epsilon_t^{\mathbb{Q}} = \epsilon_t^{\mathbb{P}} + \lambda_{t-1}. \tag{7}$$

Likewise, the relations of $K_{0X}^{\mathbb{Q}}$ and $K_{1X}^{\mathbb{Q}}$ are also given by equations

$$K_{0X}^{\mathbb{Q}} = K_{0X}^{\mathbb{P}} - \lambda_0, \tag{8a}$$

$$K_{1X}^{\mathbb{Q}} = K_{1X}^{\mathbb{P}} - \lambda_1. \tag{8b}$$

Just as in Bauer (2017) measurement errors should be included to the observed bond yields because a N dimensional factor model is not able to perfectly price $J > N$ yields. Those are measured by

$$Y_t = \widehat{Y}_t + e_t, \tag{9}$$

where e_t represent the vector with measurements errors at time t and Y_t the vector of observed yields. σ_e^2 will denote the variance of this measurement error that will be used in the remainder of this research as a free parameter that needs to be estimated.

Given all previous definitions in this section, the set of parameters to be estimated is divided in five different groups: (λ, γ), $K_{0X}^{\mathbb{Q}}$, $K_{1X}^{\mathbb{Q}}$, Σ and σ_e^2. In the following sections it is explained in which way this set is estimated. Also in the remainder of this paper the full set of parameters to be estimated is be denoted by θ.

3.1.2 Bond pricing

When the model, given in equations (1a) till (2), is estimated for the remaining parameters the model-implied yields can be determined by

$$y_{t,m} = A_m(\theta_X^{\mathbb{Q}}) + B_m(\theta_X^{\mathbb{Q}})X_t, \tag{10}$$

where $\theta_X^{\mathbb{Q}} = (K_{0X}^{\mathbb{Q}}, K_{1X}^{\mathbb{Q}}, \Sigma_X, \rho_{0X}, \rho_{1X})$ is the set of parameters under the $\mathbb{Q}$ measure that needs to be estimated. Furthermore, $y_{t,m}$ represents the model-implied yield at time t for maturity m, and A_m and B_m are the implied loadings of the yields on the risk factors X_t which satisfy Ricatti difference equations, and are computed by the recursive equations

$$\begin{aligned} \mathbb{A}_{m+1} &= \mathbb{A}_m + (K_{0X}^{\mathbb{Q}})'\mathbb{B}_m + \frac{1}{2}\mathbb{B}_m'\Sigma\Sigma'\mathbb{B}_m - \rho_{0X}, \\ \mathbb{B}_{m+1} &= (K_{1X}^{\mathbb{Q}})'\mathbb{B}_m - \rho_{1X}. \end{aligned} \tag{11}$$

These equations have starting values of $\mathbb{A}_0 = 0$ and $\mathbb{B}_0 = 0$, and the relations $A_m = -m^{-1}\mathbb{A}_m$, $B_m = -m^{-1}\mathbb{B}_m$. As I have shown in equations 11, the loadings are fully determined by the $\mathbb{Q}$ parameters and will not be affected by changes in the $\mathbb{P}$ measure.

3.1.3 Likelihood function

Following Bauer (2017) I use the same likelihood function in this paper for the model without Macro-Economic Factors. Nonetheless, the full conditional likelihood function is defined as

$$\begin{aligned} L(Y_t|Y_{t-1};\theta,\gamma) =& L(Y_t|X_t; k_\infty^{\mathbb{Q}}, \Sigma, \sigma_e^2) \\ & \times L(X_t|X_{t-1}; k_\infty^{\mathbb{Q}}, K_{1X}^{\mathbb{Q}}, \lambda, \gamma, \Sigma). \end{aligned} \tag{12}$$

Equation (12) can be parted in two parts, namely the first term that represents the $\mathbb{Q}$ measure part of the equation and the second term that is $\mathbb{P}$ measure part. Both parts are then evaluated separately which allows for an easily implementation of the Macro-Economic Factors. Firstly, the part of the $\mathbb{Q}$ measure is given by

$$\begin{aligned} \log\left(L(Y_t|X_t; k_\infty^{\mathbb{Q}}, \Sigma, \sigma_e^2)\right) =& const - (J-N)\log(\sigma_e) \\ & - 0.5||W \perp e_t||/\sigma_e^2, \end{aligned} \tag{13}$$

where *const* stands for a constant part which will drop out of the equation when the likelihood is evaluated, $||v||$ stands for the Euclidean norm of vector v, and e_t is the residual from equation (10). The second part of the likelihood function is affected by solely the $\mathbb{P}$ measure, it is denoted by

$$\begin{aligned} \log\left(L(X_t|X_{t-1}; k_\infty^{\mathbb{Q}}, K_{1X}^{\mathbb{Q}}, \lambda, \gamma, \Sigma)\right) = \\ const - 0.5\log(|\Sigma\Sigma'|) - 0.5||\Sigma^{-1}(X_t - K_{0X}^{\mathbb{P}} - K_{1X}^{\mathbb{P}}X_{t-1})||. \end{aligned} \tag{14}$$

Together with equations (13) and (14) the joint log likelihood function can be established by taking the sum over the two different evaluated log likelihoods.

3.2 Bayesian Approach

In this paper, I use a Bayesian approach for estimating the term structure posed in section 3.1. In this section, the different steps that are needed for this approach are elaborated on. First, I will explain the methods used to generate drawings from the posterior distribution for the different groups of parameters except for λ. Second, the different model selection algorithms that generate drawings for only the λ parameters are discussed. Finally, the manner of choosing priors for the parameters is given.

3.2.1 Markov Chain Monte Carlo Algorithm

The model in equations (1a) until (2) can be estimated by solving the state space model by applying a Kalman Filter. However, Bauer (2017) showed, along with others, that this state space model also can be estimated by applying the MCMC algorithm to the posterior distribution

$$P(\theta|X,\gamma) \quad \propto \quad P(X|\theta,\gamma)P(\theta|\gamma), \tag{15}$$

where X represents the principal components of data on observed yields. The advantage of using a Bayesian approach versus a Kalman Filter is two-sided in this setting. By the usage of γ a lot of different models need to be compared with each other if the Kalman Filter is applied. Bayesian estimation can estimate the values of γ, which leads to a reduced amount of models to be compared. Additionally, the selection of which elements of γ to restrict results in model uncertainty. Techniques that involve Bayesian estimation cope with this problem by also investigating the Bayes factors of the different specifications. Disadvantages of Bayesian estimation are of the nature of complexity and computationality. For most researchers Bayesian techniques do not belong to standard methods of estimation regression models or even state space models. Also, by design the estimation of a Bayesian model takes longer by drawing a sufficient amount of parameters to form the distribution.

If all restrictions on the risk price parameters should be evaluated, then in total $2^{(N+N^2)}$ different models need to be estimated. However, Bauer (2017) uses his Bayesian approach to cope with this problem by letting γ be a vector of indicator variables where each entry corresponds with an element of $\lambda \equiv (\lambda_0', vec(\lambda_1)')'$. This definition gives the possibility to only validate the most plausible restrictions by Bayesian variable selection if γ is included in the set of parameters.

Draws for one block of parameters are made from this posterior distribution given the current values for other blocks of parameters. Following Bauer (2017), a block-wise Metropolis-Hastings algorithm will be used to iteratively sample draws for the following blocks: $K^{\mathbb{Q}}_{0X}$, $K^{\mathbb{Q}}_{1X}$ and Σ. The other groups (λ and σ_e^2) can be sampled by a Gibbs sampler because their full conditional posterior distribution is known. In the following sections I will elaborate more on the procedures to obtain draws for the different blocks of parameters.

3.2.1.1 Generating draws for λ

For generating draws of the parameters for λ a convenient step can be taken due to the formulation of the model. Namely, conditional on all other parameters the drawing of λ boils down to the estimation of a restricted vector autoregression model. Therefore, a Gibbs step is used to generate a drawing for λ which tends to be an efficient way of sampling parameters. If a natural conjugate prior for λ is chosen with $\mathcal{N}(\underline{\lambda}_\gamma, \underline{V}_\gamma)$ the posterior distribution can be found by using similar derivations present in most Bayesian literature. The kernel is a Normal distribution with mean and covariance

$$\begin{aligned}\overline{\lambda}_\gamma &= \overline{V}_\gamma(\underline{V}_\gamma^{-1}\underline{\gamma}_\gamma + S'(Z \otimes \Omega^{-1})z),\\ \overline{V}_\gamma &= (\underline{V}_\gamma^{-1} + S'(X_{full}X'_{full} \otimes \Omega^{-1})S)^{-1},\end{aligned} \tag{16}$$

where the objects with an underline represent prior values and object with an overline posterior values. The subscript γ represent those values for which the corresponding value of γ is equal to 1. Further, S represents a matrix of ones and zeros with dimension $N(N+1) \times a$, where a stands for the amount of unrestricted risk prices (equal to the sum of γ). Additional X_{full} is the full set of regressors so $X_{full,t} = (1, X_t')'$, and $z = (Z' \otimes I_N)S\lambda_\gamma + u$ stands for the

3.2.1.2 Generating draws for $k^{\mathbb{Q}}_\infty$ and $K^{\mathbb{Q}}_{1X}$

Due to their high correlation $k^{\mathbb{Q}}_\infty$ and $K^{\mathbb{Q}}_{1X}$ are drawn in the same block and therefore also in the same manner. Just as in Chib and Ergashev (2009) I use the same Independence Metropolis-Hastings sampler approach. For this approach I use the same proposal distribution, namely a t-distribution with 5 degrees of freedom. Next to that, parameters for this proposal density are needed, which are represented by the ML estimates. They will be used as the mean for the proposal density and the negative of the inverse of the Hessian matrix of the conditional posterior at the proposed values. The resulting acceptance probability is

$$\alpha(\chi^{j-1}, \chi*) = \min\{\frac{P(X|\chi^*, \theta_-, \gamma)P(\chi*, \theta_-)f(vech(\chi^{j-1}))}{P(X|\chi^{j-1}, \theta_-, \gamma)P(\chi^{j-1}, \theta_-)f(vech(\chi^*))}, 1\}, \tag{17}$$

where f stands for the proposal density which is the t-distribution and χ^* stands for the proposed values of $k_\infty^{\mathbb{Q}}$ and $K_{1X}^{\mathbb{Q}}$.

3.2.1.3 Generating draws for Σ

This block of parameters is efficiently drawn by only drawing values for the Choleski decomposition of the variance matrix Σ, this decomposition can then be vectorized and is denoted by $vech(\Sigma)$. In this way, the amount of values that has to be drawn is further reduced. The method of generating draws then for the vectorized triangular matrix is the same as for the $k_\infty^{\mathbb{Q}}$, $K_{1X}^{\mathbb{Q}}$ block of parameters. So again, the proposal distribution is a multivariate t-distribution with 5 degrees of freedom. Also its mean is equal to the ML estimation of the vectorized triangular matrix, and its covariance matrix equal to the negative inverse of the Hessian matrix of the conditional posterior. This should be only done for the first iteration however and it turns out that the resulting matrix was not positive definite. This problem was overcome by adding very small increments to this matrix in order to make it positive definite. After this correction was performed 2 times at the beginning of the algorithm the resulting covariance matrix was usable to generate drawings for the covariance matrix.

After a draw Σ^* has been made its acceptance probability is given as

$$\alpha(\Sigma^{j-1}, \Sigma^*) = \min\{\frac{P(X|\Sigma^*, \theta_-)P(\Sigma*, \theta_-)f(vech(\Sigma^{j-1}))}{P(X|\Sigma^{j-1}, \theta_-)P(\Sigma^{j-1}, \theta_-)f(vech(\Sigma^*))}, 1\}, \tag{18}$$

where θ_- represents the all parameters except for the draw of Σ, and f represents the proposed multivariate t-distribution.

3.2.1.4 Generating draws for σ_e^2

The drawings for the last block of parameters can also be generated by using a Gibbs step. It is possible because the measurement errors can be seen as regression residuals conditioned on all other parameters. Then again the closed-form posterior of σ_e^2 is known and therefore the Gibbs step can be used. Just as Bauer (2017) shows that with an Inverse-Gamma prior consisting of shape parameter 0 and scale parameter 0 (the prior of σ_e^2 is taken completely diffuse), the conditional posterior becomes also an Inverse-Gamma distribution with shape n and scale ssr. The shape n is equal to $T(J-N)$ and ssr is the sum of squared residuals of Y.

3.2.2 Model Selection Samplers

For the posterior draws for the parameter pair (γ^i, λ^i), different extensions of the MCMC algorithm can be used. In this paper 3 different alternatives to the basic MCMC algorithm are considered to investigate if another method can make robust improvements to the model. Those algorithms will be explained in this section briefly, also for each algorithm the model priors for γ are given. The 3 different sampling algorithms are in the same way used as Bauer (2017) used.

3.2.2.1 Gibbs Variable Selection

The first adaptation, Gibbs Variable Selection (GVS) treats all different models (for different $\gamma' s$) as nested models, therefore the product-space is only the space of the unrestricted model. This space can become large as the amount of parameters increases but it will be used intensively for all other nested models. It is first fully described in Dellaportas et al. (2002) and based on the work of Carlin and Chib (1995). They define the model prior

$$P(\lambda|\gamma) = P(\lambda_\gamma|\gamma)P(\lambda_{\backslash\gamma}|\lambda_\gamma, \gamma), \tag{19}$$

where λ_γ represent the elements of λ which are defined so where γ is set to 1, and likewise $\lambda_{\backslash\gamma}$ represents the elements of λ which are not included. The second term in equation (19) is the so-called pseudo-prior because it is used to describe the elements not present in the current evaluated model specification. Following the reasoning of Bauer (2017) for the pseudo-prior specification I use independent Normal distributions with the parameters equal to the conditional moments of λ, given the ML estimates of all other parameters.

In the GVS algorithm for each iteration first a draw of λ_γ is made from its conditional posterior distribution. This distribution is already explained in equation (16). For the other parameters, $\lambda_{\backslash\gamma}$, the drawings are directly taken from the pseudo-prior. To determine the likelihood of the drawing the acceptance rate for γ is determined. This rate is a success probability for a Bernoulli conditional posterior of γ^i and defined by

$$\begin{aligned} &\frac{P(\gamma_i^{(j)} = 1|\lambda^{(j)}, \theta_-^{(j-1)}, \gamma_{-i}^{(j)}, Y)}{P(\gamma_i^{(j)} = 0|\lambda^{(j)}, \theta_-^{(j-1)}, \gamma_{-i}^{(j)}, Y)} \\ &= \frac{P(Y|\gamma_i^{(j)} = 1, \lambda^{(j)}, \theta_-^{(j-1)}, \gamma_{-i}^{(j)})}{P(Y|\gamma_i^{(j)} = 0, \lambda^{(j)}, \theta_-^{(j-1)}, \gamma_{-i}^{(j)})} \times \frac{P(\lambda_i^{(j)}|\gamma_i^{(j)} = 1)}{P(\lambda_i^{(j)}|\gamma_i^{(j)} = 0)} \times \frac{P(\gamma_i^{(j)} = 1, \gamma_-^{(j)})}{P(\gamma_i^{(j)} = 0, \gamma_{-i}^{(j)})}. \end{aligned} \tag{20}$$

The first term in equation (20) is the ratio of likelihoods of both included and excluded elements of γ, the second term is the ratio of model priors for the two sets of elements, and the third term is the is the ratio of priors for γ. To determine the first term on the right-hand side of the equation only the $\mathbb{P}$-likelihoods need to be evaluated because the $\mathbb{Q}$-likelihoods are not affected by γ and therefore cancels out. The middle term The last term cancels out as a whole because for both sets of elements the same prior probabilities are used (probability of 0.5).

3.2.2.2 Reversible-jump Markov Chain Monte Carlo

The Reversible-jump Markov Chain Monte Carlo (RJMCMC) model selection sampler, developed by Green (1995), is designed to evaluate different parameter spaces by making jumps in the drawings of γ. A lot of different approaches can be taken in modelling the jumps but in this paper the local reversible-jump sampler of Dellaportas and Forster (1999) is used. The current state of the chain is denoted by $(\lambda^{(j)}, \gamma^{(j)}, \theta_{-}^{(j)})$. For each jump two possibilities are available, a null move where γ does not change or a jump move where only 1 element of γ changes. The probability that the model doesjump to another parameter space is set equal to 75%. The current element of γ that changes is chosen randomly, and is set to either 0 or 1 dependant of its initial value. Whenever an element of γ is added to the model its proposal is $\lambda' = g(\lambda^{(j)}, u)$, where $g(\dot{})$ is an identity transformation to ensure that $\lambda_i' = u$, and u represents a scalar drawn from the proposal density equal to $\mathcal{N}(\mu_i, \sigma_i^2)$ and denoted by $q_i(u)$. Given that a element of γ is added to the model, the acceptance probability is

$$\begin{aligned}\alpha(\lambda^{(j)}, \gamma^{(j)}, \theta_{-}^{(j)}, \lambda', \gamma') &= \frac{P(Y|\lambda', \gamma', \theta^{(j)})}{P(Y|\lambda^{(j)}, \gamma^{(j)}, \theta_{-}^{(j)})} \times \frac{P(\lambda'|\gamma')}{P(\lambda^{(j)}|\gamma^{(j)})} \times \frac{P(\gamma')}{P(\gamma^{(j)})} \times \frac{1}{q_i(u)} \\ &= \frac{P(Y|\lambda', \gamma', \theta^{(j)})}{P(Y|\lambda^{(j)}, \gamma^{(j)}, \theta_{-}^{(j)})} \times \frac{v_i^{0.5} \exp(-u^2/v_i)}{\sigma_i \exp(-(u-\mu_i)^2/\sigma_i^2)},\end{aligned} \tag{21}$$

where the derivation is reasoned in the same way as with the GVS sampler because the prior model probability is the same. Moreover, the prior of γ is conditional independent and therefore together with the third term the second line of equation (21). When instead an element of γ is excluded from the model, the expressions do change, $(\lambda', u') = g^{-1}(\lambda^{(j)}$ which implies that λ' has a 0 at the i-th position. The acceptance probability for excluding an element of λ then becomes

$$\alpha(\lambda^{(j)}, \gamma^{(j)}, \theta_{-}^{(j)}, \lambda', \gamma') = \frac{P(Y|\lambda', \gamma', \theta^{(j)})}{P(Y|\lambda^{(j)}, \gamma^{(j)}, \theta_{-}^{(j)})} \times \frac{\sigma_i \exp(-(u'-\mu_i)^2/\sigma_i^2)}{v_i^{0.5} \exp(-u'^2/v_i)}. \tag{22}$$

3.2.2.3 Stochastic Search Variable Selection

The last sampler used in this paper is the first MCMC algorithm ever designed for variable selection purposes. The Stochastic Search Variable Selection (SSVS) approach was developed by George and McCulloch (1993). It uses the idea that the drawings for the parameters, excluded from the model, are drawn from a tight distribution around 0. This results in a parameter prior defined by

$$P(\lambda_i|\gamma_i) = (1-\gamma_i)N(0,\tau_{0,i}^2) + \gamma_i N(0,\tau_{1,i}^2), \tag{23}$$

where τ_0^2 is the prior variance for λ_j of excluded elements and τ_1^2 the variance for included elements. The usage of the τ parameters are in the same way used as in Bauer (2017) who ensures a high value for τ_1^2 and a value close to 0 for τ_0^2. In more detail, they are defined as $\tau_{k,i} = c_k\hat{\sigma}_{\lambda i}$ for $k = 0, 1$ with c_0 or c_1 two tuning parameters. I set $c_1 = \sqrt{g}$ and $c_0 = \frac{1}{c_1}$ with g the hyperparameter for λ.

Just as with the GVS sampler the success probability of a Bernoulli distribution is used to evaluate the likelihood of the drawing for γ. This probability is given as

$$\begin{aligned}\frac{P(\gamma_i^{(j)}=1|\lambda^{(j)},\theta_-^{(j-1)},\gamma_{-i}^{(j)})}{P(\gamma_i^{(j)}=0|\lambda^{(j)},\theta_-^{(j-1)},\gamma_{-i}^{(j)})} &= \frac{P(\lambda^{(j)}|\gamma_i^{(j)}=1,\theta_-^{(j-1)},\gamma_{-i}^{(j)})}{P(\lambda^{(j)}|\gamma_i^{(j)}=0,\theta_-^{(j-1)},\gamma_{-i}^{(j)})} \times \frac{P(\gamma_i^{(j)}=1)}{P(\gamma_i^{(j)}=0)} \\ &= \frac{\tau_{1i}^{-1}\exp(-0.5(\lambda_i^{(j)}/\tau_{1i})^2)}{\tau_{0i}^{-1}\exp(-0.5(\lambda_i^{(j)}/\tau_{0i})^2)},\end{aligned} \tag{24}$$

where again the same derivations of prior conditional independence and equal prior model probabilities are used to derive the second line of equation (24). For this sampler the ratio does not depend on any data because the likelihood of γ is fully described by λ.

3.2.3 Prior Selection

Recall that the set of parameters given in section 3.1 needs to be estimated and this set will be denoted by θ_{Full} in the remainder of this paper. For each different group of parameters in the full set a different choice can be made on the priors to use. Some Bayesian researchers choose to use priors that heavily influence the outcomes of the results and also to guide the drawings of the parameters to a certain degree. Especially, priors for the Macro-Economic Factors can be used because of an abundance of forecasts on these factors. An example for this approach that relies on priors to predict Macro-Economic variables is Wright (2013). Other researchers choose for largely uninformative or even diffuse priors. This way little to none information is imposed on the parameters and the resulting coefficients are mostly derived by the data itself. The manner of using uninformative priors mirrors the priors Bauer (2017) proposes and this approach is also used in this paper. Therefore, the priors for the $k_\infty^{\mathbb{Q}}$, Σ and σ_e^2 are diffuse and totally uninformative. However the prior for the eigenvalues of $K_{1X}^{\mathbb{Q}}$ needs to ensure the stationarity of this autoregressive component, thus the elements of this vector are a priori uniformly distributed over the $0 - 1$ interval. Only the priors for the pair (λ, γ) are left then. I assume an a priori uniform distribution between the different model specification that are defined by γ, and this leads to a prior of independent Bernoulli distributions with a success probability of 0.5.

3.3 Macro Economic Factors

Besides the Bayesian estimation of the term structure model, an extension will be made by adding Macro-Economic Factors to the model. This extension will, following the literature, help to better describe the short-end of the yield curve (Ang and Piazzesi, 2003). If those factors are constructed they can be added to the term structure model by following Joslin et al. (2014) in modifying the specification of the pricing measure of equation (1a) by

$$Z_{t+1} = \begin{pmatrix} X_{t+1} \\ M_{t+1} \end{pmatrix} = \begin{pmatrix} K^{\mathbb{P}}_{0X} \\ K^{\mathbb{P}}_{0M} \end{pmatrix} + \begin{bmatrix} K^{\mathbb{P}}_{XX} & K^{\mathbb{P}}_{XM} \\ K^{\mathbb{P}}_{MX} & K^{\mathbb{P}}_{MM} \end{bmatrix} \begin{pmatrix} X_t \\ M_t \end{pmatrix} + \sqrt{\Sigma_Z} \begin{pmatrix} \epsilon^{\mathbb{P}}_t \\ \epsilon^{M}_t \end{pmatrix}, \tag{25}$$

most variables in equation (25) are already defined but M_t stands for the Macro-Economic Factors at time t. Just as in equation (1a), $\epsilon^{\mathbb{P}}_t$ and ϵ^{M}_t are multivariate standard normal distributed. With the adding of these factors the shape of all parameters described in section 3.1 under the pricing measure $\mathbb{P}$ change accordingly.

The adding of these factors does however increase the dimension of parameters that needs to be estimated. This should be kept in mind when applying the MCMC algorithm, and it needs to be examined if the algorithm is still able to converge for the set of parameters. In this research the approach of Ang and Piazzesi (2003) will be followed in using two Macro-Economic Factors, namely one that represents inflation growth rate and another that stands for the economic growth rate. Just as for the parameters groups of $k^{\mathbb{Q}}_{\infty}$, Σ and σ^2_e diffuse priors will be assumed for the added elements in order to ensure that the results are mostly driven by the data.

3.3.1 Change of Measure

Due to the adding of the Macro-Economic Factors the relation between the $\mathbb{P}$ and $\mathbb{Q}$ measure has also changed. Fortunately this relation is only slightly changed as is shown in Joslin et al. (2014). The same derivation of the SDF is applied and therefore only equations (8) changed to the new set

$$
\begin{aligned}
\lambda_1 &\equiv K^{\mathbb{P}}_{XZ} - [K^{\mathbb{Q}}_{1X}, 0_{3x2}], \\
\lambda_0 &\equiv K^{\mathbb{P}}_{0X} - K^{\mathbb{Q}}_{0X},
\end{aligned}
\tag{26}
$$

where $K^{\mathbb{P}}_{XZ}$ are the first N rows of $K^{\mathbb{P}}_{Z}$ and therefore the size of λ_1 has become (3×5). The measure relation for the intercept terms in the model is not changed by the new variables and for the autoregressive component only the dimension is adjusted for the $\mathbb{P}$ measure.

3.3.2 Likelihood Function

By adding the Macro-Economic Factors to the term structure model also the likelihood function changes. However, due to the fact that only the part under the pricing measure $\mathbb{P}$ in equation (25) changes, only the part in equation (14) of the likelihood function is adjusted. The new log likelihood is then given by

$$
\begin{aligned}
&\log\left(L(Z_t|Z_{t-1}; k^{\mathbb{Q}}_{\infty}, K^{\mathbb{Q}}_{1X}, \lambda, \gamma, \Sigma_Z)\right) = \\
const - 0.5\log(|\Sigma_Z\Sigma_Z'|) &- 0.5||\Sigma^{-1}(Z_t - K^{\mathbb{P}}_{0Z} - K^{\mathbb{P}}_{1Z}Z_{t-1})||,
\end{aligned}
\tag{27}
$$

where K_{0Z} stands for the level matrix and K_{1Z} for the autoregressive matrix of Z_t from equation (25).

3.3.3 Bayesian estimation

By adding the Macro-Economic Factors to the term structure model, the dimension of the covariance matrix is increased from 3 to 5, and therefore the amount of parameters that are drawn each iteration also increases from 6 to 15 due to the vectorization of the Choleski decomposition. This increase results in some inefficiencies in the drawing procedures. This is mostly caused by the fact that the estimation of the Hessian matrix became less accurate. In order to overcome this problem the parameters of Σ are parted in two groups, with the first group the same size of the original model (so without the Macro-Economic Factors). The second group of $vech(\Sigma)$ represents the covariances between the Macro-Economic Factors and risk factors and also the variances of only the Macro-Economic Factors, and therefore this group only consists of 9 parameters. However, I still evaluate them as one block of parameters due to their high posterior correlation because they jointly explain the cross-sectional covariance of the model. This approach results in drawings that are only accepted if both groups have reasonable drawings.

3.4 Extreme Scenarios

For the purpose of extreme scenario calibration a method has to be used that is able to approximate the distribution of possible outcomes for the yield curve. In this paper scenarios are considered that will consider yields over a year. Following the approach of Chib and Ergashev (2009), the following algorithm is used to obtain the predictive density of the yields.

Algorithm 1 Algorithm to generate extreme scenarios for the yield curve

for $j = 1, 2, ..., M$ **do**
    Determine the loads: $A^j \& B^J$ and estimate the coefficients in equations (1a) till (4).
    for $h = 1, 2, ..., H$ **do**
        $\hat{X}^{\mathbb{P},j}_{T+h} = K^{\mathbb{P},j}_{X} X^{\mathbb{P},\hat{j}}_{T+h-1} + \eta^j_{T+h}$
        with $\eta^j_{T+h} \sim \mathcal{N}(0, \Sigma^j_X)$
        $y^j_{T+h} = A^j + B^j \hat{X}^{\mathbb{P},j}_{T+h} + e^j_{T+h}$
        with $e^j_{T+h} \sim \mathcal{N}(0, diag(\sigma^j_e))$
    Store y^j_{T+H}
Return $y^1_{T+H}, ..., y^M_{T+H}$

If all the draws are collected a distribution is made of the predicted yields. The two different type of risks, considered in this paper, use different approaches to establish the quantiles that determine the height of the shocks to the yield curve. For non-parallel Gap Risk different quantiles are determined for each different maturity and a new yield curve is constructed from these quantiles. This way the entire yield curve gets a different shock for different maturities. For parallel shocks different approaches can be taken. I will show the different outcomes when the shock is determined by only the short rate, only the 10 years rate, or using a weighted average of the shocks for each maturity. After the shock is determined the same effect is applied to the entirety of the yield curve. Following this approach only the level of the yield curve changes.

4 Results

In this section results of the model without and the model with Macro-Economic Factors will be discussed. First they will be examined in a simulation study. After that, historical data of 2008 till 2017 will be used to examine their performance.

4.1 Simulation Study

The same way of simulating data is used as in Bauer (2017). In short, the data is sampled by first estimating a term-structure model with 2 risk-factors. This model is estimated by maximum-likelihood estimation and from the resulting point estimates, the data on yields is simulated by using the autoregressive model from equation (1a). Then for each of 100 iterations a new dataset is generated of 300 observations and for this new set the Bayesian estimation is performed. For the extended model with Macro-Economic Factors, the same approach is used but then a term-structure model is estimated with 2 risk-factors and 2 Macro-Economic Factors. I choose for this different model, that results in different simulated data, because the data on the Macro-Economic factors also needs to be simulated. Moreover, in this way the dataset allows for retrieving the Data-Generating Parameters (DGP) that are used to generate the simulated data. In this section the robustness of the different model selection samplers are evaluated by this ability of retrieving the DGP, and therefore the simulated data should allow for that.

4.1.1 Base Model

Table 4: The following table shows whether the different model selection samplers are able to retrieve the correct model accordingly to the DGP. The first row depicts the vector of γ that is used to generate the simulated data. For the row with MCMC the values represent the percentage of iterations where the corresponding element of λ was significant in a Bayesian way, and this means that its confidence interval did to contain zero. The other rows represent the posterior means for the elements of γ because they do in fact estimate the element of γ in contrary to the unrestricted model which is represented by the second row. The most right column of the table gives the frequencies of iterations that the different methods were able to retrieve the exact same specification used for the DGP.

	Element of γ						Freq. of
	(1)	(2)	(3)	(4)	(5)	(6)	corr. model
DGP	0	0	1	0	0	0	
MCMC	0.07	0.05	0.99	0.09	0.06	0.04	76%
SSVS	0.05	0.01	0.84	0.02	0.06	0.01	89%
GVS	0.14	0.07	0.91	0.09	0.17	0.06	88%
RJMCMC	0.13	0.08	0.90	0.10	0.13	0.07	90%

In table 4 the first results are depicted, they show if the different model selection samplers are able to retrieve the model specification that is used to generate the data. The table shows that indeed all different algorithms are mostly able to generate the correct model. However, the MCMC sampler is less efficient in doing this, because in only 76% of the iterations each element of λ was correctly significant or not according to the DGP. This could be due to the amount of draws that are less than for the other samplers (15.000 versus 50.000 draws). The other samplers, that use γ to determine which elements of λ to use have percentages of around the 90% of the iterations that they correctly estimated

γ according to the DGP.

The results of table 4 for individual elements of γ are also in line with the percentages found in the most right column. For the elements that are used in the DGP the posterior means for γ of the different samplers are close to 1, the GVS sampler yields a mean of 0.91 for the third element of γ which correspond with the element of λ that is term premium of the autoregressive element for the risk-factor. Also the posterior means for the elements of γ, that are not used in the DGP and therefore equal to zero, are close to zero. The means are between 0.01 till 0.17 which is rather low. The table shows that the applied framework so the parameters are either rather conclusively taken 1 or 0 throughout the iterations. This indicates that for the dataset that is used in this research, the framework of estimating the term structure is sufficient in discovering the underlying structures. Throughout the sample it is rather clear that there is not much variation between different models, which is also due to the fact that over the last few years the variances in yield data in Europe are quite low. Therefore, in the simulation approach that I show in this study the resulting dataset does also not contain a lot of volatility.

Table 5: The results in this table depict the maximum absolute eigenvalue of the persistence component ($K_{1X}^{\mathbb{Q}}$) of equation (1a) for the different samplers. Also the values for the impulse-response function are given that represent the effect of shocks to the level factor at a 5-year horizon. The last two columns contain the implied volatilities for the 5-10 year risk-neutral forward rates (which are the expectations of the short rates) and also for the forward term premium. Both are scaled to annualized percentage points. The rows below the rows with posterior means depict the frequencies of iterations that the value of the DGP was present in the confidence interval.

		Persistence		Volatilities	
		max. eigenv.	IRF(5y)	$\Delta\tilde{f}_t$	Δftp_t
DGP		0.9674	0.1538	0.0255	0.2105
MCMC	posterior mean	0.9606	0.1273	0.0253	0.2156
	CI contains DGP	99%	93%	91%	92%
SSVS	posterior mean	0.9674	0.2842	0.0634	0.1729
	CI contains DGP	94%	94%	94%	86%
GVS	posterior mean	0.9629	0.2108	0.0432	0.1935
	CI contains DGP	95%	95%	94%	90%
RJMCMC	posterior mean	0.9650	0.2219	0.0457	0.1921
	CI contains DGP	99%	99%	98%	97%

In table 5 results are given for the autoregressive components in the model for the simulated data. It shows the values that are used in the DGP in the first row, and shows if the different samplers are able to approximate those specific values. The unrestricted MCMC sampler is already able to almost entirely retrieve the autoregressive component. This sampler is also able to correctly retrieve the implied volatilities of the forward rates.

The other samplers however are even better at it for this value, the posterior mean of the SSVS sampler is coincidentally the same as the DGP value. In contrary to this finding, the unrestricted model is better able at estimating the other aspects of the model. And therefore the values of the impulse-response function and both volatilities are better estimated when all elements of γ are set to 1.

4.1.2 Extended Model with Macro-Economic Factors

For the simulation study of the model with Macro-Economic Factors again the same 2-factor specification is used as with the base model. This means that γ has now 4 more elements included that represent the dependence of the two yield factors from the two Macro-Economic Factors. All other specifications such as the number of iterations and the different sampling algorithms were kept the same as with the base model.

During the simulation procedure all algorithms showed difficulties with retrieving correct acceptance rates for the different draws. This resulted in estimations that were not stable and were sometimes just incorrect. In this section I will illustrate this problem by showing the same results as for the specification without Macro-Economic Factors. Also the tuning parameters for the drawing algorithm of Chib and Ergashev (2009) could not solve the problem because the simulation study showed that for each different iteration a different set of tuning parameters was needed. Later on, in the section where I will discuss the empirical results I will show that for the data a suitable set of parameters can easily be found and used. However, this set was not applicable for the simulation study because each iteration a new dataset is used.

Nonetheless in table 6 the results are given for the different sampling algorithms. Immediately this table points out the implications of the problem with the simulation that is described before. All different sampling algorithms were not able to adequately retrieve the DGP parameters. In all different iterations this proved to be too difficult and therefore the percentages in the most right column of the table are equal to 0. The table also shows a probable cause of the problem. The amount of risk-parameters that the DGP includes is much higher than with the base model. For the base model only one element of λ was included to generate the data, but for the extended model now 7, out of possible 10, elements of λ are included in the DGP. The different algorithms clearly failed at retrieving this rather unrestricted specification, which is also shown by the posterior means of the individual parameters. The posterior means are much further away from 1 when an element is included in the DGP or from 0 when the element is not included. Also the different sampling algorithms are not consistent for which elements the posterior means are close to their DGP values. The RJMCMC algorithm however, used the full unrestricted model for each iteration. Apparently this algorithm ensured a very strong specification of the model for which it did not deviate in each iteration or drawing. One could argue that this specification is more in line with the DGP specification than for the other algorithms but it is not preferable that only one specification is evaluated for each drawing in each iteration.

Table 6: The following table shows whether the different model selection samplers are able to retrieve the correct model accordingly to the DGP when also Macro-Economic Factors are included in the model. The first row depicts the vector of γ that is used to generate the simulated data. For the row with MCMC the values represent the percentage of iterations where the corresponding element of λ was significant in a Bayesian way, and this means that its confidence interval did not contain zero. The other rows represent the posterior means for the elements of γ because they do in fact estimate the element of γ in contrary to the unrestricted model which is represented by the second row. The most right column of the table gives the frequencies of iterations that the different methods were able to retrieve the exact same specification used for the DGP.

	Element of γ										Freq. of
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	corr. model
DGP	1	1	1	0	0	1	1	1	1	0	
MCMC	0.13	0.26	0.96	0.05	0.09	0.57	0.25	0.20	0.55	0.20	0%
SSVS	0.22	0.31	0.76	0.41	0.20	0.43	0.24	0.63	0.89	0.52	0%
GVS	0.38	0.58	0.49	0.62	0.55	0.61	0.36	0.59	0.75	0.15	0%
RJMCMC	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0%

Some additional results of the simulation study are given in table 7 that provide more details on the posterior persistence and volatilites of the different models. Whereas with the base model all different sampling algorithms were able to retrieve the values from the DGP, now only the unrestricted model is able of estimating those values in some degree. These values are however somewhat lower than before, for example the percentage of iterations where the confidence interval of the maximum eigenvalue contained the DGP value is equal to 87% instead of 99%. Also for all different algorithms except the RJMCMC sampler, the impulse-response function became very extreme. This is another indication that the simulations are not stable because the effect of shocks to the level factor were too large. For the RJMCMC algorithm the posterior means seem to be rather close to the DGP values, but because the different drawings of this algorithm are all closely dependant on each other, which was already shown in table 6, the percentage of confidence interval that included the DGP value is very low. This lead to percentages that were equal to 1% for both volatilities.

4.1.3 Conclusion of Simulation Study

The simulation study shows that the Bayesian estimation of the term structure model is able to retrieve the DGP parameters rather well when the parameter space is not too large. When the Macro-Economic Factors are added to the specification the estimation requires tuning parameters that need to be different for each dataset in order to produce stable results. With this simulation study it was not possible to manually adjust these parameters each iteration and therefore the results, discussed in this section, show a negative image of the approach.

Table 7: The results in this table depict the maximum absolute eigenvalue of the persistence component ($K_{1X}^{\mathbb{Q}}$) of equation (25) for the different samplers. Also the values for the impulse-response function are given that represent the effect of shocks to the level factor at a 5-year horizon. The last two columns contain the implied volatilities for the 5-10 year risk-neutral forward rates and also for the forward term premium. Both are scaled to annualized percentage points. The rows below the rows with posterior means depict the frequencies of iterations that the value of the DGP was present in the confidence interval.

		Persistence		Volatilities	
		max. eigenv.	IRF(5y)	$\Delta \tilde{f}_t$	Δftp_t
DGP		1.0050	-0.0254	0.0219	0.1832
MCMC	posterior mean	1.1122	− inf	0.0286	0.2165
	CI contains DGP	87%	93%	94%	74%
SSVS	posterior mean	1.0563	+ inf	0.1126	0.1843
	CI contains DGP	12%	20%	10%	17%
GVS	posterior mean	1.0339	− inf	0.1164	0.1737
	CI contains DGP	34%	31%	25%	32%
RJMCMC	posterior mean	1.0039	-0.0630	0.0267	0.2088
	CI contains DGP	52%	68%	1%	1%

4.2 Empirical Data

In this section, real empirical data is used, as described in section 2. The data is then used to construct principal components of it. First only these factors are used to model to term-structure. Later in this section also the empirical results for the extended model with Macro-Economic Factors will be discussed.

4.2.1 Base Model

The results in table 8 show us what happens to the parameters when the full model is evaluated with only 3 risk factors. Most posterior means of the different parameters are significant in the Bayesian framework. Examples are the autoregressive components in the risk-neutral part of the model. Those all turned out to be significant and so shows the strong persistence in the risk factors. Also the acceptance percentages seem to be quite in line in what literature suggests to be good acceptance rates (Gelman et al., 1996).

Table 8: This table contains the posterior results for the unrestricted model without Macro-Economic Factors included, so each element of γ is equal to 1. Significant results are printed bold and they are significant when their confidence interval does not contain zero. The standard errors of the posterior means are given in parentheses in the row below. In the last two columns of the table the average acceptance rates and an inefficiency factor are given. The acceptance rates are calculated by the various equation given in section 3.2.1. The inefficiency factors are calculated by determining the Newey-West adjusted first autocorrelations of the drawings. The parameters $k_\infty^{\mathbb{Q}}$, λ_0, Σ and σ_e are all scaled by 1200 to represent annualized percentage points.

Parameter	Prior			Posterior			Acc.	Ineff.
$k_\infty^{\mathbb{Q}}$				**0.0586**			27.6	11.9
				(0.0095)				
$K_{1X}^{\mathbb{Q}}$	0.5000	0.5000	0.5000	**0.9872**	**0.9751**	**0.9751**	27.6	10.3
	(0.2887)	(0.2887)	(0.2887)	(0.0025)	(0.0028)	(0.0028)		
λ_0	0	0	0	**0.1028**	-0.0083	**0.0391**		1.0
	(0.5065)	(0.7695)	(0.1538)	(0.0509)	(0.0760)	(0.0155)		
λ_1	0	0	0	**-0.0233**	-0.0120	**-0.4123**		1.0
	(0.0942)	(0.1644)	(1.7126)	(0.0094)	(0.0166)	(0.1736)		
	0	0	0	0.0007	-0.0103	0.0002		
	(0.1430)	(0.2497)	(2.6019)	(0.0141)	(0.0238)	(0.2623)		
	0	0	0	0.0053	-0.0010	**-0.1801**		
	(0.0286)	(0.0499)	(0.5201)	(0.0029)	(0.0050)	(0.0527)		
Σ				**0.1411**	0	0	23.2	13.1
				(0.0052)				
				0.0502	**0.2068**	0		
				(0.0125)	(0.0088)			
				-0.0120	0.0029	**0.0402**		
				(0.0023)	(0.0023)	(0.0017)		
σ_e				**0.0896**				0.8
				(0.0029)				

Furthermore, the restrictions on the risk premium are moderately restrictive because 5 out of 12 parameters tend to be significant for the unrestricted model. This indicates that quite some restrictions are needed in order to define the term premium. Bauer (2017) finds an opposite finding in this and comes to the conclusion that his data supports a specification where many risk-prices are set to 0. Apparently the data of yields on 6-months Euribors needs a different kind of model. A model where the relations between the risk-neutral and the pricing measure are more present and have a greater effect on the outcomes of the model.

In table 9 results are given for each individual element of γ, the vector that consist of indicator variables indicating if the corresponding element of λ is taken into account or not. The differences between the various sampling algorithms already becomes visible.

Especially, the SSVS sampler is not able to accurately estimate the γ parameters which resulted in high inefficiency factors. This leads to the conclusion that this sampler does not evaluate many different models, a result which will be further illustrated by table 10. The other two samplers have overall lower inefficiencies and have visited more different versions of the model. In this way they are better at estimating the different parameters and this is also shown by the fact that they produced lower standard errors than the SSVS sampler.

Table 9: This table contains posterior results for the γ parameter for each different sampling algorithm. The first three rows of γ correspond with the intercept of the risk factors (λ_0) and the other rows the change in measures for the autoregressive component (λ_1) in equations (8). Posterior means and their standard deviations are given in each first two columns. Also inefficiency factors are given that are determined by taking the Newey-West adjusted first order autocorrelations. These factors indicate if different types of drawings are evaluated, a high factor means that consecutive drawings were similar to each other.

	SSVS			GVS			RJMCMC		
	Mean	MCSE	Ineff.	Mean	MCSE	Ineff.	Mean	MCSE	Ineff.
1	0.164	0.034	410.8	0.189	0.013	51.1	0.117	0.014	91.6
2	0.003	0.002	63.1	0.045	0.001	2.0	0.040	0.005	30.9
3	0.170	0.037	475.0	0.887	0.025	303.4	0.917	0.025	402.8
4	0.063	0.024	468.8	0.668	0.009	18.1	0.611	0.027	151.7
5	0.007	0.003	54.4	0.070	0.001	1.6	0.072	0.007	31.8
6	0.015	0.008	214.3	0.318	0.006	7.0	0.316	0.018	73.9
7	0.154	0.035	467.4	0.118	0.003	3.6	0.127	0.011	53.3
8	0.010	0.005	109.5	0.076	0.001	1.1	0.081	0.007	31.4
9	0.007	0.003	72.2	0.092	0.002	2.6	0.097	0.008	33.2
10	0.743	0.043	477.5	0.463	0.016	48.5	0.445	0.029	175.3
11	0.003	0.002	89.3	0.044	0.001	1.8	0.046	0.006	39.6
12	0.164	0.036	482.7	0.908	0.023	313.3	0.937	0.022	400.3

After some results for specific parameters are considered in the previous tables now results will be discussed of all different models that were visited in the algorithm. In table 10 those results are given and one can easily notice that the problems with the SSVS sampler as discussed for table 9 are further illustrated. This sampler has only visited a rather small fraction of possible models (only 47 out of 4096). Moreover, the models that it did visit do not correspond with the performance of the other two samplers in terms of their frequency. This can either indicate that the SSVS sampler finds other restrictions that are more important, or that this sampler is not able to find the correct specifications. When looking at the criterions given for each different model, one could only conclude that the SSVS sampler is not finding the right restrictions. This is due to the fact that the ordering of best models, accordingly to the GVS sampler, corresponds with the ordering of the ML estimates of the model. Together with the findings found at 9 and the fact that the GVS and RJMCMC sampler visited more than 10% of the total models, shows that the SSVS sampler tends to give the wrong restrictions a lot of weight. Also the restrictions that it does enforce ensure that the sample tends to overdraw them and therefore is not able to visit an adequately fraction of the total models.

The GVS sampler seems to produce also more stable results than the RJMCMC sampler which can be concluded by for example how the posterior odds are given for each different model. Where for the GVS sampler the odds gradually decline as the frequencies decline, the RJMCMC sampler does show a slightly more steep decline. The exception is off course the model where the first, fourth and tenth elements of γ are set to 1. Apparently that set of parameters is extremely not favored by the RJMCMC sampler and in approximation only visited a total of 5 times throughout the algorithm. Just as in Bauer (2017) the GVS sampler shows the most stable results and as I will show later on the same conclusion can be drawn for the extended model. Therefore the results for this sampler will be further discussed throughout the remainder of this paper.

Table 10: In this table posterior results are given on the different specifications of γ. In total there are 2^{12} different models, but only the 10 best models are given here in terms of the frequency they are visited by the algorithm. The ordering of specifications is given according to the GVS sampler. The numbers in the first column correspond with the elements of γ that are equal to 1. Each first column represent the fraction of drawings the sampler visited that model. The second row for that sampler gives then the relative odds ratios compared with the model that is visited the most. In the two most right columns Aikake information criterion and the Schwartz-Bayes information criterion are given that are retrieved from the ML estimates for that model. Below the table the amount of different specifications that are evaluated are given.

	GVS		SSVS		RJMCMC		AIC	SBIC
3,4,12	0.1632	1.0000	0.0562	1.0000	0.1481	1.0000	-13286.9494	-13250.7120
3,10,12	0.1229	1.3282	0.0846	0.6638	0.1464	1.0115	-13288.1642	-13251.9268
3,4,6,12	0.1145	1.4247	0.0000	Inf	0.1220	1.2140	-13289.0947	-13250.0698
3,4,10,12	0.0356	4.5786	0.0026	21.6000	0.0311	4.7553	-13288.6795	-13249.6546
3,6,10,12	0.0224	7.2979	0.0016	36.0000	0.0287	5.1668	-13287.6062	-13248.5813
3,4,7,12	0.0217	7.5337	0.0000	Inf	0.0239	6.1855	-13286.5169	-13247.4920
1,3,4,10,12	0.0187	8.7169	0.0000	Inf	0.0087	17.0599	-13289.8852	-13248.0729
1,3,10,12	0.0183	8.9169	0.0055	10.2482	0.0144	10.2976	-13287.6076	-13248.5827
1,4,10	0.0182	8.9561	0.0000	Inf	0.0001	2468.0000	-13281.6479	-13245.4105
3,4,9,12	0.0165	9.9017	0.0013	42.5455	0.0143	10.3264	-13285.1524	-13246.1275
models	773 / 4096		47 / 4096		436 / 4096			
visited	18.9 %		1.1 %		10.6 %			

In table 11 the stability of the GVS sampler is further accessed. Namely for the stability for the prior of λ which is the same prior as used in Bauer (2017). A higher value for the g-value leads to less different models that are visited but a higher frequency for the model that is favored the most. This also leads to a lower amount of the posterior mean of the amount of parameters for λ that are included. However this table shows again with the previous models that the data that is used in this research asks for a moderate amount of risk-price parameters to be taken into account. For example, when the posterior probabilities are considered that at most 3 parameters are included in the model, the percentages quickly drop to low amounts. This implies that indeed more than 3 parameters are needed to efficiently estimate the term structure.

Table 11: This table contains the posterior results for different model priors of λ. These results are examined by looking at the GVS model selection sampler. The first column depicts the frequency that the best model according to table 10 is visited for the different priors. The second column represent the amount of model specifications that were visited for a minimum of at least 1% percent. The third column gives the total amount of models visited. The last two column give results for the amount of restrictions that the GVS sampler imposes. The first column gives the average amount of elements of γ that were equal to 1, and the last column gives the percentage of drawings where the amount of included risk-prices were at most equal to 3.

		Models visited		Posterior	
g	Frequency $\mathbb{M}_1$	freq $\geqslant$ 1%	total	$E(a)$	$P(a \leqslant 3)$
10,000	32.3%	10	176	1.7	96.9%
1,000	26.1%	13	364	2.9	77.4%
100	16.3%	23	773	3.9	39.2%
10	4.8%	14	1409	5.5	7.0%

4.2.2 Extended Model with Macro-Economic Factors

In this section the same results will be shown as for the model without Macro-Economic Factors. However, only some differences will be discussed to avoid that similar conclusions will be drawn. I also want to point that I endured some difficulties in estimating the extended models with those factors. For the setting that I used the algorithm of Chib and Ergashev (2009) did not work optimal anymore. One of the motivations of that algorithm was that it did not had to use tuning parameters anymore to generate draws from the proposal distribution. However, I experienced that this resulted in draws that gave extremely low acceptance rates. Therefore I had to reinstate tuning parameters in the algorithm to come up with the results I will present in the following sections.

Again in table 12 posterior results are given for the full unrestricted model. For the results not much has changed for the amount of parameters that are significant for this model. All elements of the covariance matrix are now significant and roughly the same percentage of risk-price parameters are significant (8/18 instead of 5/12). However the acceptance rate for the both the autoregressive component of the risk-neutral measure and the short-rate have increased. They are still in what many researchers find to be sufficient. What we also can notice is that their inefficiency has also greatly increased. This is rather surprising because the procedure of estimating those parameters has not changed when compared to the model without Macro-Economic Factors. Apparently due

Table 12: This table contains the posterior results for the unrestricted model that is extended to also include 2 Macro-Economic Factors. Each element of γ is equal to 1 because the unrestricted model is discussed with this table. Significant results are printed bold and they are significant when their confidence interval does not contain zero. The standard errors of the posterior means are given in parentheses in the row below. In the last two columns of the table the average acceptance rates and an inefficiency factor are given. The acceptance rates are calculated by the various equation given in section 3.2.1. The inefficiency factors are calculated by determining the Newey-West adjusted first autocorrelations of the drawings. The parameters $k_\infty^{\mathbb{Q}}$, λ_0, Σ and σ_e are all scaled by 1200 to represent annualized percentage points.

Parameter	Prior					Posterior					Acc.	Ineff.
$k_\infty^{\mathbb{Q}}$						**0.0013**					42.8	355.1
						(0.0003)						
$K_{1X}^{\mathbb{Q}}$	0.5000	0.5000	0.5000			**0.9831**	**0.9831**	**0.9769**			42.8	372.0
	(0.2887)	(0.2887)	(0.2887)			(0.0002)	(0.0002)	(0.0002)				
λ_0	0	0	0	0	0	0.1423	**0.2174**	**0.0583**				1.1
	(0.7703)	(1.1012)	(0.2134)			(0.0752)	(0.1065)	(0.0210)				
λ_1	0	0	0	0	0	**-0.0440**	0.0005	**-0.4738**	0.0226	0.0012		1.1
	(0.1663)	(0.2102)	(2.1420)	(0.1270)	(0.1305)	(0.0159)	(0.0205)	(0.2117)	(0.0122)	(0.0126)		
	0	0	0	0	0	-0.0286	**-0.0779**	-0.4690	-0.0002	**-0.0755**		
	(0.2377)	(0.3005)	(3.0625)	(0.1816)	(0.1866)	(0.0229)	(0.0297)	(0.2994)	(0.0178)	(0.0183)		
	0	0	0	0	0	0.0006	-0.0003	**-0.2109**	0.0054	-0.0019		
	(0.0461)	(0.0582)	(0.5933)	(0.0352)	(0.0362)	(0.0045)	(0.0058)	(0.0590)	(0.0035)	(0.0036)		
Σ						**0.1486**	0	0	0	0	38.6	3.7
						(0.0043)						
						0.0985	**0.1911**	0	0	0		
						(0.0093)	(0.0052)					
						-0.0148	**0.0030**	**0.0391**	0	0		
						(0.0008)	(0.0003)	(0.0018)				
						0.0031	**0.0008**	**0.0004**	**0.3613**	0		
						(0.0001)	(0.0001)	(0.0001)	(0.0001)			
						-0.0006	-0.0000	0.0001	**0.0581**	**0.3590**		
						(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)		
σ_e						**0.0987**						17.7
						(0.0033)						

to the extended parameter space in the historic measure part also the parameters in the risk-neutral measure became harder to estimate.

On the other hand, table 12 also shows us that for the unrestricted model the effect of the Macro-Economic Factors have a significant effect on the risk-factors. The first factor, that represents the inflation, has an positive effect on the first risk-factor which is the one that is linked with the level of the yield curve. The other factor, which stands for the real economic growth, has a significant negative effect on the second risk factor, the one that is associated with the slope of the yield curve.

Table 13: This table contains posterior results for the γ parameter for each different sampling algorithm. The first three rows of γ correspond with the intercept of the risk factors (λ_0) and the other rows the change in measures for the autoregressive component (λ_1) in equation (26). Posterior means and their standard deviations are given in each first two columns. Also inefficiency factors are given that are determined by taking the Newey-West adjusted first order autocorrelations. These factors indicate if different types of drawings are evaluated, a high factor means that consecutive drawings were similar to each other.

	SSVS			GVS			RJMCMC		
	Mean	MCSE	Ineff.	Mean	MCSE	Ineff.	Mean	MCSE	Ineff.
1	0.026	0.009	149.0	0.112	0.009	38.4	0.129	0.019	153.2
2	0.008	0.003	54.7	0.080	0.004	8.9	0.071	0.010	73.6
3	1.000	0.000		0.906	0.027	418.3	1.000	0.000	
4	0.500	0.048	455.7	0.733	0.013	46.0	0.824	0.026	229.8
5	0.047	0.013	188.8	0.104	0.003	3.7	0.086	0.010	64.0
6	0.012	0.005	112.2	0.144	0.009	30.8	0.114	0.012	68.9
7	0.013	0.004	76.8	0.097	0.003	4.0	0.093	0.011	69.3
8	0.147	0.031	381.1	0.759	0.006	9.3	0.709	0.018	76.8
9	0.008	0.003	40.1	0.079	0.002	2.1	0.092	0.009	50.0
10	0.116	0.028	385.2	0.280	0.015	58.1	0.211	0.026	197.0
11	0.007	0.003	61.3	0.065	0.002	4.1	0.060	0.008	54.8
12	1.000	0.000		0.906	0.027	433.6	1.000	0.000	
13	0.108	0.027	391.6	0.522	0.010	19.0	0.612	0.024	125.8
14	0.037	0.011	160.7	0.153	0.002	2.1	0.147	0.011	44.2
15	0.258	0.039	401.5	0.547	0.010	21.4	0.559	0.016	52.9
16	0.241	0.038	402.2	0.252	0.010	24.1	0.205	0.018	95.1
17	0.870	0.033	470.3	0.996	0.001	9.9	0.999	0.000	13.0
18	0.008	0.002	28.9	0.084	0.002	4.0	0.071	0.008	46.5

Table 13 shows that even by increasing the dimensionality of the parameter space for the risk-prices the individual results for estimating the elements of γ have not really changed. Again the GVS seems to be most efficient in estimating those which can be reasoned by looking at the standard errors and inefficiency factors that are both lower than for the other samplers.

Table 14: In this table posterior results are given on the different models that are examined. In total there are 2^{18} different models, but only the 10 best models are given here in terms of the frequency they are visited by the algorithm. The numbers in the first column correspond with the elements of γ that are equal to 1.

	GVS		SSVS		RJMCMC		AIC	SBIC
3,4,8,12,13,17	0.0772	1.0000	0.0368	1.0000	0.1085	1.0000	-13331.3662	-13286.7663
3,4,8,12,13,14,17	0.0716	1.0783	0.0020	18.2277	0.0869	1.2493	-13334.5	-13287.2
3,4,8,12,15,17	0.0306	2.5183	0.0517	0.7116	0.0306	3.5424	-13328.3	-13283.7
3,8,10,12,15,17	0.0241	3.2043	0.0000	Inf	0.0119	9.0905	-13329.9	-13285.3
3,4,8,12,14,15,17	0.0198	3.8930	0.0141	2.6188	0.0213	5.1054	-13330.7	-13283.3
3,4,6,8,12,13,17	0.0160	4.8105	0.0000	Inf	0.0193	5.6297	-13332.1	-13284.8
3,4,12,14,16	0.0143	5.4034	0.1101	0.3344	0.0177	6.1322	-13326.5	-13284.6
3,8,10,12,15,16,17	0.0140	5.5272	0.0000	Inf	0.0101	10.7465	-13331.9	-13284.5
3,8,10,12,17	0.0131	5.8721	0.0056	6.5986	0.0059	18.4592	-13324.6	-13282.8
3,4,12,13,15,17	0.0104	7.4335	0.0079	4.6373	0.0202	5.3733	-13330.6	-13286.0
models	4312 / 262144		146 / 262144		1288 / 262144			
visited	1.6 %		0.1 %		0.5 %			

When all different models are compared with each other some effects that were earlier discussed seem to strengthen. In table 14 the percentages of models that are visited are given. The SSVS sampler seems to be even less efficient when the extra variables are added to the model. The problem of tightly restricting itself to only a few possible specifications have led to the occurrence that only 83 out of 262144 total models are evaluated. The other two samplers also seem to have this problem of visiting not that many different models. However, the models that both the GVS and the RJMCMC sampler do visit do not differ that much from each other. In terms of their criterions they seem to perform almost even well. What is also notable is the fact that for all models that are depicted in table 14 the seventeenth element of γ is equal to 1. Again this element corresponds with the effect of the second Macro-Economic Factors on the slope of the yield curve. This finding indicates that the adding of these factors does indeed change the models that are estimated.

Table 15: This table contains the posterior results for different model priors of λ. These results are examined by looking at the GVS model selection sampler. The first column depicts the frequency that the best model according to table 14 is visited for the different priors. The second column represent the amount of model specifications that were visited for a minimum of at least 1% percent. The third column gives the total amount of models visited. The last two column give results for the amount of restrictions that the GVS sampler imposes. The first column gives the average amount of elements of γ that were equal to 1, and the last column gives the percentage of drawings where the amount of included risk-prices were at most equal to 6.

		Models visited		Posterior	
g	Frequency $\mathbb{M}_1$	$freq \geqslant 1\%$	total	E(a)	$P(a \leqslant 6)$
10,000	20.4%	16	725	2.6	99.6%
1,000	5.2%	22	2068	4.2	92.3%
100	7.7%	12	4312	6.8	39.7%
10	2.2%	2	8257	9.0	4.6%

4.3 Economic Implications

In the following section more detailed results are discussed that follow from the predicted model. Namely some economic consequences will be of importance in this section. Therefore, the most robust sampler has been used and as I have showed this is the GVS sampler. Also for the three best models accordingly to table 10 and 14 the standard MCMC sampler has again been used to validate those specified models.

4.3.1 Base Model

In table 16 results are given for the autoregressive results. In order to let the implied yields be stationair all eigenvalues need to below 1. If this is not the case out of sample multiple period ahead forecasts will explode and therefore not be feasible. Correctly, all different model specifications ensure that all eigenvalue are below 1. For most of the different specifications of the MCMC sampler the persistence under pricing measure becomes lower than for the risk-neutral measure. However for the second model specification the persistence becomes higher under the pricing measure, which could be caused by the fact that for this model specification also the volatility in implied risk-neutral forward rates (which refer to the expectations of short rates) is rather high compared to the posterior means of the other model specifications.

Onwards from this section, the GVS sampler is called the Bayesian Model Average (BMA) due to the fact that it has the highest performance as shown in the previous section. For this BMA model the maximum persistence becomes almost equal for both measures in table 16. Again the volatility in the risk-neutral rates is higher than for example the unrestricted model so this indicates that even between different model specifications different economic implications can be noticed from the resulting model.

Table 16: This table contains details for the autoregressive component of the term structure model, but only the maximum eigenvalues are shown. Where the first value represents the posterior mean and the value between parentheses the posterior median. The different models depicted by M1, M2 and M3 correspond with the three best models following table 10. The M0 model stands for the unrestricted model so where all elements of γ are set to 1. Also the volatilities of changes in implied 5-10 forward rates, changes in risk-neutral forward rates and risk premiums are given in the 3 columns to the right. Below the posterior means and medians, 95% confidence intervals are given in the square brackets below.

Model	Max. eigenvalue		Volatilities		
	$\mathbb{Q}$	$\mathbb{P}$	$\Delta\hat{f}_t$	$\Delta\tilde{f}_t$	Δftp_t
M0	0.9871 (0.9873)	0.9713 (0.9728)	0.22 (0.21)	0.03 (0.02)	0.21 (0.21)
	[0.9821, 0.9918]	[0.9345, 0.9975]	[0.20, 0.24]	[0.00, 0.15]	[0.15, 0.26]
M1	0.9874 (0.9875)	0.9732 (0.9734)	0.22 (0.22)	0.04 (0.03)	0.19 (0.19)
	[0.9827, 0.9918]	[0.9579, 0.9867]	[0.21, 0.24]	[0.02, 0.08]	[0.15, 0.22]
M2	0.9839 (0.9836)	0.9989 (0.9989)	0.22 (0.22)	0.23 (0.23)	0.16 (0.16)
	[0.9801, 0.9893]	[0.9984, 0.9997]	[0.20, 0.24]	[0.21, 0.26]	[0.14, 0.20]
M3	0.9803 (0.9802)	0.9797 (0.9800)	0.20 (0.20)	0.03 (0.03)	0.18 (0.18)
	[0.9782, 0.9827]	[0.9696, 0.9891]	[0.19, 0.22]	[0.01, 0.07]	[0.15, 0.20]
BMA	0.9874 (0.9876)	0.9845 (0.9832)	0.21 (0.21)	0.11 (0.05)	0.19 (0.18)
	[0.9823, 0.9919]	[0.9618, 0.9996]	[0.19, 0.23]	[0.01, 0.34]	[0.12, 0.39]

In figure 3 graphical results for the posterior results of the implied yields are given. In the top subfigure implied yields for the risk-neutral measureare plotted along with the actual yields and implied yields by the pricing measure. The implied risk-neutral yields for the unrestricted model are very persistent and therefore stay at a stable level slightly above the 0% threshold. It remains rather unchanged by the downwards trend noticed in the actual yields. The implied yields by the pricing meausure however do follow the downwards trend and seem to fit the yield over time rather well. On the other hand the implied risk-neutral yields from the BMA model are indeed affected by the downwards trend in yields and from the year 2012 they are even fitted to be negative. Because this implication is rather not that sensible I decided to not adjust the scales in order to let them still be visible.

In the bottom panel of figure 3 the implied term-premium is plotted which follows from fitting the λ matrix. In this subfigure we can draw almost the same conclusions as from the top subfigure. Because the term premium of the unrestricted almost fits the observed yields beginning at mid 2012, the implied risk-neutral yields becomes almost zero. This can easily be argued when their relation is considered in equation (1b). The same reasoning can be taken for the observation that the risk-premium for the BMA model is at a higher level than the implied yields. This also ensures that the implied risk-neutral become negative as is again already illustrated by the top panel.

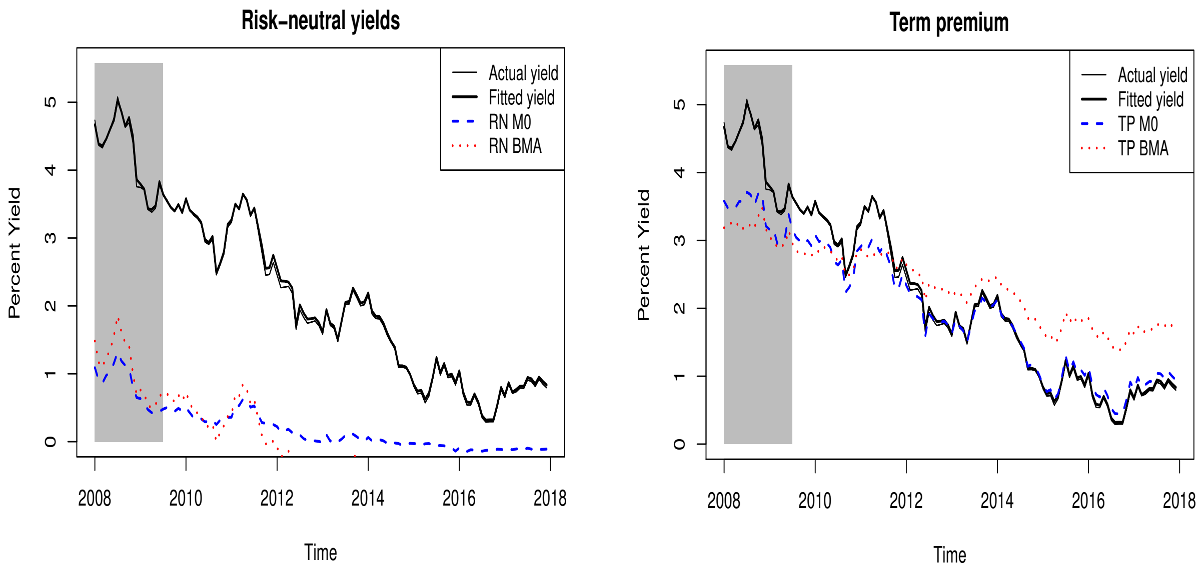

Figure 3: Figures of the expected risk-neutral yields and the term premiums over the used time period for the 10-year yield. The plotted model-implied yields are not that much visible but that is caused by the fact that the fitted yields are rather close to the actual ones. The fitted yields are obtained by the unrestricted model. In the left panel additional to the implied yields under the pricing measure $\mathbb{P}$ also the implied yields under the risk-neutral measure $\mathbb{Q}$ are plotted. In the right panel the values of the term premium is plotted, which is calculated by taking the difference between the two implied yields under measures $\mathbb{Q}$ and $\mathbb{P}$. The time periods that are depicted gray represent the crisis period, to further illustrate the movements during this period.

In figure 4 actual and fitted yields are plotted for a maturity of 5 years. For this maturity the implied risk-neutral yields are more substantial than for a maturity of 10 years. This also implies that the term premium has become lower which the lower sub-figure also shows. However eventually when the yields become lower, beginning in 2013, the risk-neutral yields again end up at a lower bound of almost zero for the unrestricted model and negative for the BMA model.

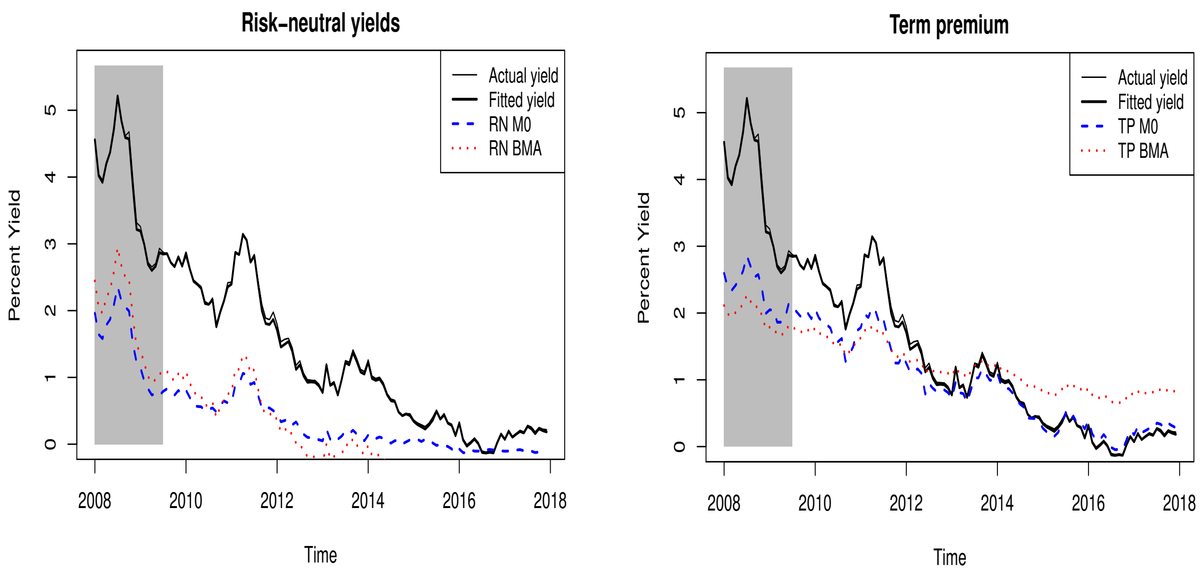

Figure 4: Same figure as for figure 3 but in this figure yields are plotted with a maturity of 5 years.

During the crisis of 2008 the yields with a maturity of 6 months declined the most steepest as is shown in figure 5. This figure also shows that the implied yield for a 6-

months maturity is almost fully declared by the implied risk-neutral yield. For both the unrestricted model as the BMA model is this the case.

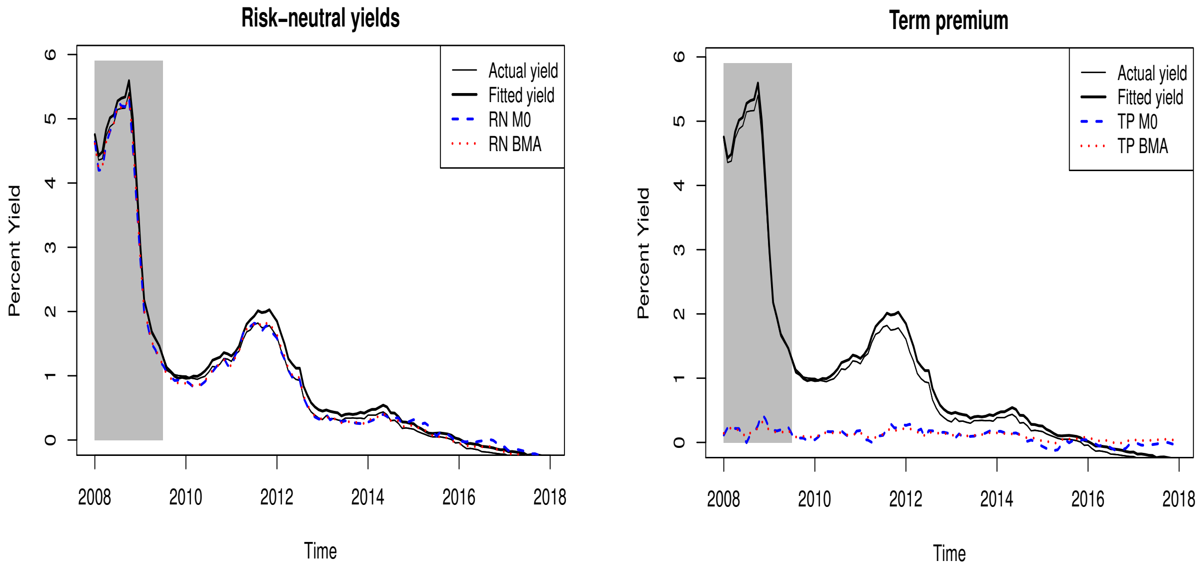

Figure 5: Same figure as for figure 3 but in this figure yields are plotted with a maturity of 6 months.

In figure 6 also the posterior results for the estimated volatilities are plotted. In the upper subfigure results for the unrestricted model are given and in the lower one results results for the BMA model are given. The unrestricted model is able to produce a rather tight confidence interval of the risk-neutral yields which even slightly becomes more tighter as maturities increases. The posterior mean of the volatility is for both models almost the same, very stable throughout the different maturities and only slowly increasing when the horizon increases. The BMA model however, does implicate a larger confidence interval around the risk-neutral yields. The lower bound of this interval seems to be similar for both models but the upper bound is approximately two times higher. This could be attributed to the fact that the unrestricted model uses all elements of λ which leads to a stricter specification where there is less room for deviation for the implied yields. Also the length of this interval for the BMA model seems to increases for longer maturities than for the unrestricted model. For the unrestricted model the interval becomes tighter when the maturity is 18 months or longer, but for the BMA model the interval only becomes tighter when the the maturity is 36 months or longer.

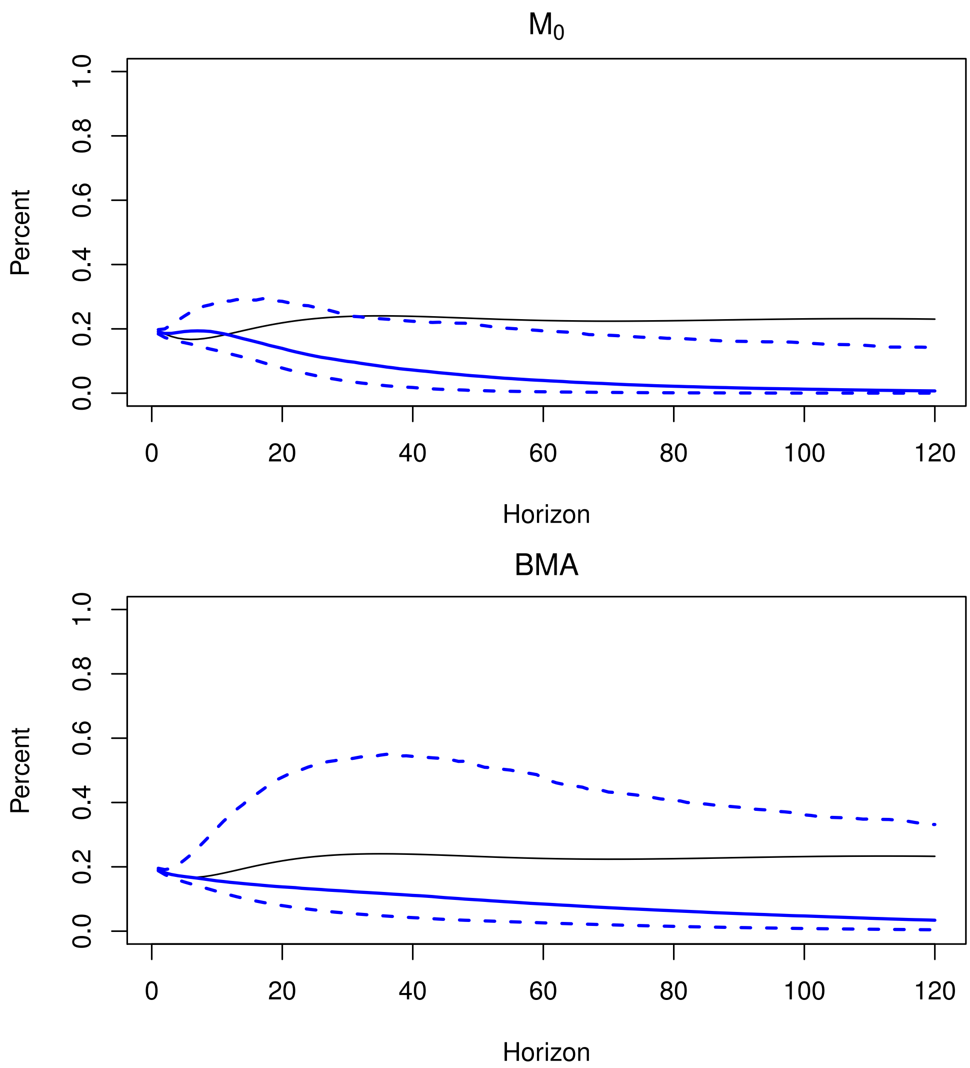

Figure 6: Figures of the posterior means of the volatilities in changes over the fitted forward rates, that are denoted by a solid black line for the different maturities that are given in months. In the top panel these means are given for the unrestricted model and in the lower panel for the BMA model. Also risk-neutral forward rates are plotted in a blue line, together with their credibility intervals that are represented by the blue dashed lines.

At last, the results in table 17 represent the ability of the different models in forecasting the yields. Because a rather small dataset is used in this research of only 10 years, their ability has also been accessed by simulating yields and evaluating if the model is able to declare the variation in bond returns. A good model should be able to explain the variation in the yields, as is also shown in other papers that research term structure models (Dai and Singleton, 2002).

What we can actually see from the results in table 17 is that the simulating procedure is quite necessary to see the potential of the proposed method in this paper. This is proven by the fact that the amount of variation that the models can actual declare is much higher when one looks at the simulated samples. This is also the case for the different maturities. However where the observed yield data is better able at predicting bond returns for higher maturities, the other models show the opposite effect.

Table 17: This table contains the predictability for some selected maturities. These predictabilities are denoted by the R^2 implied by the different models in declaring the variance in bond returns. The *Data* column represents the percentage of variation in bond returns that the observed yield data can declare. The *Pop.* columns represent the odds for the used sample period and the *Smpl.* columns are generated by simulating yields. For these simulated returns also their standard errors are given in brackets.

		M0		M1		M2		M3		BMA	
Maturity	Data	Pop.	Smpl.	Pop.	Smpl.	Pop.	Smpl.	Pop.	Smpl.	Pop.	Smpl.
3 years	0.09	0.23	0.45	0.27	0.47	0.01	0.35	0.31	0.49	0.18	0.42
			(0.17)		(0.16)		(0.18)		(0.15)		(0.17)
5 years	0.13	0.22	0.46	0.21	0.45	0.00	0.36	0.24	0.47	0.14	0.42
			(0.16)		(0.15)		(0.17)		(0.15)		(0.16)
7 years	0.18	0.20	0.45	0.15	0.42	0.00	0.36	0.18	0.44	0.11	0.41
			(0.16)		(0.16)		(0.17)		(0.15)		(0.16)
10 years	0.23	0.21	0.45	0.12	0.41	0.05	0.36	0.13	0.41	0.10	0.40
			(0.15)		(0.16)		(0.17)		(0.16)		(0.16)

4.3.2 Extended Model with Macro-Economic Factors

Again in this section not all results will be discussed very detailed. Only differences in respect to the model without Macro-Economic Factors will be handled. As I also will show in this section, the problems with stationarity for the extended model became were also present in some degree just as with the simulation study. However due to the use of the tuning parameters the implications were far less severe for this section.

In table 18 results are again given of the persistence of the extended model with Macro-Economic Factors. These show that indeed the persistence in the risk-neutral measure has increased very highly, even to the point where some are basically equal to 1. For the M2 model specification it even ensured that one eigenvalue always stayed at exactly 1. One would maybe expect those changes to happen in the yields for the other measure because only that measure is directly affected by the Macro-Economic Factors (see equation (25)). This table shows that those effects indeed affect the structure of the estimates in the risk-neutral measure. In the results for the volatilities of model implied forward rates not much changes can be seen. Except for the fact that we already can see the effects of stationarity of the second model specification. Because the persistence is equal to one for that specification the implied rates tend to vary much more which seems problematic and not a good estimate of the volatility.

Table 18: This table contains details for the autoregressive component of the term structure model with the Macro-Economic Factors added, but only the maximum eigenvalues are shown. Where the first value represents the posterior mean and the value between parentheses the posterior median. The different models depicted by M1, M2 and M3 correspond with the three best models following table 10. The M0 model stands for the unrestricted model so where all elements of γ are set to 1. Also the volatilities of changes in implied 5-10 forward rates, changes in risk-neutral forward rates and risk premiums are given in the 3 columns to the right. Below the posterior means and medians, 95% confidence intervals are given in the square brackets below.

Model	Max. eigenvalue		Volatilities		
	$\mathbb{Q}$	$\mathbb{P}$	$\Delta\hat{f}_t$	$\Delta\tilde{f}_t$	Δftp_t
M0	0.9831 (0.9831)	0.9452 (0.9465)	0.23 (0.23)	0.02 (0.00)	0.23 (0.23)
	[0.9829, 0.9835]	[0.8937, 0.9927]	[0.22, 0.24]	[0.00, 0.12]	[0.19, 0.25]
M1	0.9816 (0.9816)	0.9658 (0.9657)	0.22 (0.22)	0.01 (0.01)	0.20 (0.21)
	[0.9815, 0.9818]	[0.9516, 0.9789]	[0.21, 0.22]	[0.00, 0.04]	[0.19, 0.21]
M2	0.9824 (0.9823)	0.9642 (0.9641)	0.22 (0.22)	0.01 (0.01)	0.21 (0.21)
	[0.9822, 0.9827]	[0.9515, 0.9774]	[0.21, 0.22]	[0.00, 0.03]	[0.19, 0.21]
M3	0.9860 (0.9859)	0.9708 (0.9709)	0.22 (0.22)	0.02 (0.02)	0.20 (0.20)
	[0.9846, 0.9873]	[0.9597, 0.9815]	[0.22, 0.22]	[0.01, 0.05]	[0.18, 0.21]
BMA	0.9830 (0.9830)	0.9773 (0.9744)	0.22 (0.22)	0.07 (0.03)	0.20 (0.20)
	[0.9830, 0.9832]	[0.9528, 0.9993]	[0.21, 0.23]	[0.01, 0.25]	[0.13, 0.27]

Not much can be said about the effects of adding the Macro-Economic Factors by looking at the figures in figure 7, and this is also the case when implied yields with a maturity of 5 years and 6 months are regarded. Therefore only the figure for the fitted 10-year yields is given in this section.

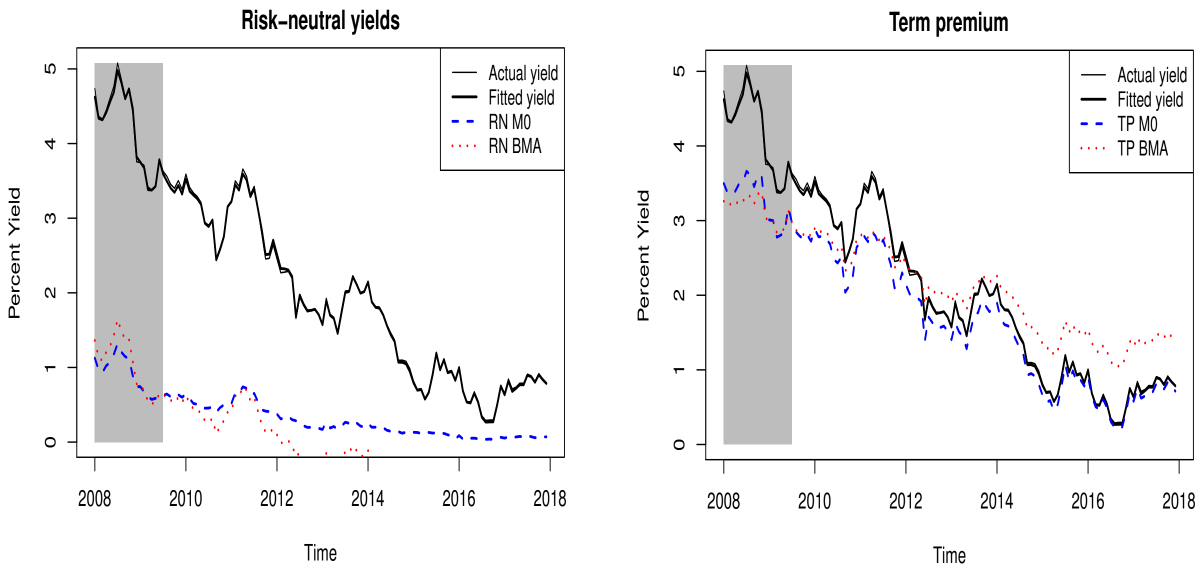

Figure 7: Figure of the plotted term premiums over the used time period with a maturity of 10 years. The plotted actual yields are not that much visible but that is caused by the fact that plotted fitted yields are rather close to the actual ones. The fitted yields are obtained by the unrestricted model.

Next to the implied yields in figure 7 also the posterior results of the volatilities are plotted in figure 8. The first thing that can be noticed from this figure is that the confidence intervals of the implied volatility for the unrestricted model is much lower than for the BMA model. Only for short-term yields the BMA sampler is able to produce estimates that are not too volatile. Quickly for maturities of longer than 1 year the interval widens greatly. The unrestricted model is better able at producing stable results for the volatilities of the model. This relation stays in place even for long-term rates. Apparently, the Macro-Economic Factors give the model some extra room to let the estimates be more volatile instead of the stable results in the base model. This effect becomes even stronger when also a model selection sampler is applied which can more precisely denote the relations between the two different meausures in the model.

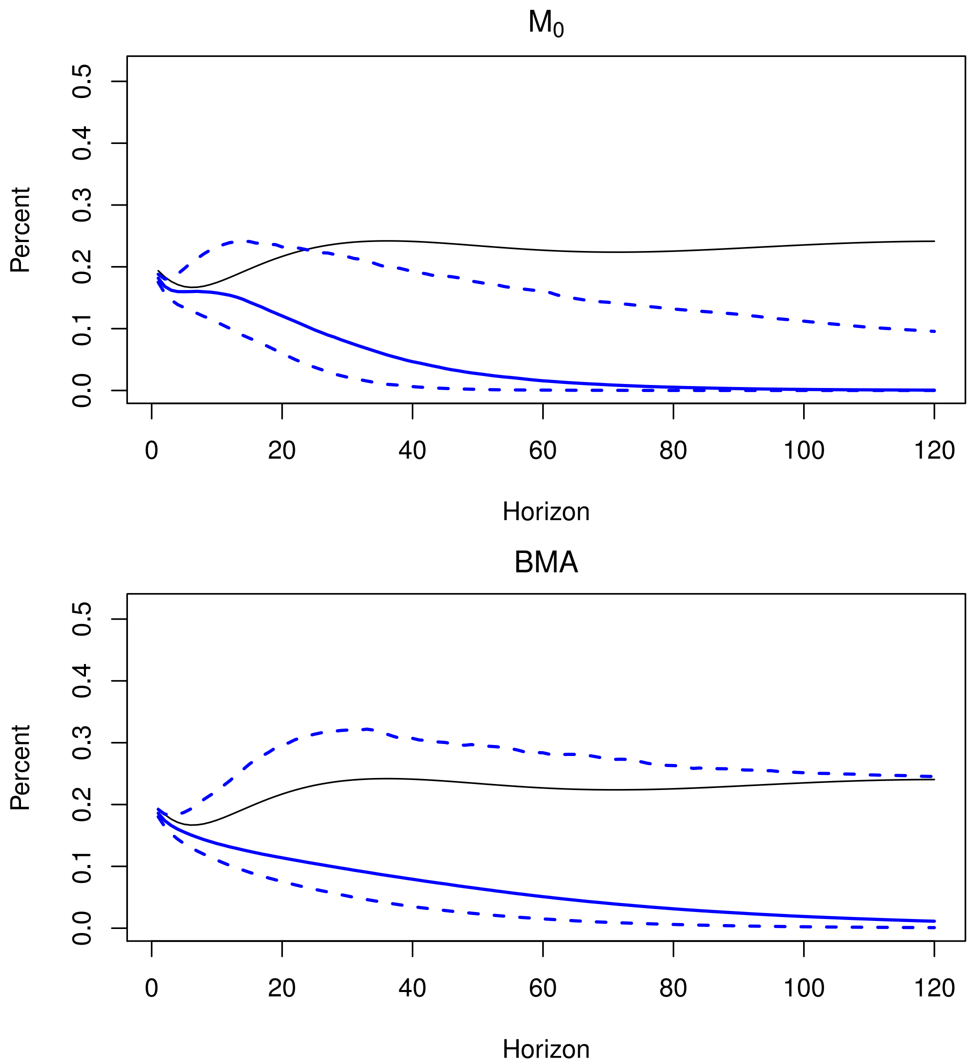

Figure 8: Figures of the posterior means of the volatilities in changes over the fitted forward rates. Also risk-neutral forward rates are plotted, together with their credibility intervals in squared brackets.

In table 19 again results on the predictabilities of the different models are given. If this table is compared with table 17 it is shown that the amount of variation in excess bond returns that is declared is slightly higher when the Macro-Economic Factors are included. This is as expected because more regressors are now used (5 instead of 3) but it also shows a more stable pattern across the different specifications and for the different maturities. For example the R^2 for the unrestricted model is equal to 0.24 for the maturities of 3,5 and 7 years and equal to 0.26 for a maturity of 10 years. Due to the earlier described problems with simulating data from the extended model with Macro-Economic Factors the R^2s are flawed. More specifically for the $M1$ specification the model is fully able to declare all variation in the excess bond returns. For the other models these percentages are at more normal levels but then again they are retrieved from fitting the model on simulated data that proved to have several problems.

Table 19: This table contains the predictability for some selected maturities. These predictabilities are denoted by the R^2 implied by the different models in declaring the variance in bond returns. The *Data* column represents the percentage of variation in bond returns that the observed yield data can declare. The *Pop.* columns represent the odds for the used sample period and the *Smpl.* columns are generated by simulating yields. For these simulated returns also their standard errors are given in brackets. Implications of the problems of overfitting that are described earlier are also visible in this table. Therefore the R^2 with simulated yields is very high for the $M1$ model.

		M0		M1		M2		M3		BMA	
Maturity	Data	Pop.	Smpl.	Pop.	Smpl.	Pop.	Smpl.	Pop.	Smpl.	Pop.	Smpl.
3 years	0.15	0.24	0.45	0.28	1.00	0.28	0.90	0.25	0.47	0.20	0.54
			(0.15)		(0.00)		(0.04)		(0.16)		(0.17)
5 years	0.17	0.24	0.45	0.23	1.00	0.24	0.88	0.21	0.45	0.16	0.52
			(0.15)		(0.00)		(0.05)		(0.15)		(0.17)
7 years	0.21	0.24	0.45	0.19	1.00	0.20	0.87	0.17	0.43	0.13	0.50
			(0.15)		(0.00)		(0.05)		(0.15)		(0.17)
10 years	0.25	0.26	0.45	0.17	1.00	0.18	0.86	0.14	0.41	0.12	0.49
			(0.15)		(0.00)		(0.06)		(0.15)		(0.18)

4.4 Extreme Scenarios

As a last part of the results, the scenarios that are generated by the algorithm described in section 3.4 will be discussed. In this section I will discuss the results obtained by making forecasts for a year ahead of the yield curve. Those are plotted in figure 9. For each different maturity a distribution has been made and from the 7 different quantiles the resulting yield curve. An important observation that I draw from the figure is that of the effect of the Macro-Economic Factors on the yield curve. These factors are able to change the shape of the yield curve as a whole. This even leads to a slightly inverse relation at the short-end of the yield curve which is a rather strange phenomenon but became less unlikely in recent periods. It should be noted however, that in the used dataset this pattern was not visible yet and that it is driven by the Macro-Economic Factors.

The base model without Macro-Economic Factors is not able to change the shape of the yield curve as a whole, largely because it is only dependant on historic data. This can cause quite a problem because the model can only then predict a difference in the structure of the yield curve when those changes have actually already been happened. Obviously it will always predict a change in structure too late by default. When the model also uses other variables such as those factors it can better predict those changes in the structure as I have shown in figure 9. This stresses the importance of using also other kind of variables instead of only using historic yield data when models are needed for the purpose of predicting forecasts or establishing extreme scenarios. In the figure also observations for the yield curve are given of the first of June 2018, because they illustrate how the yield curve develops in 2018. It shows that the yield curve is very persistent and does not abbreviate much from the latest observation in 2017.

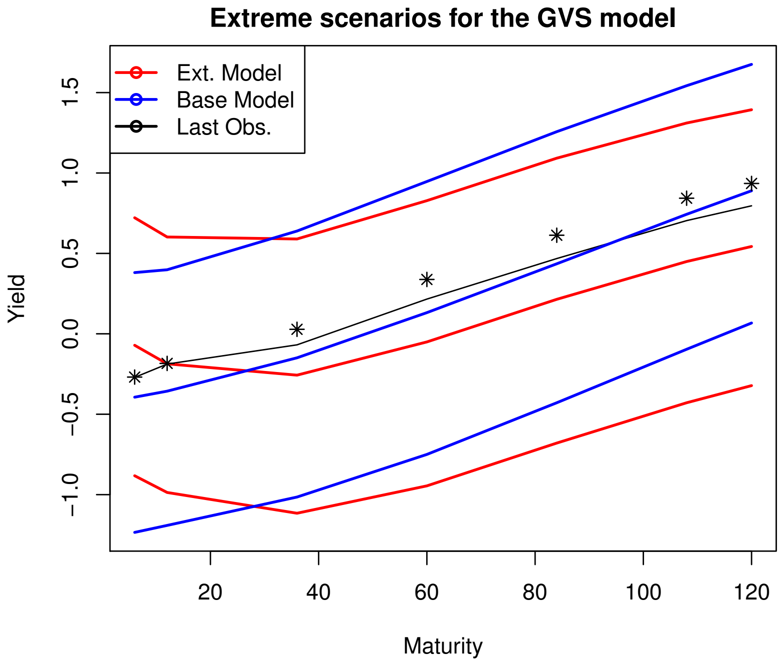

Figure 9: This figure shows the extreme scenarios for the GVS sampler, the quantiles that were used are 0.9, 0.5 and 0.1. The highest pair of lines corresponds with the highest quantile. The blue lines depict the scenarios for the base model and the red lines for the scenarios resulting from the models that also included the Macro-Economic Factors. Also the last observation of the yield curve is plotted as the black line. The asterisks that are plotted are the latest available observations namely that of the first of June 2018. These are not used in the sample of the model but are given to illustrate how the yield curve develops on wards from the latest observation in 2017.

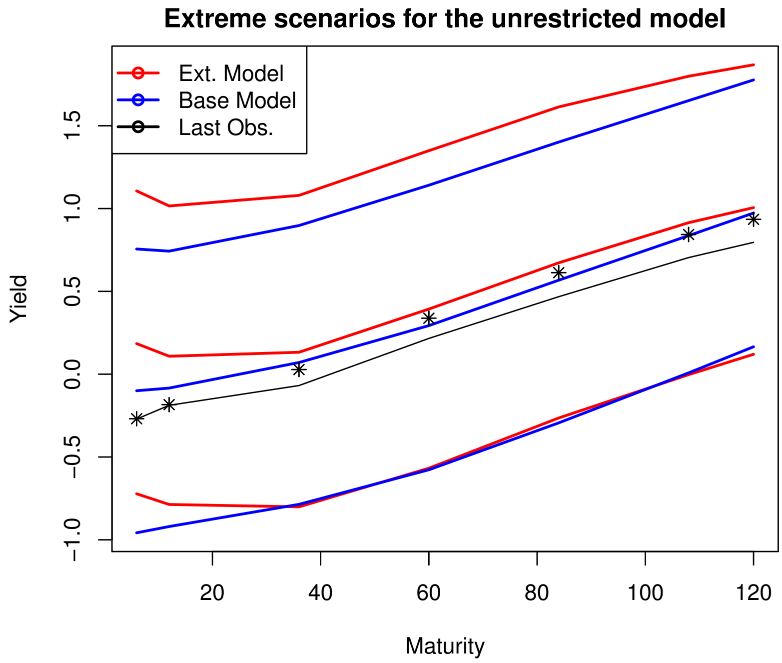

Figure 10: Figure showing the extreme scenarios for the GVS sampler, the quantiles that were used are 0.9, 0.5 and 0.1. The blue lines depict the scenarios for the base model and the red lines for the scenarios resulting from the models that also included the Macro-Economic Factors. Also the last observation of the yield curve is plotted as the black line. The asterisks that are plotted are the latest available observations namely that of the first of june 2018. These are not used in the sample of the model but are given to illustrate how the yield curve develops onwards from the latest observation in 2017.

In figures 11 until 13 quantiles are plotted of separate forecasts for yields of some maturities. These include 6 months, 5 years and 10 years to compare the same data as in section 4.3. Forecasts that are made with the GVS sampler are plotted next to forecasts implied by the unrestricted model. The unrestricted model obviously uses more parameters and therefore it seems to overfit the predictions to the fact that the adding of Macro-Economic Factors has no clear effect anylonger. Only for the more volatile maturity, namely that of 6 months, the effects of these factors are still visible.

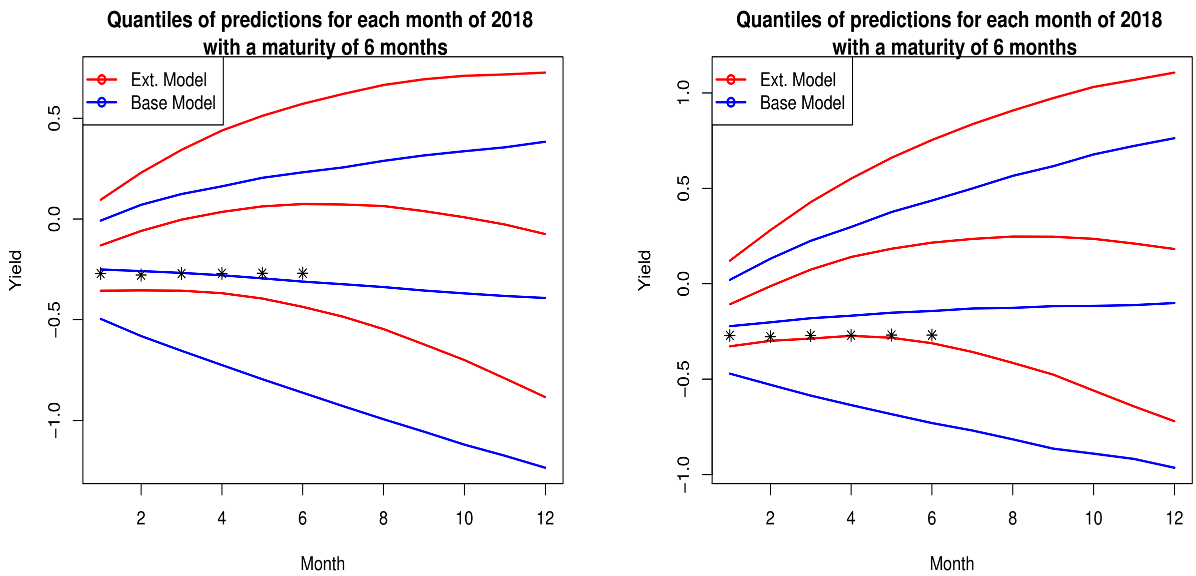

Figure 11: These figures have plotted the quantiles of individual forecasts for yields per month as out-of-sample forecasts with a maturity of 6 months. The used quantiles are 0.1, 0.5 and 0.9. The left panel gives the forecasts for the GVS sampler while the right panel is constructed with the unrestricted model. The blue lines depict the basis model and the red lines depict the extended model with Macro-Economic Factors. Also the observations for the first 6 months are given with asterisks used as symbols.

Figure 11 shows that the unrestricted model has a higher trend throughout the predictions, because the forecasts increase more for each quantile when out-of-sample period increases. For both different methods the adding of the Macro-Economic Factors leads to higher predictions of the 6 month yield. Also the predictions tend to follow a non-linear trend in contrary to the base model that only uses historic yield data. However, when looking at the actual observations of 2018 it seems that the base model is better able to predict the short term yields, when the 50^{th} percentile is used. This pattern is mosltly due to the fact that the actual data is very persistent and does not change much which is also reflected in line for the 50^{th} quantile of the base model.

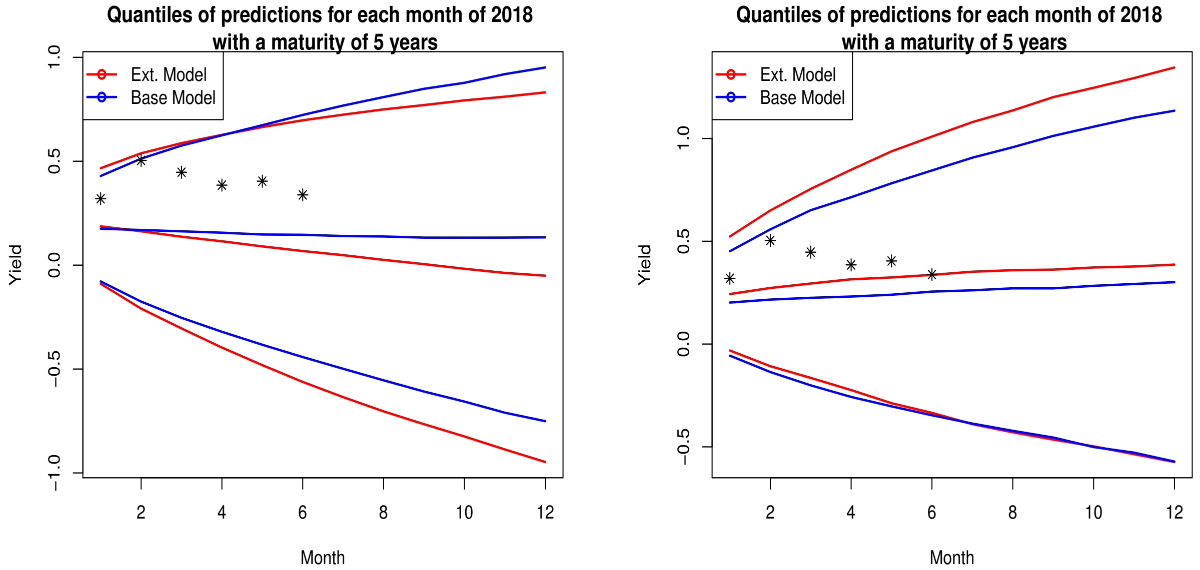

Figure 12: These figures have plotted the quantiles of individual forecasts for yields per month as out-of-sample forecasts with a maturity of 5 years. The used quantiles are 0.1, 0.5 and 0.9. The left panel gives the forecasts for the GVS sampler while the right panel is constructed with the unrestricted model. The blue lines depict the basis model and the red lines depict the extended model with Macro-Economic Factors. Also the observations for the first 6 months are given with asterisks used as symbols.

In figure 12 the individual forecasts of yields with a maturity of 5 years are plotted. For the unrestricted model the adding of the Macro-Economic Factors does not have a large effect anymore, which is due to the overfitting that has occurred. The unrestricted model without Macro-Economic Factors already uses 26 different parameters. For the GVS sampler however the adding of those factors still does have an impact. It diminishes the level of all three quantiles and this difference becomes larger when a prediction has to be made further away in the future. Also the pattern of the actual observations in 2018 is different than the one shown in figure 11. According to the models the more recent observations all fall below the 0.9 quantile, which makes them not much extreme.

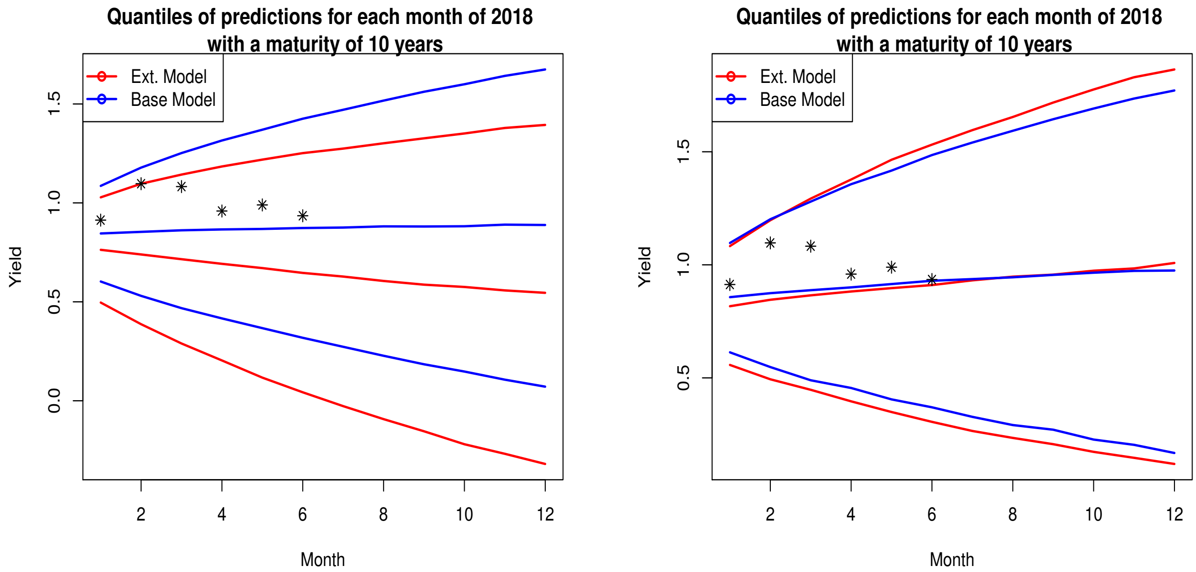

Figure 13: These figures have plotted the averages of individual forecasts for yields per month as out-of-sample forecasts with a maturity of 10 years. The used quantiles are 0.1, 0.5 and 0.9. The left panel gives the forecasts for the GVS sampler while the right panel is constructed with the unrestricted model. The blue lines depict the basis model and the red lines depict the extended model with Macro-Economic Factors. Also the observations for the first 6 months are given with asterisks used as symbols.

Figure 13 does not have many differences with figure 12 except for the fact that the effect of adding Macro-Economic Factors in combination with the GVS sampler is strengthened. Therefore the difference in for example the 0.1 quantile between the two models is around 0.8 percentage point. Again the more recent observations of yields with a maturity of 10 years fall below the 0.9 quantile of both models.

4.4.1 Parallel Scenarios

In this section I show some results of the different models for parallel extreme scenarios. These scenarios are not realistic to actually happen because the same shock is applied to the entire yield curve. However, this approach of establishing extreme scenarios is widely used by both regulators and private parties such as banks. Therefore, I show the outcomes of these kind of scenarios implied by the proposed models, of equations 1a and 25, in this paper. In the figures, discussed in this section, also scenarios are plotted that are generated by estimating a simple AR(1) model with the data. For the quantiles a Normal distribution is assumed with parameters $\mu = 12 * \mu_{sample}$ and $\sigma^2 = \sum_{i=1}^{12} \sum_{j=1}^{12} \hat{\phi}^{|i-j|} * \sigma^2_{sample}$. The two sample parameters are estimated from the used dataset and the ϕ parameter is estimated from the AR(1) model. For a parallel shift of the yield curve a method needs to be used when determining an equal shock for all maturities. I chose for an approach by determining the shock for just one certain maturity and then applying that shock to the entire curve. Other methods such as using averages of non-parallel shocks can also be used. For both a maturity of 6 months and 10 years scenarios are generated in this section.

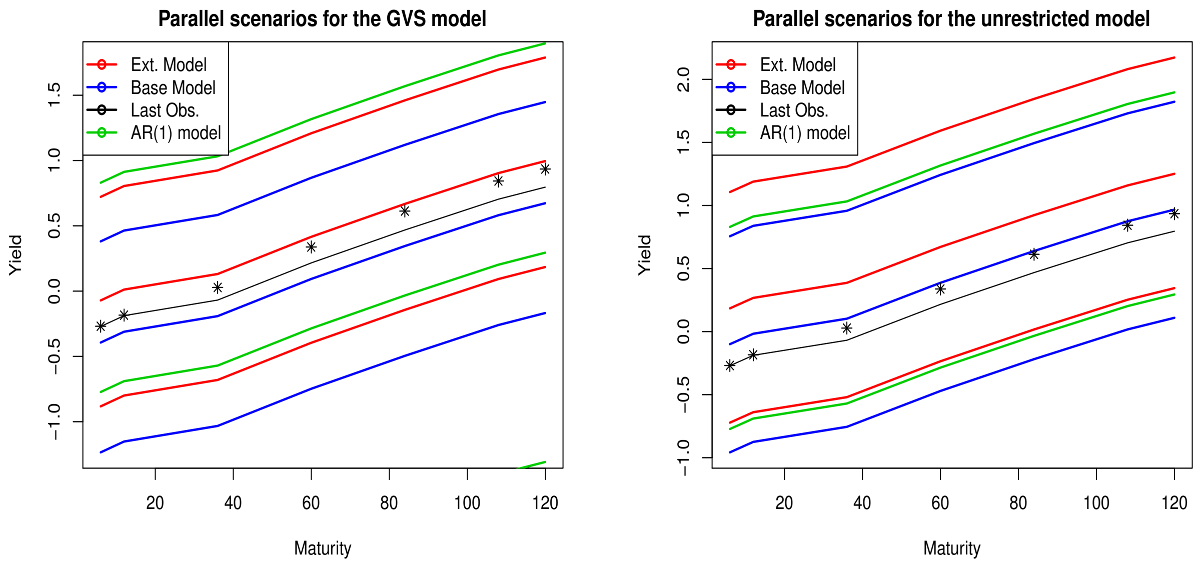

Figure 14: This figure shows the parallel extreme scenarios for both the GVS sampler and the unrestricted model. The shock is determined by looking at the quantiles of the forecasts for yields with a maturity of 6 months. The used quantiles are 0.1, 0.5 and 0.9 where the highest set of lines corresponds with the highest quantile. The blue lines again stand for the basis model and the red lines for the model with Macro-Economic Factors. Next to that, the asterisks depict the actual observations of the yield curve on the first of June 2018. In both panels the green lines depict the scenarios that are generated by the AR(1) method described in this section. For this method the outcomes varied much and therefore the lines for the 0.1 quantile are not visible.

In figure 14 scenarios are plotted that are generated by determining the shock for the 6 month maturity and then applying that shock to the entire curve. The adding of the Macro-Economic Factors contributes to a higher level of the curve for the different quantiles for both the GVS sampler as for the unrestricted model. Also the figure shows that when only historical data is used to model the short rate a downwards trend in the yield curve is to be expected. The adding of Macro-Economic Factors helps to fight off this trend and to actually model a upwards movement in the yield curve. The AR(1) model tends to model scenarios at a whole different level than the term structure models. Those implied scenarios look less realistic, as for example the 0.5 quantile scenario is way off the last observation and also the observation of June 2018. Also the width of the different quantiles is much higher which also contributes to the lower performance of the more basic method.

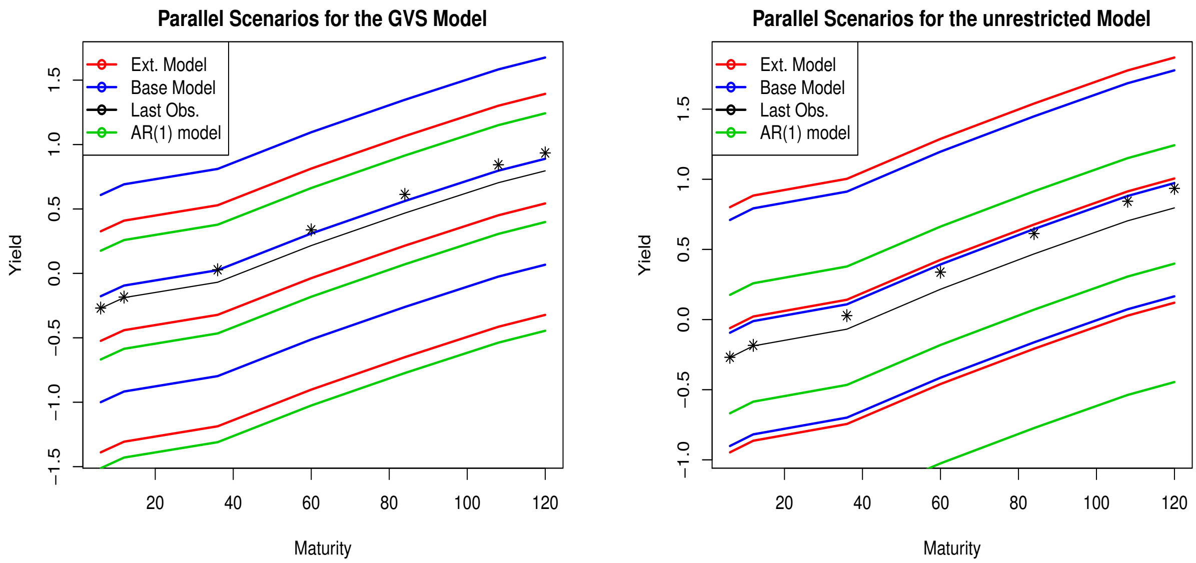

Figure 15: This figure shows the parallel extreme scenarios for both the GVS sampler and the unrestricted model. The shock is determined by looking at the quantiles of the forecasts for yields with a maturity of 10 years. The used quantiles are 0.1, 0.5 and 0.9 where the highest set of lines corresponds with the highest quantile. The blue lines again stand for the basis model and the red lines for the model with Macro-Economic Factors. Next to that, the asterisks depict the actual observations of the yield curve on the first of June 2018. In both panels the green lines depict the scenarios that are generated by the AR(1) method described in this section.

In figure 15 the parallel scenarios are shown that are determined by forecasting yields with a maturity of 10 years. It shows the opposite effect when compared with figure 14, namely that the scenarios for the base model are at higher level for the GVS sampler. So when a longer maturity is used to obtain a parallel shock the Macro-Economic Factors actually lower the level of the scenarios. However, for the unrestricted model the levels of both term structure models are quite in vicinity of each other. Again, the AR(1) model predicts the scenarios at a lower level than the other models.

5 Conclusion and Discussion

In this paper I showed that the Bayesian framework of estimating a Gaussian Dynamic Term Structure Model produced stable results. The approach performed well in a simulation study by being able to mostly correctly retrieve the parameters on which the data was simulated. Furthermore, different model selection samplers were used and compared to each other. The simulation study showed that the three different methods all improved the performance of retrieving the data generating process parameters when compared to an unrestricted model where all risk price parameters were taken into account. There were only small differences between the methods that did not give enough evidence to conclude on which sampler method performs best.

When Macro-Economic Factors are added to the model the simulation study shows a different conclusion. Due to the extended parameter space, the different drawing algorithms are not able to produce stable results anymore. For each different simulated dataset a different set of tuning parameters is needed for making the draws of the different parameters. Because these parameters should be created manually, in a way that these fit the dataset, a simulation study is not able to generate stable results for the extended model.

After the simulation study, both models are evaluated with a dataset on Euribor yield data. Section 4.2 shows that the estimated model without Macro-Economic Factors already asks for a reasonable amount of restrictions on risk prices. This leads to a model that is able to estimate the structure of the yield data rather well. Moreover, the results also uncover the differences between the different model selection samplers. I explained the differences by looking at the posterior results of the γ parameter, which is an indicator variable for λ. The GVS and RJMCMC samplers both generated consistent results for the elements of γ. The SSVS sampler was not able to perform as consistently as the other samplers and this resulted in posterior estimates of γ that were overall inefficient for all elements. Furthermore, the different samplers were evaluated by looking at how many models they did evaluate and how well those particular models are in terms of Bayes factors. It showed why the SSVS sampler is performing worse than the other samplers, because it only evaluated a very small amount of different model specifications. The top ten best performing models for the GVS algorithm also show that there is a consistent pattern in the different specifications. Due to delivering the most consistent performance, the GVS sampler is used in other sections as well to benchmark its performance against the three best model specifications.

The adding of the Macro-Economic Factors did not only contribute to the amount of parameters, but it also led to significant parameters of the effects on the risk prices of those economic variables. However, the estimated model became less stable by adding the two factors. This resulted in extreme low fractions of different model specifications that were evaluated. Also, between the top candidate specifications, in terms of their frequency being visited by the algorithm, the differences disappeared and were almost performing as good as each other. Additionally, the model also became even more persistent because the eigenvalues of the autoregressive component were even closer to 1 than before. Inevitably this also led to higher volatilities and more fluctuating in-sample forecasts.

When looking at economic implications the results showed that the problems of stability for the extended model did not automatically lead to weaker forecasts of this model when compared to the model without Macro-Economic Factors. The differences were rather small between the two different models and the forecast performance slightly improved by adding the Macro-Economic Factors. The extended model also used a more unrestricted specification which led to a lower volatility in the fitted forward rates. Also the extended model became less stationary with maximum eigenvalues of the autoregressive component that were closer to 1.

In the last results section the different models were used to generate extreme scenarios of the yield curve with a forecast horizon of 12 months. This application showed the main difference when using Macro-Economic Factors instead of only historic yield data. When the model only used yield data, it was only able to establish scenarios that led to curves that still had the same shape as the last observed curve. However, for the extended model the extreme scenarios indicated a large increase in the short end of the yield curve. The quantiles for the part of the yield curve with longer maturities were more similar for both models and therefore the shape of the implied scenarios changed. Furthermore, because of the characteristics of the historic data on Euribor yields, any model that solely uses these yields is not able to predict any changes in patterns of the data. When other variables, such as inflation or economic growth, are used to model the term structure, changing patterns could be predicted and modelled. The extended model in this paper showed that yield factors are dependant on factors of these variables and that they contribute in predicting a changing shape of the yield curve.

A main limitation of the research, presented in this paper, was that of the stability of the model that also included Macro-Economic Factors. It led to a Bayesian estimation that was not stable, so that tuning parameters, that were not needed for other datasets, were needed for an algorithm (Bauer, 2017; Chib and Ergashev, 2009). Further research is therefore needed in which the drawing procedure can be set up in a way that for several different datasets no tuning parameters are needed. If such manner could be established, then the proposed method of estimating a term structure could lead to suitable acceptance rates when making draws of the parameter distribution. Also, the use of more sophisticated computer systems could attribute to this, as I found that due to the expanded parameter space the estimation of several matrix computations became less accurate. Additional, the scope of this paper did not focus on the use of priors for the different parameters of the term structure model. Possible a more extensive use of priors can help to address some om the problems that were present in this paper. Especially for the Macro-Economic Factors a more specified set of priors could be used such as democratic priors proposed by Wright (2013).

References

Aiolfi, M., Capistran, C., and Timmermann, A. G. (2010). Forecast combinations.

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.

Bates, J. M. and Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468.

Bauer, M. D. (2017). Restrictions on risk prices in dynamic term structure models. *Journal of Business & Economic Statistics*, pages 1–16.

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

Bauer, M. D. and Rudebusch, G. D. (2016b). Resolving the spanning puzzle in macro-finance term structure models. *Review of Finance*, 21(2):511–553.

Carlin, B. P. and Chib, S. (1995). Bayesian model choice via markov chain monte carlo methods. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 473–484.

Chib, S. and Ergashev, B. (2009). Analysis of multifactor affine yield curve models. *Journal of the American Statistical Association*, 104(488):1324–1337.

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.

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

Dai, Q. and Singleton, K. J. (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of financial Economics*, 63(3):415–441.

De Pooter, M., Ravazzolo, F., and Van Dijk, D. J. (2010). Term structure forecasting using macro factors and forecast combination.

Dellaportas, P. and Forster, J. J. (1999). Markov chain monte carlo model determination for hierarchical and graphical log-linear models. *Biometrika*, 86(3):615–633.

Dellaportas, P., Forster, J. J., and Ntzoufras, I. (2002). On bayesian model and variable selection using mcmc. *Statistics and Computing*, 12(1):27–36.

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

Diebold, F. X. and Pauly, P. (1987). Structural change and the combination of forecasts. *Journal of Forecasting*, 6(1):21–40.

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

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

Elerian, O., Chib, S., and Shephard, N. (2001). Likelihood inference for discretely observed nonlinear diffusions. *Econometrica*, 69(4):959–993.

Eraker, B. (2001). Mcmc analysis of diffusion models with application to finance. *Journal of Business & Economic Statistics*, 19(2):177–191.

Gelman, A., Roberts, G. O., Gilks, W. R., et al. (1996). Efficient metropolis jumping rules. *Bayesian statistics*, 5(599-608):42.

George, E. I. and McCulloch, R. E. (1993). Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88(423):881–889.

Green, P. J. (1995). Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732.

Hamilton, J. D. and Wu, J. C. (2012). The effectiveness of alternative monetary policy tools in a zero lower bound environment. *Journal of Money, Credit and Banking*, 44(s1):3–46.

Jones, C. S. (1998). Bayesian estimation of continuous-time finance models. *manuscript University of Rochester*.

Jones, C. S. (2003). Nonlinear mean reversion in the short-term interest rate. *The review of financial studies*, 16(3):793–843.

Joslin, S., Priebsch, M., and Singleton, K. J. (2014). Risk premiums in dynamic term structure models with unspanned macro risks. *The Journal of Finance*, 69(3):1197–1233.

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.

Kim, D. H. and Singleton, K. J. (2012). Term structure models and the zero bound: an empirical investigation of japanese yields. *Journal of Econometrics*, 170(1):32–49.

Koopman, S. J. and van der Wel, M. (2013). Forecasting the us term structure of interest rates using a macroeconomic smooth dynamic factor model. *International Journal of Forecasting*, 29(4):676–694.

Laurini, M. P. and Hotta, L. K. (2010). Bayesian extensions to diebold-li term structure model. *International Review of Financial Analysis*, 19(5):342–350.

Monfort, A., Pegoraro, F., Renne, J.-P., and Roussellet, G. (2017). Staying at zero with affine processes: an application to term structure modelling. *Journal of Econometrics*, 201(2):348–366.

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

Sanford, A. D. and Martin, G. M. (2005). Simulation-based bayesian estimation of an affine term structure model. *Computational statistics & data analysis*, 49(2):527–554.

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

Wright, J. H. (2013). Evaluating real-time var forecasts with an informative democratic prior. *Journal of Applied Econometrics*, 28(5):762–776.