Erasmus University Rotterdam

Erasmus School of Economics

Global Yield Curve Arbitrage with Term Structure Modeling

Jiawei Xu

380392

University supervisor:

Dick van Dijk

External supervisor:

Casper Zomerdijk

Co-reader:

Michael Gong

A thesis submitted in partial fulfillment of the requirements for the degree of

Master in Econometrics and Management Science

Date final version: October 17, 2017

Abstract

This study investigates the profitability of trading strategies based on yield curve arbitrage opportunities in international fixed-income markets. For constructing the yield curve, this paper adopts three- and four-factor Nelson-Siegel models, cubic B-splines, and kernel smoothing methods as candidate models, and evaluates the model fit by their ability of identifying the mispricing of the government bonds with a generic yield curve arbitrage strategy that longs all the over-priced bonds and shorts all the under-priced bonds. Next the paper continues with the most feasible yield curve models selected and constructs a modified yield curve arbitrage strategy with improved strategy specifications that maximize exposure to the arbitrage opportunities while limiting effect of other risk factors. The results show that there is no superior yield curve model globally when the pricing power of the yield curve is examined but the four-factor Nelson-Siegel models are preferred by more countries. A yield curve arbitrage strategy is profitable for all countries and is able to generate a positive risk premium. A more refined strategy specification is able to further enhance the performance by reducing volatilities and eliminating exposures to other risk factors.

Keywords: Fixed-income arbitrage, government bonds, yield curve arbitrage, term structure modeling, trading strategy

Contents

1 Introduction

Fixed-income arbitrage, as the name suggests, exploits the misvaluations during the pricing processes of fixed-income securities, and has become one of the most popular choices for hedge funds and arbitrageurs since the hedge fund crisis of 1998. Within the fixed-income arbitrage sector, yield curve arbitrage is one of the well-known arbitrage strategies that show good performance and profitability (see Duarte et al., 2006). Essentially, a yield curve arbitrage strategy identifies under- and over-priced government bonds along the yield curve and sets up portfolios accordingly. Duarte et al. (2006) conclude that the better performance of the yield curve arbitrage strategy boils down to the fact that this strategy requires most "intellectual capital" as it involves modeling the yield curve. Thus for a successful strategy implementation, modeling the term structure is as important as designing the strategy specifications.

This research follows the path of the yield curve arbitrage literature and aims to go one step further by seeking for a trading strategy specification that harvests the potential mispricings of the government bonds, and exploring the profitability and significance of yield curve arbitrage as an anomaly. Along the way, this study also investigates different term structure models and intends to identify the most feasible ones among them.

For this study, the price information of the government bonds are observed at the end of the month, and the yield curve is then modeled for setting up the positions of the arbitrage portfolio at the beginning of the next month. As neither forecasting the term structure nor the dynamics of the term structure over time is of interest, affine term structure models that describe the stylized time-series properties of the term structure (see, for example, Vasicek, 1977; Cox et al., 1985) are not considered for this research, and only the curve fitting category of the term structure models are ideal candidates for modeling the yield curves. Term structure modeling with curve fitting methods has been studied intensively by fixed-income literature and great progress has been made over the last decades. Nelson and Siegel (1987) propose to fit the term structure with a parametric function of three exponential components that are capable of following many of the typically observed shapes that the yield curve assumes. Since then various extensions based on this framework have been explored to increase the flexibility. Björk and Christensen (1999) and Svensson (1994) incorporate one extra factor to capture the yield movements in the short and medium maturities, respectively, and both their results show improvements in terms of yield curve fitting. Although with no restrictions regarding arbitrages imposed, the

Nelson-Siegel model and its extensions have been the workhorse of estimating the yield curve in financial practice due to their parsimony and good empirical performance. The Bank of International Settlements (BIS, 2005) reports that among the thirteen central banks that report their curve estimation methods to the BIS, nine use either the Nelson-Siegel or the Svensson model to construct zero-coupon yield curves. On the other hand, McCulloch (1971) estimates the discount factors using regression splines, making the spline principles a popular approach. Following this path Chambers et al. (1984) fit the yield curve itself with polynomial splines, and Vasicek and Fong (1982) show that exponential splines provide better yield curve estimations compared to polynomial splines. Linton et al. (2001) propose to use kernel smoothing methods for estimating the term structure and conclude that the estimations generated by their methods have sensible interpretations even under misspecification. Sediva and Marek (2015) show that yield curves estimated by kernel smoothing are comparable to those generated by parametric methods.

This study adopts the Nelson-Siegel models and its extensions, cubic regression splines, and kernel smoothing for fitting the term structure. In order to overcome the numerical difficulties the Nelson-Siegel type of models suffer from due to the existence of non-linearity (see, for example, Gimeno and Nave, 2006) and achieve a robust fit, the estimation of the Nelson-Siegel model and its extensions is conducted using Bayesian inferences with Markov chain Monte Carlo (MCMC) simulation techniques.

The goodness-of-fit of the yield curves generated by different methods is afterwards evaluated to select a most feasible yield curve model. Followed by most existing literature, the standard approach is to evaluate the models by means of their statistical properties (see, for example, Bliss, 1996; de Pooter, 2007) such as the (adjusted) R-squared or the pricing (yield) errors. Nevertheless, these measures are from a curve fitting point of view. For the purpose of evaluating term structure models that a trading strategy will be based on, one aspect that is probably even more relevant yet often neglected is the economic values behind the modeled curves. For example, any line that connects or goes through all the data points may give satisfactory statistics as it indeed produces no residuals, but it is not very likely a trading strategy or monetary policy based on this yield curve model will achieve convincing results. This paper proposes to evaluate the yield curve models with a yield curve arbitrage strategy that delves into how well the modeled curves identify the mispricings in the bonds. The intuition here is that if the fitted yield curve by a method is superior, that is, it captures the underlying “true” term structure precisely over time, an arbitrage-based trading strategy that goes long in the bonds that are undervalued by the market and short in the ones overvalued should be profitable. Following this approach, a generic yield curve arbitrage strategy is formed to construct portfolios that longs all the underpriced bonds and shorts the over-valued ones along the modeled yield curve, and the weight of each bond in the portfolio is decided by the corresponding yield error. To avoid duration bias, the weights of the bond are also scaled according to the corresponding durations. and the most feasible term structure modeling method is then selected according to the performance of this yield curve arbitrage strategy, with the criteria involving both the returns and the risks over the full sample period and the subsamples. As a byproduct, such a fixed-income arbitrage strategy can serve as a foundation of finding a best yield curve arbitrage strategy that further exploits and benefits from the mispricings of the bonds.

The results of the generic strategy show that the Nelson-Siegel models are preferred by most countries as a yield curve model, and within the Nelson-Siegel family, the four-factor models are with advantage. On top of providing a yield curve model selection, the results also give insight on some characteristics of the strategy performance. Firstly, such an arbitrage-based strategy is profitable in all countries. Moreover, the profit can be mostly attributed to periods when the market is with more turbulence or during crisis. Hu et al. (2013) conclude that the magnitude of the government bond mispricing is larger during crisis for the U.S., as the arbitrage strategy is based on mispricings, this result is reasonable. To further understand the drives of arbitrage strategy returns, a market portfolio and two portfolios based on the yield curve slope and curvature changes are formed as common risk factors and the arbitrage strategy returns' exposure to these factors are examined. The risk-adjusted performance show that although being exposed to the risk factors, especially to the yield curve positions, the arbitrage strategy is able to generate a positive alpha in all countries, and this risk premium based on arbitrage opportunities is significant for most countries. Finally, the Nelson-Siegel-model-based strategy shows more profitability in most countries in general, but it is also observable that the strategy with more flexible non-parametric methods (B-splines and kernel smoothing) has an advantage when the market is volatile and with much movements.

With the selected modeled yield curve, this paper continues studying yield curve arbitrage by looking for improvements in the trading strategy settings. The yield curve arbitrage strategy for evaluating the yield curve fitting method is a generic starting point that considers only maximizing the exposure to the arbitrage opportunities along the yield curve. The results of the risk-adjusted performance of the generic strategy, however, suggest that positioning neutral to the risk factors can enhance the performance such that the risk profile of the strategy can be improved. Moreover, lowered exposure to risk factors that influence the bond returns can also make the strategy itself more practically applicable. To improve upon the generic strategy, this research considers the setting of Beekhuizen et al. (2016) and constructs a modified strategy that groups the available bonds into maturity buckets of 1-3 year, 3-5 year, 5-7 year, 7-10 year, 10-15 year and 15+ year according their corresponding remaining time-to-maturities. For a maturity-bucket-based strategy specification, as each bond within the maturity bucket has similar maturity, the long- and short- positions are better hedged so the strategy is less exposed to market movements. Also, within a maturity bucket, the slope and curvature of the yield curve is much reduced such that the impact of the yield curve change is less influential. Similar sub-portfolios as the generic strategy are then constructed within the maturity buckets: each sub-portfolio has one duration of long position and one duration of short position, and the weight of each bond is decided by the yield error. Finally the positions of bonds are scaled by the number of buckets entering trade such the total position of the whole portfolio is one duration long and one duration short. The results show that this modified strategy is able to achieve at least as good profitability as the generic strategy with much lower volatilities. The exposures to other risk factors are much lowered at the same time. Sensitivity analysis with the holding period of the portfolios and incorporating transaction costs also shows that the modified strategy is more robust against alternative trading strategy specifications and is thus more practically implementable.

This study contributes to both academic literature and market practice in several ways. For academic research, this study compares comprehensively various versions of the Nelson-Siegel models estimated using Bayesian techniques with different non-parametric curve fitting methods for the purpose of static term structure estimation. The sample for this research covers the major government bond market in the world, which is a wider range of countries than most of the existing literature. Moreover, evaluating the yield curve models with a fixed-income arbitrage strategy that examines the yield curve model's ability of identifying the mispricings is innovative for the fixed-income literature. Since the trading strategy and the returns of the portfolios are based on arbitrage opportunities, this study also enriches the factor investing literature with the arbitrage factor.

The remainder of this thesis starts with the developing the data in section 2. The yield curve models considered in this study are elaborated and examined in section 3 and section 4. Section

5 is dedicated for exploring possible improvements for the yield curve arbitrage strategy, and section 6 completes with conclusions and discussions.

2 Data Construction and Analysis

The bond data used for this research is retrieved from Barclays database[1]. More than 80 characteristics of 78,676 bonds from all over the world are given on a monthly basis, starting from September 1988 to October 2016.

To keep the bonds included in the analysis as homogeneous as possible, bonds bearing different characteristics that may influence the yield patterns (for example, credit risk or illiquidity) are not desirable. For this purpose, I only consider treasuries issued in local currency by governments from Australia (AU), Canada (CA), Germany (DE), the U.K. (UK), Japan (JP), Sweden (SE), and the U.S. (US), as the government bond markets of these countries are relatively large, safe, liquid and developed[2]. The remaining data from these seven countries are also given treatments for the same purpose of keeping homogeneity. Following existing literature (see, for example, Diebold and Li, 2006; Gürkaynak et al., 2007; Diebold et al., 2008), the procedure is as follows

1. Calculate the all-in dirty price as the sum of price and accrued interest as the bond price is only given as the clean price.

2. Exclude floating rate or deferred-coupon and floating coupon frequency bonds[3], bonds with option features[4], bonds extended beyond the original redemption date, bonds with non-regular-bond security types[5], bonds that changed their characteristics (coupon type,

[1]There are two versions of the data available for each month, one from the statistics universe reflecting the daily changes of the bonds and another from the return universe presenting the characteristics of the bonds at the end of each month. Those two versions have identical data for the majority of cases, exceptions happen when a bond drops out or enters the index in the middle of the month, when this happens it will be documented in the statistics universe, but will not be used for return calculation of that month, causing difference in those two versions of the data. This study focuses on the return universe, as it may happen that a bond the portfolio holds drops out of the index during the holding period for some reason, and the return universe in this case will document its last observed information, which can be used for rebalancing with the assumption that the position is cleared when it is last observed, the statistics universe, however, will no longer have any information for this bond at the end of the month, which causes problems for the strategy construction and return calculation.

[2]About 51.91% of the observations belong to these seven countries.

[3]Coupon type of 'Euribor Floater', 'GBP Floater', 'Pure Libor Floater', 'Step-up Once', 'Step-up Multiple', 'Compound coupon paid at Maturity', 'Partly Paid', 'Pay-In-Kind' and undefined coupon type.

[4]Put/call type of 'Clean-up Call Feature', 'Callable, Refundable', 'Callable, Never Refundable', 'European Call (Callable Only on Scheduled Dates)', 'Putable with a Fixed Schedule', and undefined put/call type.

[5]Security type of 'Collateralized Strip', 'Credit Default Swap', 'Global Issue', 'Locally Issued Bonds', 'RETAIL', 'Treasury Inflation Protection Securities' and undefined security type.

security type, and currency) over time, and bonds with market value smaller than 50 million. About 28.13% of the observations are left after this procedure.

3. Exclude observations with not-available price, observations with non-trader-quoted price[6], observations with zero or not-available amount outstanding, observations of bonds with yield-to-maturity lower than -1% or higher than 50%, observations with zero modified duration and observations of bonds maturing in less than three months, where a month is defined as 30.4375 days as in Diebold and Li (2006)[7]. About 25.38% of the observations are left after this procedure.

4. Exclude observations with data errors: observations where the bond is priced by stale pricing[8], observation with uncleaned price[9], observations with a issuing date later than the maturing date, bonds issued by corperates while classified as government bonds, observations with yields differing greatly from yields at nearby maturities[10], and observations with excess month-to-day return per duration more than 150 basis points. About 21.63% of the observations are left after this procedure.

After the data cleaning, there are 4,111 bonds with 105,802 observations left for the use of modeling the term structure of the seven countries. An overview of how the data is distributed across countries is summarized in Table 1. One can observe that when looking at sample size, Japan and the U.S. are the markets with larger number of bonds while Australia and Sweden have relatively less bonds. The number of bonds available per month are not stable and vary substantially for all countries, thus numerical problems may appear when the yield curve model to be estimated has a large number of parameters, especially for Australia and Sweden.

---

[6]Only trader-quoted price is the traded price of the bond.

[7]Such a specification is a weighted average of all possible numbers of days in a month: $30.4375 = \frac{7}{12} \times 31 + \frac{4}{12} \times 30 + \frac{1}{12}(\frac{3}{4} \times 28 + \frac{1}{4} \times 29)$. Similarly, a year is defined as 365.25 days throughout this study.

[8]This is done by checking if the price stays the same up to 4 decimals for two consecutive months.

[9]This is done by checking if there is a shoot up of the clean price and a zero accrued interest when the observation date is not around a coupon payment day.

[10]This is done by checking if there is an extreme and unexplainable outlier of the backed-out Fama-Bliss spot rates for the corresponding bond observation used, see the text below.

	Australia	Canada	Germany	the U.K.	Japan	Sweden	the U.S.
# of bonds	52	153	284	184	894	44	2500
# of obs.	2923	8339	12368	7204	52684	2565	64197
Avg. # of bonds per month	14.03	42.63	56.523	32.73	127.20	11.586	145.90
Min. # of bonds per month	9	25	44	24	43	8	97
Max. # of bonds per month	22	50	82	46	152	15	202

Table 1: Number of bonds and observations available per country. The first and the second row show the total number of bonds and observations available for each country. The third to fifth row show the average, the minimum, and the maximum number of bonds observed per month for each country, respectively.

Table 2 provides summary statistics of the (cleaned) bond data. Only the results for U.S. bonds are presented in the text for demonstration as the U.S. is one of the largest and most developed government bond markets, and similar tables for bonds from the other countries can be found in the Appendix. The bonds are grouped and analyzed according to their maturities for the ease of illustration. For the U.S., most market shares are given to the long-term bonds, although more than a half of the bonds available are with maturity of less than five years. The yield of the bonds show that longer term bonds in general have higher yield as well as lower volatilities, which is in line with the stylized fact that the yield curve general takes an upward shape and that the long end of the yield curve is less volatile. Note that most of the bonds with 10 to 15 years till maturity are earlier issued as 30 or longer term bonds, which explains the older age of the 10-15 year maturity bucket.

Maturity	Avg. # Bonds	Avg. Duration	Avg. Maturity (year)	Avg. Market value (bln.)	Avg. Age (year)	Avg. Yield (%)	Std. Yield (%)	Skew. Yield	Kurt. Yield	JB Stats.
US (09/88 - 10/16)										
0-1 year	36.08	0.55	0.56	21.20	2.28	3.34	2.64	0.26	1.95	19.18
1-3 year	53.68	1.76	1.85	19.85	2.75	3.71	2.60	0.18	1.90	18.88
3-5 year	34.91	3.62	4.00	18.28	2.85	4.20	2.42	0.12	1.95	16.27
5-7 year	17.01	5.14	6.02	19.21	3.93	4.54	2.26	0.13	2.03	14.18
7-10 year	14.31	6.86	8.52	26.24	2.96	4.84	2.11	0.14	2.10	12.48
10-15 year	8.20	8.41	12.64	12.49	14.29	4.99	2.09	0.30	2.14	13.25
15+ year	25.73	12.96	23.78	20.06	6.11	5.47	1.83	0.15	2.11	12.32

Table 2: Summary statistics for government bonds of the U.S., from September 1988 to October 2016, grouped according to the maturity of the bonds. Column 2 to column 6 contain the characteristics of the bonds in the maturity buckets over time, measured by the average: number of bonds, duration, maturity, market value and age of the bonds; column 7 to column 11 report the statistics of the yields of the bonds in the maturity buckets: mean, standard deviation, skewness, kurtosis and Jarque-Bera statistic. The results for yields are on annual basis.

Next the spot rates are extracted from the filtered data using the method documented by Fama and Bliss (1987) for bootstrapping the unsmoothed zero-coupon yields. As measuring the zero-coupon rates precisely is a crucial foundation for modeling the yield term structure, instead of following most existing literature (see, for example, Diebold and Li, 2006; Diebold et al., 2008) and pooling the spot rates to get a regular maturity schedule, the unsmoothed zero-coupon yields and maturity schedule are kept, resulting in an irregular maturity schedule with different numbers of maturity-yield pairs for each month. For a random month, the number of maturity-yield pairs coincides with the number of bonds observed.

The Fama-Bliss unsmoothed spot rates and the (indicated) term structure over time are presented in Figure 1. Similar figures for the other countries are reported in the Appendix. The stylized facts about term structures are clearly present: the yield curve can take various shapes, which creates difficulties for modeling the yield curves; from both Figure 1 and Table 2, one can also observe that the short end of the yield is more volatile and in general has lower yield than the long end; also Figure 1 shows that yield dynamics are persistent. Figure 1 also shows that there often exists a gap along the maturity spectrum due to the non-issuance of bonds with long maturities, which is in line with the older age of the bonds in the 10-15 year maturity bucket as presented in Table 2.

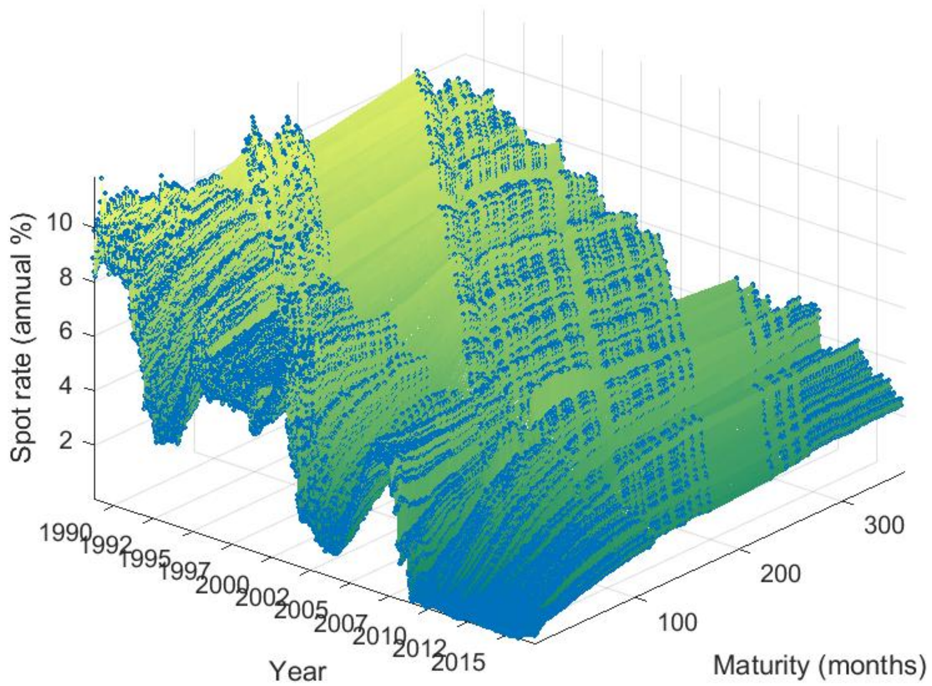


Figure 1: Term structure and bootstrapped spot rates over time of the U.S. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is the linearly interpolated term structure across time and maturities.

Another thing worth noticing from Figure 1 and Figure 6 in the Appendix is that the term structure takes different shapes and are with varying characteristics across countries, which is especially the case for the earlier sample period. For example, the term structure of Australia went through quite some shape changes over time while the Japanese one stayed rather stable at the same time. This indicates that using one yield curve model to fit the term structures of all countries can be difficult. During more recent years, however, the term structure becomes relatively calm and takes similar shapes for all countries.

3 Yield Curve Fitting Methods

This section discusses the yield curve modeling methods used in this study. Section 3.1 explains in detail the yield curve models considered for this research and the model evaluations are described in section 3.2.

3.1 Yield curve models

For this study, both the Nelson-Siegel type of models as well as non-parametric methods are considered. These methods are shown robust, performing well in capturing the dynamics of the term structure along different maturities, and are thus popular both in academics and among practitioners (see Hagan and West, 2006; de Pooter, 2007). As discussed earlier, the term structure for each of the seven countries considered in the research in general presents different features, thus modeling the bonds from different countries as one universe may be too restrictive. Therefore this research treats different countries as individual markets and models the term structure independent of each other. At the end of every month, the term structure is modeled for each country after observing the bond pricing information.

3.1.1 Nelson-Siegel models

Nelson and Siegel (1987) propose to fit the forward curve using Laguerre polynomials with an exponential decay term. The resulting function for the instantaneous forward rate curve is as follows

$$f_t[\tau] = \beta_{1,t} + \beta_{2,t} e^{-\frac{\tau}{\lambda_t}} + \beta_{3,t} \frac{\tau}{\lambda_t} e^{-\frac{\tau}{\lambda_t}} \tag{3.1}$$

where $f_t[\tau]$ is the instantaneous forward rate for a given maturity τ at time t. Averaging out the forward rates as given above by taking the integral with respect to τ, one gets the spot curve

$$r_t[\tau] = \beta_{1,t} + \beta_{2,t}\frac{1-e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} + \beta_{3,t}(\frac{1-e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} - e^{-\frac{\tau}{\lambda_t}}) \tag{3.2}$$

where $r_t[\tau]$ is the spot rate of maturity τ at time t, $\beta_{1,t}, \beta_{2,t}$, and $\beta_{3,t}$ are parameters for the three components interpreted as long-, short- and medium-term component, respectively. The long-term component $\beta_{1,t}$ takes loading 1 for all maturities; the loading of $\beta_{2,t}$ takes value 1 when approaching maturity zero and decays exponentially as the maturity increases, thus $\beta_{2,t}$ has most impact on the short maturities; the loading of $\beta_{3,t}$ takes value zero when approaching maturity zero, reaches 1 around medium-term maturities and decays to 0 again when the maturity approaches infinity, thus $\beta_{2,t}$ is interpreted as the medium-term component. And finally λ_t governs the decay of the β parameters towards zero. This model is denoted as **NS1** for the rest of this research.

The Nelson-Siegel model is popular in both academics and practice for several reasons. First, the approximation of the term structure with only a small number of parameters makes the model itself parsimonious. Second, it generates a smooth and continuous forward curve by construction at the same time. Lastly, it has clear interpretations for its components such that the variations in the yields can be explained by these terms. Much research and extensions have been done regarding this framework to improve its performance since Nelson and Siegel (1987). One of the various ways to increase the flexibility of the Nelson-Siegel model is by introducing a fourth component. Björk and Christensen (1999) extend the Nelson-Siegel model by introducing a fourth term that resembles the second term:

$$f_t[\tau] = \beta_{1,t} + \beta_{2,t}e^{-\frac{\tau}{\lambda_t}} + \beta_{3,t}\frac{\tau}{\lambda_t}e^{-\frac{\tau}{\lambda_t}} + \beta_{4,t}e^{-\frac{2\tau}{\lambda_t}} \tag{3.3}$$

$$r_t[\tau] = \beta_{1,t} + \beta_{2,t}\frac{1-e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} + \beta_{3,t}(\frac{1-e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} - e^{-\frac{\tau}{\lambda_t}}) + \beta_{4,t}\frac{1-e^{-\frac{2\tau}{\lambda_t}}}{\frac{2\tau}{\lambda_t}} \tag{3.4}$$

the fourth term, together with the second term, affects the short-term maturities. Diebold et al. (2006) confirm that this model marginally improves the in-sample fit compared to the original Nelson-Siegel model. This model is denoted as **NS2**.

Alternatively, Svensson (1994) proposes to add a second hump-shape factor with a separate decay parameter, resulting in a forward rate curve and a zero-coupon yield curve as follow:

$$f_t[\tau] = \beta_{1,t} + \beta_{2,t} e^{-\frac{\tau}{\lambda_{1,t}}} + \beta_{3,t} \frac{\tau}{\lambda_{1,t}} e^{-\frac{\tau}{\lambda_{1,t}}} + \beta_{4,t} \frac{\tau}{\lambda_{2,t}} e^{-\frac{\tau}{\lambda_{2,t}}} \tag{3.5}$$

$$r_t[\tau] = \beta_{1,t} + \beta_{2,t} \frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} + \beta_{3,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} - e^{-\frac{\tau}{\lambda_{1,t}}} \right) + \beta_{4,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{2,t}}}}{\frac{\tau}{\lambda_{2,t}}} - e^{-\frac{\tau}{\lambda_{2,t}}} \right) \tag{3.6}$$

the fourth term is of the same structure of the third term that affects mainly the medium-term maturities, but with a different decay parameter. This specification can fit the term structures that take shape of more than one hump along the maturity spectrum, which is one of the features the unsmoothed spot rates exhibit as shown in earlier sections. However, as shown and discussed in existing literature, the estimation of this model suffers from multicolinearity problems when the decay parameters $\lambda_{1,t}$ and $\lambda_{2,t}$ take similar values (see, for example, de Pooter, 2007; Gimeno and Nave, 2006). Thus instead of estimating the original Svensson model, this study considers an "adjusted" Svensson model proposed by de Pooter (2007) that guarantees the two medium-term components take different values when $\lambda_{2,t}$ approaches $\lambda_{1,t}$:

$$f_t[\tau] = \beta_{1,t} + \beta_{2,t} e^{-\frac{\tau}{\lambda_{1,t}}} + \beta_{3,t} \frac{\tau}{\lambda_{1,t}} e^{-\frac{\tau}{\lambda_{1,t}}} + \beta_{4,t} \left[\frac{\tau}{\lambda_{2,t}} e^{-\frac{\tau}{\lambda_{2,t}}} + \left(\frac{2\tau}{\lambda_{2,t}} - 1 \right) e^{\frac{2\tau}{\lambda_{2,t}}} \right] \tag{3.7}$$

$$r_t[\tau] = \beta_{1,t} + \beta_{2,t} \frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} + \beta_{3,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} - e^{-\frac{\tau}{\lambda_{1,t}}} \right) + \beta_{4,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{2,t}}}}{\frac{\tau}{\lambda_{2,t}}} - e^{-\frac{2\tau}{\lambda_{2,t}}} \right) \tag{3.8}$$

This model, denoted as **NS3**, has the fourth component that still affects the medium-term maturities but increases at a faster rate than the other medium-term component, and therefore multicolinearity is no longer an issue for the estimation.

The specifications from the Nelson-Siegel family considered in this research can be written in a more general form:

$$\mathbf{r}_t = \Phi_t[\lambda_t]\beta_t + \varepsilon_t \tag{3.9}$$

$$\varepsilon_t \sim N(\mathbf{0}, \sigma_t^2 I_{N,t}) \tag{3.10}$$

where r_t is a vector of length N containing all the spot rates observed in month t, $\Phi_t[\lambda_t]$ is a $N \times k$ matrix, where N is the number of bonds observed in month t and k is the length of the parameter vector β_t, and the value of the elements depends on λ_t. Note that N takes different values each month. ε_t collects the error terms and I_p is a p-dimensional identity matrix. Typically, elements in ε_t are assumed to be independent but with different variances for different maturities. Considering the setting of this research that the maturities are not pooled such that τ_{i_t} varies over time, maturity-specific variance for each maturity for every month casts obstacles for both the estimation and the interpretations of the variance parameters, thus the error terms are are assumed independent and having the same variance σ_t^2.

Due to the existence of the parameters λ_t, the estimation for the Nelson-Siegel type of models suffers from non-linearity and can generally only resort to non-linear least squares (NLS). However, for most of the countries the term structure only has a rather small number of observations along the maturity spectrum. Estimation with NLS can thus be problematic, presenting more than one local minimum for the λ_t parameters, and leading to a poor fit of the yield curve and extreme values for the estimated β_t parameters. Existing literature also address and discuss the numerical problems and difficulties in the estimation of the Nelson-Siegel type of models (see, for example, Bolder and Gusba, 2002; Gimeno and Nave, 2006; de Pooter, 2007). To overcome these issues, Diebold and Li (2006) keep λ_t constant over time and fix it to a pre-determined value and proceed to estimate the β_t parameters using ordinary least squares. Given that it is of particular interest of this study to get a precise measure of the yields and forward rates, maintaining the model specification is preferred over using approximations or fixing parameters beforehand. Thus the estimation is performed through Bayesian inferences with Markov chain Monte Carlo (MCMC). Chib and Ergashev (2009) point out that estimation using Bayesian techniques has advantages over dimensionality, identification and inference problems in multi-factor term structure models. In the Bayesian approach, the non-linearity is no longer a concern as the parameters are simulated from the conditional posterior distributions, which means when sampling β_t parameters, the other parameters including λ_t are treated as given such that (3.10) become linear. Moreover, since the parameters estimations are taken as the mean value of the draws from the posterior distributions, parameter uncertainty caused by uncertainties in the data is also accounted for. The detailed methodology regarding the estimation of the Nelson-Siegel models are discussed in the Appendix.

3.1.2 Non-parametric models

Next to the Nelson-Siegel methods, model-free methods are also a popular choice for fitting the yield curve. Such non-parametric methods, as the name suggests, do not necessarily provide economic interpretations for its components or the outputs but rather a best fit based on the data, and are relatively easy to implement.

For this study, cubic regression splines are the first type of methods considered. Bliss (1996) states that a yield curve modeled with with spline method performs at least as well as other methods. Moreover, Beim (1992) reports that cubic splines perform at least as well as other spline orders. Following Steeley (1991), the splines are constructed with a B-spline basis for this study such that each piece-wise polynomial takes form of a B-spline. Nawalkha and Soto (2009) report that B-splines have advantage in incorporating the possible heterogeneity for different maturities compared to the simple piece-wise regression splines. Lin (2002) also confirms that this approach is satisfactory in obtaining reliable term structure and not very sensitive to some *ad hoc* choices in the model estimation. Because these cubic splines are defined piece-wise, the estimation is straightforward and numerically stable (see Bolder and Gusba, 2002). In the rest of the paper this method is denoted as **BS**.

Suppose the relationship between time-to-maturity and spot rates is approximated by a k-piece cubic spline parameterized with a B-spline basis. The theoretical spot rate with maturity τ by the spline is given by

$$\mathrm{r_t}[\tau] = \sum_{\mathrm{i}} \alpha_{\mathrm{i}} \mathrm{B}_{\mathrm{i},3,\mathrm{t}}[\tau] \tag{3.11}$$

where $\mathrm{B}_{\mathrm{i},3,\mathrm{t}}$ is the basis function for a B-spline of order three defined over the interval between the i^{t}h and $\mathrm{i}+1^{\mathrm{th}}$ knot, $[\mathrm{k_i}, \mathrm{k_{i+1}}]$, and α_{i} is the coefficient to be estimated. The basis functions of order n, $\mathrm{B}_{\mathrm{i},\mathrm{n},\mathrm{t}}$, can be derived by means of Cox-de Boor recursion formula (see de Boor, 1978)[11]

$$\mathrm{B}_{\mathrm{i},0,\mathrm{t}}[\tau] = \begin{cases} 1 & \text{if } \mathrm{k} \leq \tau < \mathrm{k_{i+1}} \\ 0 & \text{otherwise} \end{cases} \tag{3.12}$$

$$\mathrm{B}_{\mathrm{i},\mathrm{n},\mathrm{t}}[\tau] = \frac{\tau - \mathrm{k_i}}{\mathrm{k_{i+n}} - \mathrm{k_i}} \mathrm{B}_{\mathrm{i},\mathrm{n}-1,\mathrm{t}}[\tau] + \frac{\mathrm{k_{i+n+1}} - \tau}{\mathrm{k_{i+n+1}} - \mathrm{k_{i+1}}} \mathrm{B}_{\mathrm{i}+1,\mathrm{n}-1,\mathrm{t}}[\tau], \qquad \mathrm{n} = 1, 2, 3 \tag{3.13}$$

From (3.12) and (3.13), a cubic B-spline, namely n = 3, is defined in the following manner for arbitrary τ

[11] The worked-out formulation is discussed in the Appendix with more details.

$$B_{i,3,t}[\tau] = \begin{cases} B_{i,3,t}[\tau] & \text{if } k_i \leq \tau < k_{i+4} \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

Such a specification satisfies the restriction that the fitted curve is C^2. Despite that most studies prefer a similar amount of data points for each sub-interval for the ease of estimation (see, for example, McCulloch, 1971), with the aim of capturing the yield dynamic of different maturities, this study splits the whole maturity spectrum into three sub-intervals such that the resulting cubic spline has three pieces and four knots similar as in existing literature (see a discussion in for example, Steeley, 1991; Nawalkha and Soto, 2009)[12]. The three sub-intervals are: maturities of less than 48 months (four years), between 48 and 120 months (four to 10 years), and longer than 120 months (10 years)[13], representing short, medium and long maturities, respectively. And the knots are correspondingly defined as $[0, 48, 120, \tau_{max}]$[14]. Such an interval division gives flexibility of capturing the yield dynamic at different part of the maturity spectrum, also the number of observations in each interval do not very drastically in most cases. The α parameters are estimated by minimizing the mean squared errors:

$$\underset{\alpha}{\text{argmin}} \quad \frac{1}{N}\sum_{i=1}^{N}(r_i - \hat{r}_i)^2 \tag{3.15}$$

In order to avoid over-smoothing, the restriction that there should be no less than six data points available for estimating each piece-wise polynomial is also imposed. Therefore in some cases there are not enough data points for some sub-interval. When this happens to a short or long maturity interval, the missing data points are made up by taking the points in the medium maturity interval that are closest to the short or long maturity interval; when this happens to a medium interval, the points are taken from the long maturity interval, as the short end of the yield curve usually contains more variations and should be treated more carefully.

Next, the kernel smoothing framework is the other focus of the non-parametric methods

[12]Nawalkha and Soto (2009) further splits the short-term interval at one-year maturity. For the sample used in this study, it happens that the earliest maturing bond is with maturity of more than two years, thus I do not consider intervals smaller than this the specification described.

[13]A month is defined as 30.4375 days and a year is defined as 365.25 days.

[14]Experimentations show that a flexible selection for the second knot around 48 do not have large impact on the resulting curve, due to the fact that there are usually ample short- and medium-term bonds; altering the third knot can, however, change the shape of the estimated yield curve at the long end when the long-bonds are spread further away from each other along the maturity spectrum.

for this research. This non-parametric regression approach assumes no function form of the relationship between $r_t[\tau]$ and τ such that the yield curve can be expressed as

$$r_t[\tau] = g_t[\tau; \theta] + \varepsilon_t \tag{3.16}$$

where θ is the parameters in the function $g_t[\cdot]$, ε_t is the error term and $g_t[\cdot]$ is the term of interest. Note that this approach assumes that the elements in ε_t are independent and $E[\varepsilon_{it}] = 0$ and $Var[\varepsilon_{it}] = \sigma^2$ for all $i = 1, ..., N$. One of the most popular forms $g_t[\cdot]$ takes is the Nadaraya–Watson estimator (see Nadaraya, 1964; Watson, 1964). Sediva and Marek (2015) report that the yield curve fitted by kernel regressions with the Nadaraya–Watson estimator can achieve comparable results as those by the Nelson-Siegel models. The function form $g[\cdot]$ is then given by

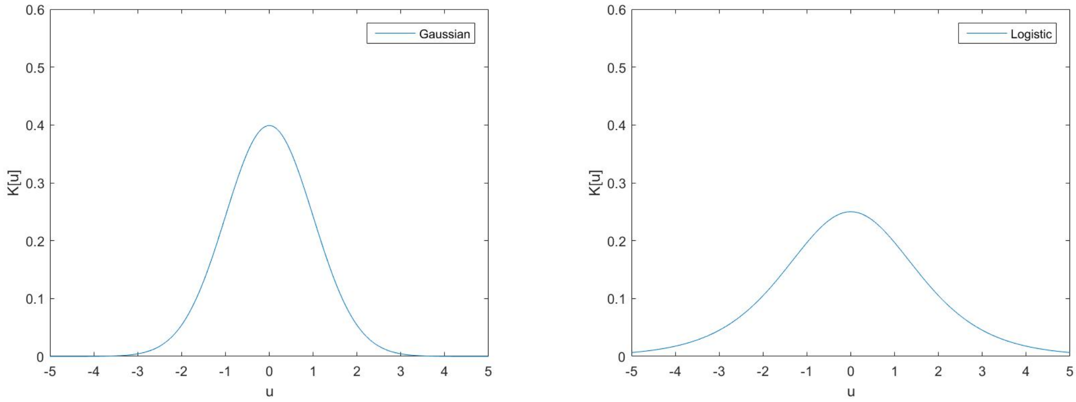
$$g[\tau; h] = \frac{\sum_{i=1}^{N} r[\tau_i] K[\frac{\tau - \tau_i}{h}]}{\sum_{i=1}^{N} K[\frac{\tau - \tau_i}{h}]} \tag{3.17}$$

where $K[\cdot]$ is a kernel function and h is the bandwidth parameter that controls the smoothness. Note that the fitted function values derived from such a specification can be viewed as the weighted average of the data, and the weight of each data point is decided by the kernel function. The choice of the kernel function and the bandwidth parameter has a substantial effect on the curve fitting. For the kernel function this paper follows the methods by Dupačová et al. (1997) and Sediva and Marek (2015) and considers two continuous kernel functions, namely Gaussian: $K[u] = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$, and Logistic: $K[u] = \frac{1}{e^{-u} + 2 + e^{u}}$, denoted as **KS1** and **KS2**, respectively[15].

The main difference of these two kernel functions are that the Gaussian kernel is more peaked and has thinner tails than the Logistic kernel, which essentially means that when the bandwidth parameter is the same, the Gaussian kernel mainly considers a smaller subset of the data for evaluating the function value than the Logistic kernel, as the Gaussian kernel approaches to zero at a faster rate. Therefore the curve generated with the Gaussian kernel turns to be more oscillating. The selection of bandwidth parameter is performed every month using cross-validation techniques (see Simonoff, 2012) that compute the optimal bandwidth value again by minimizing the mean squared errors as given in (3.15), where $\hat{r}_j = \hat{g}_{-j}[\tau_j]$ is calculated as the Nadaraya–Watson estimator computed without the point (τ_j, r_j) and evaluated at $\tau = \tau_j$.

In some cases, the gap between two consecutive maturities along the maturity spectrum can

[15]Experimentations with discontinuous kernel functions (Epanechnikov and Quartic) sometimes failed in the estimation for the bandwidth parameters, especially for countries with less data points (Australia and Sweden).



(a) Gaussian kernel (b) Logistic kernel

Figure 2: Kernel function for the Gaussian kernel (left panel) and the Logistic kernel (right panel).

be large (for example, when a country started reissuing long term bonds after some time), if a maturity of interest falls in this gap, it could happen that Gaussian kernel takes zero value for all $u = \frac{\tau - \tau_i}{h}$ when the bandwidth parameter is estimated to have a small absolute value[16]. To avoid this, I imposed the restriction that for the Gaussian kernel, the bandwidth parameter should be large enough that, within the maturity spectrum, the maturity sits farthest to all data points can get a fitted value with the Nadaraya–Watson estimator[17]. One can show that this maturity is actually the mid-point of the largest gap along the maturity spectrum. Written in mathematical notation, suppose for a random month, the largest gap among all consecutive maturities are between τ_p and τ_q, let $\hat{\tau}$ be the maturity sits farthest to all data points, that is, $\hat{\tau} = \frac{\tau_q - \tau_p}{2} + \tau_p = \tau_q - \frac{\tau_q - \tau_p}{2}$

$$g[\hat{\tau}; h] = \frac{\sum_{i=1}^{N} r[\tau_i] K[\frac{\hat{\tau} - \tau_i}{h}]}{\sum_{i=1}^{N} K[\frac{\hat{\tau} - \tau_i}{h}]} > \xi \tag{3.18}$$

where ξ is an arbitrarily small positive number[18]. One can show that the equivalence of (3.18)

[16]The Logistic kernel takes form of a combination of an exponential function and an inverse function and is thus free from this problem.

[17]Only estimating the yield curves for Japan and the U.S. requires this restriction, as the larger gaps along the maturity spectrum mostly happen to these two countries. A large part of the yield curve estimation of the two countries are bound by this restriction to have a valid fitted yield for all maturities.

[18]For the estimation, ξ is set to the machine number floating-point relative accuracy, approximated by 2.22×10^{-16}.

is the following

$$K[\frac{\hat{\tau}-\tau_p}{h}] > \xi \quad \vee \quad K[\frac{\hat{\tau}-\tau_q}{h}] > \xi \tag{3.19}$$

$$\tag{3.20}$$

where $\vee$ stands for the logical operation "or", namely at least one of the constraints should be satisfied. However, since $\hat{\tau} - \tau_p = \hat{\tau} - \tau_q$, and $K[\cdot]$ is symmetric around zero for the Gaussian kernel, a one sided inequality is sufficient, this means,

$$K[\frac{\frac{\tau_q-\tau_p}{2}}{h}] > \xi \tag{3.21}$$

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{(\frac{\tau_q-\tau_p}{2h})^2}{2}} > \xi \tag{3.22}$$

after manipulating, the restriction for the parameter h is then

$$h > \frac{\frac{\tau_q-\tau_p}{2}}{\sqrt{-2\ln(\xi\sqrt{2\pi})}} \tag{3.23}$$

3.2 Model evaluation

3.2.1 Statistical properties

To examine the goodness-of-fit of the models proposed above, one can compare the curve fitting with respect to the data points in terms of statistical properties by means of R-squared (R^2). To assess the influence of the increased number of parameters estimated, one can also consider the adjusted R-squared ($\bar{R}^2$) as a modified version of R^2 that adjusts for the number of parameters in the model[19]. For each month, the R^2 is calculated as follows

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(r[\tau_i] - \hat{r}[\tau_i])^2}{\sum_{i=1}^{N}(r[\tau_i] - \bar{r})^2} \tag{3.24}$$

[19]Another common used measure is the mean squared errors (MSE), calculated as the average value of the squared residuals, $MSE = \frac{1}{N}\sum_{i=1}^{N}(r[\tau_i] - \hat{r}[\tau_i])^2$. As the MSE statistic has a one-to-one correspondence to the R^2 defined as (3.24), here I only report the R^2 statistics.

where $\hat{r}[\tau_i]$ denotes the fitted value of $r[\tau_i]$, $\bar{r}$ is the average of the observed spot rates over all maturities, $|\cdot|$ the absolute value function, and N the number of bonds. And the $\bar{R}^2$ is

$$\bar{R}^2 = 1 - \frac{(1 - R^2)(N - 1)}{N - k - 1} \tag{3.25}$$

where R^2 is as defined in (3.24) and k is the number of parameters estimated in the model.

Another important measurement for goodness-of-fit for models pricing derivatives is the pricing error. Here the mean squared pricing error (MSPE) and the mean squared yield error (MSYE) generated by the each curve are computed as

$$MPSE = \frac{1}{N}\sum_{i=1}^{N}(P_i - \hat{P}_i)^2 \tag{3.26}$$

$$\hat{P}_i = \sum_{j=1}^{k} CF_{i,j}e^{-\tau_{i,j}r[\tau_{i,j}]}, \quad i = 1, .., N \tag{3.27}$$

where P_i is the price of the i^{th} bond, $\hat{P}_i$ is the theoretical price by yield curve calculated as the sum of the discounted value of future cash flows, $CF_{i,j}$ and $r_t[\tau_{i,j}]$ are the j^{th} cash flow and corresponding discount rate for the j^{th} cash flow of the i^{th} bond and this bond has k remaining cash flows; and

$$MPYE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{3.28}$$

$$P_i = \sum_{j=1}^{k} CF_{i,j}e^{-\tau_{i,j}y_i}, \quad i = 1, .., N \tag{3.29}$$

$$\hat{P}_i = \sum_{j=1}^{k} CF_{i,j}e^{-\tau_{i,j}\hat{y}_i}, \quad i = 1, .., N \tag{3.30}$$

where y_i stands for the actual yield of the bond while $\hat{y}_i$ stands for the theoretical yield by the model, and both of them need to be solved with numerical tools.

3.2.2 Economic values

For a setting of yield curve fitting, a necessary yet often neglected aspect of goodness-of-fit of the curves is the economic values behind. This paper proposes an innovative approach with a yield curve arbitrage strategy that examines the modeled yield curves' ability to identify the mispricings in the bonds as a measure for the economic values of the yield curve models, as if a model is able to capture the maturity-yield dynamics well, it should be able to pick up the latent mispricings. Following this intuition, a trading strategy based on yield curve arbitrage opportunities, that is, going long in the bonds that are "cheap" (traded at a lower price or implying higher yield) compared to the theoretical value indicated by the term structure and/or shorting the ones that are more expensive (traded at a higher price or implying lower yield), should generate consistent profits. To my best knowledge, this is the first study that compares the Nelson-Siegel models and various non-parametric methods for in-sample yield curve fitting by examining the economic values behind in terms of the pricing power with a fixed-income arbitrage strategy. The results of this section serve as a foundation and a generic start for constructing a most profitable yield curve arbitrage strategy.

In order to fully reflect the modeled yield curve, the generic arbitrage strategy should be constructed in a way that the information from the modeled yield curves are used as much as possible. Thus the strategy sets up portfolios with both long and short positions to include all arbitrage opportunities. Ideally, the magnitude of the positions of the bonds are decided by the yield error each bond bears, however, since bonds with longer duration bear more interest rate risk, a better approach to eliminate the effect of the interest rate risk is to scale the positions by yield error per duration rather than the yield error itself. Therefore, in each month, the strategy longs all the bonds that have a higher actual yield than theoretical yield and shorts all the bonds that have a lower actual yield, and both the long part and the short part have one unit of duration so the portfolio itself is duration neutral. Written in mathematical formulation, suppose in a random month, there are N bonds observed, the weight for a bond i can be decided as

$$\mathrm{w_i} = \begin{cases} \frac{1}{\mathrm{D_i}} \frac{\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i}}{\sum_{\mathrm{j}=1}^{\mathrm{N}} (\mathrm{y_j} - \hat{\mathrm{y}}_\mathrm{j}) I_{\hat{\mathrm{y}}_\mathrm{j} \geq \mathrm{y_j}}} & \text{if } \mathrm{y_i} \geq \hat{\mathrm{y}}_\mathrm{i} \\ -\frac{1}{\mathrm{D_i}} \frac{\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i}}{\sum_{\mathrm{j}=1}^{\mathrm{N}} (\mathrm{y_j} - \hat{\mathrm{y}}_\mathrm{j})(1 - I_{\mathrm{y_j} \geq \hat{\mathrm{y}}_\mathrm{j}})} & \text{otherwise} \end{cases} \tag{3.31}$$

where $I_{[\cdot]}$ is an indicator function, $\mathrm{D_i}$ stands for the duration of bond i, and $\hat{\mathrm{y}}_\mathrm{i}$ and $\mathrm{y_i}$ are the theoretical yield by the modeled spot curve (see (3.30)) and the actual yield (see (3.29)) of the bond, respectively. If all the bonds are considered cheap or expensive[20], no trade is set up for

[20]This only happens to Australia and Sweden during the time when the number of bond existing in the country is small. The yield of a coupon bearing bond will lie below a zero-coupon yield with the same maturity due to the discounting of the coupon payments, which holds for both the modeled yields and the actual yields. This makes it possible that all the bonds along the maturity spectrum are evaluated as "rich" or "expensive".

this month, and the strategy invests in cash and earns zero excess return. This is due to the fact that in this case a duration neutral portfolio is not possible. Note that such a portfolio does not necessarily have zero net positions but is only duration neutral. The portfolio is held for one month and rebalanced every month. The analysis throughout the following sections are conducted based on excess portfolio returns in stead of gross returns.

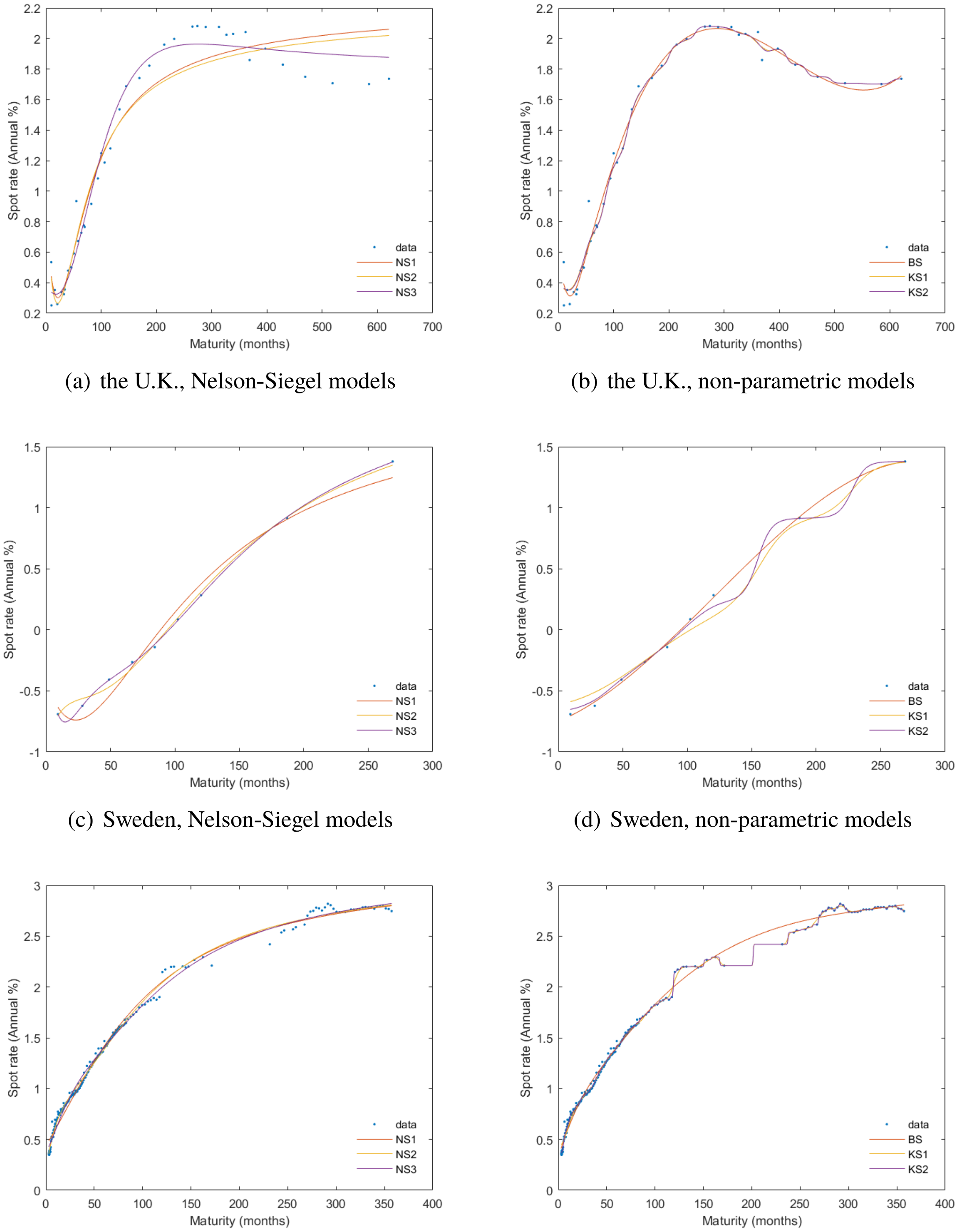
4 Yield Curve Model Selection

This section discusses the model fit of term structure models. The performance in terms of statistical measures and economic values elaborated in the last section are reported in section 4.1 and section 4.2, respectively. The final yield curve model selection is presented in section 4.3.

4.1 Model fit

By definition, more flexible models give better fit in term of statistics when the model parameters are properly estimated, due to the increase in the degrees of freedom. Thus the hypothesis is that comparing the statistical properties would favor the non-parametric methods.

Figure 3 gives an example of a visualization of the yield curve models. The upper, middle, and lower panel depicts the zero-coupon curve of the U.K., Sweden and the U.S., respectively. The spot rate data the modeled curves are based on is observed at the end of October 2016. Similar figures for Australia, Canada, Germany, and Japan can be found in the appendix. It is easy to notice that the curves estimated by kernel smoothing are more oscillating, which is to be expected. This is especially the case for the U.K., and the U.S. As explained before, the Nadaraya-Watson estimator fits every point on the curve as an weighted average of the data, since the U.K. and the U.S. have a larger number of data points available around each point to be evaluated compared to Sweden, the estimator turns to choose a smaller bandwidth parameter, meaning the estimation depends only on values close to the point of evaluation, which also explains the "stepping" feature of the curves by kernel smoothing the right panels of Figure 3 exhibit: when the point of interest falls at a region where there are only limited amount of data, the kernel function reaches to zero quickly for data points further away, generating wobbly curves, and this is especially the case for the KS1 model. Also, the curves generated by kernel smoothing follow the data more closely and are therefore less smooth. Due to the existence of such features, the kernel smoothing methods are much less smooth than the other models.



(a) the U.K., Nelson-Siegel models
(b) the U.K., non-parametric models
(c) Sweden, Nelson-Siegel models
(d) Sweden, non-parametric models
(e) the U.S., Nelson-Siegel models
(f) the U.S, non-parametric models

Figure 3: Yield curve modeled by different methods for the U.K., Sweden, and the U.S. The zero-coupon yields are bootstrapped based on the bond price information observed at the end of September, 2016.

Figure 3 and Figure 14 also show that the shape that the term stricture takes also influences the model fit. For the U.K. and Canada, the very long end of the yield curve is observed to bend downwards. The non-parametric methods can easily adapt this feature but the Nelson-Siegel models encounter more difficulties when trying to fit this shape, as the loadings of Nelson-Siegel models for the $\beta_2, \beta_3, \beta_4$ parameters go to zero when the maturity increases, resulting in a flatter long end of the yield curve.

Looking at the statistical properties, all the models are able to explain most of the variations in the data according to the large R-squared (R^2), except that the NS3 model does not do well in capturing the yield-maturity dynamics for Australia and Sweden. As described in earlier sections NS3 has six parameters to be estimated, and unlike the spline method, the function form of the Nelson-Siegel models is highly non-linear. Thus it is reasonable that poor estimations are likely to happen when the sample size is small, which is the case for Australia and Sweden. The time series plots of the R^2 (Figure 15) reported in the Appendix also confirm that numerical numerical problems happen to the estimation of NS3 for Australia and Sweden more often. The adjusted R-squared ($\bar{R}^2$) statistics show that within the Nelson-Siegel models, adding an extra slope factor (NS2) is more likely to benefit the model fit than an extra curvature factor (NS3). The lower $\bar{R}^2$ compared to the R^2 also shows that BS following NS3 suffers from efficiency loss due to the larger number of parameters estimated. The KS1 and KS2 model in general give the best statistics, especially in terms of $\bar{R}^2$, mean squared pricing errors (MSPE), and mean squared yield errors (MSYE). This is reasonable as these statistics are to a big extent directly related to how closely can the model follow the data. From Figure 15 and Figure 16 in the Appendix one can also observe that while KS1 and KS2 remaining the models with better fit over time, for most models, fitting the yield curve during the crisis (for example, around 2008) when the market is more turbulent is more difficult than when the market is quiet, regardless of the country.

To further understand how data size influences the statistical properties of the models, the countries are also grouped and analyzed by the amount of the bonds available in the market. Canada, Germany, the U.K., Japan, and the U.S. have larger numbers of bonds observed throughout the sample period and are thus labeled together as large market, Australia and the Sweden are on the contrary labeled as small market. For the countries with larger markets, the results are in line with the hypothesis that the non-parametric methods, especially the kernel smoothing models, give better statistics. Nevertheless, when looking at Australia and Sweden, the Nelson-Siegel models are often able to generate comparable results as those of the non-parametric methods, sometime even outperform. Larger difference between the modeled yields and the actual yields are more likely to happen at the long end of the yield curve. In larger markets, the long-term bonds usually have different characteristics (for example, a bond with 20-year-till maturity issued as 30-year-bond before tends to have much higher coupons thus lower yield than a recently issued 20-year bond), this could cause the larger pricing errors for the bigger countries as presented in the lower panels in Table 3.

	NS1	NS2	NS3	BS	KS1	KS2		NS1	NS2	NS3	BS	KS1	KS2
R^2							$\bar{R}^2$						
AU	0.81	**0.85**	0.66	0.83	0.82	**0.83**	AU	0.70	0.72	0.08	0.56	**0.80**	**0.81**
CA	0.78	0.79	0.78	0.85	**0.86**	**0.86**	CA	0.75	0.74	0.72	0.81	**0.85**	**0.86**
DE	0.91	0.92	0.90	0.94	**0.95**	**0.95**	DE	0.90	0.91	0.89	0.93	**0.94**	**0.95**
UK	0.80	0.84	0.81	**0.88**	0.86	**0.87**	UK	0.76	0.80	0.76	0.84	**0.86**	**0.86**
JP	0.99	0.99	0.99	1.00	**1.00**	**1.00**	JP	0.99	0.99	0.99	1.00	**1.00**	**1.00**
SE	**0.79**	**0.82**	0.64	0.78	0.74	0.74	SE	0.67	0.69	0.23	0.54	**0.71**	**0.72**
US	0.85	0.87	0.88	0.93	**0.96**	**0.97**	US	0.84	0.86	0.87	0.92	**0.96**	**0.97**
Average	0.85	0.87	0.81	**0.88**	0.88	**0.89**	Average	0.80	0.82	0.65	0.80	**0.88**	**0.88**
Avg. Large	0.86	0.88	0.87	0.92	**0.92**	**0.93**	Avg. Large	0.85	0.86	0.85	0.90	**0.92**	**0.93**
Avg. Small	0.80	**0.84**	0.65	**0.80**	0.78	0.78	Avg. Small	0.69	0.70	0.15	0.55	**0.76**	**0.76**
MSPE							**MSYE**						
AU	0.60	0.56	0.86	0.58	**0.53**	**0.52**	AU	2.86	**2.67**	3.70	2.81	2.73	**2.30**
CA	2.54	2.42	2.52	1.61	**1.34**	**1.28**	CA	5.12	5.16	5.96	5.00	**4.67**	**4.58**
DE	1.13	1.07	0.94	0.50	**0.34**	**0.34**	DE	1.91	1.82	2.24	1.65	**1.44**	**1.40**
UK	2.40	2.08	2.98	1.22	**1.19**	**1.04**	UK	3.14	2.94	3.87	**2.54**	2.62	**2.31**
JP	0.49	0.34	0.37	0.21	**0.09**	**0.09**	JP	0.69	0.40	0.45	0.33	**0.27**	**0.25**
SE	0.67	0.57	1.72	0.78	**0.52**	**0.50**	SE	**3.20**	**2.79**	4.97	3.52	4.04	3.96
US	1.76	1.76	1.85	1.90	**1.61**	**1.60**	US	6.84	6.62	6.83	6.76	**6.37**	**6.22**
Average	1.37	1.26	2.01	0.97	**0.80**	**0.77**	Average	3.39	3.20	4.67	3.23	**3.16**	**3.00**
Avg. Large	1.66	1.53	2.10	1.09	**0.91**	**0.87**	Avg. Large	3.54	3.39	4.42	3.26	**3.07**	**2.95**
Avg. Small	0.64	0.57	1.79	0.68	**0.52**	**0.51**	Avg. Small	**3.03**	**2.73**	5.28	3.17	3.39	3.13

Table 3: Statistical properties examined of different yield curve fitting models across countries. The two upper panels report the average R-squared (R^2) and the average adjusted R-squared ($\bar{R}^2$) over time, and the two lower panels contain the average mean squared pricing errors (MSPE) and the average mean squared yield errors (MSYE) over time. The bottom three rows of each panel report the average value of all countries (Average), countries with ample number of bonds, namely Canada, Germany, the U.K., Japan, and the U.S. (Avg. Large), and countries with limited number of bonds, Australia and Sweden (Avg. Small). The R^2, $\bar{R}^2$ and the MSPE are reported in real numbers and the the MSYE is in basis point. The two best-performing models for each country-property pair are marked bold face.

Thus one might want to conclude that based on the statistical properties, the kernel smoothing method is more suitable for modeling the term structure in general; if the number of bonds in the market is also taken into account, one should consider the Nelson-Siegel model with the extension of Björk and Christensen (1999) or a B-spline as alternatives. However, these statistical measurements do not fully penalize for the number of parameters estimated. As a matter of fact, the non-parametric methods used in this paper function to a large extent as numerical methods, which makes it difficult to justify the effect of the number of parameters or the actual number of parameters. Moreover, as the yield curve models serve as a foundation on which a trading strategy will be built upon, a more relevant and more important measure would be how well the models perform in an economic content rater than a curve fitting content. Therefore it is of importance to further examine economic values of the models to a conclusion on which yield curve model performs best and on which yield curve model should a trading strategy be based on.

4.2 Economic values

Unlike for the statistical properties, a more relative aspect for a yield curve arbitrage strategy would be the ability of capturing the underlying market information rather than following the noisy up and down movements along the maturity spectrum closely. While the Nelson-Siegel type of models generate a smooth curve, the numerical models are more prone to over-fitting problems, especially for markets with a smaller data size or more volatilities. Hence one would expect that the returns of a strategy based on the Nelson-Siegel models will be less volatile, and that least for smaller and more volatile countries, the Nelson-Siegel models would perform better in terms of yield curve arbitrage strategy returns.

4.2.1 Returns

The results of the trading strategy constructed in section 3.2.2 and (3.31) are reported in Table 4, and the figures showing the strategy cumulative performance of each country with different models are reported in Figure 17 in the Appendix.

As expected, the best performing models under almost all measures are now spread over and differ across countries rather than concentrated around the kernel smoothing models and the lower volatilities often belong to the Nelson-Siegel models. Figure 17 in the Appendix reveals that while the strategy performs moderately or even suffers from losses when the market is quiet, the profit is mostly attributed to more volatile market periods (for example, the dot-com bubble around 1999-2003 and the financial crisis of 2007–2008). Since the strategy benefits from arbitrage opportunities, this is in line with the conclusion of Hu et al. (2013) that the magnitude of mispricing of the U.S. government bonds is larger during crisis periods.

The Sharpe ratios for the majority of the country-model pairs are positive, meaning the models are able to identify cheap and expensive bonds and capture the arbitrage opportunities. Moreover, it is noticeable that the trading strategy in general generates higher volatilities for Australia and Sweden than for the other countries, regardless of the model the strategy is based on. Overall, the most profitable country for an arbitrage strategy is Canada, as most of the models are able to give high returns with low volatilities.

	NS1	NS2	NS3	BS	KS1	KS2		NS1	NS2	NS3	BS	KS1	KS2
Sharpe							**Vol.**						
AU	*-0.34*	*-0.01*	**0.17**	*-0.27*	**0.11**	*-0.27*	AU	**27.16**	28.56	37.47	30.09	**26.86**	27.83
CA	1.14	0.98	0.07	0.94	**1.21**	**1.14**	CA	**12.25**	**12.15**	28.47	17.94	18.04	17.92
DE	0.08	**0.15**	0.05	**0.13**	0.08	0.07	DE	**16.44**	**15.79**	23.86	20.72	25.12	24.74
UK	**0.54**	**0.53**	0.40	0.34	0.40	0.28	UK	**25.69**	**26.47**	34.11	26.65	27.98	28.54
JP	*-0.16*	**0.18**	**0.32**	0.08	*-0.15*	*-0.23*	JP	16.36	11.80	**11.10**	**11.23**	13.80	14.09
SE	**0.44**	0.25	**0.56**	0.31	0.42	0.40	SE	49.77	**47.83**	55.61	**49.18**	51.47	53.48
US	0.18	**0.52**	**0.59**	0.49	0.32	0.30	US	**23.19**	**20.57**	25.63	24.12	27.28	26.81
Mean							**Skew.**						
AU	-9.36	-0.18	**6.24**	-8.10	**2.91**	-7.40	AU	0.22	0.35	0.45	0.37	0.63	0.14
CA	13.92	11.89	1.92	16.95	**21.76**	**20.43**	CA	1.02	0.22	1.68	0.89	0.66	0.77
DE	1.38	**2.39**	1.11	**2.75**	2.01	1.75	DE	0.97	-1.00	-0.36	0.74	2.36	2.19
UK	**13.93**	**14.01**	13.61	9.13	11.12	8.00	UK	0.60	-0.52	-0.25	0.32	0.08	0.23
JP	-2.63	**2.13**	**3.51**	0.89	-2.12	-3.22	JP	0.53	0.17	0.17	0.20	-0.02	-0.31
SE	**22.03**	12.06	**31.24**	15.49	21.72	21.62	SE	4.81	5.78	3.71	5.04	3.37	3.12
US	4.22	10.65	**15.18**	**11.82**	8.84	8.15	US	0.35	0.94	0.96	1.06	0.70	0.72
VaR$_{.95}$							**D.D.**						
AU	**-14.41**	-15.45	-16.48	-14.59	**-10.91**	-14.80	AU	53.38	40.71	81.03	43.05	**29.69**	**34.12**
CA	**-3.83**	**-4.31**	-9.34	-6.55	-5.81	-5.89	CA	**11.76**	**15.06**	68.15	20.54	17.25	16.96
DE	**-6.67**	**-7.91**	-11.40	-9.39	-10.09	-11.33	DE	59.70	70.65	69.37	61.77	**56.31**	**56.52**
UK	-10.41	**-7.16**	-13.79	**-10.01**	-12.15	-12.58	UK	**40.98**	**55.07**	80.02	75.36	77.37	79.84
JP	-8.14	**-5.56**	**-4.95**	-5.65	-7.12	-7.18	JP	76.76	**47.83**	**36.98**	49.85	66.41	69.19
SE	**-12.25**	**-12.70**	-15.67	-13.91	-15.88	-19.43	SE	**65.53**	71.09	69.80	**63.68**	70.74	70.61
US	**-9.55**	**-6.79**	-11.05	-10.23	-11.20	-11.19	US	89.08	**63.48**	**53.44**	70.15	84.04	85.00

Table 4: Generic yield curve arbitrage strategy excess returns of different yield curve fitting models across countries. The strategy longs one duration of "cheap bonds" and shorts one duration of "expensive" bonds along the yield curve. The two upper panels report the annualized Sharpe ratios (Sharpe) and the annualized volatilities (Vol.), the two middle panels contain the annualized mean returns (Mean) and the skewness (Skew.) of the returns, and the two bottom panels are the empirical value-at-risk with confidence level of 0.95 (VaR$_{.95}$) and the maximum drawdown (D.D.) of the returns. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point. For the Sharpe ratios, the volatilities, and the mean returns, the two best-performing models for each country-property pair are marked bold face. Negative Sharpe ratios are marked italic.

The results for Australia, Germany, and Japan are weaker compared to the other countries. For Japan, this could be explained by the lower returns and volatilities. As Japan has one of the most mature government bond markets in the world, it is reasonable that the arbitrage opportunities do not occur as often in Japan as in the other countries, and the magnitude is probably smaller when they do occur. For Australia and Germany, this is more likely due to the difficulties in identifying the rich and expensive points along the yield curve since the performance of all models is not satisfactory.

Another aspect of the strategy to examine is the risk performance, which are evaluated in terms of value-at-risk (VaR) and drawdown analysis in the bottom panels of Table 4. The VaR statistic is the empirical VaR with confidence level of 0.95 (VaR$_{.95}$) of the returns and resembles a lower bound for the monthly returns. The maximum drawdown, on the other hand, measures the maximum loss the strategy has experienced. The risk results show that the strategy based on the Nelson-Siegel models in general generates less risky returns. Also, the VaR values show that the riskier countries for an arbitrage strategy are Austalia and Sweden, and the drawdowns indicate that the huge portfolio losses are likely to happen to Germany, the U.K., Sweden, and the U.S., although it is also plausible that strategies with more radical performance are also more likely to have larger losses while moderate strategies do not suffer from huge losses but also are not very profitable. These results together point out that Sweden is the more volatile and riskier market for fixed-income arbitrageurs.

Based on the information in Table 4, especially the Sharpe ratios, and Figure 17, the candidate models for each country are as follows: NS3 and KS1 are the only models that generate profit over time for Australia, however since NS3 is numerically unstable for countries with small sample size like Australia, it is less favorable than KS1; for Canada, while KS1 and KS2 are with advantage, the performance of all the other models but NS3 is also satisfying; for Germany, NS2 is the better choice followed by BS; for the U.K., NS1 and NS2 perform better while NS3 and KS1 are comparable, too; NS3, NS2, and BS achieved positive Sharpe ratios and are chosen for Japan; for Sweden, NS3 is the best performing model, but for a similar reason as selecting KS1 for Australia, NS1, KS1, and KS2 are selected for Sweden; and finally for the U.S., NS2, NS3, and BS perform much better. These results show that the four factor Nelson-Siegel models are more widely applicable for different market settings, giving both more profitable and less risky performance.

To further understand if there is estimation bias towards under- or over-valued bonds, I decomposed the generic strategy portfolio into a long-only portfolio and a short-only portfolio. Intuitively, the under-valued bonds are more attractive as they have higher values, and the contrary holds for the over-valued bonds. Thus if the market portfolio stands for a preference neutral portfolio of the market as a whole, the performance of the long-only portfolio should be better than the market and the short-only portfolio should perform slightly worse. In order to make the portfolios comparable, I scaled the market portfolio by the market duration such that the portfolio has one duration as the long- and short- portfolios. The results of this analysis are presented in Table 5.

Short							Market							Long
	NS1	NS2	NS3	BS	KS1	KS2		NS1	NS2	NS3	BS	KS1	KS2	
AU	**0.09**	**0.09**	**0.09**	**0.11**	**0.02**	**0.11**	0.13	0.06	0.10	0.12	0.08	0.03	0.09	AU
CA	**0.24**	**0.23**	0.25	**0.24**	**0.24**	**0.24**	0.24	**0.29**	**0.28**	**0.27**	**0.29**	**0.30**	**0.30**	CA
DE	0.17	0.17	0.17	**0.17**	0.18	0.17	0.17	**0.18**	**0.18**	**0.19**	**0.18**	**0.17**	0.17	DE
UK	**0.14**	**0.14**	**0.12**	**0.14**	**0.14**	**0.14**	0.16	**0.19**	**0.19**	**0.16**	**0.17**	**0.18**	0.15	UK
JP	**0.22**	**0.22**	**0.21**	**0.23**	**0.24**	**0.25**	0.26	0.24	0.25	0.25	0.25	0.24	0.23	JP
SE	**0.18**	0.19	0.19	0.20	**0.16**	**0.18**	0.18	**0.25**	**0.22**	**0.29**	**0.24**	**0.23**	**0.24**	SE
US	**0.17**	**0.17**	**0.16**	**0.17**	**0.17**	**0.17**	0.18	**0.19**	**0.20**	**0.21**	**0.20**	**0.20**	**0.20**	US

Table 5: Sharpe ratio of the short-only portfolio (the bonds and the corresponding positions are decided by the short part of generic strategy portfolio), market portfolio, and long-only portfolio (the bonds and the corresponding positions are decided by the long part of generic strategy portfolio) across countries. For the short-only portfolio, the performance worse than the market portfolio is market bold and for the long-only portfolio, the performance better than the market portfolio is market bold.

Here one can observe the weaker performance of the strategy in Australia, Germany, and Japan again: identifying the overpricing is difficult for Germany while identifying the underpricing has troubles in Australia and Japan. The strategy in the other countries, on the other hand, forms long portfolios that perform better than the market and short portfolios performing no better than the market, indicating that the models can identify both the overpriced and underpriced bonds.

Subsample results: 2007-2009

The three years that this subsample consists of are the years when the financial crisis of 2008 has the most impact on. As mentioned earlier, Hu et al. (2013) find that pricing errors are larger during crisis, therefore one would expect the results of the arbitrage strategy to be stronger in terms of returns and volatilities with larger magnitudes during this sample period. Table 6 confirms this result. When looking at excess returns, the majority of the country-model pairs achieved higher Sharpe ratios with the cost of a higher volatility in most cases. Figure 17 also shows that the strategy achieves large returns together with more volatilities across countries during this period.

Another difference can be observed is the negative skewness, which are not the case for the full sample results. This indicates that the returns are more likely to be smaller than the mean value and and the downside risk is higher, which is in line with the background that this subsample period is during the crisis when the market is volatile. Note that for Australia, the KS2 model identifies all the bonds as expensive during this period and thus did not set up any trades. The other model specifications also enter trade for very limited number of times and the results for Australia are therefore less reliable.

	NS1	NS2	NS3	BS	KS1	KS2		NS1	NS2	NS3	BS	KS1	KS2
Sharpe							**Vol.**						
AU	0.10	**2.23**	1.44	0.48	**2.97**	-	AU	**37.21**	64.64	55.80	**55.46**	55.81	-
CA	1.71	0.77	1.04	1.55	**1.99**	**1.92**	CA	**12.23**	**13.33**	40.12	18.25	19.19	19.85
DE	0.36	*-0.55*	0.20	**0.64**	**0.54**	0.35	DE	**25.27**	**26.00**	40.30	34.48	47.07	46.26
UK	*-0.03*	0.08	0.21	0.85	**0.87**	**0.98**	UK	38.38	42.64	54.68	35.68	**34.38**	**34.29**
JP	0.41	**0.96**	0.66	**0.83**	0.25	0.15	JP	21.25	16.83	**11.17**	**15.37**	19.36	20.58
SE	**0.58**	0.00	**0.01**	*-0.27*	*-0.39*	*-0.14*	SE	**27.78**	30.12	36.18	**24.93**	47.34	54.84
US	0.06	0.41	**1.01**	**0.97**	0.50	0.30	US	34.21	26.33	**18.70**	20.57	**20.30**	22.37
Mean							**Skew.**						
AU	3.73	**144.40**	80.13	26.62	**165.75**	-	AU	1.24	-0.09	0.20	0.94	-0.55	-
CA	20.87	10.26	**41.87**	28.25	**38.16**	38.07	CA	1.10	0.16	3.40	1.30	0.84	0.97
DE	9.13	-14.34	8.19	**22.14**	**25.51**	15.98	DE	1.65	-0.91	-0.22	0.85	2.03	2.00
UK	-0.96	3.32	11.29	**30.30**	29.89	**33.66**	UK	-0.23	-1.73	-1.01	-0.83	-1.46	-1.31
JP	8.69	**16.15**	7.41	**12.84**	4.85	3.04	JP	-0.08	-0.03	-0.11	0.44	0.19	-0.26
SE	**16.05**	0.04	**0.37**	-6.72	-18.44	-7.56	SE	-0.68	-0.25	-0.41	-1.28	-1.34	-0.75
US	2.19	10.76	**18.89**	**19.88**	10.06	6.65	US	-2.19	-1.61	-0.18	-0.28	-0.78	-1.57
$\text{VaR}_{.95}$							**D.D.**						
AU	-9.75	**-9.16**	-13.00	-9.97	**-7.83**	-	AU	13.72	**9.16**	13.00	9.97	**7.83**	-
CA	**-4.72**	-7.90	-5.93	-7.03	**-5.41**	-5.57	CA	**7.19**	15.06	10.49	10.77	**8.12**	8.52
DE	**-14.56**	-18.47	-28.55	**-18.34**	-22.10	-20.69	DE	**31.90**	55.33	38.54	32.14	**28.52**	41.99
UK	**-31.50**	-43.77	-43.18	**-30.96**	-33.70	-31.52	UK	39.51	55.07	70.63	37.37	**35.92**	**33.39**
JP	-14.88	-9.34	**-5.67**	**-8.71**	-10.52	-15.51	JP	14.88	11.35	**9.06**	**11.22**	22.48	25.73
SE	**-24.59**	-25.24	**-24.00**	-25.66	-49.24	-49.24	SE	**41.78**	50.62	**40.42**	44.56	57.00	57.00
US	-43.17	-30.05	**-11.05**	**-15.60**	-17.98	-24.68	US	65.26	46.73	**17.16**	**27.39**	35.44	40.75

Table 6: Generic yield curve arbitrage strategy excess returns of different yield curve fitting models across countries, from January, 2007 to September, 2009. The strategy longs one duration of "cheap bonds" and shorts one duration of "expensive" bonds along the yield curve. The two upper panels report the annualized Sharpe ratios (Sharpe) and the annualized volatilities (Vol.), the two middle panels contain the annualized mean returns (Mean) and the skewness (Skew.) of the returns, and the two bottom panels are the empirical value-at-risk with confidence level of 0.95 ($\text{VaR}_{.95}$) and the maximum drawdown (D.D.) of the returns. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point. For the Sharpe ratios, the volatilities, and the mean returns, the two best-performing models for each country-property pair are marked bold face. Negative Sharpe ratios are marked italic.

For this subsample, Canada remains the most profitable country. Japan and the U.S. with the best performing models, on the other hand, have comparable performance. While Australia and Sweden continues to be more volatile, the increased volatilities make the the U.K. one of the most volatile markets. Looking at the risk measures, the U.K. and Sweden are the markets that face larger downside risks. From Figure 17, most countries experience a drop around 2008 when the crisis hit, but the magnitude of this drop of the selected model is often smaller. The subsample results also show that in most cases, the better performing models according to the full sample results can achieve better subsample performance, too. However, compared to the full sample results, the non-parametric methods are preferred more often, indicating an arbitrage strategy based on non-parametric methods is more profitable when the market is turbulent.

Subsample results: 2012-2016

This subsample consists the most recent five years in the data set, thus the performance of the generic yield curve models under a quiet market setting that is similar to the current economic environment is examined. During this period, the interest rates for almost all countries are kept at a low level, making government bonds a more attractive alternative for investing in cash, thus one would expect the arbitrage opportunities are more driven away during this period, which is also one of the conclusions of Hu et al. (2013). Also, the interest rates and the term structure are relatively stable over the past few years (see Figure 6). One reasonable assumption would then be that the investments on government bonds are less volatile, as one of the main drives determining the returns for holding bonds, the yield curve changes, do not happen as often.

	NS1	NS2	NS3	BS	KS1	KS2		NS1	NS2	NS3	BS	KS1	KS2
Sharpe							**Vol.**						
AU	*-0.25*	**0.02**	*-0.80*	*-0.47*	*-0.18*	*-0.17*	AU	24.96	**20.28**	35.25	25.62	**20.32**	23.96
CA	0.63	1.07	*-0.27*	0.62	**1.37**	**1.23**	CA	**9.86**	**9.84**	22.51	10.42	11.08	10.67
DE	*-0.69*	**0.68**	*-0.37*	*-0.48*	*-0.41*	*-0.36*	DE	9.04	**6.64**	11.22	**6.67**	7.85	8.44
UK	**0.24**	**0.17**	*-0.21*	*-0.90*	*-0.73*	*-0.87*	UK	**9.84**	**9.13**	16.59	14.27	15.01	14.62
JP	*-1.24*	*-0.87*	*-0.36*	*-0.83*	*-1.15*	*-1.16*	JP	15.27	**7.92**	**4.46**	8.54	10.50	10.58
SE	0.61	0.44	**0.94**	0.25	**0.82**	0.77	SE	**24.02**	**14.06**	47.54	27.05	36.01	36.08
US	*-0.18*	*-0.09*	*-0.22*	*-0.23*	*-0.58*	*-0.58*	US	17.04	**11.43**	17.34	12.39	11.92	**11.46**
Mean							**Skew.**						
AU	-6.36	**0.35**	-28.26	-12.07	**-3.71**	-4.10	AU	0.69	-0.47	0.93	0.30	0.37	0.69
CA	6.25	10.55	-6.10	6.42	**15.19**	**13.09**	CA	1.15	0.31	-0.15	1.02	0.94	1.02
DE	-6.24	**4.49**	-4.14	-3.22	-3.22	**-3.00**	DE	0.51	1.39	-1.34	0.15	-0.91	-1.13
UK	**2.34**	**1.56**	-3.49	-12.86	-11.02	-12.67	UK	-0.39	-0.15	0.10	-0.44	-0.13	-0.05
JP	-18.93	**-6.86**	**-1.59**	-7.06	-12.10	-12.28	JP	1.38	0.59	0.52	-0.34	-0.74	-0.75
SE	14.64	6.19	**44.68**	6.77	**29.54**	27.78	SE	0.76	1.42	1.06	0.64	0.92	0.93
US	-3.11	**-1.08**	-3.81	-2.90	-6.86	-6.69	US	0.39	-0.32	-0.70	-0.37	-0.22	-0.38
VaR$_{.95}$							**D.D.**						
AU	-12.93	-13.43	-19.05	-14.03	**-12.39**	**-12.43**	AU	53.38	**31.81**	81.03	43.05	**29.69**	34.12
CA	-4.13	-4.54	-13.59	-5.57	**-3.70**	**-3.31**	CA	11.26	9.20	53.75	16.93	**7.76**	**6.74**
DE	-5.26	**-2.36**	-5.80	-3.92	-3.99	**-3.87**	DE	31.84	**6.80**	**20.36**	23.05	21.98	20.86
UK	**-5.10**	**-5.80**	-9.14	-9.30	-10.26	-10.32	UK	**18.09**	**11.33**	39.04	53.13	49.39	53.24
JP	-8.01	**-5.30**	**-3.22**	-6.23	-7.47	-7.52	JP	64.69	**31.82**	11.68	32.31	48.00	48.35
SE	**-11.40**	**-6.23**	-19.93	-14.69	-19.08	-19.43	SE	34.50	**12.97**	**44.85**	47.92	43.14	**32.65**
US	-8.82	**-6.66**	-13.47	-8.31	**-7.26**	-7.92	US	39.93	**27.17**	47.23	**31.71**	41.25	39.52

Table 7: Generic yield curve arbitrage strategy excess returns of different yield curve fitting models across countries, from January, 2012 to September, 2016. The strategy longs one duration of "cheap bonds" and shorts one duration of "expensive" bonds along the yield curve. The two upper panels report the annualized Sharpe ratios (Sharpe) and the annualized volatilities, the two middle panels contain the annualized mean returns and the skewness of the returns, and the two bottom panels are the empirical value-at-risk with confidence level of 0.95 (VaR$_{.95}$) and the maximum drawdown (D.D.) of the returns. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point. For the Sharpe ratios, the volatilities, and the mean returns, the two best-performing models for each country-property pair are marked bold face.

Table 7 presents the results of the sabsample. Indeed, the average returns of the strategies are in most cases much lower compared to the full sample results or the results of 2007 to 2009, if not taking negative values. This indicates that the magnitude of the the arbitrage opportunities are small and identifying the arbitrage opportunities are in general more difficult. As expected, the strategy volatilities are also lower in most cases. These return characteristics together result in a weaker performance across countries and model specifications. This can also be observed from Figure 17 that the general trend of the strategy performance for most of the countries is going downwards, except for Canada and Sweden, which are the countries that still have strong results in this subsample. Therefore it is sensible to conclude that this subsample period with a rather stable market setting does not contribute much to the performance of the yield curve arbitrage strategy. The risk measures also confirm that the performance of the strategies are moderate during subsample period, which also means that the downside risk an investor bears is lower. Again here Australia, the U.K., and Sweden appear to be riskier than the other markets when looking at the VaR statistics.

Regarding the robustness of the model performance over this sample period, for most countries the best performing model based on the full sample performance is doing well. Nevertheless, the subsample results suggest that the term structure of the countries might have gone through some changes, so as the feasible model specifications. For Australia and Germany, NS2 becomes the only model that achieves a positive Sharpe ratio, and for Japan and the U.S., none of the models is able to make a profit. These results show that special treatment and further study can be done for the structural changes of the yield curves.

4.2.2 Risk-adjusted performance

The risk-adjusted returns might be an even more important criterion for the model selection process. Controlling for other risk factors not only allows a better understanding of the drive for the returns of the strategy, but also distinguishes the models that actually profit from the mispricings from the ones mimic other risk factors or simply bet on luck.

A common risk factor in fixed-income literature is the market portfolio, the return of which is constructed as a market value weighted average of the return of all the bonds in the market. On top of this, a slope and a curvature factor are also formed. These two factors are considered since for government bond holders, the returns in the short run are particularly influenced by yield curve changes, which are mostly reflected in the shape changes including slope changes and curvature changes, and the up and down shifts of the yield curve. The latter is covered by setting both long and short positions but not yet the former. The construction of the slope and curvature factor follows the principle of Fama and French (1992) with a similar approach as in

Duarte et al. (2006). For the slope factor, the returns are calculated as the difference of the market value weighted average return of the bonds maturing between 10 to 15 years[21], resembling a long-term bond's return, and the market value weighted average return of the bonds maturing between one to three years, resembling a short-term bond's return. This factor reflects a fixed-income strategy that profits from yield curve slope changes. Analogously for the curvature factor that captures the benefits of yield curve curvature changes, the returns are calculated as the difference of the market value weighted average return of the bonds maturing between seven to 10 years, resembling a medium-term bond's return, and the sum of the market value weighted average return of the bonds maturing between one to three years and between 10 to 15 years, resembling a short-term and a long-term bond's return. To make the slope and curvature factor comparable to the arbitrage portfolio returns, the positions of the slope and curvature portfolio are also scaled to one duration long and one duration short only, and the portfolios themselves are duration neutral.

Market				**Slope**				**Curvature**			
	Mean	Vol.	t-score		Mean	Vol.	t-score		Mean	Vol.	t-score
AU	1.42	3.53	1.63	AU	0.19	0.55	1.39	AU	**0.13**	0.21	2.50
CA	**2.83**	3.52	3.25	CA	-0.10	0.55	-0.74	CA	0.04	0.34	0.50
DE	**2.26**	3.56	2.68	DE	0.13	0.56	0.97	DE	**0.15**	0.31	2.03
UK	**2.82**	5.66	2.13	UK	-0.04	0.61	-0.25	UK	0.07	0.24	1.22
JP	**1.76**	1.86	3.82	JP	**0.19**	0.31	2.54	JP	**0.13**	0.21	2.55
SE	**2.57**	3.70	2.97	SE	-0.03	0.58	-0.24	SE	0.07	0.30	1.04
US	**2.09**	3.48	3.19	US	-0.03	0.62	-0.25	US	0.07	0.39	0.97

Table 8: Characteristics of the excess returns of the market portfolio (left panel), slope portfolio (middle panel), and curvature portfolio (right panel). The mean and volatility of the excess returns are reported in annualized percentage. The t-statistics are of the coefficients of the regressing the portfolio excess returns on a constant, $r_t = \beta + \epsilon_t$, where r_t stands for the monthly portfolio excess return in percentage. The reference critical value is 1.97 for a confidence level of 0.95 and 1.65 for a confidence level of 0.90. The mean excess returns significantly different from zero at a confidence level of 0.95 are marked boldface.

Table 8 summarizes the characteristics of the three risk factors considered. The market portfolio has strong performance for all countries but Australia, which rises doubt that the bad performance of the arbitrage strategy for Australia can be tracked back to that the Australian government bond market does not generate much profit itself. Positioning on the yield curve in general is not very profitable for most countries. Riding on the slope of the yield curve gives significant returns only in Japan, and a strategy betting on the yield curve curvature changes

[21] In case there is no bond with maturity of 10 to 15 years (for example, Canada in the early 2000's), bonds maturing between seven to 10 years are used instead.)

generates significant returns for Australia, Germany, and Japan. These statistics show that a yield curve strategy conducted in Japan is likely to be exposed to the risk factors.

Next the risk factors are used to test the risk-adjusted performance of the arbitrage portfolios. Following Fama and French (1992), the excess returns of the strategies are regressed on the excess returns of the risk factor based portfolios. The regression used for the analysis is then

$$\mathrm{R_a} = \alpha + \beta_\mathrm{m}\mathrm{R_m} + \beta_\mathrm{s}\mathrm{R_s} + \beta_\mathrm{c}\mathrm{R_c} + \varepsilon \tag{4.1}$$

where $\mathrm{R_a}$, $\mathrm{R_m}$, $\mathrm{R_s}$, $\mathrm{R_c}$ stand for the excess returns of the arbitrage strategy, market portfolio, slope portfolio, and curvature portfolio, respectively, and ε is the residual term of the regresion.

The results of the regression analysis are reported in Table 9. The bad performance of the strategy for Australia, Germany, and Japan can also be observed from the risk-adjusted returns that after accounting for other risk factors, the arbitrage factor barely explains the strategy return. As none of the models achieves a significant alpha, it might worth trying other model specifications for Australia and Japan for further research. The high profitability of the arbitrage strategy in Canada is also presented here such that the alphas of most arbitrage portfolios based on different yield curve models are positive and significant.

For the majority of the country-model pairs, following the market (oppositely) can explain the arbitrage portfolio performance. The arbitrage strategy returns have a significant negative loading on the slope factor for most countries regardless of the model setting except for Sweden, indicating that the strategy for most countries sits on the short and/or short end of the yield curve. However, since the slope portfolio returns are only significant from zero for Japan, for the majority of the countries, positioning oppositely as the slope factor does not have much influence for the performance of the arbitrage portfolio. The loadings on the curvature factor are significantly positive for Canada, the U.K., and Japan, meaning that the arbitrage portfolio for these three countries also turns to position following the yield curve curvature while for Germany, Sweden, and the U.S., the positions are opposite to the curvature. As the curvature portfolio returns are significant for Germany and Japan, they also have an influence on the stand-alone arbitrage factor. These results also show that a yield curve arbitrage strategy can be improved by choosing strategy specifications that reduce exposure to the risk factors to further benefit from the arbitrage opportunities.

		α	R_m	R_s	R_c		α	R_m	R_s	R_c		α	R_m	R_s	R_c		α	R_m	R_s	R_c
AU						**CA**					**DE**					**UK**				
coef.	NS1	-0.06	**-1.35**	**-29.73**	0.80	NS1	**1.12**	0.26	-1.76	-10.41	NS1	0.40	-0.06	**-8.83**	**-14.07**	NS1	**1.19**	-0.10	-0.79	-1.99
t-score		-0.11	-2.24	-6.30	0.05		4.53	0.84	-0.48	-1.60		1.40	-0.19	-3.02	-2.37		2.29	-0.26	-0.22	-0.21
coef.	NS2	0.81	-1.20	**-24.96**	-19.70	NS2	**0.98**	0.18	2.04	-2.75	NS2	0.35	-0.29	**9.33**	**-16.53**	NS2	**1.13**	-0.26	**-14.51**	13.87
t-score		1.35	-1.76	-4.73	-1.17		3.75	0.56	0.53	-0.40		1.12	-0.79	2.87	-2.51		2.21	-0.69	-4.03	1.52
coef.	NS3	1.43	-1.44	**-17.54**	-37.63	NS3	0.23	**-2.20**	**-36.47**	**40.68**	NS3	0.51	-0.20	-0.05	**-31.06**	NS3	0.82	**1.81**	**-22.39**	-17.88
t-score		1.68	-1.59	-2.28	-1.51		0.41	-3.17	-4.42	2.77		1.15	-0.38	-0.01	-3.34		1.29	3.94	-5.08	-1.61
coef.	BS	0.75	**-1.35**	**-40.62**	-15.14	BS	**0.87**	0.86	-33.04	18.83	BS	**0.55**	**0.99**	**-22.81**	**-19.23**	BS	0.53	-0.08	**-35.89**	**33.28**
t-score		1.68	-2.76	-10.00	-1.16		3.30	2.63	-8.54	2.73		2.13	3.26	-8.52	-3.55		1.38	-0.29	-13.39	4.91
coef.	KS1	0.67	0.12	**-41.25**	6.93	KS1	**1.23**	**1.01**	**-29.94**	**25.88**	KS1	0.55	**1.30**	**-25.90**	**-25.91**	KS1	0.68	-0.11	**-31.83**	**42.09**
t-score		1.42	0.21	-9.61	0.50		3.94	2.62	-6.51	3.16		1.71	3.42	-7.74	-3.83		1.51	-0.34	-10.15	5.27
coef.	KS2	0.51	-0.46	**-37.17**	**-31.27**	KS2	**1.11**	**0.93**	**-32.15**	**28.46**	KS2	0.48	**1.29**	**-27.05**	**-20.96**	KS2	0.51	0.30	**-38.28**	**38.88**
t-score		1.21	-0.92	-8.16	-2.20		3.69	2.48	-7.24	3.59		1.48	3.35	-7.98	-3.05		1.18	0.96	-12.60	5.09
JP						**SE**					**US**									
coef.	NS1	0.28	**1.79**	**-54.03**	**12.63**	NS1	**2.04**	0.33	**-15.38**	**-53.33**	NS1	0.48	0.72	**7.34**	**-37.93**					
t-score		1.26	2.87	-13.54	2.31		2.26	0.36	-2.10	-3.68		1.44	1.67	2.05	-5.85					
coef.	NS2	0.25	**2.88**	**-47.42**	**26.91**	NS2	1.24	-0.07	-9.64	**-45.44**	NS2	**0.83**	**1.06**	0.20	**-19.12**					
t-score		1.49	6.15	-15.78	6.54		1.39	-0.08	-1.34	-3.19		2.63	2.60	0.06	-3.14					
coef.	NS3	0.31	**1.84**	**-19.90**	3.52	NS3	**2.99**	-0.69	0.51	-32.98	NS3	**0.94**	**1.55**	**-25.22**	3.74					
t-score		1.40	2.92	-4.91	0.64		2.67	-0.60	0.05	-1.78		2.72	3.49	-6.84	0.56					
coef.	BS	0.26	**1.86**	**-45.94**	**28.00**	BS	1.49	0.98	3.41	**-75.52**	BS	**0.69**	**1.68**	**-21.73**	-3.47					
t-score		1.80	4.60	-17.72	7.89		1.62	1.05	0.46	-5.13		2.17	4.12	-6.40	-0.56					
coef.	KS1	0.06	**1.29**	**-54.45**	**43.62**	KS1	1.89	1.42	16.36	**-59.64**	KS1	0.53	**1.36**	**-23.51**	-8.97					
t-score		0.35	2.50	-16.49	9.65		1.87	1.33	1.88	-3.67		1.51	3.02	-6.29	-1.33					
coef.	KS2	-0.07	**1.37**	**-56.30**	**48.89**	KS2	1.93	1.29	16.70	**-63.83**	KS2	0.48	**1.20**	**-22.52**	-6.77					
t-score		-0.36	2.61	-16.76	10.63		1.76	1.10	1.77	-3.69		1.36	2.64	-5.94	-0.99					

Table 9: Estimated coefficients and corresponding t-statistic for the regression of monthly percentage excess returns of the generic yield curve arbitrage strategy on the the other risk factors (see (4.1)). R_m, R_S, and R_C stand for the market portfolio, slope factor, and curvature factor, respectively. The α coefficient values are in basis point and the β coefficient values are in percentage. The reference critical value is 1.97 for a confidence level of 0.95 and 1.65 for a confidence level of 0.90. The coefficients significant at a confidence level of 0.95 are marked bold-face.

On top of the indicating the positioning of the arbitrage portfolio and the influence to the arbitrage factor performance, the loadings on the slope and curvature portfolios can also give indications on the estimation bias of the term structure. The significantly negative slope factor coefficients suggest that the short end of the yield curve for most countries are likely to be estimated as lying above the actual spot rates, resulting in the short-term bonds being considered as "cheap" and/or the long term bonds considered "expensive". Similarly, the significantly positive loadings on the curvature portfolio implies the medium-term bonds are likely to be estimated as under-valued while the long- and/or short-term bonds are "over-priced" for Canada, the U.K., and Japan, and for Germany, Sweden, and the U.K., the opposite is likely to hold.

4.3 Model selection

Although the pattern of the model performance based on the full sample arbitrage strategy results are robust across subsample periods for Canada and Sweden, there are some differences worth noticing for Australia, Germany, the U.K., Japan, and the U.S. For Canada and Sweden, the strategy returns over different samples all agree that KS1 and NS1 are the better specifications, respectively. And since these two models are also able to generate a significant alpha for Canada and Sweden, they are selected as the most feasible models for Canada and Sweden.

The selected model for Australia is KS1, the subsample results of 2012 to 2016, on the contrary, show this models is not desired but NS2 is. Since none of the models gives a significant risk premium for the arbitrage factor, it is more difficult to draw a conclusion if NS2 is really more feasible than KS1, thus I follow the full sample results and do not change the model specification; for Germany, BS is the model preferred by the full sample results and the subsample results of 2007 to 2009, but the only model performs well during 2012 to 2016 is NS2. As discussed earlier, the period 2012 to 2016 barely attributes to the arbitrage strategy performance. Also since BS achieves a significant risk-premium, BS is selected for Germany. For Japan and the U.S., while NS2 and NS3 are the preferred specification based on the full sample results and the subsample results of 2007 to 2009, all the models perform bad during the most recent years. For Japan, none of the models gives significant risk-adjusted returns, I follow the strategy returns' results and select NS2. For the U.S., NS3 is chosen for that is gives a significant alpha. For the U.K., the only model seletced based on full sample results that performs well over the two subsample periods is NS2. As it is also one of the two specifications with a significant alpha, NS2 is selected for the U.K. These results indicate structural changes in the yield curves of Australia, Germany, the U.K., and Japan might worth noticing and further research.

To conclude, the final model selection for each country is as follows

	AU	CA	DE	UK	JP	SE	US
Best performing	KS1	KS1	BS	NS2	NS2	NS1	NS3

Table 10: Final model selection for each country based on the profitability of the strategy, the significance of the arbitrage factor, and the robustness across subsamples.

The final selection of models shows that the Nelson-Siegel type of models are preferred by the majority of the countries except Australia. For Canada, although KS1 is selected, NS1 and NS2 are also able to achieve very comparable results. Within the Nelson-Siegel family models, clear advantage is shown by the four-factor models, especially for countries with more developed government bond markets.

In order to examine how sensitive these results are to the weighting scheme of the strategy, I also formed similar trading strategies but with equal weighting and pricing error weighting. The results show that the selected models also perform well in these strategy settings. Detailed results can be found in the Appendix. The paper continues with the selected model for each country to further explore the profitability of the yield curve arbitrage strategies.

5 Yield Curve Arbitrage

In this section, the yield curve arbitrage is studied in more details. The possible strategy specifications that can potentially improve upon the generic strategy are discussed in section 5.1 and the results are analyzed in section 5.2.

5.1 Modified trading strategy

In section 3.2, the generic fixed-income arbitrage strategy is constructed with the aim of evaluating the term structure models, the strategy specifications are thus also formed combining all the information from the modeled yield curves, in this way the returns of the generic strategy reflect the characteristics of the models. However, a satisfactory arbitrage strategy should not only be able to capture the arbitrage opportunities, but also limit the exposure to other risk factors. The only risk factor considered by the generic strategy so far is the interest risk. Duarte et al. (2006) form a yield curve arbitrage strategy that positions as a "butterfly" trade, that is, shorting the short and long end and longing the middle part of the yield curve, which to a large extent is analogous to the construction of the curvature factor as describe in section 4.2.2. This trading specification, as reported by Perold (1999), is also a most frequently executed yield curve arbitrage strategy. Nevertheless, it is not clear if the returns of such a "butterfly" trade are due to the arbitrage opportunities or the benefit from positioning on the yield curve's curvature, neither the exposure to other risk factors is fully taken into account and examined.

As discussed earlier, the generic strategy is in most cases largely exposed to market movements as well as yield curve movements, to further improve the strategy, here I consider an alternative strategy specification as described by Beekhuizen et al. (2016) that groups the bonds into maturity buckets instead of looking at individual bonds. Doing so allows better hedging for the positions as the strategy searches for opportunities within a small segment of the yield curve and the market is likely to have less impact. Also the slope and the curvature of the yield curve within each maturity bucket are reduced compared to those of the whole curve, yield curve movements are also expected to influence the strategy performance less.

Thus to construct the maturity bucket based arbitrage trading strategy, in each month, the observed bonds are divided in to six maturity buckets according to their remaining time-to-maturity: 1-3 year, 3-5 year, 5-7 year, 7-10 year, 10-15 year and 15+ year, and within each maturity bucket, the strategy longs all the bonds that have a lower theoretical yield than actual yield and shorts the rest. Both the total portfolio has one unit duration of long positions as well as one unit duration of short positions to make it duration neutral, however, as the trade for a maturity bucket is only formed when both cheap and rich bonds exist, it is possible that for some month, some country only enters trade with two maturity buckets, in this case both the two buckets are assigned with a half year duration of long positions and a half year duration of short positions. Written in mathematical formulation, suppose, in a random month, there are N_i bonds observed in maturity bucket i, the weight for a bond j in the bucket can be decided in a similar manner in section 3.2.2 as

$$w_i = \begin{cases} \frac{1}{N}\frac{1}{D_j}\frac{y_j - \hat{y}_j}{\sum_{k=1}^{N_i}(y_k - \hat{y}_k) I_{\hat{y}_k \geq y_k}} & \text{if } y_j \geq \hat{y}_j \\ -\frac{1}{N}\frac{1}{D_j}\frac{y_j - \hat{y}_j}{\sum_{k=1}^{N_i}(y_k - \hat{y}_k)(1 - I_{y_k \geq \hat{y}_k})} & \text{otherwise} \end{cases} \tag{5.1}$$

where N stands for the number of maturity buckets that enters the portfolio, y_j and $\hat{y}_j$ stand for the actual and theoretical yield of the bonds, D_i is the duration of bond j, and $I_{[\cdot]}$ is an indicator function. The yields are calculated based on the selected models for each country as in section 4.3, (3.29) and (3.30). The portfolios are again set up at the beginning of the month after observing the bond price information and held for a month. This modified strategy is expected to enter trade less often, especially for small markets with limited numbers of bonds like Australia and Sweden. Since the portfolio is formed only when bonds that are identified as cheap and expensive both exist, and searching within maturity buckets are naturally more restrictive than taking all the bonds available into account. Following this argument, one would also expect a lower volatility for a maturity bucket based strategy. Also, as the modified strategy is designed to reduce the impact of yield curve changes, one would expect reduced exposure to the slope and curvature factor. Moreover, as the portfolio now is not only duration neutral as a whole but also duration neutral within each segment of the yield curve, the impact of the market is also expected to be less.

5.2 Results

5.2.1 Returns

The results of the modified yield arbitrage strategy are reported in the right panel of Table 11, and the cumulative returns of the strategy are plotted in Figure 4. For an easy comparison, the results of the generic strategy used for evaluating the yield curve models are also presented in the left panel of Table 11. As expected, for most countries the volatility of the strategy is significantly reduced with the modified strategy settings, which can also be seen from Figure 4 that the return series are more stable for the modified strategy for almost all countries. For Sweden and the U.S., the major spikes that exist in the generic strategy (for Sweden, around 2000[22] and 2005[23]; for the U.S., around 1989[24]) continue showing up in the modified strategy, causing a similar level of volatilities as the generic strategy.

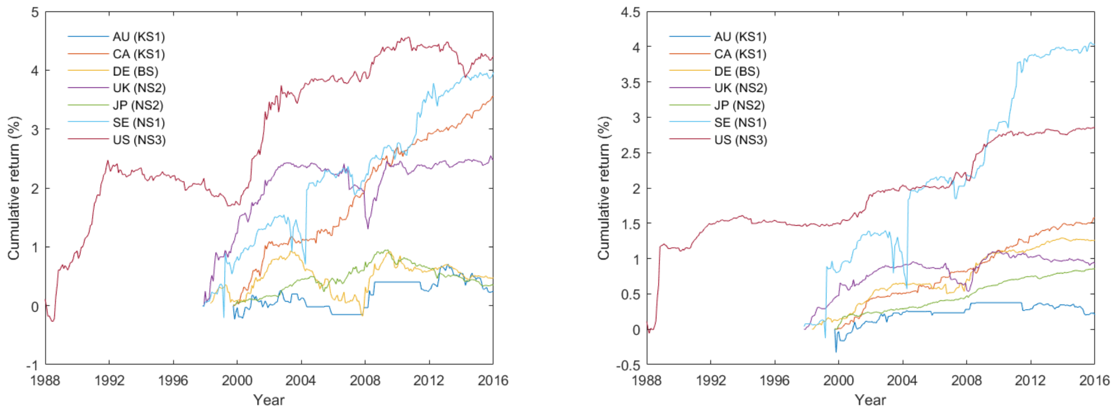
[22]This spike comes from bond “SE0000148454” with month-to-day return of 1.47% for January 2000, both the generic and the modified strategy long 0.73 unit of this bond, making a profit of 0.86%.

[23]This spike comes from bond “SE0000197535” with month-to-day return of 3.79% and bond “SE0000346272” with month-to-day return of 2.41% for February 2005, the generic longs 0.15 unit of “SE0000197535” and 0.30 unit of “SE0000346272”, together making a profit of 1.21%.; the modified strategy longs 0.15 unit of “SE0000197535” and 0.31 unit of “SE0000346272”, together making a profit of 1.24%.

[24]This spike comes from bond “912810BF” with month-to-day return of -16.94% for July 1989, the generic shorts 0.009 unit of “912810BF”, making a profit of 0.17%.; the modified strategy shorts 0.08 unit of “912810BF”, making a profit of 1.48%.

Generic							Modified						
	Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.
AU	0.11	2.91	26.86	0.63	-10.91	29.69	AU	0.09	1.25	13.40	-1.65	-3.74	32.85
CA	1.21	21.76	18.04	0.66	-5.81	17.25	CA	1.18	9.67	8.17	1.45	-2.35	8.12
DE	0.13	2.75	20.72	0.74	-9.39	61.77	DE	0.86	7.12	8.25	0.23	-2.44	14.90
UK	0.53	14.01	26.47	-0.52	-7.16	55.07	UK	0.48	5.06	10.44	1.38	-3.60	35.82
JP	0.18	2.13	11.80	0.17	-5.56	47.83	JP	1.43	5.25	3.66	0.86	-1.08	3.93
SE	0.44	22.03	49.77	4.81	-12.25	65.53	SE	0.49	22.09	44.87	6.17	-12.03	64.81
US	0.59	15.18	25.63	0.96	-11.05	53.44	US	0.68	10.18	14.94	6.88	-2.72	16.24

Table 11: Generic and modified yield curve arbitrage strategy excess returns of different yield curve fitting models across countries. The generic strategy longs one duration of "cheap bonds" and shorts one duration of "expensive" bonds along the yield curve while the modified strategy positions within each maturity bucket. The annualized Sharpe ratios (Sharpe), the annualized mean returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 (VaR$_{.95}$), and the maximum drawdown (D.D.) of the returns. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.



(a) Generic strategy (b) Modified strategy

Figure 4: Cumulative return for generic and the modified strategy across countries. The selected yield curve model for each country is indicated in the legend.

When looking at the Sharpe ratios, the modified strategy is able to achieve at least as good performances as the generic strategy, which is a result of the lowered volatilities together with the more modest returns. From Figure 4, another major improvement of the modified strategy is that it is able to achieve an increase in the cumulative returns after 2012 for most of the countries while the generic strategy is not. The decreased VaR values point out that the modified strategy is also less risky. Japan benefits most from the altered strategy settings; Canada and the U.K., however, end up with a much lower total wealth compared to the generic strategy. These results indicate that the generic strategy benefits from large yield errors that few bonds bear, and that it might worth considering different trading strategy settings for different countries.

Subsample results: 2007-2009

Similar as in section 4, here the subsample analysis is first conducted for the sample period of 2007 to 2009. As shown in Figure 4, this subsample period is rather volatile and both strategies experienced numbers of losses for almost all countries, especially the generic portfolios for the U.K. Compared to the full sample, for most countries, the strategy performance for this subsample is more radical, which is in line with the background that this is a time period when the crisis hit.

Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	$\text{VaR}_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	$\text{VaR}_{.95}$	D.D.
AU	0.10	3.73	37.21	1.24	-7.83	7.83	AU	1.00	14.64	14.59	3.51	-0.02	0.02
CA	0.77	10.26	13.33	0.16	-5.41	8.12	CA	0.47	3.41	7.22	-1.23	-5.84	8.12
DE	0.54	25.51	47.07	2.03	-18.34	32.14	DE	0.85	16.36	19.15	1.00	-13.42	13.42
UK	0.85	30.30	35.68	-0.83	-43.77	55.07	UK	0.36	6.08	16.73	1.39	-10.65	34.53
JP	0.66	7.41	11.17	-0.11	-9.34	11.35	JP	0.94	3.03	3.24	0.37	-2.90	2.90
SE	0.58	16.05	27.78	-0.68	-24.59	41.78	SE	0.61	15.16	24.81	-1.06	-26.46	30.85
US	0.06	2.19	34.21	-2.19	-11.05	17.16	US	0.51	12.00	23.41	-2.11	-10.98	12.23

Table 12: The annualized Sharpe ratios (Sharpe), the annualized mean returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 ($\text{VaR}_{.95}$), and the maximum drawdown (D.D.) of the returns of the generic and modified yield curve arbitrage strategy returns across countries, for the period of January 2007 to December 2009. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.

When comparing the results between the two strategies for this subsample, one can observe that this subsample results exhibit comparable features as the ones for the full sample. The level for the volatilities of the modified strategy are much lower than those of the generic strategy. Moreover, most countries achieved higher Sharpe ratios with the modified strategy. Australia and the U.S. even have higher returns for the modified strategy. The VaR and the maximum drawdowns are also more extreme in most cases for the generic strategy than for the modified strategy, which indicates that the for most countries, the generic strategy specification is more sensitive to market environment changes. Similar to the full sample, for Canada and the U.K. again the generic strategy outperform the modified strategy, while the results of Australia, Japan, and the U.S. is significantly improved.

Subsample results: 2012-2016

For this subsample, the argument in explaining the subsample results for selecting the yield curve models also holds here, that the government bonds are more attractive in the period since the interest rates are kept low by the central banks and investing in cash is no longer desirable, and that the market is rather quiet at the same time. Therefore again one would expect that both the generic and the modified strategy to achieve more moderate returns with lower volatilities; the Sharpe ratios should also be higher as a benefit of a more stable performance.

Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.
AU	-0.25	-6.36	24.96	0.69	-12.39	29.69	AU	-0.27	-2.64	9.73	-3.36	-5.90	17.08
CA	1.07	10.55	9.84	0.31	-3.70	7.76	CA	1.06	6.77	6.40	2.11	-3.02	6.35
DE	-0.41	-3.22	7.85	-0.91	-3.92	23.05	DE	0.27	0.69	2.51	0.19	-2.11	4.15
UK	-0.90	-12.86	14.27	-0.44	-5.80	11.33	UK	-0.79	-4.70	5.93	-0.28	-3.02	16.77
JP	-0.36	-1.59	4.46	0.52	-5.30	31.82	JP	0.84	2.42	2.88	0.97	-1.37	3.93
SE	0.61	14.64	24.02	0.76	-11.40	34.50	SE	0.39	5.43	13.87	-0.27	-7.47	16.15
US	-0.18	-3.11	17.04	0.39	-13.47	47.23	US	0.09	0.37	4.14	0.39	-2.30	7.86

Table 13: The annualized Sharpe ratios (Sharpe), the annualized mean returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 (VaR$_{.95}$), and the maximum drawdown (D.D.) of the returns of the generic and modified yield curve arbitrage strategy returns across countries, for the period of January 2012 to September 2016. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.

The statistics in Table 13 show that, both the generic and modified strategy are with returns at a lower level as well as lower volatilities compared to the full sample and the subsample of 2007 to 2009. The risk measure also indicates that the generic strategy is riskier than the modified strategy even under a rather calm market environment, which confirms that the modified strategy is less sensitive to the changes of market settings. As shown in Figure 4, the generic strategy experiences mainly portfolio losses for most countries, the returns and the Sharpe ratios are therefore also correspondingly negative. The modified strategy is able to generate profit for most of the countries, but the bad performance of Australia and the U.K. is still not improved.

5.2.2 Risk-adjusted returns

The modified strategy is based on maturity buckets and is designed to reduce the exposure of yield curve changes and market movements. To test if this is indeed the case, similar regression analysis as described in section 4.2.2 and (4.1) are conducted with the full sample excess returns of the modified strategy such that the risk factors examined are market portfolios, and slope and curvature of the yield curves. The results of the risk-adjusted returns are presented in Table 14. For the ease of comparison, here again the results of the generic strategy used for evaluating the yield curve models are reported together.

Generic						Modified					
		α	R_m	R_S	R_C			α	R_m	R_S	R_C
AU	coef.	0.67	0.12	**-41.25**	6.93	AU	coef.	0.14	0.04	**-5.54**	4.12
	t-score	1.42	0.21	-9.61	0.50		t-score	0.51	0.15	-2.22	0.60
CA	coef.	**1.23**	**1.01**	**-29.94**	**25.88**	CA	coef.	**0.74**	0.18	-1.49	2.09
	t-score	3.94	2.62	-6.51	3.16		t-score	4.27	0.81	-0.58	0.46
DE	coef.	**0.55**	**0.99**	**-22.81**	**-19.23**	DE	coef.	**0.69**	0.19	**-3.28**	**-8.12**
	t-score	2.13	3.26	-8.52	-3.55		t-score	4.64	1.10	-2.13	-2.59
UK	coef.	**1.13**	-0.26	**-14.51**	13.87	UK	coef.	0.39	0.05	-1.40	2.01
	t-score	2.21	-0.69	-4.03	1.52		t-score	1.89	0.35	-0.99	0.55
JP	coef.	0.25	**2.88**	**-47.42**	**26.91**	JP	coef.	**0.40**	**0.49**	**-2.75**	1.35
	t-score	1.49	6.15	-15.78	6.54		t-score	5.11	2.22	-1.97	0.71
SE	coef.	**2.04**	0.33	**-15.38**	**-53.33**	SE	coef.	**1.98**	0.78	0.79	**-50.53**
	t-score	2.26	0.36	-2.10	-3.68		t-score	2.31	0.89	0.11	-3.69
US	coef.	**0.94**	**1.55**	**-25.22**	3.74	US	coef.	**0.73**	**0.67**	-2.23	-0.19
	t-score	2.72	3.49	-6.84	0.56		t-score	3.07	2.20	-0.88	-0.04

Table 14: Estimated coefficients and corresponding t-statistic for the regression of monthly percentage excess returns on the the other risk factors. R_m, R_S, and R_C stand for the market, slope and curvature portfolio, respectively. The α coefficient values are in basis point and the β coefficient values are in percentage. The reference critical value is 1.97 for a confidence level of 0.95 and 1.65 for a confidence level of 0.90. The coefficients significant at a confidence level of 0.95 are marked bold-face.

First thing comes to notice is that the market exposure is no longer significant for most countries when looking at the modified strategy returns. This was not the case for the generic strategy, which means forming maturity buckets indeed helps with being neutral towards the market movements. Although Japan and the U.S. still load on the market, the coefficient is significantly smaller. One reasonable argument for the significant market effect for Japan and the U.S. is that these two countries are with the highly developed government bond markets and the issuance of long term bonds is frequent. This results in that the 15+ year maturity bucket is a relatively wide maturity division and that the market effect is still present within the bucket.

For most of the countries, the alpha value of the generic strategy is higher than those of the modified strategy, which is reasonable, as the the returns of the modified strategy are at a smaller magnitude. For the U.K., the significant alpha of the generic strategy becomes insignificant when switched to the modified strategy, which is in line with the results in the previous section that the performance of the modified strategy is weaker for the U.K. Although the exposure to the slope factor is eliminated, Japan is still exposed to the yield curve slope changes even with a maturity bucket strategy, which indicates that the yield curve slope changes together with following the market could be the strongest drive of the arbitrage strategy returns for Japan. And finally, Australia has an insignificant risk premium for the arbitrage factor for both the generic and modified strategy, indicating other strategy specifications or yield curve models may be considered for future research.

There results of the risk-adjusted returns also show that country-wise arbitrage strategies might worth investigating. For example, for the U.K. the modified strategy is able to generate a significant risk premium on top of other risk factors while the generic strategy can not, but for Japan it is the other way around.

5.3 Sensitivity analysis

This section tests the profitability of the generic and modified yield curve arbitrage strategy against altered strategy settings by varying the holding period of the strategies and incorporating the transaction costs into the trades. The analysis is based on full sample results, and the subsample results can be found in the Appendix.

5.3.1 Holding period

Duarte et al. (2006) use a holding period of up to 12 month for their yield curve arbitrage strategy, as it often takes longer than one month for a mispricing to converge to zero such that the traded price of a bond converges to the theoretical price. Such a strategy specification is also common seen in practice since it also helps to reduce turnovers. Thus to test how sensitive the strategies are to the changes of the holding period of the portfolios, the same generic and modified yield curve arbitrage strategy are formed with holding period of six months and twelve months. The analysis is conducted based on full sample, and the subsample results are reported in the Appendix.

Holding period: 6 months													
Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.
AU	-0.45	-14.56	32.04	-1.81	-15.38	74.83	AU	-0.28	-4.70	16.86	-6.08	-3.56	65.77
CA	1.15	27.77	24.06	1.59	-6.57	25.56	CA	1.20	1.64	1.37	1.42	-0.38	1.36
DE	-0.31	-6.62	21.37	-1.57	-9.42	85.94	DE	0.72	0.96	1.35	0.17	-0.52	2.66
UK	0.40	8.95	22.21	0.55	-7.46	52.21	UK	0.67	1.37	2.03	1.99	-0.58	7.07
JP	-0.08	-0.91	11.69	0.13	-5.71	53.59	JP	1.28	0.96	0.75	1.41	-0.18	0.86
SE	0.25	10.52	42.27	4.71	-14.12	62.45	SE	0.41	3.82	9.35	2.60	-3.04	18.70
US	0.38	8.40	22.33	1.02	-8.05	51.45	US	0.55	1.97	3.56	6.40	-0.46	7.60
Holding period: 12 months													
Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.
AU	-0.23	-7.38	31.57	-1.54	-13.08	62.94	AU	-0.29	-5.01	17.07	-5.83	-5.73	67.63
CA	0.64	17.54	27.29	0.13	-7.49	57.50	CA	1.23	0.85	0.69	1.34	-0.19	0.68
DE	-0.42	-8.43	20.31	-1.12	-10.72	85.36	DE	0.59	0.47	0.79	-1.09	-0.26	1.66
UK	0.37	7.82	20.95	0.58	-10.04	56.11	UK	0.73	0.93	1.28	3.42	-0.29	3.59
JP	-0.23	-2.63	11.54	0.10	-6.11	55.28	JP	0.99	0.60	0.61	3.00	-0.09	0.86
SE	0.08	3.32	39.42	4.58	-12.80	65.48	SE	-0.12	-1.04	8.68	-6.23	-2.05	42.59
US	0.16	3.44	21.90	0.89	-9.01	57.03	US	0.41	1.16	2.85	7.87	-0.23	7.60

Table 15: The annualized Sharpe ratios (Sharpe), the annualized mean returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 (VaR$_{.95}$), and the maximum drawdown (D.D.) of the returns of the generic and modified yield curve arbitrage strategy returns across countries with holding period of six months (upper panel) and 12 months (lower panel). The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.

The results of the returns of the strategies with altered holding periods are presented in Table 15. For the generic strategy, the Sharpe ratios in general decrease while the time of holding the portfolios increases. This boils down to a deteriorated return with a similar level of volatility such that the return per unit volatility is worsened. On the other hand, the modified strategy in most cases is able to achieve a similar magnitude of Sharpe ratio when the holding period is prolonged, and the decreased returns correspond to the decreased strategy volatilities. For Sweden, however, a longer holding period is not preferred, as the reduction in the volatility does not compensate for the drop in the returns. These results show that the maturity bucket strategy is more robust against holding period changes, which makes it more applicable and attractive than the generic strategy in practice.

5.3.2 Transaction costs

Another relative aspect to examine for a trading strategy to be practically applicable is the profitability after taking into account the transaction costs. This study considers the bid-ask spread as the main source as the transaction costs and incorporates it with a similar approach by Hu et al. (2013). After taking market difference into account, the bid-spread is set to one basis point of the yield of the bonds for the big countries (Canada, Germany, the U.K., Japan, and the U.S.) and two basis points for the small countries (Australia and Sweden). The transaction costs are then incorporated in the following manner: the strategy only sets up a trade when the magnitude of the yield error of the bond is above the bid-ask spread. And if a trade is set up, the return earned is the month-to-day return of the bond minus the transaction cost. The analysis is based on the full sample and subsample results can be found in the Appendix.

Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	$\text{VaR}_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	$\text{VaR}_{.95}$	D.D.
AU	-0.13	-3.73	28.46	0.08	-16.09	32.13	AU	-0.04	-0.48	13.50	-2.26	-4.39	33.70
CA	1.16	20.99	18.03	0.64	-6.04	17.96	CA	1.16	9.54	8.22	1.43	-2.37	8.42
DE	0.00	-0.10	21.12	0.65	-10.15	68.15	DE	0.72	7.26	10.05	0.44	-3.07	21.83
UK	0.46	12.15	26.64	-0.52	-7.51	54.73	UK	0.41	4.59	11.23	1.12	-4.74	37.05
JP	0.09	1.07	12.16	0.15	-5.86	52.84	JP	1.48	5.77	3.90	0.55	-1.38	4.60
SE	0.41	20.91	50.41	4.60	-13.30	66.05	SE	0.55	25.14	45.65	5.87	-12.02	64.09
US	0.55	14.00	25.67	0.91	-11.18	55.71	US	0.65	9.83	15.08	6.69	-2.96	16.20

Table 16: The annualized Sharpe ratios (Sharpe), the annualized mean excess returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 ($\text{VaR}_{.95}$), and the maximum drawdown (D.D.) of the excess returns of the generic and modified yield curve arbitrage strategy returns across countries with transaction cost restrictions. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.

As the strategies already assign weights to the bonds according to the yield error the bonds bear, bonds with smaller yield error naturally have less influence on the portfolio returns, thus one would expect that the transaction costs do not have a large impact on the results, especially for the generic strategy. The results presented in Table 16 confirms this hypothesis. Most countries are able to achieve a similar level of Sharpe ratio and strategy returns as when there is no transaction cost restriction for both the generic strategy and the modified strategy. Weaker results appear for Australia and Germany, where the arbitrage strategy in general does not profit much as discussed in section 4.2. Nevertheless, the risk performance show that the risk level of the generic strategy is reduced while that of the modified strategy remains the same. This indicates that the numerous small positions of the generic strategy do have influence on the strategy performance and the more restricted maturity bucket strategy is robust when considering transaction costs.

6 Conclusions

This study investigates yield curve arbitrage opportunities in the international government bond markets by means of constructing a trading strategy based on a most feasible yield curve model. The main contribution to the literature of this research is evaluating different yield curve models by their economic values in terms of the ability of identifying the mispricings of the bonds and using a generic yield curve arbitrage strategy to summarize the pricing information in the modeled yield curve, and the yield curve models considered involve a comprehensive comparison of the Nelson-Siegel type of models estimated with Bayesian techniques and various popular non-parametric models. Also, this paper uses a database that covers the seven major government bond markets in the world while most previous studies focus on the term structure or the profitability of a trading strategy of one country. The design of the maturity-bucket-based yield curve arbitrage trading strategy is innovative for the existing fixed-income arbitrage literature as well, which makes the arbitrage factor constructed in this study arbitrage driven while similar studies construct the arbitrage factor resembling other risk factors such as the yield curve curvature.

The yield curve modeling of this study considers three specifications of the Nelson-Siegel models, a cubic spline with B-spline basis, and kernel smoothing with two kernel functions. The results show that when the evaluation criteria are statistical properties, the most flexible kernel smoothing methods are preferred across all countries. However, in terms of the ability of identifying bond mispricings, which is examined by a generic yield curve arbitrage strategy that goes one duration long in the under-priced bonds and one duration short in the over-priced bonds, none of the methods is universally feasible. The non-parametric methods remain favored for Australia, Canada, and Germany, while for the other countries, the Nelson-Siegel models are preferred. Within the Nelson-Siegel type of models, the more flexible four factor models by Björk and Christensen (1999) or by Svensson (1994) show better performance. Nevertheless, though the Nelson-Siegel models are preferred overall, it is observable that the non-parametric methods have an advantage when the market is volatile and there are much movements in the term structure.

On top of selecting the yield curve models, the results of the generic yield curve arbitrage strategy used to evaluate the yield curve models also show that a trading strategy based on arbitrage opportunities is able to profit in all countries, although the profitability of each countries varies, with Canada being the most profitable and Germany the least. Moreover, such a strategy is shown to have stronger performance during periods when the market is turbulent, for example during crisis, than when the market is stable, for example, the recent few years. After adjusting for other risk factors, the arbitrage strategy is able to achieve a positive risk premium, and the significant loadings on the risk factors indicate positioning bias of the strategy as well as possibilities for improvements, that is, positioning neutral to the risk factors to reduce the volatilities and improve the applicabilities of the strategy. To achive this, a modified strategy that collects the bonds into maturity buckets is formed. The results confirm that this strategy is able to achieve at least as good results as the generic strategy while reducing the volatilities and and limiting exposures to other risk factors. The subsample results and the sensitivity analysis show that the strategy specification is robust across sample period and holding periods as well as the transaction cost restriction, making the modified strategy to be more practically applicable than the generic strategy.

Due to time and other limits, the smoothness of the yield curve generated by different models is not formally measured, as it is difficult to find a good measure that is applicable for the kernel smoothing methods. This study does not consider model combinations to further improve the model fit both in terms of the statistical properties and the generic yield curve arbitrage strategy performance, which can be one of the directions for future research. Also, it might worth investigating a sample-period-specific yield curve model selection as the results. Finally, identifying the cheap and expensive points along the yield curve can be applied to trading strategy based on other factors to examine if buying (selling) the "cheap" ("expensive") bonds only will further improve the strategy's profitability.

Acknowledgement

I want to thank my Robeco supervisor Casper Zomerdijk, Martin Martens and Johan Duyvesteyn, and other colleagues from Quant Allocation Department of Robeco for very useful discussions, suggestions and comments. I also thank my supervisor professor Dick van Dijk and co-reader Michael Gong for very helpful discussions and comments. Finally I thank Yanzhong Xu, Dongfang Yang and Michael Hogers for their support.

References

Beekhuizen, P., Duyvesteyn, J. G., Martens, M., and Zomerdijk, C. (2016). Carry investing on the yield curve. Working paper.

Beim, D. (1992). Term structure and the non-cash values in bonds. Working papers, Columbia - Graduate School of Business.

BIS (2005). Zero-coupon yield curves: Technical documentation. *Bank for International Settlements*.

Björk, T. and Christensen, B. J. (1999). Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4):323–348.

Bliss, R. R. (1996). Testing term structure estimation methods. Technical report, Working paper, Federal Reserve Bank of Atlanta.

Bolder, D. J. and Gusba, S. (2002). Exponentials, polynomials, and fourier series: More yield curve modelling at the bank of canada.

Chambers, D. R., Carleton, W. T., and Waldman, D. W. (1984). A new approach to estimation of the term structure of interest rates. *Journal of financial and quantitative analysis*, 19(03):233–252.

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). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407.

de Boor, C. (1978). *A practical guide to splines*, volume 27. Springer-Verlag New York.

de Pooter, M. (2007). Examining the nelson-siegel class of term structure models. Technical report, Tinbergen Institute Discussion Paper.

De Pooter, M., Ravazzolo, F., and van Dijk, D. J. (2007). Predicting the term structure of interest rates: Incorporating parameter uncertainty, model uncertainty and macroeconomic information.

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., Li, C., and Yue, V. Z. (2008). Global yield curve dynamics and interactions: a dynamic nelson–siegel approach. *Journal of Econometrics*, 146(2):351–363.

Diebold, F. X., Rudebusch, G. D., and Aruoba, S. B. (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of econometrics*, 131(1):309–338.

Duarte, J., Longstaff, F. A., and Yu, F. (2006). Risk and return in fixed-income arbitrage: Nickels in front of a steamroller? *The Review of Financial Studies*, 20(3):769–811.

Dupačová, J., Abaffy, J., Bertocchi, M., and Hušková, M. (1997). On estimating the yield and volatility curves. *Kybernetika*, 33(6):659–673.

Fama, E. F. and Bliss, R. R. (1987). The information in long-maturity forward rates. *The American Economic Review*, pages 680–692.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741.

Gimeno, R. and Nave, J. M. (2006). Genetic algorithm estimation of interest rate term structure.

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

Hagan, P. S. and West, G. (2006). Interpolation methods for curve construction. *Applied Mathematical Finance*, 13(2):89–129.

Hastie, T., Tibshirani, R., and Friedman, J. (2013). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics. Springer New York.

Hu, G. X., Pan, J., and Wang, J. (2013). Noise as information for illiquidity. *The Journal of Finance*, 68(6):2341–2382.

Lin, B.-H. (2002). Fitting term structure of interest rates using b-splines: the case of taiwanese government bonds. *Applied Financial Economics*, 12(1):57–75.

Linton, O., Mammen, E., Nielsen, J. P., and Tanggaard, C. (2001). Yield curve estimation by kernel smoothing methods. *Journal of Econometrics*, 105(1):185–223.

McCulloch, J. H. (1971). Measuring the term structure of interest rates. *The Journal of Business*, 44(1):19–31.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142.

Nawalkha, S. K. and Soto, G. M. (2009). Term structure estimation.

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of business*, pages 473–489.

Perold, A. F. (1999). *Long-Term Capital Management, LP (C)*. Harvard Business School Pub.

Ritter, C. and Tanner, M. A. (1992). Facilitating the gibbs sampler: the gibbs stopper and the griddy-gibbs sampler. *Journal of the American Statistical Association*, 87(419):861–868.

Sediva, B. and Marek, P. (2015). Term structure of interest rates: comparison of the czech republic and germany. Conference proceedings, Mathematical Methods in Economics 2015.

Simonoff, J. S. (2012). *Smoothing methods in statistics*. Springer Science & Business Media.

Steeley, J. M. (1991). Estimating the gilt-edged term structure: Basis splines and confidence intervals. *Journal of Business Finance & Accounting*, 18(4):513–529.

Svensson, L. E. (1994). Estimating and interpreting forward interest rates: Sweden 1992-1994. Technical report, National Bureau of Economic Research.

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

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

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372.

Appendix A Bootstrapping Fama-Bliss unsmoothed yields

For each month, the Fama-Bliss unsmoothed yields are bootstrapped as follows: the bonds are first sorted by their maturities. Suppose the one-day forward rates are calculated from day 0 to T and the next bond matures at day T + k. Coupons before time T of this bond can be priced with the backed out forward rates and the remaining cash flows and their schedule can be used to calculate the one-day yield from T to T + k, which is the one-day forward rate for this time interval, too. Such calculations generate a term structure with constant one-day forward rates between two successive maturities. By induction, it is trivial that the spot rate is the average of these one-day forward rates with corresponding maturities.

Cash Flow Scheduling As this part of information is not directly available from the data, the cash flow schedule is constructed from the coupon frequency and maturity date for each bond. The last cash flow date is set to be the maturity date and second last cash flow date is one year (half a year, or a quarter of a year) ago on the same date, if the coupon frequency of the bond is one (two or four), and proceed further backwards until the issuing date. If any of the cash flow date falls on a Friday or a national holiday thus a non-trading day[25], it is set to the next trading day. One coupon payment falls on each of these coupon paying dates, and on the last cash flow date the principle is also paid, giving a cash flow of 100 + coupon payment. All coupon payments are discounted according to the coupon frequency of the bond.

Ex-Coupon Effect The bonds in the data follows different ex-coupon schedules, which makes fitting one ex-coupon rule for all bonds not desirable. Thus the following steps are used to decide if the ex-coupon effect is present for each observation of each bond: if there is drop in the accrued interest compared to last month and if the observation date is before or on the next coupon payment date, the observation date is marked with ex-coupon effect. When ex-coupon effect is present, the first future cash flow is not considered therefore the current price of the bond is treated as the sum of the discounted value of the second to the last future cash flows.

[25] The non-traidng date data is retrieved from Bloomberg.

Appendix B Bayesian inference for the Nelson-Siegel models

The joint posterior distributions of the parameters for the Nelson-Siegel models are of unknown distribution and do not have a known closed-form expression, thus the marginal densities of the parameters can not be computed analytically and should be simulated using Markov chain Monte Carlo (MCMC).

Prior Specification For that the modeled curves should reflect how the market moves in reality as well as to achieve an easy posterior simulation, the priors are chosen to be flat or from the conjugate prior family when possible. For λ_t, a uniform distribution is used

$$\lambda_t \sim U(\underline{\lambda}, \overline{\lambda}) \tag{A.1}$$

λ parameters determine the maturity at which the loading of the medium-term component is maximized, thus following existing literature (see, for example, Diebold and Li, 2006; de Pooter, 2007; De Pooter et al., 2007), λ_t for NS1 and NS2 model and $\lambda_{1,t}$ are chosen such that the six- to 60- month (half a year to five years) maturities achieve maximal loading for the third (medium maturity) term, and $\lambda_{2,t}$ for the NS3 model, the six- to 60-month maturities achieve maximal loading for the fourth (the second medium maturity) term, resulting in the following prior specification

$$\begin{aligned} \lambda_t(\lambda_{1,t}) &\sim U(3.3458, 33.4400) \\ \lambda_{2,t} &\sim U(5.8981, 58.9756) \end{aligned} \tag{A.2}$$

in the NS3 model, $\lambda_{1,t}$ and $\lambda_{2,t}$ are assumed independent for simplicity. An extra restriction of $\lambda_{1,t} \leq \lambda_{2,t}$ is also imposed such that the model is well-identified. For some countries with smaller numbers of bonds available (for example, Australia, Sweden, and Canada and the U.K. in the early 2000's), it happens that the shortest term bond already has a maturity of more than 20 months so that the domain for the decay parameters given in (A.2) is no longer reasonable. Thus I also give an alternative prior for the λ_t parameters that only applies to Australia and Sweden, or when the maturity spectrum starts from longer than six months (half a year) for the other countries[26]

$$\lambda_t(\lambda_{1,t}) \sim U(13.3824, 33.4400)$$
$$\lambda_{2,t} \sim U(23.5937, 58.9756) \tag{A.3}$$

This specification is derived by maximizing the loading of the medium-term component at 24- to 60- month maturities so that the interval is reasonable and sufficiently wide in order not to be too restrictive (see de Pooter, 2007).

For the variance parameter σ_t^2, the prior is specified with a flat prior

$$p[\sigma_t^2] \propto \sigma^{-2} \tag{A.4}$$

β_t takes form of a multivariate Normal prior

$$\beta_t|\mu_\beta, \Sigma_\beta \sim MN(\mu_\beta, \Sigma_\beta) \tag{A.5}$$

with μ_β and Σ_β predetermined by the mean and the covariance matrix the OLS estimator $\hat{\beta}_t$ with λ_t fixed and evaluated with a grid of 200 points in the interval as defined in (A.2). For the NS3 model the grid is set to be from the mesh of the 200×200 points of which satisfy the restriction of $\lambda_{1,t} \leq \lambda_{2,t}$.

Posterior simulation Posterior results are obtained using MCMC simulation with the Gibbs sampler of Geman and Geman (1984). The conditional posterior distributions are needed for sampling using the Gibbs sampler. Collecting the parameters in $\theta = \lambda, \sigma_t^2, \beta_t$, the joint posterior is given as follows

$$\begin{aligned} p[\theta|\mathbf{r}_t] &\propto p[\mathbf{r}_t|\theta]p[\theta] \\ &\propto \sigma^{-N} e^{-\frac{1}{2\sigma^2}(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)'(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)} |\Sigma_\beta|^{-\frac{1}{2}} e^{-\frac{1}{2}(\beta_t - \mu_\beta)'\Sigma_\beta^{-1}(\beta_t - \mu_\beta)} \\ &\quad \sigma^{-2} I_{\underline{\lambda} \leqslant \lambda_t \leqslant \overline{\lambda}} \end{aligned} \tag{A.6}$$

where $I_{[\cdot]}$ is an indicator function.

[26]Experimentation with wider domains as in (A.2) especially when together with a small sample size indeed resulted in fairly extreme estimates for both the λ_t and the β_t parameters.

To sample $\sigma^2|\mathbf{r}_t, \beta_t, \lambda_t, \mu_\beta, \Sigma_\beta$, one can consider

$$p[\sigma^2|\mathbf{r}_t, \beta_t, \lambda_t, \mu_\beta, \Sigma_\beta] \propto \sigma^{-(N+2)} e^{-\frac{1}{2\sigma^2}(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)'(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)} \tag{A.7}$$

which is a kernel of an Inverted Gamma-2 distribution with scale $(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)'(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)$ and N degrees of freedom.

To sample $\beta_t|\mathbf{r}_t, \sigma^2, \lambda_t, \mu_\beta, \Sigma_\beta$, one can use

$$\begin{aligned} p[\beta_t|\mathbf{r}_t, \sigma^2, \lambda_t, \mu_\beta, \Sigma_\beta] &\propto e^{-\frac{1}{2\sigma^2}(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)'(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)} e^{-\frac{1}{2}(\beta_t - \mu_\beta)'\Sigma_\beta^{-1}(\beta_t - \mu_\beta)} \\ &\propto e^{-\frac{1}{2\sigma^2}[(\mathbf{r}_t - \Phi[\lambda_t]\hat{\beta}_t)'(\mathbf{r}_t - \Phi[\lambda_t]\hat{\beta}_t) + (\beta_t - \hat{\beta}_t)'\Phi[\lambda_t]'\Phi[\lambda_t](\beta_t - \hat{\beta}_t)]} \\ &\quad e^{-\frac{1}{2}(\beta_t - \mu_\beta)'\Sigma_\beta^{-1}(\beta_t - \mu_\beta)} \\ &\propto e^{-\frac{1}{2\sigma^2}(\beta_t - \hat{\beta}_t)'\Phi[\lambda_t]'\Phi[\lambda_t](\beta_t - \hat{\beta}_t)} e^{-\frac{1}{2}(\beta_t - \mu_\beta)'\Sigma_\beta^{-1}(\beta_t - \mu_\beta)} \\ &\propto e^{-\frac{1}{2\sigma^2}[(\beta_t - \hat{\beta}_t)'\Phi[\lambda_t]'\Phi[\lambda_t](\beta_t - \hat{\beta}_t) + (\beta_t - \mu_\beta)'\Sigma_\beta^{-1}\sigma^2(\beta_t - \mu_\beta)]} \\ &\propto e^{-\frac{1}{2\sigma^2}(\beta_t - \tilde{\beta})'V'V(\beta_t - \tilde{\beta})} \end{aligned} \tag{A.8}$$

$\hat{\beta}_t$ denotes the OLS estimator for β_t. This is a kernel of a multivariate normal distribution with mean $\tilde{\beta} = (\Phi[\lambda_t]'\Phi[\lambda_t] + \sigma^2\Sigma_\beta^{-1})^{-1}(\Phi[\lambda_t]'r_t + \sigma^2\Sigma_\beta^{-1}\mu_\beta)$ and covariace matrix $\sigma^2(V'V) = \sigma^2(\Phi[\lambda_t]'\Phi[\lambda_t] + \sigma^2\Sigma_\beta^{-1})^{-1}$.

Lastly, the conditional posterior for λ_t is given by

$$p(\lambda_t|\mathbf{r}_t, \sigma^2, \beta_t, \mu_\beta, \Sigma_\beta) \propto e^{-\frac{1}{2\sigma^2}(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)'(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)} I_{\underline{\lambda} < \lambda_t < \bar{\lambda}} \tag{A.9}$$

which is of an unknown density, following De Pooter et al. (2007) $\lambda_t|\mathbf{r}_t, \sigma^2, \beta_t, \mu_\beta, \Sigma_\beta$ is sampled using the Griddy Gibbs algorithm developed by Ritter and Tanner (1992) with a grid of 200 points since the space of λ_t is well-defined.

The simulation for the posterior results can be performed by the running the following steps:

- Initializing: set β_t, θ to starting values.
- Sampling: run the following steps for M iterations, of which the initial M_0 draws are discarded
  1. Sample σ^2 from Inverted Gamma-2 distribution with scale $(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)'(\mathbf{r}_t - \Phi[\lambda_t]\beta_t)$ and N degrees of freedom.

2. Sample μ_β from $MN(\beta_t,\Sigma_\beta)$.

3. Sample β_t from $MN((\Phi[\lambda_t]'\Phi[\lambda_t]+\sigma^2\Sigma_\beta^{-1})^{-1}(\Phi[\lambda_t]'r_t+\sigma^2\Sigma_\beta^{-1}\mu_\beta),\sigma^2(\Phi[\lambda_t]'\Phi[\lambda_t]+\sigma^2\Sigma_\beta^{-1})^{-1})$.

4. Sample λ_t with the Griddy Gibbs sampler: Divide the interval $[\underline{\lambda},\overline{\lambda}]$ into a grid of 200 points $x_i, i = 1, ..., 200$, and evaluate $p(\lambda_t|\mathbf{r}_t,\beta_t,\theta)_{\lambda_t=x_i}$ to obtain $v_i = p(\lambda_t = x_i|\mathbf{r}_t,\beta_t,\theta)$. These $v_i, i = 1, ..., 200$ are used to approximate the empirical CDF of $p(\lambda_t|\mathbf{r}_t,\beta_t,\theta)$. Then sample from $U(0,1)$ and take value of the inverse CDF as the the new draw for λ_t.

Appendix C Estimation of the cubic B-Spline

The B-spline algorithm used for constructing cubic splines for this reseach follows Hastie et al. (2013). Here the estimation for the cubic splines is discussed using the three-piece case with the knots defined as $[0, 48, 120, \tau_{\max}]$ for illustration, where $\tau_{\max}$ stands for the maturity of the bond with the longest maturity observed in a random month. The two- and one-piece cubic splines with alternative knot specifications can be derived analogously.

With (3.12) and (3.13), the basis function of order three defined over the interval $[k_i, k_{i+1}]$ can be written out explicitly into polynomial form

$$
B_{i,3,t}[\tau] = \begin{cases}
\frac{(\tau-k_i)^3}{(k_{i+3}-k_i)(k_{i+2}-k_i)(k_{i+1}-k_i)} & \text{if } k_i \leq \tau < k_{i+1} \\[2em]
\frac{(\tau-k_i)^2(k_{i+2}-\tau)}{(k_{i+3}-k_i)(k_{i+2}-k_i)(k_{i+1}-k_i)} + \frac{(\tau-k_i)(\tau-k_{i+1})(k_{i+3}-\tau)}{(k_{i+3}-k_i)(k_{i+3}-k_{i+1})(k_{i+2}-k_{i+1})} & \\
\quad + \frac{(\tau-k_{i+1})^2(k_{i+4}-\tau)}{(k_{i+4}-k_{i+1})(k_{i+2}-k_{i+1})(k_{i+3}-k_{i+1})} & \text{if } k_{i+1} \leq \tau < k_{i+2} \\[2em]
\frac{(\tau-k_i)(k_{i+3}-\tau)^2}{(k_{i+3}-k_i)(k_{i+3}-k_{i+1})(k_{i+3}-k_{i+2})} + \frac{k_{i+4}-\tau}{k_{i+4}-k_{i+1}} & \\
\quad [\frac{(\tau-k_{i+1})(k_{i+3}-\tau)}{(k_{i+3}-k_{i+1})(k_{i+3}-k_{i+2})} + \frac{(\tau-k_{i+2})(k_{i+4}-\tau)}{(k_{i+4}-k_{i+2})(k_{i+3}-k_{i+2})}] & \text{if } k_{i+2} \leq \tau < k_{i+3} \\[2em]
\frac{(k_{i+4}-\tau)^3}{(k_{i+4}-k_{i+1})(k_{i+4}-k_{i+2})(k_{i+4}-k_{i+3})} & \text{if } k_{i+3} \leq \tau < k_{i+4} \\[2em]
0 & \text{otherwise}
\end{cases} \tag{A.1}
$$

It is easy to notice from (3.14) and (A.1) that for each $[k_i, k_{i+1}]$ interval to have a valid basis function, at least three intervals to the right ($[k_{i+1}, k_{i+2}]$, $[k_{i+2}, k_{i+3}]$, and $[k_{i+3}, k_{i+4}]$) are required. Actually, the procedure of iteratively calculating the basis function of order three from lower order basis functions can be expressed by a diagram as in Figure 5 below

$$
\begin{tikzcd}
B_{0,0,t} \arrow[r] & B_{0,1,t} \arrow[r] & B_{0,2,t} \arrow[r] & B_{0,3,t} \\
B_{1,0,t} \arrow[r] \arrow[ur] & B_{1,1,t} \arrow[r] \arrow[ur] & B_{1,2,t} \arrow[r] \arrow[ur] & B_{1,3,t} \\
B_{2,0,t} \arrow[r] \arrow[ur] & B_{2,1,t} \arrow[r] \arrow[ur] & B_{2,2,t} \arrow[r] \arrow[ur] & B_{2,3,t} \\
B_{3,0,t} \arrow[r] \arrow[ur] & B_{3,1,t} \arrow[r] \arrow[ur] & B_{3,2,t} \arrow[ur] & \\
B_{4,0,t} \arrow[r] \arrow[ur] & B_{4,1,t} \arrow[ur] & & \\
B_{5,0,t} \arrow[ur] & & &
\end{tikzcd}
$$

Figure 5: Tableau for iteratively calculating the basis function of order three. $B_{i,n,t}$ denotes the basis function of order n defined over the interval between knot k_i and k_{i+1}.

Thus for the setting of this study, $[k_0, ..., k_3, ..., k_6]$ should be well defined while the data only provides $[k_0, ..., k_3]$, where $k_0 = 0, ..., k_3 = \tau_{max}$ as defined earlier. However, another fact can be observed from (A.1) is that, since $B_{i,3,t}$ is defined over the mode the interval $[k_i, k_{i+1}]$ instead of the value of k_i and k_{i+1} itself, $[k_0, k_1, ...k_N]$ gives the exact same basis functions $B_{i,3,t}$ as $[k_0 + c, k_1 + c, ...k_N + c]$, where c is an arbitrary constant. Using this property of the basis function, the auxiliary parameters $[k_4, k_5, k_6]$ can be constructed as

$$k_i = k_{i-1} + (k_{i-3} - k_{i-4}) \quad \text{for } i > 3 \tag{A.2}$$

this specification for the auxiliary knots completes the paramization of the B-splines.

To fit the yield curve of a given month t, the a regression analysis for the spot rates can be constructed and written into the following matrix notation

$$\begin{bmatrix} B_{0,3,t}[\tau_1] & B_{1,3,t}[\tau_1] & B_{2,3,t}[\tau_1] \\ \cdots & \cdots & \cdots \\ B_{0,3,t}[\tau_N] & B_{1,3,t}[\tau_N] & B_{2,3,t}[\tau_N] \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} r_t[\tau_1] \\ \cdots \\ r_t[\tau_N] \end{bmatrix} \tag{A.3}$$

or in a more compact form

$$B_{3,t}\alpha = \mathbf{r}_t \tag{A.4}$$

where τ_i, $i = 1, ..., N$ are the maturities of the bonds observed in a random month, N denotes the number of observations and the α parameters are of interest. Such a matrix notation reveals the coefficients α can be solved by regression theory, $\hat{\alpha} = (\mathrm{B}'_{3,t}\mathrm{B}_{3,t})^{-1}(\mathrm{B}'_{3,t}\mathbf{r}_t)$. $\mathrm{B}_{0,3,t}[\tau]$ takes non-zero values for τ belongs to the interval $[k_0, k_1]$, $[k_1, k_2]$, and $[k_2, k_3]$; $\mathrm{B}_{1,3,t}[\tau]$ takes non-zero values for τ in the interval $[k_1, k_2]$, and $[k_2, k_3]$; and $\mathrm{B}_{2,3,t}[\tau]$ takes non-zero values only for τ in $[k_2, k_3]$. Thus cubic spline consists of six pieces of polynomials over the whole maturity spectrum and the dimension of α is correspondingly 6.

Appendix D Tables

Maturity bucket	Avg. # Bonds	Avg. duration	Avg. Maturity (year)	Avg. Market value (bln)	Avg. Age (year)	Avg. Yield (%)	Std. Yield (%)	Skew. Yield	Kurt. Yield	JB Stats.
Australia (06/00 [a]-10/16)										
0-1 year	1.20	0.61	0.63	7.46	9.62	4.18	1.48	-0.18	1.82	11.75
1-3 year	3.36	1.89	2.04	8.30	8.00	4.43	1.50	-0.43	1.97	16.24
3-5 year	2.92	3.50	3.97	8.45	6.42	4.58	1.47	-0.60	2.11	20.01
5-7 year	1.71	4.97	5.92	9.09	6.22	4.65	1.35	-0.71	2.17	23.01
7-10 year	1.99	6.78	8.48	9.61	4.14	4.77	1.27	-0.75	2.25	23.26
10-15 year	1.95	8.58	11.38	6.49	1.78	4.89	1.19	-0.83	2.46	25.09
15+ year	0.52	13.00	17.92	6.01	0.96	3.65	0.66	-0.22	1.92	2.98
Canada (06/00-10/16)										
0-1 year	4.21	0.59	0.60	6.83	5.52	1.85	1.41	1.20	3.84	44.52
1-3 year	11.59	1.79	1.87	8.29	4.50	2.42	1.48	0.43	2.09	12.81
3-5 year	6.91	3.64	3.96	9.08	4.49	2.82	1.48	0.13	1.91	10.40
5-7 year	4.55	5.14	5.86	9.54	5.07	3.10	1.46	0.02	1.84	11.11
7-10 year	5.18	7.13	8.46	10.36	3.53	3.40	1.41	-0.07	1.87	10.64
10-15 year	3.27	8.71	12.01	5.02	11.13	3.63	1.33	-0.09	1.89	10.04
15+ year	6.61	14.22	23.73	12.73	8.23	3.94	1.26	-0.06	1.89	10.28
Germany (01/99 [b]-10/16)										
0-1 year	8.41	0.62	0.61	17.27	4.92	1.92	1.68	0.09	1.63	17.20
1-3 year	15.83	1.83	1.88	19.83	4.69	2.06	1.70	-0.02	1.65	16.43
3-5 year	10.65	3.72	3.97	22.74	4.01	2.39	1.70	-0.25	1.74	16.38
5-7 year	5.05	5.37	5.98	26.81	4.23	2.69	1.66	-0.40	1.90	16.40
7-10 year	7.43	7.40	8.45	26.51	2.05	3.00	1.58	-0.52	2.09	16.88
10-15 year	1.76	9.58	12.53	14.45	14.86	3.01	1.48	-0.39	2.07	10.48
15+ year	7.95	15.10	24.13	22.79	6.61	3.69	1.46	-0.63	2.40	17.45
UK (07/98-10/16)										
1- year	2.71	0.59	0.61	26.21	9.70	2.78	2.25	0.08	1.36	23.89
1-3 year	4.77	1.89	2.01	31.34	7.32	3.05	2.14	-0.09	1.36	25.05
3-5 year	4.71	3.61	3.99	31.62	5.98	3.34	1.88	-0.24	1.53	21.99
5-7 year	3.32	5.09	5.94	30.58	7.52	3.54	1.67	-0.34	1.68	20.08
7-10 year	4.30	6.93	8.48	31.84	6.03	3.74	1.43	-0.51	1.94	19.84
10-15 year	3.20	9.01	12.30	33.41	8.72	3.90	1.21	-0.70	2.26	23.14
15+ year	9.74	15.14	25.92	34.43	5.42	4.03	0.89	-1.07	3.28	42.41
Japan (06/00-10/16)										
1- year	25.66	0.63	0.64	23.77	4.62	0.15	0.20	0.90	3.65	29.87
1-3 year	54.30	1.90	1.93	27.39	4.48	0.23	0.28	0.99	3.34	33.33
3-5 year	39.24	3.90	4.01	34.61	3.85	0.43	0.36	0.58	2.72	11.66
5-7 year	23.90	5.74	6.03	32.32	4.88	0.65	0.43	0.09	2.48	2.44
7-10 year	35.53	7.97	8.47	36.60	2.72	0.95	0.48	-0.48	2.69	8.39
10-15 year	27.83	10.96	12.72	11.04	7.37	1.32	0.50	-0.95	3.31	30.12
15+ year	60.96	16.53	20.40	14.20	2.85	1.74	0.47	-1.47	5.12	108.14
Sweden (07/98-10/16)										
1- year	1.26	0.62	0.62	6.51	7.05	2.24	1.62	-0.12	1.72	11.80
1-3 year	2.36	1.88	1.97	9.56	8.25	2.44	1.68	-0.19	1.84	13.61
3-5 year	1.97	3.66	3.99	10.19	7.03	2.80	1.66	-0.36	2.03	13.26
5-7 year	1.75	5.26	5.95	8.72	5.78	3.05	1.60	-0.42	2.13	13.43
7-10 year	2.38	7.24	8.51	8.49	3.62	3.27	1.55	-0.45	2.19	13.54
10-15 year	1.13	9.20	11.48	7.44	2.73	3.47	1.54	-0.43	2.18	11.65
15+ year	0.81	16.06	22.86	5.83	2.86	2.99	1.15	0.04	2.01	5.08

Table 17: Summary statistics of the bonds per country, grouped in maturity buckets. The sample periods are indicated in the brackets. Column 2 to column 6 contain the characteristics of the bonds in the maturity buckets over time, measured by the average: number of bonds, modified duration, maturity, market value and age of the bonds; column 7 to column 11 report the statistics of the yields of the bonds in the maturity buckets: mean, standard deviation, skewness, kurtosis and Jarque-Bera statistic. The results for yields are on annual basis.

[a] The original bond data starts from January 1999 with limited number of observations, thus the analysis starts from June 2000.

[b] The original bond data starts from July 1998 with limited number of observations and is discontinuous, thus the analysis starts from January 1999.

	NS1	NS2	NS3	BS	KS1	KS2		NS1	NS2	NS3	BS	KS1	KS2
Sharpe							**Vol.**						
AU	0.07	**0.24**	**0.34**	0.08	0.13	*-0.09*	AU	12.17	11.38	11.93	10.67	**9.73**	**8.25**
CA	0.95	1.02	0.93	**1.06**	**1.11**	1.05	CA	**5.65**	**5.64**	6.21	6.69	6.48	6.07
DE	**1.13**	**0.78**	0.62	0.74	0.52	0.36	DE	**5.24**	6.44	6.04	6.65	7.84	9.00
UK	0.35	**0.49**	**0.60**	0.38	0.42	0.26	UK	**9.55**	9.60	11.05	10.98	**9.55**	11.45
JP	0.94	**1.44**	0.97	**1.37**	1.11	0.69	JP	3.92	2.94	3.47	**2.28**	**2.12**	2.33
SE	**0.49**	0.28	0.42	0.39	0.44	**0.49**	SE	39.81	41.51	38.87	41.27	**32.86**	**33.07**
US	0.35	**0.77**	**0.66**	0.30	0.35	0.22	US	7.73	6.04	7.02	7.67	**6.67**	7.93
Mean							**Skew.**						
AU	0.89	**2.72**	**4.02**	0.90	1.24	-0.78	AU	0.15	-1.53	0.60	-0.04	-0.44	-2.04
CA	5.37	5.78	5.75	**7.11**	**7.17**	6.39	CA	0.76	0.42	0.97	0.76	1.38	1.07
DE	**5.92**	**4.99**	3.75	4.91	4.08	3.20	DE	1.86	-0.59	0.06	-0.23	2.91	1.00
UK	3.31	**4.67**	**6.60**	4.15	3.98	2.96	UK	0.94	1.25	1.18	3.12	0.05	2.30
JP	**3.69**	**4.24**	3.38	3.13	2.34	1.62	JP	1.63	1.29	2.25	0.20	0.38	0.29
SE	**19.48**	11.45	**16.50**	16.00	14.57	16.11	SE	6.41	5.97	7.09	6.01	3.90	3.84
US	2.67	**4.64**	**4.65**	2.33	2.34	1.71	US	-3.69	-0.26	0.38	-0.50	-0.61	-4.20
$VaR_{.95}$							**D.D.**						
AU	-17.13	**-16.69**	-17.05	-18.48	**-13.22**	-18.37	AU	45.75	51.49	79.69	40.62	**32.74**	**38.22**
CA	**-3.75**	**-3.19**	-7.61	-9.14	-9.36	-9.49	CA	**14.33**	**15.81**	56.19	51.68	34.98	39.49
DE	**-7.64**	**-8.52**	-12.87	-14.14	-16.25	-16.74	DE	**59.68**	64.34	**64.18**	75.95	71.58	71.49
UK	**-8.97**	**-6.27**	-14.02	-14.71	-15.65	-16.08	UK	**33.83**	**30.15**	82.88	51.27	64.51	68.79
JP	**-4.69**	**-4.68**	-5.10	-5.44	-8.03	-8.42	JP	40.74	**20.03**	**18.21**	20.79	49.48	52.70
SE	**-12.24**	-15.02	-16.00	**-12.11**	-14.97	-17.10	SE	**67.22**	73.60	69.13	**66.34**	69.31	69.09
US	**-6.98**	**-6.98**	-11.53	-12.13	-15.05	-14.08	US	69.51	**50.87**	**52.84**	60.29	81.93	82.04

Table 18: Generic yield curve arbitrage strategy excess returns of different yield curve fitting models across countries. The strategy longs one duration of "cheap bonds" and shorts one duration of "expensive" bonds along the yield curve. The weight of the bonds are derived analogously as in (3.31): $w_i = w$ if $y_i \geq \hat{y}_i$ and $-w$ otherwise, where $w = \frac{1}{D_i}\frac{1}{\sum_{i=1}^{N} I_{\hat{y}_i \geq y_i}}$ and y_i and $\hat{y}_i$ are the actual yield and theoretical yield of the bond, respectively. The two upper panels report the annualized Sharpe ratios (Sharpe) and the annualized volatilities (Vol.), the two middle panels contain the annualized mean returns (Mean) and the skewness (Skew.) of the returns, and the two bottom panels are the empirical value-at-risk with confidence level of 0.95 ($VaR_{.95}$) and the maximum drawdown (D.D.) of the returns. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point. For the Sharpe ratios, the volatilities, and the mean returns, the two best-performing models for each country-property pair are marked bold face. Negative Sharpe ratios are marked italic.

	NS1	NS2	NS3	BS	KS1	KS2		NS1	NS2	NS3	BS	KS1	KS2
Sharpe							**Vol.**						
AU	*-0.29*	*-0.15*	**0.08**	-0.39	**0.02**	*-0.34*	AU	**32.26**	**33.26**	39.69	34.85	33.45	34.31
CA	**1.24**	**1.29**	0.28	0.81	0.87	0.83	CA	**12.10**	**9.38**	23.51	23.16	23.85	24.01
DE	**0.26**	**0.31**	0.13	0.10	0.07	0.06	DE	**17.10**	**16.86**	22.04	26.60	35.51	35.68
UK	**0.77**	**0.78**	0.50	0.58	0.53	0.39	UK	**26.25**	**25.28**	35.62	32.55	35.82	36.48
JP	0.46	**0.67**	**0.89**	0.47	0.08	0.04	JP	12.45	**11.40**	**11.00**	12.09	16.63	17.43
SE	**0.34**	0.20	**0.54**	0.33	0.29	0.30	SE	46.45	46.68	51.56	46.52	**43.30**	**44.00**
US	0.55	**0.76**	**0.64**	0.60	0.46	0.45	US	**20.22**	**20.67**	29.92	30.89	34.46	34.11
Mean							**Skew.**						
AU	-9.25	-5.15	**3.26**	-13.50	**0.53**	-11.59	AU	0.58	0.63	0.70	0.51	1.07	0.11
CA	14.98	12.09	6.68	18.69	**20.79**	**19.82**	CA	1.03	0.70	2.98	0.94	0.75	0.76
DE	**4.48**	**5.19**	2.85	2.53	2.43	1.97	DE	-0.71	-0.69	-0.34	0.07	1.59	1.50
UK	**20.20**	**19.74**	17.87	18.93	19.09	14.37	UK	0.99	0.92	-0.18	0.14	0.29	0.16
JP	5.68	**7.64**	**9.81**	5.71	1.41	0.73	JP	1.14	0.25	-0.13	0.24	-0.42	-0.59
SE	**15.71**	9.16	**27.66**	15.38	12.43	13.36	SE	5.23	5.90	3.44	5.45	3.04	2.79
US	11.03	15.69	**19.07**	**18.65**	15.70	15.27	US	-0.28	0.22	1.21	0.96	0.78	0.82
$\mathrm{VaR}_{.95}$							**D.D.**						
AU	-17.13	**-16.69**	-17.05	-18.48	**-13.22**	-18.37	AU	45.75	51.49	79.69	40.62	**32.74**	**38.22**
CA	**-3.75**	**-3.19**	-7.61	-9.14	-9.36	-9.49	CA	**14.33**	**15.81**	56.19	51.68	34.98	39.49
DE	**-7.64**	**-8.52**	-12.87	-14.14	-16.25	-16.74	DE	**59.68**	64.34	**64.18**	75.95	71.58	71.49
UK	**-8.97**	**-6.27**	-14.02	-14.71	-15.65	-16.08	UK	**33.83**	**30.15**	82.88	51.27	64.51	68.79
JP	**-4.69**	**-4.68**	-5.10	-5.44	-8.03	-8.42	JP	40.74	**20.03**	**18.21**	20.79	49.48	52.70
SE	**-12.24**	-15.02	-16.00	**-12.11**	-14.97	-17.10	SE	**67.22**	73.60	69.13	**66.34**	69.31	69.09
US	**-6.98**	**-6.98**	-11.53	-12.13	-15.05	-14.08	US	69.51	**50.87**	**52.84**	60.29	81.93	82.04

Table 19: Generic yield curve arbitrage strategy excess returns of different yield curve fitting models across countries. The strategy longs one duration of "cheap bonds" and shorts one duration of "expensive" bonds along the yield curve. The weight of the bonds are derived analogously as in (3.31): $w_i = w$ if $P_i \leq \hat{P}_i$ and $-w$ otherwise, where $w = \frac{1}{D_i} \frac{P_i - \hat{P}_i}{\sum_{i=1}^{N} (P_i - \hat{P}_i) I_{\hat{P}_i \leq P_i}}$ and P_i and $\hat{P}_i$ are the traded price and theoretical price of the bond, respectively. The two upper panels report the annualized Sharpe ratios (Sharpe) and the annualized volatilities (Vol.), the two middle panels contain the annualized mean returns (Mean) and the skewness (Skew.) of the returns, and the two bottom panels are the empirical value-at-risk with confidence level of 0.95 ($\mathrm{VaR}_{.95}$) and the maximum drawdown (D.D.) of the returns. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point. For the Sharpe ratios, the volatilities, and the mean returns, the two best-performing models for each country-property pair are marked bold face.

		α	R_m	R_s	R_c		α	R_m	R_s	R_c		α	R_m	R_s	R_c		α	R_m	R_s	R_c
AU						**CA**					**DE**					**UK**				
coef.	NS1	0.25	-0.42	**-5.40**	0.92	NS1	**0.58**	0.28	-1.49	-0.30	NS1	**0.75**	0.18	-2.32	**-9.27**	NS1	0.28	0.03	-1.17	3.81
t-score		0.82	-1.38	-2.00	0.12		3.88	1.51	-0.68	-0.08		6.09	1.27	-1.82	-3.57		1.37	0.20	-0.83	1.04
coef.	NS2	0.31	-0.44	**-4.55**	-2.20	NS2	**0.56**	**0.36**	-1.54	-0.33	NS2	**0.68**	0.08	**-3.10**	**-6.40**	NS2	0.39	0.05	-1.40	2.01
t-score		1.25	-1.72	-2.02	-0.35		3.94	2.04	-0.74	-0.09		4.81	0.47	-2.13	-2.15		1.89	0.35	-0.99	0.55
coef.	NS3	0.29	-0.22	-1.68	-3.87	NS3	**0.51**	0.28	-1.82	-0.78	NS3	**0.67**	0.36	**-4.11**	-6.87	NS3	**0.65**	0.09	0.22	-0.05
t-score		0.95	-0.69	-0.61	-0.51		3.09	1.36	-0.75	-0.18		3.83	1.77	-2.28	-1.86		2.60	0.48	0.13	-0.01
coef.	BS	0.18	-0.24	-3.92	-3.31	BS	**0.71**	0.15	-1.27	-1.15	BS	**0.69**	0.19	**-3.28**	**-8.12**	BS	0.37	-0.12	-1.24	3.16
t-score		0.64	-0.84	-1.56	-0.48		4.12	0.71	-0.50	-0.26		4.64	1.10	-2.13	-2.59		1.57	-0.72	-0.77	0.76
coef.	KS1	0.14	0.04	**-5.54**	4.12	KS1	**0.74**	0.18	-1.49	2.09	KS1	**0.57**	0.20	**-3.45**	**-7.96**	KS1	0.32	-0.16	-1.00	3.79
t-score		0.51	0.15	-2.22	0.60		4.27	0.81	-0.58	0.46		3.44	1.02	-2.02	-2.28		1.56	-1.12	-0.72	1.05
coef.	KS2	0.04	0.00	-3.54	-1.32	KS2	**0.70**	0.18	-1.38	1.93	KS2	**0.53**	0.17	-2.81	**-10.24**	KS2	0.31	-0.24	-1.29	7.33
t-score		0.19	-0.02	-1.85	-0.25		4.11	0.83	-0.55	0.43		2.77	0.75	-1.42	-2.53		1.28	-1.40	-0.79	1.73
JP						**SE**					**US**									
coef.	NS1	**0.38**	**0.81**	-2.20	**-5.26**	NS1	**1.98**	0.78	0.79	**-50.53**	NS1	**0.46**	**0.51**	**5.79**	**-11.78**					
t-score		3.83	2.88	-1.22	-2.12		2.31	0.89	0.11	-3.69		2.68	2.30	3.13	-3.53					
coef.	NS2	**0.40**	**0.49**	**-2.75**	1.35	NS2	1.35	-0.95	0.42	-25.29	NS2	**0.59**	**0.56**	-1.54	-0.49					
t-score		5.11	2.22	-1.97	0.71		1.65	-1.14	0.06	-1.94		4.06	2.99	-0.99	-0.17					
coef.	NS3	**0.21**	0.33	2.23	0.43	NS3	1.67	0.08	2.02	**-45.43**	NS3	**1.07**	0.07	2.23	-0.15					
t-score		2.32	1.26	1.35	0.19		1.92	0.09	0.29	-3.27		2.42	0.12	0.47	-0.02					
coef.	BS	**0.29**	**0.60**	**-3.64**	2.16	BS	**1.93**	-0.22	-0.54	**-49.17**	BS	**0.64**	0.59	5.15	**-12.70**					
t-score		4.90	3.51	-3.36	1.45		2.20	-0.24	-0.08	-3.52		2.36	1.68	1.77	-2.41					
coef.	KS1	**0.21**	**0.48**	**-5.42**	**7.32**	KS1	**1.41**	-0.89	0.07	**-27.00**	KS1	**0.62**	0.40	3.68	**-9.62**					
t-score		3.56	2.88	-5.14	5.06		2.20	-1.36	0.01	-2.64		2.66	1.34	1.48	-2.14					
coef.	KS2	0.11	**0.59**	**-7.24**	**9.50**	KS2	**1.52**	-0.76	0.86	**-25.49**	KS2	**0.51**	0.49	**6.00**	**-12.48**					
t-score		1.88	3.63	-6.96	6.65		2.32	-1.14	0.16	-2.45		2.18	1.64	2.39	-2.75					

Table 20: Estimated coefficients and corresponding t-statistic for the regression of monthly percentage excess returns of the modified yield curve arbitrage strategy on the the other risk factors. R_m, R_S, and R_C stand for the market portfolio, slope factor, and curvature factor, respectively. The α coefficient values are in basis point and the β coefficient values are in percentage. The reference critical value is 1.97 for a confidence level of 0.95 and 1.65 for a confidence level of 0.90. The coefficients significant at a confidence level of 0.95 are marked bold-face.

Holding period: 6 months													
Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.
AU	0.24	10.05	41.09	1.09	-60.63	74.83	AU	-0.13	-2.24	17.09	0.55	-46.90	51.11
CA	1.22	20.56	16.89	-0.90	-21.07	21.07	CA	0.45	0.54	1.20	-1.20	-0.98	1.36
DE	0.54	14.99	28.00	0.08	-10.09	23.04	DE	0.88	2.39	2.72	-0.38	-1.12	1.80
UK	0.94	28.46	30.42	-0.36	-25.83	46.86	UK	0.34	0.94	2.77	1.50	-1.73	6.63
JP	-0.08	-0.87	10.30	0.18	-10.05	11.24	JP	0.89	0.51	0.58	0.54	-0.48	0.48
SE	0.06	1.50	26.74	-1.39	-23.43	39.86	SE	0.21	1.97	9.29	-1.27	-9.26	14.15
US	-0.34	-8.95	26.44	-1.73	-8.24	14.59	US	0.52	2.02	3.87	-2.14	-1.83	2.04
Holding period: 12 months													
Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	VaR$_{.95}$	D.D.
AU	0.23	8.90	38.85	0.52	-60.63	60.63	AU	0.35	4.79	13.58	3.17	-46.90	51.11
CA	0.76	13.29	17.44	-1.83	-51.11	57.50	CA	0.45	0.27	0.60	-1.20	-0.49	0.68
DE	0.39	9.37	24.31	0.13	-18.33	37.74	DE	0.87	1.19	1.37	-0.37	-0.56	0.90
UK	1.16	29.14	25.20	0.50	-19.82	42.68	UK	0.34	0.47	1.38	1.50	-0.86	3.36
JP	-0.40	-3.96	9.99	-0.32	-9.94	14.18	JP	0.89	0.26	0.29	0.54	-0.24	0.24
SE	-0.60	-17.46	28.98	-1.12	-25.66	61.44	SE	-0.74	-3.89	5.26	-1.44	-4.91	15.90
US	-0.51	-13.85	27.11	-2.44	-8.21	13.48	US	0.52	1.01	1.93	-2.14	-0.92	1.02

Table 21: The annualized Sharpe ratios (Sharpe), the annualized mean returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 (VaR$_{.95}$), and the maximum drawdown (D.D.) of the returns of the generic and modified yield curve arbitrage strategy returns across countries with holding period of six months (upper panel) and 12 months (lower panel). The sample period is January 2007 to December 2009. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.

Holding period: 6 months													
Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	$VaR_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	$VaR_{.95}$	D.D.
AU	-0.54	-12.14	22.35	1.35	-12.39	57.69	AU	-0.56	-1.06	1.89	-3.24	-11.08	28.32
CA	1.03	8.47	8.23	0.79	-3.16	6.79	CA	1.04	1.11	1.07	2.12	-0.50	1.07
DE	-0.74	-5.71	7.71	-0.33	-9.63	23.36	DE	0.95	0.60	0.63	-0.31	-0.79	1.18
UK	-0.65	-10.99	16.99	-0.63	-4.26	8.50	UK	-0.80	-0.77	0.96	-0.17	-0.50	2.97
JP	-0.72	-2.83	3.92	-0.08	-4.77	35.75	JP	0.84	0.40	0.47	0.82	-0.23	0.66
SE	0.45	9.27	20.75	0.64	-11.55	32.63	SE	0.18	0.60	3.35	-1.37	-3.35	3.88
US	-0.26	-4.42	16.72	0.55	-5.04	29.92	US	0.08	0.05	0.69	0.40	-0.40	1.03
Holding period: 12 months													
Generic							**Modified**						
	Sharpe	Mean	Vol.	Skew.	$VaR_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	$VaR_{.95}$	D.D.
AU	-0.77	-16.02	20.84	0.69	-12.61	62.94	AU	-0.63	-0.84	1.33	-3.76	-11.08	28.00
CA	0.89	6.94	7.82	0.77	-3.37	7.36	CA	1.04	0.56	0.53	2.12	-0.25	0.54
DE	-0.86	-8.15	9.44	-0.41	-8.90	21.49	DE	0.95	0.30	0.31	-0.31	-0.39	0.59
UK	-0.65	-11.56	17.90	-0.44	-3.79	7.57	UK	-0.80	-0.39	0.48	-0.17	-0.25	1.49
JP	-0.67	-2.46	3.66	-0.41	-4.44	36.52	JP	0.84	0.20	0.24	0.82	-0.11	0.33
SE	0.30	4.78	15.93	0.32	-9.09	20.80	SE	0.23	0.40	1.75	-1.51	-1.91	1.96
US	-0.29	-4.69	16.44	0.59	-6.42	30.03	US	0.08	0.03	0.34	0.40	-0.20	0.51

Table 22: The annualized Sharpe ratios (Sharpe), the annualized mean returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 ($VaR_{.95}$), and the maximum drawdown (D.D.) of the returns of the generic and modified yield curve arbitrage strategy returns across countries with holding period of six months (upper panel) and 12 months (lower panel). The sample period is January 2012 to September 2016. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.

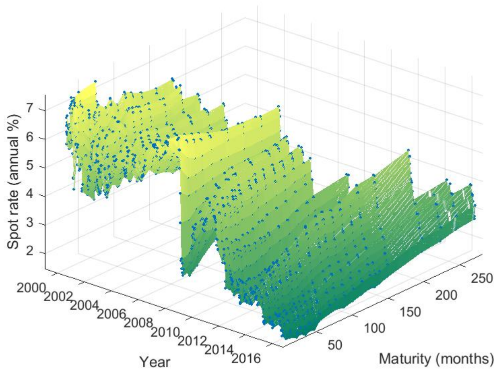
Generic							Modified						
	Sharpe	Mean	Vol.	Skew.	$VaR_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	$VaR_{.95}$	D.D.
AU	-	-	-	-	-	-	AU	0.58	1.23	2.14	5.75	0.00	0.00
CA	1.89	26.80	19.18	0.84	-5.50	8.30	CA	0.83	6.95	8.40	0.44	-5.50	8.30
DE	0.52	18.56	35.78	0.74	-28.26	38.32	DE	0.83	16.47	19.83	0.35	-13.90	14.69
UK	0.05	2.04	42.87	-1.72	-43.57	54.73	UK	0.35	6.90	19.99	1.08	-43.57	54.73
JP	0.84	14.56	17.42	0.01	-9.74	11.79	JP	1.94	7.73	3.98	-0.40	-9.74	11.79
SE	0.43	12.46	29.16	-0.75	-26.14	45.19	SE	0.60	14.73	24.64	-1.08	-26.14	45.19
US	1.05	19.58	18.72	-0.19	-11.18	17.08	US	1.40	21.40	15.25	-0.34	-11.18	17.08

Table 23: The annualized Sharpe ratios (Sharpe), the annualized mean returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 ($VaR_{.95}$), and the maximum drawdown (D.D.) of the returns of the generic and modified yield curve arbitrage strategy returns across countries with transaction cost restrictions. The sample period is January 2007 to December 2009. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.

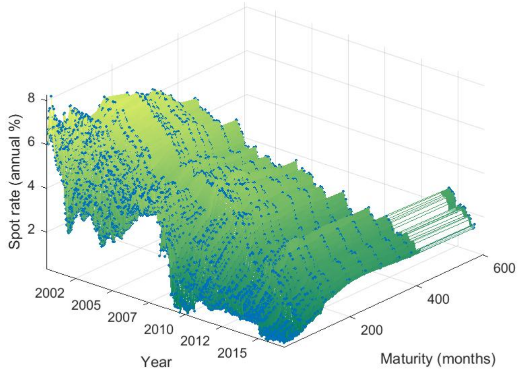
Generic							Modified						
	Sharpe	Mean	Vol.	Skew.	$VaR_{.95}$	D.D.		Sharpe	Mean	Vol.	Skew.	$VaR_{.95}$	D.D.
AU	0.34	8.78	25.47	0.84	-12.77	32.13	AU	-0.27	-2.02	25.47	0.07	-12.77	20.33
CA	1.41	15.69	11.14	0.94	-3.66	7.44	CA	0.99	7.39	11.14	0.07	-3.66	6.05
DE	-0.61	-4.08	6.66	0.01	-5.88	21.44	DE	0.98	3.46	6.66	0.04	-1.62	4.05
UK	0.18	1.62	9.17	-0.13	-5.70	11.11	UK	-0.69	-3.82	-3.82	0.06	-5.70	20.46
JP	-0.95	-7.72	8.15	0.50	-5.44	34.62	JP	1.17	-7.72	3.57	0.03	-5.44	4.60
SE	0.53	13.58	25.45	0.53	-13.30	37.94	SE	0.84	13.58	17.78	0.21	-13.30	15.80
US	-0.29	-5.04	17.67	-0.69	-13.50	48.68	US	0.13	-5.04	0.61	0.05	-13.50	8.95

Table 24: The annualized Sharpe ratios (Sharpe), the annualized mean returns (Mean), the annualized volatilities (Vol.), skewness (Skew.), empirical value-at-risk with confidence level of 0.95 ($VaR_{.95}$), and the maximum drawdown (D.D.) of the returns of the generic and modified yield curve arbitrage strategy returns across countries with transaction cost restrictions. The sample period is January 2012 to September 2016. The Sharpe ratios and the skewness are reported in real numbers, the maximum drawdowns are in percentage, and the volatilities, the mean returns and the VaR are in basis point.

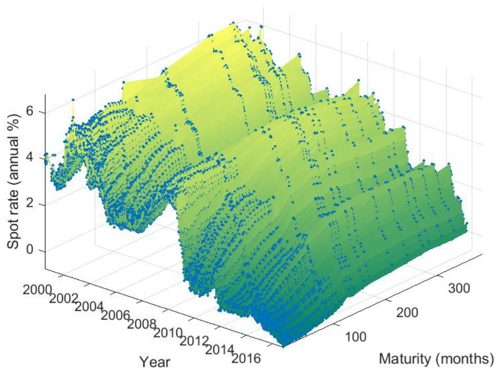
Appendix E   Figures



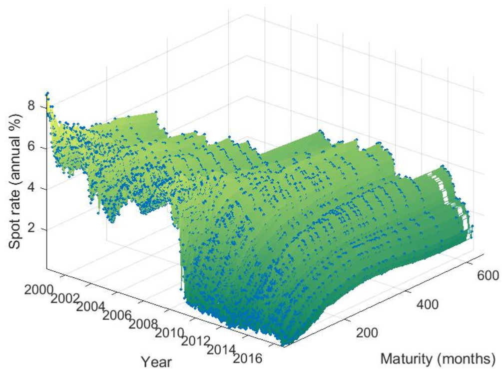
(a) Australia



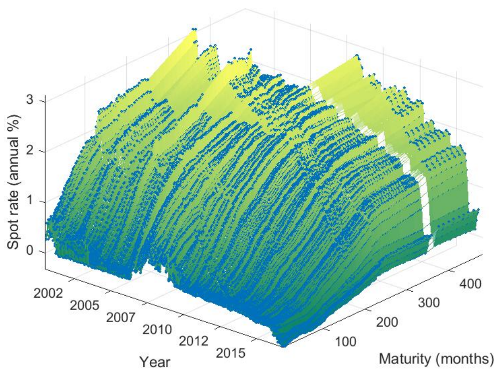
(b) Canada



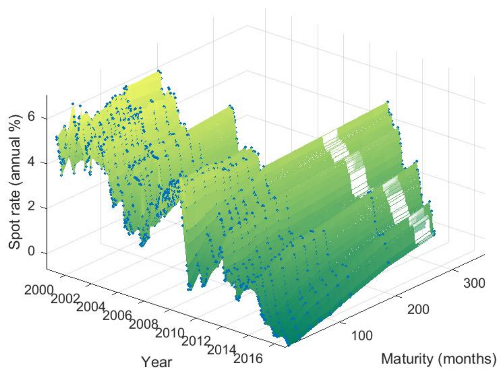
(c) Germany



(d) U.K.

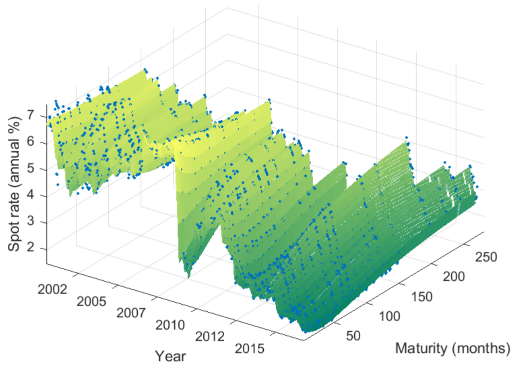


(e) Japan

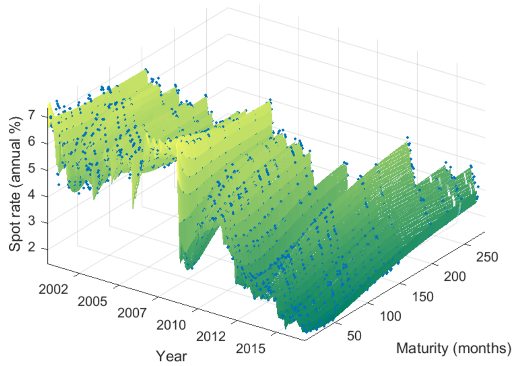


(f) Sweden

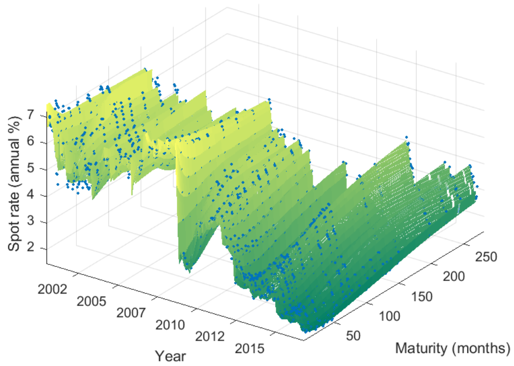
Figure 6: Term structure changes over time per country. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is linearly interpolated across time and maturities.



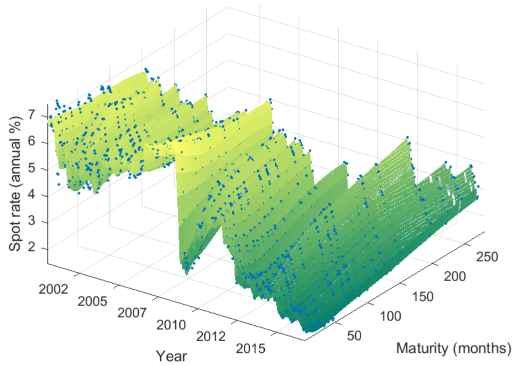
(a) NS1



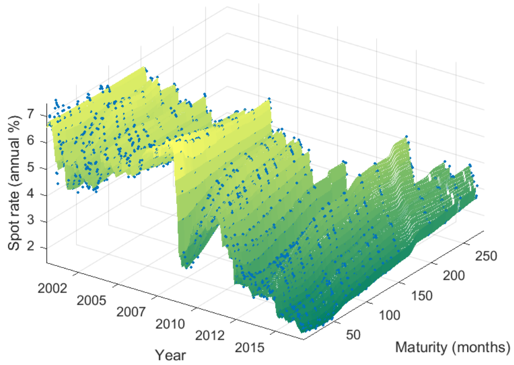
(b) NS2



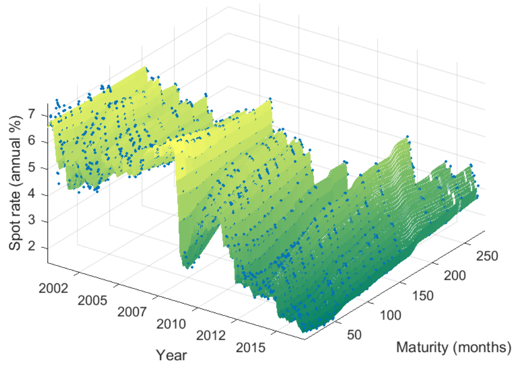
(c) NS3



(d) BS

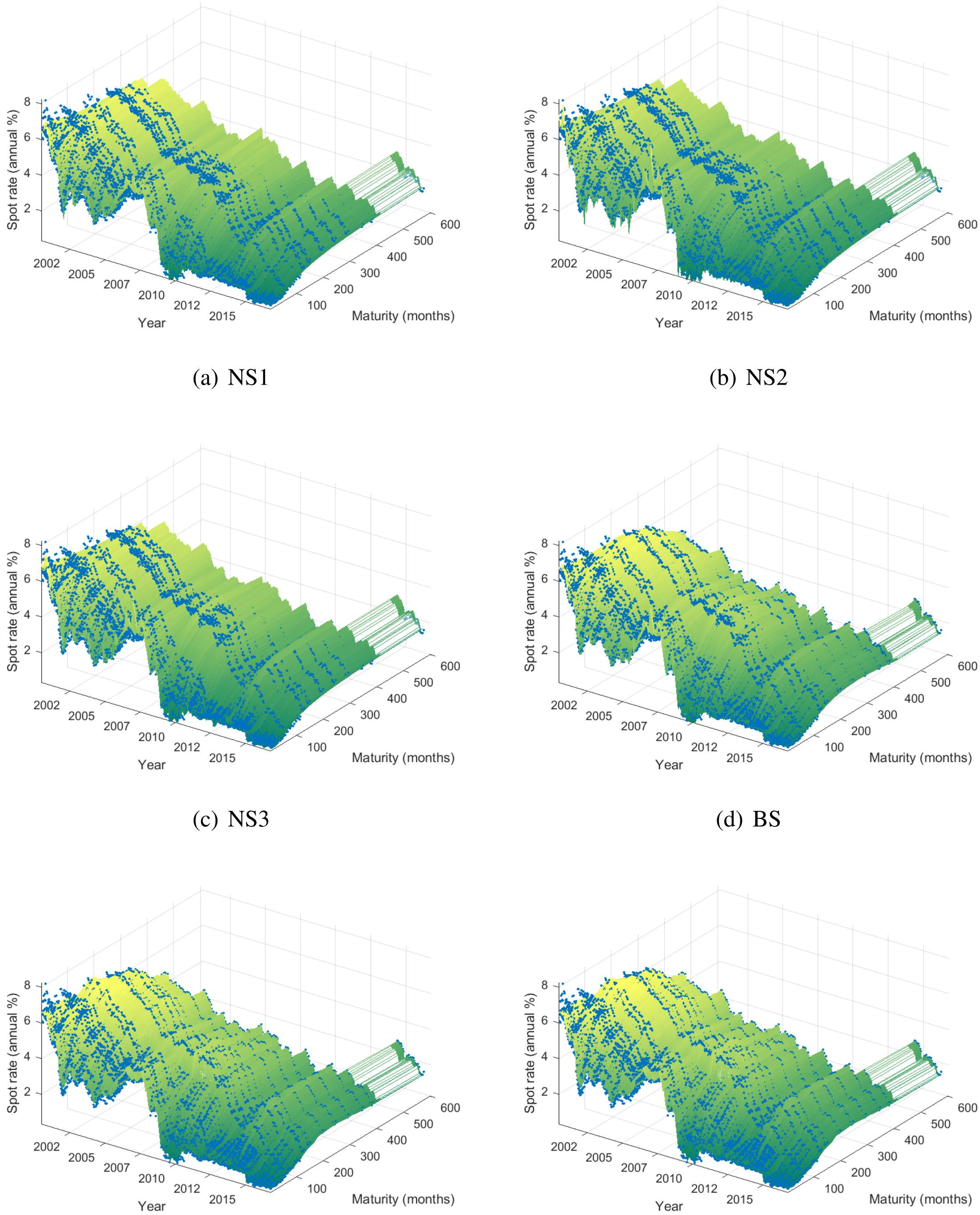


(e) KS1



(f) KS2

Figure 7: Term structure modeled by different yield curve models for Australia. Panel (a) to (f) each contains the term structure modeled by NS1, NS2, NS3, BS, KS1, and KS2, respectively. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is the modeled yield curve across time and maturities.

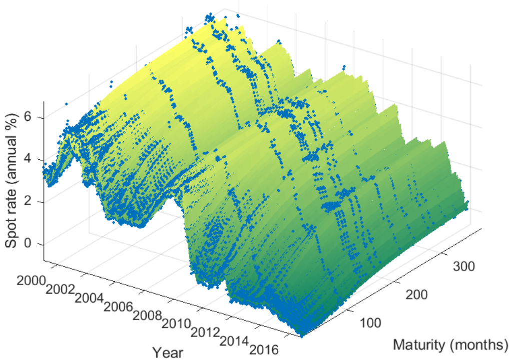


(a) NS1 (b) NS2

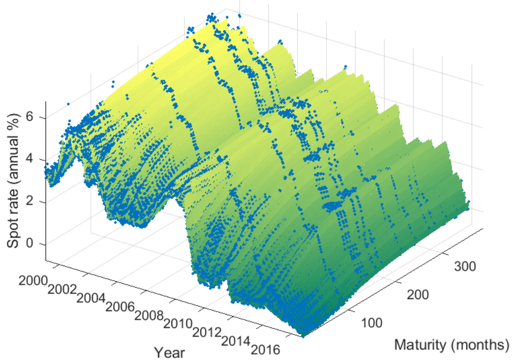
(c) NS3 (d) BS

(e) KS1 (f) KS2

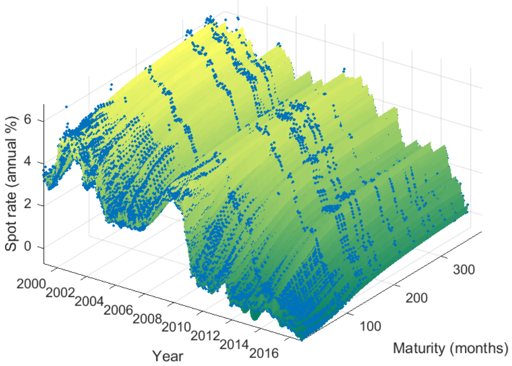
Figure 8: Term structure modeled by different yield curve models for Canada. Panel (a) to (f) each contains the term structure modeled by NS1, NS2, NS3, BS, KS1, and KS2, respectively. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is the modeled yield curve across time and maturities.



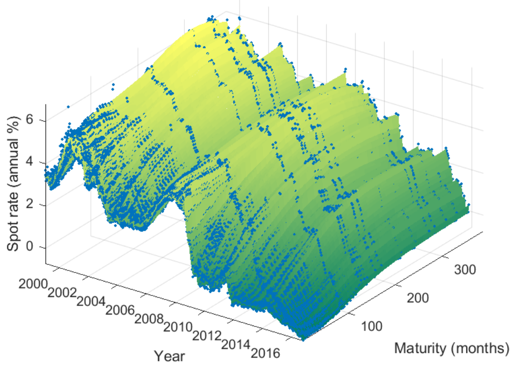
(a) NS1



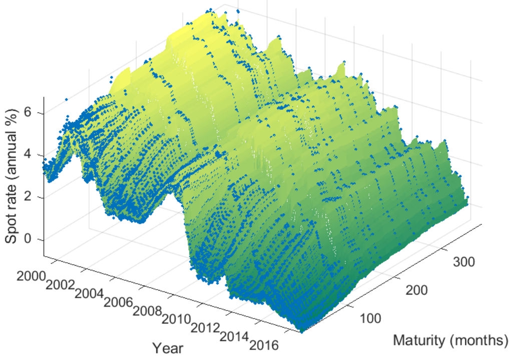
(b) NS2



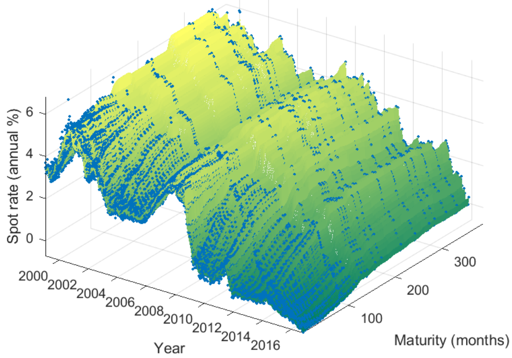
(c) NS3



(d) BS

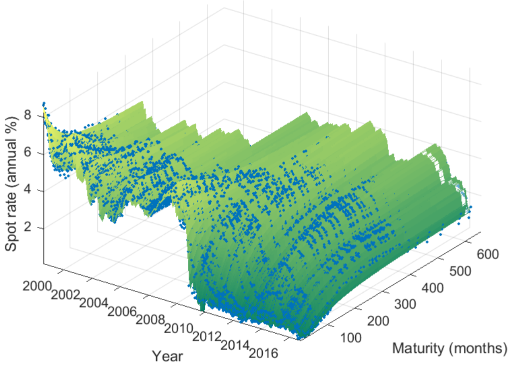


(e) KS1

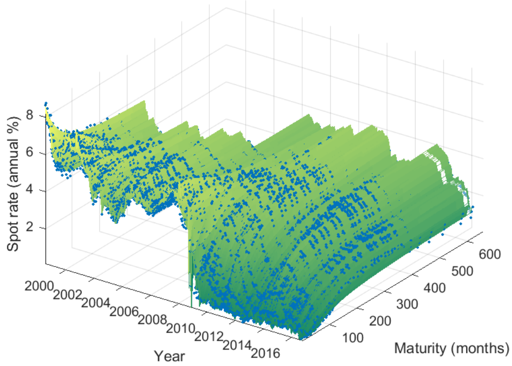


(f) KS2

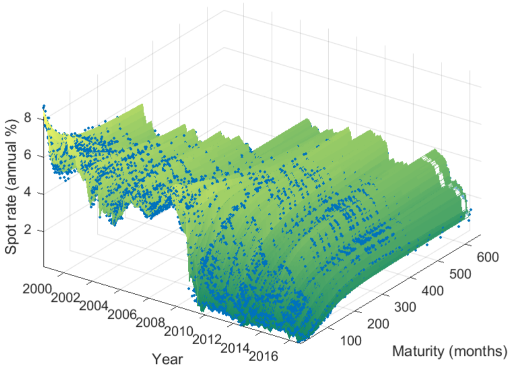
Figure 9: Term structure modeled by different yield curve models for Germany. Panel (a) to (f) each contains the term structure modeled by NS1, NS2, NS3, BS, KS1, and KS2, respectively. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is the modeled yield curve across time and maturities.



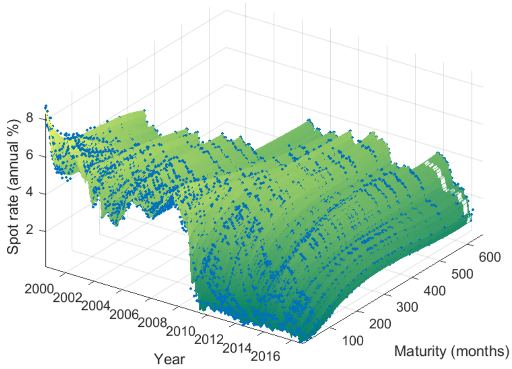
(a) NS1



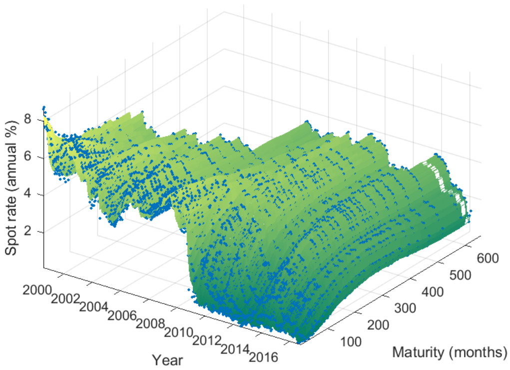
(b) NS2



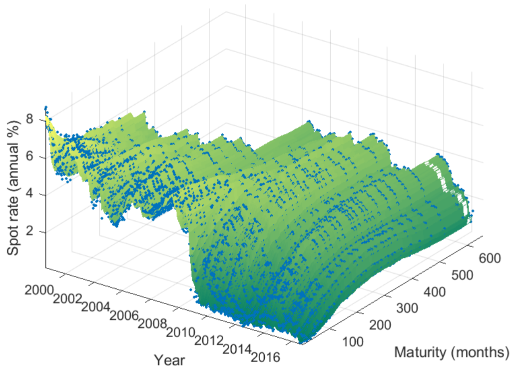
(c) NS3



(d) BS

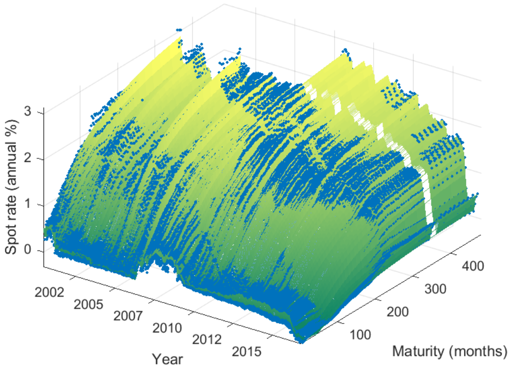


(e) KS1

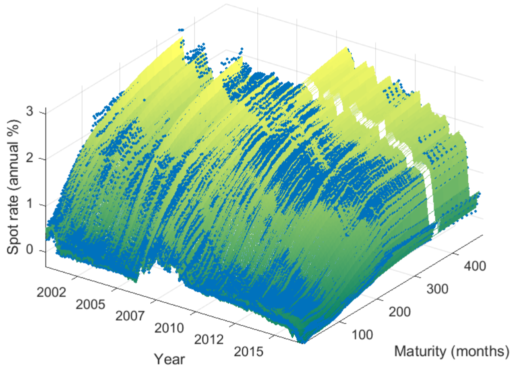


(f) KS2

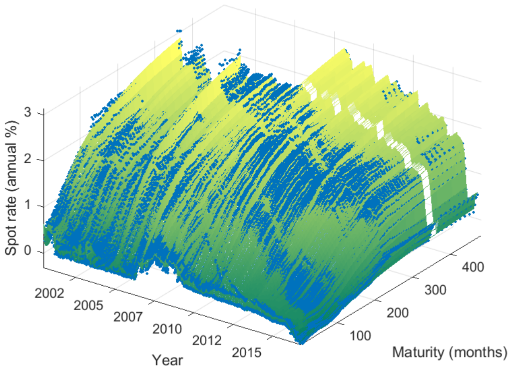
Figure 10: Term structure modeled by different yield curve models for the U.K. Panel (a) to (f) each contains the term structure modeled by NS1, NS2, NS3, BS, KS1, and KS2, respectively. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is the modeled yield curve across time and maturities.



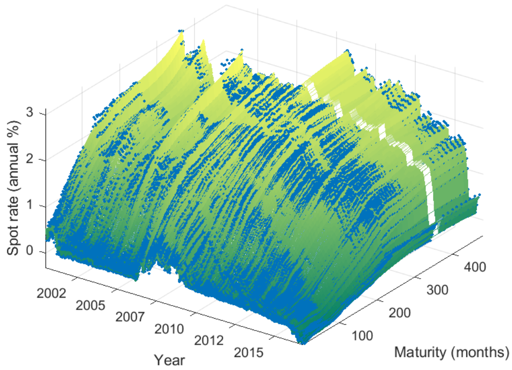
(a) NS1



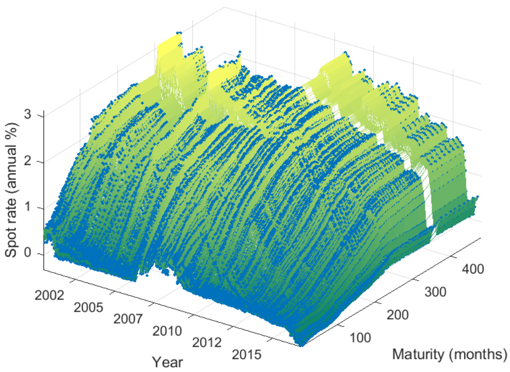
(b) NS2



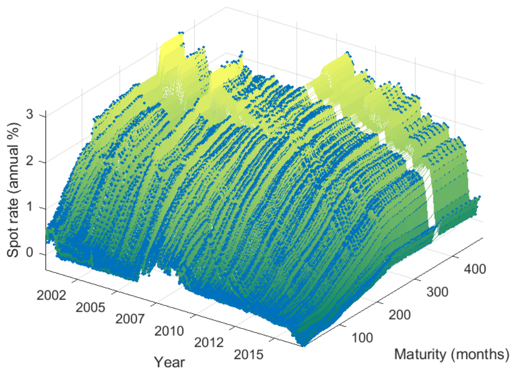
(c) NS3



(d) BS

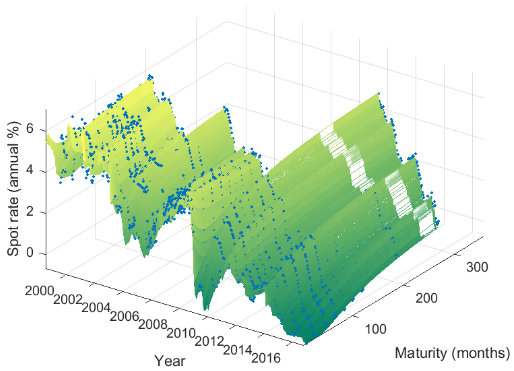


(e) KS1

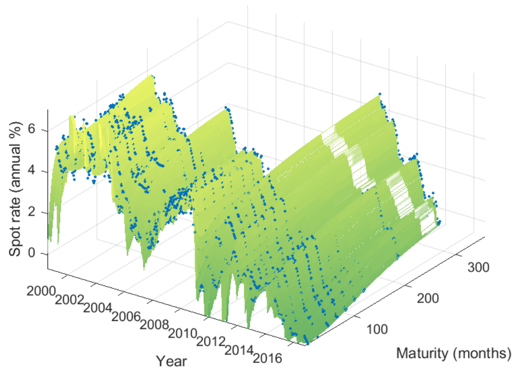


(f) KS2

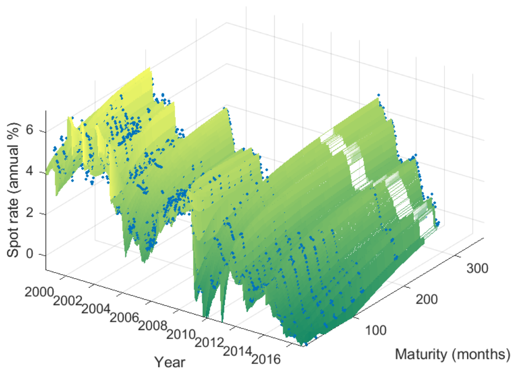
Figure 11: Term structure modeled by different yield curve models for Japan. Panel (a) to (f) each contains the term structure modeled by NS1, NS2, NS3, BS, KS1, and KS2, respectively. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is the modeled yield curve across time and maturities.



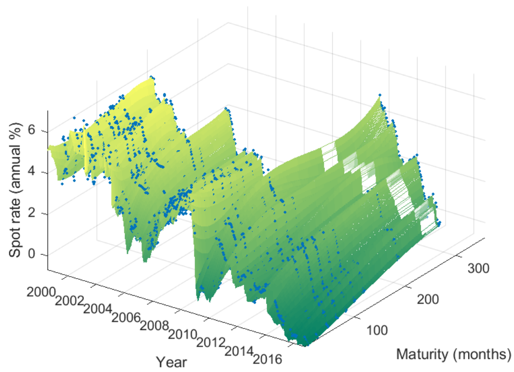
(a) NS1



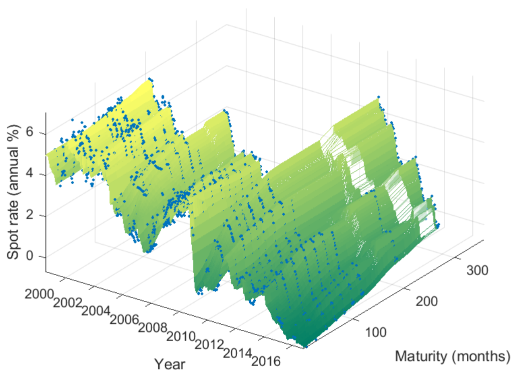
(b) NS2



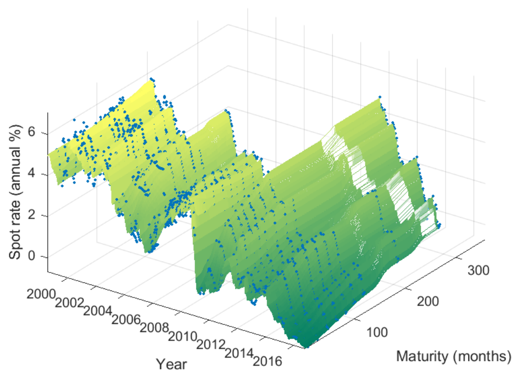
(c) NS3



(d) BS

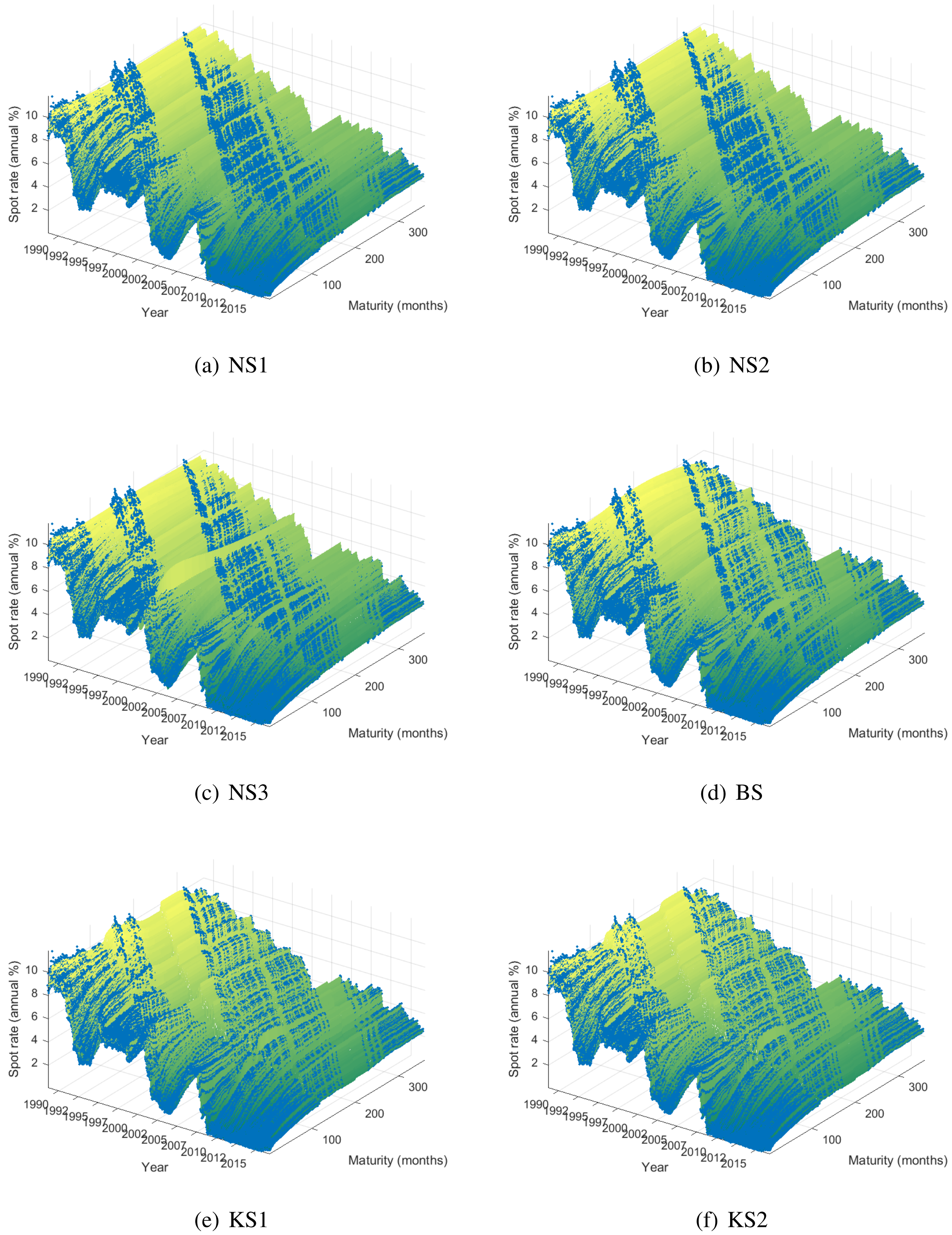


(e) KS1



(f) KS2

Figure 12: Term structure modeled by different yield curve models for Sweden. Panel (a) to (f) each contains the term structure modeled by NS1, NS2, NS3, BS, KS1, and KS2, respectively. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is the modeled yield curve across time and maturities.



(a) NS1 (b) NS2

(c) NS3 (d) BS

(e) KS1 (f) KS2

Figure 13: Term structure modeled by different yield curve models for the U.S. Panel (a) to (f) each contains the term structure modeled by NS1, NS2, NS3, BS, KS1, and KS2, respectively. The blue dots are the end-of-month unsmoothed Fama-Bliss zero-coupon yields and the green mesh is the modeled yield curve across time and maturities.

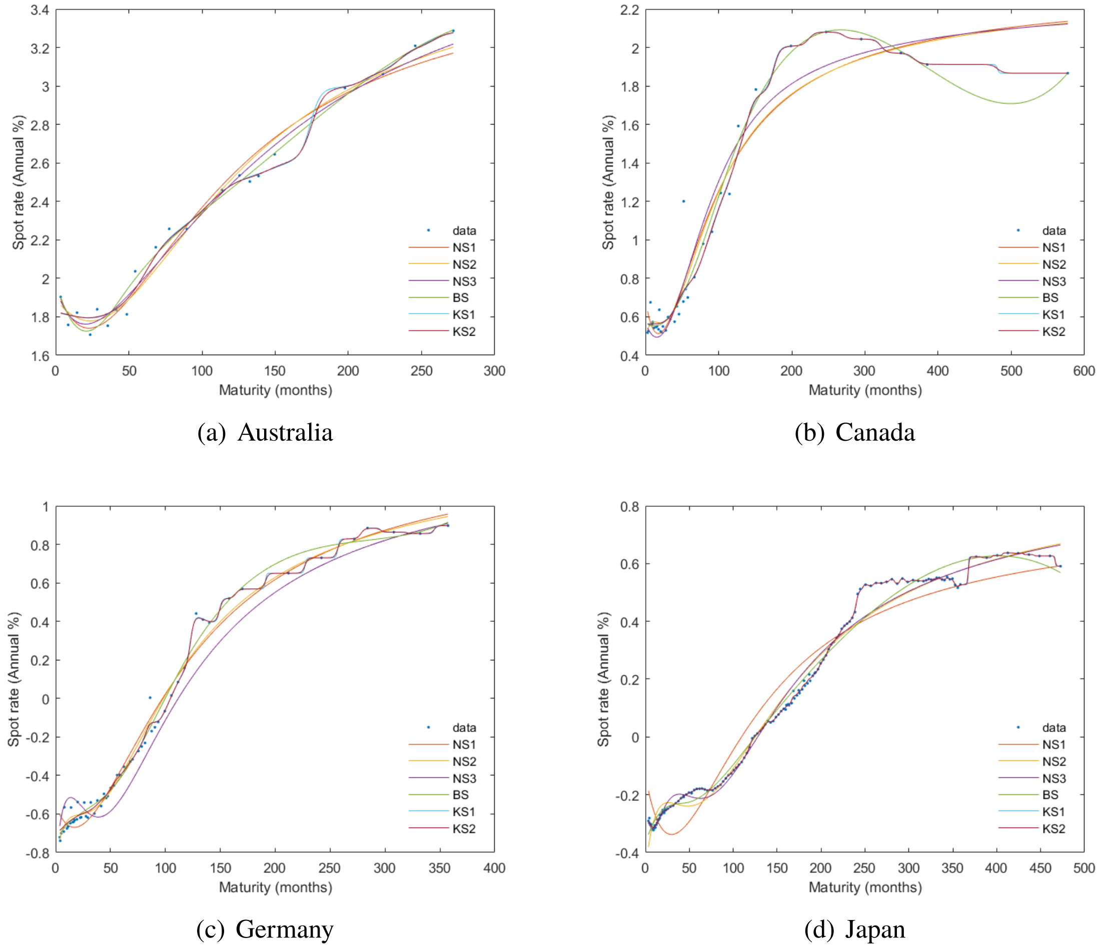
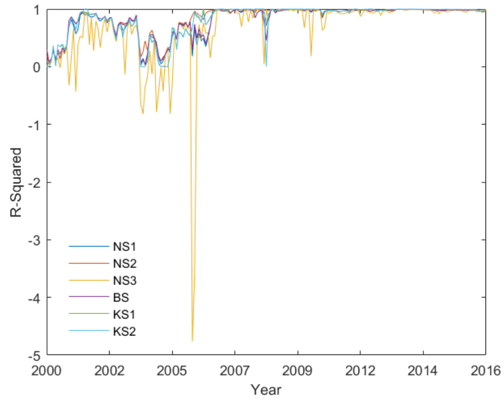
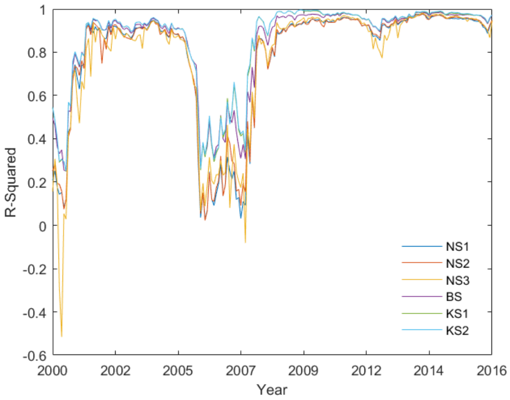


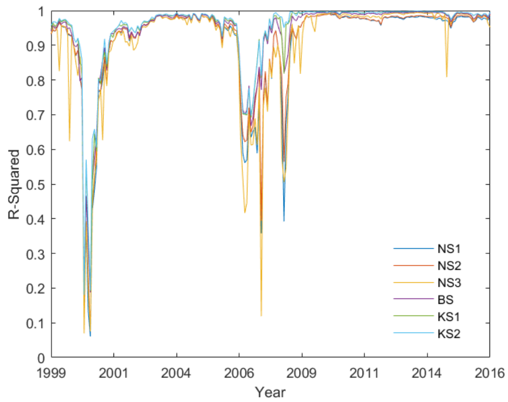
Figure 14: Yield curve modeled by different methods, observed at the end of October, 2016.



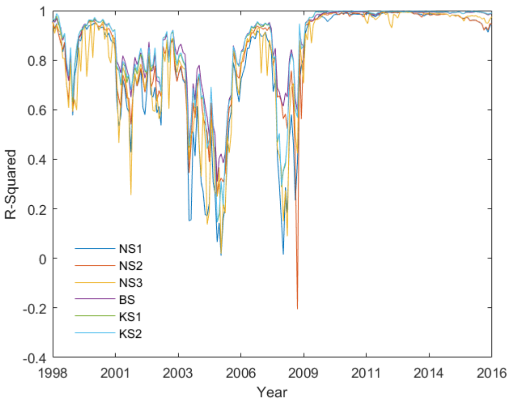
(a) Australia



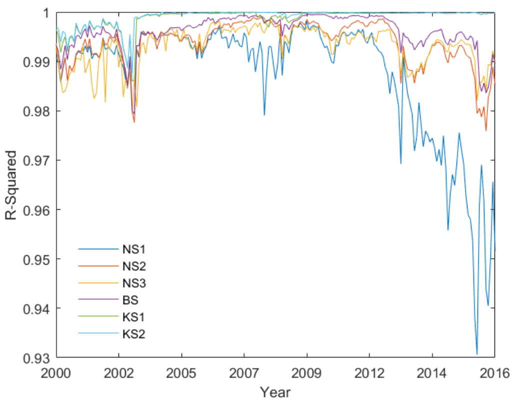
(b) Canada



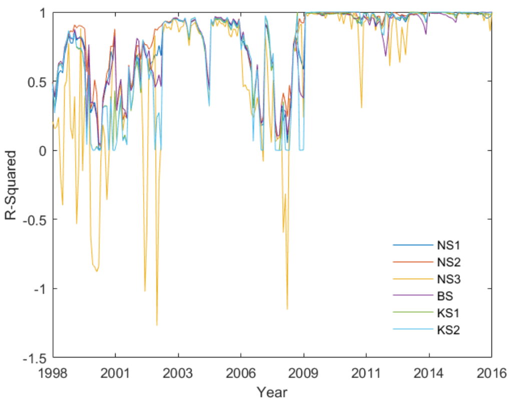
(c) Germany



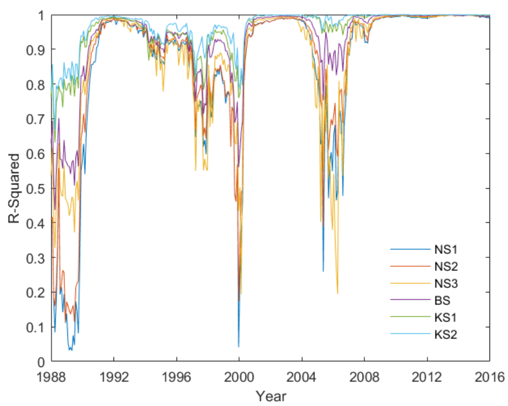
(d) th eU.K.



(e) Japan

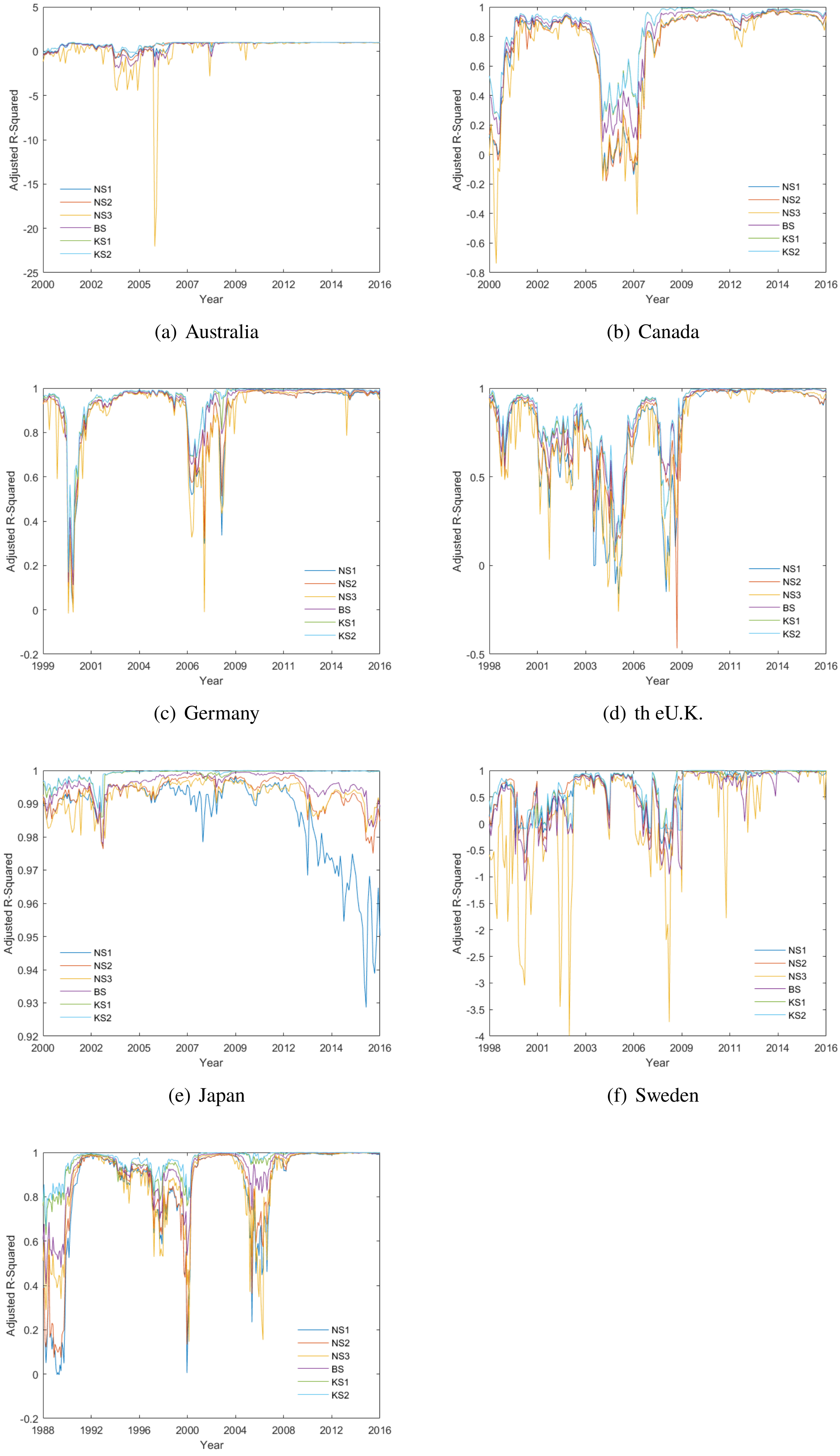


(f) Sweden



(g) the U.S.

Figure 15: R-Squared.



(a) Australia

(b) Canada

(c) Germany

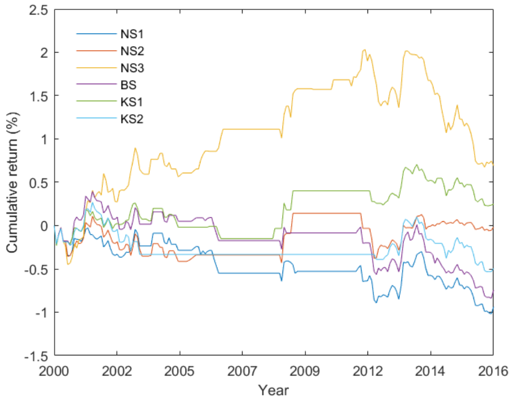
(d) th eU.K.

(e) Japan

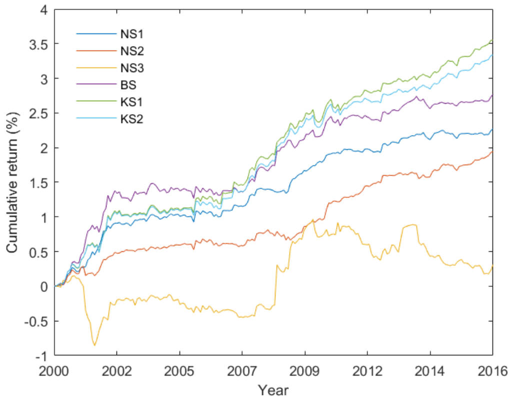
(f) Sweden

(g) the U.S.

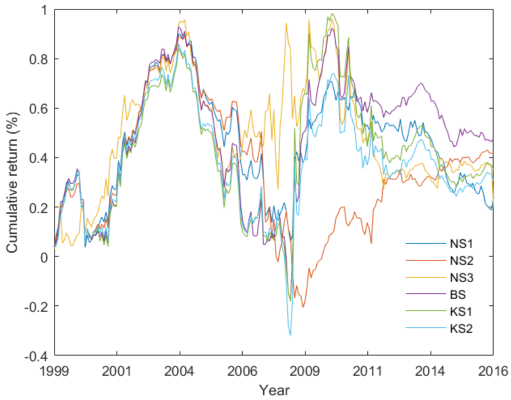
Figure 16: Adjusted R-Squared.



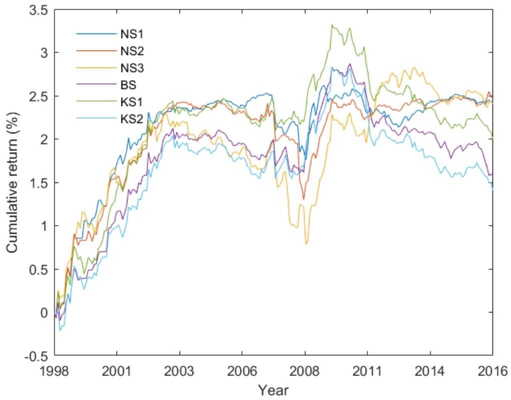
(a) Australia



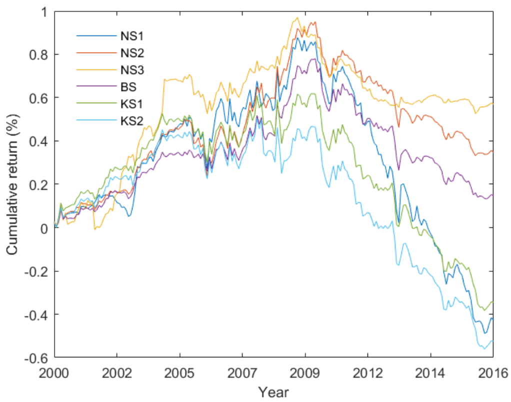
(b) Canada



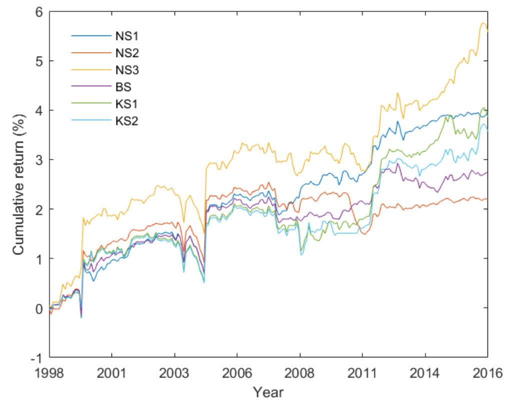
(c) Germany



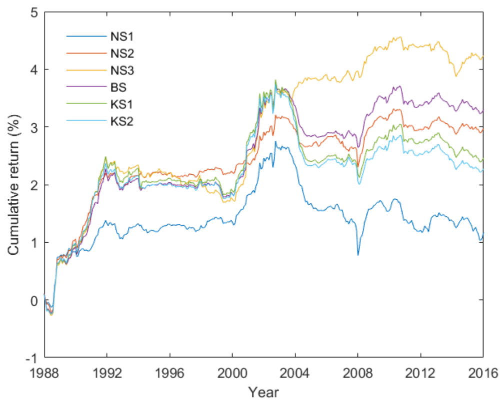
(d) the U.K.



(e) Japan



(f) Sweden



(g) the U.S.

Figure 17: Cumulative returns of generic strategy per country based on different yield curve models.