Nelson-Siegel vs. Constant Spread Bond Price Prediction

Martha Zaverdinos

A Thesis
in
The Department
of
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts at
Concordia University
Montreal, Quebec, Canada

August 2020

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Martha Zaverdinos

Entitled: Nelson-Siegel vs. Constant Spread Bond Price Prediction

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Mathematics & Statistics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Mélina Mailhot

_____ Examiner
Dr. Cody Hyndman

_____ Supervisor
Dr. Frédéric Godin

Approved by: _____
Dr. Galia Dafni, Graduate Program Director
Department of Mathematics and Statistics

_____ 2020 _____
Dr. Francesca Scala, Associate Dean
Faculty of Arts and Science

ABSTRACT

Nelson-Siegel vs. Constant Spread Bond Price Prediction

Martha Zaverdinos

This thesis contributes to the topic of developing methods for bond price predictions over a time horizon of one to five days. This is done by first representing yields as a term structure of spot rates by introducing an alternative Nelson-Siegel model which includes a linearized estimation procedure provided with a new approach in determining a $\lambda - parameter$. We then proceed by comparing this representation of yields to a flat spread structure across maturities. This is followed by applying the latter approaches to the U.S. Treasury, Goldman Sachs and Verizon Communications during the years 2017 and 2018. In order to assess the accuracy of bond price predictions, we compare three distinct approaches for the representation of interest rates and yields. In addition, we want to know if more complex yield curve modeling provides better price predictions than more naive approaches.

Acknowledgement

I would like to thank primarily my supervisor Frédéric Godin for his valuable advice and helpful comments throughout preparing and writing this thesis. He has contributed with impressive knowledge, high expectations, and very supportive and helpful attitude. I am also extremely thankful to my family for the ongoing support during my studies. Finally, I would like to acknowledge the Institut des Sciences Mathématiques for funding support through the ISM Graduate Scholarship.

Contents

List of Figures

List of Tables

1. Introduction

The objective of this thesis is to develop methods for bond price predictions over a time horizon of one to five days. Government or corporate bonds oblige an issuer to repay a bondholder borrowed money at a given point in time. The bond issuer makes payments to the bondholder at a fixed or variable interest rate for a specified period. With respect to a short time horizon, the prediction of the bond market performance is particularly very important to issuers of debt as a useful indicator in order to time their issuance and issue the debt at the right price.

Corporate bond markets have more than doubled in size over the last ten years, reaching a size exceeding that of the Treasury markets (Ericsson, & Reneby, 2005). During the years 1997 to 2000, approximately 70% of new capital raised by the U.S. corporations was in the form of debt.[1] Given the corporate bond market growth, modeling and forecasting the term structure of interest rates are important in financial economics for the reasons of pricing financial assets and their derivatives, allocating portfolios and managing financial risk. With concern to all these key challenges, researchers have developed several models and techniques which involve the dynamic evolution of yield curves (Härdle, & Majer, 2012).

A yield curve is an instrument for portfolio management and for pricing securities (Diebold, & Li, 2006). It is a functional representation of the relationship between spot rates and the time to maturity where a spot rate is the yield quoted for immediate settlement at a given maturity. The yield curve represents the discount rates to a bullet bond maturing

[1]Retrieved from the Securities Industry and Financial Markets Association at www.sifma.org

on a given date. The connection between maturities and spot rates is used as the core assumption when building yield curves. An important tool in the development of financial theory is the term structure of interest rates where the relationship between rates and maturity has proved to be critical to market practitioners (Ioannides, 2003).

According to Stojanovic and Vaughan (1997), since yields for bonds of all maturities change every day due to market fluctuations, the shape of the yield curve is always changing. In light of this, the process of fitting yield curves is a demanding task. The most popular approaches of modeling the term structure of interest rates are equilibrium and no-arbitrage models (Härdle, & Majer, 2012). Longstaff and Schwartz (1992) developed a two-fractor general equilibrium model of the term structure. They derive closed-form expressions for discount bonds and study the properties of the term structure implied by the model. The equilibrium framework provides exact fits to the observed term structure (Longstaff, & Schwartz, 1992). One of the main contributors to no-arbitrage models are Hull and White (1990), where they present two one-state variable models of the short-term interest rate. However, according to Härdle and Majer (2012), both of the above models do not provide a good predictive performance since forecasting is not the main goal of these approaches.

Nevertheless, thanks to the econometricians Nelson and Siegel, the related literature provides an efficient solution to solving the problem of poor predictive performance through a three-factor latent model (Nelson, & Siegel, 1987). Nelson and Siegel (1987) mention that the fitting of yield curves date back to at least the first half of the 20^{th} century. The majority of yield curve models which followed were based on the methodology of multilinear regression and spline interpolation. Without being entirely complex and efficient, these models were able to capture the basic shapes observed in yield curves which include monotonic, humped and S-shaped (Nelson & Siegel, 1987). The Nelson-Siegel model is now widely used, mostly as the foundation for other models. Diebold and Li (2006) proposed the Nelson-Siegel curve with time varying parameters which has gained popularity among financial market practitioners and central banks (Härdle, & Majer, 2012).

Throughout this research, we have used many variations of the Nelson-Siegel model, in-

troduced by Nelson and Siegel (1987), in order to fit the spot rate curves. In light of the fact that the Nelson-Siegel model leads to easy estimation and exhibits computational tractability, both Nelson and Siegel (1987) and Diebold and Li (2006) implement the Nelson-Siegel model on zero coupon bonds only. Generally, coupon bearing bond prices differ tremendously from zero coupon bond prices. In order to better deliver forward price trends, all debt securities used must first be converted to zero-coupon bonds. Gauthier and Simonato (2012) present an alternative approach to estimating the Nelson-Siegel parameters which relies on linearization where the steps to this alternative model are implemented on coupon-bearing bonds. Gauthier and Simonato (2012) describe how the Nelson-Siegel parameter estimation which relies on linearization compares favorably to the original model which includes shorter computing time.

The structure to this thesis starts with summarizing the data used within the research described further in Chapter 2. To follow, Chapter 3 breaks down into two main topics. The first being our initial method which is based on representing the risk-free rates and spread as a term structure of spot rates. This is done by initially including a detailed description of the Nelson-Siegel model followed by applying Gauthier and Simonato's (2012) linearized estimation procedure to both government and corporate bonds. Finally, we provide a new insight which introduces our alternative model for determining a $\lambda - parameter$ similarly done in Diebold and Li (2006). The second main topic introduces a method which uses spreads that are idiosyncratic for all securities and flat across maturities compared to a single curve applied to all tenors. We then proceed by comparing three distinct approaches for the representation of interest rates and yields in Chapter 4. To obtain such predictions, we forecast bond yields instead of prices since the latter are more stationary. The first two approaches are done by using the two aforementioned representation of bond yields within our forecasting procedures. Finally, the last approach involves directly forecasting corporate yields-to-maturity. Although this last approach is not theoretically sound, it is used only for comparison with the forecasts generated by sophisticated techniques as in the first two approaches.

Results are then provided in Chapter 5. For both yield curve estimation and price forecasting we compare all approaches among themselves and express our findings via framework

of tables and figures with short comments on each. In terms of price forecasting errors, we take advantage of five forecasting methods. The first method entails a naive benchmark model with the simple assumption that the best forecast corresponding to the quantity of interest in subsequent days is its current value. Next we consider two different linear approaches to model the relationship between the time differences of a process or between values of a process itself. Finally, similarly to the linear approaches, we implement an AR(1) model to the time difference of a process or alternatively a process itself. Results show that constructing a single curve for all bonds is the least favorable. In addition, modeling the linear relationship of a process through Ordinary Least Squares regression performs best in comparison to the other four price forecasting methods. Finally, we end this thesis with Chapter 6, labelled as Conclusion, where we provide a short discussion on obtained results and put them into wider economical context.

2. Data

Data used for this research was retrieved by Overbond, an Artificial Intelligence (AI) company. In this study we employ various variables retrieved from Overbond's centralized hub which are pooled from different data sources. The data used was extracted from the following sources: Trade Reporting and Compliance Engine (TRACE), Bloomberg and Refinitiv. Firstly, TRACE reports over-the-counter transactions for the eligible fixed-income securities. Secondly, the Bloomberg Terminal is a computer software system through which users can monitor and analyze real-time financial market data. Finally, Refinitiv is a global provider of financial market data and infrastructure.

As the frequency of corporate bond transactions are low, in order to get a reasonably large sample, we chose bonds that traded daily. In light of this, we use daily traded securities issued during the years 2017 and 2018 by a multinational American investment bank and financial services, the Goldman Sachs Group, Inc. All data used includes dates within January 3^{rd}, 2017 to December 31^{st}, 2018 with a total of 499 trading days. In addition to the firm data, to implement the Nelson-Siegel model, the risk-free interest rates are needed in order to discount the cash flows. Therefore, we also use the U.S. Treasury (UST) daily traded securities during the same dates within the years 2017 and 2018.

The observed data, in both the corporate and Treasury datasets, include the clean daily closing bid and ask yields-to-maturity, coupon class, coupon frequency, principal amount (all \$100), currency, seniority, instrument type, indicator for convertible bonds, indicator for callable bonds, issue date and maturity date. When collecting data, bonds that did not meet the following criteria were excluded from our dataset:

- Fixed coupons;
- USD currency;
- Unsecured and Senior Unsecured notes and bonds pertaining to the Treasury and Goldman Sachs datasets respectively;
- Nonconvertible bonds;
- Noncallable bonds;
- Bonds with a total initial amount issued more than $1 million;
- Bonds which are traded on each 499 days previously mentioned;
- Yields-to-maturity below 700 basis points;
- More than 0.5 years between the trade date and the maturity date;
- Bonds that were issued after 1999.

After all data cleaning was applied, a total of 99 UST bonds and 409 Goldman Sachs bonds were ultimately retained in the final dataset which are traded on every day of the sample. This being said, the convention used in the current thesis is that the i^{th} bond on day t is equivalent to the i^{th} bond on day $t+1$.

For each bond traded daily, we use the closing mid yields which are computed as the average between the clean closing bid and ask yields-to-maturity. Since the bond prices are not observed, they are computed by using the formula below which is used strictly for recovering prices as per quoting convention from the data provider Refinitiv:

$$B = \frac{C}{n} \times 100 \times \frac{\left[1 - \left(1 + \frac{YTM}{n}\right)^{-nT}\right]}{YTM/n} + 100 \left(1 + \frac{YTM}{n}\right)^{-nT}. \tag{2.1}$$

In (2.1), for each corresponding bond, C represents the coupon rate, n is the coupon frequency, YTM is the clean closing mid yield-to-maturity observed on the trade date and T represents the time in years between the trade date and the maturity date based on a 365 day-count convention. These computed mid prices correspond to the clean prices.

In order to obtain a true reflection of the bond value, accrued interest is added to (2.1). This leads to the prices used within this research and otherwise known as the dirty price.

3. Term Structure Representation of Bond Yields

The ultimate objective of this thesis is to obtain price predictions for bonds over a time horizon of one to five days. According to Bauer and Rudebusch (2017), standard finance models of the yield curve assume that interest rates are stationary, where the expectations component of long-term interest rates varies only modestly. Given this, to obtain predictions, we proceed by forecasting bond yields instead of prices since the latter are more stationary. The current chapter considers two methods in representing the term structure of bond yields for such purposes. To begin, the yields are represented as a term structure of spot rates. Furthermore, spot rates are broken down into a risk-free rate component and a spread component. Under both methods, the whole term structure of the risk-free component of spot rates is considered and modeled through treasury securities. Conversely, the spread component is obtained by using corporate bonds. The first method decomposes the spread into a term structure, whereas it is idiosyncratic to each bond in the second method. The second method involves a flat spread across maturities.

To begin, since the datasets used cover multiple days where each day includes a cross-section of bonds, the price of the i^{th} bond on day t can be defined as

$$B_{i,t} = \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} c_{i,t}^{(j)} \times P(t, \tau_{i,t}^{(j)}), \tag{1}$$

where

$$P(t, \tau_{i,t}^{(j)}) = e^{-y_i(t,\tau_{i,t}^{(j)})\tau_{i,t}^{(j)}}. \tag{2}$$

The vector $\boldsymbol{\tau}_{i,t} = \left(\tau_{i,t}^{(1)}, \ldots, \tau_{i,t}^{(m_{i,t})}\right)$ represents the times-to-maturity in years corresponding to all remaining cash flows from the bond, where the j^{th} cash flow among these is denoted $c_{i,t}^{(j)}$, $j = 1, \ldots, m_{i,t}$ and $m_{i,t}$ is the total number of remaining cash flows for the i^{th} bond in the cross-section of a given day t. In addition, $\boldsymbol{c}_{i,t} = \left(c_{i,t}^{(1)}, \ldots, c_{i,t}^{(m_{i,t})}\right)$ represents the vector of cash flows corresponding to the i^{th} bond on day t. Finally, (2) represents the discount factor[2] and $y_i(t, \tau_{i,t}^{(j)})$ is the spot rate corresponding to the j^{th} time-to-maturity, denoted $\tau_{i,t}^{(j)}$, of the i^{th} bond[3] on day t.

In addition, the times-to-maturity corresponding to each cash flow per bond is obtained in years by working back from the maturity date of each bond depending on the coupon frequency. We work backwards in order to compute the timing of cash flows since the last cash flow is paid on the maturity date.

For the various models considered, a breakdown of the corporate spot rates is viewed by separation between a risk-free and a spread component denoted as

$$y_i(t, \boldsymbol{\tau}_{i,t}) = y_i^{(gov)}(t, \boldsymbol{\tau}_{i,t}) + y_i^{(spread)}(t, \boldsymbol{\tau}_{i,t}).$$

Within our dataset, the spot rates associated to each unique bond i for all days included in our sample are not observed and must be obtained through modeling. A two-step procedure is followed by first obtaining risk-free spot rates, and only then obtaining spreads based on results from the first step. These steps are done by applying the Nelson-Siegel model. Since the bonds in our datasets are coupon bearing bonds, we then proceed with a detailed estimation procedure provided by Gauthier & Simonato (2012) which incorporates an estimation relying on linearization. Gauthier & Simonato (2012) show that the linearized estimation procedure saves on computational time when running a simulation. It is also highlighted by Sundaram & Das (2010) that a problem encountered with Nelson-Siegel modeling is the possibility of mulitple local optima due to the nonlinearities in the parameters. In this regard, the possibility of finding several "optimal" solutions is

[2]The second argument in the discount function $P(t, \tau_{i,t}^{(j)})$ represents the time-to-maturity and not the maturity date itself.

[3]Under the first method provided in this section, spot rates are the same for all bonds in a given day t. Therefore, dependence on i can de dropped for that specific method.

not uncommon. In light of these issues, we provide a new approach in order to stabilize our daily Nelson-Siegel paramaters.

3.1 The Nelson-Siegel Model

This section reviews the Nelson-Siegel model in order to fit daily risk-free curves. We first describe the original Nelson-Siegel base case model introduced by Nelson and Siegel (1987) and then proceed by illustrating the linearized estimation approach proposed by Gauthier & Simonato (2012) in order to optimize the curve parmaters. Finally, in comparison to the approach in Diebold and Li (2006) of fixing the λ parameter to a predetermined value, we introduce an alternative way in order to optimize bond price predictions.

To begin, the Nelson-Siegel function, which will subsequently be used in diverse ways to model spot rates, is defined as

$$y(\tau;\boldsymbol{\theta}) = \beta_0 + \beta_1\phi_1(\tau;\lambda) + \beta_2\phi_2(\tau;\lambda), \tag{3}$$

where τ is the time-to-maturity expressed in years, $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \lambda)$ is a vector of unknown parameters,

$$\phi_1(\tau;\lambda) = \tfrac{\lambda}{\tau}\left(1 - e^{\frac{-\tau}{\lambda}}\right),$$

and

$$\phi_2(\tau;\lambda) = \left(\tfrac{\lambda}{\tau}\left(1 - e^{\frac{-\tau}{\lambda}}\right) - e^{\frac{-\tau}{\lambda}}\right).$$

As described by Gauthier & Simonato (2006), the parameters in $\boldsymbol{\theta}$ can be interpreted through the behavior of $\phi_1(\tau;\lambda)$ and $\phi_2(\tau;\lambda)$ as τ varies. Indeed, as $\tau \to \infty$, $\phi_1(\cdot)$ and $\phi_2(\cdot)$ tend to 0, and therefore, $y(\tau;\boldsymbol{\theta})$, tends to β_0. Hence, β_0 can be viewed as the long-term factor and the level of the spot rate curve. Using the same methodology as above, we can conclude that β_1 represents the short-term factor and is characterized as

the slope of the spot rate curve. Thirdly, β_2 represents the medium-term factor and the curvature of the spot rate curve. Finally, as stated by Diebold and Li (2006), λ governs the exponential decay rate. Small values for λ produce low decay and can better fit the curve for larger τ values whereas large values for λ can better fit the curve for smaller τ values.

3.1.1 Nelson-Siegel Linearized Estimation

Gauthier & Simonato (2012) describe how a Nelson-Siegel parameter estimation which relies on linearization compares favorably to the original estimation approach which includes a naive multivariate optimization of paramaters. For the purpose of this work, the linearization procedure is only used to obtain starting parameters in our optimization. Gauthier & Simonato (2012) demonstrate that this procedure easily provides acceptable starting parameters that saves on computational time when running a simulation. To summarize, once we obtain reasonable starting parameters we then proceed with a numerical optimization, which does not rely on the linear approximation in order to achieve further accurate results.

As described in Gauthier & Simonato (2012), the Nelson-Siegel price associated to the i^{th} bond on day t can be defined as

$$B(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t}; \boldsymbol{\theta}) = \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} c_{i,t}^{(j)} \times P(t, \tau_{i,t}^{(j)}; \boldsymbol{\theta}), \tag{4}$$

In comparison to (2), $P(t, \tau_{i,t}^{(j)}; \boldsymbol{\theta}) = e^{-y(t,\tau_{i,t}^{(j)};\boldsymbol{\theta})\tau_{i,t}^{(j)}}$ represents the discount factor where the spot rate $y(t, \tau_{i,t}^{(j)}; \boldsymbol{\theta})$ is now a function of the vector of parameters $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \lambda)$.

For each day t, a separate estimation of paramaters is computed, and therefore we introduce the time-dependent vector of paramaters $\boldsymbol{\theta}_t = (\beta_{0,t}, \beta_{1,t}, \beta_{2,t}, \lambda_t)$. Parameters in $\boldsymbol{\theta}_t$ are estimated by minimizing the weighted average of squared errors for each day t defined as

$$Q(\boldsymbol{\theta}_t) = \frac{1}{N}\left(\sum_{i=1}^{N}\left[\left(B(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t}; \boldsymbol{\theta}_t) - B_{i,t}^{(obs)}\right) \times \frac{1}{D_{i,t}}\right]^2\right), \tag{5}$$

where

$$D_{i,t} = \left[\frac{1}{B_{i,t}^{(obs)}} \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} \tau_{i,t}^{(j)} c_{i,t}^{(j)} \left(1 + \frac{YTM_{i,t}}{n_i}\right)^{-n_i \tau_{i,t}^{(j)}}\right] \times \left(1 + \frac{YTM_{i,t}}{n_i}\right)^{-1} \tag{6}$$

and N denotes the total number of bonds traded daily. In (6), $D_{i,t}$ represents the modified duration, $B_{i,t}^{(obs)}$ is the observed bond price, $YTM_{i,t}$ is the yield-to-maturity and n_i represents the coupon frequency[4] of the i^{th} bond on day t. According to Gauthier & Simonato (2012), weighting the prices by inverse durations produces a better fit in spot rates and decreases the computing time by avoiding the extra numerical work required by the spot rate conversion. Bonds with higher durations have prices which are more sensitive to its yield-to-maturity and thus are more likely to produce higher errors. Therefore, in order to avoid the mean squared error being driven more heavily by this category of bonds, a weighting by duration is performed.

Gauthier & Simonato (2012) adapt a parameter estimation approach which cuts down the number of parameters over which a non-linear search must be performed. This approach is obtained through an approximation of the Nelson-Siegel price in (4). In order to linearize the discount factor $P(t, \tau; \boldsymbol{\theta}) = exp\{-y(t, \tau; \boldsymbol{\theta}) \times \tau\}$ in (4), Gauthier & Simonato (2012) incorporate the Taylor approximation defined as $e^x \approx 1 + x$, where this approximation is true only if x is in the neighborhood of zero. In this case, x is defined as $y(t, \tau; \boldsymbol{\theta}) \times \tau$. Since τ is not in the neighborhood of zero, especially for long maturities, Gauthier & Simonato (2012) proceed by introducing a parameter φ incorporated into the discount factor which is then rewritten as

$$P(t, \tau; \boldsymbol{\theta}) = exp\{-\varphi\tau\}\, exp\{(\varphi - y(t, \tau; \boldsymbol{\theta}))\, \tau\}. \tag{7}$$

In general, the parameter φ is meant to represent a proper centering point for our Taylor

[4]Coupon frequency is represented as number of coupons per year.

expansion in (7). To successfully linearize the discount function in (4), the expression $(\varphi - y(t, \tau; \boldsymbol{\theta}))\,\tau$ in the second exponential of (7) must be as close as possible to zero in order to incorporate the Taylor approximation previously defined. In light of this, Gauthier & Simonato (2012) propose to fix $\varphi_{0,t} = \frac{1}{N}\sum_{i=1}^{N} YTM_{i,t}$, where $YTM_{i,t}$ represents the observed yield-to-maturity of the i^{th} bond on each given day t.

By fixing the parameters φ and λ to some values φ_0 and λ_0, the Nelson-Siegel price in (4) is now of linear form and defined as

$$\begin{aligned} B(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t}; \boldsymbol{\theta}_t) &= \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} c_{i,t}^{(j)} exp\left\{-y\left(t, \tau_{i,t}^{(j)}; \boldsymbol{\theta}_t\right)\tau_{i,t}^{(j)}\right\}, \\ &= \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} c_{i,t}^{(j)} exp\left\{-\varphi_0 \tau_{i,t}^{(j)}\right\} exp\left\{\left[\varphi_0 - y\left(t, \tau_{i,t}^{(j)}; \boldsymbol{\theta}_t\right)\right]\tau_{i,t}^{(j)}\right\}. \end{aligned}$$

where the Nelson-Sigel spot rate function, defined in (3), is now linear since λ_0 is set as a fixed value. Finally, the approximated linearized transformation of the Nelson-Siegel bond price allows for the optimization of the other three paramaters in $\boldsymbol{\theta}$.

As defined in Gauthier & Simonato (2012), Appendix A.1 shows how the linearization previously introduced leads to an approximate system of the form $\mathbf{Y}_{\varphi_0} = \mathbf{X}_{\varphi_0,\lambda_0}\boldsymbol{\theta}_{\varphi_0,\lambda_0} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} = \top(\varepsilon_1, \ldots, \varepsilon_N)$ is a vector of error terms and expressions $\mathbf{Y}_{\varphi_0}$ and $\mathbf{X}_{\varphi_0,\lambda_0}$ are further defined in Appendix A.1.

Finally, given the fixed values for φ_0 and λ_0, the least squares estimator for $\boldsymbol{\theta}_{\varphi_0,\lambda_0}$ is:

$$\hat{\boldsymbol{\theta}}_{\varphi_0,\lambda_0} = \left(\mathbf{X}_{\varphi_0,\lambda_0}^T \mathbf{X}_{\varphi_0,\lambda_0}\right)^{-1} \mathbf{X}_{\varphi_0,\lambda_0}^T \mathbf{Y}_{\varphi_0},$$

where $\boldsymbol{\theta}_{\varphi_0,\lambda_0}$ now depends on parameters φ_0 and λ_0.

As previously mentioned, because this linearized approach relies on an approximation, we proceed with a full optimization using the parameters $\hat{\boldsymbol{\theta}}_{\varphi_0} = (\beta_{0,\varphi_0,\lambda_0}, \beta_{1,\varphi_0,\lambda_0}, \beta_{2,\varphi_0,\lambda_0}, \lambda_0)$ as initial values which then minimizes the weighted average squared errors in (5). In such

optimization, bonds are priced in (5) according to their exact formula (4) which does not rely on linearization. This being said, the optimization used to obtain $\hat{\boldsymbol{\theta}} = \left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\lambda}\right)$, which is an implementation of that of Nelder & Mead (1965), gives further accurate results. The above steps are repeated for each day t in order to obtain an optimal set of parameters, denoted as $\hat{\boldsymbol{\theta}}_t = \left(\hat{\beta}_{0,t}, \hat{\beta}_{1,t}, \hat{\beta}_{2,t}, \hat{\lambda}_t\right)$, corresponding to each trading day.

3.1.2 New Approach in Determining a λ-Parameter

When estimating the Nelson-Siegel paramaters using a linearized estimation procedure as defined by Gauthier & Simonato (2012) and described in Section 3.1.1, the results lead to daily instability in the paramater estimate $\hat{\lambda}_t$ for each day t. This can be seen in Figure 3.1.1 where each estimated spot rate curve parameter associated to the U.S. Treasury data are plotted. In light of this, a new approach is applied by using a grid search to obtain a single fixed optimal λ^* parameter used for all days within the dataset. In order to obtain a fixed optimal λ^*, the steps described in Section 3.1.1 are repeated daily as whole for various fixed values of λ.

To begin, we repeat all steps provided in Section 3.1.1 given a fixed value λ_j, where $j = 1, 2, \ldots, k$ and k represents the total number of fixed λ values within our grid. With a fixed value for φ_0, previously defined as the average observed yields-to-maturity, and λ_j, the least squares estimator for $\boldsymbol{\theta}_{\varphi_0,\lambda_j}$ is

$$\hat{\boldsymbol{\theta}}_{\varphi_0,\lambda_j} = \left(\mathbf{X}^T_{\varphi_0,\lambda_j}\mathbf{X}_{\varphi_0,\lambda_j}\right)^{-1}\mathbf{X}^T_{\varphi_j,\lambda_0}\mathbf{Y}_{\varphi_0},$$

Similar to Section 3.1.1, the estimated parameters $\hat{\boldsymbol{\theta}}_{\varphi_0} = \left(\hat{\beta}_{0,\varphi_0,\lambda_j}, \hat{\beta}_{1,\varphi_0,\lambda_j}, \hat{\beta}_{2,\varphi_0,\lambda_j}\right)$, which now does not include λ_j since this value stays fixed, are used as initial parameters and optimized using an implementation of that of Nelder & Mead (1965) to obtain $\hat{\boldsymbol{\theta}} = \left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \lambda_j\right)$. This is done for each day t using the same fixed λ_j in order to obtain the daily parameter estimates denoted as $\hat{\boldsymbol{\theta}}_t = \left(\hat{\beta}_{0,t}, \hat{\beta}_{1,t}, \hat{\beta}_{2,t}, \lambda_j\right)$.

Finally, we compute the estimated price MSE for all bonds $i = 1, \ldots, N$ associated to each day t as

$$MSE_{\lambda_j,t} = \frac{1}{N}\left(\sum_{i=1}^{N}\left[\left(B(\boldsymbol{c}_{i,t},\boldsymbol{\tau}_{i,t};\boldsymbol{\theta}_{\lambda_j,t}) - B_{i,t}^{(obs)}\right)\times\frac{1}{D_i}\right]^2\right).$$

Therefore, given there are T days in the sample dataset, the total estimated price MSE associated to fixed λ_j is defined as

$$MSE_{\lambda_j}^{(TOT)} = \frac{1}{T}\sum_{t=1}^{T} MSE_{\lambda_j,t}.$$

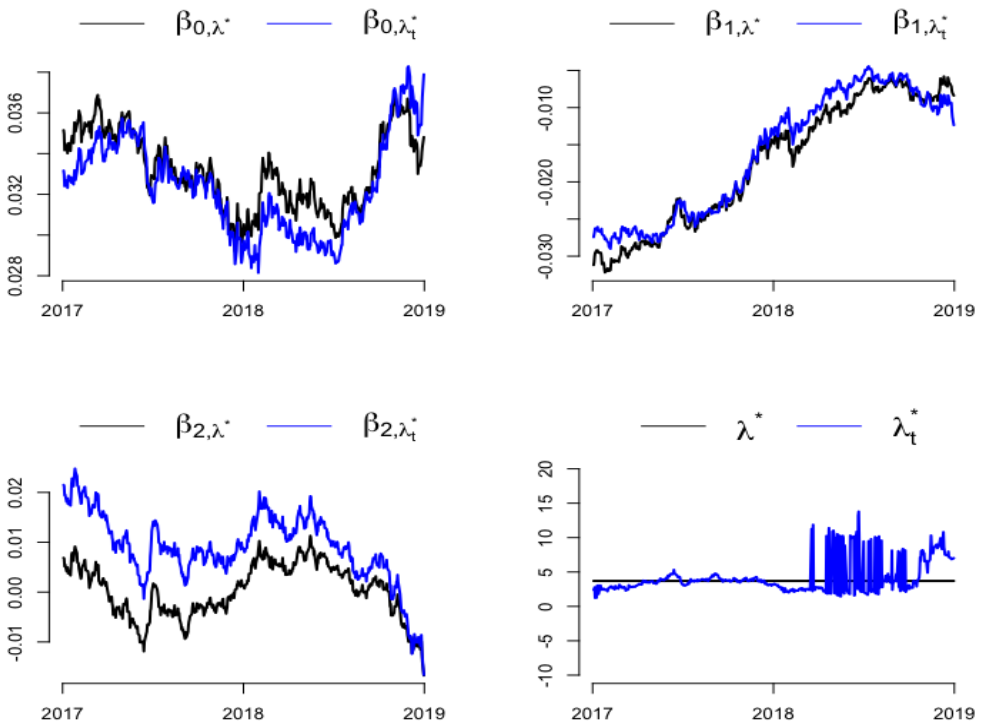
The optimal paramater λ^* is then defined as

$$\lambda^* = \underset{\lambda_j}{\operatorname{argmin}}\, MSE_{\lambda_j}^{(TOT)},$$

for all $j = 1,\ldots,k$.

Figure 3.1.1 shows the estimated spot rate curve parameters for the U.S. Treasury data during years 2017 and 2018. All graphs show a comparison between parameter estimates using a fixed optimal λ^* parameter versus a daily optimal λ_t^* for each day t. Comparing both methods, there is minimal change in the level, slope and curvature of the estimated spot rate curve paramaters. In terms of λ, there is noticeable instability during the year 2018 compared to the constant optimal $\lambda^* = 3.7$. Finally, 85.97% of the daily estimated price MSE associated to the fixed optimal λ^* are smaller compared to a daily change in λ. In light of these results and for stability purposes, we chose to subsequently use a fixed λ parameter across all days in our dataset which can be selected according to our proposed methodology in Section 3.1.2.

Fig. 3.1.1: *Estimated spot rate curve parameters associated to U.S. Treasuries 2017-2018*



Above plots show a comparison between parameter estimates using a fixed optimal λ^* *versus a daily optimal* λ_t^* *for each day* t*.*

3.2 Nelson-Siegel Spread Curve

A credit spread is the yield differential between a corporate and Treasury bond of the same maturity. In other words, the amount of adding the spread to one point on the treasury yield curve represents the discount factor that will make a corporate security's cash flow equal to its current market price. This being said, corporate yields are a function of the risk-free rate defined as

$$YTM = r_f + S,$$

where r_f represents the risk-free yields and S represents the credit spread. Corporate yield fluctuations are commonly due to change in firm creditworthiness, demand for investment within particular markets and economic conditions such as interest rates, inflation, and economic growth[5]. The corporate yield fluctuations due to economic conditions indicate

[5]Retrieved from *www.investopedia.com*

that efforts to forecast corporate spot rates may be driven by economic factors that impact Treasury yields where the risk associated to the Treasury yield will affect all the bonds within the country. Given this, removing the risk-free rate from the firms risk in order to obtain the credit spread is necessary since it is a key metric to gauge the level of expense for corporate bonds.

Sections 3.1.1 and 3.1.2 focused on representing yields as a term structure of spot rates. Applying the steps described in these previous sections on Treasury data leads to the Treasury spot rate parameters for each day t denoted as $\boldsymbol{\theta}^{(gov)}_{\lambda^*,t} = \left(\beta^{(gov)}_{0,t}, \beta^{(gov)}_{1,t}, \beta^{(gov)}_{2,t}\right)$. Given these Treasury spot rate parameters, this section focuses on decomposing the credit spread into term structures of spot rates.

To begin, the Nelson-Siegel estimated corporate bond price is defined as

$$B(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t}; \boldsymbol{\theta}^{(gov)}, \boldsymbol{\theta}^{(spread)}) = \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} c^{(j)}_{i,t} \times P(t, \tau^{(j)}_{i,t}; \boldsymbol{\theta}^{(gov)}, \boldsymbol{\theta}^{(spread)}), \tag{8}$$

where

$$P(t, \tau^{(j)}_{i,t}; \boldsymbol{\theta}^{(gov)}, \boldsymbol{\theta}^{(spread)}) = exp\left\{-\left[y(t, \tau^{(j)}_{i,t}; \boldsymbol{\theta}^{(gov)}) + y(t, \tau^{(j)}_{i,t}; \boldsymbol{\theta}^{(spread)})\right] \times \tau^{(j)}_{i,t}\right\}.$$

Equation (18) represents the discount factor which is now a function of the two sets of paramaters $\boldsymbol{\theta}^{(gov)}$ and $\boldsymbol{\theta}^{(spread)}$. Since parameters $\boldsymbol{\theta}^{(gov)}_{\lambda^*,t}$, previously described, are already estimated for each day, the Nelson-Siegel spread parameters are the only values in (8) left unknown.

The steps involved in the linearization of equation (8), defined in Gauthier & Simonato (2012), are shown in Appendix A.2. These equations lead again to an approximate system of the form $\mathbf{Y}^{(corp)}_{\varphi_0} = \tilde{\mathbf{X}}_{\varphi_0,\lambda_0}\boldsymbol{\theta}^{(spread)}_{\varphi_0,\lambda_0} + \boldsymbol{\epsilon}$, where each component is defined in Appendix A.2 and $\boldsymbol{\epsilon} = \top(\epsilon_1, \ldots, \epsilon_N)$ is a vector of error terms. A described in Section 3.1.1, φ_0 is fixed to represent the average yields-to-maturity corresponding to a given day t and λ_0 represents a fixed arbitrary value. Finally, given the fixed values for φ_0 and λ_0, the least squares estimator for $\boldsymbol{\theta}^{(spread)}_{\varphi_0,\lambda_0}$ is

$$\hat{\boldsymbol{\theta}}^{(spread)}_{\varphi_0,\lambda_0} = \left(\tilde{\mathbf{X}}^T_{\varphi_0,\lambda_0}\tilde{\mathbf{X}}_{\varphi_0,\lambda_0}\right)^{-1}\tilde{\mathbf{X}}^T_{\varphi_0,\lambda_0}\mathbf{Y}^{(corp)}_{\varphi_0}. \tag{9}$$

As previously stated in Section 3.1.1, the estimated parameters in $\hat{\boldsymbol{\theta}}^{(spread)}_{\varphi_0,\lambda_0}$ are then optimized using an implementation of that of Nelder & Mead (1965) to obtain $\hat{\boldsymbol{\theta}}^{(spread)}$ which minimizes the sum of squared errors defined as

$$Q(\theta) = \sum_{i=1}^{N}\left[\left(B^{(corp)}(m_i, \boldsymbol{c}_i, \boldsymbol{\tau}_i; \boldsymbol{\theta}^{(gov)}, \boldsymbol{\theta}^{(spread)}) - B^{(obs)}_i\right) \times \frac{1}{D_i}\right]^2. \tag{10}$$

The next steps involve obtaining an optimal $\tilde{\lambda^*}$ parameter as described in Section 3.1.2. In order to do so, we apply (9) with $\lambda_0 = \lambda_j$ to obtain the estimated parameters in $\hat{\boldsymbol{\theta}}^{(spread)}_{\varphi_0,\lambda_j}$ where $j = 1, 2, \ldots, k$ and k represents the total number of fixed λ values within our grid. Finally, these estimated parameters are optimized to obtain $\hat{\boldsymbol{\theta}}^{(spread)} = \left(\hat{\beta}^{(spread)}_0, \hat{\beta}^{(spread)}_1, \hat{\beta}^{(spread)}_2, \lambda_j\right)$ which minimizes the sum of squared errors in (10). This is done for each day t using the same fixed λ_j in order to obtain the time-dependent parameter estimates denoted as $\hat{\boldsymbol{\theta}}^{(spread)}_t = \left(\hat{\beta}^{(spread)}_{0,t}, \hat{\beta}^{(spread)}_{1,t}, \hat{\beta}^{(spread)}_{2,t}, \lambda_j\right)$.

This process is continued by computing the estimated price MSE for all bonds $i = 1, \ldots, N$ associated to each day t as

$$MSE^{(spread)}_{\lambda_j,t} = \frac{1}{N}\left(\sum_{i=1}^{N}\left[\left(B^{(corp)}(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t}; \boldsymbol{\theta}^{(spread)}_{\lambda_j,t}, \boldsymbol{\theta}^{(gov)}_{\lambda^*,t}) - B^{(obs)}_{i,t}\right) \times \frac{1}{D_{i,t}}\right]^2\right).$$

where $D_{i,t}$ is as defined in (6). Finally, given there are T days in the sample dataset, the optimal parameter $\tilde{\lambda^*}$ is defined as

$$\tilde{\lambda^*} = \underset{\lambda_j}{\operatorname{argmin}}\left\{\frac{1}{T}\sum_{t=1}^{T} MSE^{(spread)}_{\lambda_j,t}\right\},$$

for all $j = 1, \ldots, k$.

3.3 Zero-Volatility Spread

The current section introduces an alternative approach for representing credit spreads. Contrarily to the previous approach in Section 3.2, which models daily spread curves applying to all securities of the same issuer, here we use spreads that are idiosyncratic to each security. The flat structure is being considered in order to avoid over-parametrization and provide unique solutions to spread determination. By modeling daily spread curves in the approach of Section 3.2, we are making the assumption that all cash flows of various securities on a given day are priced using the same set of discount factors which leads to over- or under-pricing securities, as seen in subsequent sections providing empirical tests. Various bonds with identical cash flows might be priced differently due for instance to different bond liquidity. The assumption that all cash flows of various securities on a given day are priced using the same set of discount factors would not hold anymore for the current Z-spread representation.

3.3.1 Methodology

The flat spread structure is referred to in practice as the zero-volatility spread (Z-spread) which measures the spread that an investor will receive over the entire Treasury spot rate curve of a given day t. It is the Z-spread that makes the price of a security equal to the present value of its cashflows when added to the yield at each point on the spot rate Treasury curve where cashflow is received (Yen et al., 2007). In other words, each cashflow is discounted at the appropriate Treasury spot rate plus the Z-spread. To summarize, in order to calculate the Z-spread, an investor must take the Treasury spot rate at each relevant maturity, add the Z-spread to this rate, and then use this combined rate to compute the price of the corporate bond.

Since the correct Z-spread is the one that makes the present value of cashflows equal to the price of the bond, we may represent the price of the i^{th} bond for a given day t as,

$$B^{(corp)}\left(c_{i,t}, \boldsymbol{\tau}_{i,t}; \boldsymbol{\theta}^{(gov)}_{\lambda^*,t}, S^{(i)}_t\right) = \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} c^{(j)}_{i,t} \times exp\left\{-\left(y(t, \tau^{(j)}_{i,t}; \boldsymbol{\theta}^{(gov)}_{\lambda^*,t}) + S^{(i)}_t\right)\tau^{(j)}_{i,t}\right\}. \quad (11)$$

The set of parameters $\boldsymbol{\theta}_{\lambda^*,t}^{(gov)}$ are provided by the methodology from Section 3.2 and $S_t^{(i)}$ represents the Z-spread for the i^{th} bond on day t. In general, this method allows for obtaining a set of Z-spreads denoted as $S_t^{(i)}$ for $i = 1, ..., N$ bonds and $t = 1, ..., M$ days.

Z-spreads may then be computed for all $i = 1, ..., N$ and $t = 1, ..., M$ by using a numerical solver on

$$B^{(corp)}\left(c_{i,t}, \boldsymbol{\tau}_{i,t}; \boldsymbol{\theta}_{\lambda^*,t}^{(gov)}, S_t^{(i)}\right) = B_{i,t}^{(obs)}$$

which is straightforward due to bond prices being a monotonic decreasing function of the spread.

4. Price Forecasting

This chapter describes the methodology to construct forecasts of bond prices based on the approaches summarized in Chapter 3. The first section of this chapter outlines all five forecasting models considered in this research. The forecasting models are described in a general format by using an arbitrary time series. Finally, the forecasting models are then applied to three different representation methods for yields. The first entails forecasting the treasury and spread curves. This approach relies on the representation of both the risk-free rate and the spread as a term structure seen in Sections 3.1 and 3.2. The second approach entails forecasting the Z-spreads, as seen in Section 3.3, followed by forecasting the corresponding treasury curves as in the first approach. The last being a more naive approach which involves no curve fitting and depends upon directly forecasting observed corporate yields-to-maturity.

4.1 Forecasting Models

This section outlines the five forecasting methods used in this research in order to predict future bonds prices based on their price history. All methods are described in a general format by using an arbitrary time series process. We start by describing a more naive benchmark method with the simple assumption that the best forecast corresponding to the quantity of interest in subsequent days is its current value. Table 4.1 contains the descriptive statistics associated to the daily change in the estimated Nelson-Siegel parameters associated to the UST and Goldman Sachs data respectively for years 2017 and 2018. As we could see, according to the standard deviation of all parameters, there is

very minimal change between days. In light of this, the naive benchmark method might be thought to be efficient. Next we consider two different linear regression methods to model the relationship between successive values of the considered process or its time differences. Finally, we consider a first order autoregressive (AR(1)). Given our interest in quantifying potential momentum in the market, AR(1) models are typical in such circumstances. Similarly to the linear methods, we implement an AR(1) model to the time difference of the process or alternatively the process itself.

Table 4.1: *Descriptive statistics and sample autocorrelations of the daily change in estimated U.S. treasury Nelson-Siegel and Goldman Sachs coefficients for 2017 and 2018.*

Coefficient	$\Delta\hat{\beta}_0$	$\Delta\hat{\beta}_1$	$\Delta\hat{\beta}_2$
	U.S. Treasury: Descriptive statistics		
Mean	0.03314	-0.01738	0.00033
Std. Dev.	0.00183	0.00853	0.00536
Minimum	0.02920	0-0.03219	-0.01673
Maximum	0.03697	-0.00581	0.01129
	U.S. Treasury: Sample Autocorrelations		
$\hat{\rho}(1)$	0.076(0.0895)	0.136(0.0024)	0.015(0.7376)
$\hat{\rho}(2)$	-0.030(0.5029)	-0.052(0.2452)	0.047(0.2937)
$\hat{\rho}(3)$	-0.097(0.0302)	-0.087(0.0520)	0.021(0.6391)
$\hat{\rho}(4)$	-0.005(0.9108)	0.007(0.8760)	0.006(0.8934)
$\hat{\rho}(5)$	-0.075(0.0939)	-0.080(0.0739)	-0.036(0.4214)
	Goldman Sachs: Descriptive statistics		
Mean	0.02126	-0.00549	-0.02534
Std. Dev.	0.00088	0.00558	0.01249
Minimum	0.01942	-0.02583	-0.04179
Maximum	0.02349	0.00241	0.01011
	Goldman Sachs: Sample Autocorrelations		
$\hat{\rho}(1)$	-0.072(0.1078)	-0.047(0.2937)	-0.030(0.5029)
$\hat{\rho}(2)$	-0.053(0.2364)	-0.016(0.7211)	-0.056 (0.2109)
$\hat{\rho}(3)$	0.060(0.1802)	0.014(0.7534)	0.008(0.8579)
$\hat{\rho}(4)$	0.009(0.8407)	0.077(0.0854)	0.106(0.0179)
$\hat{\rho}(5)$	-0.045(0.3149)	0.006(0.8934)	-0.006(0.8934)

This table contains the mean, standard deviation, minimum value, maximum value and the sample autocorrelations $\hat{\rho}(j)$ of lag j associated to the daily changes in the Nelson-Siegel estimated parameters for the U.S. Treasury and Goldman Sachs data respectively for the years 2017 and 2018. In addition, the p-values associated to the sample autocorrelations are in brackets where they are computed given the assumption that $\Delta\hat{\beta}_i$ for $i = 1, 2$ and 3 are i.i.d. and Gaussian.

To begin, we define $\boldsymbol{X} = \{X\}_{t=1}^{T}$ to be a arbitrary time series of T discrete observations $X_1, X_2, \ldots, X_T$ with mean μ_X. In the forecasting models described below, the time series $\boldsymbol{X}$ will be substituted by the various quantities of interest among the different spot rate curve representation models.

The first method considered is the naive benchmark method and as previously mentioned, this method assumes that the best forecast corresponding to the quantity of interest in subsequent days is its current value. In general, the predicted observation as of time t is defined as

$$\hat{X}_{t,h} = X_t,$$

where $h = 1, \ldots, 5$ is the forecasting day horizon and for $t = 1, \ldots, T$.

Secondly, the first linear approach considered consists of regressing X_{t+h} on X_t where $t = 1, \ldots, T - h$. Separate regressions are performed for the various prediction horizons $h = 1, \ldots, 5$. The predicted observation at time $t + h$ is then defined as

$$\hat{X}_{t,h} = \hat{\gamma}_h + \hat{\alpha}_h X_t, \tag{12}$$

where $\hat{\gamma}_h$ and $\hat{\alpha}_h$ represent estimated ordinary least squared (OLS) linear regression coefficients with respect to the day horizon h.

The third forecasting model considered also involves a linear approach. This model focuses on regressing the time difference of the process $\boldsymbol{X}$. In general, this is defined by regressing $X_{t+h} - X_t$ on $X_t - X_{t-1}$ for $t = 1, \ldots, T - h$ with the intercept set to 0. The predicted observation at time $t + h$ is then defined as

$$\hat{X}_{t,h} = X_t + \tilde{\alpha}_h \left(X_t - X_{t-1}\right) \tag{13}$$

where $\tilde{\alpha}_h$ represents the estimated OLS linear regression slope parameter which depends on the day horizon h.

The fourth forecasting method consists of implementing an AR(1) model to the process $\boldsymbol{X}$. Firstly, μ_X, defined as the expected value of x_t, is estimated by the same mean of all the observation from time $t = 1$ to T where its estimated value is denoted as $\hat{\mu}_{X_{1:T}}$. The fourth forecasting method is then defined by regressing $(X_t - \hat{\mu}_{X_{1:T}})$ on $(X_{t-1} - \hat{\mu}_{X_{1:T}})$ with the intercept set to 0 and for $t = 1, \ldots, T$. Therefore, the predicted observation at time $t + h$ is then defined as

$$\hat{X}_{t,h} = \hat{\alpha}^h \left(X_t - \hat{\mu}_{X_{1:T}}\right) + \hat{\mu}_{X_{1:T}}, \tag{14}$$

where $\hat{\alpha}$ represents the estimated OLS linear regression slope parameter.

Finally, the last forecasting method similarly consists of implementing an AR(1) model but this time it is implemented on the time difference of the process $\boldsymbol{X}$. The last forecasting method is then defined by regressing V_{t+1} on V_t, where $V_t = X_t - X_{t-1}$ for all $t = 1, \ldots, T - 1$. In general, the predicted observation at time $t + h$ is defined as

$$\hat{X}_{t,h} = X_t + \sum_{j=1}^{h} \hat{V}_{t,j}, \tag{15}$$

where

$$\hat{V}_{t,j} = \tilde{\alpha}^j V_t,$$

for $h = 1, \ldots, 5$ and where $\tilde{\alpha}$ represents the estimated OLS linear regression slope parameter with the intercept set to 0. In addition, when the day horizon $h = 1$, this forecasting method coincides with the second forecasting method.

4.2 Forecasting Approaches

This section focuses on applying the forecasting methods described in the previous Section 4.1 to three different approaches applied for the representation of interest rates and yields described in Chapter 3. As earlier described, the first forecasting approach depends on the representation of both the risk-free rate and the spread as a term structure. The second approach requires separately forecasting the Z-spreads. Finally, the last forecasting approach involves no curve fitting and entails directly forecasting the observed yields-to-maturity.

To begin, the Nelson-Siegel curve bond price prediction demonstrates our first methodology to construct forecasts of bond prices which requires forecasting the treasury and spread curve parameters. The first step entails replacing the process $\boldsymbol{X}$ in equations (12) to (15) with each estimated parameter in $\hat{\boldsymbol{\theta}}^{(gov)}_{\lambda^*,t} = \left(\hat{\beta}^{(gov)}_{0,t}, \hat{\beta}^{(gov)}_{1,t}, \hat{\beta}^{(gov)}_{2,t}\right)$ and in $\hat{\boldsymbol{\theta}}^{(spread)}_{\lambda^*,t} = \left(\hat{\beta}^{(spread)}_{0,t}, \hat{\beta}^{(spread)}_{1,t}, \hat{\beta}^{(spread)}_{2,t}\right)$ for each day t. This is done in order to obtain the predicted government and spread parameters on day $t+h$ denoted as $\tilde{\boldsymbol{\theta}}^{(gov)}_{\lambda^*,t,h,k}$ and $\tilde{\boldsymbol{\theta}}^{(gov)}_{\lambda^*,t,h,k}$ respectively for each day horizon $h = 1, \ldots, 5$ and where $k = 1, \ldots, 5$ represents the five forecasting methods described in Section 4.1. Finally, the predicted bond price on day $t+h$ corresponding to the k^{th} forecasting method of the i^{th} bond as of time t is given by

$$\hat{B}^{(corp)}_{t+h,k}(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t+h}; \hat{\boldsymbol{\theta}}^{(gov)}_{\lambda^*,t,h,k}, \hat{\boldsymbol{\theta}}^{(spread)}_{\lambda^*,t,h,k}) = \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} c^{(j)}_{i,t} exp\left\{-\left[y_k\left(\tau^{(j)}_{i,t+h}; \hat{\boldsymbol{\theta}}^{(gov)}_{\lambda^*,t,h,k}\right) + y_k\left(\tau^{(j)}_{i,t+h}; \hat{\boldsymbol{\theta}}^{(spread)}_{\lambda^*,t,h,k}\right)\right] \times \tau^{(j)}_{i,t+h}\right\}.$$

Next, the second approach consists of constructing forecasts of bond prices which requires forecasting estimated Z-spread contrarily to the previous approach where spread curves were considered. We first replace the process $\boldsymbol{X}$ in equations (12) to (15) with each estimated Z-spread $S^{(i)}_t$ for each day t. This is done in order to obtain the i^{th} predicted Z-spread as of time t, denoted as $\hat{S}^{(i)}_{t,h,k}$ where h and k are as previously defined. Finally the predicted bond price on day $t+h$ corresponding to the k^{th} forecasting method of the

i^{th} bond as of time t is given by,

$$\tilde{B}^{(corp)}_{t+h,k}(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t+h}; \hat{\boldsymbol{\theta}}^{(gov)}_{\lambda^*,t,h,k}, \hat{S}^{(i)}_{t,h,k}) = \sum_{j=1}^{\#\boldsymbol{\tau}_{i,t+h}} c^{(j)}_{i,t} exp\left\{-\left[y_1\left(\tau^{(j)}_{i,t+h}; \hat{\boldsymbol{\theta}}^{(gov)}_{\lambda^*,t,h,k}\right) + \hat{S}^{(i)}_{t,h,k}\right] \times \tau^{(j)}_{i,t+h}\right\},$$

where the predicted parameter $\hat{\boldsymbol{\theta}}^{(gov)}_{\lambda^*,t,h,k}$ is computed using the first forecasting method previously described.

Finally, the last bond price prediction approach involves directly forecasting observed yields-to-maturity (YTM). Spot rate curve modeling is often used in financial theory including numerous theoretically sound research approaches such as the popular Nelson-Siegel model. Applying theoretical algorithms in practice might be too complex or complicated. In light of this, we introduce this very simple benchmark which does require any modeling steps in comparison to the more sophisticated methods from previous sections.

The steps to this naive approach begins by replacing the process $\boldsymbol{X}$ in equations (12) to (15) with each observed $YTM^{(i)}_t$ for each day t. We then obtain the i^{th} predicted YTM on day $t+h$ denoted as $\widehat{YTM}^{(i)}_{t,h,k}$ where h and k are as previously defined. Finally, the predicted bond price on day $t+h$ corresponding to the k^{th} forecasting method of the i^{th} bond as of time t is given by,

$$\begin{aligned}\bar{B}^{(corp)}_{t+h,k}(C_i, n_i, T_{i,t,h}; \widehat{YTM}^{(i)}_{h,k}) = &\frac{C_i}{n_i} \times \frac{\left[1 - \left(1 + \frac{\widehat{YTM}^{(i)}_{h,k}}{n_i}\right)^{-n_i T_{i,t,h}}\right]}{\widehat{YTM}^{(i)}_{h,k}/n_i} \\ &+ 100\left(1 + \frac{\widehat{YTM}^{(i)}_{h,k}}{n_i}\right)^{-n_i T_{i,t,h}}. \end{aligned} \tag{16}$$

In (16), C_i and n_i represent the coupon rate in percentage and frequency, respectively, associated to the i^{th} bond where the sum of discounted coupons are calculated using the formula for a flat annuity with a constant rate. In addition, $T_{i,t,h}$ represents the time-to-

maturity in years computed by the difference between the maturity date and day $t+h$ of the i^{th} bond.

5. Results

This chapter applies the theoretical framework given in Chapters 3 and 4 to the data described in Chapter 2. We take both U.S. Treasury and Goldman Sachs data, one after another, and evaluate both fitting and forecasting performance of the approaches previously described. We will first begin with a short description of the metrics used in order to measure the precision of our bond price forecasts. The following two sections applies the Nelson-Siegel curve, Z-spread and naive bond price prediction results corresponding to the U.S. Treasury and Goldman Sachs data respectively. In addition to the results, we have included analysis pertaining to the company Verizon Communications where results are shown in Appendix B.

5.1 Precision Metrics

To begin, as a measure of success, we use two different statistical indicators. The first precision metric that is used is the Root Mean Square Error (RMSE). RMSE is applied instead of Mean Square Error (MSE) since the latter will produce results in dollar units instead of squared dollars. Seeing that our sample datasets include a total of 249 trading days in 2018, the predicted bond price RMSE is computed as

$$RMSE_h^{(j)} = \sqrt{\frac{1}{N_j \times 249} \sum_{t=1}^{249} \left\{ \sum_{i=1}^{N_j} \left(\hat{B}_{i,t,h}^{(j)}(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t,h}) - B_{i,t,h}^{(obs)} \right)^2 \right\}}, \tag{17}$$

In (17), $h = 1, \ldots, 5$ represents the predicted day horizon, $j = 1$ represents the data associated to the U.S. Treasury and $j = 2$ represents the data obtained by Goldman Sachs. In

addition, N_j represents the number of bonds traded daily corresponding to the j^{th} dataset.

The second precision metric used is the Mean Absolute Error (MAE). Although optimization is based on the MSE objective function, MAE is still a relevant precision metric given that it is a more intuitive interpretation of price errors since it does not rely on dollars squared. Finally, using the same methodology as above, the MAE is computed as

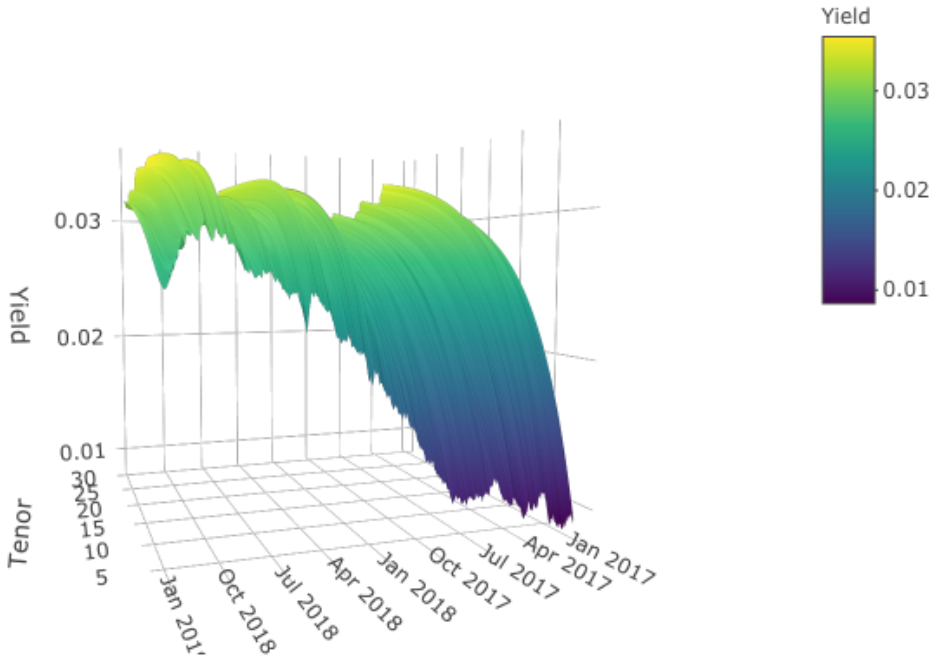
$$MAE_h^{(j)} = \frac{1}{N_j \times 249} \sum_{t=1}^{249} \left\{ \sum_{i=1}^{N_j} \left| \hat{B}_{i,t,h}^{(j)}(\boldsymbol{c}_{i,t}, \boldsymbol{\tau}_{i,t,h}) - B_{i,t,h}^{(obs)} \right| \right\}.$$

5.2 U.S. Treasury Bond Price Prediction Results

This section summarizes the bond price prediction results obtained by applying the methodology described in Section 4.1 and 4.3 to the U.S. Treasury data detailed in Chapter 2. We begin by summarizing the estimated spot rates obtained by following the steps detailed in Section 3.1. This is followed by the bond price prediction results derived by forecasting the estimated treasury yield curves and comparing this to the naive approach of directly forecasting the observed UST yields-to-maturity. Note that the Z-spread bond price prediction methodology described in Section 4.2 cannot be applied to UST bonds given this approach is only for corporate bonds including a spread component.

We begin by first constructing U.S. Treasury (UST) yield curves pertaining to each trading day in our sample. This step is done before computing any forecasts. The UST yield curves are constructed by applying the methodologies described in Section 3.1 to obtain the daily estimated parameters. The estimated parameters are then input into (3) with respect to the constant optimal Nelson-Siegel parameter $\lambda^{*(UST)} = 3.7$. Figure 5.2.1 shows a visual representation of these estimated spot rates with a 3D plot. We see smoothness in the evolution of the estimated UST spot rate curves which is consistent with the assumption that Government bonds are usually viewed as low-risk investments. In addition, Table 5.1 contains the UST spot rate estimates in percentage corresponding to standard tenors of 1 to 30 years where the sample dates include the first day of every month between 2017-01-03 and 2018-12-31.

Fig. 5.2.1: *3D plot of the U.S. Treasury estimated spot rate curves for the years 2017 and 2018*



Next, Table 5.2 contains the 1 to 5 day horizon UST bond price prediction MSE, with MAE in brackets, averaged for all dates associated to the trading days in 2018 from our sample dataset where values are in dollar units over a face value of $100. The rows correspond to the five forecasting methods previously discussed in Section 4.1. These forecasting methods include a naive benchmark, two linear methods and two AR(1) models which both model the relationship between time differences of a time series and on a time series itself. The table also breaks down into sections associated to two of the three representation of yield approaches described in Section 4.2. The first approach includes predicting the Nelson-Siegel estimated parameters and the second approach is predicting the observed yields-to-maturity. With respect to all forecasting day horizon, the smallest RMSE of 0.276 and associated MAE of 0.157 arise under the naive benchmark approach of predicting the observed yields-to-maturity. The results obtained when directly forecasting the observed yields-to-maturity show greater improvement with respect to predicting Nelson-Siegel paramaters under all forecasting day horizons. In conclusion, this implies that the most naive approach described in Section 4.2 performs best.

Table 5.1: *UST spot rate estimates for years 2017-2018*

Date	1-YR	5-YR	7-YR	10-YR	15-YR	20-YR	30YR
2017-01-03	0.8592%	2.0032%	2.3196%	2.6293%	2.9135%	3.0640%	3.2149%
2017-02-01	0.8574%	1.9999%	2.3202%	2.6367%	2.9308%	3.0884%	3.2477%
2017-03-01	1.0473%	2.0541%	2.3461%	2.6414%	2.9240%	3.0795%	3.2391%
2017-04-03	0.9844%	1.9177%	2.2031%	2.5016%	2.7991%	2.9682%	3.1455%
2017-05-01	1.0345%	1.8974%	2.1796%	2.4863%	2.8055%	2.9931%	3.1933%
2017-06-01	1.0820%	1.8190%	2.0737%	2.3589%	2.6647%	2.8483%	3.0467%
2017-07-03	1.1320%	1.9749%	2.2230%	2.4765%	2.7221%	2.8585%	2.9996%
2017-08-01	1.1022%	1.8536%	2.1031%	2.3767%	2.6639%	2.8338%	3.0158%
2017-09-01	1.1472%	1.7903%	2.0195%	2.2799%	2.5634%	2.7355%	2.9223%
2017-10-02	1.2698%	1.9788%	2.2055%	2.4488%	2.6987%	2.8440%	2.9983%
2017-11-01	1.4296%	2.0568%	2.2586%	2.4760%	2.7001%	2.8309%	2.9699%
2017-12-01	1.6415%	2.1467%	2.3045%	2.4718%	2.6410%	2.7383%	2.8410%
2018-01-02	1.7595%	2.2638%	2.4127%	2.5651%	2.7131%	2.7954%	2.8807%
2018-02-01	1.9709%	2.5961%	2.7444%	2.8722%	2.9682%	3.0084%	3.0419%
2018-03-01	2.0612%	2.6152%	2.7629%	2.9036%	3.0280%	3.0914%	3.1535%
2018-04-02	2.1367%	2.5790%	2.6973%	2.8103%	2.9106%	2.9620%	3.0124%
2018-05-01	2.3765%	2.8504%	2.9581%	3.0471%	3.1085%	3.1310%	3.1471%
2018-06-01	2.3417%	2.7838%	2.8858%	2.9713%	3.0322%	3.0557%	3.0737%
2018-07-02	2.4557%	2.7892%	2.8676%	2.9347%	2.9843%	3.0046%	3.0212%
2018-08-01	2.5729%	2.9135%	2.9955%	3.0674%	3.1228%	3.1469%	3.1677%
2018-09-04	2.5859%	2.8192%	2.8897%	2.9628%	3.0351%	3.0759%	3.1185%
2018-10-01	2.7340%	2.9976%	3.0719%	3.1457%	3.2147%	3.2518 %	3.2895%
2018-11-01	2.7825%	3.0107%	3.0989%	3.2028%	3.3199%	3.3926%	3.4725%
2018-12-03	2.8173%	2.8837%	2.9489%	3.0455%	3.1746%	3.2630%	3.3647%

This table contains a sample of the UST spot rates in percentage. The dates from the sample include the first day of every month from January 3^{rd}*, 2003 to December* 31^{st}*, 2018.*

5.3 Goldman Sachs Bond Price Prediction Results

This section summarizes the bond price prediction results obtained by applying the methodology described in Chapter 4 to the Goldman Sachs data. To begin, we estimate the Goldman Sachs daily spread yield curves by following the methodology in Section 3.2 along with the estimated Z-spreads as described in Section 3.3. This is followed by obtaining the bond price prediction results derived by forecasting the estimated U.S. Treasury and Goldman Sachs spread Nelson-Siegel estimated parameters with respect to the day horizon of interest for all dates within the year 2018. These results are then compared to the Z-spread approach and to the naive predictions approach which directly forecasts the observed Goldman Sachs corporate yields-to-maturity.

Table 5.2: *UST average 1-5 day horizon bond price prediction errors for all of 2018.*

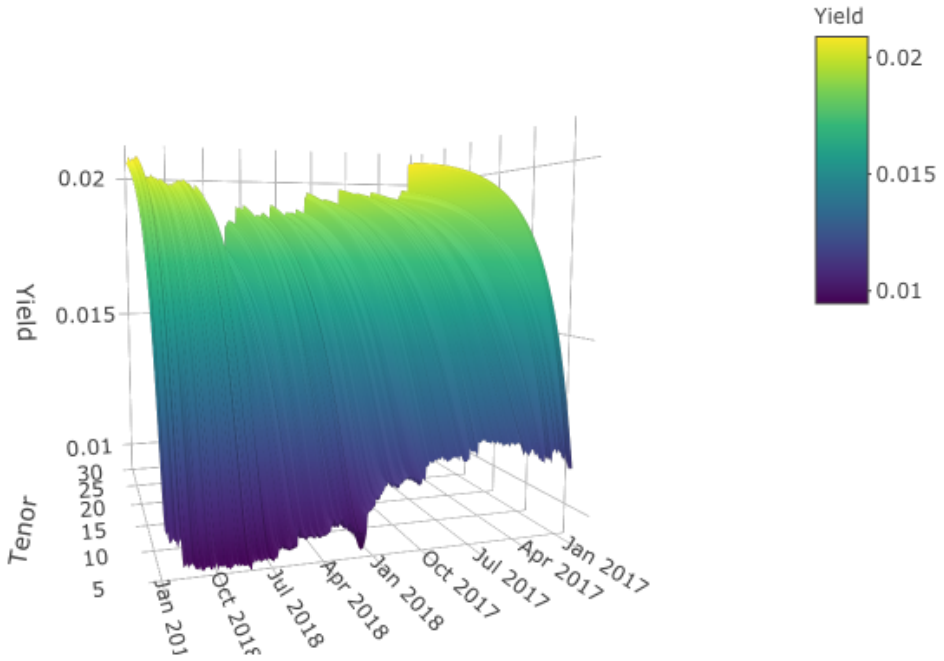
Forecast Method	1-DAY	2-DAY	3-DAY	4-DAY	5-DAY
			Nelson-Siegel Predicted Parameters		
Benchmark	0.300 (0.174)	0.404 (0.236)	0.486 (0.282)	0.559 (0.322)	0.628 (0.366)
$OLS^{(1)}$	0.299 (0.173)	0.399 (0.233)	0.478 (0.277)	0.547 (0.317)	0.612 (0.360)
$OLS^{(2)}$	0.301 (0.174)	0.404 (0.236)	0.485 (0.281)	0.559 (0.322)	0.627 (0.367)
$AR(1)^{(1)}$	0.300 (0.175)	0.401 (0.237)	0.482 (0.282)	0.555 (0.324)	0.623 (0.370)
$AR(1)^{(2)}$	0.301 (0.174)	0.403 (0.236)	0.487 (0.282)	0.560 (0.322)	0.627 (0.365)
			Predicted Yields-to-Maturity		
Benchmark	0.276 (0.157)	0.384 (0.224)	0.470 (0.270)	0.545 (0.312)	0.615 (0.357)
$OLS^{(1)}$	0.279 (0.159)	0.390 (0.227)	0.478 (0.276)	0.555 (0.319)	0.625 (0.364)
$OLS^{(2)}$	0.277 (0.158)	0.387 (0.225)	0.472 (0.271)	0.548 (0.314)	0.618 (0.359)
$AR(1)^{(1)}$	0.280 (0.160)	0.394 (0.232)	0.485 (0.282)	0.567 (0.328)	0.644 (0.378)
$AR(1)^{(2)}$	0.277 (0.158)	0.385 (0.225)	0.471 (0.271)	0.546 (0.313)	0.616 (0.358)

This table contains the 1 to 5 day horizon UST bond price prediction RMSE, with MAE in brackets. The rows correspond to the five forecasting methods previously discussed in Section 4.1. The table also breaks down into sections associated to the three forecasting approaches described in Section 4.2. In addition, all units are in dollar unit over a bond face value of \$100.
(1) Forecast method associated to the process itself.
(2) Forecast method associated to the time differences of the process.

As described in the previous paragraph, we first construct the Goldman Sachs (GS) corporate yield curves pertaining to each trading day in our sample. This step is done by applying the methodologies described in Section 3.2 along with incorporating the estimated UST spot rates. The estimated UST and Goldman Sachs spread parameters are then inputted into (3) with respect to the constant optimal Nelson-Siegel parameter $\lambda^{*(UST)} = 3.7$ and $\lambda^{*(GSspread)} = 2$. Figure 5.3.1 shows us a visual representation of the estimated Goldman Sachs spread spot rates with a 3D plot. Similarly to Figure 5.2.1, we see smoothness in the evolution for all 499 trading days between 2017-01-03 and 2018-12-31.

Once the Goldman Sachs spread yield curves pertaining to each trading day in our sample are constructed, we move onto the next methodology of estimating the associated Z-spreads as described in Section 3.3. The estimated Z-spreads are computed by inputting the estimated UST Nelson-Siegel paramaters into (11). Figure 5.3.2 contains two estimated Z-spread times series for the years 2017 and 2018. The first Z-spread time

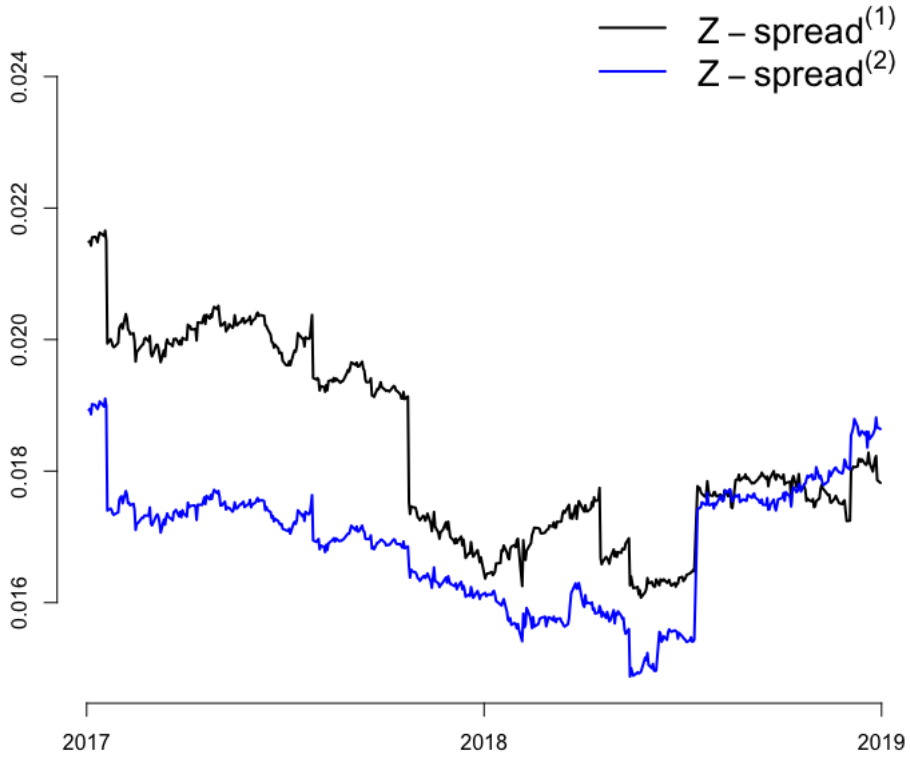
Fig. 5.3.1: *3D plot of the Goldman Sachs estimated spread spot rate curves for the years 2017 and 2018*



series is associated to a bond with a short maturity, i.e. of 2.8 years on January 3^{rd}, 2017, whereas the second corresponds to a bond with a long maturity of 28.8 years on January 3^{rd}, 2017. As a result, during the year 2017, the Z-spread time series associated to a shorter maturity time follows the same pattern in comparison but with higher Z-spread values. As time progresses, we see that both Z-spread time series merge closer to value while still keeping a similar pattern to eachother.

Table 5.3 reports the 1 to 5 day horizon Goldman Sachs bond price prediction MSE (and MAE in brackets) averaged for all dates associated to the trading days in 2018 from our sample dataset. Similarly to Table 5.2, the rows correspond to the same five forecasting methods previously discussed in Section 4.1. The table also breaks into sections associated to all three of the representation of yield approaches described in Section 4.2. The first approach includes predicting the Nelson-Siegel estimated parameters, the second approach predicts the estimated Z-spreads and the third predicts the observed yields-to-maturity. In addition, for all three methods, the same methodology is used for the risk-free rate

Fig. 5.3.2: *Plot of the Goldman Sachs estimated Z-spreads for two bonds during the years 2017 and 2018*



(1) Z-spread time series associated to the bond with a maturity time of 2.8 years on January 3^{rd}, 2017.
(2) Z-spread time series associated to the bond with a maturity time of 28.8 years on January 3^{rd}, 2017.

and the spread curves. In comparison to Table 5.2, forecasting both treasury and spread curves is also the least favorable approach. These results imply that the first methodology of using a single curve for all bonds of Goldman Sachs might be missing idiosyncraticities of bonds, which for instance have different liquidity characteristics. Given this, there is a much bigger difference between this approach and the predicted Z-spreads and predicted yields-to-maturity approach. In analyzing the last two approaches, the errors associated to predicting Z-spreads are very similar to those errors associated to predicting yields-to-maturity. Given the similarities between these two approaches, for all day horizons, the most favorable results fall under the linear forecasting approach of regressing the Z-spreads time series with itself. In addition, we have computed the difference between

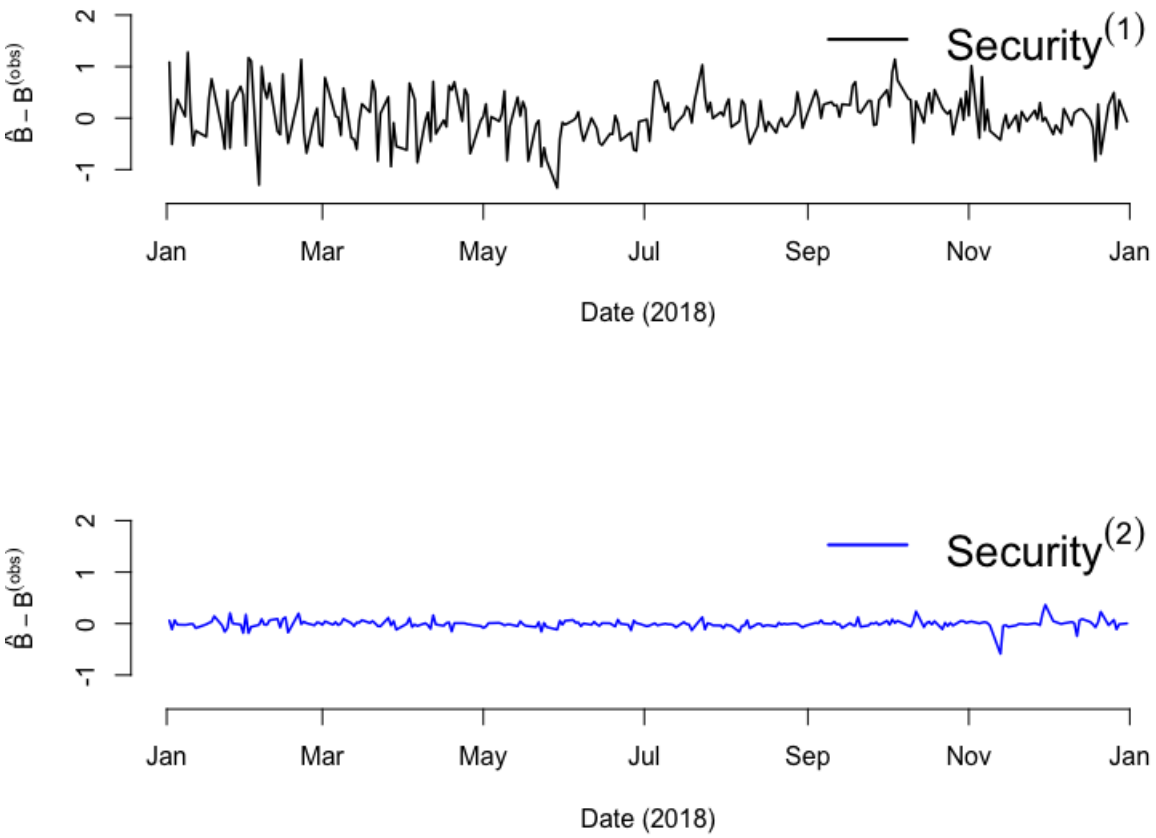
the one day horizon predicted prices of two securities for each day in 2018 associated to the linear forecasting approach of regressing the Z-spreads time series with itself and the corresponding observed prices. The time series of such prediction errors is provided in Figure 5.3.3. As we could see, there is an increase in the prediction errors for the security with a larger time to maturity.

Table 5.3: *Goldman Sachs average 1-5 day horizon bond price prediction errors for all of 2018.*

Forecast Method	1-DAY	2-DAY	3-DAY	4-DAY	5-DAY
			Nelson-Siegel Predicted Parameters		
Benchmark	1.627 (1.180)	1.648 (1.193)	1.666 (1.206)	1.687 (1.220)	1.708 (1.234)
$OLS^{(1)}$	1.627 (1.180)	1.647 (1.193)	1.666 (1.206)	1.688 (1.220)	1.711 (1.236)
$OLS^{(2)}$	1.627 (1.180)	1.648 (1.193)	1.666 (1.206)	1.687 (1.220)	1.708 (1.235)
$AR(1)^{(1)}$	1.627 (1.180)	1.647 (1.193)	1.668 (1.207)	1.691 (1.222)	1.715 (1.239)
$AR(1)^{(2)}$	1.627 (1.180)	1.648 (1.193)	1.666 (1.206)	1.687 (1.220)	1.708 (1.234)
			Predicted Z-Spreads		
Benchmark	0.308 (0.204)	0.421 (0.290)	0.505 (0.345)	0.584 (0.399)	0.652 (0.451)
$OLS^{(1)}$	0.306 (0.203)	0.418 (0.290)	0.498 (0.344)	0.574 (0.396)	0.639 (0.449)
$OLS^{(2)}$	0.308 (0.202)	0.420 (0.289)	0.504 (0.344)	0.583 (0.399)	0.652 (0.453)
$AR(1)^{(1)}$	0.308 (0.204)	0.420 (0.292)	0.504 (0.348)	0.584 (0.402)	0.654 (0.457)
$AR(1)^{(2)}$	0.308 (0.202)	0.420 (0.289)	0.504 (0.346)	0.584 (0.399)	0.652 (0.451)
			Predicted Yields-to-Maturity		
Benchmark	0.308 (0.204)	0.421 (0.290)	0.505 (0.345)	0.584 (0.398)	0.653 (0.452)
$OLS^{(1)}$	0.311 (0.207)	0.427 (0.296)	0.514 (0.352)	0.595 (0.409)	0.664 (0.465)
$OLS^{(2)}$	0.307 (0.204)	0.422 (0.290)	0.506 (0.346)	0.585 (0.401)	0.655 (0.455)
$AR(1)^{(1)}$	0.311 (0.208)	0.430 (0.299)	0.522 (0.359)	0.609 (0.421)	0.688 (0.483)
$AR(1)^{(2)}$	0.307 (0.204)	0.420 (0.289)	0.504 (0.345)	0.583 (0.399)	0.652 (0.451)

This table contains the 1 to 5 day horizon Goldman Sachs bond price prediction RMSE, with MAE in brackets. The rows correspond to the five forecasting methods previously discussed in Section 4.1. The table also breaks down into sections associated to the three forecasting approaches described in Section 4.2. In addition, all units are in dollar unit over a bond face value of \$100.
(1) Forecast method associated to the process itself.
(2) Forecast method associated to the time differences of the process.

Fig. 5.3.3: *One day horizon prediction error time series associated to the linear forecasting approach of regressing the Z-spreads time series with itself for two bonds during the year 2018.*



(1) Bond price forecasting error associated to the bond with a maturity time of 2.8 years on January 3^{rd}, 2017.
(2) Bond price forecasting error associated to the bond with a maturity time of 28.8 years on January 3^{rd}, 2017.

5.4 Verizon Communications Bond Price Prediction Results

This section includes a short summary on the results associated to the data retrieved from the company Verizon Communications. After following the data criteria described in Chapter 2, a total of 23 Verizon bonds were ultimately retained in the final data set which are traded on everyday of the sample. In comparison to the Goldman Sachs data,

there are much less Verizon bonds traded daily. Even in the context of scarce data, we further analyze if the results associated to the Verizon data point in the same direction as the results obtained from the larger data set of Goldman Sachs.

Table 6.1 in Appendix B reports the 1 to 5 day horizon Verizon Communications bond price predictions MSE and MAE in brackets averaged for all dates associated to the trading days in 2018 from our sample data set. Similarly to Tables 5.1 and 5.2, the rows correspond to the same five forecasting methods previously discussed in Section 4.1. In addition, the tables break down into the same sections as in Table 5.2 which includes the three forecasting approaches described in Section 4.2.

To begin, the results obtained from the Nelson-Siegel predicted paramaters approach is also the least favorable approach. In comparison to the results obtained from Goldman Sachs, we see that the results from Verizon follow a similar direction with the errors associated to the predicted Z-spreads and predicted yields-to-maturity being very similar. For all day horizons, the most favorable approach falls under the results obtained by the linear forecasting method which consists of regressing the Z-spreads with past Z-spreads depending on the day horizon of interest. Even though the Verizon Communications data set consists of less bonds traded daily, we can conclude the results point in the same direction as the results obtained from the larger data set of Goldman Sachs.

Table 5.4: *Verizon Communications average 1-5 day horizon bond price prediction errors for all of 2018.*

Forecast Method	1-DAY	2-DAY	3-DAY	4-DAY	5-DAY
			Nelson-Siegel Predicted Parameters		
Benchmark	1.633 (1.230)	1.683 (1.257)	1.727 (1.281)	1.774 (1.309)	1.818 (1.334)
$OLS^{(1)}$	1.657 (1.250)	1.709 (1.279)	1.761 (1.312)	1.819 (1.349)	1.872 (1.383)
$OLS^{(2)}$	1.635 (1.231)	1.683 (1.257)	1.726 (1.281)	1.772 (1.307)	1.818 (1.334)
$AR(1)^{(1)}$	1.661 (1.252)	1.739 (1.304)	1.813 (1.356)	1.886 (1.402)	1.957 (1.448)
$AR(1)^{(2)}$	1.635 (1.231)	1.683 (1.257)	1.728 (1.281)	1.775 (1.309)	1.818 (1.334)
			Predicted Z-Spreads		
Benchmark	0.589 (0.413)	0.781 (0.556)	0.912 (0.658)	1.031 (0.739)	1.129 (0.811)
$OLS^{(1)}$	0.584 (0.408)	0.767 (0.546)	0.888 (0.637)	0.996 (0.710)	1.082 (0.777)
$OLS^{(2)}$	0.590 (0.412)	0.778 (0.555)	0.907 (0.652)	1.026 (0.732)	1.121 (0.801)
$AR(1)^{(1)}$	0.585 (0.407)	0.770 (0.546)	0.896 (0.639)	1.010 (0.721)	1.102 (0.790)
$AR(1)^{(2)}$	0.590 (0.412)	0.778 (0.554)	0.910 (0.656)	1.029 (0.736)	1.123 (0.804)
			Predicted Yields-to-Maturity		
Benchmark	0.589 (0.413)	0.781 (0.556)	0.912 (0.658)	1.031 (0.739)	1.128 (0.811)
$OLS^{(1)}$	0.591 (0.417)	0.779 (0.557)	0.907 (0.655)	1.022 (0.735)	1.114 (0.809)
$OLS^{(2)}$	0.585 (0.411)	0.774 (0.552)	0.908 (0.653)	1.028 (0.737)	1.122 (0.803)
$AR(1)^{(1)}$	0.591 (0.417)	0.792 (0.567)	0.938 (0.677)	1.071 (0.772)	1.185 (0.864)
$AR(1)^{(2)}$	0.585 (0.411)	0.773 (0.551)	0.907 (0.652)	1.025 (0.734)	1.121 (0.803)

This table contains the 1 to 5 day horizon Verizon Communications bond price prediction RMSE, with MAE in brackets. The rows correspond to the five forecasting methods previously discussed in Section 4.1. The table also breaks down into sections associated to the three forecasting approaches described in Section 4.2. In addition, all units are in dollar unit over a bond face value of $100.
(1) Forecast method associated to the process itself.
(2) Forecast method associated to the time differences of the process.

6. Conclusion

This thesis aims for an overall development of bond price prediction models over a time horizon of one to five days. The purpose is to test several bond pricing methods including various forecasting approaches to then compare the performance of each method individually. Pricing bonds was done by comparing three distinct approaches for the representation of interest rates and yields. To summarize, the first approach consisted of forecasting the estimated Nelson-Siegel paramaters which includes a new insight in incorporating a grid search in order to obtain a fixed optimal λ^* parameter. The second approach entails forecasting estimated Z-spreads which incorporates a flat spread structure. The last approach involves directly forecasting observed yields-to-maturity. In addition to the three approaches of the representation of yields, five forecasting methods were considered and used within this research. These methods include a naive benchmark method and two different linear regression approaches to model the relationship between the time differences of a time series and between the values of the time series itself. Finally, similarly to the linear approaches, we implement an AR(1) model to the time difference and between the values of a time series.

From the results in Chapter 5, the approach of constructing a single curve for all bonds is the least favorable. These results are consistent with the UST government data and both Goldman Sachs and Verizon Communications corporate data sets. This implies a lack of bond liquidity features in pricing by using a single curve across all securities of the same issuer. We also see that, in comparing the last two approaches of predicting Z-spreads and yields-to-maturity, the results are very similar and much more superior compared to constructing yield curves. This being said, modeling the linear relationship through OLS

regression between past estimated Z-spreads performed best under both corporate issuers. All this considered, applying theoretical algorithms in practice can be complex or complicated. From the results provided in this thesis, using more simple albeit less theoretically sound approaches is more efficient than a complex model to predict the evolution of bond prices using historical data.

In conclusion, it is generally difficult to overperform very naive approaches in order to predict the evolution of liquid bond prices using only information contained in historical bond prices over a very short time horizon. Results show that the more naive method of predicting Z-spreads works best for corporate issuers. In addition, comparing the two approaches of predicting Z-spreads or yields-to-maturity directly, the results are very similar even though directly forecasting observed yields-to-maturity is not fully theoretically sound due to non-stationarities caused by the modification of cash flow maturities through time. In addition, for better precision in bond price predictions, additional information might be needed. For example, information on who is buying or selling the bonds, signals from the equity market or trade volumes to analyze clusters in bond liquidity could prove useful. Further research may show how this additional information may be beneficial in improving bond price prediction accuracy.

Bibliography

Ang, A., Piazzesi, M., & Wei, M. (2004). What Does the Yield Curve Tell us about GDP Growth? *National Bureau of Economic Research.*

Bauer, M. D., & Rudebusch, G. D. (2017). Interest Rates Under Falling Stars. *SSRN Electronic Journal.*

Bedock, N., & Stevanović, D. (2017). An empirical study of credit shock transmission in a small open economy. *Canadian Journal of Economics/Revue Canadienne D'économique*, 50(2), 541-570.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting.* Cham: Springer.

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

Enders, W. (2018). *Applied econometric time series.* New Delhi: Wiley.

Ericsson, J., & Reneby, J. (2005). Estimating Structural Bond Pricing Models. *The Journal of Business*, 78(2), 707-735.

Gauthier, G., & Simonato, J.-G. (2012). Linearized Nelson-Siegel and Svensson models for the estimation of spot interest rates. *European Journal of Operational Research*, 219(2), 442-451.

Gurkaynak, R. S., Sack, B. P., & Wright, J. H. (2006). The U.S. Treasury Yield Curve: 1961 to the Present. *SSRN Electronic Journal.*

Hardle, W. K., & Majer, P. K. (2012). Yield Curve Modeling and Forecasting Using Semiparametric Factor Dynamics. *The European Journal of Finance*, 1109-1129.

Hull, J., & White, A. (1990). Pricing Interest-Rate-Derivative Securities. *Review of Financial Studies*, 3(4), 573-592.

Ioannides, M. (2003). A comparison of yield curve estimation techniques using UK data. *Journal of Banking & Finance*, 27(1), 1-26.

Lewellen, J. (2002). Momentum and Autocorrelation in Stock Returns. *Review of Financial Studies*, 15(2), 533-564.

Longstaff, F. A., & Schwartz, E. S. (1992). Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model. *The Journal of Finance*, 47(4), 1259-1282.

Munk, C. (2015). *Fixed income modelling.* Oxford: Oxford University Press.

Nelder, J. A., & Mead, R. (1965). *A Simplex Method for Function Minimization.* The Computer Journal, 7(4), 308-313.

Nelson, C. R., & Siegel, A. F. (1987). *Parsimonious Modeling of Yield Curves.* The Journal of Business, 60(4), 473.

Pooter, M. D. (2007). Examining the Nelson-Siegel Class of Term Structure Models: In-Sample Fit versus Out-of-Sample Forecasting Performance. *SSRN Electronic Journal.*

Stojanovic, D., & Vaughan, M. D. (1997, January 1). *Yielding clues about recessions: the yield curve as a forecasting tool.* Retrieved from https://ideas.repec.org/a/fip/fedlre/y1997ioctp10-11.html

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

Yen, J., Elliott, R. J., Jarrow, R. A., Fu, M., & Madan, D. B. (2007). *Advances in mathematical finance.* Boston: Birkhäuser.

APPENDIX A.1

Linearized Nelson-Siegel model for coupon bonds

To simplify notation in this appendix, the subscript t and i are dropped. To begin, we introduce a new paramater φ to obtain the following coupon bond pricing equation:

$$\begin{aligned} B(\boldsymbol{c}, \boldsymbol{\tau}; \boldsymbol{\theta}) &= \sum_{j=1}^{m} c^{(j)} exp\left\{-y\left(\tau^{(j)}; \boldsymbol{\theta}\right) \times \tau^{(j)}\right\}, \\ &= \sum_{j=1}^{m} c^{(j)} exp\left\{-\varphi\tau^{(j)}\right\} exp\left\{\left[\varphi - y\left(\tau^{(j)}; \boldsymbol{\theta}\right)\right] \times \tau^{(j)}\right\}. \end{aligned}$$

The equation above is then linearized in the second exponential to obtain:

$$\begin{aligned} B(\boldsymbol{c}, \boldsymbol{\tau}; \boldsymbol{\theta}) &\approx \sum_{j=1}^{m} c^{(j)} exp\left\{-\varphi\tau^{(j)}\right\}\left(1 + \varphi\tau^{(j)} - y\left(\tau^{(j)}; \boldsymbol{\theta}\right) \times \tau^{(j)}\right), \\ &= \sum_{j=1}^{m} c^{(j)} \psi_0\left(\tau^{(j)}\right) - \beta_0 \sum_{j=1}^{m} c^{(j)} \psi_1\left(\tau^{(j)}\right) - \beta_1 \sum_{j=1}^{m} c^{(j)} \psi_2\left(\tau^{(j)}\right) - \beta_2 \sum_{j=1}^{m} c^{(j)} \psi_3\left(\tau^{(j)}\right) \end{aligned}$$

where

$$\psi_0\left(\tau^{(j)}\right) = exp\left\{-\varphi\tau^{(j)}\right\}\left(1 + \varphi\tau^{(j)}\right), \qquad \psi_1\left(\tau^{(j)}\right) = exp\left\{-\varphi\tau^{(j)}\right\}\tau^{(j)},$$

$$\psi_2\left(\tau^{(j)}\right) = exp\left\{-\varphi\tau^{(j)}\right\}\tau^{(j)}\phi_1\left(\tau^{(j)};\lambda\right) \quad and \quad \psi_3\left(\tau^{(j)}\right) = exp\left\{-\varphi\tau^{(j)}\right\}\tau^{(j)}\phi_2\left(\tau^{(j)};\lambda\right).$$

We could then approximate the bond price with

$$\hat{B}(\boldsymbol{c},\boldsymbol{\tau}) = \sum_{j=1}^{m} c^{(j)}\psi_0\left(\tau^{(j)}\right) - \beta_0\sum_{j=1}^{m} c^{(j)}\psi_1\left(\tau^{(j)}\right) - \beta_1\sum_{j=1}^{m} c^{(j)}\psi_2\left(\tau^{(j)}\right) - \beta_2\sum_{j=1}^{m} c^{(j)}\psi_3\left(\tau^{(j)}\right) + \varepsilon.$$

By setting the parameters φ and λ to some fixed values φ_0 and λ_0, the above equation is linear in the other three paramaters β_0, β_1 and β_2. Assuming that N bond prices are observed during a given day, the bond prices can be rewritten as a linear system with the $N \times 1$ vectors

$$\boldsymbol{X} = \begin{pmatrix} \frac{1}{D_1}\sum_{j=1}^{m_1} c_1^{(j)}\psi_0\left(\tau_1^{(j)}\right) & \frac{1}{D_1}\sum_{j=1}^{m_1} c_1^{(j)}\psi_1\left(\tau_1^{(j)}\right) & \frac{1}{D_1}\sum_{j=1}^{m_1} c_1^{(j)}\psi_2\left(\tau_1^{(j)}\right) \\ \vdots & \vdots & \vdots \\ \frac{1}{D_N}\sum_{j=1}^{m_N} c_N^{(j)}\psi_0\left(\tau_N^{(j)}\right) & \frac{1}{D_N}\sum_{j=1}^{m_N} c_N^{(j)}\psi_1\left(\tau_N^{(j)}\right) & \frac{1}{D_N}\sum_{j=1}^{m_N} c_N^{(j)}\psi_2\left(\tau_N^{(j)}\right) \end{pmatrix},$$

and

$$\boldsymbol{Y}_{\varphi_0} = \begin{pmatrix} \frac{1}{D_1}\left[\sum_{j=1}^{m_1} c_1^{(j)}\psi_0\left(\tau_1^{(j)}\right) - B_1^{(obs)}\left(\boldsymbol{c}_1,\boldsymbol{\tau}_1\right)\right] \\ \vdots \\ \frac{1}{D_N}\left[\sum_{j=1}^{m_n} c_N^{(j)}\psi_0\left(\tau_N^{(j)}\right) - B_N^{(obs)}\left(\boldsymbol{c}_N,\boldsymbol{\tau}_N\right)\right] \end{pmatrix},$$

where X_k represents the k^{th} column and each elements of the system of equations are now weighted with the inverse of the modified duration.

APPENDIX A.2

Linearized Nelson-Siegel spread model for coupon spread bonds

To simplify notation in this appendix, the subscript t and i are dropped. To begin, we introduce a new parameter φ to obtain the following coupon bond pricing equation:

$$B^{(corp)}(\boldsymbol{c}, \boldsymbol{\tau}; \boldsymbol{\theta}^{(gov)}, \boldsymbol{\theta}^{(spread)}) = \sum_{j=1}^{m} c^{(j)} exp\left\{-\left[y\left(\tau^{(j)}; \boldsymbol{\theta}^{(gov)}\right) + y\left(\tau^{(j)}; \boldsymbol{\theta}^{(spread)}\right)\right] \times \tau^{(j)}\right\},$$

Since $\boldsymbol{\theta}^{(gov)}$ is a set of parameter that are fixed due to being previously estimated, this implies that $y\left(\tau^{(j)}; \boldsymbol{\theta}^{(gov)}\right)$ is also known by equation (1) for all $j = 1, \ldots, m$. Therefore, let $K^{(j)} := c^{(j)} exp\left\{-y^{(gov)}\left(\tau^{(j)}; \boldsymbol{\theta}^{(gov)}\right) \times \tau^{(j)}\right\}$ and we have:

$$B^{(corp)}(\boldsymbol{K}, \boldsymbol{\tau}; \boldsymbol{\theta}^{(gov)}, \boldsymbol{\theta}^{(spread)}) = \sum_{j=1}^{m} K^{(j)} exp\left\{-\varphi\tau^{(j)}\right\} exp\left\{\left[\varphi - y\left(\tau^{(j)}; \boldsymbol{\theta}^{(spread)}\right)\right] \times \tau^{(j)}\right\}.$$

The equation above is then linearized in the second exponential to obtain:

$$B^{(corp)}(\boldsymbol{K},\boldsymbol{\tau};\boldsymbol{\theta}^{(spread)}) \approx \sum_{j=1}^{m} K^{(j)} exp\left\{-\varphi\tau^{(j)}\right\}\left(1+\varphi\tau^{(j)} - y\left(\tau^{(j)};\boldsymbol{\theta}^{(spread)}\right)\times\tau^{(j)}\right),$$

$$= \sum_{j=1}^{m} K^{(j)}\psi_0\left(\tau^{(j)}\right) - \beta_0^{(spread)}\sum_{j=1}^{m} K^{(j)}\psi_1\left(\tau^{(j)}\right) - \beta_1^{(spread)}\sum_{j=1}^{m} K^{(j)}\psi_2\left(\tau^{(j)}\right)$$

$$- \beta_2^{(spread)}\sum_{j=1}^{m} K^{(j)}\psi_3\left(\tau^{(j)}\right)$$

where

$$\psi_0\left(\tau^{(j)}\right) = exp\left\{-\varphi\tau^{(j)}\right\}\left(1+\varphi\tau^{(j)}\right), \qquad \psi_1\left(\tau^{(j)}\right) = exp\left\{-\varphi\tau^{(j)}\right\}\tau^{(j)},$$

$$\psi_2\left(\tau^{(j)}\right) = exp\left\{-\varphi\tau^{(j)}\right\}\tau^{(j)}\phi_1\left(\tau^{(j)};\lambda^{(spread)}\right) \ \ and \ \ \psi_3\left(\tau^{(j)}\right) = exp\left\{-\varphi\tau^{(j)}\right\}\tau^{(j)}\phi_2\left(\tau^{(j)};\lambda^{(spread)}\right).$$

We could then approximate the bond price with

$$\hat{B}^{(corp)}(\boldsymbol{K},\boldsymbol{\tau}) = \sum_{j=1}^{m} K^{(j)}\psi_0\left(\tau^{(j)}\right) - \beta_0^{(spread)}\sum_{j=1}^{m} K^{(j)}\psi_1\left(\tau^{(j)}\right) - \beta_1^{(spread)}\sum_{j=1}^{m} K^{(j)}\psi_2\left(\tau^{(j)}\right)$$

$$- \beta_2^{(spread)}\sum_{j=1}^{m} K^{(j)}\psi_3\left(\tau^{(j)}\right) + epsilon.$$

By setting the parameters φ and $\lambda^{(spread)}$ to some fixed values φ_0 and λ_0, the above equation is linear in the other three paramaters $\beta_0^{(spread)}$, $\beta_1^{(spread)}$ and $\beta_2^{(spread)}$. Assuming that N corporate bond prices are observed during a given day, the corporate bond prices can be rewritten as a linear system with the $N \times 1$ vectors

$$\tilde{\boldsymbol{X}} = \begin{pmatrix} \frac{1}{D_1}\sum_{j=1}^{m_1} K_1^{(j)}\psi_0\left(\tau_1^{(j)}\right) & \frac{1}{D_1}\sum_{j=1}^{m_1} K_1^{(j)}\psi_1\left(\tau_1^{(j)}\right) & \frac{1}{D_1}\sum_{j=1}^{m_1} K_1^{(j)}\psi_2\left(\tau_1^{(j)}\right) \\ \vdots & \vdots & \vdots \\ \frac{1}{D_N}\sum_{j=1}^{m_N} K_N^{(j)}\psi_0\left(\tau_N^{(j)}\right) & \frac{1}{D_N}\sum_{j=1}^{m_N} K_N^{(j)}\psi_1\left(\tau_N^{(j)}\right) & \frac{1}{D_N}\sum_{j=1}^{m_N} K_N^{(j)}\psi_2\left(\tau_N^{(j)}\right) \end{pmatrix},$$

and

$$\boldsymbol{Y}_{\varphi_0}^{(corp)} = \begin{pmatrix} \frac{1}{D_1}\left[\sum_{j=1}^{m_1} K_1^{(j)}\psi_0\left(\tau_1^{(j)}\right) - B_1^{(obs)}\left(\boldsymbol{c}_1, \boldsymbol{\tau}_1\right)\right] \\ \vdots \\ \frac{1}{D_N}\left[\sum_{j=1}^{m_n} K_N^{(j)}\psi_0\left(\tau_N^{(j)}\right) - B_N^{(obs)}\left(\boldsymbol{c}_N, \boldsymbol{\tau}_N\right)\right] \end{pmatrix},$$

where $\tilde{X}_k$ represents the k^{th} column and each elements of the system of equations are now weighted with the inverse of the modified duration.