Pricing the Term Structure with Linear Regressions

Authors:

Tobias Adrian

Richard K. Crump

Emanuel Moench

Studied by:

Xiyu Xue

Arthikan Rajaratnam

Joudy Benkaddour

Abdoulaye Gaye

Paris Dauphine University-PSL

Saturday, February 3, 2024

Introduction

Estimating the term structure of interest rates is pivotal within the financial sector, guiding the strategic decisions of investors, companies, and financial institutions. This term structure, commonly depicted through the yield curve, offers insights into market expectations regarding future interest rate movements. Grasping and foreseeing these shifts is imperative for prudent portfolio management and well-informed decision-making.

The advent of affine term structure models represented a notable breakthrough. These models propose linear relationships between coefficients and state variables, providing analytical tractability and efficient estimation. The ease of parameter estimation and the interpretability of affine models contributed to their broad adoption in academic and industrial settings. However, these models hinge on specific assumptions that give rise to certain outcomes, ultimately influencing their conclusions.

The paper we are about to present aligns with an alternative to affine term structure models by employing a three-step estimation method to model (with a low margin of error, as we will later demonstrate) an interest rate curve based on data that may span various types, especially without observing a zero-coupon yield curve. This confers an advantage to this estimation method over those found in the existing literature.

To illustrate their estimation methodology, a selected database employed by the authors showcases its efficacy. This database comprises zero-coupon yield information curated by Gurkaynak, Sack, and Wright (2007, GSW), widely recognized as a prominent dataset for zero-coupon Treasury yield curves. Encompassing zero-coupon yields across various maturities, the GSW dataset also contains crucial parameters necessary for implementing the Nelson-Siegel model. Utilizing these parameters, we will calculate excess returns on cross sections of yields spanning from January 1987 to January 2024.

In order to accurately construct a yield curve, it is necessary to obtain yield data on bonds across a spectrum of different maturities. Therefore, we will use the Nelson-Siegel-Svensson formula (developed by Andrew Nelson, Robert Siegel, and Pierre-Olivier Gourinchas) to generate yield data and further calculate the excess returns:

$$f(t, t+m) = \beta_0 + \beta_1 e^{-\frac{m}{\tau_1}} + \beta_2 \frac{m}{\tau_1} e^{-\frac{m}{\tau_1}} + \beta_3 \frac{m}{\tau_2} e^{-\frac{m}{\tau_2}}$$

with :

β_0 : Constant, representing the long-term level of interest rates.

β_1 : Coefficient associated with the first exponential component, influencing the slope of the yield curve.

β_2 : Coefficient associated with the second exponential component, contributing to the curvature of the yield curve.

β_3 : Coefficient associated with the third exponential component, influencing a possible second hump or inflection point in the yield curve.

The parameters τ_1 and τ_2 are time constants controlling the rate at which the exponential components decrease over time. Specifically, τ_1 is associated with the decrease of the slope, and τ_2 with the decrease of the curvature.

Once we computed the new yields using the Nelson-Siegel method, it seemed prudent to compare them with the original base, that of GSW. Appendix in Table 1 provides a summary of this comparison. From this analysis, several points can be observed. Initially, we note a percentage difference in yields. Across various maturities, we observe that, on average, the difference between yields is around 1% increasing with the maturity. However, as indicated in the second column, this comparison exhibits a variance which also increase with the maturity. Consequently, we can infer that the utilization of the Nelson-Siegel method to calculate the yields of our zero-coupon bonds slightly alters the initial yields. Nonetheless, this adjustment is made in alignment with the methodology outlined in the article.

Factor Retrieval

Preliminary Step

Initially, we will convert the GSW database using the Nelson-Siegel-Svensson formula. This involves generating multiple cross-sections (for each observation) of bond yields with different maturities (in months). We transform this dataset by calculating the prices of zero-coupon bonds, where the price is determined by the yield and maturity:

$$P = \exp(-y \cdot T)$$

After obtaining the prices, we create a new dataset containing the log returns of zero-coupon bonds, effectively reducing the number of observations. The log-return dataset will later enable us to compute the excess returns required for a future regression.

Within the log-return dataset, we designate the first column (as indicated in the article) as the column representing the risk-free return. Subsequently, we calculate excess returns (subsequent columns) relative to the first column.

The calculation of excess return is based on the formula below:

$$rx_{t+1}^{(n-1)} = \ lnP_{t+1}^{(n-1)} - lnP_t^{(n)} - r_t$$

with :
$rx_{t+1}^{(n-1)}$: the log excess holding return of a bond maturing in n periods.
$lnP_t^{(n)}$: the log zero-coupon Treasury bond price with maturity n at time t.
r_t : the continuously compounded risk-free rate.

Factor Extraction

For this purpose, we will employ Principal Component Analysis (PCA): an analysis that transforms correlated variables into new variables that are uncorrelated with each other, termed "principal components."

It is crucial to center and scale the data points for various reasons: to prevent a variable with high variance from exerting undue influence on the PCA. Subsequently, we compute the variance-covariance matrix of our scaled returns. This matrix, being square, symmetric, and real-valued, is diagonalizable. We retrieve its eigenvectors and eigenvalues. Next, we extract the principal components of returns by centering and normalizing the product of the matrix of standardized returns by the eigenvectors of the covariance matrix of standardized returns.

This yields a matrix composed of observations for each of our principal components, serving as our factors.

After obtaining the factors, we conducted an analysis to assess how much variance each principal component captures from the yield data. Typically, the first principal component accounts for the largest possible variance in the data set. The second principal component, being uncorrelated with the first principal component, accounts for the subsequent highest variance, and so on.

We first computed the explained_variance_ratio of each principal component. It is evident that the first component explains approximately 84.32% of the variance, signifying its dominant role in representing the dataset's variability. Subsequent components contribute much less, with values of 15.67%, 0.01%, 0.0006%, and 0.0000005%, respectively. The steep drop in the variance explained by subsequent components implies the diminishing contribution to capturing yield data variance. Furthermore, we calculated the cumulative_variance_ratio to show the cumulative variance explained by the first K factors. The results revealed that all five components collectively account for an overwhelming majority (approximately 99%) of the variance as shown in Table 2. In summary, the PCA model is quite effective in capturing the total variance of the dataset with 5 components.

Figure 6 and 7 depicts the evolution of the factor weights across maturities with two calculus methods. Indeed, in Figure 6, Factor 1 has the biggest weight for maturities up to 50 months. From 80 months of maturity, the factor 2 present the biggest weight among all of them.

Explanation of the Estimation Procedure

The estimation procedure for the yield curve is conducted in three steps:

Firstly, it is necessary to decompose the factors we have determined into two components: a predictable component and an innovation component. To achieve this, we perform a regression using Ordinary Least Squares (OLS) of the factors against their lagged values, following the equation (mu and phi are reported in Tables 3 and 4):

$$X_{t+1} = \mu + \Phi X_t + v_{t+1}$$

.

By hypothesis, the shocks v_{t+1} conditionally follow a Gaussian distribution. Therefore, after the regression, we tested the innovations with a Jarque-Bera test for all factors, and they indeed follow it (see Table 5).

We thus construct an estimator of the state variance-covariance matrix:

$$\hat{\Sigma} = \frac{\hat{V}\hat{V}'}{T}$$

The innovations are then stored in the matrix $\mathbf{V}$, representing the error term of this equation.

Secondly, we conduct a second regression, this time regressing excess returns against the lagged values of the factors and also against the factors we refer to as "innovation," which we previously stored in the matrix $\mathbf{V}$. The equation is as follows:

$$rx = \mathbf{a}'\iota_T + \beta'\hat{V} + \mathbf{c}X^- + E$$

where:

- ι_T represents a $T \times 1$ matrix of ones.
- X^- represents the matrix of lagged factors $[X_0, X_1, \ldots, X_{T-1}]$.
- E represents the error term.

We then collect the regressors $[\hat{\mathbf{a}}, \hat{\beta}, \hat{\mathbf{c}}]$, where $\hat{\beta}$ signifies the innovation term.

In a third step, and with the help of the following equations, we calculate the estimators for the risk parameters:

$$\bullet\ \hat{\lambda}_0 = (\hat{\beta}\hat{\beta}')^{-1}\hat{\beta}(\hat{a} + \frac{1}{2}(\hat{B}^* \text{vec}(\hat{\Sigma}) + \hat{\sigma}^2 \iota_N))$$

$$\bullet\ \hat{\lambda}_1 = (\hat{\beta}\hat{\beta}')^{-1}\hat{\beta}\hat{c}$$

$$\text{where : } \mathrm{B}^* = \begin{bmatrix} \text{vec}(\beta(1)\beta(1)') \\ \vdots \\ \text{vec}(\beta(N)\beta(N)') \end{bmatrix} \text{ N } \times K^2 \text{ matrix.}$$

Before implementing the recursive approach to generate the zero coupon yield curve, we conducted a regression analysis of the one-month T-bill against the pricing factors to obtain the short-rate parameters, denoted as δ_0 and δ_1. These two parameters assist us in obtaining recursive pricing parameters, which are then utilized in the generation of the yield curve.

Ultimately, we calculate the bond prices through the application of an affine model with estimated model parameters.

This model has the property that all the log bond yields are linear or affine in the state variable. We assume that the yield on a bond of any maturity is linear or affine in the state variable X_t. The form of the price function for an n-period bond is :

$$lnP_t^{(n)} = \ A_n + \ B_n' X_t + u_t^{(n)}$$

A_n and B_n are two coefficients that help establish the relationship between yields of different maturities and the state variables that influence the dynamics of the yield curve. They are derived from the recursive formulas with homoskedastic shocks.

$$A_n = \ A_{n-1} + \ B_{n-1}'(\mu - \lambda_0) + \frac{1}{2}(B_{n-1}'\Sigma B_{n-1} + \sigma^2) - \delta_0$$

$$B_n' = \ B_{n-1}'(\Phi - \lambda_1) - \delta\text{`}_1$$

$$(when\ t = 0,\ A_0 = 0,\ B_0' = 0;\ when\ t = 1,\ A_1 = -\delta_0\ and\ B_1 = -\delta_1.)$$

$$\beta'^{(n)} = \ B_n'$$

Empirical Results

Table 6 presents the time-series properties of both yield pricing errors and return pricing errors implied by the five-factor model. We calculated their averages, standard deviations, skewness, and kurtosis. The autocorrelation coefficients of order one and six are calculated to check the existence of serial correlation in yields and return pricing errors. Our empirical findings indicate a significant degree of serial correlation in yield fitting errors, while there is minimal to no autocorrelation observed in return pricing errors.

The three panels on the left side of Figure 1 present the time series of observed and fitted yields for the two-, five-, and ten-year Treasury notes. We also chose to plot the time series of term premiums across those maturities, and we will elaborate on how they are derived. Initially, we set our risk premias, λ_0 and λ_1, equal to 0 in the previous equations. This yields the risk neutral yield curve. Subsequently, we take the difference between the risk neutral yield curve and the fitted yield curve, resulting in the time series of term premiums. This series oscillates around zero, and its values tend to become more negative as both the maturity and the observation year increase. These visuals demonstrate that the five-factor specification provides a generally accurate fit to the observed yields. However, upon closer inspection, as depicted in Figure 1, variations become more noticeable with increasing maturity. To illustrate it, we can consider Figures 2 and 3 which depict the behaviour of the yields across the maturities. Indeed, we can observe that the average fitted yield differs from the observed one of 0.025 to 0.040 across the maturities. Moreover, fitted yields have in general a more significant unconditional volatility than the observed ones, as shows Figure 3. Despite adhering to the procedures outlined in the paper, we observe discrepancies that deviate from the expected perfect fit, as suggested by the paper's methodology.

Figure 4 presents a three-dimensional representation of the fitted yields. We depict these fitted yields based on their maturity and observation period. At the beginning of the observation period (considering the relative maturities of our bonds), we observe that the yields were higher than those seen today. It is also noteworthy that during our observation period, the yield curve experienced numerous inversions, indicating various economic cycles that unfolded during this time. This once again underscores the accuracy of our fitted yields, which align perfectly with the observation period, not just the maturity of the bonds.

Extensions

We obtained the monthly data for the 1y, 2y, 3y, 5y, and 10y UK zero coupon bond yields spanning from January 2000 to December 2023 via Bloomberg. Then, we used the Nelson-Siegel-Svensson methodology on the acquired yield to deduce the NSS parameters, thereby generating the set of yields for maturities n = 1,...,120 months. The process to determine NSS parameters involved three critical functions. *NelsonSiegelSvensson* is used to calculate yields based on model parameters. *NSSResidu* is used to measure model accuracy and manage outliers as we calculate the sum of squared residuals between predicted and observed yields. Then *NSSOptimize* is set to optimize the parameter by minimizing the residuals, using the built-in minimize function from Python's scipy package. The optimized parameters obtained from this process will then be integrated into our five-factor models for further analysis. In a subsequent phase, all five principal components were extracted from these yields. Then, we adopted the three-step estimation approach to estimate all parameters and generate the yield curve. The result is shown in Figure 5.

Appendix

Table 1: **Comparison of GSW yields over NSS yields for maturities up to thirty months**

	Mean Error	Variance	Min Error	Max Error	Mean Absolute Error
SVENY01	0.002202	0.000023	-0.011540	0.032812	0.003768
SVENY02	0.004056	0.000038	-0.015009	0.025929	0.005664
SVENY03	0.005816	0.000053	-0.017622	0.024707	0.007501
SVENY04	0.007462	0.000070	-0.018690	0.028588	0.009166
SVENY05	0.008971	0.000087	-0.018844	0.031481	0.010606
SVENY06	0.010331	0.000103	-0.018458	0.033399	0.011877
SVENY07	0.011538	0.000118	-0.017755	0.034578	0.012973
SVENY08	0.012595	0.000131	-0.016873	0.035206	0.013916
SVENY09	0.013511	0.000143	-0.015900	0.037259	0.014721
SVENY10	0.014293	0.000152	-0.015920	0.039012	0.015404
SVENY11	0.014954	0.000160	-0.015816	0.040427	0.015974
SVENY12	0.015504	0.000166	-0.015543	0.041536	0.016442
SVENY13	0.015954	0.000170	-0.015147	0.042366	0.016820
SVENY14	0.016313	0.000173	-0.014664	0.042944	0.017116
SVENY15	0.016592	0.000174	-0.014123	0.043294	0.017339
SVENY16	0.016799	0.000175	-0.013545	0.043439	0.017500
SVENY17	0.016941	0.000175	-0.012945	0.043400	0.017606
SVENY18	0.017026	0.000174	-0.012338	0.043196	0.017659
SVENY19	0.017061	0.000173	-0.011730	0.043101	0.017667
SVENY20	0.017051	0.000171	-0.011130	0.043296	0.017633
SVENY21	0.017001	0.000168	-0.010543	0.043396	0.017562
SVENY22	0.016916	0.000166	-0.009970	0.043411	0.017460
SVENY23	0.016800	0.000164	-0.009415	0.043351	0.017330
SVENY24	0.016656	0.000161	-0.008880	0.043225	0.017175
SVENY25	0.016489	0.000158	-0.008366	0.043040	0.016998
SVENY26	0.016301	0.000156	-0.007871	0.042806	0.016802
SVENY27	0.016094	0.000153	-0.007643	0.042526	0.016591
SVENY28	0.015872	0.000151	-0.007618	0.042207	0.016366
SVENY29	0.015635	0.000149	-0.007591	0.041853	0.016135
SVENY30	0.015387	0.000147	-0.007562	0.041468	0.015898

Table 2: **Variance explanation as a function of the number of principal components**

PC1	PC2	PC3	PC4	PC5
0.843202788284	0.9998983047	0.9999938131	0.9999999463	0.9999999965

Table 3: **Mu values**

PC1	PC2	PC3	PC4	PC5
-0.002675	-0.004496	-0.00281	-0.002497	0.002394

Table 4: **Phi values**

PC1	PC2	PC3	PC4	PC5
0.992728	-0.005171	-0.035938	-0.015666	0.004177
-0.004646	0.982815	0.027305	0.017392	-0.010973
0.010172	-0.041402	0.866463	-0.064606	-0.06003
0.005693	-0.027243	-0.110928	0.663697	-0.219919
-0.020344	-0.025581	-0.097565	-0.25753	0.555779

Table 5: **Jarque and Bera test over the residuals**

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Statistic	5.964	57.323	275.791	3767.026	6448.890
P-value	0.051	0.000	0.000	0.000	0.000

Table 6: **Return pricing errors vs. Yield pricing errors**

	12	24	36	60	84	110
mean_error	0.003	0.003	0.003	0.002	0.001	-0.000
std_error	0.005	0.013	0.023	0.042	0.052	0.048
skew_error	0.079	-0.228	-0.347	-0.449	-0.439	-0.315
kurtosis_error	2.240	3.363	4.267	5.444	5.947	5.940
autocorr_1_error	0.188	0.015	-0.026	-0.066	-0.090	-0.085
autocorr_6_error	0.153	-0.030	-0.052	-0.054	-0.042	-0.003

	12	24	36	60	84	110
mean_error	-0.002	-0.000	0.004	0.007	-0.012	-0.071
std_error	0.008	0.025	0.048	0.096	0.133	0.161
skew_error	-0.223	-0.258	-0.262	-0.182	0.156	0.560
kurtosis_error	1.413	1.828	1.483	0.724	0.172	-0.515
autocorr_1_error	0.865	0.874	0.886	0.907	0.924	0.955
autocorr_6_error	0.400	0.436	0.513	0.635	0.716	0.810

This table depicts characteristics of yield pricing errors and yield pricing errors accross maturities

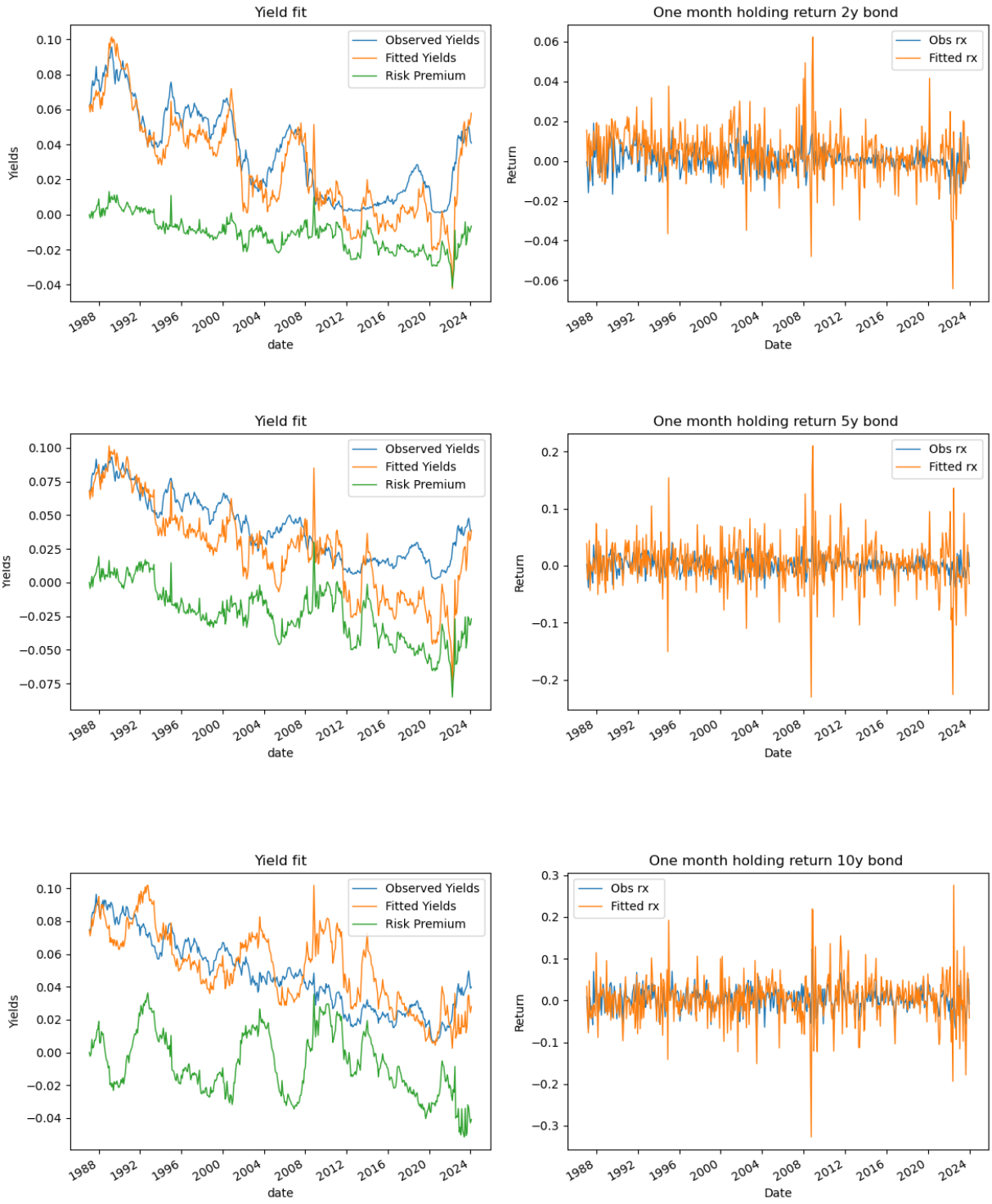

Figure 1: **Time series of observed and fitted yields and one month holding for the two-, five-, and ten-year bonds**

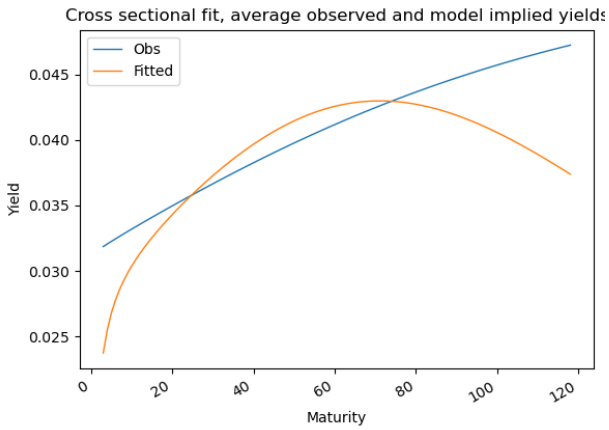

Figure 2: **Average yields**

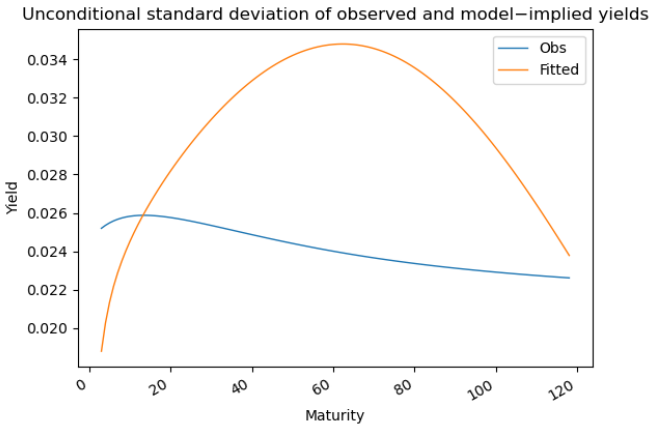

Figure 3: **Unconditional standard deviation of yields**

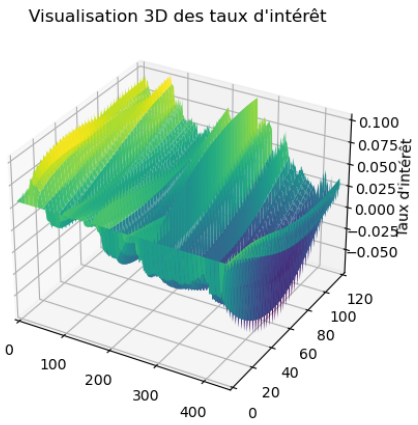

Figure 4: **3D View of Interest Rates**

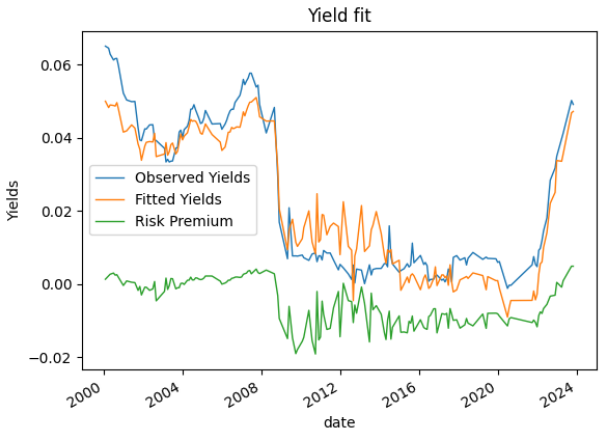

Figure 5: **UK Zero Coupon Yields 1Y**

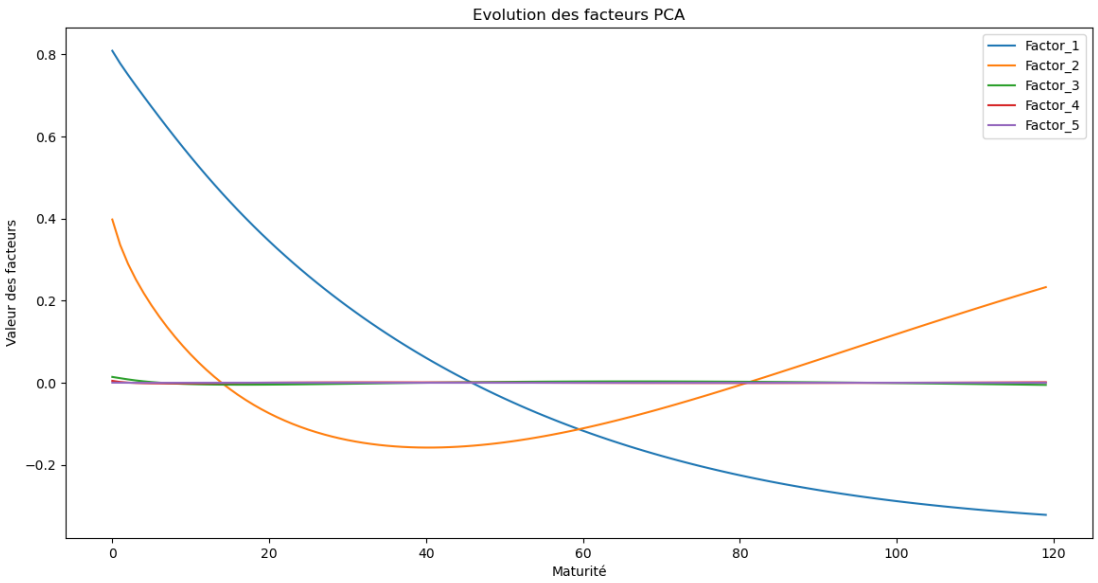

Figure 6: **Factors' evolution across maturities**

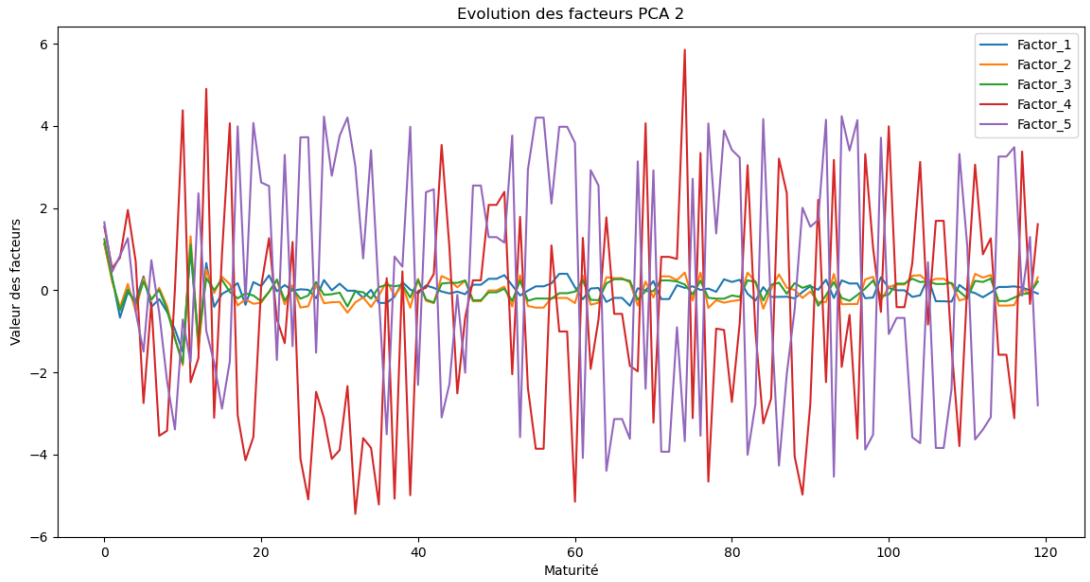

Figure 7: **Factor's evolution across maturities by centering and normalizing the product of the matrix of standardized returns by the eigenvectors of the covariance matrix of standardized returns**