CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.

DEFINING ACCURATE: A LOOK INTO THE NELSON-SIEGEL MODEL FOR YIELD CURVE FORECASTING

TESINA

QUE PARA OBTENER EL TÍTULO DE

LICENCIADO EN ECONOMÍA

PRESENTA

CARLOS ANDRÉS PÉREZ PÉREZ

DIRECTOR DE TESINA: DR. JUAN RAMÓN HERNÁNDEZ GONZÁLEZ

LECTOR DE TESINA: M.E. EDWIN SALIM TAPIA ALVARADO

CIUDAD DE MÉXICO 2024

Dedicatoria

Dedico este trabajo de investigación, el cual representa múltiples noches sin sueño y el conocimiento acumulado durante cuatro años, a las dos personas que lo hicieron posible. A mi madre, Marianela, quien me enseñó el verdadero valor del esfuerzo y la dedicación, dándole así una razón a todo lo que hago y quien me amó incondicionalmente hasta formar al hombre que soy hoy en día. A mi padre, Luis, quien hizo todo lo que estaba en sus manos y más para darme las herramientas y oportunidades que necesitaba para acercarme cada vez más a todos mis sueños y ambiciones.

Agradecimientos

Mi familia siempre ha sido un lugar seguro para mí, un pilar de estabilidad en el que puedo apoyarme. Por esto, quiero agradecer a mis padres, Luis Pérez Suárez y Marianela Pérez Mejía. Me han dado todo lo que un hijo podría pedir y, aunque intentaré, sé que nunca podré pagarselos por completo. Mis hermanos, Luis Fernando Pérez Pérez y Juan Ramon Pérez Pérez, quienes son los mejores amigos que mi madre me dio. Ellos son las personas que siempre estarán a mi lado, pase lo que pase, y a quienes siempre amaré. Mis primas, o más bien hermanas, Ana Isabella Rosete Pérez y Marianella Rosete Pérez, quienes comparten un amor y una conexión especial conmigo. Por último, tengo que agradecer a Bruno, mi compañero de cuatro patas, un ser tan cercano a mí que ha estado en mis mejores y peores momentos. Te amo y siempre lo haré.

Mi vida en el CIDE fue placentera gracias a las maravillosas personas que me hicieron crecer como estudiante, pero lo más importante, como persona. Necesito agradecer a Alejandro Lapuente, una de las personas más inteligentes e intuitivas que conozco, un verdadero amigo confiable y buen oyente. Andro Asatashvili, una persona verdaderamente culta y conocedora, siempre tenía algo interesante que decir y grandes consejos cuando uno está perdido. Emiliano Pérez, una persona llena de amor y curiosidad, enriqueciste la vida de una manera unica. Raúl Cepeda, mi compañero de cuarto, siempre recordaré todas las risas y conversaciones, un verdadero trabajador incansable y una de las personas más motivadas que conozco. Javier Lapuente, uno de los mejores oyentes que alguien pueda conocer, una persona graciosa y confiable. Valeria Ovalle, le agradezco por ser una amiga en tiempos difíciles y por la confianza que nos brindamos mutuamente. Eduardo Valencia, una persona extremadamente brillante, siempre fue un placer discutir mis ideas contigo. Mariano Álvarez, alguien verdaderamente amable cuya compañía siempre fue agradecida. Sergio Tovar, por todas las risas y los comentarios. Maia Esquer, por apoyarme en las últimas instancias de este proyecto y mi carrera.

Hay personas que siempre han estado a mi lado, lugares seguros a los que siempre he podido recurrir. Tengo que agradecer a Ana Laura por siempre escucharme; en mis peores momentos, acudí a ella y siempre estuvo allí. Ana Belén Rugarcía, nunca he conocido a alguien que crea tanto en mí como ella, incluso cuando yo no lo hacía. Raúl Egremy, la persona con la que más disfruto pasar el tiempo, un amigo leal y confiable. Alain Pellegrin, por todos esos viajes en la carretera y esas conversaciones profundas. Santiago Amezcua y Alfredo Martínez, mis amigos más antiguos, la definición de lealtad y amor. Gracias por siempre estar allí. Mi admiración recide en cada uno de ustedes.

Los profesores son la mejor guía hacia los sueños que quieres cumplir. Durante mi estancia en el CIDE, conocí a las mentes más capaces y brillantes que guiaron mi formación. Quiero agradecer a Fausto Hernández Trillo, quien me dio la primera oportunidad para trabajar con él y siempre estuvo dispuesto a ayudarnos y escucharnos, un profesor verdaderamente interesado en el futuro de sus alumnos y de la universidad. Edwin Salim Tapia Alvarado, quien me dio la oportunidad de trabajar con él, una persona curiosa y llena de bondad, que me apoyó a desarrollar mis habilidades profesionales y de quien nació el tema de esta tesis. Juan Ramón Hernández González, el director de esta tesis y la persona que me guió en la investigación y redacción de la misma, un profesor cuya capacidad y habilidades son admirables para cualquiera en este campo. A todos los profesores que me acompañaron en esta institución, les agradezco: Raciel, Itza, Sonia, Marcelo, Francisco, Raúl. Siempre los recordaré en mi futuro.

Abstract

This thesis examines the Nelson-Siegel model, with a focus on the Mexican and United States yield curves over a 13-year period. The study investigates the model's accuracy and adaptability in the face of exogenous shocks that produce sudden changes to the curve. Utilizing the dynamic Nelson-Siegel model and a cross-validation process, we optimize the λ parameter to enhance forecasting precision. The results demonstrate that while the Nelson-Siegel model is generally accurate, it faces challenges during periods of yield curve stress. Our findings indicate that short-term maturities are more accurately forecasted for the United States, whereas long-term maturities yield better forecasts for Mexico. This research contributes to the literature by identifying the model's limitations and proposing an improved forecasting methodology that adjusts to sudden yield curve changes. The implications of these findings are significant for central bankers and private investors, offering insights into better monetary policy and investment decisions.

Contents

List of Figures

List of Tables

1. Introduction

The significance of the yield curve is indisputable, despite occasionally being overlooked. The information in this instrument is crucial for any decision you make in today's world of credit. Suppose you want to get a credit card, a small loan, a car loan, or a mortgage, or invest your money, buying bonds, stocks, lending, or a savings account. The interest rate that you will pay if you borrow or gain if you lend will depend on the time to maturity (repayment), and thus, it will depend on the yield curve of said financial instrument. For most of your financial decisions, you will have to look at the yield curve, and this is true for governments too.

Notably, the yield curve of sovereign bonds is an important indicator of a country's financial health, and its shape can detail the market's expectations of the macroeconomic outlook. Additionally, the yield curve is a strong predictor of economic recessions, according to Benzoni et al. (2018) several factors that influence the rate premiums are summarized within the yield curve, and therefore inversions of the curve are related with expectations of economic downturn.

Knowing the yield curve's importance, forecasting its behavior is extremely attractive. Several methods exist for this; one is particularly interesting for the descriptiveness of its factors: the Nelson-Siegel model. Diebold and Li (2006) dynamic modification of the model and its recommendations allows us to understand the factors as descriptors of the curve; level, slope and curvature. Additionally, the model demonstrates the relevancy of the exponential decay factor which determines the point of maximum curvature.

This thesis follows the yield curve throughout 13 years of data in Mexico and the United States. The choice for these countries is based on the intersection of their economies and monetary policy, the recent economic downturns that they have faced and the differences on the health and risk of their economy. This enriches the analysis by illustrating differences in behaviour and expectations between developed and developing economies that go through similar economic cycles and crisis.

This research analyzes the effectiveness of the model by looking at the shapes that it creates, and its forecasting capabilities. Nelson and Siegel (1987) explain the effectiveness of the forecasts made by their models, comparing them to AR(1), VAR(1), random walks, and other models that forecast the yields directly. Diebold and Li (2006) modify the model so that its factors are dynamic and change over time. Additionally, they emphasize the exponential decay factor and recommend using the value 0.0609 for forecasting. While these two papers are some of the most significant advances in this field, the Nelson-Siegel model presents shortcomings while producing specific yield curve shapes. Diebold and Li's recommendation of an exponential decay factor of 0.0609 is not an accurate one-fits-all solution.

Therefore the research objective is to answer how well the Nelson-Siegel can fit the shapes of the yield curve throughout different points of stress and how can the forecasting capabilities of the model become adaptable to sudden changes on the yield curve's behaviour and periods of stress.

This work contributes to the existing literature on the yield curve by decomposing the shortcomings of the Nelson-Siegel model to fit yield curve and identifying times of stress. Moreover, it introduces a cross-validation method for forecasting the yield curve that gives better results than those from the Diebold and Li modifications according to the Diebold-Mariano statistic.

The structure is as follows: Chapter 2 discusses the relevancy and history of the yield curve, its importance, and various models for its forecasting. Chapter 3 covers the Nelson-Siegel model and the literature. Once we understand the yield curve's relevance and the Nelson-Siegel model, Chapter 4 examines the data structure used in this thesis. Chapter 5 describes the methodology that will yield results. In Chapter 6, discusses the analysis results and their meaning. Chapter 7 concludes the findings of this research.

2. The Yield Curve

The yield curve is a graph that shows the relationship between the interest rates (yields) of bonds and their different maturities. It illustrates how the payout structure changes over time and can be interpreted differently depending on the financial instrument. For sovereign bonds offers numerous practical applications. Private agents can leverage it to identify optimal investment opportunities and secure the best loan interest rates. Additionally, the yield curve reveals the economy's phase, indicating whether it is expanding or approaching a recession, according to Estrella and Mishkin (1996) the yield curve is a simple and reliable indicator for the probability of recession in the United States. Furthermore, as a primary policy tool, the sovereign yield curve illustrates debt construction strategies. Constructing a yield curve involves more than plotting yields against maturities; it entails creating a spectrum of financial instruments across various maturities. This approach grants the government control over the average life of its debts and provides lenders with diverse investment options.

Campbell (1995) explained the yield curve's importance while recounting how, in 1994, the United States Government took a high interest in the term structure. He used the yield curve to explain why the government had taken an approach to reduce the average maturity of the sovereign debt. This describes how manipulation of the maturity of financial instruments is a tool for controlling how a government issues debt. Haubrich et al. (1996) examined the forecast capabilities of the Yield Curve and its significance in predicting economic growth and recession, particularly finding that the yield spread derived from the yield curve was the best predictor four quarters ahead of time, additionally in analyzing the yield curve they discovered that the economy became close to the yield curve only after 1994.

The expectations hypothesis is rooted in the yield curve's ability to predict economic states.

According to Mishkin (1990) this theory posits that long-term interest rates are simply an average of the expected short-term interest rates in the future. Therefore, the yield curve provides insights into economic expectations. For instance, a low long-term interest rate is often associated with a recession, making an inverted curve a potential predictor of an economic downturn. Additionally, the Yield Curve reflects market participants' expectations of the central bank's actions, as monetary policy significantly influences yields of many maturities.

The sovereign yield curve is constructed using the rates for the zero-coupon bonds, which are debt instruments that only repay the face value plus the yield. The basic formula for calculating the price of a zero-coupon bond is $P = \frac{F}{(1+r)^n}$, where P is the bond's price, F is its face value, r is the zero-coupon rate, and n is the number of periods until maturity. The spot rates derived from this formula are the pure discount rates for various maturities, which are foundational in constructing the yield curve.

In addition to spot rates, the yield curve facilitates the derivation of forward rates, which reflect market expectations about future interest rates. The forward rate, calculated using the formula $f_{1,2} = \left(\frac{(1+r_2)^{t_2}}{(1+r_1)^{t_1}}\right)^{\frac{1}{t_2-t_1}} - 1$, where r_1 and r_2 are the spot rates for maturities t_1 and t_2, respectively, provides information about the conditions of loans or investments that will commence at a future date. The shape and slope of the yield curve influence forward rates and, consequently, investors' expectations about future interest rate movements. An upward-sloping yield curve suggests that short-term interest rates are expected to be lower than long-term rates, while a downward-sloping yield curve implies the opposite, reflecting expectations of an expanding or contracting economy respectively.

The term structure of interest rates is also used to infer information on other financial instruments that are not only related to sovereign debt and the overall macroeconomic conditions. Umar et al. (2021) used the three components of the yield curve (level, slope and curvature) as well as sectorial equity indices to find a connection and give additional interpretation to these factors. They conclude that the level component is a transmitter of return spillovers, and the curvature is a transmitter of volatility spillovers. While this is an specific use of the components of the yield curve, these have general importance as a macroeconomic and expectations indicators. These components can be calculated using the components from the Nelson-Siegel model or using the actual rates for different maturities and performing arithmetic operations. These operations and the difference with the Nelson-Siegel components will be discussed in the methodology section.

In Mexico, the yield curve has experienced a complex history. According to Meade Kuribreña et al. (2012), since the 1990's there was an intention to construct a yield curve with longer and more diverse maturities for government bonds. However, the 1994 crisis temporarily halted this project. Constructing the yield curve for Mexico involved extending the average life of Mexican debt and making the country more attractive to investors by offering a wider range of maturities and instruments. Most importantly, it creates a yield curve with fixed rates protected against risks that could increase the costs of adjustable bonds and debt instruments.

The yield curve's significance extends beyond the information it conveys to the public; it is also a vital tool for structuring government debt and ensuring the proper functioning of public and private debt markets. For example, Meade Kuribreña et al. (2012) explains that in 2012, the yield curve reflected market participants' positive sentiment toward Mexico's debt and their growth expectations.

This thesis studies the Mexican and United States yield curves. While the financial markets of both countries are vastly different, they experience similar macroeconomic shocks at specific points in time. Analyzing both curves allows us to observe the differences in yield curve movement between two countries with distinct market risks and monetary policies. The sovereign bonds that construct the yield curves for both countries have identical maturities, except for the 2-year maturity bond, which only appears in the United States. Thus, the main differences are the bond yields and the premiums they comprise. Additionally, the United States developed the instruments that construct their yield curve earlier than Mexico.

The yield curve serves as a versatile information instrument. It is highly correlated with a country's macroeconomic conditions, making it an effective information transmitter. Developed countries require a proper yield curve with various maturities, and market participants must examine it for insights into investments and credit structuring.

2.1. Calibrating the Model

This thesis describes the relevance and multiple uses of the Nelson-Siegel Model. This model and its components can forecast and explain the yield curve dynamics. Within this research, we will look at how the Nelson-Siegel dynamic factors model has behaved in the last 13 years for Mexico and the United States; our objective is to identify any possible shortcomings of the model and suggest some guidelines on the appropriate way to calibrate to have the most accurate forecast possible.

3. The Nelson-Siegel Model

The forward rate, $f(t, \tau)$, is the expected future interest rate for a particular period starting at time t and lasting for τ periods. According to the expectations hypothesis, the long-term interest rates can be viewed as the geometric average of the expected future short-term rates. This hypothesis underpins much of the theoretical framework for yield curve modeling.

Nelson and Siegel created a model to capture the yield curve's movements using a small number of parameters by modeling the instantaneous forward rate as a function of time to maturity (τ). The forward rate function is:

$$f(t, \tau) = \beta_0 + \beta_1 exp(-\lambda\tau) + \beta_2 \tau exp(-\lambda\tau) \tag{3.1}$$

Here, β_0 represents the long-term level of interest rates, β_1 captures the short- to medium-term slope and β_2 reflects the curvature of the yield curve. The decay parameter λ controls the rate at which the influence of β_1 and β_2 diminishes with maturity.

The yield to maturity, $y(t, \tau)$, which is the average of the forward rates up to maturity τ, can be derived by integrating the forward rate function:

$$y(t, \tau) = \frac{1}{\tau}\int_0^\tau f(t, u)du = \beta_0 + \beta_1(\frac{1 - exp(-\lambda\tau)}{\lambda\tau}) + \beta_2(\frac{1 - exp(-\lambda\tau)}{\lambda\tau} - exp(-\lambda\tau)) + \epsilon_t(\tau) \tag{3.2}$$

This model reduces the complexity of the yield curve to a few interpretable parameters, providing a robust tool for both fitting historical yield curve data and forecasting future interest rates. The parameters β_0, β_1, and β_2 can be interpreted as follows: β_0: The long-term level factor influencing the overall height of the yield curve. β_1: The slope factor affecting the difference between short-term and long-term interest rates. β_2: The curvature factor, which measures the extent to which the yield curve bends or arches. Building on the Nelson-Siegel framework, Diebold and Li (2006) introduced a dynamic version of the model to better capture the time-varying nature of the yield curve parameters. They proposed treating β_0, β_1, and β_2 as time-varying factors, modeled using auto-regressive processes. This dynamic factor model allows the parameters to evolve in time, more accurately representing the changing economic environment.

The state-space representation of the Diebold and Li model is as follows:

$$y_t(\tau) = \beta_{0,t} + \beta_{1,t}(\frac{1 - exp(-\lambda\tau)}{\lambda\tau}) + \beta_{2,t}(\frac{1 - exp(-\lambda\tau)}{\lambda\tau} - exp(-\lambda\tau)) + \epsilon_t(\tau) \tag{3.3}$$

$$\beta_{0t} = \alpha_0 + \phi_0\beta_{0,t-1} + \eta_{0t}$$

$$\beta_{1t} = \alpha_1 + \phi_1\beta_{1,t-1} + \eta_{1t} \tag{3.4}$$

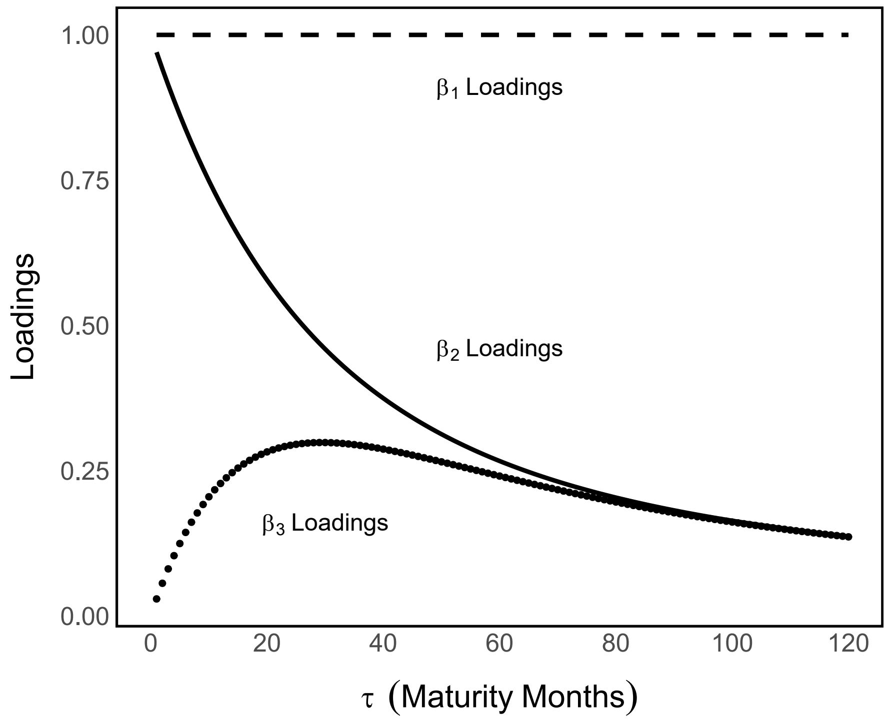
$$\beta_{2t} = \alpha_2 + \phi_2\beta_{2,t-1} + \eta_{2t}$$

Where $\epsilon_t(\tau)$ represents measurement errors, and η_{it} are white noise processes capturing the innovations in the factors. This formulation allows the factors to follow a first-order auto-regressive process (AR(1)), making it possible to forecast the future values of the yield curve.

A critical component of the Nelson-Siegel and Diebold-Li models is the decay parameter λ. The choice of λ affects how quickly the influence of the slope and curvature factors decays with maturity. Diebold and Li fixed λ at 0.0609, a value that provided a good fit for the United States yield curve data. Figure 3.1 represents how the beta loadings behave given this λ value. If we focus on the β_3 loading we can see that it reaches maximum curvature on the 30 months mark, this goes by what Diebold and Li (2006) propose. However it seems logical that not all actual yield curves will behave this way.

Diebold and Li used a two-step procedure to estimate the parameters. First, they estimated the parameters β_{0t}, β_{1t}, and β_{2t} at each point in time using non-linear least squares. Second, they

Figure 3.1: Nelson-Siegel loadings with $\lambda = 0.0609$.



Source: Own elaboration based on Diebold and Li (2006).

modeled these estimated parameters as auto-regressive processes and estimated the parameters of these processes using AR(1).

The dynamic Nelson-Siegel model enables forecasting future yield curves by predicting the future values of the latent factors β_{0t}, β_{1t}, and β_{2t}. This is achieved by applying the estimated autoregressive models to project the factors forward. The forecasted yield curve at time $t + h$ is then constructed using the predicted values of the factors:

$$\hat{y}_{t+h}(\tau) = \hat{\beta}_{0,t+h} + \hat{\beta}_{1,t+h}\left(\frac{1 - exp(-\lambda\tau)}{\lambda\tau}\right) + \hat{\beta}_{2,t+h}\left(\frac{1 - exp(-\lambda\tau)}{\lambda\tau} - exp(-\lambda\tau)\right) + \epsilon_t(\tau) \quad (3.5)$$

Diebold and Li demonstrated that their dynamic Nelson-Siegel model provides a more accurate long-term forecasts than traditional benchmarks such as random walk and VAR models. The ability to accurately forecast the yield curve has significant implications for bond portfolio management, monetary policy analysis, and financial risk management.

3.1. The Literature

The yield curve is essential for understanding macroeconomic outlooks, debt dynamics, and investment strategies. Forecasting the yield curve enables us to anticipate future economic conditions and make informed decisions. Over the years, various models for yield curve forecasting have emerged. This section explores these diverse models, highlighting their mechanisms and implications for accurate yield curve prediction. Moreover, we review the applications and modifications of the Nelson-Siegel model throughout the literature.

Initially, the field was impacted by the innovative work of Nelson and Siegel (1987), who introduced a parsimonious approach to modeling the yield curve. They aimed to create a model that could capture the different shapes the yield curve can take—monotonic, humped, and S-shaped—summarizing it in just a few parameters. Their empirical results demonstrated that the model could explain 96% of the variation in U.S. Treasury Bill yields across different maturities during the period 1981-1983. Additionally, the model had high predictive power when forecasting long-term Treasury bonds. Based on its results, the Nelson-Siegel model has been adopted in academic research and practical financial analysis.

Campbell (1995) examined the expectations hypothesis and its relation to bond yields using the Nelson-Siegel model. They tested the hypothesis that long-term interest rates reflect market expectations of future short-term rates. By fitting the model to historical yield data, they assessed how bond yields incorporated expectations of future economic conditions. The study found that the Nelson-Siegel model effectively captured the expectations embedded in bond yields, supporting the hypothesis and demonstrating the model's practical value for financial forecasting and policy analysis.

Kim and Wright (2005) utilized a three-factor term structure model to estimate the decomposition of U.S. Treasury yields, expanding on the work of Duffie and Kan (1996) and Duffee (2002). Their model included no-arbitrage conditions to ensure correct bond pricing across different maturities and incorporated inflation data to better understand the factors driving interest rates. Applied to U.S. Treasury yields since 1990, the model indicated that a significant portion of the decline in long-term yields and distant-horizon forward rates since mid-2004 could be attributed to a decrease in term premiums, with approximately two-thirds of the decline in nominal-term premiums due to a reduction in real-term premiums.

Building on the Nelson-Siegel model, Diebold and Li (2006) introduced a dynamic factor modification that innovates the Nelson-Siegel exponential components framework. They focused on parsimoniously representing the yield curve through three parameters: level, slope, and curvature. By transforming the yield curve into three time-varying factors and forecasting using an AR(1) process, they found that at the one-year horizon, the model was quite effective in forecasting the structure of the curve, outperforming traditional VAR(1), AR(1), random walks models and other methodologies used at the time.

Elizondo (2017) proposed forecasting the yield curve in Mexico using an affine model that integrates exogenous factors derived through principal component analysis for forecasting interest rates over one to sixty months. This model employs an affine framework aligning with the no-arbitrage condition. Using state variables following a VAR(1) process, Elizondo established a linear relationship essential for forecasting across different maturities. Her methodology showed better performance than traditional models such as random walks, AR(1), and VAR(1) models.

Haubrich et al. (1996) explored the yield curve's predictive power for real economic growth using the Nelson-Siegel model. They analyzed the relationship between the yield curve's components and future economic activity, assessing the model's performance in forecasting real Gross Domestic Product (GDP) growth. The results indicated that the Nelson-Siegel model strongly predicted economic growth, outperforming naive forecasts and other traditional indicators.

Koopman et al. (2010) extended the traditional Nelson-Siegel framework by introducing time-varying parameters, allowing the level, slope, and curvature factors to evolve. This modification provided a more flexible and accurate representation of interest rate movements. The study highlighted the Dynamic Nelson-Siegel model's superior performance in forecasting interest rates, particularly in capturing the effects of macroeconomic shocks and policy changes on the yield curve.

Christensen et al. (2011) developed an affine arbitrage-free version of the Nelson-Siegel model, ensuring adherence to the no-arbitrage condition for a more consistent and theoretically sound representation of the term structure. This model provided a superior fit to historical yield data compared to the traditional model, leading to greater consistency with financial theory and improved predictive accuracy.

Coroneo et al. (2016) explored the presence of macroeconomic factors in the yield curve that are not captured by traditional components. By integrating macroeconomic data into the Nelson-Siegel framework, they enhanced the model's explanatory power and forecasting accuracy. The study revealed that inflation and output growth significantly influenced the yield curve beyond the traditional components, improving the model's fit to historical data and predictive performance.

Kim et al. (2020) combined the Nelson-Siegel model with machine learning techniques to predict the dynamics of credit default swap (CDS) spreads. Leveraging the parsimonious nature of the Nelson-Siegel model, they captured the term structure of CDS spreads through its level, slope, and curvature components. This hybrid approach incorporated non-linear relationships and complex patterns in the data, significantly improving the accuracy of CDS spread forecasts.

Lastly, Umar et al. (2022) utilized the Nelson-Siegel model to investigate the interplay between sovereign yield curves and oil price shocks. By decomposing the yield curve into its fundamental components, they analyzed how oil price shocks influenced the yield curve's level, slope, and curvature across different maturities. The study found significant spillover effects, particularly on the level component, highlighting a direct link between commodity prices and long-term economic expectations.

4. Data

For this analysis we will use daily yields for sovereign bonds in the United States and Mexico. The data for the daily par treasury yields of the United States was gathered from the Treasury Department and contains a total of 3254 observations, this accounts for all working days from January 4th, 2010, up until January 3rd, 2023, and contains information for the 1, 3 and 6 months maturities and the 1, 2, 3, 5, 7, 10, 20, and 30-year maturities for a total of 11 maturities. In the Mexican case the data comes from PiP Latam a valuation and pricing corporation. It contains 679 observations that account for the weekly outputs of this information and the data goes from January 5th 2010 up until January 3rd 2023 and contains information on the 1, 3 and 6 months maturities and the 1, 3, 5, 7, 10, 20, and 30-year maturities for a total of 10 maturities.

There is an initial discrepancy of maturities, with only ten different ones for Mexico and 11 for the United States; we solve this by interpolating the 2-year maturity using spline. We also do this for the United States for two reasons: first, we want to apply the same methodology for both countries to ensure the validity of out results, and second, since we have information on the two-year maturity for the United States, we can check the accuracy of our spline interpolation

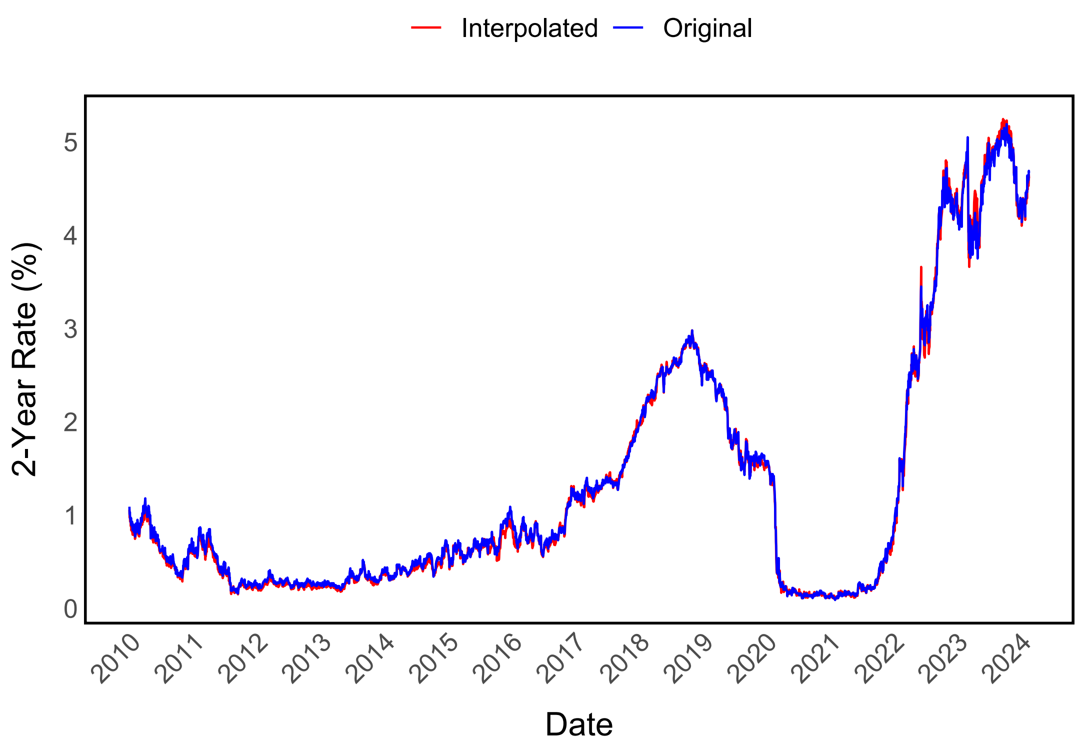
Spline interpolation is a mathematical technique that constructs a smooth curve that passes through a given set of points. It is beneficial for estimating intermediate values within the range of the data points. The general equation for spline interpolation involves fitting a piece wise polynomial function to the data points. For a cubic spline, the interpolation function $S(x)$ is defined as:

$$S(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 \tag{4.1}$$

For each interval$[x_i, x_{i+1}]$, where $a_i, b_i, c_i,$ and d_i are the coefficients determined by the interpolation process. These coefficients are calculated so that the spline is continuous and smooth at the data points, ensuring that the first and second derivatives of the spline are also continuous. Appendix A.5 shows the applied spline methodology.

The following graph we compare the 2-year maturity yield over time for the United States, alongside its interpolated counterpart.

Figure 4.1: 2-year Spline interpolation for The United States over time.



Source: Own elaboration.

Figure 4.1 shows that the spline interpolation closely matches the actual data, ensuring accurate results. We will now move to descriptive statistics and facts about the data.

Table 4.1: Descriptive Statistics for the United States Yields

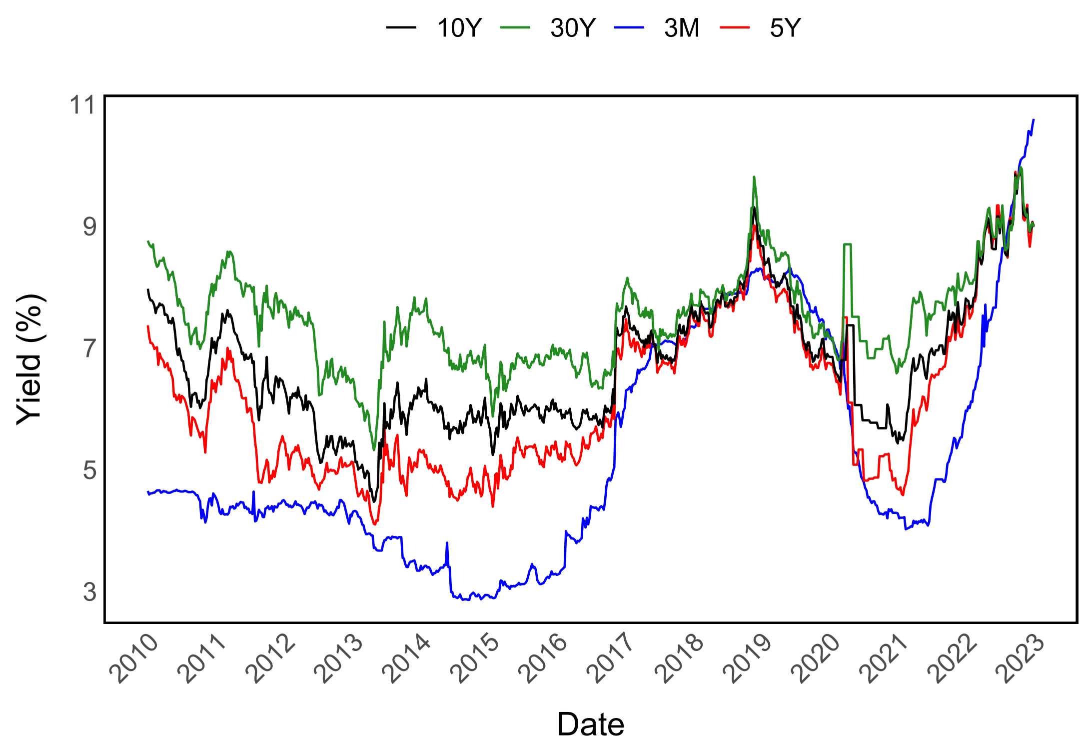
Statistic	1M	3M	6M	1Y	2Y	3Y	5Y	7Y	10Y	20Y	30Y
Mean	0.5740	0.6347	0.7286	0.8238	1.0001	1.2129	1.6065	1.9521	2.2561	2.7573	2.9702
SD	0.8760	0.9408	1.0000	1.0162	0.9676	0.9016	0.7900	0.7426	0.7326	0.7768	0.7671
Min	0.0000	0.0000	0.0200	0.0400	0.0996	0.1000	0.1900	0.3600	0.5200	0.8700	0.9900
Max	4.1700	4.5300	4.7800	4.8000	4.8023	4.6600	4.4500	4.3600	4.2500	4.6900	4.8500
Median	0.1000	0.1200	0.1800	0.2750	0.6245	0.9700	1.5700	1.9800	2.2400	2.7100	2.9800

Table 4.1 presents the descriptive statistics of United States yields over the past 13 years. The data reveal significant fluctuations in yields, with changes exceeding four percentage points across most maturities, indicating multiple economic cycles. Despite this volatility, the yields remain relatively low, with means ranging from 0.57% to 2.97%.

Table 4.2: Descriptive Statistics for the Mexican Rates

Statistic	1M	3M	6M	1Y	2Y	3Y	5Y	7Y	10Y	20Y	30Y
Mean	5.1107	5.2835	5.4282	5.5694	5.7172	5.8402	6.2054	6.4328	6.7275	7.3224	7.4950
SD	1.8020	1.8613	1.9046	1.9195	1.7805	1.5596	1.3185	1.1967	1.0608	0.8594	0.8069
Min	2.2900	2.8600	2.9600	2.9900	3.2216	3.5500	4.1000	4.3000	4.4700	5.0600	5.3200
Max	10.4900	10.7600	10.9000	10.9700	10.7875	10.2100	9.8900	9.8600	9.8700	9.9800	9.9700
Median	4.3500	4.4700	4.6000	4.7500	5.0768	5.3700	5.8600	6.2000	6.5500	7.2400	7.5100
IQR	3.0300	3.0300	3.0450	2.9850	2.7428	2.3800	2.0550	1.8100	1.5500	1.2100	1.1000

Figure 4.2: Mexican sovereign rates over time.

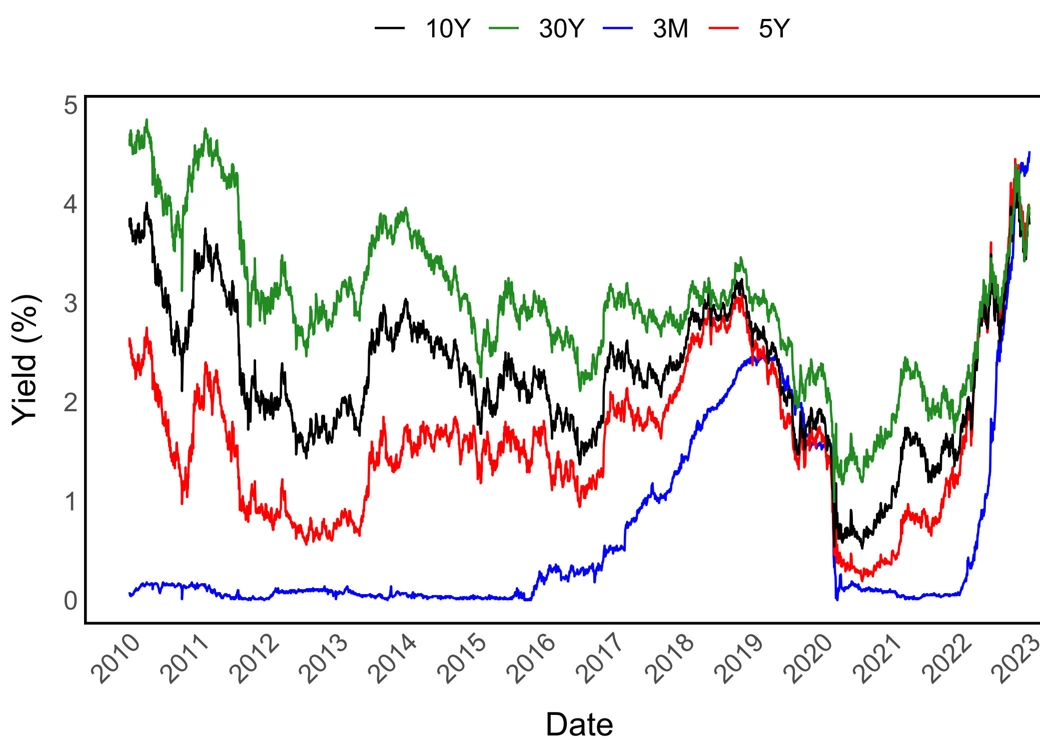


Source: Own elaboration.

Table 4.2 presents the descriptive statistics of Mexican yields over the past 13 years. The fluctuations are more pronounced than in the United States, with changes of up to 7 percentage points. The yields range from 2.9% to 10.48%. Notably, the maximum yields are higher at shorter maturities, consistent with the 2022 monetary policy and the inverted yield curve during that period.

In Figure 4.2, we observe the yield curve movement over time for Mexico. The 3-month rate represents the short term, the 5-year yield reflects the medium term, and the 30-year rate indicates the long term. Additionally, the 10-year yield helps illustrate the inversion dynamic. The graph shows that the most significant movements occur with the 3-month rate. Yield curve inversions are primarily driven by sharp upward movements in the 3-month yield, rather than declines in the 10-year yield, which aligns with longer-term maturities.

Figure 4.3: United States Treasury rates over time.



Source: Own elaboration.

Figure 4.3 illustrates the yield curve over time for the United States. The 3-month yield shows minimal movement, reflecting the Federal Reserve's monetary policy and its goal of maintaining a 0% interest rate. However, from 2016 to 2020, significant movements occur, paralleling Mexico's pattern and indicating yield curve stress. The 30-year yield exhibits a downward trend until 2022, when rates begin to rise due to inflation. Similar to Mexico, the United States' medium and long-term yields, including the 10-year yield, follow the same dynamic, however the 3 month-yield follow its own dynamic. The moments of inversion are less clear and less frequent compared to the Mexican case.

The yield spread is defined as the difference between the 10-year yield and the 3-month yield. When this difference is negative, the yield curve inverts. Therefore, it is relevant to examine the dynamics and processes leading to these inversions.

Given the extensive and rich dataset, we can undertake a thorough analysis of the yield curve's behavior across various economic cycles, thereby identifying and examining its different configurations over time and its responses to external shocks. The use of spline interpolation ensures data completeness, which is instrumental for applying the methodologies delineated in the subsequent section.

5. Methodology

5.1. Nelson-Siegel

We employ the dynamic Nelson-Siegel model to fit the yield curve data for both the U.S. and Mexico. The specific formulation of the model is as follows:

$$y_t(\tau) = \beta_{0,t} + \beta_{1,t}(\frac{1 - exp(-\lambda\tau)}{\lambda\tau}) + \beta_{2,t}(\frac{1 - exp(-\lambda\tau)}{\lambda\tau} - exp(-\lambda\tau)) \tag{5.1}$$

Where $y(t)$ represents the yield at maturity t, and β_0, β_1, and β_2 are the level, slope, and curvature factors, respectively. The parameter λ governs the exponential decay rate and is crucial for capturing the different maturity segments of the yield curve. Calculation of the parameters and betas follows a linear regression with the following form:

$$rate \sim 1 + \frac{1 - exp(-\lambda\tau)}{\lambda\tau} + (\frac{1 - exp(-\lambda\tau)}{\lambda\tau} - exp(-\lambda\tau)) \tag{5.2}$$

The factors are calculated for each maturity, and subsequently the data frame containing the maturities and factor values is processed through linear regression. The coefficients resulting from this process are the Nelson-Siegel parameters for that given point in time.

However, the λ value necessary for the Nelson-Siegel model is ad-hoc to the maturity in which the curvature or the β_2 factor reaches its peak. While Diebold and Li (2006) recommend using 0.0609 which corresponds to 30 months, this value is not optimal for fitting all datasets. We estimate the parameters by running a loop that tests 2,500 different λ values, ranging from 0.001 to 25 in increments of 0.01. We choose the parameters and λ that minimize the sum of squared residuals (SSR), defined by the following equation:

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{5.3}$$

Where y_i are the observed values of the rates, while $\hat{y}_i$ are the rates calculated from the Nelson-Siegel model parameters, the λ that minimizes this value compared to other λ values ensures the best fit. This allows us to analyze the model's limitations and identify the yield curve shapes it cannot accurately represent.

Once the parameters and λ values are obtained, we use equation 5.1 to calculate the fitted yields and compare them to the actual values. To evaluate the goodness of fit of the Nelson-Siegel model, we utilize the R^2 metric, which quantifies the proportion of variance in the yield data explained by the model. The R^2 value is calculated using the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2} \tag{5.4}$$

where y_i represents the actual yields, $\hat{y}_i$ represents the fitted yields, and $\bar{y}$ is the mean of the actual yields. This metric provides insight into the model's explanatory power and adequacy in capturing the underlying yield curve structure.

Additionally to understand the fitted values from the R^2 we provide several comparison of the parameters of the yield curve. We know that β_0, β_1, and β_2 are the level, slope, and curvature, but for the actual yields, we calculate these factors as follows:

$$\begin{gathered} level = 30 yearyield \\ slope = 10 yearyield - 3 monthyield \\ curvature = 2 \times 2 yearyield - (3 monthyield + 10 yearyield) \end{gathered} \tag{5.5}$$

While calculating the factors using equations 5.5 provides an approximation of the yield curve's dynamics, this approach fails to consider the dynamics across all maturities, as the equations only include certain maturities. The β factors of the Nelson-Siegel model ensure that all maturities forming the curve are considered, offering a more comprehensive economic interpretation, as supported by the literature.

This approach allows us to assess how well the Nelson-Siegel model approximates the yields and how accurately the parameters describe the characteristics of the yield curve over time. This forms the basis of our analysis, providing insights into the precision with which the dynamic Nelson-Siegel model can recreate the yield curve.

5.2. Time Series Cross-Validation

To evaluate the Nelson-Siegel forecasting abilities, we propose a different method for selecting a λ value appropriate for forecasting the yield curve at different horizons. As we have discussed, the traditional forecasting process for the yield curve follows the next steps:

1. Transform the yield data to the Nelson-Siegel parameters using linear regression
2. Forecast the β parameters using an AR(1) process
3. Transform the parameters to yields using $\lambda = 0.0609$

Once we have obtained the parameters and forecast the beta values, it is crucial to make a deliberate decision on which λ to use, as yield curves should vary in curvature and shape depending on economic expectations.

We propose using time series cross-validation to identify a suitable λ value. This methodology allows for precise parameter estimation by testing the forecasting process in the month preceding the actual forecast period. The reason behind using this methodology is that it provides a simple parameter estimation process that is based on previous data. The factors of the yield curve are sequential, meaning that the current curve is similar to the previous one and therefore the λ value is also similar. Cross-validation selects the best fitting λ in time $t - 1$ which should prove a more accurate value for time t.

First, we divide the dataset into three different subsets: the training set, the testing set, and the evaluation set. The training set spans one year, the testing set spans one month, and the evaluation set varies according to the forecast horizon (1, 6, or 12 months).

Additionally, we evaluate this methodology over time using a rolling window approach, moving one month at a time from the start to the end of the data. After creating the time sets, the cross-validation process follows these steps:

1. Transform the actual yield data into Nelson-Siegel parameters as done previously.

2. Use the training data (previously calculated parameters) to forecast the test set using an AR(1) process.

3. Iterate through λ values to find the one that minimizes the root mean squared error (RMSE) and store that λ value.

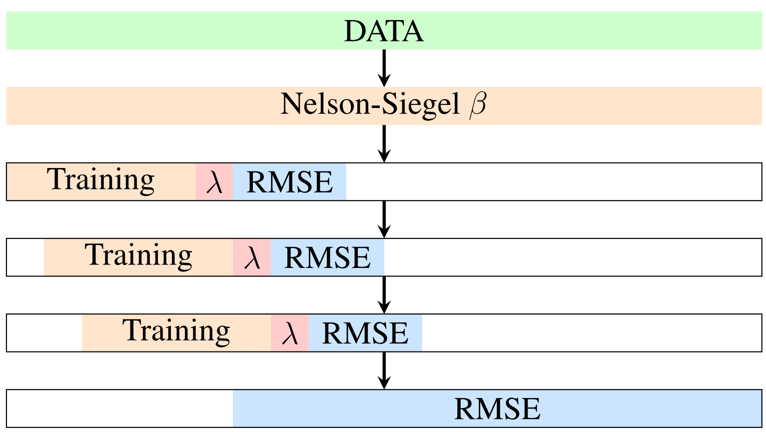
4. Extend the training set β's to the size of the training set plus the test set.

5. Forecast the evaluation set (1, 6, or 12 months) using an AR(1) process, then calculate the yields using the stored λ.

6. Calculate the RMSE to assess the effectiveness of the methodology.

Figure 5.1 graphically illustrates the cross-validation process, showing that the final result is a time series of the RMSE for the forecast.

Figure 5.1: Cross-validation process.



Source: Own elaboration.

The AR(1) model's specification, comes from Diebold and Li (2006). The AR(1) model is independently applied to the three components ($\beta_0, \beta_1, \beta_2$), forecasting one year ahead. The AR(1) process is defined as:

$$\beta_{t+1,i} = \alpha_i + \phi_i \beta_{t,i} + \epsilon_{t,i} \tag{5.6}$$

where α is the intercept, ϕ is the auto-regressive parameter, and ϵ_t is the error term. This approach employs a rolling window methodology, where the initial training period comprises three years of historical data. Subsequently, the training window is extended annually, and forecasts are generated iteratively. This technique ensures the model adapts to new information, providing updated forecasts each year.

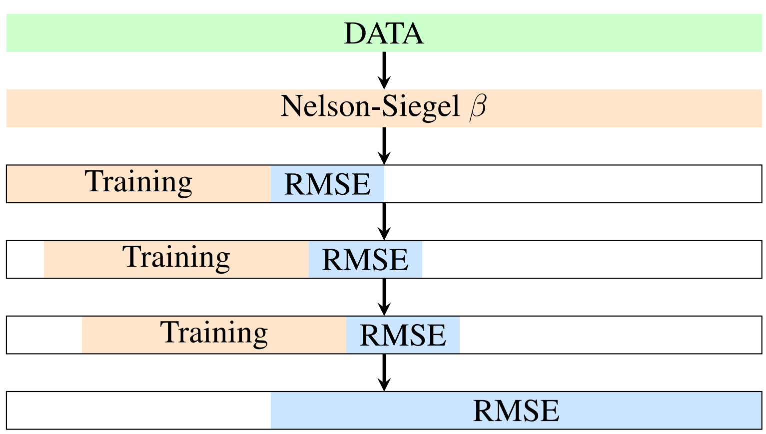
To assess the accuracy of the forecasts, we calculate the Root Mean Squared Error (RMSE) for each model and each forecast observation, segmented by maturity. The RMSE metric is calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{5.7}$$

where $\hat{y}_i$ are the forecast yields and y_i are the actual yields. This metric provides a measure of the average deviation between the forecast and actual yields, allowing us to evaluate the precision of the forecasts across different maturities and periods.

To compare with λ set to 0.0609, we follow a similar rolling window forecast process to extract the RMSE of this specification over time. Figure **??** shows the forecast process using fixed λ; we omit the testing process and instead directly forecast and evaluate the RMSE, resulting in another RMSE time series that we will use tom compare accuracy.

Figure 5.2: Cross-validation process using fixed λ.



Source: Own elaboration.

5.3. Diebold-Mariano test

Finally, we apply the Diebold-Mariano test to compare the forecast accuracy between models using the estimated λ and the fixed λ. The Diebold-Mariano test, initially proposed by Diebold and Mariano (2002), evaluates whether the differences in forecast accuracy between two competing models are statistically significant. The test statistic also provides information about the accuracy of the forecast model. This thesis uses the Diebold-Mariano test to determine which model specification is more accurate and which maturities are forecasted more precisely. Since this thesis compares two model specifications, the Diebold-Mariano test is appropriate for comparing their forecasting performances. The Diebold-Mariano test statistic is calculated as follows:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} \tag{5.8}$$

Where $\overline{d}$ is the mean loss differential, $\hat{f}_d(0)$ is the spectral density of the loss differential at frequency zero, and T is the sample size. This test provides a robust framework for assessing the relative performance of the yield forecasting methodologies employed in this study.

The following steps summarize the methodology employed in this study: First, we apply the Nelson-Siegel model to fit the yield curve data for the U.S. and Mexico, capturing the level, slope, and curvature parameters. Next, we evaluate the goodness of fit by calculating the R^2 values.

Subsequently, we use an AR(1) model to forecast the Nelson-Siegel parameters one year ahead, employing a rolling window methodology. We determine the optimal λ values for yield calculations using the Newton-Raphson method and a fixed value of 0.0609. We then assess forecast accuracy by calculating RMSE for each model and observation. Finally, we compare forecast performance using the Diebold-Mariano test. We do this for 1, 6 and 12 month ahead forecasts.

6. Results

Following the previous methodology, we first examine the fitted values of the Nelson-Siegel Model. We utilized a variation of the Nelson-Siegel function from the YieldCurve library in R Studio.[1] The function computes for larger λ values, calculating factors using ordinary least squares while iterating through different λ values to minimize the residual sum of squares. Table 6.1 presents the descriptive statistics of λ confirming our proposition that λ values vary from 0.0609, tending to be higher and closer to one. The mean and maximum values indicate that the curvature through the data peaks within the first 30 months. It is relevant to note that the table shows a maximum value of 21.51939. Indeed, λ can take values greater than one. Large values of λ cause the exponential terms to decay more rapidly. This means the influence of β_1 and β_2 in the yield curve diminishes more quickly as maturity increases. We allowed for these large values to ensure more accurate results.

Table 6.1: Descriptive Statistics for λ Values for Mexico and the United States

Statistic	Mean	Median	Variance	SD	Min	Max
λ (US)	0.7121973	0.5004548	2.206892	1.485561	0.06061652	21.51939
λ (MX)	0.9939696	0.352779	7.607143	2.758105	0.06061652	21.51939

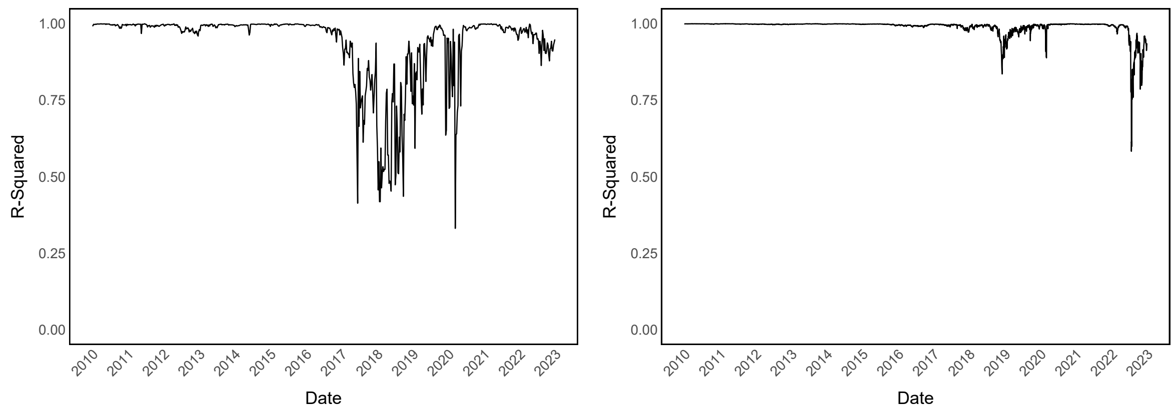
Once we calculate the factors, we use the dynamic Nelson-Siegel formula to derive the yields from the factors and compute the R-squared between the actual and calculated yields. Table 6.2 presents the descriptive statistics of the R^2. The minimum value highlights a shortcoming of the Nelson-Siegel Model in fitting certain yield curves, which becomes evident in the time plot of the R^2 values.

[1] Credit for the YieldCurve package and functions goes to Segio Salvino Guirreri at https://github.com/cran/YieldCurve.git

Table 6.2: Descriptive Statistics for R-squared Values from the Nelson-Siegel Model for the US and Mexico

	Mean	Median	Variance	SD	Min	Max
R^2 (US)	0.9900744	0.9979249	0.0007421216	0.02724191	0.5826492	0.9999346
R^2 (MX)	0.936052	0.9895473	0.01446142	0.1202556	0.3306261	0.9998465

Figure 6.1: Comparison of Mexican and United States R squared over time.



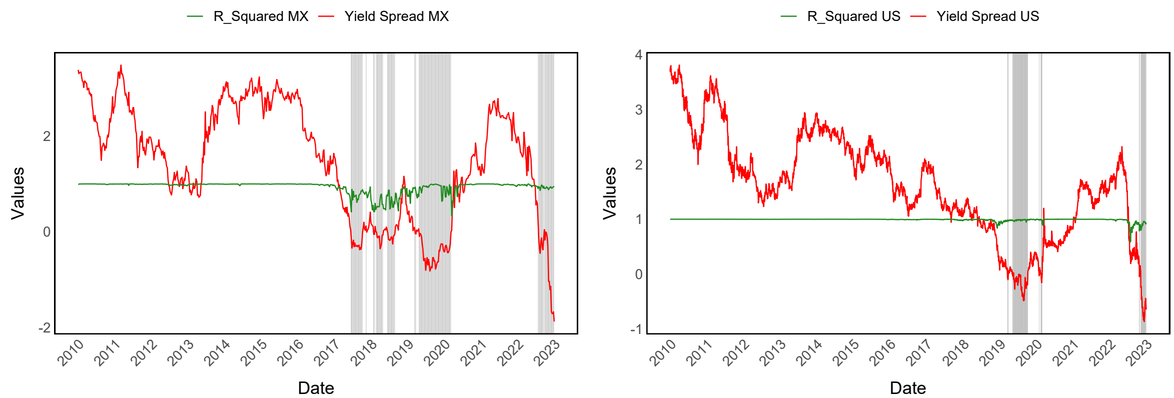
(a) Mexico **(b)** United States

Source: Own elaboration.

Figure 6.1 illustrates that, despite optimizing the λ value for a better model fit, the Nelson-Siegel dynamic factors fail to recreate the observed yield curve at specific points. This indicates the model's inability to capture all types of yield curves in reality. Notably, the model loses predictive power on similar dates for the United States and Mexico. We will now examine the yield curve shapes within the data more closely.

Figure A.1 depicts the various mean shapes of the yield curve and the model's ability to reproduce them. We generated plots for the mean yield curve, the mean yield curve when inverted (i.e., the 3-month yield exceeds the 10-year yield), and yield curves with an R-squared less than 0.75. Inconsistencies are evident, which changes in the yield spread may explain. Therefore, we create a series of yield spreads, calculated as the difference between the 10-year and 3-month yields.

Figure 6.2: Comparison of Mexican and United States Yield Spread over the R Squared



(a) MX Yield Spread **(b)** US Yield Spread

Source: Own elaboration.

Figure 6.2 shows that the loss in R-squared aligns with the yield spread's decline. In the Mexican case, a low R-squared is particularly evident during periods of yield curve inversion (grey bars). This suggests that model re-calibration is necessary to avoid significant forecasting errors during these periods.

We can employ factor decomposition to understand the model's shortcomings comprehensively. The Nelson-Siegel model stands out in its ability to describe the yield curve through its factors; level, slope, and curvature. To derive the equivalent factors for the actual yields, we will utilize equation 5.5

Figure A.2 reveals that the primary source of turbulence is the curvature component of the yield curve or β_2 in our model. Given its significance, we must observe the behavior of λ for our forecast. We will thus focus on how the λ value, which maximizes curvature at different maturities, affects forecast accuracy.

For the AR(1) forecast of the Nelson-Siegel components, Diebold and Li (2006) use a fixed value of 0.0609. This value maximizes the λ at 30 months and reduces the model's computational stress. However, we observe large λ values from the fitted values, indicating that 0.0609 may not always be optimal.

Therefore, we must carefully select the λ value for out-of-sample forecasts. Following our cross-validation methodology, we will calculate the λ value that minimizes the RMSE. The cross-validation was conducted for three different forecast horizons: 1, 6, and 12 months, and compared using the Diebold-Mariano statistical test.

Appendix A.1 presents the results of the Diebold-Mariano test conducted for each maturity at each horizon and for both countries. The optimized λ from our cross-validation process consistently improves over the fixed λ at the highest statistical significance across all maturities and forecast horizons. This finding is crucial as it introduces a new step to the Nelson-Siegel forecasting process, resulting in a more accurate forecast and, consequently, more precise deductions from its parameters.

However, there is a second significant finding in the Diebold-Mariano test statistic. The tables in Appendix A.1 show that for the United States, the test statistic is lower for shorter maturities across all forecast horizons, while for Mexico, the opposite is true. It is also notable that the spread of the Diebold-Mariano test statistic is larger for the United States than for Mexico. These findings indicate that it is more accurate to forecast short maturity yields in the United States, whereas, for Mexico, it is better to forecast long maturity yields. Additionally, the cross-validation process returns two RMSE series, allowing us to observe forecast accuracy over time.

Figure A.3 demonstrates the evolution of the RMSE over time. The RMSE values for the United States are generally lower. In 2017, at a point of low R-squared, the cross-validation process corrected the model, resulting in better forecasts during stress periods. This is clearer in the Mexican case, where the period of yield curve stress is longer. Our optimized λ produces significantly lower RMSE for the forecasts. The improvement in accuracy is more evident in the 1-month and 6-month ahead forecasts compared to the 12-month ahead forecast, observable in the forecast spikes and the last observations during the yield curve inversion process.

Two additional observations are noteworthy. The forecast process for the United States shows a relatively high RMSE at the series' start, particularly from 2012 to 2016. In the Mexican case, two spikes in RMSE values exist, while for the United States high RMSE seem to indicate specific yield curve cycles that the model struggles to forecast. For Mexico, sudden changes in the yield curve hinder accurate cross-validation forecasts.

The results could be explained by various factors. One possible explanation is the market's expectations of monetary policy and the macroeconomic outlook. These expectations can rapidly change due to economic surprises or exogenous shocks, causing RMSE spikes as the curve changes quickly from one month to another. However, the adaptability of the cross-validation process allows the lambda value to adjust to the new state of the curve in the following month. From our RMSE statistics, we can also deduce why some countries are more accurately forecasted in short-term maturities rather than long-term, and vice versa. A possible cause would be the monetary policy of the country. In the case of the United States, for most of the data analyzed, the yields were close to zero, possibly making the short term more stable and easier to forecast than the long term, where yields rise. For Mexico, continuous changes in monetary policy could make the short and medium term less accurate compared to the stability of long-term rates. Additionally, the lower variance in Mexico's long-term yields compared to those of the United States may also contribute to this phenomenon.

Ultimately, the particularities of the yield curves' behavior could have different explanations than those theorized in this thesis. However, the objective of this work is not to explain the sudden movements in RMSE, but rather to make the model more adaptable to shocks of any nature.

7. Conclusions

The yield curve is a versatile instrument that conveys information about the economy, market sentiment, and government debt structure. The Nelson-Siegel model is critical as it decomposes the yield curve into descriptive parameters, enhancing forecasting accuracy. As reviewed in Section 3.1, the Nelson-Siegel model offers multiple uses and benefits, making its accuracy crucial.

This thesis exhibits 13 years of fitted value analysis and forecasting evaluation. The results indicate that while the model is highly accurate at times, its precision decays at certain points, rendering it unable to fit specific yield curve shapes.

The initial analysis of the fitted and λ values reveals that the yield curve's curvature varies in maturity at the maximum point. Therefore, Diebold and Li's (2006) recommendation of using a fixed λ may not be universally applicable. This thesis proposes a cross-validation process for selecting a λ value. The findings show that cross-validation corrects high RMSE points during yield curve stress and provides overall better forecasts according to the Diebold-Mariano test. Additionally, we find that short-term maturities are more accurately forecasted for the United States, while long-term maturities are more accurately forecasted for Mexico.

These findings hold significant relevance for central bankers and private individuals. For central bankers, a more accurate forecast of the yield curve, especially one that adapts to yield curve inversions, is invaluable for monetary policy and anticipating changes in the term structure of government bonds. Private individuals can also benefit, as an improved yield curve forecast aids in making better credit and investment decisions and provides insights into a country's financial health.

Precise forecasting is an extensive field with continuously emerging models and specifications. The Nelson-Siegel model, evolving through academic contributions, exemplifies this progress. While this thesis presents a valuable modification to the forecasting process, unanswered questions still need to be answered, particularly during high RMSE or spike periods and the role of volatility in each market at different maturities. Addressing these issues is a challenge and an opportunity for further research and academic contributions.

Bibliography

Benzoni, L., Chyruk, O., and Kelley, D. (2018). Why does the yield-curve slope predict recessions? *Chicago Fed Letter*, (404).

Campbell, J. Y. (1995). Some lessons from the yield curve. *Journal of economic perspectives*, *9*(3), 129–152.

Christensen, J. H., Diebold, F. X., and Rudebusch, G. D. (2011). The affine arbitrage-free class of nelson–siegel term structure models. *Journal of Econometrics*, *164*(1), 4–20.

Coroneo, L., Giannone, D., and Modugno, M. (2016). Unspanned macroeconomic factors in the yield curve. *Journal of Business & Economic Statistics*, *34*(3), 472–485.

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

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, *20*(1), 134–144.

Elizondo, R. (2017). Forecasting the term structure of interest rates in mexico using an affine model. *Estudios Económicos*, *32*(2), 213–253.

Estrella, A. and Mishkin, F. S. (1996). The yield curve as a predictor of us recessions. *Current issues in economics and finance*, *2*(7).

Haubrich, J. G., Dombrosky, A. M., et al. (1996). Predicting real growth using the yield curve. *Economic Review*, *32*(1), 26–35.

Kim, D. H. and Wright, J. H. (2005). *An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates* (Finance and Economics Discussion Series 2005-33). Federal Reserve Board.

Kim, W. J., Jung, G., and Choi, S.-Y. (2020). Forecasting cds term structure based on nelson–siegel model and machine learning. *Complexity*, *2020*(1).

Koopman, S. J., Mallee, M. I., and Van der Wel, M. (2010). Analyzing the term structure of interest rates using the dynamic nelson–siegel model with time-varying parameters. *Journal of Business & Economic Statistics*, *28*(3), 329–343.

Meade Kuribreña, J. A., Rodríguez Regordosa, G., González Anaya, J. A., Treviño Medina, C. A., Laynez Potisek, J., Espinosa Cantellano, I., and Montaño Reyes, L. M. (2012). *El Crédito Público en la Historia Hacendaria de México*. Secretaría de Hacienda y Crédito Público.

Mishkin, F. S. (1990). *Yield Curve* (Technical Report, 3550). National Bureau of Economic Research.

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

Umar, Z., Aharon, D. Y., Esparcia, C., and AlWahedi, W. (2022). Spillovers between sovereign yield curve components and oil price shocks. *Energy Economics*, *109*.

Umar, Z., Yousaf, I., and Aharon, D. Y. (2021). The relationship between yield curve components and equity sectorial indices: Evidence from china. *Pacific-Basin Finance Journal*, *68*.

A. Appendix

A.1. Deibold-Marinao Tests

Table A.1: Diebold-Mariano Test Results for US - 1 Month Ahead

Test Statistic	P Value	Better Method	Significance Level	Maturity
2.959	0.004	Optimized λ	***	1M
3.877	0.000	Optimized λ	***	3M
5.424	0.000	Optimized λ	***	6M
7.734	0.000	Optimized λ	***	1Y
9.545	0.000	Optimized λ	***	2Y
10.100	0.000	Optimized λ	***	3Y
10.219	0.000	Optimized λ	***	5Y
10.206	0.000	Optimized λ	***	7Y
10.192	0.000	Optimized λ	***	10Y
10.233	0.000	Optimized λ	***	20Y
10.337	0.000	Optimized λ	***	30Y

Table A.2: Diebold-Mariano Test Results for US - 6 Months Ahead

Test Statistic	P Value	Better Method	Significance Level	Maturity
2.476	0.015	Optimized λ	**	1M
4.444	0.000	Optimized λ	***	3M
6.610	0.000	Optimized λ	***	6M
8.967	0.000	Optimized λ	***	1Y
10.001	0.000	Optimized λ	***	2Y
10.159	0.000	Optimized λ	***	3Y
10.054	0.000	Optimized λ	***	5Y
9.946	0.000	Optimized λ	***	7Y
9.868	0.000	Optimized λ	***	10Y
9.940	0.000	Optimized λ	***	20Y
10.045	0.000	Optimized λ	***	30Y

Table A.3: Diebold-Mariano Test Results for US - 12 Months Ahead

Test Statistic	P Value	Better Method	Significance Level	Maturity
2.358	0.020	Optimized λ	**	1M
3.845	0.000	Optimized λ	***	3M
5.322	0.000	Optimized λ	***	6M
7.466	0.000	Optimized λ	***	1Y
9.474	0.000	Optimized λ	***	2Y
10.067	0.000	Optimized λ	***	3Y
10.116	0.000	Optimized λ	***	5Y
9.989	0.000	Optimized λ	***	7Y
9.864	0.000	Optimized λ	***	10Y
9.855	0.000	Optimized λ	***	20Y
9.921	0.000	Optimized λ	***	30Y

Table A.4: Diebold-Mariano Test Results for MX - 1 Month Ahead

Test Statistic	P Value	Better Method	Significance Level	Maturity
9.244	0.000	Optimized λ	***	1M
8.332	0.000	Optimized λ	***	3M
8.009	0.000	Optimized λ	***	6M
7.682	0.000	Optimized λ	***	1Y
7.335	0.000	Optimized λ	***	2Y
7.146	0.000	Optimized λ	***	3Y
6.695	0.000	Optimized λ	***	5Y
6.530	0.000	Optimized λ	***	7Y
6.260	0.000	Optimized λ	***	10Y
5.835	0.000	Optimized λ	***	20Y
5.852	0.000	Optimized λ	***	30Y

Table A.5: Diebold-Mariano Test Results for MX - 6 Months Ahead

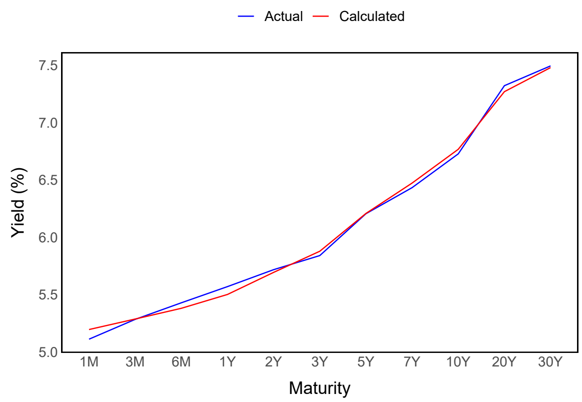
Test Statistic	P Value	Better Method	Significance Level	Maturity
8.964	0.000	Optimized λ	***	1M
8.446	0.000	Optimized λ	***	3M
8.005	0.000	Optimized λ	***	6M
7.547	0.000	Optimized λ	***	1Y
7.247	0.000	Optimized λ	***	2Y
7.236	0.000	Optimized λ	***	3Y
7.035	0.000	Optimized λ	***	5Y
7.062	0.000	Optimized λ	***	7Y
6.893	0.000	Optimized λ	***	10Y
6.552	0.000	Optimized λ	***	20Y
6.297	0.000	Optimized λ	***	30Y

Table A.6: Diebold-Mariano Test Results for MX - 12 Months Ahead

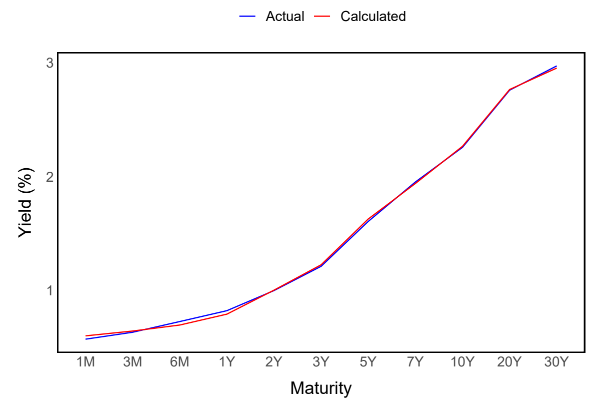
Test Statistic	P Value	Better Method	Significance Level	Maturity
8.367	0.000	Optimized λ	***	1M
7.813	0.000	Optimized λ	***	3M
7.338	0.000	Optimized λ	***	6M
6.904	0.000	Optimized λ	***	1Y
6.723	0.000	Optimized λ	***	2Y
6.815	0.000	Optimized λ	***	3Y
6.705	0.000	Optimized λ	***	5Y
6.793	0.000	Optimized λ	***	7Y
6.754	0.000	Optimized λ	***	10Y
6.644	0.000	Optimized λ	***	20Y
6.462	0.000	Optimized λ	***	30Y

A.2. Yield Curve Shapes

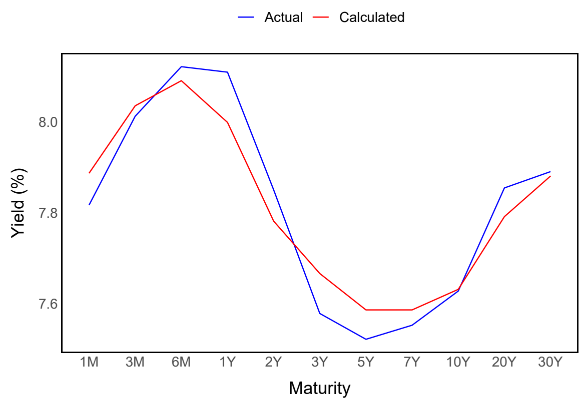
Figure A.1: Comparison of Mexican and United States Yield Curves.



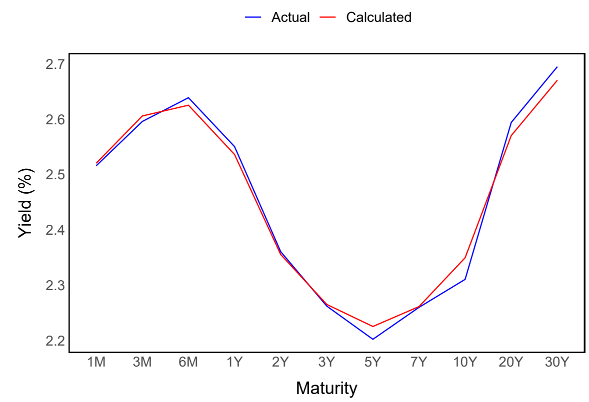
(a) MX Mean Yield Curve



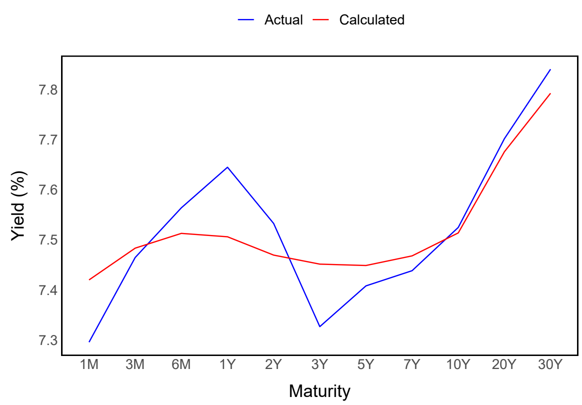
(b) US Mean Yield Curve



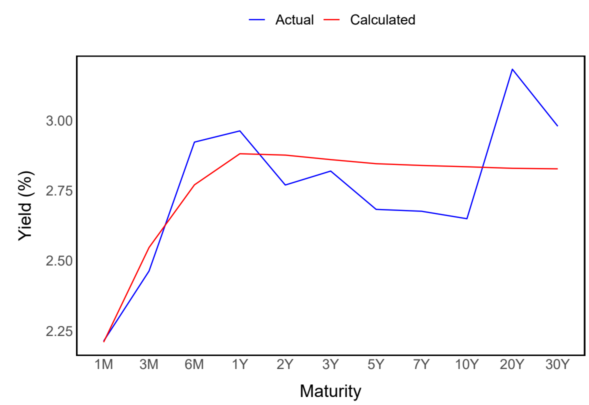
(c) MX Mean Inverted Yield Curve



(d) US Mean Inverted Yield Curve



(e) MX Low R-Squared Mean Curve

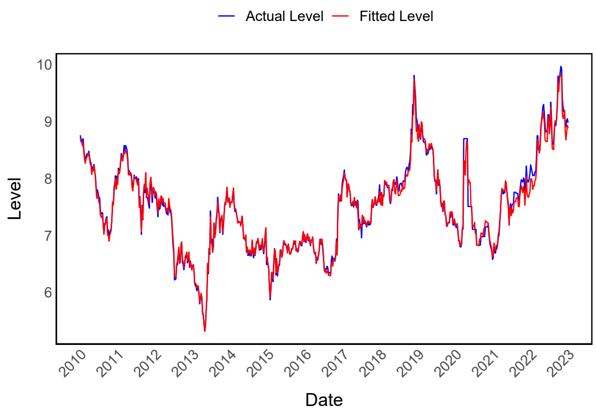


(f) US Low R-Squared Mean Curve

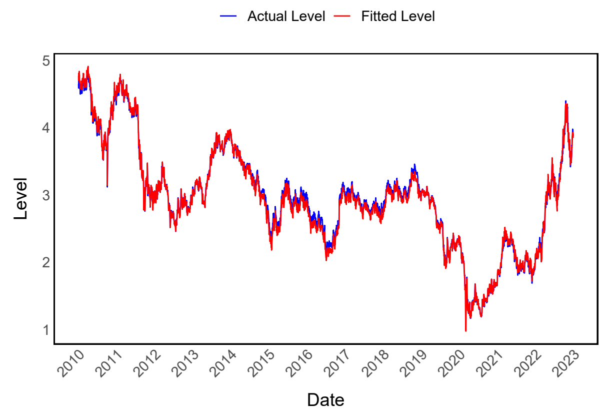
Source: Own elaboration.

A.3. Factor Decomposition

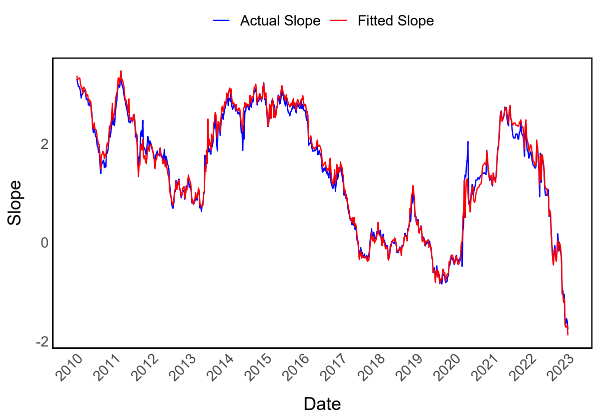
Figure A.2: Comparison of Mexican and United States Yield Curve Components.



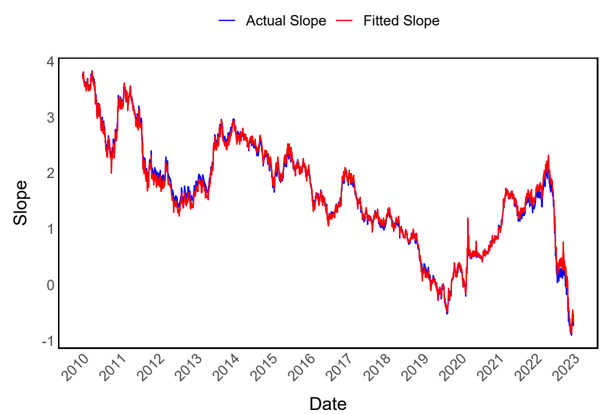
(a) MX Level



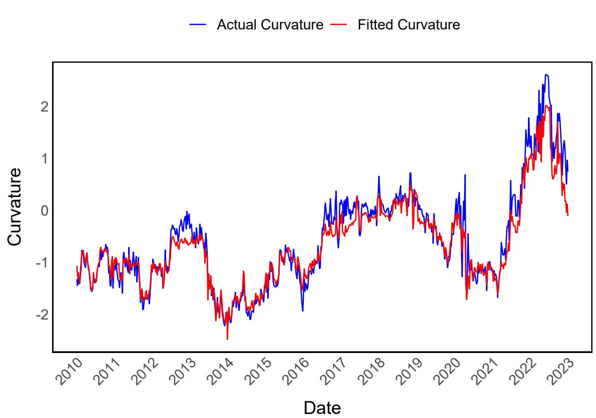
(b) US Level



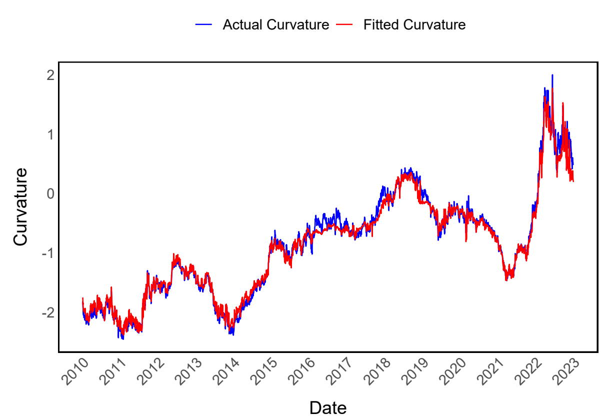
(c) MX Slope



(d) US Slope



(e) MX Curvature

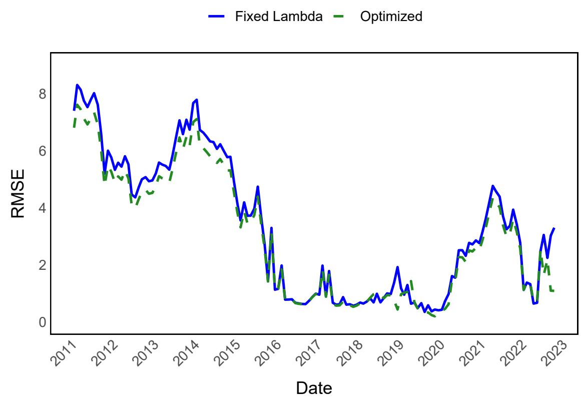


(f) US Curvature

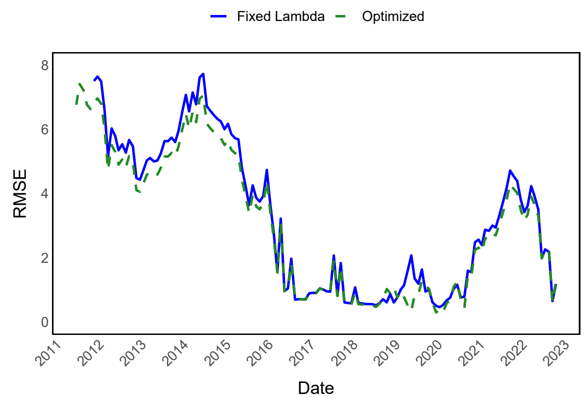
Source: Own elaboration.

A.4. RMSE

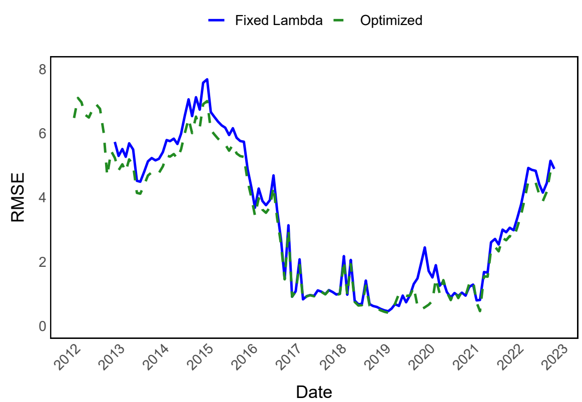
Figure A.3: RMSE Comparison for US and MX - Fixed Lambda vs Optimized.



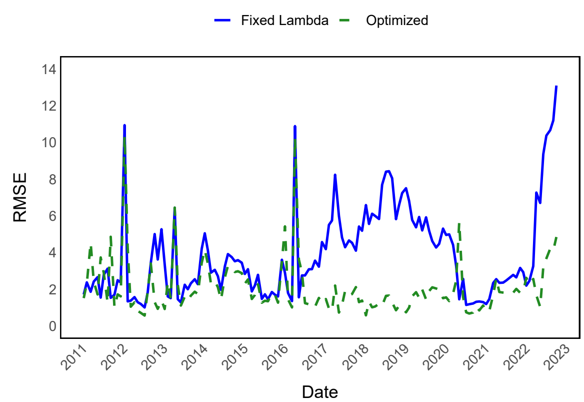
(a) RMSE for US - 1 Month Ahead



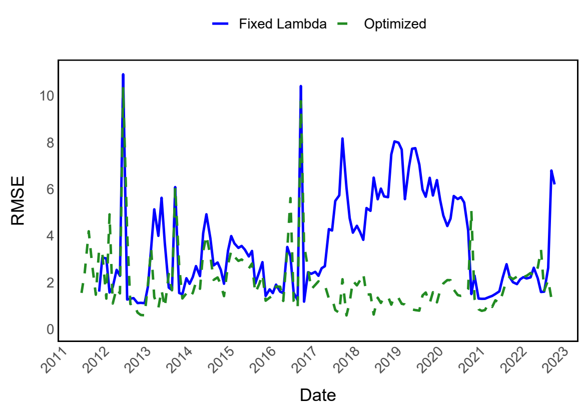
(b) RMSE for US - 6 Months Ahead



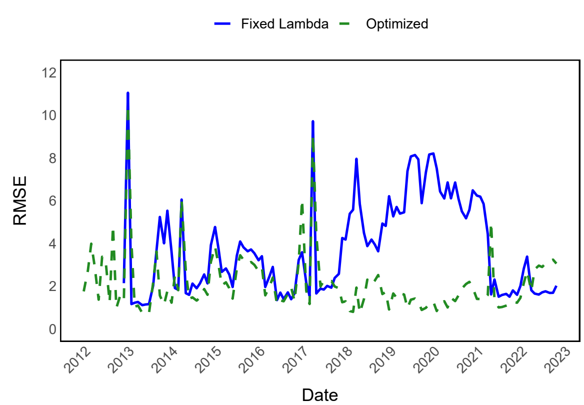
(c) RMSE for US - 12 Months Ahead



(d) RMSE for MX - 1 Month Ahead



(e) RMSE for MX - 6 Months Ahead



(f) RMSE for MX - 12 Months Ahead

Source: Own elaboration.

A.5. Spline

The spline interpolation process for the two year yield, involves setting up the data points, such that x_i represents the maturity in months and y_i represents the corresponding interest rates. The spline cubic function $S(x)$, defined as:

$$S(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3$$

It must follow three conditions to derive the coefficient from the system of equations. The first is the interpolation condition, ensuring the spline passes through the given data points, $S(x_i) = y_i$ and $S(x_{i+1}) = y_{i+1}$. The second is the continuity condition, ensuring the first and second derivatives of the spline are continuous at each interior data point, $S'(x_{i+1}^-) = S'(x_{i+1}^+)$ and $S''(x_{i+1}^-) = S''(x_{i+1}^+)$. The third is the boundary condition, required to uniquely determine the spline. Examples include natural splines, where the second derivative at the endpoints is zero, $S''(x_1) = 0$ and $S''(x_n) = 0$, or clamped splines, where the first derivative at the endpoints is specified, $S'(x_1) = f'(x_1)$ and $S'(x_n) = f'(x_n)$.

Using the conditions outlined, the coefficients a_i, b_i, c_i, and d_i for each interval are computed. The spline function $S(x)$ is then evaluated at the maturity $x = 24$ months to obtain the interpolated rate.

The interval $[x_i, x_{i+1}]$ where $x_i \leq 24 \leq x_{i+1}$ is identified as $[12, 36]$ in this case. Using the data points, the coefficients for the cubic polynomial in this interval are computed by solving the system of equations defined by the interpolation, continuity, and boundary conditions. Substituting $x = 24$ into the cubic polynomial for the interval $[12, 36]$:

$$S(24) = a + b(24 - 12) + c(24 - 12)^2 + d(24 - 12)^3$$

where a, b, c, d are the computed coefficients. Resulting in the interpolated rate for the 24 months maturity.