A Primer on Yield Curve Construction

Introduction

A yield curve depicts the yields of fixed income securities against their term, or time to maturity. In this article we explain the role that the discount function and forward rates have in the creation of a yield curve and we compare different methodologies to fit the yield curve, namely, bootstrapping, spline-based methods, parametric and nonparametric models. Finally, we conduct an analysis specifically on the Nelson-Siegel and Wiseman models.

Discount Function and Yield Curve

The most important function used to understand the term structure of interest rates is the discount function δ, which represents the present value of \$1 repayable in m years. We expect this function to be continuously differentiable and monotonically decreasing, due to the time value of money. The function can be represented as follows:

$$\delta(i) = \frac{1}{(1+r_i)^i}.$$

This is the discrete version of the formula where i represents time in years and r_i is the risk-free rate at time i. By deriving the limit as time intervals go to 0, we obtain the formula for continuous discounting, which allows for a smoother and more precise curve:

$$\delta(0, m) = e^{-r(m)m},$$

where the first variable is the start time of the discounting and the second one is the time when the payment is received. From this continuous discount function we can derive, for instance, the spot 1Y discount factor $\delta(0,1)$, and the spot 2Y discount factor $\delta(0,2)$. Hence, we can compute the forward 1Y1Y discount factor, another central component to the construction of a yield curve:

$$\delta(1,2) = \frac{\delta(0,2)}{\delta(0,1)}.$$

This comes from the no-arbitrage condition which states that $\delta(0,2)$ must equal $\delta(0,1) \cdot \delta(1,2)$. We can clearly see that the forward rate f is positive, which is consistent with the discount function being decreasing. It can be expressed more clearly with this notation:

$$f(1,2) = \frac{r(m_2)m_2 - r(m_1)m_1}{m_2 - m_1}.$$

This formula can also be expressed in the general form $f(m, m + \epsilon)$, which is crucial to derive the instantaneous forward rate function:

$$f(m) = \lim_{\epsilon \to 0} f(m, m + \epsilon) = \frac{d}{dm} r(m)m.$$

With this relationship, we can observe that when the yield curve is in its "normal" upwards sloping state, the forward rates will lie above the yield curve, and that they will lie below the yield curve when it is inverted. By integrating on both sides, we can isolate r and obtain a formula for the yield curve:

$$r(m) = \frac{1}{m}\int_0^m f(s)ds = -\frac{1}{m}\ln\big(\delta(m)\big).$$

The Importance of the Yield Curve

To understand why these calculations are made, it is pivotal to comprehend the importance of the yield curve. Firstly, the yield curve serves as an indicator of the overall environment of an economy. When the yield curve is in an upwards sloping, "normal" state, it indicates the economy is expanding. On the other hand, an inverse yield curve has historically been a predecessor to a recession, as discussed in this article. We are currently witnessing a period where the yield curve is inverted, and so far, markets have responded accordingly, with 2022 being a disastrous year for equities world-wide.

Another possibly more important and relevant use of the yield curve is bond pricing across all possible maturities, not only the on the main liquidity points such as the 2Y, 5Y, or 10Y Treasury. This is relevant due to the fact that having the ability to accurately price a bond at any maturity lays the foundation for Relative Value strategies, one of the 2 main strategies implemented in rates trading. Relative Value strategies rely on the evaluation of whether a bond is trading too "rich" or too "cheap" at any given maturity, which is why constructing a yield curve is the foremost important tool in making these evaluations. When a bond is trading "rich" it means that it has a lower yield, or a higher price than bonds with similar credit and time to maturity, or term. The opposite is true for when a bond is trading "cheap".

Spline-based Methods

One of the earliest methods for fitting the term structure of interest rates was proposed by McCulloch (1971). It consists of a linear regression of the discount function δ which is assumed to be expressed as a linear combination of k differentiable functions $f_j(m)$ plus a constant term α_0, that is:

$$\delta(m) = \alpha_0 + \sum_{j=1}^k a_j f_j(m).$$

The present value of \$1 paid back instantly is \$1, that is $\delta(0) = 1$, which implies that $\alpha_0 = 1$ and $f_j(0) = 0$. Recall that the price p of a bond with constant and continuous coupon c principal F and maturity m_0 is:

$$p = F\delta(m_0) + c\int_0^{m_0}\delta(x)dx$$

combining the two equations above yields:

$$y = \sum_{j=1}^k a_j x_j$$

where $y = p - F - cm_0$ and $x_j = F f_j(m_0) + c \int_0^{m_0} f_j(x) dx$. Now, let us move to multiple dimensions and consider N observations of bonds. To take into account errors in the bond pricing formula due to factors such as transaction costs, callability and tax exemption, we introduce an error term in the formula for the bond i:

$$\overline{p_i} = F_i \delta(m_0) + c_i \cdot \int_0^{m_i} \delta(x) dx + \epsilon_i,$$

where $\overline{p_i}$ is the average between bid and ask, and ϵ_i is the error term. The standard error (S.E.) of ϵ_i is assumed to be of the form σv_i, where v_i is the bid-ask spread. Consequently, y for a bond i becomes:

$$y_i = \sum_{j=1}^{k} a_j x_{ij} + \epsilon_i.$$

The only unknowns are the a_j's and σ, which can both be estimated through a weighted least-squares regression of the vector $Y = (y_i)$ on the matrix $X = (x_{ij})$, obtaining $\widehat{a_1}, \cdots, \widehat{a_k}, \widehat{\sigma_k}$. We can now estimate the discount function:

$$\hat{\delta}(m) = \alpha_0 + \sum_{j=1}^{k} \widehat{a_j} f_j(m)$$

and finally the yield curve:

$$\hat{r}(m) = -\frac{1}{m} \ln \hat{\delta}(m).$$

As for the choice of k, if it is too low it might not fit the discount function well, while if it is too high it will overfit outliers, compromising the smoothness of the curve. The best choice is the one that minimizes the variance of the residuals. The functions $f_j(m)$ can be chosen arbitrarily, as long as they are continuously differentiable and $f_j(0) = 0$. Since maturities are not uniformly distributed, ideally $f_j(m)$ provides more resolution when maturities are clustered, which in the case of Treasuries occurs in the short term. The simplest approach would be:

$$f_j(m) = m^j, \quad j = 0, \cdots, k.$$

This choice of functions makes $\delta(m)$ a k-th degree polynomial, but it does not respect the criteria on resolution. For this reason, McCulloch (1971) and McCulloch (1975) adopt quadratic and cubic splines, which are piecewise polynomial functions that interpolate a set of nodes, i.e. yields at known maturities. In particular, cubic splines are twice continuously differentiable and exhibit great flexibility when fitting the yield curve.

Later on, new methods have been developed, such as Vasicek and Fong (1982), which relies on exponential splines to replicate the exponential nature of the discount function, as well as Adams and Van Deventer (1994), which maximizes the smoothness criterion $Z(m) = \int_0^m f''(s)^2 ds$ by fitting the forward rates curve with a fourth-order spline without the cubic term.

Parametric Models

Parametric models define a specific functional form to fit the term structure. One of the earliest models was presented by Cohen, Kramer and Waugh (1966), and it consists of a multilinear regression of the yield on the days to maturity and on the squared log of days to maturity. It was later revisited by Echols and Elliot (1976), that performs a linear regression of the form:

$$\ln\big(1 + r(m)\big) = a\frac{1}{m} + bm + c + \epsilon_m$$

Currently, one of the most used parametric models was proposed by Nelson and Siegel (1987). It specifies a functional form for the instantaneous forward rates, given by the solution of a second-order differential equation:

$$f(m) = \beta_0 + \beta_1 \exp\left(-\frac{m}{\tau}\right) + \beta_2 \frac{m}{\tau}\exp\left(-\frac{m}{\tau}\right).$$

This model is consistent with the three factors of the term structure proposed by Litterman and Scheinkman (1991): level, slope, and curvature. Moreover, its formula captures some key properties of the forward rates curve, namely being monotonic, humped, and S-shaped, by modeling the following parameters:

- Level β_0: long-run level of interest rates (strength of long-term component);

- Slope β_1: spread between short term and long-term interest rates (strength of short-term component);

- Curvature β_2: magnitude and direction of the hump (strength of medium-term component);

- τ: shape parameter, position (in time) of the hump.

The yield curve can then be derived:

$$r(m) = \beta_0 + (\beta_1 + \beta_2)\left[1 - \exp\left(-\frac{m}{\tau}\right)\right] \cdot \left(\frac{\tau}{m}\right) - \beta_2 \exp\left(-\frac{m}{\tau}\right),$$

where the estimators of β_0, β_1 and β_2 can be calculated through Ordinary Least Squares (OLS), while the shape τ is the optimal one among a grid of values (grid search): this procedure is preferred to the one which estimates all parameters simultaneously through nonlinear regression. Ridge regression is usually employed to address multicollinearity, as proposed by Annaert (2012).

One of the key points of the Nelson-Siegel model is its asymptotic behavior in the long run, which reflects the fact that real yields tend to converge to a constant level at long maturities. Moreover, the model describes the typical hump that can be found in a forward rates curve. On this regard, Svensson (1995) proposes an improvement of the Nelson-Siegel model with two humps, consequently including new parameters β_3 and τ_2 which represent the magnitude and the position of the second hump:

$$f(m) = \beta_0 + \beta_1 exp\left(-\frac{m}{\tau_1}\right) + \beta_2 \frac{m}{\tau_1} exp\left(-\frac{m}{\tau_1}\right) + \beta_3 \frac{m}{\tau_2} exp\left(-\frac{m}{\tau_2}\right)$$

Other parametric models include Wiseman (2012):

$$f(m) = y_0 + y_1 exp\left(-\frac{m}{z_1}\right) + \cdots + y_4 exp\left(-\frac{m}{z_5}\right)$$

where an OLS is performed after fixing values for z_i, for instance, 12 years, 5 years, 2 years, 6 months, and 1 month. Intuitively, as in Nelson Siegel, each y_i affects different regions of the curve, depending on the values of the z_i.

LMNT Kernel Estimation Method

A substantial improvement in performance compared to the McCulloch spline methods (1971, 1975) comes from the non-parametric kernel smoothing procedure developed by Linton, Mammen, Nielsen and Tanggaard (2000). To give some mathematical background needed to understand the model, we briefly discuss the definition of a Kernel Density Estimator (KDE). We define KDE for a bandwidth $h > 0$ as follows:

$$\hat{f}_n(x) := \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)$$

Where our function K is called a Kernel and is defined as Lebesgue integrable function that satisfies the condition $\int_{-\infty}^{\infty} K(u) du = 1$. Although there is a wide variety of usable Kernels, one of the preferred ones in practical uses is the Gaussian Kernel, which is also adopted in the LMNT model. We define the Gaussian Kernel as $K(u) = (2\pi)^{-\frac{1}{2}} e^{-\frac{u^2}{2}}$.

The LMNT model studies in depth two main estimators, the "local constant" and "local linear" methods, which locally approximate the discount function as a constant and a linear function of maturity. Once again, we define the yield curve for a maturity m as: $r(m) = -\frac{1}{m}\ln(\delta(m))$. This approach displays a few fundamental advantages, namely that $\delta(0) = 1$ that $\delta(m) > 0 \; \forall m$ and finally that it is closer to being log-linear than linear. According to the model, we can approximate the yield curve as a linear function $r(v) + (m - v)r'(v)$ where v is known maturity close to m. From this relationship and from the formula of the present value of a bond i, our estimated present value is as follows:

$$\widehat{PV^i} = \sum_{j=1}^{m^i} b_j^i e^{-\left(r(v_j) + \left(m_j^i - v_j\right)r'(v_j)\right)m_j^i}.$$

Our estimation model is based on minimizing the sum of squared pricing errors: particularly, we want to find a function $\hat{r}$ and its first derivative such that with respect to the actual yield curve the following criterion is minimized:

$$Q_N(r, r') = \sum_{i=1}^{N} \int \cdots \int \left(P^i - \sum_{j=1}^{m^i} b_j^i e^{-\left(r(v_j) + \left(m_j^i - v_j\right)r'(v_j)\right)m_j^i}\right)^2 \prod_{k=1}^{m^i} K_h\left(v_k - m_k^i\right) dv_k,$$

with K being our Kernel function, h as our bandwidth and $K_h(\cdot) = \frac{K\left(\frac{\cdot}{h}\right)}{h}$.

A simpler way to obtain a result is by solving the two first order conditions that derive from the minimization problem above for our actual yield curve and its first derivative, namely:

$$1. \sum_{i=1}^{N} \sum_{k=1}^{m^i} X_k^i\big(v; r(\cdot), r'(\cdot)\big) = 0$$

$$2. \sum_{i=1}^{N} \sum_{k=1}^{m^i} X_k^i\Big(v; r(\cdot), r'(\cdot)(v - \mathrm{m}_k^i)\Big) = 0,$$

with $X_k^i(v; r, r') = \Big(K_h\big(v_k - \mathrm{m}_k^i\big)b_k^i \mathrm{m}_k^i d_k^i(v)\Big)\Big(P^i - b_k^i d_k^i(v) - \sum_{j=1, j \neq k}^{m^i}\big(\int K_h\big(x - \mathrm{m}_j^i\big)b_j^i d_j^i(x)dx\big)\Big)$.

Given the following framework for our model, we need to pick a few elements that are necessary to carry out the estimations of $\hat{r}$ and its derivative: in this decision, we will report the methodology of Jeffrey et al. (2000). Firstly, as briefly mentioned, we will use the Gaussian Kernel for our density estimation; secondly, in order to pick the appropriate bandwidth, we observe that cash flows generated by bonds of longer maturities tend to be more sparse, thus calling for h to be a function of maturity when computing the present value of the cashflow n. The paper suggests an increasing linear relationship of the form $h(m_n^i) = a + bm_n^i$ with $h(0) = 2/12$ and $h(10) = 1$, allowing for greater flexibility in the front-end of the curve, where more information is available. Lastly, we pick a finite set of maturities $v_1, \cdots, v_k$ to be used to compute our estimated $\hat{r}$ and $\widehat{r'}$. Our set of maturities is calculated as $v_1 = 0$ and $v_i = v_{i-1} + \frac{1}{2}h(v_{i-1})$.

Theoretically, given our first order conditions we could compute any possible v_i, but for practical reasons we need to work with a finite set and interpolate between the calculated points. The interpolation procedure aims to minimize our original criterion obtaining an estimated pure zero-coupon bond:

$$\hat{Y}(m) = -\frac{1}{m}\ln\left(\frac{\sum_{k=1}^{k} K_h(v_i - m)e^{-\left(\hat{r}(v_i) + (m - v_i)\widehat{r'}(v_i)\right)}}{\sum_{i=1}^{k} K_h(v_i - m)}\right)$$

Fama-Bliss Bootstrapping

One of the more classical methods to estimate the yield curve, which displays a tremendous accuracy in the front-end of the curve (but it's fairly imprecise in estimating maturities longer than one year) is the bootstrapping procedure elaborated by Fama and Bliss (1987).

The process is of iterative nature, where starting from the relationship between the forward rate f and the discount function $\delta(m)$ we then assume that the forward curve is constant between successive observed bond maturities. That is, for a time-to-maturity interval $(m^{i-1}, m^i]$, $f(m) = F^i$ for the bond i.

Given this assumption, our discount function takes the following form: $\delta(m) = e^{-F^K(m - m^{K-1}) - \sum_{k=1}^{K-1} F^k(m^k - m^{k-1})}$. We can then extract the price of the bonds, starting from the shortest maturity one P^1, which is $\sum_{j=1}^{m^1} b_j^1 e^{-F^1 \times m_j^1}$; from this, $P^2 = \sum_{j=1}^{m^2} b_j^2 \delta(m_j^2)$ given F^1 and so on. The great advantage of the method is that all in-sample bonds will be perfectly priced and, generally speaking, the method displays a high degree of reliability for out-of-sample ones, even though this is limited to the short-term contracts (<1 year).
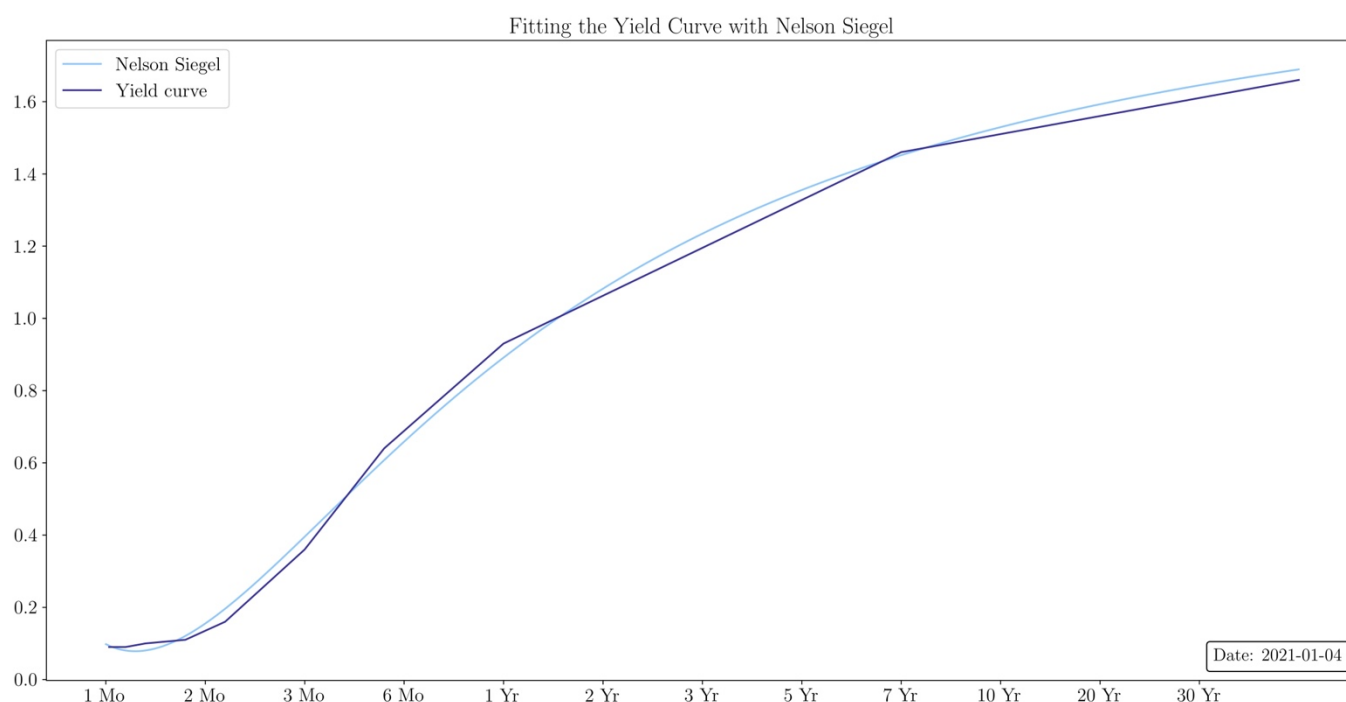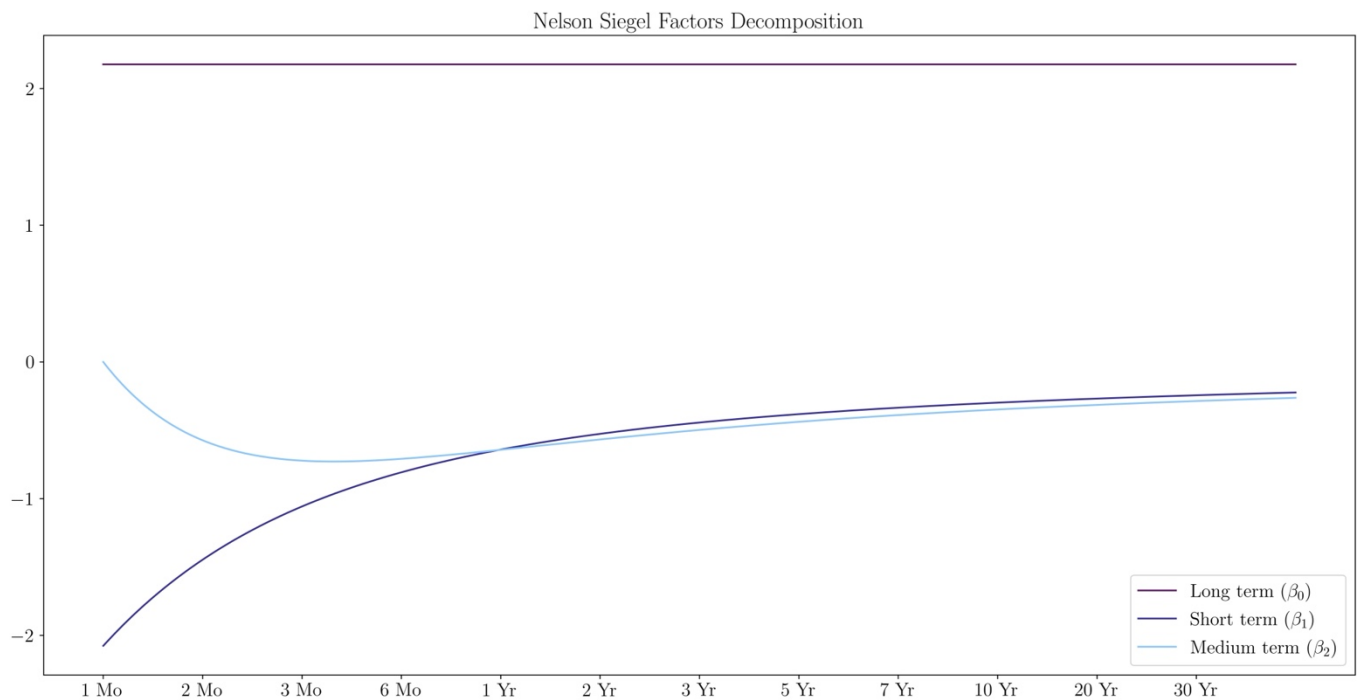
Analysis of Parametric Models

The dataset on which we conduct our analysis is published by the *U.S. Department of the Treasury*, and it comprises 2 years of daily treasury par yield curve rates, with maturities ranging from one month to 30 years. The par yield curve assumes that the price of the bonds that lie in the curve is equal to their face value, not their market value. On the par yield curve, the coupon rate will match yield to maturity, which is why the bond will trade at "par". This curve is used to determine the coupon rate that a newly emitted bond, with a given maturity will pay in order to sell at par today.

For each trading day we fit the curve with the Nelson Siegel model, by estimating the parameters β_0, β_1 and β_2 with OLS, and τ with a grid search that ranges from 1 to 2000 and minimizes Mean Absolute Percentage Error (MAPE).

Source: *U.S. Department of the Treasury, Bocconi Students Investment Club*

Source: *U.S. Department of the Treasury, Bocconi Students Investment Club*

In the pictures above we observe how the Nelson Siegel model fits the yield curve as well as the decomposition of the effects of its parameters. We can expand this analysis on the whole dataset by studying how the model's parameters have changed over time.

Source: *U.S. Department of the Treasury, Bocconi Students Investment Club*

The shape parameter has decreased steadily in the last two years, expect for a spike at the end of 2021. This means that the hump of the yield curve gets nearer to the short-term maturities. It is important to note that a high value of the shape parameter may signal the absence of a hump or the presence of two humps, a case which is covered only by the Svensson model.
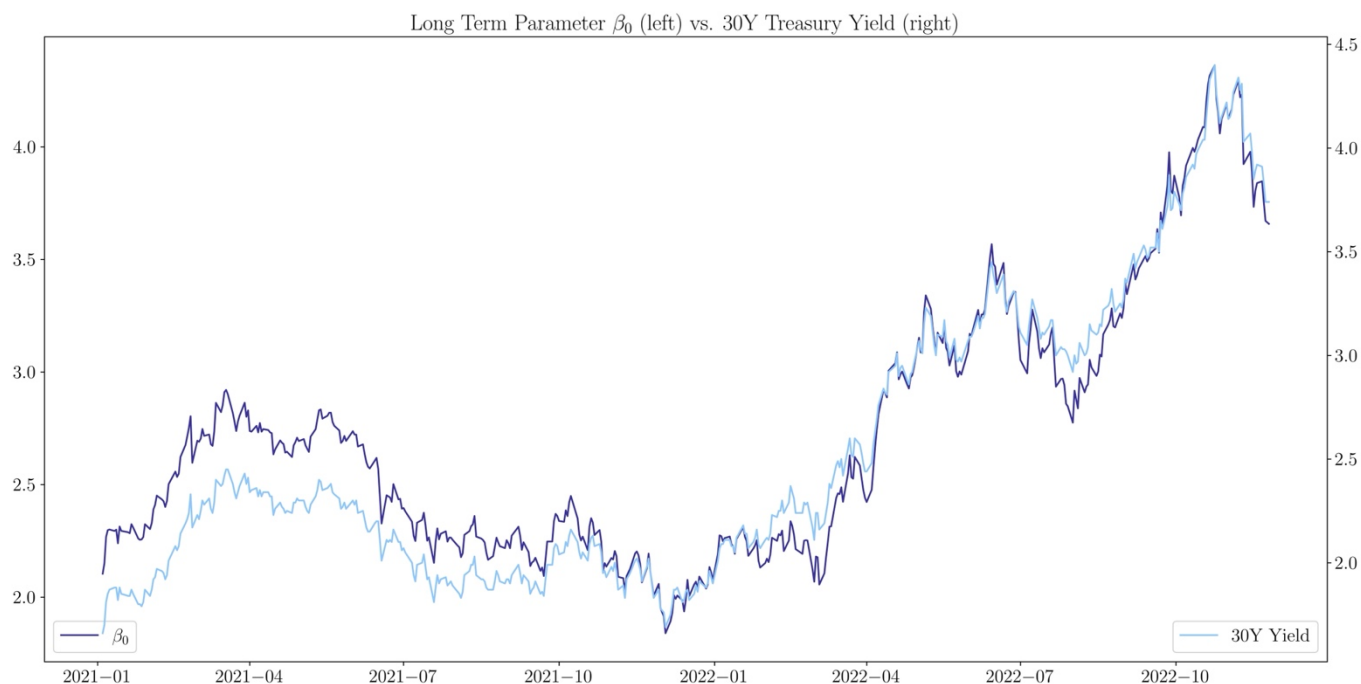
Source: *U.S. Department of the Treasury, Bocconi Students Investment Club*

The long-term parameter essentially represents the asymptotic behavior of the yield curve, therefore it has high correlation with the 30Y Treasury Yield, with a Pearson coefficient of 0.96 .

Source: *U.S. Department of the Treasury, Bocconi Students Investment Club*

The short-term parameter depicts the slope of the curve, when it is negative the curve is in its normal increasing shape, when positive the curve is inverted. It exhibits a strong correlation (Pearson coefficient 0.7) with the 2Y10Y Treasury spread, which breaks from January 2022 to May 2022.
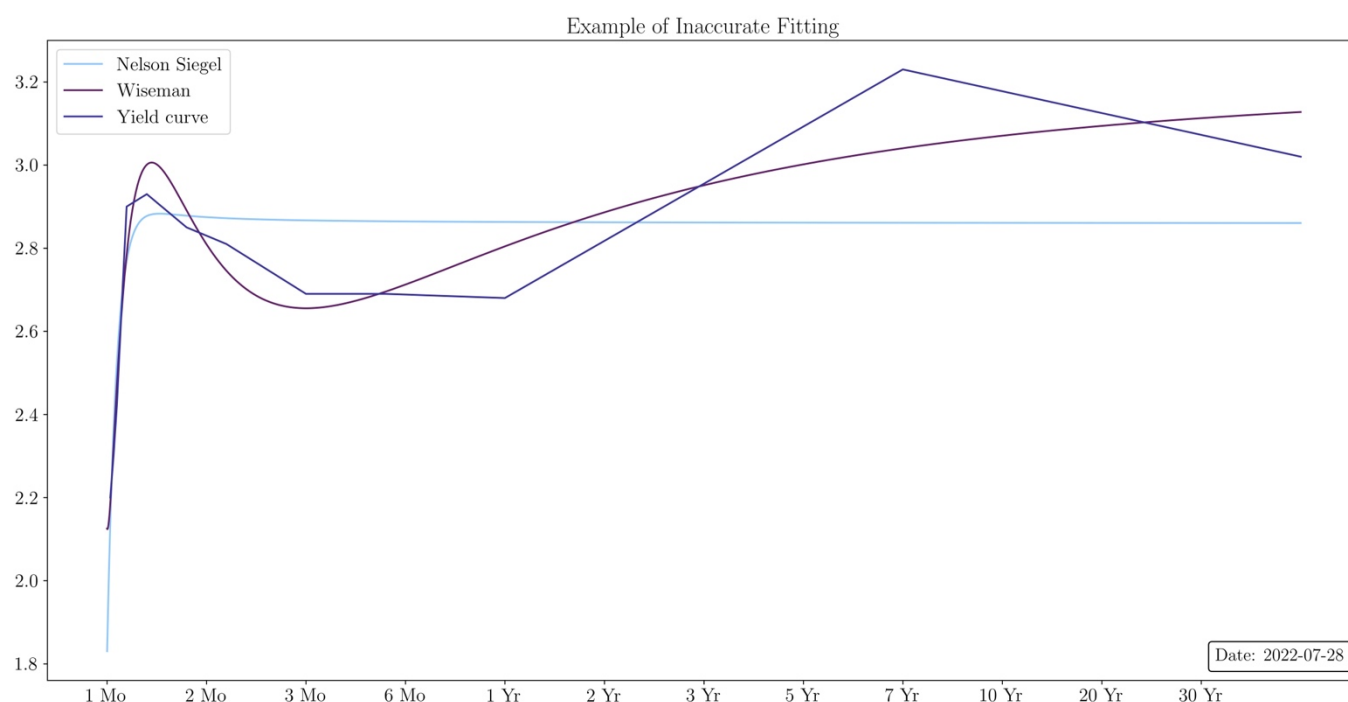

Source: *U.S. Department of the Treasury, Bocconi Students Investment Club*

The medium-term parameter is the most difficult to interpret due to its oscillatory behavior. We observe that in the period from January 2022 to March 2022 the parameter is almost constantly set to 0, resulting in a simplified model that generates a monotonic yield curve, without curvature effect.

Source: *U.S. Department of the Treasury, Bocconi Students Investment Club*

Source: *U.S. Department of the Treasury, Bocconi Students Investment Club*

To conclude, we test the residuals of the model by computing the Mean Squared Error (MSE), and we compare them with those of Wiseman (2012). Overall, both models exhibit greater accuracy in 2021, whereas the residuals in 2022 are higher and more volatile. This is probably due to the abnormal shapes of the yield curve this year (see above). Wiseman's model has consistently lower residuals in the period considered.

TAGS: yield curve, modeling, fixed income

Sources

- Cohen, Kramer, Waugh, 1966. "Regression Yield Curves for U.S. Government Securities".

- Echols, Elliot, 1976. "A Quantitative Yield Curve Model for Estimating the Term Structure of Interest Rates".

- McCulloch, 1971. "Measuring the Term Structure of Interest Rates".

- Vasicek and Fong, 1982. "Term Structure Modeling Using Exponential Splines".

- Nelson, Siegel, 1987. "Parsimonious Modeling of Yield Curves".

- Litterman, Scheinkman, 1991. "Common Factors Affecting Bond Returns".

- Adams, Van Deventer, 1994. "Fitting Yield Curves and Forward Rate Curves with Maximum Smoothness".

- Svensson, 1995. "Estimating forward interest rates with the extended Nelson & Siegel method".

- Jeffery, Linton, Nguyen, 2000. "Flexible Term Structure Estimation: Which Method Is Best?"

- Hagan, West, 2008. "Methods for Constructing a Yield Curve".

- Wiseman, 2012. "The Magpie Yield Curve Model at SG".

- Annaert et alii, 2012. "Estimating the Yield Curve Using the Nelson-Siegel Model: A Ridge Regression Approach".

- ECB, 2018. "Yield curve modelling and a conceptual framework for estimating yield curves: evidence from the European Central Bank's yield curves".