

The Twist Factor of Yields

Xuyang Ma

PanAgora Asset Management, USA

Copyright©2017 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract The twist factor is the fourth latent yield factor following the level, slope, and curvature factors. In this paper, we obtain a four factor composed model with an additional “twist” factor based on the three factor composed Nelson-Siegel model. Specifically, we derive the kernel function for the twist factor and use it to fit the in-sample yields and to forecast future yields. We show that, the twist factor derived within the Nelson-Siegel framework has comparable variance with the curvature factor and helps reducing the fitting errors both in-sample and out-of-sample. Furthermore, we derive portfolios using future yields and yield factors forecasts. The twist factor performs the second best following the slope factor in terms of Sharpe Ratio of formed portfolios.

Keywords Yield Factors, Nelson-Siegel Model, The Twist Factor, Factor Model, Portfolio Construction, Out-of-sample Forecast Evaluations

1. Introduction

In this paper, we propose a new factor --- the twist factor. The twist factor is the fourth latent factor of the yield curve following the level, slope and curvature factor. The kernel function of this factor adds high loadings to the long maturities of the yield curve, thus improves the in-sample and out-of-sample fitting at the long ends of the yield curve. We propose the “twist” factor as a new useful parameter both to forecast future yields and to generate profitable portfolios. Our estimation method is in line with the work of Diebold and Li (2006) [1], in which a single loading parameter λ is used and all latent factors are predicted based on AR(1) models. Lastly we test the portfolio construction using both yields forecast and each single yield factor forecast. Judged by Sharpe Ratio, the twist factor is shown to be a better indicator to form portfolio than the level and the curvature factors.

Since high order yield factors derived using Principle Components Analysis (PCA) in general explain very small portion of the variation of the forward rates, they are normally deemed to be insignificant in practice. Nevertheless,

in this paper we show that, the fourth factor derived using the Nelson-Siegel framework has a comparable variance with the third factor and is a very useful factor in fitting and forecasting future yields, and is also useful in portfolio constructions. To the best of our knowledge, our paper is the first work studying the twist factor within the Nelson-Siegel framework.

Researchers have been trying to use latent factors to predict future interests or yields. Some of them use two factors and others use three factors or some modifications of three factors. For example, Bomfim (2003) [2] and Rudebusch and Wu (2008) [3] considered two-factor affine models to explain interest rate dynamics. Diebold, Piazzesi, and Rudebusch (2005) [4] examined a two-factor Nelson-Siegel model with the first two factors, the level and the slope factor. Bjork and Christensen (1999) [5] proposed to add another slope factor on top of the existing slope factor. The new slope component had a slightly different form with the original slope factor and was designed to mainly affect short-term maturities. Bliss (1996) [6] allowed the slope factor and the curvature factor to have different decay parameters, λ_1 and λ_2 . Svensson (1995) [7] added a second curvature factor with a second value of λ but let the second curvature factor keep the same functional form with the first curvature. De Pooter (2007) [8] adjusted the Svensson (1995) [7] model by slightly changing the formula of the second curvature factor to ensure that multicollinearity was no longer an issue. All these work has only focused on the first two or three factors. After they sense the limitation of the first three factors, they did some variations on the existing factors but never to add a totally new factor. Our research makes up this gap by introducing the twist factor.

Our work can be viewed as a further extension of Diebold and Li (2006) [1] and also an application of the Nelson-Siegel term structure model. Diebold and Li (2006) [1] used the Nelson-Siegel model to fit the in-sample yields and to forecast the out-of-sample yields. Before them, the Nelson-Siegel model has been employed in the Svensson generalization. In particular, Svensson (1995) [7] estimated the yield curve with the level, slope, and two curvature factors. His application is still widely used by practitioners and central banks nowadays.

We follow the convention of Diebold and Li (2006) [1] in

terms of model estimations. We estimated the level, slope, curvature and twist factor with a single loading parameter λ with no more restrictions. Previous literature tested different restrictions or extensions to the model. For example, Christensen, Diebold, and Rudebusch (2009) [9] extended the Nelson-Siegel model by imposing an arbitrage-free condition. Bianchi, Mumtaz, and Surico (2006) [10] studied the assumption of time-varying variance of the latent yield factors instead of constant variance. Koopman, Mallee, and Wel (2010) [11] extended the model in two directions: one is to consider time-varying loading parameter λ_t , and the other one is to allow changes in the overall volatility of interest rates. These changes on model specifications could be very useful but are not the current focus of this paper. For now we intend just to introduce the twist factor and emphasize its usefulness. The variation of model specifications may be a good research topic in the future.

We are not the first to examine the fourth yield factor though. Previous research has tried to test it, not in the form of Nelson-Siegel but derived from principle components. For example, Cochrane and Piazzesi (2005) [12] derived a “return-forecasting factor” which is a linear combination of five forward rates. This “return-forecasting factor” can be recast as a linear combination of the first five yield factors, whereas the yield factors can be derived using principle components. Later, Cochrane and Piazzesi (2009) [13] further examined the fourth and the fifth factors from principle components and noticed that even though the variance of the fifth factor is almost negligible; it adds a significant amount of R^2 to the return forecast regression. Further, Joslin, Priebsch, and Singleton (2014) [14] also included the first five principle components of yield curve into the test when developing their term structure model. None of previous work has tried using Nelson-Siegel framework to derive the fourth yield factor though.

Our application focuses on the future yields forecast and portfolio constructions. Previous work has shown that they can apply the yield model to other aspects as well. For example, Diebold, Li, and Yue (2008) [15] applied the framework to global yield curve forecast including U.S, Germany, Japan and U.K. Yu and Zivot (2008) [16] tested data of corporate bonds and credit rating. It would be interesting to test the application of the twist factor on forecasting other economic data in the future.

The rest of our paper is organized as follows. In section 2, we model the yield curve with the additional twist factor and derive the kernel function of the twist factor. In section 3, we describe the statistics of data. In section 4, the loading parameters are estimated which represent the level, slope, curvature and twist factor. We further characterize their empirical approximations. In section 5, we present the statistics of in-sample yield curve fitting. In section 6, we focus on the portfolio construction using each yield factor. Lastly, in section 7, we discuss our results and conclude the paper.

2. Modeling the Yield Curve: with an Additional Twist Factor

In this section, we introduce the fourth yield factor, hereafter referred as the “twist” factor. We also derive the kernel function for the twist factor presented in the extended Nelson-Siegel function.

2.1. Constructing Yields from Discount Prices and forward Rates

We follow the notation in Diebold and Li (2006) [1] and make use of the relationships among three curves: the discount curve, the forward curve and the yield curve. Let $P_t(\tau)$ denote the price of a τ -period ahead discount bond at time t , i.e., the present value of a one-dollar zero-coupon bond received at τ period ahead. Correspondingly, $y_t(\tau)$ denotes the continuously compounded nominal yield to maturity of this bond.

The discount price curve can be generated by the yield curve using the following equation

$$P_t(\tau) = e^{-\tau y_t(\tau)}.$$

Moreover, the forward rate curve can be generated by the discount price curve using

$$f_t(\tau) = -P'_t(\tau)/P_t(\tau).$$

The relationship between the yield curve and the forward curve is thus given by

$$y_t(\tau) = \frac{1}{\tau} \int_0^\tau f_t(u) du,$$

which states that the yield rate can be derived from the forward rates as an average of the forward rates. Diebold and Li (2006) [1] showed three approaches to construct yields. Two popular approaches to construct yields proceed by estimating a smooth discount curve and then converting to yields at the relevant maturities using equations above. The third approach is due to Fama and Bliss (1987) [18], who constructed yields not via an estimated discount curve, but rather via estimated forward rates at the observed maturities. Their method constructed the “unsmoothed Fama-Bliss” forward rates necessary to price successively long-maturity bonds, and then constructed “unsmoothed Fama-Bliss yields” by averaging the appropriate unsmoothed Fama-Bliss forward rates. The unsmoothed Fama-Bliss yields exactly price the included bonds. Diebold and Li (2006) [1] adopted this third method to construct yields and so do we in this paper.

2.2. Modeling Yields: Nelson-Siegel (1987) Functional Form and its Extension

At any given time, if we have a set of yields with different maturities, we can fit a parametric curve for the purpose of modeling and forecasting. The popular Nelson and Siegel (1987) [17] functional form works with the forward rate curve,

$$f_t(\tau) = \beta_{1t} + \beta_{2t}e^{-\lambda_t\tau} + \beta_{3t}\lambda_t\tau e^{-\lambda_t\tau},$$

which can be viewed as a constant plus a Laguerre function.

The corresponding yield curve is

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right).$$

Since the Nelson and Siegel (1987) [17] forward rate curve includes a polynomial times an exponential decay term which is a mathematical approximating function, we extend the forward curve with an additional term and obtain

$$f_t(\tau) = \beta_{1t} + \beta_{2t}e^{-\lambda_t\tau} + \beta_{3t}\lambda_t\tau e^{-\lambda_t\tau} + \beta_{4t}(\lambda_t\tau)^2 e^{-\lambda_t\tau}.$$

The corresponding yield curve becomes

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right) + \beta_{4t}\left(\frac{2(1-e^{-\lambda_t\tau})}{\lambda_t\tau} - e^{-\lambda_t\tau}(\lambda_t\tau + 2)\right). \quad (1)$$

The formula $(2(1 - e^{-\lambda_t\tau})/\lambda_t\tau - e^{-\lambda_t\tau}(\lambda_t\tau + 2))$ then serves as a kernel function for the fourth latent dynamic factor β_{4t} , namely the “twist” factor. The shape of the kernel function shows up as a twisted “s” shape in several data sets we tested, thus we name it “twist”. As analyzed by Diebold and Li (2006) [1], β_{1t} can be viewed as a long-term factor since the loadings on β_{1t} is one, which is a constant that does not decay as the maturity τ increases. The loading on β_{2t} is $(1 - e^{-\lambda_t\tau})/\lambda_t\tau$, a function that starts at one but decays monotonically and quickly to zero as the maturity τ increases; hence it can be viewed as a short-term factor. The loading on β_{3t} is $((1 - e^{-\lambda_t\tau})/\lambda_t\tau - e^{-\lambda_t\tau})$, which first starts at zero, increases, and then again decays to zero as the maturity τ increases. It has a humped shape and has its maximum loading under a short-median maturity, therefore it may be viewed as a medium-term factor. However, from Figure 1 of the four factor loadings we observe that by construction, the loading curve for β_{3t} is always below the loading curve for β_{2t} , even at β_{3t} 's maximum loading point under a short-median maturity. Thus we lack a high loading under longer maturity. To address this issue, Svensson (1995) [7] added a second curvature factor with a second value of λ ,

such that it can add high loadings on the median-long maturities. In our case, as we observe from Figure 1, by adding the twist factor, we naturally add high loadings on the median-long maturities while using a single λ throughout.

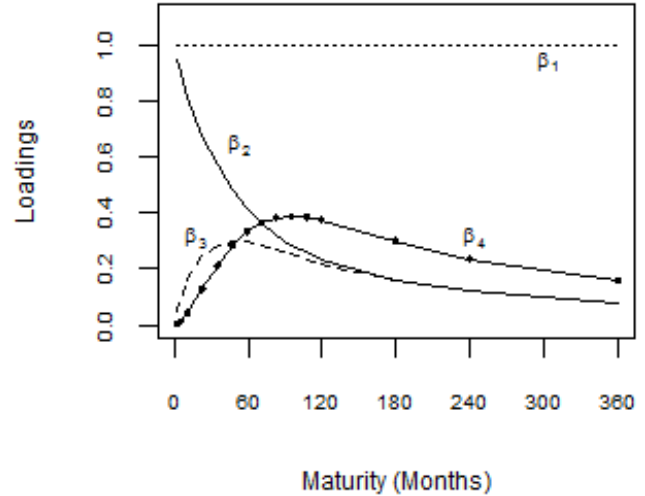


Figure 1. Loadings of four yield factors

Figure 1 plots the loadings of four estimated factors, i.e., the kernel functions of the level, slope, curvature and twist factor. The high loadings that the twist factor brings at median-long maturities explain part of the intuition why we want to incorporate this factor in. The data and λ used to plot Figure 1 will be explained in subsequent sections.

3. Description of Data

We use non-smoothed, zero coupon monthly yields data from Bloomberg. The yields data has fifteen maturities ranged from three months to thirty years. The maturities are 3M, 6M, 1Y, 2Y, 3Y, 4Y, 5Y, 6Y, 7Y, 8Y, 9Y, 10Y, 15Y, 20Y and 30Y, respectively. Data ranges from Dec. 1989 till Sep. 2015. The descriptive statistics of the data are in Table 1.

Table 1. Descriptive statistics of yields

Maturity (Months)	Mean	Std	Min	Max	Acf(1)	Acf(12)	Acf(30)
3	3.03	2.39	0.00	8.05	0.99	0.73	0.31
6	3.16	2.41	0.03	8.39	0.99	0.73	0.33
12	3.31	2.43	0.10	8.73	0.99	0.74	0.37
24	3.65	2.44	0.20	9.18	0.99	0.76	0.45
36	3.89	2.36	0.29	9.25	0.99	0.77	0.49
48	4.12	2.27	0.44	9.23	0.99	0.78	0.53
60	4.35	2.19	0.59	9.22	0.98	0.78	0.56
72	4.50	2.12	0.80	9.22	0.98	0.78	0.57
84	4.66	2.05	1.00	9.23	0.98	0.78	0.58
96	4.78	2.00	1.19	9.24	0.98	0.78	0.59
108	4.90	1.95	1.38	9.24	0.98	0.77	0.60
120	5.00	1.91	1.54	9.25	0.98	0.77	0.60
180	5.16	1.84	1.83	9.23	0.98	0.77	0.62
240	5.35	1.77	2.00	9.27	0.98	0.78	0.63
360	5.82	1.69	2.37	9.48	0.98	0.77	0.63

From Table 1 we can see that yields at each maturity are very persistent, at lag 1 they generally have auto-correlation close to 0.98 and auto-correlations decay with higher lags.

For each yield with a specific maturity, its value fluctuates quite a bit over the past 10 years. To get a more direct view, we plot Figure 2.

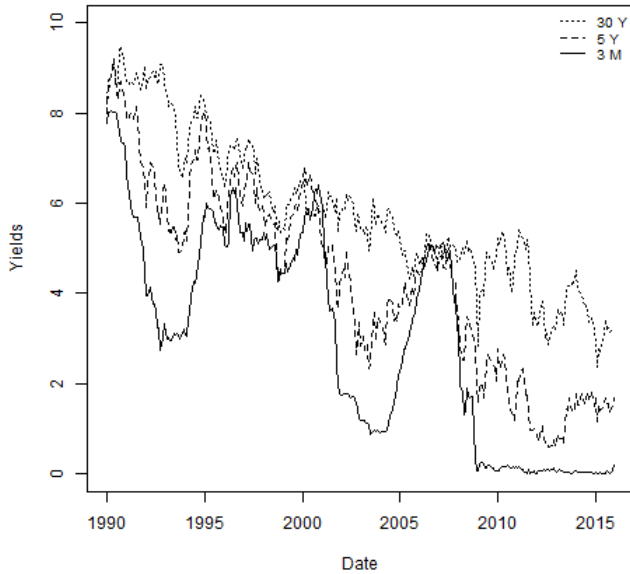


Figure 2. Time-series plot of yields of 3M, 5Y and 30Y.

Figure 2 plots the time series of yields with maturities of 3M, 5Y and 30Y. From 1990 to 2015, the general trend of yields is decreasing. Three-month yield starts from around eight percent in 1990 and drops to around zero percent in 2015 with big fluctuations. Thirty-year yield also drops in a similar but steadier way. Yields with any other maturities also follow a similar down-sloping trend.

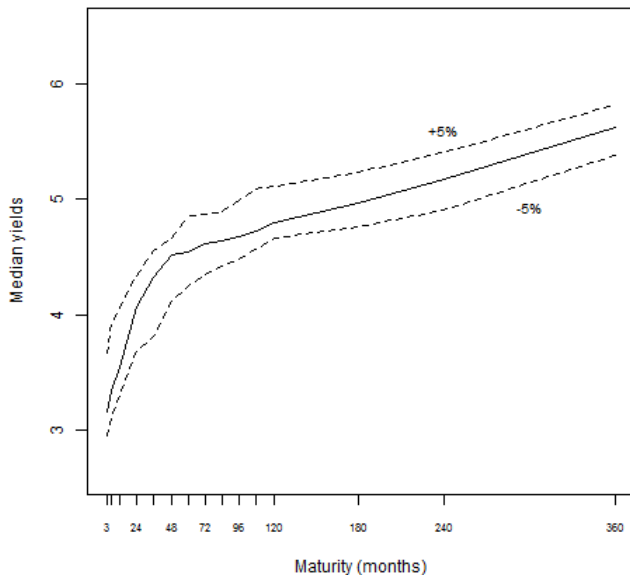


Figure 3. Average yield curve and $\pm 5\%$ quantile.

At a certain time t , yields increase with maturities. In Figure 3, we plot the median yield curve which connects the median value of yields at different maturities, together with the 45th and 55th percentiles of these median yields.

4. Estimation of $\hat{\beta}$ and their Empirical Approximations

4.1. Value of λ and Estimation of $\hat{\beta}$

The value of λ determines at which maturity the loadings on the yield factors achieve their maximums. Diebold and Li (2006) [1] used 0.0609 as its value. This number corresponds to the maximum loading of the curvature factor at thirty months. We adopt their choice of λ into test. We also test the case where λ equals to 0.035. It corresponds to the maximum loading of the curvature factor at four years and the maximum loading of the twist factor at eight years, respectively. This choice of λ corresponds to our understanding of the most reasonable value of λ for the twist factor. Theoretically, we can set λ to any other value and it should not impact the character of the estimation of β_t , but we would want to choose a value of λ so that the loadings for β_4 can be high at median-long maturities; otherwise we are not fully exploring the benefits of this new factor. As can be seen in Figure 1, which is plotted with λ equal to 0.035, the loadings for β_4 goes above the loading of β_2 and β_3 at median-long maturities. The formula of loading for β_3 confines it to be always below the loading of β_2 ; the loading of β_2 decreases monotonously with larger maturities. Thus at median-long maturities, only the loadings of β_4 can possibly be high.

To estimate the time series value of β_t , we regress yields data with 15 maturities onto three or four factor kernels at each month and get the estimated values $(\beta_{1t}, \beta_{2t}, \beta_{3t}, \beta_{4t})$. After we run all the regressions, we have a monthly estimation of β_t , also a time series of regression residuals.

Four different models are compared, two are three factors models and the other two are four factors models. The first three factors model includes the level, slope and curvature factor with λ equal to 0.0609; the second three factors model includes the level, slope, twist factor but no curvature factor with λ equal to 0.035. These two three factor models can help us to compare the curvature and the twist factor. The two four factors models include all four factors with λ equal to 0.0609 and 0.035. The two four factor models will help us to illustrate the impact of different values of λ on the estimation of β_t . As shown in the following tables, different choices of λ give slightly different estimation of β_t but the implications are very similar.

Table 2. Descriptive statistics of $\hat{\beta}_t$.

3 factors, no "twist", $\lambda=0.0609$									
Factors	Mean	Std	Min	Max	Acf_1	Acf_12	Acf_30	ADF	P_value
β_1	5.77	1.69	2.38	9.53	0.98	0.76	0.60	-1.50	0.27
β_2	-2.65	1.67	-5.77	0.79	0.97	0.51	-0.20	-0.96	0.29
β_3	-2.88	2.79	-9.01	3.70	0.95	0.57	0.16	-1.77	0.21
3 factors, no "curvature", $\lambda=0.035$									
Factors	Mean	Std	Min	Max	Acf_1	Acf_12	Acf_30	ADF	P_value
β_1	6.13	1.54	2.58	9.62	0.98	0.74	0.57	-1.42	0.28
β_2	-3.35	2.11	-7.51	0.86	0.98	0.60	-0.12	-0.77	0.17
β_4	-1.13	1.50	-5.21	2.16	0.91	0.38	0.19	-2.85	0.01
4 factors, $\lambda=0.0609$									
Factors	Mean	Std	Min	Max	Acf_1	Acf_12	Acf_30	ADF	P_value
β_1	5.93	1.63	2.44	9.61	0.98	0.75	0.61	-1.49	0.33
β_2	-3.07	1.74	-6.50	0.72	0.97	0.54	-0.18	-0.76	0.74
β_3	-1.21	2.32	-6.10	6.46	0.93	0.48	0.02	-2.56	0.01
β_4	-1.33	1.15	-4.72	1.25	0.86	0.31	-0.09	-2.49	0.0003
4 factors, $\lambda=0.035$									
Factors	Mean	Std	Min	Max	Acf_1	Acf_12	Acf_30	ADF	P_value
β_1	6.06	1.63	2.50	9.63	0.98	0.76	0.62	-1.46	0.31
β_2	-3.14	1.77	-6.96	0.65	0.97	0.54	-0.20	-0.74	0.71
β_3	-0.96	2.62	-6.39	7.37	0.92	0.49	0.34	-3.09	0.00
β_4	-0.40	1.03	-3.82	2.17	0.86	0.34	0.08	-3.74	0.000003

Table 2 presents the descriptive statistics of the estimated yield factors ($\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\beta}_{4t}$). The estimated factors under different models present similar characters. The estimated level factor has a positive mean and the other three factors have negative means. The higher level of the factor, the less absolute value it has. For example, the estimated level factor has an average value around 6 and the estimated twist factor has an average value around -1. The estimated curvature factor has the highest volatility around 2.5. The estimated twist factor has the lowest volatility around 1.

We also present autocorrelations at three lags of one, twelve and thirty. The autocorrelations are high at lag one around 0.9 and decrease with higher lags. The last two columns present statistics of the Augmented Dickey-Fuller (ADF) unit-root test. The column ADF presents the value of the ADF test; the column P_value presents the P_value of the test. The null hypothesis of a unit root is not rejected if the P_value is higher than a critical value of five percent or one percent. For simplicity, we set test formula with lag equals one. The critical values for rejection of the hypothesis of a unit root are then -2.58 at one percent level, -1.95 at five percent level and -1.62 at ten percent level. From the table we can see that the ADF tests suggest that $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ may have unit roots, and $\hat{\beta}_{3t}$ and $\hat{\beta}_{4t}$ may not.

4.2. Empirical Approximation for the Twist Factor

Diebold and Li (2006) [1] gave very clear interpretations on the empirical approximations for the estimated yield factors using yields data with maturities up to ten years. In this section we do a similar exercise using yields data with maturities up to thirty years. Our empirical approximations are slightly different due to longer data sets. Additionally, we find the empirical approximation for the twist factor.

To measure the level factor β_{1t} , we use the longest maturity available since theoretically $y_t(\infty) = \beta_{1t}$. That is to use the thirty-year yield to approximate β_{1t} . Similarly, since $y_t(\infty) - y_t(0) = -\beta_{2t}$, the appropriate measure for the slope factor would be the difference between the longest maturity yield and the shortest maturity yield. That is the difference between the thirty-year yield and the three-month yield, i.e.

$y_t(360) - y_t(3) = -0.87\beta_{2t} + 0.03\beta_{3t} + 0.16\beta_{4t}$. By definition, β_{1t} governs the level of the yield curve and β_{2t} governs its slope.

The instantaneous yield curve $y_t(0)$ is often modeled as an affine function of two state variables. Under the Nelson-Siegel three factor model, we have $y_t(0) = \beta_{2t} + \beta_{2t}$, which matches the two-factors specification in the literature. Under the four factor model, we still have $y_t(0) = \beta_{2t} + \beta_{2t}$.

The empirical approximation of the curvature factor requires three maturities including the short end, the long end and the middle maturity. The short end and the long end can obviously be set to three months and thirty years. We choose three years as the middle maturity. In result, we have $2y_t(36) - y_t(3) - y_t(360) = 0.11\beta_{2t} + 0.44\beta_{3t} + 0.26\beta_{4t}$.

If we interpret the slope factor as a measurement for the steepness of the whole yield curve, and the curvature factor as a measurement for the difference between the steepness of long maturity yields and the steepness of the short maturity yields, then, the twist factor would represent the difference of slope change under the short-, median- and long-maturity. If

we choose two-year and six-year as two median maturities then the empirical twist factor becomes $2(y_t(72) - y_t(24) - (y_t(360) - y_t(72)) - (y_t(24) - y_t(3)))$, or $y_t(3) - 3y_t(24) + 3y_t(72) - y_t(360) = -0.06\beta_{2t} + 0.09\beta_{3t} + 0.56\beta_{4t}$.

Figure 4 plots the time series of estimated yield factors $(\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\beta}_{4t})$ and their empirical approximations. The correlations between regression estimated factors and their empirical approximations are 0.996, 0.997, 0.954 and 0.906, respectively.

Figure 4. $\hat{\beta}_t$ and their empirical approximations

5. Yield Curve Fitting and Yields Factors Modeling

5.1. In-Sample Yield Curve Fitting

We fit the two three-factor models and the two four-factor models to compare the in-sample fitting of the yield curve. The formulas for the two three factor models are $y_t(\tau) = \beta_{1t} + \beta_{2t} \left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} \right) + \beta_{3t} \left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau} \right)$ with λ equal to 0.0609 and $y_t(\tau) = \beta_{1t} + \beta_{2t} \left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} \right) + \beta_{4t} \left(\frac{2(1-e^{-\lambda_t\tau})}{\lambda_t\tau} - e^{-\lambda_t\tau}(\lambda_t\tau + 2) \right)$ with λ equal to 0.035. The formulas for the two four factor models are $y_t(\tau) = \beta_{1t} + \beta_{2t} \left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} \right) + \beta_{3t} \left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau} \right) + \beta_{4t} \left(\frac{2(1-e^{-\lambda_t\tau})}{\lambda_t\tau} - e^{-\lambda_t\tau}(\lambda_t\tau + 2) \right)$ with λ equal to 0.0609 and 0.035, respectively. The descriptive statistics of the fitting residuals at each maturity for the four factor model where λ equals to 0.035 are presented in table 3. The first four columns present the mean, standard deviation, minimum and maximum of the residuals. The in-sample fitting errors increase with maturities. MAE stands for mean absolute error. Table 3 also presents residual sample autocorrelations at the displacement of one, twelve and thirty months. Autocorrelations become less significant

with higher displacement and residuals of longer maturities have higher autocorrelations. At displacement of one, autocorrelations range from 0.58 to 0.96. At displacement of thirty, autocorrelations can become negative. The last four columns compare the root mean square errors (RMSE) of all four models. RMSE_3_006 stands for root mean square error for the three factor models without the twist factor with λ equal to 0.0609; RMSE_3_003 stands for root mean square error for the three factor models without curvature factor with λ equal to 0.035; RMSE_4_006 stands for root mean square error for the four factor models with λ equal to 0.0609; RMSE_4_003 stands for root mean square error for the four factor models with λ equal to 0.035.

From the last four columns of table 3 where RMSEs are presented we can see that the four factor models generally fit better than the three factor models with smaller errors. Figure 5 plots the RMSEs in the last four columns of table 3. Comparing the plots of the two different three factor models, we can see that the twist factor is useful at reducing fitting errors at long maturities; the curvature factor is useful at reducing fitting errors at median maturities. Comparing two four factor models with different values of λ , we can see that $\lambda = 0.035$ performs better than $\lambda = 0.0609$ in terms of reduce fitting errors at the long-ends.

Table 3. The descriptive statistics of yield curve in-sample fitting residuals, mainly for the four factor model with $\lambda = 0.035$.

Maturity (M)	Mean	Std	Min	Max	MAE	Acf_1	Acf_12	Acf_30	RMSE_3_006	RMSE_3_003	RMSE_4_006	RMSE_4_003
3	-0.001	0.07	-0.31	0.11	0.05	0.78	0.24	-0.09	0.12	0.26	0.05	0.07
6	0.02	0.05	-0.12	0.28	0.04	0.58	0.23	-0.06	0.05	0.13	0.05	0.05
12	-0.02	0.08	-0.24	0.31	0.06	0.83	0.28	-0.11	0.13	0.11	0.07	0.08
24	-0.0001	0.05	-0.11	0.17	0.04	0.80	0.19	-0.08	0.14	0.24	0.05	0.05
36	-0.03	0.03	-0.15	0.07	0.03	0.76	0.13	-0.09	0.06	0.23	0.04	0.04
48	-0.02	0.03	-0.12	0.09	0.03	0.80	0.28	0.18	0.04	0.16	0.04	0.04
60	0.02	0.05	-0.13	0.17	0.04	0.86	0.15	-0.17	0.07	0.09	0.06	0.06
72	0.01	0.03	-0.08	0.17	0.02	0.81	0.04	-0.16	0.09	0.03	0.02	0.03
84	0.03	0.05	-0.06	0.18	0.04	0.92	0.45	0.06	0.10	0.09	0.05	0.06
96	0.02	0.02	-0.03	0.09	0.02	0.88	0.57	0.29	0.09	0.10	0.03	0.03
108	0.03	0.03	-0.02	0.13	0.03	0.79	0.28	0.07	0.07	0.12	0.04	0.04
120	0.04	0.05	-0.06	0.23	0.05	0.86	0.20	-0.04	0.06	0.15	0.05	0.06
180	-0.13	0.10	-0.48	0.10	0.14	0.96	0.49	-0.10	0.14	0.12	0.17	0.17
240	-0.13	0.09	-0.48	0.07	0.13	0.96	0.53	-0.01	0.09	0.14	0.14	0.16
360	0.15	0.12	-0.12	0.63	0.16	0.95	0.45	-0.11	0.36	0.16	0.25	0.19

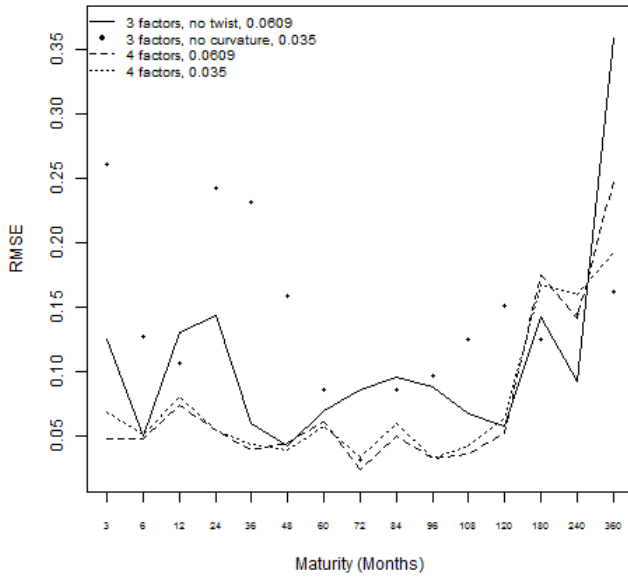


Figure 5. RMSE of the in-sample fitting of four models

5.2. Yields Factors Modeling

Diebold and Li (2006) [1] compared ten out-of-sample yield forecasting models and found that the Nelson-Siegel model outperformed all the other models at 12-month-ahead forecasts. Since in this paper we do not focus on comparing different forecasting models, we only use the 1-month-ahead

forecasts for the latent yield factors forecasts since 1-month-ahead forecasts give the most precise forecasts compared to the 6-month-ahead and the 12-month-ahead forecasts.

We use univariate AR(1) model to model and forecast yield factors. The future yields forecasts are based on forecasted betas. The AR(1) model for forecasting yield factors is

$$\hat{\beta}_{i,t+1/t} = \hat{c}_i + \hat{\gamma}_i \hat{\beta}_{i,t}, i = 1, 2, 3, 4.$$

\hat{c}_i and $\hat{\gamma}_i$ are obtained by regressing $\hat{\beta}_{i,t}$ on an intercept and $\hat{\beta}_{i,t-1}$. We use 1-month-ahead regressions and run regressions with overlapping data. Note that here we use a constant value for λ_t , so λ_t equals to λ and equals to 0.0609 or 0.035. Both Figure 6 and Figure 7 are plotted based on the four factor model with λ equal to 0.035.

The univariate AR(1) model is chosen because of its simplicity and it is reasonable as discussed in Diebold and Li (2006) [1]. To check whether the AR(1) model is appropriate, we plot the autocorrelations of yields factors and of estimation residuals. What we want to see is that autocorrelations of yields factors are large which means yield factors are persistent, but the autocorrelations of regression residuals are small which means residuals are not as persistent. The regression model is:

$$\hat{\beta}_{i,t+1} = \hat{c}_i + \hat{\gamma}_i \hat{\beta}_{i,t} + \hat{\varepsilon}_{i,t+1}, i = 1, 2, 3, 4.$$

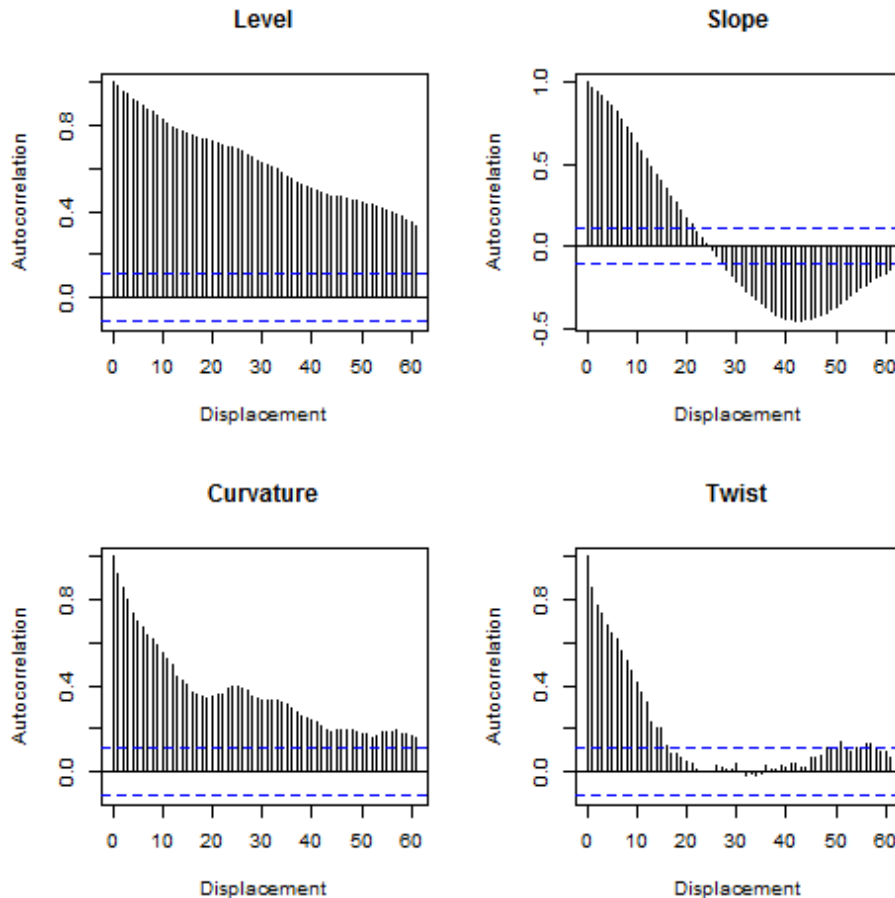


Figure 6. Autocorrelations of $\hat{\beta}_t$.

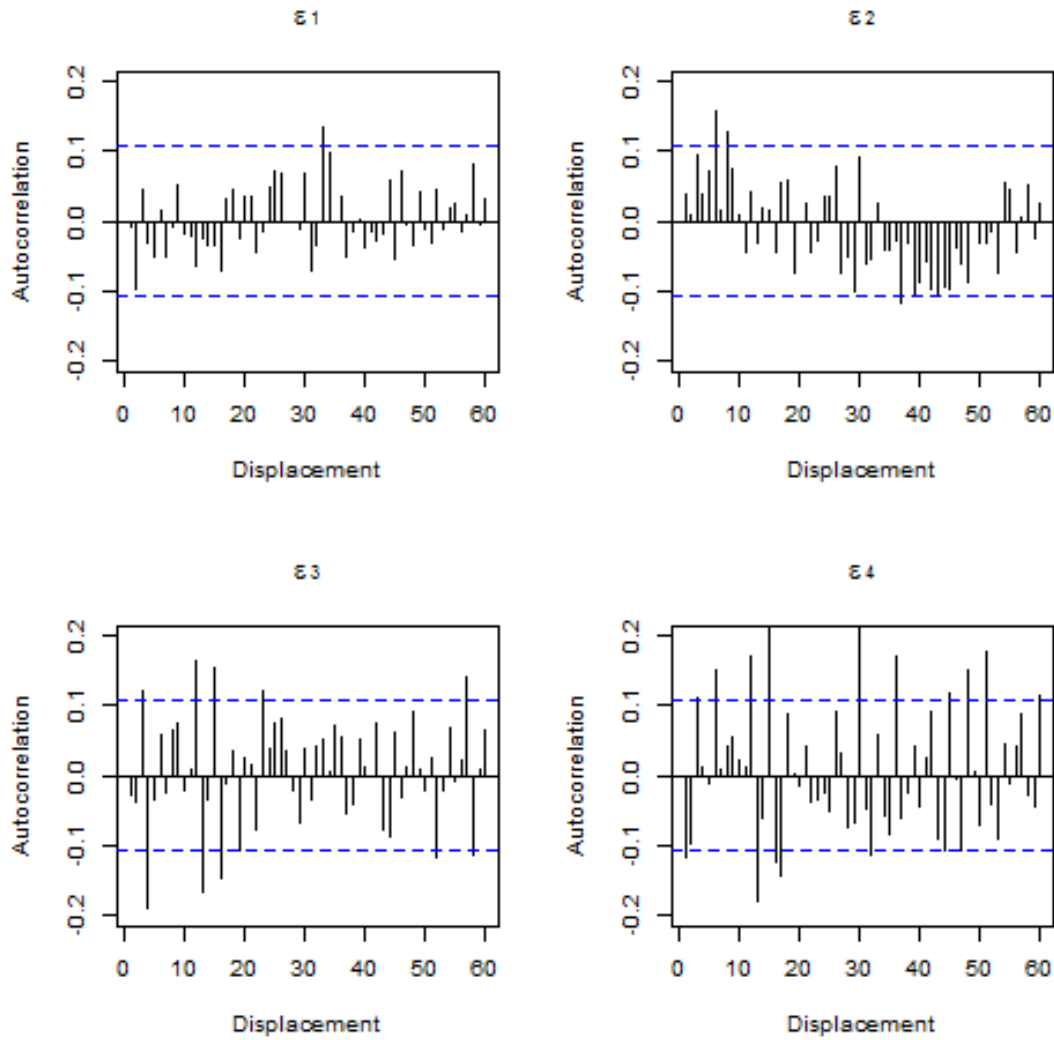


Figure 7. Autocorrelations of $\hat{\varepsilon}_t$.

We plot the autocorrelations of $(\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\beta}_{4t})$, along with Bartlett’s approximate 95% confidence bands in Figure 6. From Figure 6 we can see that the yield factors $\hat{\beta}_t$ are very persistent. We plot the autocorrelations of the regression residuals $(\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t}, \hat{\varepsilon}_{3t}, \hat{\varepsilon}_{4t})$ in Figure 7, along with Bartlett’s approximate 95% confidence bands. The autocorrelations of regression residuals are small and negligible compared to that of $\hat{\beta}_t$. This confirms that the AR(1) model for yields factors is appropriate.

5.3. Out-of-sample Yield Curve Fitting

The in-sample fitting in previous section illustrates the point that the twist factor improves yields fitting at the long-ends, in this section we further present the statistics of

out-of-sample yield curve fitting to emphasize this point.

Table 4 presents the statistics of the out-of-sample fitting residuals for the four factor model with λ equal to 0.035. The first four columns present the mean, standard deviation, minimum and maximum. For the four factor model, the fitting error is large at the median maturities and small at the short ends and long ends. MAE stands for mean absolute error. We again present the auto-correlations at displacement of one, twelve and thirty. The auto-correlations at displacement one are between 0.80 and 0.90 . The auto-correlations decrease when displacements go up. The last four columns present the root mean square error of all models which are plot in Figure 8 for comparison.

Table 4. The descriptive statistics of yield curve out-of-sample fitting residuals, mainly for the four factor model with $\lambda = 0.035$.

Maturity (M)	Mean	Std	Min	Max	MAE	Acf_1	Acf_12	Acf_30	RMSE_3_006	RMSE_3_003	RMSE_4_006	RMSE_4_003
3	0.02	0.08	-0.31	0.18	0.07	0.80	0.18	0.02	0.11	0.26	0.10	0.08
6	0.09	0.09	-0.18	0.28	0.11	0.81	0.28	0.24	0.19	0.13	0.18	0.13
12	0.08	0.14	-0.30	0.38	0.13	0.87	0.40	0.18	0.41	0.12	0.24	0.16
24	0.02	0.06	-0.13	0.20	0.05	0.86	0.52	0.19	0.55	0.31	0.34	0.06
36	-0.07	0.06	-0.22	0.11	0.08	0.90	0.31	-0.29	0.51	0.32	0.37	0.09
48	-0.12	0.12	-0.41	0.22	0.14	0.92	0.42	-0.01	0.45	0.25	0.39	0.17
60	-0.13	0.18	-0.55	0.33	0.17	0.90	0.40	0.09	0.43	0.16	0.45	0.22
72	-0.17	0.20	-0.58	0.41	0.21	0.91	0.34	-0.04	0.35	0.09	0.44	0.26
84	-0.16	0.20	-0.54	0.45	0.21	0.90	0.25	-0.16	0.34	0.06	0.47	0.26
96	-0.17	0.21	-0.55	0.47	0.23	0.90	0.29	-0.07	0.29	0.05	0.46	0.27
108	-0.15	0.20	-0.51	0.46	0.21	0.90	0.30	0.02	0.29	0.07	0.48	0.26
120	-0.13	0.19	-0.47	0.45	0.20	0.90	0.30	0.09	0.30	0.11	0.50	0.23
180	-0.29	0.25	-0.78	0.43	0.32	0.93	0.43	0.02	0.08	0.15	0.22	0.38
240	-0.25	0.21	-0.70	0.31	0.27	0.94	0.48	0.07	0.10	0.19	0.18	0.33
360	0.08	0.06	-0.03	0.21	0.08	0.92	0.30	-0.28	0.45	0.11	0.44	0.10

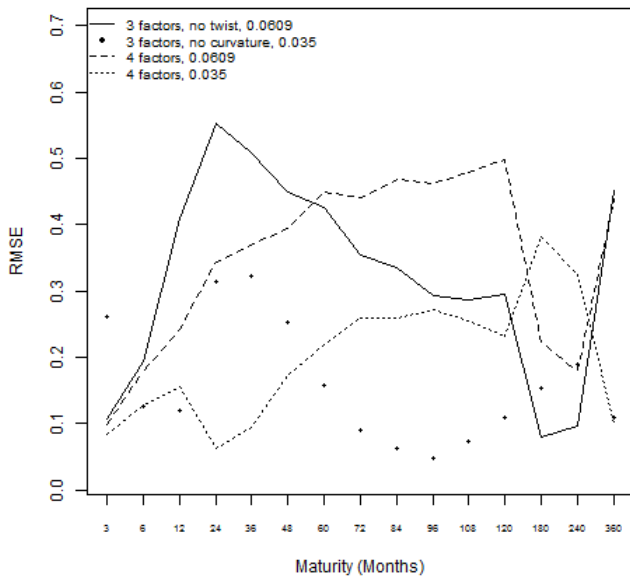


Figure 8. RMSE of the out-of-sample fittings of four models.

Figure 8 plots the last four columns of table 4, the RMSE of the out-of-sample fittings of four models. Comparing the two three factors models with different parameters, we can see that the model with the twist factor on average have smaller fitting errors than the model with the curvature factor. Comparing the two four factor models, we can see that the model with λ equal to 0.035 on average have smaller fitting errors than the model with λ equal to 0.0609.

6. Portfolio Constructions Based on Yields Forecast and Yield Factors Forecast

In this section we test yield factors’ abilities in making profitable portfolios. As shown in Equation 1, the total yield equals to the level factor multiply the level kernel which is constant one, plus the slope factor multiply the slope kernel,

plus the curvature factor multiply the curvature kernel, plus the twist factor multiply the twist kernel. Thus, whenever yield factors increase/decrease, the total yields would increase/decrease too. In other words, there is a linear relationship between yield factors and yields. Since yields have a negative correlation with future bond returns, increase in yields or yield factors should indicate decrease in future bond returns. This negative correlation serves as the intuition for our portfolio construction and we form portfolios based on forecast of the future changes of yields. We also form portfolios based on each individual yield factor’s forecasted change because we want to compare each single factor’s ability in forming profitable portfolios.

The forecasted yields take into account all four yield factors’ forecast. Take the slope factor as an example: suppose at time t the forecasted next period change of slope is positive, then at time t at each maturity we multiply the forecasted change with the kernel function of slope and also with a negative sign, thus a positive forecasted change of slope would become a negative value. We add the four values for the four factors together and the final value would serve as the indicator for yields change. We divide the final values by the durations of their corresponding maturities and then change them proportionally into weighting scores. The weighting method we use is a long-only transformation by ranking all the values stepwise and then scales them linearly. This way the final weights would be in equal stepwise increments.

Next we multiply the final weighting score for each maturity with the next period future return to have a forecasted return series. We then sum the returns of different maturities at t to make one single return and that would serve as the forecasted portfolio’s return at t .

Mathematically, we first use AR(1) model to forecast next period’s latent factors, then get the forecasted change by using the forecasted value to minus the current value, i.e., $E_t(\Delta\beta_{i,t+1}) = E_t(\beta_{i,t+1} - \beta_{i,t}) = \hat{\beta}_{i,t+1/t} - \beta_{i,t}$, where $\hat{\beta}_{i,t+1/t} = \hat{c}_i + \hat{\gamma}_i\hat{\beta}_{i,t}$, $i = 2, 3, 4$. We multiply $E_t(\Delta\beta_{i,t+1})$

with the kernel function at each maturity and then change these values into weights by using linear long-only methods mentioned above. Finally we multiply these weights with future returns of next period to generate returns and sum up the returns of different maturities at t to get the portfolio's return at t .

The future return data we use is from Bloomberg and the data includes nine maturities. They are monthly data from Jan 1992 to Aug 2015. The statistics of data is summarized in Table 5. As maturities increases, the averages of returns increase and the standard deviations also increase.

The generated portfolios start from Nov. 1999 and end by Aug. 2015. In total we have 190 monthly observations of data. Statistics of the formed portfolios are summarized in Table 6 and Table 7. Table 6 presents statistics of portfolios using four different models, model 1 is the three factors model with no twist factor while λ is set to 0.0609, model 2 is the three factors model with no curvature factor while λ is

set to 0.035, model 3 is the four factors model with λ equal to 0.0609, model 4 is the four factors model with λ equal to 0.035. Table 7 presents statistics of portfolios formed using each individual yield factor. Based on Sharpe Ratio in Table 7, the slope factor generates the most efficient portfolio with a Sharpe Ratio of 1.43, and the twist factor generates the second most efficient portfolio with a Sharpe Ratio of 0.79. The level factor generates the least efficient portfolio with a negative Sharpe Ratio. The information revealed from portfolios formed by total yields is less interpretable because the four factor models do not necessarily perform better than the three factor models in terms of Sharpe Ratio.

Overall, the twist factor serves as a reliable indicator for portfolio constructions. Even though the absolute variation that the twist factor contributes to the yield curve is small, but it contains important and unique information of the yield curve. The information contained is useful for forming profitable portfolios.

Table 5. Descriptive statistics of future returns data.

Maturity (Months)	Mean	Std	Min	Max	Acf(1)	Acf(12)	Acf(30)
3	0.01	0.05	-0.14	0.35	0.11	-0.02	-0.05
6	0.03	0.11	-0.26	0.67	0.19	-0.02	-0.08
12	0.06	0.26	-0.65	1.15	0.22	0.00	-0.10
24	0.14	0.59	-1.52	1.77	0.20	0.02	-0.08
36	0.20	0.92	-2.58	2.77	0.17	0.02	-0.05
48	0.26	1.24	-4.13	4.21	0.15	0.01	-0.03
60	0.32	1.54	-5.69	5.71	0.14	0.01	-0.02
120	0.36	1.71	-5.80	8.68	0.08	0.03	0.06
360	0.44	2.70	-9.99	13.65	0.05	-0.03	0.09

Table 6. Performances of portfolios based on forecasted yields change.

Yields	Mean Y	Std Y	Sharpe Ratio Y	Min M	Max M	Skewness M	Kurtosis M
Model 1	2.02	2.43	0.83	-2.48	2.03	-0.28	1.01
Model 2	2.03	3.29	0.62	-3.50	2.45	-0.69	1.27
Model 3	1.79	2.43	0.73	-2.28	2.02	-0.55	1.10
Model 4	1.45	2.96	0.49	-3.19	2.22	-0.98	2.15

Table 7. Performances of portfolios based on forecasted yields factors change.

Factors	Mean Y	Std Y	Sharpe Ratio Y	Min M	Max M	Skewness M	Kurtosis M
Level	-0.17	3.50	-0.05	-4.88	2.63	-1.04	3.55
Slope	5.30	3.70	1.43	-1.97	6.18	1.16	4.22
Curvature	2.28	4.01	0.57	-4.88	6.18	0.08	5.82
Twist	2.65	3.35	0.79	-3.41	3.69	-0.12	1.38

7. Summary

We have extended the yield curve model of Nelson and Siegel (1987) [17] and Diebold and Li (2006) [1] to incorporate a new latent factor, the twist factor. The twist factor is the fourth yield factor following the level, slope and curvature factor. We derive its kernel function and show that it naturally adds loadings to yields at long maturities and thus help fitting and forecasting the yield curve at the long ends. The estimated twist factor can be closely approximated using short maturity minus median-short maturity plus median-long maturity minus long maturity, and that is why we named it the “twist” factor. Using monthly data from Dec. 1989 to Sep. 2015, we show that the twist factor can also help generating profitable portfolio.

REFERENCES

- [1] Diebold FX, Li C. Forecasting the term structure of government bond yields. *Journal of econometrics*. 2006 Feb 28;130(2):337-64.
- [2] Bomfim AN. Monetary policy and the yield curve.
- [3] Rudebusch GD, Wu T. A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy. *The Economic Journal*. 2008 Jul 1;118(530):906-26.
- [4] Diebold FX, Piazzesi M, Rudebusch G. Modeling bond yields in finance and macroeconomics. *National Bureau of Economic Research*; 2005 Jan 31.
- [5] Björk T, Christensen BJ. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*. 1999 Oct 1;9(4):323-48.
- [6] Bliss RR. Testing term structure estimation methods. Working paper, Federal Reserve Bank of Atlanta; 1996.
- [7] Svensson LE. Estimating forward interest rates with the extended Nelson & Siegel method. *Sveriges Riksbank Quarterly Review*. 1995 Sep;3(1):13-26.
- [8] De Pooter M. Examining the Nelson-Siegel class of term structure models: In-sample fit versus out-of-sample forecasting performance.
- [9] Christensen JH, Diebold FX, Rudebusch GD. An arbitrage-free generalized Nelson–Siegel term structure model. *The Econometrics Journal*. 2009 Nov 1;12(3).
- [10] Bianchi F, Mumtaz H, Surico P. The UK great stability: a view from the term structure of interest rates. Working Paper; 2006.
- [11] Koopman SJ, Mallee MI, Van der Wel M. Analyzing the term structure of interest rates using the dynamic Nelson–Siegel model with time-varying parameters. *Journal of Business & Economic Statistics*. 2010 Jul 1;28(3):329-43.
- [12] Cochrane JH, Piazzesi M. Bond Risk Premia. *American Economic Review*. 2005 Mar;95(1):138-60.
- [13] Piazzesi M, Cochrane J. Decomposing the Yield Curve. In 2009 Meeting Papers 2009 (No. 18). Society for Economic Dynamics.
- [14] Joslin S, Priebsch M, Singleton KJ. Risk premiums in dynamic term structure models with unspanned macro risks. *The Journal of Finance*. 2014 Jun 1;69(3):1197-233.
- [15] Diebold FX, Li C, Yue VZ. Global yield curve dynamics and interactions: a dynamic Nelson–Siegel approach. *Journal of Econometrics*. 2008 Oct 31;146(2):351-63.
- [16] Yu W, Zivot E. Forecasting the term structures of treasury and corporate yields: Dynamic nelson-siegel models evaluation.
- [17] Nelson CR, Siegel AF. Parsimonious modeling of yield curves. *q Journal of Business*. 1987;60(4):473-89.
- [18] Fama EF, Bliss RR. The information in long-maturity forward rates. *The American Economic Review*. 1987 Sep 1:680-92.