# Hedging Interest Risk with a Score-Driven Nelson-Siegel Model

Preliminary version | March 2024

Caio César Rostirolla [1]
Sicredi & PPGE/UFRGS

Marcelo Savino Portugal [2]
PPGE/UFRGS

## Abstract

In this paper we implement dynamic extensions to the Nelson-Siegel term structure model and investigate hedging performance using the time-varying parameters of these extensions. Using data from brazilian interest rate futures market contracts, we estimate different specifications of the Nelson-Siegel model with special attention to the econometric treatment of decay parameter and common stochastic volatility. Since the model is non-linear, we propose the use of the Kalman Filter combined with GAS-type dynamics (*Generalized Autoregressive Score*) to estimate the parameters of interest. The great advantage of this methodology is that the likelihood function of the state space model with time-varying parameters is available in closed form, which facilitates the estimation of parameters by maximum likelihood. An innovative feature of this work is the empirical application of Nelson-Siegel models to hedge long-term exposures. We demonstrate how the different estimated specifications perform in the context of immunizing a fixed-income portfolio of long-term credit assets using short-term derivative instruments. Our estimates indicate that considering dynamics in the decay parameter combined with stochastic volatility significantly improves hedge performance. These results are of great interest to financial institutions in the Brazilian market, which do not find sufficient liquidity to hedge long maturities and need to adopt other strategies that increase transaction costs.

**Key-Words**: Score-Driven Nelson-Siegel; Yield Curve; Time-Varying Parameters; Fixed-Income Portfolio Immunization.

[1]Banco Cooperativo Sicredi and Phd Candidate at PPGE/UFRGS.
[2]Professor Titular at PPGE/UFRGS.

# 1 Introduction

Fluctuating interest rates bring with them changes in the prices of assets and liabilities, generating unexpected losses and gains. Therefore, financial institutions must immunize their portfolios to protect their portfolio values against interest rate risk. The term "immunization" describes the steps taken by a manager to construct and manage a portfolio in such a way that this portfolio achieves a predetermined objective. One of these steps requires modeling how the prices of contracts with different maturities react to shocks from different risk factors. There is recognized literature that models both the time series of the yield curve and the cross-section of the term structure of returns and prices (Diebold and Rudebusch (2011)). As the practice of hedging requires the minimization of errors in predicting the returns of a large set of maturities, the imposition of a structure of risk factors on the yield curve, to reduce the dimensionality of the problem, proves to be an attractive solution to this type of problem.

One line of action in portfolio immunization involves the use of parametric models of the term structure of interest rates. In this type of formulation it is assumed that at each moment the temporal structure of interest rates accepts a particular functional form, which is expressed as a function of time and a limited number of parameters. In this line of research, Nelson and Siegel (1987) proposed a seminal statistical model that presents good empirical performance to the data. The methodology received considerable attention with the work of Diebold and Li (2006), where the authors proposed a dynamic version, in which the yield curve is described by a model of just three time-varying factors. The three factors are level, slope and curvature, in which the decay is governed by a static parameter $\lambda$. An extension of this model is the Svensson (1994) specification, used internationally by central banks and for measuring reference and mark-to-market curves. In Brazil, several articles estimated Nelson-Siegel and Svensson models for different interest curves and with different objectives[1]. For example, the parameters of the Nelson-Siegel-Svensson model for Inflation-Linked Bonds (IPCA), Treasury Yields and fixed income private bonds market are published daily by Anbima.

Since Diebold and Li (2006), the empirical yield curve literature has moved towards different generalizations of the dynamic factors model and mainly investigated the inclusion of non-linearities and how they impact the forecasting capacity of these models. Initially,

[1]See, for example, for Exchange Rate Coupon Curve Pinheiro, Almeida, and Vicente (2007) and Laurini and Hotta (2009); for Real Interest Curve and Implicit Inflation Almeida and Lund (2014) and Fernandes and Thiele (2015); for Interest Curve and Macroeconomic Variables Vieira, Fernandes, and Chague (2017) and Fernandes, Nunes, and Reis (2020); for Central Bank Communication and Interest Curve Andrade Alves, Joseph Abraham, and Poletti Laurini (2023); for Yield Curve Forecast Caldeira, Laurini, and Portugal (2010) Caldeira, Moura, and Santos (2016b) and Caldeira et al. (2023); for Portfolio Optimization Caldeira, Moura, and Santos (2016a).

Koopman, Mallee, and Wel (2010) proposed the inclusion of dynamics in $\lambda$ and volatility, using non-linear techniques such as the Extended Kalman Filter. Another initiative emerged with Hautsch and Yang (2012) and Caldeira, Laurini, and Portugal (2010), where the authors also proposed the introduction of temporal variability in $\lambda$ and volatility using Bayesian techniques. Another innovation in these works was the inclusion of stochastic volatility for the covariance matrix of dynamic factors. The third generalization of this literature was the inclusion of macroeconomic covariates in the model (Diebold, Rudebusch, and Aruoba (2006)). In summary, a large number of articles have shown that there are significant gains in predictive power when considering $\lambda$ or volatility as time-varying parameters (TVP) or when macroeconomic variables are included in the model (see, for example, Caldeira et al. (2023)).

This paper contributes to the literature that considers the introduction of time-varying parameters into Nelson-Siegel yield curve models. However, while most articles evaluated the predictive capacity of different generalizations, in this work we investigated improvements that the inclusion of dynamics in $\lambda$ and heterocesticity can bring for hedging purposes. The Nelson-Siegel model is particularly suitable for this problem, since its formulation offers additional risk parameters beyond the hedging technique via *duration*, that is, it allows controlling parallel and non-parallel shocks, and its parsimony allows identifying the temporal variation of the covariance structure of interest rates.

An innovative feature of this work is the empirical application of hedging to long-term exposures. We highlight that in other articles in Brazilian literature, such as Almeida and Lund (2014) and Meirelles and Fernandes (2018), parametric immunization techniques are applied to durations that have instruments with approximate maturities. In the present study, we simulate a credit portfolio that grows monthly with an approximate *duration* greater than 20 years. Aiming at immunization with interest rate derivatives, we know that the Brazilian DI Future market does not present liquidity for these maturities. Thus, we estimate different variations and extensions of the Nelson-Siegel dynamic model using reference rates from the Brazilian nominal interest market and use the results of the parameters of interest to replicate the excess returns *targets* with liquid derivative instruments with terms between 1 up to 12 years adjusted to minimize portfolio variance, as Litterman and Scheinkman (1991). This empirical problem of hedging long-term exposures is of great interest to financial institutions such as Pension Funds, Banks and other Credit Institutions, which generally face index mismatches in their balance sheets. Furthermore, a large part of the Brazilian fixed income market immunizes portfolios via *duration*, considering in most cases only level fluctuations[2]

[2]A popular approach in the Brazilian market, called '*key rate duration* ' assumes that changes in interest rates can be accurately described by changes in the level of a limited number of segments (vertices or drivers of the yield curve) into which the term structure is subdivided, thus generalizing the concepts of duration and convexity for a multivariate context, considering the portfolio's joint exposure to these '*key rates*'. See, for example, Ho (1992).

For this proposal, our empirical strategy brings important methodological elements that differentiate this article from the literature applied to the Brazilian market. In the Nelson-Siegel model, the factor structure is fully controlled by the $\lambda$ parameter, as it defines the decay rate of the exponential polynomial. In other words, different values of the decay parameter provide different time series for the dynamic factors and therefore different predictions. This parameter has received little attention, and is often not even estimated, but rather defined as an *ad hoc* value. Diebold and Li (2006) set $\lambda$ at 0.0609, imposing a maximum point of curvature of the term structure in 30 months. In Brazilian literature, Almeida et al. (2009) define a fixed $\lambda$ of 3.58 for the Brazilian curve in a four-factor model between November 2004 and December 2006. In turn, Caldeira, Laurini, and Portugal (2010) found a value fixed to $\lambda$ of 1.255 between January 2006 and February 2009 for a three-factor model. Meirelles and Fernandes (2018) minimized the mean squared error with moving regressions and documented an optimal $\lambda$ ranging between 0.532 and 1.532. These empirical estimates show that the works differ mainly in the ability to predict short and long maturities of the interest curve, since small values of $\lambda$ cause a smooth reduction in the exponential coefficients and serve to adjust the curve well in long terms, while high values of $\lambda$ result in a more pronounced reduction of the exponential coefficients and serve to adjust the curve in shorter periods.

In this work we focus on estimating dynamic factors and propose an intuitive version of modeling the temporal variation of the model's dynamic parameters. The inclusion of dynamics in $\lambda_t$ makes the model non-linear, so the Kalman Filter cannot be applied in the conventional way. Koopman, Mallee, and Wel (2010) proposed the Extended Kalman Filter formulation to deal with this situation. Here, we use the econometric methodology GAS (*Generalized Autoregressive Score*)[3] combined with Kalman Filter, so that we impose a rule of movement for the parameters from the conditional likelihood *score* function, following Creal, Koopman, and Lucas (2013) and Harvey (2013). We present the analytical derivations of the Nelson-Siegel model with the GAS methodology in state space form, which can be called *Score-Driven Nelson-Siegel*. The great advantage of this methodology is that the likelihood function of the state space model with time-varying parameters is available in closed form, which facilitates the estimation of parameters by maximum likelihood and substantially reduces the computational cost in relation to other econometric techniques.

We know of only three articles applied to the Brazilian market that modeled the temporal variation in $\lambda$. Considering Svensson's formulation, Laurini and Hotta (2010) estimated decay values between 2004 and 2006. Caldeira, Laurini, and Portugal (2010) estimated it for the period 2007 to 2009 and also found considerable variability. Franciscangelo (2015) applied the Extended Kalman Filter for the period 2012-2015. According to these references, the decay factor $\lambda_t$ reached levels above 2.00 in 2006 and 2008 and was situated

[3]See `https://www.gasmodel.com/`.

in the range 0.30-0.90 between 2012 and 2015.

Regarding international literature, Koopman, Mallee, and Wel (2010) considered temporal variation in $\lambda$ through the Extended Kalman Filter, while Creal, Koopman, and Lucas (2008) considered the GAS approach. Recently, Quaedvlieg and Schotman (2022) also considered variability in $\lambda$ using a *dynamic conditional score* (DCS) specification and investigated the hedging capacity of different Nelson-Siegel specifications for immunizing long-term liabilities, a similar problem to the treaty in this work. Other articles that documented relevant variability in $\lambda$ were Cordeiro et al. (2019) and Caldeira et al. (2023).

Additionally, we allow for a stochastic volatility framework with GAS dynamics to control for the presence of conditional heteroskedasticity observed in interest rates. Koopman, Mallee, and Wel (2010) considered in the form of state space a specification where the main component of the error terms of the Nelson-Siegel equation follow a GARCH process. This same strategy was used by Quaedvlieg and Schotman (2022) and Caldeira et al. (2023). Laurini and Hotta (2010), Hautsch and Ou (2012), Hautsch and Yang (2012) and Caldeira, Laurini, and Portugal (2010) proposed estimation of stochastic volatility with Bayesian techniques in Nelson-Siegel specifications. Using the GAS approach, Koopman, Lucas, and Zamojski (2017) showed that the inclusion of multivariate stochastic volatility combined with the $t-$Student distribution considerably improves model fit. We follow the GAS approach for modeling the common volatility of *yields*, while the latent factor structure is assumed homoscedastic. Like the yield curve parameters, the log-volatility factors also have an analytical scoring structure and the likelihood functions are available in closed form.

In our model, the temporal variation in $\lambda$ and the variance of the error terms are interconnected, as the covariance matrix directly influences the updates of $\lambda$. If we consider homoscedasticity of the error terms, that is, if volatility is not modeled, there is a considerable loss of predictive quality, which is quite relevant when we deal with a hedging problem. The empirical hedging exercise shows, step by step, how to build a portfolio for immunization, where mark-to-market (MtM) errors are considerably reduced when we compare models with constant $\lambda$ and with homocesticity of the error terms (such as in Diebold and Li (2006)).

The article is organized as follows. In the next section we introduce the Nelson-Siegel model with GAS dynamics. We then present the data treatment and the model results, such as the estimation of the time-varying parameters $\lambda_t$ and the dynamic factors. Finally, we present the results of the immunization exercise.

## 2 Score-Driven Nelson-Siegel

In this section we present the Nelson-Siegel model for estimating the term structure of interest rates and formulate the time-varying parameters with GAS-type dynamics, giving rise to the name Score-Driven Nelson-Siegel.

### 2.1 Nelson-Siegel Dynamic Model for the Term Structure of Interest Rates

Nelson and Siegel (1987) originally derived the following specification to describe the dynamics of *forward* curves:

$$f(\tau) = \alpha_1 + \alpha_2 \exp(-\lambda\tau) + \alpha_3(\lambda\tau)\exp(-\lambda\tau) \tag{1.1}$$

so that, if we use the following relationship between *forward* rates and spot rates (*spot*), it is possible to calculate the *spot* curve using the NS model:

$$y(\tau) = \frac{1}{\tau}\int_0^\tau f(\tau)d\tau \tag{1.2}$$

At a given period in time, $t$, the interest rate curve, denoted by $y_t(\tau)$, represents the interest rate as a function of maturities $\tau$. The exponential model proposed by Nelson and Siegel (1987) and reinterpreted by Diebold and Li (2006) considers a parametric structure for the evolution of the term structure of interest rates over time, in which the coefficients can be interpreted as level, slope and curvature:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3,t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right) + \epsilon_t \tag{1.3}$$

In the equation above, the shape of the yield curve is determined by the three parameters ($\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$) and by the factor loadings associated with they. The parameter $\lambda_t$, treated as a constant in Diebold and Li (2006), determines the decay rate of the exponential polynomial, so that small (large) values of $\lambda_t$ are associated with a decay smooth (fast), and provides better grip for longer maturities (short). The weight of the first component is equal to 1 (constant) for all maturities, so $\beta_{1,t}$ is interpreted as the interest curve level factor, which equally influences short-term and long-term rates , an important characteristic for hedging purposes. The charge of the second component, $\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right)$, starts equal to 1 and converges to zero monotonically and quickly, so that $\beta_{2,t}$ is interpreted as the interest curve slope factor, or short-term factor, given that this factor has a strong influence on shorter-term interest rates . The charge of the third component, $\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right)$, is a concave function, in which it takes on a value equal to zero for zero maturity, increases with maturity, and then monotonically converges to zero at longer maturities. Thus, $\beta_{3,t}$ is interpreted as the curvature factor, or medium-term factor, as it has a strong influence on the rates at the core of the curve.

In each period $t$, there are interest rates $y_t(\tau)$ for different maturities $\tau$. Thus, the above equation can be estimated for each period in time $t$, obtaining time series for the parameter vector $\beta_t$. As an example, Diebold and Li (2006) and Almeida and Lund (2014) used a two-step procedure to estimate the model. In the first step, the time series are obtained by moving OLS/NLS regressions at every $t$, and then the three series of $\beta_{1,t}, \beta_{2,t}$

and $\beta_{3,t}$ are modeled in a second step as a VAR model (1).

An alternative approach is to represent the model in state-space form and estimate all parameters simultaneously through the Kalman Filter. The system, composed of a measurement equation and a transition equation, is given by:

$$\begin{aligned} y_t(\tau) &= \Lambda(\lambda_t)\beta_t + \epsilon_t, \qquad \mathcal{N} \sim (0, \Sigma_\epsilon \\ \beta_t &= (I - \Phi)\mu + \Phi\beta_{t-1} + \eta_t \end{aligned} \tag{1.4}$$

where the measurement equation defines the vector $y_t(\tau)$ of interest rates with dimension $(T \times N)$ as the sum of the dynamic factors $\beta_t$ multiplied by the factor loadings $\Lambda(\lambda_t)$ is a vector $\epsilon_t$ of Gaussian errors that are independent between maturities. The vector $\beta_t$ $(3 \times 1)$ represents the dynamic factors, and $\Phi$ is the VAR coefficient matrix that determines the dynamics of the states over time. Regarding $\Lambda(\lambda_t)$, it is a non-linear $(N \times 3)$ matrix of weights/loadings, in which it varies in time only if the decay parameter $\lambda_t$ is variable.

To estimate the model in linear form with the Kalman Filter, for example, we need to consider that the weight matrix is constant over time for each maturity, implying $\lambda_t = \lambda$. However, $\lambda$ is a crucial parameter for building a hedge portfolio, as it governs exposure to different factors across maturities. Therefore, considering $\lambda_t$ as a variable over time can bring gains in hedge performance and predictive accuracy. The difficulty, however, is that the inclusion of variability in $\lambda$ makes the model non-linear and other approaches need to be applied, such as Bayesian techniques (Hautsch and Yang (2012) and Caldeira, Laurini, and Portugal (2010)) or Kalman Filter Extended (Koopman, Mallee, and Wel (2010); Caldeira et al. (2023)).

Recently, some works that considered GAS-type dynamics on the parameters of the Nelson-Siegel model emerged as an alternative (Creal, Koopman, and Lucas (2008); Mesters, Schwaab, and Koopman (2014); Koopman, Lucas, and Zamojski (2017); Quaedvlieg and Schotman (2022); Vleuten, Lange, and Wel (2023)). Despite this, this methodology has not yet been tested with data from the Brazilian market[4].

Therefore, we will follow the GAS approach to estimate the time-varying factors $\beta_t$ and the decay parameter $\lambda_t$. Additionally, we consider stochastic volatility to control for conditional heteroskedasticity. In the next section we will introduce some state space concepts under the GAS optimum, in line with Creal, Koopman, and Lucas (2013) and Harvey (2013). The following two sections derive the Nelson-Siegel formulation considering GAS-type dynamics with only the three factors $\beta_t$ varying in time ($\lambda$ fixed), which we call NS3F, and with the four time-varying parameters $(\beta_t, \lambda_t)$, which we call NS4F. Finally, in a third section, we present the inclusion of common stochastic volatility in the error matrix $\Sigma_\epsilon$ also following GAS dynamics. This last case will result in generalizations called

[4] Santos, Ribeiro, and Sanfins (2019) applied the GAS model to estimate the level factors of the interest curve. However, the authors adopted a more simplified empirical strategy than that exposed in the present work.

NS3F-V and NS4F-V.

## 2.2 GAS models in State Space Form

In this subsection we introduce the technical aspects of GAS models from the perspective of state space. We follow the Harvey (1990) notation.

Consider a time series model that has the following measurement and transition equations under state space representation:

$$\begin{aligned} y_t &= Z_t\alpha_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, H_t), \\ \alpha_t &= T_t\alpha_{t-1} + \eta_t \qquad \eta_t \sim \mathcal{N}(0, Q_t), \qquad t = 1, ..., n, \end{aligned} \tag{1.5}$$

where $y_t$ is a vector of dimension $N \times 1$ of observable variables, such as interest rates for different maturities, $\epsilon_t$ is a vector of dimension $N \times 1$ of measurement errors, $\alpha_t$ is a $m \times 1$ vector of state variables and $\eta_t$ is a $m \times 1$ vector of transition errors. The two error terms $\epsilon_t$ and $\eta_t$ are assumed to be Gaussian distributed and uncorrelated for all time periods, that is, $\mathbb{E}(\epsilon_t \ eta_s') = 0 \ \forall \ t, s$ and present covariance matrices $H_t$ and $Q_t$, respectively. The initial value of the state vector is also assumed to be Gaussian distributed, $\alpha_0 \sim \mathcal{N}(\alpha_0, P_0)$ and uncorrelated $\forall t$ with $\epsilon_t$ and $\eta_t$.

Following Harvey (1990), it is generally assumed that the matrices $Z_t$, $H_t$, $T_t$ and $Q_t$ are non-stochastic. As a result, the system (1.5) is linear with respect to the state vector. Conditioned on the informational set $Y_{t-1} = \{y_{t-1}, ..., y_1\}$ and a vector of parameters $\theta$, the state vector and the vector of observations are Gaussian; i.e., $y_t|Y_{t-1}; \theta \sim \mathcal{N}(Z_t a_t, F_t)$ and $\alpha_t|Y_{t-1}; \theta \sim \mathcal{N}(a_t, P_t)$, and the log-likelihood function on $t$ is:

$$l_t = \log p(y_t|Y_{t-1}; \theta) \propto -\frac{1}{2}\Big(\log|F_t| + v_t' F_t^{-1} v_t\Big) \tag{1.6}$$

so that prediction error $v_t$, its covariance matrix $F_t$, the conditional mean of the state vector $a_t$, and the quadratic error matrix $P_t$ (MSE), are recursively estimated by the Kalman Filter. The update steps are given by:

$$\begin{aligned} v_t &= y_t - Z_t a_t \\ a_{t|t} &= a_t + P_t Z_t' F_t^{-1} v_t \\ F_t &= Z_t P_t Z_t' + H_t \\ P_{t|t} &= P_t - P_t Z_t' F_t^{-1} Z_t P_t \end{aligned} \tag{1.7}$$

and the prediction steps are:

$$\begin{aligned} a_{t+1} &= T_{t+1} a_{t|t} \\ P_{t+1} &= T_{t+1} P_{t|t} T_{t+1}' + Q_{t+1} \end{aligned} \tag{1.8}$$

for $t = 1, ..., n$. Specifically, we have that $a_t = \mathbb{E}(\alpha_t|Y_{t-1}, \theta)$ is called a predictive filter with MSE matrix $P_t = \mathbb{E}[(a_t - \alpha_t)(a_t - \alpha_t)'|Y_{t-1}, \theta]$ while $a_{t|t} = \mathbb{E}(\alpha_t|Y_t, \theta)$ is called real-time filter with MSE matrix $P_{t|t} = \mathbb{E}[(a_{t|t} - \alpha_t)(a_{t|t} - \alpha_t)'|Y_{t-1}, \theta]$. The state space model in

(1.5) is defined in contemporary form as in Harvey (1990). We chose the Harvey (1990) notation so that the system matrices present the same temporal structure as the vector of time-varying parameters.

In GAS family models, it is assumed that variations in the matrix system over time are endogenous and depend on past observations. Therefore, although stochastic, the matrix system is predetermined, which means that, conditional on past observations, the system matrices can be considered fixed. As a result, the model is still considered conditionally Gaussian as in Harvey (1990).

This configuration has three attractive features. First, both the state vector and the observation vector are conditionally Gaussian. Second, the likelihood function can be written in the form of the decomposition of prediction errors (Equation 1.6) and calculated using the Kalman Filter in (1.7). Third, although the model is not linear in observations, the Kalman Filter results in optimal state vector estimates as in Harvey (1990). The main analytical challenge here is represented by the joint updating of the system matrices and the state vector. To solve this problem, we derive a different set of recursions that work in parallel with the Kalman Filter.

In the GAS model, the time-varying elements of the matrix system in (1.5) are collected in a vector $f_t$, called the TVP vector. As in Creal, Koopman, and Lucas (2008), Creal, Koopman, and Lucas (2013) and Harvey (2013), the following motion rule is defined for the TVP vector:

$$f_{t+1} = c + Af_t + Bs_t, \qquad s_t = \mathcal{S}_t \nabla_t, \qquad t = 1, ..., n, \tag{1.9}$$

with

$$\nabla_t = \frac{\partial l_t}{\partial f_t}, \qquad \mathcal{S}_t = \mathcal{I}_{t|t-1}^{-k} = -\mathbb{E}_t \left( \frac{\partial l_t^2}{\partial f_t f_t'} \right)^{-1}$$

where $\nabla_t$ is the *score* (gradient) function of the log-likelihood function $l_t$ with respect to the TVP vector $f_t$, the scale matrix, $\mathcal{S}_t$, is defined as the inverse of the Fischer information matrix $\mathcal{I}_{t|t-1}^{k}$,, $k = \{0, 1/2, 1\}$. In this case, $s_t$ has a conditional mean equal to zero and a conditional variance equal to the inverse of the information matrix. The parameters in $B$ determine the sensitivity of the TVP parameters in relation to the conditional likelihood score, and thus in relation to the information contained in the prediction error. The matrix system can contain both time-varying and constant elements. The constant parameters are collected in the vector $\theta_m$. Therefore, in each period $t$, the matrix system depends on $f_t$ and $\theta_m$, denoting $Z_t = Z(f_t, \theta_m)$, $T_t = T(f_t, \theta_m)$, $H_t = H(f_t, \theta_m)$, and $Q_t = Q(f_t, \theta_m)$. The scores vector $s_t$ is calculated conditionally on the information up to the period $t$, so the vector $f_t$ is entirely determined by past observations and the static parameters vector $\theta_f = \{c, A, B\}$. Since the matrix system is purely *observation-driven*, i.e., entirely determined by past observations and the vector $\theta = (\theta_f', \theta_m')'$, the model is defined as conditionally Gaussian and the log-likelihood function (1.6) can be evaluated recursively through the Kalman Filter.

The gradient $\nabla_t$ and the Fisher information matrix $\mathcal{I}_t$ can be calculated analytically. Given the expressions (1.5)-(1.9), we have:

$$\begin{aligned}\nabla_t &= \frac{1}{2}\Big[\check{F}_t(F_t \otimes F_t)^{-1}vec(v_t v_t^{'} - F_t) - 2\check{V}_t^{'} F_t^{-1} v_t\Big] \\ \mathcal{I}_t &= \frac{1}{2}\Big[\check{F}_t^{'}(F_t \otimes F_t)^{-1}\check{F}_t + 2\check{V}_t^{'} F_t^{-1}\check{V}_t\Big]\end{aligned} \tag{1.10}$$

where $\check{V}_t = \partial v_t/\partial f_t^{\prime}$ and $\check{F}_t = \partial vec(F_t)/\partial f_t^{\prime}$ measures the sensitivity of the prediction error $v_t$ and its variance $F_t$ with respect to $f_t$.

Note that all elements in the Fischer information matrix in $\mathcal{I}_t$ are calculated using information up to $t-1$. On the other hand, the gradient $\nabla_t$ contains information contemporaneous with the observation vector $y_t$ via the error vector $v_t$. The terms $\check{V}_t$ and $\check{F}_t$ are fundamental to the gradient $\nabla_t$. They measure the sensitivity of the first and second moments of the state vector in relation to $f_t$, respectively. Together with the variance of the prediction error ($F_t$) and the curvature of the conditional likelihood (proportional to $\mathcal{I}_t$), they determine the impact that new information, summarized in $v_t$, has over the TVP vector. Note that $v_t$ and $F_t$ are calculated recursively via the Kalman Filter in (1.7), while the Jacobian counterparts $\check{V}_t$ and $\check{F}_t$, are obtained recursively through the following filtering.

The Jacobian counterparts of the Kalman Filter result in the following set of expressions:

$$\begin{aligned}\check{V}_t &= -\Big[(a_t^{'} \otimes I_N)\check{Z}_t + (a_{t-1|t-1} \otimes Z_t)\check{T}_t\Big] \\ \check{F}_t &= 2N_N(Z_t P_t \otimes I_N)\check{Z}_t + 2(Z_t \otimes Z_t)N_m(T_t P_{t-1|t-1} \otimes I_m)\check{T}_t + \check{H}_t + (Z_t \otimes Z_t)\check{Q}_t\end{aligned} \tag{1.11}$$

where $\check{Z}_t = \partial vec(Z_t)/\partial f_t^{'}$, $\check{H}_t = \partial vec(H_t)/\partial f_t^{'}$, $\check{T}_t = \partial vec(T_t)/\partial f_t^{'}$ and $\check{Q}_t = \partial vec(Q_t)/\partial f_t^{'}$ are the Jacobians of the matrix system with respect to $f_t$, and $N_m$ is a symmetrizing matrix (i.e., for any matrix $S$ of dimension $n \times n$, we have that $N_n vec(S) = vec[\frac{1}{2}(S + S^{'})]$).

Establishing the formulation for $\nabla_t$, $\mathcal{I}_t$, $\check{V}_t$ and $\check{F}_t$ together, we can calculate the *score* $s_t = \mathcal{S}_t\nabla_t$ and estimate the vector $f_t$ recursively using the filter (1.9 ).

The Kalman Filter combined with GAS filtering (Kalman-GAS) can be described according to the algorithm described below:

1. Initialize $a_{0|0}$, $a_1$, $P_{0|0}$, $P_1$, $f_1$, e $Z_1$, $T_1$, $H_1$, $Q_1$.
2. Para $t = 1, ..., n$:
   (a) Calculate $\check{Z}_t$, $\check{T}_t$, $\check{H}_t$, $\check{Q}_t$
   (b) Calculate $v_t$, $F_t$, $\check{V}_t$, $\check{F}_t$, $l_t$
   (c) Calculate $a_{t|t}$, $P_{t|t}$, $\nabla_t$, $\mathcal{I}_t$, $s_t$
   (d) Calculate $f_{t+1}$
   (e) Calculate $Z_{t+1}$, $T_{t+1}$, $H_{t+1}$, $\mathcal{Q}_{t+1}$

(f) Calculate $a_{t+1}$, $P_{t+1}$

Finally, the parameter vector $\theta$ can be estimated by maximum likelihood (ML), that is, $\widehat{\theta} = \arg\max \sum_{t=1}^{n} l_t(\theta)$. Given the above algorithm, the evaluation of the log-likelihood function is straightforward and the maximization can be obtained numerically.

### 2.3 GAS-Kalman Filter - Nelson-Siegel Model with 3 Factors (NS3F)

The Nelson-Siegel 3-Factor Model (NS3F) in state space form with the factors following GAS dynamics can be described as follows. Initially, we consider the TVP factor $f_t = \beta_t$, $\beta_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t})'$ and we consider Gaussian density for $y_t - \Lambda(\lambda_t) f_{t-1} \sim N(0, \Sigma_\epsilon)$, where $f_t$ follows an equation of motion of the GAS(1,1) type:

$$f_t = \omega + A f_{t-1} + B s_t \qquad \text{with} \qquad s_t = \nabla_t S_t \tag{1.12}$$

where $\omega$ is a vector of conditional means and $A$ and $B$ elasticity matrices. As defined in (1.6) and (1.9), the vector of *scores* $s_t$ and the Fischer information matrix $\mathcal{I}_{t|t-1}$ can be derived as follows. In the NS3F model of equation (1.5), the log-likelihood function is given by:

$$\log p(y_t|\beta_t) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|F_t| - \frac{1}{2}(y_t - \Lambda(\lambda)\beta_t)' F_t^{-1}(y_t - \Lambda(\lambda)\beta_t), \tag{1.13}$$

and has a vector of *scores* denoted by

$$\begin{aligned}\nabla(y_t|\beta_t) = \frac{\partial}{\beta_t}\log p(y_t|\beta_t) &= Z' F^{-1} v_t \\ &= \Lambda(\lambda)' F^{-1}(y_t - \Lambda(\lambda)\beta_t)\end{aligned} \tag{1.14}$$

The matrix $S_t$, which is equal to the Fisher information inverse matrix, is defined as

$$\begin{aligned}S_t = \mathcal{I}_{t|t-1}^{-k} = -\mathbb{E}[\mathcal{H}(y_t|\theta_t)|\theta_{t|t-1})]^{-1} &= [Z' F^{-1} Z]^{-1} \\ &= [\Lambda' F^{-1} \Lambda]^{-1},\end{aligned} \tag{1.15}$$

where $\mathcal{H}(y_t|\theta_t)|\theta_{t|t-1})$ denotes the Hessian matrix defined by

$$\mathcal{H}(y_t|\theta_t)|\theta_{t|t-1}) = \frac{\partial^2}{\partial\beta_t \partial\beta_t'}\log p(y_t|\beta_t) = -\Lambda' F^{-1}\Lambda \tag{1.16}$$

Now we will redefine the matrix $S_t$ as a predictor function of the Fischer information matrix. This penalty matrix is defined as

$$S_t = [\rho_\beta I_{t|t-1}]^{-1} = [\rho_\beta[\Lambda' F^{-1}\Lambda]]^{-1} \tag{1.17}$$

where $\rho_b$ denotes a penalty coefficient. We considered the inclusion of this coefficient for invertibility and computational treatment purposes. This matrix has a full rank as $\lambda$ does not approach zero (resulting in convergence of the second factorial load of $\Lambda$ to the first load) or assumes a high value (resulting in convergence of the third factorial load of $\Lambda$ for the second load).

In the NS3F model, the Kalman-GAS Filter update steps are identical to the conventional Kalman Filter. Based on (1.5-1.7), these steps are given by:

$$
\begin{aligned}
\beta_{t|t} &= \beta_{t|t-1} + P_{t|t-1}Z'(ZP_{t|t-1}Z' + H_t)^{-1}v_t \\
&= \beta_{t|t-1} + P_{t|t-1}\Lambda^{'}(\Lambda P_{t|t-1}\Lambda^{'} + \Sigma_\epsilon)^{-1}(y_t - \Lambda\beta_{t|t-1}) \\
P_{t|t} &= P_{t|t-1} - P_{t|t-1}Z'(ZP_{t|t-1}Z' + H_t)^{-1}ZP_{t|t-1} \\
&= P_{t|t-1} - P_{t|t-1}\Lambda^{'}(\Lambda P_{t|t-1}\Lambda^{'} + \Sigma_\epsilon)^{-1}\Lambda P_{t|t-1}
\end{aligned}
\tag{1.18}
$$

The main changes are made to the prediction steps, so that now the state vector and its variance matrix are given as:

$$
\begin{aligned}
\beta_{t+1|t} =&\mathbb{E}[\beta_{t+1}|\mathcal{F}_t] \\
=&\mathbb{E}[\omega + \Phi\beta_t + S_t\nabla(y_t|\beta_t)|\mathcal{F}_t] \\
=&\omega + \Phi\beta_{t|t} + S_t\nabla(y_t|\beta_{t|t})
\end{aligned}
\tag{1.19}
$$

$$
\begin{aligned}
\mathbb{V}[\beta_{t+1}|\mathcal{F}_t] =&\mathbb{V}[\omega + \Phi\beta_t + S_t\nabla(y_t|\beta_t)|\mathcal{F}_t] \\
=&\Phi P_{t|t}\Phi^{'} + S_t\mathbb{V}[\Lambda^{'}F^{-1}(y_t - \Lambda\beta_t)|\mathcal{F}_t]S_t^{'} \\
=&\Phi P_{t|t}\Phi^{'} + S_t\mathbb{V}[\Lambda^{'}F^{-1}\Lambda\beta_t|\mathcal{F}_t]S_t^{'} \\
=&\Phi P_{t|t}\Phi^{'} + S_t\Lambda^{'}F^{-1}\Lambda P_{t|t}\Lambda^{'}F^{-1}\Lambda S_t^{'}
\end{aligned}
\tag{1.20}
$$

where $\mathbb{E}[\cdot]$ and $\mathbb{V}[\cdot]$ denote the expectation and variance operators respectively, and $\mathcal{F}_t$ is the informational set in the period $t$. Therefore, the *score* evaluated under the updated state at $t$ contributes to the prediction at $t+1$.

## 2.4 GAS-Kalman Filter - Nelson-Siegel Model with 4 Factors (NS4F)

In this subsection we expand the application of the Extended Kalman Filter from Koopman, Mallee, and Wel (2010) in the formulation of the Dynamic Nelson-Siegel model with 4 Factors (NS4F), which allows variability in $\lambda$. However, here we adopt a GAS dynamic over the four time-varying parameters. We define the TVP vector as $f_t^+ = (\beta_t', \log\lambda_t)$ of dimension $4 \times 1$. Now the motion rule GAS(1, 1) is given as

$$
f_t^+ = \omega + Af_{t-1}^+ + Bs_t \tag{1.21}
$$

When we include a fourth factor, $\log\lambda_t$, as a time variable in the Nelson-Siegel model, the model becomes non-linear in the state vector. Thus, the new observation equation is non-linear in $\beta_t$ as:

$$
y_t = Z(\theta_t) + \epsilon_t = \Lambda(\lambda_t)\beta_t + \epsilon_t \tag{1.22}
$$

Note that the Kalman Filter only applies to models that are linear in the state vector. This way, as in Koopman, Mallee, and Wel (2010) and Caldeira et al. (2023), we locally linearize the function $Z(\theta_t) = \Lambda(\lambda_t)\beta_t$ in $f = \theta = \theta_{t|t-1}$. The resulting update equations are identical as in the Extended Kalman Filter. The linearized model becomes:

$$y_t = Z_t(\theta_{t|t-1}) + \check{Z}_t \cdot (\theta_t - \theta_{t|t-1}) + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \Sigma_\epsilon), \quad \check{Z}_t = \left.\frac{\partial Z_t(\theta)}{\partial \theta}\right|_{\theta=\theta_{t|t-1}} \tag{1.23}$$

where

$$\check{Z}_t = \left.\frac{\partial Z_t(\theta)}{\partial \theta}\right|_{\theta=\theta_{t|t-1}} = \begin{bmatrix} \left.\frac{\partial Z_t}{\partial \beta_{1,t}}\right|_{\beta_{1,t}=\beta_{1,t|t-1}} & \left.\frac{\partial Z_t}{\partial \beta_{2,t}}\right|_{\beta_{2,t}=\beta_{2,t|t-1}} & \left.\frac{\partial Z_t}{\partial \beta_{3,t}}\right|_{\beta_{3,t}=\beta_{3,t|t-1}} & \left.\frac{\partial Z_t}{\partial \lambda_t}\right|_{\lambda_t=\lambda_{t|t-1}} \end{bmatrix} \tag{1.24}$$

where for the fourth column we made use of the chain rule such that $\frac{\partial Z(x_t)}{\partial \log \lambda_t} = \frac{\partial Z(x_t)}{\partial \lambda_t} \cdot \lambda_t$ and we find

$$\begin{aligned} \frac{\partial Z(x_t)}{\partial \lambda_t} =& \frac{\beta_{2,t} \cdot \exp(-\tau \cdot \lambda_t)(\tau\lambda_t - \exp(\tau\lambda_t) + 1)}{\tau\lambda_t^2} \\ &+ \frac{\beta_{3,t} \cdot \exp(-\tau\lambda_t)(\tau^2\lambda_t^2 + \tau\lambda - \exp(\tau\lambda) + 1)}{\tau\lambda_t^2} \end{aligned} \tag{1.25}$$

resulting in the Jacobian matrix $\check{Z}_t$

$$\check{Z}_t = \begin{bmatrix} \iota_{N\times 1} & \Lambda_2(\lambda_t) & \Lambda_3(\lambda_t) & \frac{\partial Z_t}{\partial \lambda_t} \cdot \lambda_t \end{bmatrix} \tag{1.26}$$

with $\iota_{N\times 1}$ being a vector of 1s with dimension $N \times 1$.

Now we derive the equations for $S_t$ and $\nabla_t$. In the NS4F model, the log-likelihood function of $y_t$ is given by

$$\log p(y_t|\theta_t) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|F_t| - \frac{1}{2}(y_t - \Lambda(\lambda_t)\beta_t)' F_t^{-1}(y_t - \Lambda(\lambda_t)\beta_t), \tag{1.27}$$

with vector of scores defined as

$$\nabla(y_t|\theta_t) = \begin{pmatrix} \frac{\partial}{\partial \beta_t}\log p(y_t|\theta_t) \\ \frac{\partial}{\partial \log \lambda_t}\log p(y_t|\theta_t) \end{pmatrix} = \begin{pmatrix} \Lambda(\lambda_t)\beta_t)' F_t^{-1}(y_t - \Lambda(\lambda_t)\beta_t) \\ (y_t - \Lambda(\lambda_t)\beta_t)' \check{\Lambda}(\lambda_t)\beta_t \end{pmatrix} \tag{1.28}$$

where $\check{\Lambda}(\lambda) = \frac{\partial \Lambda(\lambda)}{\partial \log \lambda}$.

The Hessian matrix $\mathcal{H}(y_t|\theta_t)$ of $\log p(y_t|\theta_t)$ is given by

$$\mathcal{H}(y_t|\theta_t) = \frac{\partial^2}{\partial \theta_t \partial \theta_t'}\log p(y_t|\theta_t) = \begin{pmatrix} \frac{\partial^2}{\partial \beta_t \partial \beta_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial \beta_t \partial \log \lambda_t'}\log p(y_t|\theta_t) \\ \frac{\partial^2}{\partial \log \lambda_t \partial \beta_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial (\log \lambda_t)^2}\log p(y_t|\theta_t) \end{pmatrix} \tag{1.29}$$

where

$$\begin{aligned} \frac{\partial^2}{\partial \beta_t \partial \beta_t'}\log p(y_t|\theta_t) &= -\Lambda(\lambda_t)' F_t^{-1}\Lambda(\lambda_t), \\ \frac{\partial^2}{\partial \beta_t \partial \log \lambda_t'}\log p(y_t|\theta_t) &= (y_t - (\Lambda(\lambda_t)\beta_t)' F_t^{-1}\check{\Lambda}(\lambda_t) - (\check{\Lambda}(\lambda_t)\beta_t)' F_t^{-1}\Lambda(\lambda_t) \\ \frac{\partial^2}{\partial (\log \lambda_t)^2}\log p(y_t|\theta_t) &= (y_t - (\Lambda(\lambda_t)\beta_t)' F_t^{-1}\tilde{\Lambda}(\lambda_t)\beta_t - (\check{\Lambda}(\lambda_t)\beta_t)' F_t^{-1}\check{\Lambda}(\lambda_t)\beta_t \end{aligned} \tag{1.30}$$

and $\check{\Lambda}(\lambda) = \frac{\partial \Lambda(\lambda)}{\partial \log \lambda}$ e $\tilde{\Lambda}(\lambda) = \frac{\partial^2 \Lambda(\lambda)}{\partial (\log \lambda)^2}$.

Again, we denote the matrix $S_t$ as a function of the Fischer information matrix. The Fischer information matrix in the NS4F model is derived as:

$$\begin{aligned}
\mathcal{I}_{t|t-1}^{-1} &= -\mathbb{E}[\mathcal{H}(y_t|\theta_t)|\theta_{t|t-1})]^{-1} \\
&= \begin{bmatrix} \Lambda(\lambda_{t|t-1})'F_t^{-1}\Lambda(\lambda_{t|t-1}) & \Lambda(\lambda_{t|t-1})F_t^{-1}(\check{\Lambda}(\lambda_{t|t-1})\beta_{t|t-1})' \\ \beta_t'(\check{\Lambda}(\lambda_{t|t-1}))'F_t^{-1}\Lambda(\lambda_{t|t-1}) & \beta_{t|t-1}'(\check{\Lambda}(\lambda_{t|t-1}))'F_t^{-1}\check{\Lambda}(\lambda_{t|t-1})\beta_{t|t-1} \end{bmatrix}^{-1}
\end{aligned} \tag{1.31}$$

where $\mathcal{H}(y_t|\theta_t)|\theta_{t|t-1}]$ denotes the Hessian matrix and $\check{\Lambda}$ is the derivative of $Lambda(\lambda_t)$ relative to $\log\lambda_t$. The resulting matrix $S_t$ is given by

$$S_t = \begin{bmatrix} \rho_\beta\Lambda(\lambda_{t|t-1})'F_t^{-1}\Lambda(\lambda_{t|t-1}) & 0 \\ 0 & \rho_\lambda\beta_{t|t-1}'(\check{\Lambda}(\lambda_{t|t-1}))'F_t^{-1}\check{\Lambda}(\lambda_{t|t-1})\beta_{t|t-1} \end{bmatrix}^{-1} \tag{1.32}$$

where $\rho_\lambda$ represents a scalar coefficient that only penalizes the factor $\log\lambda_t$.

As an Extended Kalman filter, the update steps are defined as:

$$\begin{aligned}
\theta_{t|t} &= \theta_{t|t-1} + P_{t|t-1}\check{Z}_t'(\check{Z}_tP_{t|t-1}\check{Z}_t' + \Sigma_\epsilon^+)^{-1}(y_t - Z(\theta_{t|t-1})) \\
P_{t|t} &= P_{t|t-1} - P_{t|t-1}\check{Z}_t'(\check{Z}_tP_{t|t-1}\check{Z}_t' + \Sigma_\epsilon^+)^{-1}\check{Z}_tP_{t|t-1}
\end{aligned} \tag{1.33}$$

While the prediction steps are changed as the equation of states changes. The prediction equations for $\theta$ and its estimated variance are now defined as:

$$\begin{aligned}
\theta_{t+1|t} =& \mathbb{E}[\theta_{t+1}|\mathcal{F}_t] \\
=& \mathbb{E}[\omega + \Phi\theta_t + S_t\nabla(y_t|\theta_t)|\mathcal{F}_t] \\
=& \omega + \Phi\theta_{t|t} + S_t\nabla(y_t|\theta_{t|t})P_{t+1|t} \\
\mathbb{V}[\theta_{t+1}|\mathcal{F}_t] =& \mathbb{V}[\omega + \Phi\theta_t + S_t\nabla(y_t|\theta_t)|\mathcal{F}_t] \\
=& \Phi P_{t|t}\Phi' + S_t\mathbb{V}[\nabla(y_t|\theta_t)|\mathcal{F}_t]S_t'
\end{aligned} \tag{1.34}$$

Note that for the expressions $\mathbb{E}[\nabla(y_t|\theta_t)|\mathcal{F}_t]$ and $\mathbb{V}[\nabla(y_t|\theta_t)|\mathcal{F}_t]$ it is not trivial to find an analytical solution due to the fact that the elements in $\nabla(y_t|\theta_t)$ are non-linear in $\theta_t$. Thus, we locally linearize the function $\nabla(y_t|\theta_t)$ in $\theta_t = \theta_{t|t}$. This results

$$\nabla(y_t|\theta_t) = \nabla(y_t|\theta_{t|t}) + \nabla^2(y_t|\theta_{t|t})(\theta_t - \theta_{t|t}) \tag{1.35}$$

where $\nabla^2(y_t|\theta_{t|t}) = \left.\frac{\partial\nabla(y_t|\theta_t)}{\partial\theta_t'}\right|_{\theta_t=\theta_{t|t}}$ and consequently

$$\begin{aligned}
\mathbb{E}[\nabla(y_t|\theta_t)|\mathcal{F}_t] &= \mathbb{E}[\nabla(y_t|\theta_{t|t}) + \nabla^2(y_t|\theta_{t|t})(\theta_t - \theta_{t|t})|\mathcal{F}_t] = \nabla(y_t|\theta_{t|t}) \\
\mathbb{V}[\nabla(y_t|\theta_t)|\mathcal{F}_t] &= \mathbb{V}[\nabla(y_t|\theta_{t|t}) + \nabla^2(y_t|\theta_{t|t})(\theta_t - \theta_{t|t})|\mathcal{F}_t] = \nabla^2(y_t|\theta_{t|t})(\nabla^2(y_t|\theta_{t|t}))'
\end{aligned} \tag{1.36}$$

resulting in the following prediction equations

$$\begin{aligned}
\theta_{t+1|t} &= \omega + \Phi\theta_{t|t} + S_t\nabla(y_t|\theta_{t|t}) \\
P_{t+1|t} &= \Phi P_{t|t}\Phi' + S_t\nabla^2(y_t|\theta_{t|t})P_{t|t}\nabla^2(y_t|\theta_{t|t})'S_t'
\end{aligned} \tag{1.37}$$

## 2.5 Time-Varying Common Volatility

Many extensions of the Diebold and Li (2006) model are considered and treated in the literature. Each amplification or generalization typically relaxes one or two assumptions of the seminal model. One of them concerns the variance of the error terms of the measurement equation. For example, Koopman, Mallee, and Wel (2010) introduce a GARCH(1,1) specification of conditional heterocesticity for dynamic measurement of error variance. Mesters, Schwaab, and Koopman (2014) allow errors to follow a heavy-tailed distribution with stochastic volatility. Caldeira, Laurini, and Portugal (2010), on the other hand, consider specific stochastic volatility for the error terms of the factor transition equation.

At this stage we consider that volatility follows dynamics determined by a single common factor. However, unlike the aforementioned works, we model log-volatility as a time-varying process using GAS-type dynamics. We define the general formulation according to the equations with common time-varying volatility as

$$
\begin{aligned}
y_t &= \Lambda(\lambda) f_t + \epsilon_t \qquad \epsilon_t \sim N(0, \Sigma_t) \\
\Lambda &= \begin{pmatrix} 1 & \frac{1-e^{\lambda\tau}}{\lambda\tau} & \frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \end{pmatrix} \\
\Sigma_t &= \mathrm{diag}(\sigma_1^2 e^{h_t}, ..., \sigma_N^2 e^{h_t})
\end{aligned}
\tag{1.38}
$$

To identify $h_t$, it is necessary to set one of the $\sigma_i^2$ equal to 1. We chose to normalize for the 1-year maturity, $\sigma_{12m}^2 = 1$. Thus, the factor $h_t$ adjusts up or down for the remaining maturities. The choice does not affect the results of the work.

Considering the formulation with three dynamic factors and common stochastic volatility, in which we call this version NSF3-V, the equation of states can be defined as

$$
\begin{bmatrix} f_{t+1} \\ h_{t+1} \end{bmatrix} = \begin{bmatrix} \beta_{1,t+1} \\ \beta_{2,t+1} \\ \beta_{3,t+1} \\ h_{t+1} \end{bmatrix} = \omega + A \begin{bmatrix} f_{t+1} \\ h_{t+1} \end{bmatrix} + B s_t \tag{1.39}
$$

The same can be applied to the NS4F specification, called NSF4-V:

$$
\begin{bmatrix} f_{t+1} \\ \lambda_{t+1} \\ h_{t+1} \end{bmatrix} = \begin{bmatrix} \beta_{1,t+1} \\ \beta_{2,t+1} \\ \beta_{3,t+1} \\ \lambda_{t+1} \\ h_{t+1} \end{bmatrix} = \omega + A \begin{bmatrix} f_{t+1} \\ \lambda_{t+1} \\ h_{t+1} \end{bmatrix} + B s_t \tag{1.40}
$$

For the estimation process, we consider matrices A and B of static parameters of the transition equation as block-diagonals. Consequently, we do not capture spillover effects from volatility factors to the level equation, or vice versa.

The score vector and the Fischer information matrix are derived in this case as follows. The equations for $\nabla_t$ in the NS3F-V and NS4F-V models are given, respectively, by

$$\nabla(y_t|\theta_t) = \begin{pmatrix} \frac{\partial}{\partial\beta_t}\log p(y_t|\theta_t) \\ \frac{\partial}{\partial h_t}\log p(y_t|\theta_t) \end{pmatrix} \quad \text{e} \quad \nabla(y_t|\theta_t) = \begin{pmatrix} \frac{\partial}{\partial\beta_t}\log p(y_t|\theta_t) \\ \frac{\partial}{\partial\log\lambda_t}\log p(y_t|\theta_t) \\ \frac{\partial}{\partial h_t}\log p(y_t|\theta_t) \end{pmatrix} \tag{1.41}$$

where $\nabla_h = \frac{\partial}{\partial h_t}\log p(y_t|\theta_t) = \frac{1}{2}(\epsilon_t' F_t^{-1}\epsilon_t - N)$.

The Hessian matrix $\mathcal{H}(y_t|\theta_t)$ of $\log p(y_t|\theta_t)$ in the NS3F-V model is given by

$$\mathcal{H}(y_t|\theta_t) = \frac{\partial^2}{\partial\theta_t\partial\theta_t'}\log p(y_t|\theta_t) = \begin{pmatrix} \frac{\partial^2}{\partial\beta_t\partial\beta_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial\beta_t\partial h_t'}\log p(y_t|\theta_t) \\ \frac{\partial^2}{\partial h_t\partial\beta_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial(h_t)^2}\log p(y_t|\theta_t) \end{pmatrix} \tag{1.42}$$

while in the NS4F-V model:

$$\begin{aligned} \mathcal{H}(y_t|\theta_t) &= \frac{\partial^2}{\partial\theta_t\partial\theta_t'}\log p(y_t|\theta_t) \\ &= \begin{pmatrix} \frac{\partial^2}{\partial\beta_t\partial\beta_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial\beta_t\partial\log\lambda_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial\beta_t\partial h_t'}\log p(y_t|\theta_t) \\ \frac{\partial^2}{\partial\log\lambda_t\partial\beta_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial(\log\lambda_t)^2}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial\log\lambda_t\partial h'}\log p(y_t|\theta_t) \\ \frac{\partial^2}{\partial h_t\partial\beta_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial h_t\log\lambda_t'}\log p(y_t|\theta_t) & \frac{\partial^2}{\partial h_t h_t'}\log p(y_t|\theta_t) \end{pmatrix} \end{aligned} \tag{1.43}$$

Again, we denote the matrix $S_t$ as a function of the Fischer information matrix, as in the previous specifications. Now, we just enlarge the dimension of the Fischer matrix with the inclusion of the second derivative of $h_t$, $\mathcal{I}_{hh} = -\frac{\partial^2}{\partial h_t h_t'}\log p(y_t|\theta_t) = \frac{N}{2}$ and we include a penalty parameter $\rho_h$.

## 3 Data

In this section, we describe the information used to estimate the models. For the empirical assessment of the yield curve and measurement of the mark-to-market of contracts negotiated for hedging, we used the daily adjustment prices of futures contracts in the DI1 interest rate derivatives market traded on B3 between 02/01/2015 and 31/ 12/2023 ($T$ = 2,227 observations).

The database used consists of a panel of daily time series of the settlement prices of DI futures contracts. The interbank deposit futures contract (DI futures) with maturity $\tau$ is a futures contract whose basic asset is the interest rate DI[5] accrued daily, capitalized between the trading period $t$ and $\tau$. The information provided by the rates negotiated daily in the market reflects the expectation of the average CDI from period 0 to maturity. The value of the contract is defined by its value at maturity, R$ 100,000.00, discounted according to the accumulated interest rate, negotiated between the seller and the buyer.

[5]The DI rate is the average daily rate of Brazilian interbank deposits (loans/loans), calculated by the Custody and Settlements Chamber (CETIP) for all business days. The DI rate, published daily, is expressed in terms compounded annually, based on 252 business days.

When buying a DI futures contract at the DI price at time $t$ and holding it until expiration $\tau$, the gain or loss is given by:

$$100.000\left(\frac{\Pi_{i=1}^{\xi(t,\tau)}(1+y_i)^{\frac{1}{252}}}{(1+DI^*)^{\frac{\xi(t,\tau)}{252}}}-1\right) \tag{1.44}$$

where $y_i$ denotes the DI rate, $(i-1)$ days after the trading day. The function $\xi(t,\tau)$ represents the number of business days between $t$ and $\tau$.

The DI contract is quite similar to the zero-coupon bond, except for the daily payment of marginal adjustments. Every day cash flow is the difference between the current day's settlement price and the previous day's settlement price, indexed by the previous day's DI rate.

DI futures contracts are traded on B3, which determines the number of maturities for authorized contracts. In general, there are about 20 maturities with authorized contracts every day, but not all of them are liquid. Approximately 10 maturities have contracts with greater liquidity. There are contracts with monthly maturities for the months at the beginning of each quarter (January, April, July and October). In addition, there are contracts with expiry dates for the four months following the current month. The expiration date is the first business day of the month in which the contract expires.

We separated the information used into two sets of data. In the case of the information used in estimating the Nelson-Siegel model, we will use the rates interpolated via flat-forward by B3, used as reference rates for SWAP DI x Pre contracts, and then interpolated via cubic splines for 24 fixed maturities. This panel of 24 maturities will compose the vector of observations $y_t(\tau)$. In the case of mark-to-market of the credit portfolio, we will use the curves interpolated in the previous step for maturities from 1 business day to 11,000 business days.

The Table (1.1) presents the descriptive statistics of the Brazilian interest rate curve. For each of the 24 time series we report mean, standard deviation, minimum, maximum and the last three columns contain sample autocorrelations at offsets of 21, 252 and 504 working days. The summary statistics confirm some stylized facts about the Brazilian yield curve: the sample average curve is upward-sloping and concave, volatility decreases with maturity, autocorrelations are very high.

In Figure 2 we present a three-dimensional graph of the data set and illustrate how forward interest levels vary substantially throughout the sample. Although the series vary strongly over time for each of the timeframes, a strong common pattern across the 24 series over time is apparent. It is clear from Figure 2 that not only does the level of the term structure fluctuate over time, but also its slope and curvature.

Table 1.1: Descriptive statistics of SWAP DI x Pre contracts for selected maturities (In %)

| $252 \cdot \tau$ | Mean | Variance | Minimum | Maximum | $\widehat{\rho}(21)$ | $\widehat{\rho}(252)$ | $\widehat{\rho}(504)$ |
|---|---|---|---|---|---|---|---|
| 63 | 9,236 | 4,251 | 1,879 | 14,677 | 0,992 | 0,483 | -0,164 |
| 126 | 9,280 | 4,216 | 1,870 | 15,075 | 0,991 | 0,477 | -0,169 |
| 189 | 9,318 | 4,162 | 1,871 | 15,472 | 0,990 | 0,470 | -0,173 |
| 252 | 9,341 | 4,087 | 1,967 | 15,801 | 0,989 | 0,463 | -0,177 |
| 378 | 9,386 | 3,882 | 2,230 | 16,279 | 0,986 | 0,452 | -0.179 |
| 504 | 9,447 | 3,657 | 2,545 | 16,547 | 0,983 | 0,445 | -0,176 |
| 756 | 9,634 | 3,254 | 3,312 | 16,810 | 0,977 | 0,434 | -0,160 |
| 1.080 | 9,856 | 2,982 | 3,965 | 16,888 | 0,972 | 0,426 | -0,146 |
| 1.260 | 10,730 | 2,801 | 4,566 | 16,921 | 0,968 | 0,418 | -0,138 |
| 1.512 | 10,241 | 2,659 | 5,037 | 16,878 | 0,965 | 0,410 | -0,134 |
| 1.764 | 10,364 | 2,556 | 5,345 | 16,826 | 0,962 | 0,401 | -0,132 |
| 2.016 | 10,479 | 2,480 | 5,687 | 16,800 | 0,959 | 0,392 | -0.131 |
| 2.520 | 10,665 | 2,369 | 6,213 | 16,799 | 0,956 | 0,374 | -0,125 |
| 2.772 | 10,737 | 2,326 | 6,393 | 16,807 | 0,955 | 0,366 | -0.124 |
| 3.024 | 10,795 | 2,296 | 6,486 | 16,826 | 0,953 | 0,358 | -0,122 |
| 3.780 | 10,921 | 2,224 | 6,691 | 16,935 | 0,951 | 0,343 | -0.112 |
| 5.040 | 11,027 | 2,156 | 6,781 | 17,001 | 0,948 | 0,330 | -0,105 |
| 6.300 | 11,092 | 2,118 | 6,835 | 17,000 | 0,946 | 0,319 | -0,110 |
| 7.560 | 11.135 | 2,094 | 6,869 | 17,000 | 0,944 | 0,311 | -0,113 |
| 8.820 | 11,166 | 2,077 | 6,897 | 17,000 | 0,943 | 0,305 | -0,116 |
| 10.080 | 11,190 | 2,065 | 6,914 | 17,000 | 0,942 | 0,301 | -0,118 |

Notes: The table presents descriptive statistics for daily data from SWAP DI x Pre contracts between January 2015 and December 2023.

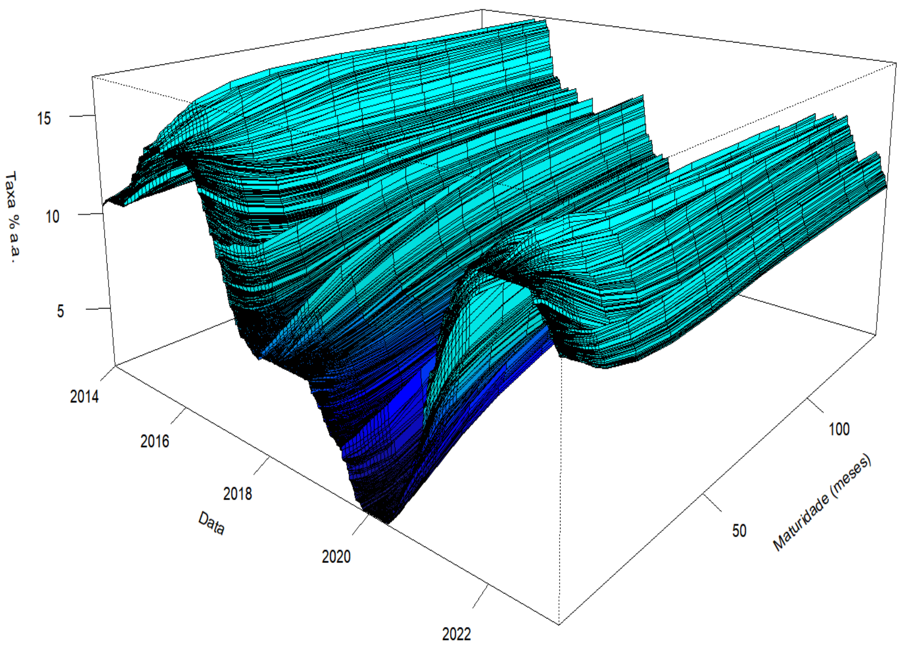


Figure 1.1: This chart details the evolution of the term structure of interest rates (based on DI futures) for the time horizon 2014:03-2023:12. The sample was composed of daily returns for maturities of 21, 42, 63, 84, 105, 126, 147, 168, 189, 210, 231, 252, 273, 294, 504, 756, 1080, 1260, 1512, 1764 , 2016, 2520, 2772 and 3024 business days.

# 4 Results

## 4.1 Dynamic Nelson-Siegel Models Estimates

We present the main results of the Nelson-Siegel models in Table (1.1). To compare the estimates of the Nelson-Siegel models with the NS3F, NS4F, NS3F-V and NS4F-V specifications, we also estimate a *benchmark* version of the model considering fixed $\lambda$, with the factors $\beta_t$ variables in time and without taking into account conditional heteroscedasticity, as originally proposed by Diebold and Li (2006). To compare model fit, the table presents statistics on RMSE measurement errors, defined as the differences between observed interest rates and their estimates

From the RMSE statistics, it can be seen that the NS4F-V model fits better in relation to the other specifications. Thus, it seems clear that treating the parameter $\lambda$ and the covariance matrix $\Sigma_\epsilon$ as fixed, despite facilitating the model estimation procedures, implies a loss in the quality of the fit. This conclusion is also valid if we only consider the inclusion of $\lambda_t$ in the NS4F model compared to the specification *benchmark*.

Table 1.2: Parameter Estimates: Dynamic Nelson-Siegel Extensions

| Param | NS3F cons | NS3F | NS4F | NS3F-V | NS4F-V |
|---|---|---|---|---|---|
| $\lambda$ | 0.587 | - | - | - | - |
| $\omega_1$ | 0.390 | 0.185 | 0.1895 | 0.191 | 0.184 |
| $\omega_2$ | 0.108 | 0.081 | 0.085 | 0.974 | 0.958 |
| $\omega_3$ | -0.018 | -0.025 | 0.038 | -0.025 | -0.028 |
| $\omega_4$ | - | - | -0.407 | - | -0.579 |
| $A_{11}$ | 0.973 | 0.985 | 0.987 | 0.9783 | 0.989 |
| $A_{12}$ | 0.097 | 0.147 | 0.088 | 0.117 | 0.107 |
| $A_{13}$ | 0.019 | 0.031 | 0.025 | 0.032 | 0.028 |
| $A_{14}$ | - | - | -0.670 | - | -0.242 |
| $A_{21}$ | 0.381 | 0.487 | 0.675 | 0.556 | 0.654 |
| $A_{22}$ | -0.091 | -0.121 | -0.105 | -0.112 | -0.081 |
| $A_{23}$ | 0.947 | 0.931 | 0.944 | 0.952 | 0.931 |
| $A_{24}$ | - | - | -0.055 | - | -0.744 |
| $A_{31}$ | 0.094 | 0.087 | 0.093 | 0.084 | 0.080 |
| $A_{32}$ | -0.041 | -0.053 | -0.045 | -0.055 | -0.048 |
| $A_{33}$ | -0.057 | -0.067 | -0.051 | -0.052 | -0.049 |
| $A_{34}$ | - | - | 0.087 | - | 0.073 |
| $A_{41}$ | - | - | 0.015 | - | 0.018 |
| $A_{42}$ | - | - | -0.055 | - | -0.072 |
| $A_{43}$ | - | - | 0.074 | - | 0.058 |
| $A_{44}$ | - | - | 0.587 | - | 0.583 |
| $\rho_\beta$ | - | 0.638 | 0.555 | 0.341 | 0.741 |
| $\rho_\lambda$ | - | - | 30.583 | - | 32.444 |
| RMSE | 0.175 | 0.126 | 0.084 | 0.086 | 0.073 |

Note: RMSE expressed in basis points.

Figure (1.3) presents the estimates of the factors, $(\beta_{1,t}, \beta_{2,t}, \beta_{3,t})$ and $\lambda_t$, which is treated as a fourth latent factor. The evolution of $\beta_{1,t}$ faithfully reflects the level interpretation of this parameter, presenting the dynamics of the average yield curve over time. The evolution of the other factors also adequately captures the evolution of the slope and curvature components of the term structure observed in interest rates. In particular, $\lambda_t$ takes on values between 0.5 and 0.9 at different times.

The stochastic volatility component, presented in Figure (1.4), shows the model's ability to capture the conditional heteroscedasticity existing in interest rates, identifying moments of changes and stress cycles, for example, occurring in 2017 (Joesley-Day), in 2018 (truck drivers' strike), in the first quarter of 2020 (pandemic), and the inflationary surprise in 2021, periods characterized by changes in the direction of the country's monetary and fiscal policy. It is also noted that, when including the volatility factor, the dynamics displayed by the $\lambda_t$ factor changes significantly.

In the analysis carried out using the *in-sample* adjustment of the yield curve, we can conclude that the model extensions estimated with GAS dynamics contribute significantly to improving the adjustment capacity of the Nelson-Siegel model in the dynamic formulation proposed by Diebold and Li (2006). The extension that treats the $\lambda_t$ parameter as a time-variant and multivariate stochastic volatility component (NS4F-V) was the one that presented the most significant gains, encouraging its use in fixed income.

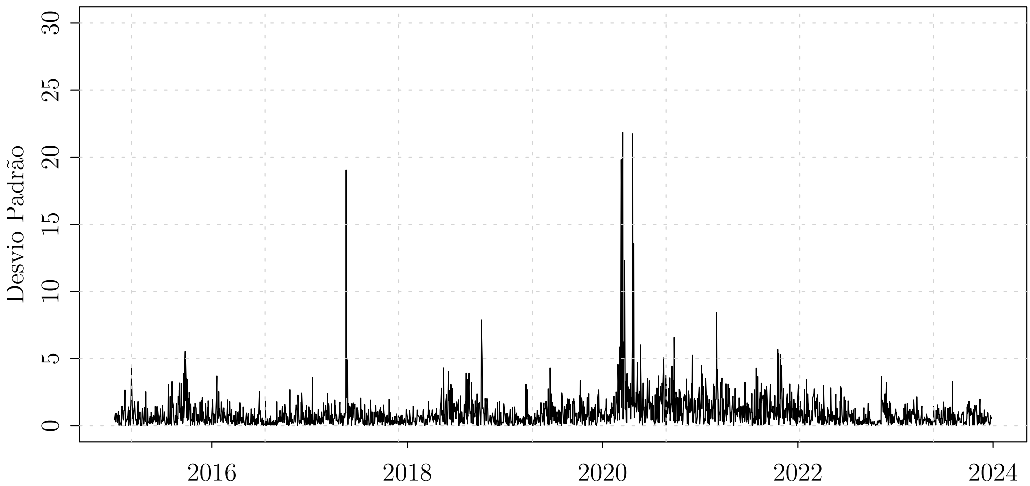


Figure 1.2: Common Stochastic Volatility Factor.
Standard deviations multiplied by 1,000.

## 5 Parametric Immunization of Long-Term Assets

Immunizing a portfolio with fixed income instruments aims to minimize the volatility of a portfolio's results, that is, to reduce losses or financial profits arising from fluctuations in interest rates.

Now suppose that in period $t$ a financial institution has a credit asset with cash flows receivable maturing in the distant future $T_0 = t + \tau_0$. If there are long-term fixed income instruments available, the institution could immunize this exposure by purchasing interest rate derivatives with maturity $\tau_0$. However, in the Brazilian market, there are typically no liquid instruments for maturities such as $\tau_0 > 20$ years, while for average maturities with $10 < \tau_0 < 15$ years, the market is insufficiently liquid . Therefore, the financial institution

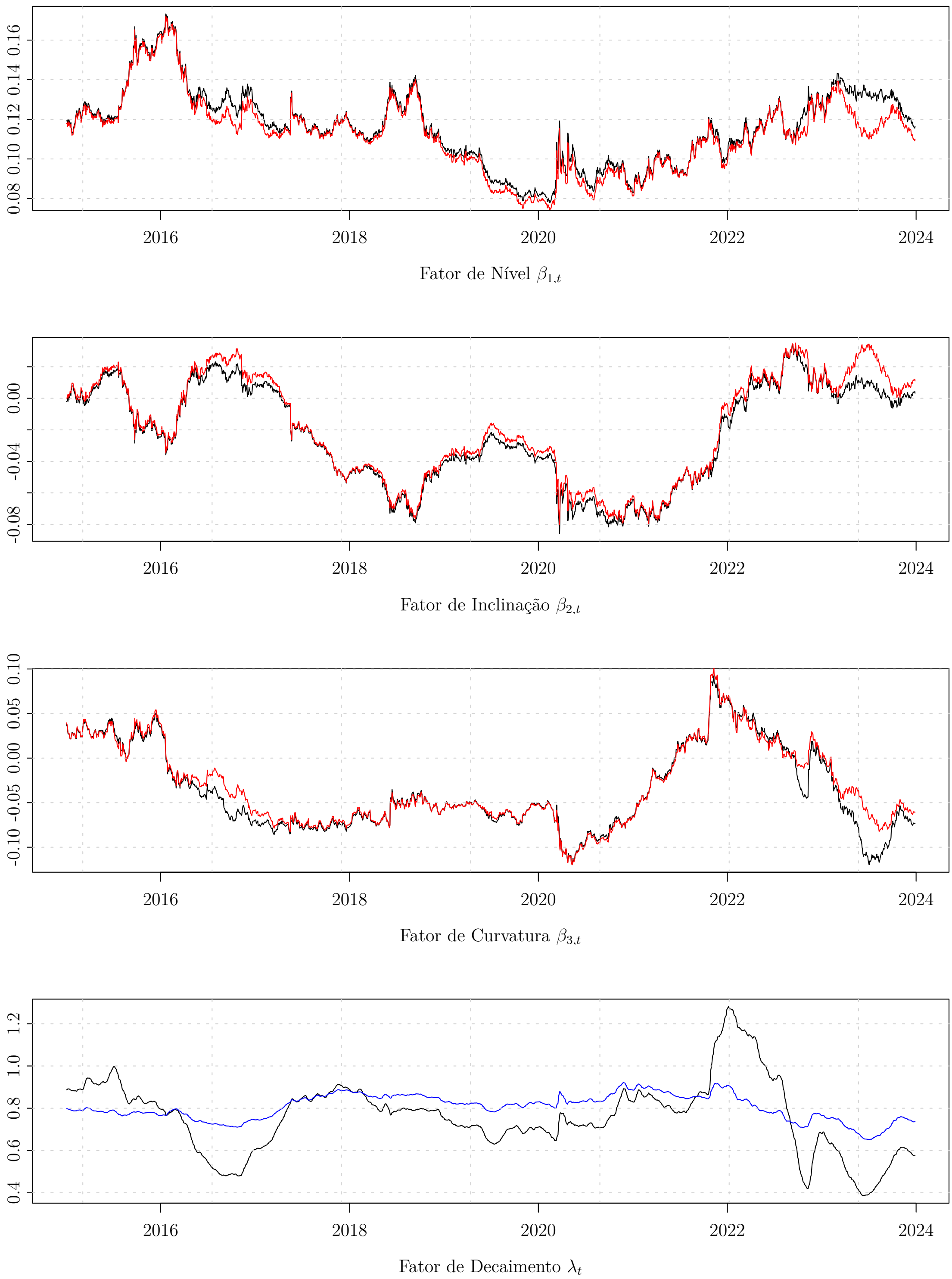


Figure 1.3: Filtered Estimates for the Dynamic Factors of Nelson-Siegel Score-Driven Models. The black series refer to the NS4F specification, while the red and blue series refer to the NS4F-V specification.

needs to hedge its assets using instruments with short maturities $\tau_i < \tau_0$.

The most common hedging technique with fixed income instruments is Hedging via *Duration*. the term *duration* in this case refers to the sensitivity of a security's price to changes in the interest rate, that is, it measures the approximate first-order change in the asset's price given an infinitesimal change in the interest rate curve. In other words, this technique consists of building a portfolio with a duration equal to the maturity of the exposure. For example, for a portfolio of $n$ derivatives with maturities $\tau_i$, the *duration* of a portfolio with weights $w_i$ is given by $D = \sum_i w_i \tau_i$, a weighted average of instrument maturities. Considering the case of an exposure longer than the instruments available on the market, hedging via *duration* inevitably involves leverage, that is, some weights need to be negative. The popularity of the duration technique derives from the property that duration measures measure the relative variation in the value of the portfolio in relation to parallel shocks in the interest rate curve. Thus, *Duration hedging* techniques apply to variations over short intervals. For long intervals, the non-linearity of the price-*yield* relationship introduces other risk factors, such as those presented in the Nelson-Siegel formulation.

## 5.1 Building a Synthetic Credit Portfolio

The empirical exercise consists of selecting the optimal number of DI1 interest derivative instruments (DI-Futuro) with a maturity lower than the *duration* of long-term assets, a combination that will reproduce a mark-to-market (MtM) differential closest to zero in the long term. However, an important step in this process is the selection of the portfolio of instruments, since they will be liquid instruments with a shorter maturity than the exposure to be immunized, a duration that does not have liquid instruments available in the futures interest market. The optimal quantities are defined by the partial derivative of the price function with respect to each of the three dynamic factors. The selection of contracts will follow an optimization problem as will be demonstrated below.

The bank credit portfolio will begin on January 3, 2018 with 30 active credit operations chosen at random, with future values between R$ 1 thousand and R$ 10 million and maturity between 1,764 and 10,080 business days. The choice of terms aims to be in line with the banking reality of portfolios where the granting of bank credit has a longer term. All operations will be pre-fixed based on the interest curve negotiated on the day the operation begins. On the first business day of each observed month, the credit portfolio will receive 30 new transactions in the same format as the initial portfolio. The portfolio rebalancing dynamics will take place on the first business day of each month. Hedging and rebalancing costs will be taken into account, aiming for greater adherence to reality[6]. The synthetic portfolio will be considered *risk-free*, without liquidity restrictions, aiming to focus on immunizing only the market risk related to the Brazilian interest rate. As mentioned previously, mark-to-market will follow the SWAP DI x Pre interest rate curve provided by B3, interpolated by Cubic Splines. Below we present some relevant information

[6]For example, to adapt to each rebalancing, we will need to sell a certain amount of contracts purchased in previous periods.

about the simulated portfolio to be immunized.

Table 1.3: Simulated Credit Portfolio

| Reference | Portfolio Duration | Portfolio MtM (R$) | Aditional Notional (R$) | Aditional Delta EVE (R$) |
|---|---|---|---|---|
| jan/18 | 25,4 | 1.606.078 | 34.557.693 | 914.551 |
| feb/18 | 26,1 | 3.049.588 | 22.221.892 | 1.045.259 |
| may/18 | 26,4 | 4.922.364 | 33.815.633 | 1.081.722 |
| apr/18 | 21,2 | 6.505.109 | 19.500.287 | 1.099.333 |
| may/18 | 23,6 | 7.657.324 | 21.103.889 | 1.024.947 |
| jun/18 | 19,9 | 7.940.623 | 31.203.690 | 669.206 |
| jul/18 | 23,6 | 7.828.185 | 42.268.446 | 673.982 |
| aug/18 | 23,7 | 9.559.316 | 33.015.008 | 724.276 |
| sep/18 | 23,1 | 10.885.290 | 50.921.994 | 529.584 |
| oct/18 | 25,1 | 10.280.534 | 35.386.349 | 632.856 |
| nov/18 | 23,8 | 14.151.054 | 28.019.962 | 996.623 |
| dec/18 | 20,8 | 21.154.876 | 17.591.006 | 1.059.101 |
| jan/19 | 23,2 | 23.814.972 | 20.549.964 | 1.217.899 |
| feb/19 | 24,9 | 28.029.502 | 15.499.070 | 1.375.970 |
| mar/19 | 25,1 | 31.793.238 | 19.871.213 | 1.264.659 |
| apr/19 | 24,7 | 31.552.550 | 21.584.261 | 1.325.014 |
| may/19 | 23,2 | 34.262.692 | 17.310.984 | 1.286.280 |
| jun/19 | 22,3 | 35.679.576 | 14.911.449 | 1.520.729 |
| jul/19 | 21,2 | 43.618.765 | 13.612.499 | 1.992.250 |
| aug/19 | 22,6 | 55.410.825 | 14.228.876 | 2.026.571 |
| sep/19 | 23.8 | 57.643.630 | 12.922.493 | 1.900.792 |
| oct/19 | 22,2 | 57.055.574 | 9.524.465 | 2.150.220 |
| nov/19 | 22,3 | 65.214.207 | 8.866.039 | 2.502.191 |
| dec/19 | 22,7 | 74.082.082 | 7.920.247 | 2.159.402 |
| jan/20 | 21,8 | 67.202.151 | 8.613.927 | 2.044.489 |
| feb/20 | 21,9 | 66.727.335 | 12.464.579 | 2.041.442 |
| mar/20 | 21,0 | 68.447.788 | 8.806.426 | 2.039.703 |
| apr/20 | 21,0 | 68.234.007 | 8.824.703 | 1.449.894 |
| may/20 | 21,7 | 53.695.835 | 11.970.016 | 1.444.197 |
| jun/20 | 20,8 | 55.542.953 | 11.562.022 | 1.631.959 |
| jul/20 | 21,8 | 62.903.699 | 10.541.131 | 1.814.881 |
| aug/20 | 22,3 | 70.105.703 | 7.494.657 | 2.021.032 |
| sep/20 | 20,5 | 76.862.909 | 9.246.934 | 1.694.282 |
| oct/20 | 22,0 | 67.711.413 | 20.525.746 | 1.400.461 |
| nov/20 | 23,9 | 59.988.575 | 16.299.765 | 1.381.919 |
| dec/20 | 22,4 | 59.622.839 | 16.044.168 | 1.452.941 |
| jan/21 | 22,8 | 63.216.974 | 10.947.406 | 1.846.935 |
| feb/21 | 21,5 | 75.776.836 | 12.677.008 | 1.703.750 |
| mar/21 | 20,7 | 69.715.425 | 9.203.812 | 1.363.475 |
| apr/21 | 19,2 | 58.107.226 | 17.687.705 | 1.154.445 |
| may/21 | 20,4 | 51.225.400 | 15.083.129 | 1.147.206 |
| jun/21 | 19,2 | 51.611.452 | 17.403.410 | 1.159.684 |
| jul/21 | 20,3 | 52.081.313 | 23.382.594 | 1.129.774 |
| aug/21 | 22,4 | 50.772.127 | 27.629.777 | 980.182 |
| sep/21 | 22,9 | 45.577.864 | 43.130.283 | 783.610 |
| oct/21 | 23,2 | 38.475.722 | 44.893.242 | 705.347 |
| nov/21 | 21,3 | 35.162.254 | 40.298.847 | 545.384 |
| dec/21 | 19,9 | 29.245.986 | 23.496.531 | 692.802 |
| jan/22 | 21,1 | 35.010.179 | 21.509.384 | 751.246 |
| feb/22 | 19,9 | 37.188.588 | 44.210.647 | 686.361 |
| mar/22 | 20,9 | 34.904.989 | 44.253.614 | 668.575 |
| apr/22 | 20,7 | 34.118.048 | 30.678.308 | 729.621 |
| may/22 | 21,5 | 36.094.823 | 40.771.215 | 598.941 |
| jun/22 | 22,5 | 31.229.550 | 41.835.465 | 563.702 |
| jul/22 | 21,1 | 29.879.854 | 53.610.996 | 518.272 |
| aug/22 | 20,5 | 28.325.375 | 61.008.676 | 536.059 |
| sep/22 | 21,1 | 29.408.289 | 34.569.031 | 633.090 |
| oct/22 | 20,9 | 33.368.284 | 22.160.436 | 678.218 |
| nov/22 | 19,7 | 35.119.921 | 24.991.686 | 669.352 |
| dez/22 | 19,2 | 34.444.666 | 55.201.943 | 529.718 |

Notes: $Duration_t = \frac{\sum_{t=1}^{n} \tau \cdot \frac{F_t}{(1+i)^\tau}}{\sum_{t=1}^{n} \frac{F_t}{(1+i)^\tau}}$.

Delta EVE is a measure of interest rate risk used by financial institutions around the world, it measures the worst loss in terms of present value in the face of a shock of 4% in the interest curve (in the case of fixed rates it is the High shock).

## 5.2 Factor Hedge Portfolio Optimization

Since we are interested in reducing the risk of a portfolio, we should look at portfolio returns, as suggested in Litterman and Scheinkman (1991). Therefore, for the portfolio optimization stage, we will transform the model structure to focus on excess returns instead of fees. In this case, excess returns have the same factor structure as rates.

To get the excess returns, we first need to get the logarithm of prices as

$$\log(P_t(\tau)) = p_t(\tau) = \log(e^{-\tau y_t(\tau)}) = -\tau y_t(\tau) \tag{1.45}$$

and define the returns as

$$r_{t+1,i} = (p_{t+1,\tau_i} - p_{t,\tau_i+1})/p_{t,\tau_i+1} \tag{1.46}$$

where, for short horizons (1 day, for example), we can approximate the above returns as

$$r_{t+1,i} \approx \log \frac{p_{t+1,\tau_i}}{p_{t,\tau_i}} = -\tau_i \Delta y_{t+1,\tau_i} \tag{1.47}$$

where $\Delta$ is the first difference operator, i.e., $\Delta y_{t+1,\tau_i} = y_{t+1,\tau_i} - y_{t,\tau_i}$ is the variation of the maturity *yield* $\tau_i$.

The excess returns of an instrument of maturity $\tau_i$ in period $t+1$ in relation to the risk-free rate is defined as

$$\begin{aligned} r_{t+1}(\tau) &= \log(\exp(-\tau_i \cdot y_{t+1,\tau_i}) - \log(\exp(-\tau_i \cdot y_{t,\tau_i}) - \log(\exp(-\tau_1 \cdot y_{t,\tau_1})) \\ &= -\tau_i \cdot (y_{t+1,\tau_i} - y_{t,\tau_i}) - \frac{y_t(1)}{252} \end{aligned} \tag{1.48}$$

so that the above returns can be calculated as the difference of the logarithm of prices multiplied by 100, and the rate *risk-free* $y_t(1)/252$, is the 1-day CDI.

Based on market data, it is known that the volatility of returns increases almost linearly with maturity. The definition $p_t(\tau) = -\tau y_t(\tau)$ also suggests that returns are proportional to maturity if *yields* move in parallel. Therefore, we weight returns by maturity, i.e., we define the excess returns weighted by maturity as $\rho_t(\tau) = r_t(\tau)/\tau$.

For the remainder of this section we assume that the factors $\beta_t$ and $\lambda_t$ are known. Working with known factors is common in immunization processes that use the Nelson-Siegel model or when factors are obtained through principal components techniques (PCA) based on historical returns. An example is the seminal work by Litterman and Scheinkman (1991) who applies hedging techniques using a three-factor model estimated by PCA.

The construction of the DI Futuro derivatives portfolio for immunization is described as follows. Once the estimated $\beta_t$ and $\lambda_t$ for SWAP DI x Pre yield curve data are known, we consider the following reformulation of the factor model:

$$\begin{bmatrix} \rho_{0t} \\ \rho_t \end{bmatrix} = \begin{bmatrix} b_{0t} \\ b_t \end{bmatrix} f_t + \begin{bmatrix} \epsilon_{0t} \\ \epsilon_t \end{bmatrix} \tag{1.49}$$

where the matrix $b_t$ of dimension $(n \times k)$ contains the loadings for the net maturities eligible for hedging, which use estimated values of $\lambda_t$ from the NS3F and NS4F models, and $b_0$ is the vector $(1 \times k)$ of factor loadings calculated each period $t$ for maturity *target*, $b_{0,t} = \Lambda_{t,0}(\lambda_t)/\tau_0$. Note that $b_t$ is also a matrix of *loadings* obtained through the estimated values of $\lambda_t$ and depends on the maturity of the derivatives to be selected at each exposure renewal. We have established that the set of liquid instruments for hedging will consist of the 14 available contracts with a maturity of more than 252 business days at each exposure renewal. The rows of the matrix $b_t$ are denoted as $b_{t,i}$, and are defined as $b_{t,i} = \Lambda_{t,i}(\lambda_t)/\tau$ to make it compatible with the weighting of excess returns. The vectors $\rho_{0t}$ and $\rho_t$ follow the same rationale described for $b_{0t}$ and $b_t$. In the case of $\rho_{0t}$, we calculated based on information from the SWAP DI x Pre reference curves, used by financial institutions for marking to market. The residues $\epsilon_t$ and $\epsilon_{0t}$ collect the respective prediction error terms and have a covariance matrix $\sigma^2 I$, in the case of specifications with constant volatility.

Due to the $\epsilon_{0t}$ and $\epsilon_t$ errors, it is no longer possible to obtain a 100% perfect hedge. Since the excess returns of long assets $\rho_0$ are exposed to the same risk factors as market-traded instruments, $\rho_t$, a portfolio holding a portfolio $w$ with the same target exposure $B_0$[7], i.e. $w^{'}B = B_0$, this portfolio will immunize all risk factors. To define $w$, we consider Litterman and Scheinkman (1991)'s concept of hedge construction. In practice, we consider the construction of a portfolio of excess returns of liquid instruments that present better replication of excess returns *target* through a problem of minimizing hedging errors.

Furthermore, since we use maturity-weighted excess returns, we need to adjust the scale of returns as

$$\widehat{r}_{0t} = w'r_t \qquad \text{with} \qquad w_i = g_i\tau_0/\tau_i \tag{1.50}$$

where the weight vector $g_t$ of dimension $n \times 1$ determines the weights of the hedge portfolio. As in Diebold and Li (2006), the portfolio $w$ presents the same Generalized *Duration* $B_0$ of the asset *target* (i.e. a perfect hedge can be constructed if we find weights $w_i$ that result in $\sum_i w_i B_i = B_0$).

We can relax the restriction that the hedge portfolio has the same Generalized *Duration* as the asset *target*. As before, let $g_t$ be the vector of maturity-weighted weights of the hedge portfolio, and let $w$ be the portfolio weights. We choose $w$ (or equivalently $g$) as the vector that minimizes the squares of the residuals of the hedging errors, that is,

$$\min_w \mathbb{E}\Big[(r_{0t} - w^{'}r_t)^2\Big] = \tau_0^2 \min_g \mathbb{E}\Big[(\rho_{0t} - g^{'}\rho_t)^2\Big] \tag{1.51}$$

[7]We can define the Generalized Duration vector of the target exposure as

$$B_0 = \sum_{h=1}^{H} \frac{p_{t,\tau_h} c_h}{\sum_{h=1}^{H} p_{t,\tau_h} c_h} \tau_h b_h$$

, where $c_h$ are payment flows to be received in $\tau_h$ terms and $b_h$ are the loadings for the respective terms, $h =, ..., H$.

where the hedging error vector is given by

$$\widehat{\epsilon}_{0t} = \rho_{0t} - g^{'}\rho_t = \varepsilon_{0t} - g^{'}\varepsilon_t + (b_0 - g^{'}b)f_t \tag{1.52}$$

and has three components: (i) the non-immunized error term $\epsilon_{0t}$; (ii) idiosyncratic noise in the cross-section of returns $g^{'}\epsilon_t$, and (iii) a bias that depends on the performance of the hedge portfolio in immunizing exposure to risk factors. Given $f_t$, we can estimate $\Sigma = \mathbb{E}[f_t f_t^{'}]$, such that the squares of the hedge residuals have expectation

$$\mathbb{E}[\widehat{\epsilon}_{0t}^2] = \sigma^2(1 + g^{'}g) + (b_0 - g^{'}b)\Sigma(b_0 - g^{'}b)^{'} \tag{1.53}$$

where minimization of the above equation with respect to $g_t$ returns the optimal predictor (MSE-optimal)

$$g = (b\Sigma b^{'} + \sigma^2 I)^{-1} b\Sigma b_0^{'} \tag{1.54}$$

. where we calculate the weights over time using the estimates from the NS models. When $\lambda$ and $\sigma^2$ vary over time, the hedge portfolio also becomes time-varying, even if the factor loadings remain stable. Therefore, we can use estimates $\hat{b}$, $\widehat{\Sigma} = \widehat{\mathbb{E}}[f_t f_t^{'}] = \frac{1}{T}\sum_t \widehat{f}_t \widehat{f}_t^{'} - \widehat{\sigma}^2(b^{'}b)^{-1}$ and $\hat{\sigma}^2$ estimated from NS3F, NS4F, NS3F-V and NS4F-V to find in each case a vector $g_t$ of optimal hedge weights.

## 5.3 Hedge Results

In this step we present the hedge performance results using the estimates from the models presented in the previous section. With the estimates of $\beta_t$, $\lambda_t$ and the variance $\sigma_t$, it is possible to use the equations on excess returns to find the optimal portfolios that replicate the synthetic credit portfolio exposures over time. We calculate the optimized portfolios for each of the NS3F, NS4F, NS3F-V, and NS4F-V specifications for each time period and evaluate the hedging errors produced by each model. According to Table (1.4), models that consider $\lambda_t$ and time-varying volatility perform better in terms of RMSE.

Table 1.4: Hedging Performance

| | *benchmark* | NS3F | NS3F-V | NS4F | NS4F-V |
|---|---|---|---|---|---|
| Bias | -0.06 | -0.05 | -0.03 | -0.03 | -0.01 |
| StDev | 1.53 | 1.49 | 1.35 | 1.24 | 1.18 |
| RMSE | 1.53 | 1.49 | 1.35 | 1.24 | 1.18 |

Notes: The above measures are obtained from the difference between the predicted values of excess returns of the synthetic portfolio and the observed values of $\rho_{0,t}$.

Figure 1.4 presents some general results of the composition of portfolios derived from the NS4F-V specification. The optimal portfolios constructed in each period $t$ are more heavily weighted in the long part, and to a lesser extent in the short part. To control for other risk factors, some contracts located at the middle of the curve have a weight equal to zero or are negative. We assume there are no liquidity restrictions and use all available contracts on the day of each monthly recalibration. Another point to be considered is that DI Future contracts are not split, and this worsens the fit to the data.

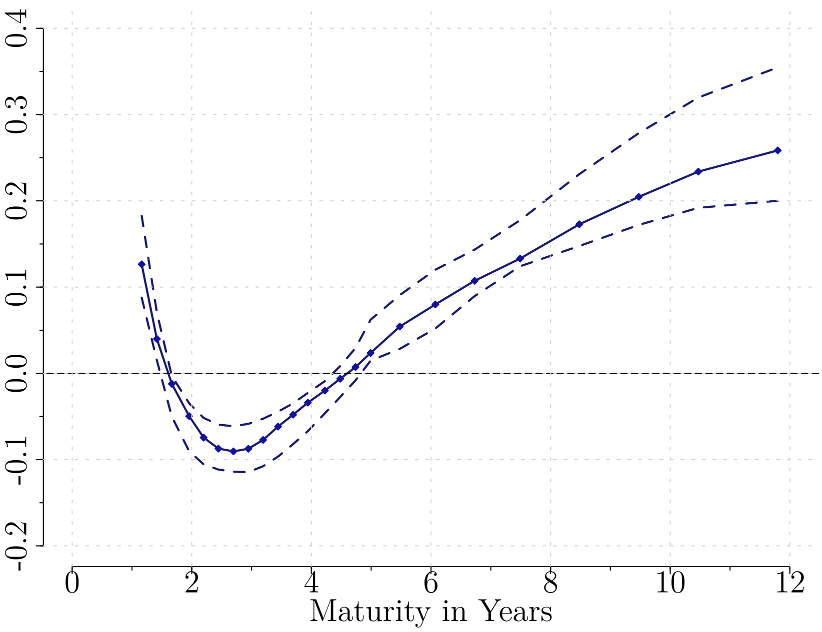


Figure 1.4: Portfolio NS4F-V: The figure presents the median of the optimal weights calculated for the respective median maturities of the futures contracts used. The dotted lines are the 5% and 95% percentiles of the weight distribution.

## 6 Final Remarks

In this article we implement extensions to the Nelson-Siegel family term structure model. Using data from futures interest market contracts traded on B3, we estimate the Nelson-Siegel model with dynamic factors and common volatility varying over time. We follow the structure proposed by Koopman, Mallee, and Wel (2010), and model the decay parameter $\lambda$ as a time-varying parameter. We also consider a common stochastic volatility component for measurement errors.

Since the model is non-linear, we propose the use of the Kalman-GAS Filter to estimate the dynamic parameters of interest. We present the analytical derivations of the Nelson-Siegel with the GAS methodology in state space format, which can be called Score-Driven Nelson-Siegel. The great advantage of this methodology is that the likelihood function of the state space model with time-varying parameters is available in closed form, which facilitates the estimation of parameters using maximum likelihood and substantially reduces the computational cost in relation to other econometric techniques. This structure allows capturing the volatility of the yield curve in a flexible and yet parsimonious way.

The results of our estimations provide evidence for improved predictive power with the modeling of $\lambda_t$ and heteroscedasticity as time-varying parameters following GAS dynamics, especially the quality of fit obtained for longer maturities. This was already expected, according to the structure of the parametric model, which depends on $\lambda$, and is in line with other evidence highlighted in the literature, for example, in Caldeira, Laurini, and Portugal (2010). However, the use of this methodology for the Brazilian yield curve is new.

We demonstrate the application of the Nelson-Siegel Dynamic model with stochastic volatility for an empirical hedging exercise. Our attention is focused on how the treatment of $\lambda$ and volatility as time-varying factors is capable of improving the predictive quality of maturities that do not have liquid instruments available for immunization. Using the results of different Nelson-Siegel specifications estimated with the Kalman-GAS Filter, we were able to find optimized portfolios and satisfactorily immunize a synthetic pre-fixed credit portfolio with a maturity greater than 20 years.

The results presented in this work are of great interest to the financial market as they indicate potential for efficiency gains in long-term hedging operations. Currently, there are no liquid instruments available to immunize pre-fixed portfolios, and this leads institutions to use swap operations with other indexers, which increases the final cost of an operation.

As for future applications, the present study can be used as a basis for a series of applications, such as the immunization of 'green' long term liabilities, for long liabilities of pension funds and for hedging with inflation-indexed bonds. Methodologically, some modeling challenges can be refined and applied to the Brazilian market. The first of these is to consider the expansion of volatility factors. As in Koopman, Lucas, and Zamojski (2017), one could estimate the stochastic volatility factors for each of the maturities of the vector of observables, but this would significantly increase the dimensionality of the matrix system. Another innovation is to consider stochastic volatility in the covariance matrix of the transition vector, as in Caldeira, Laurini, and Portugal (2010). The second refinement, in line with good practices in the yield curve literature, is to test the application of the methodology for the Svensson specification and for those specifications that consider macroeconomic factors. Finally, the analytical properties of the Kalman-GAS Filter allow other probability distributions to be considered, as in Mesters, Schwaab, and Koopman (2014) and Koopman, Lucas, and Zamojski (2017), something not yet considered in modeling the Brazilian yield curve.

It is also important to note that the results obtained for the Nelson-Siegel models are valid for the in-sample adjustment, and an analysis for out-of-sample results could obtain different results, favoring more parsimonious specifications, but we emphasize that the focus of the present work was verify which structure is most appropriate for the behavior observed in this interest curve and not directly make predictions outside the sample of this curve. Therefore, an out-of-sample prediction exercise comparing the Kalman-GAS Filter with other methodologies in the literature is future work to be considered.

## References


ALMEIDA, Caio; LUND, Bruno. Immunization of fixed-income portfolios using an exponential parametric model. **Brazilian Review of Econometrics**, v. 34, n. 2, p. 155–201, 2014.

ALMEIDA, Caio et al. Does curvature enhance forecasting? **International Journal of Theoretical and Applied Finance**, World Scientific, v. 12, n. 08, p. 1171–1196, 2009.

ANDRADE ALVES, Cassio Roberto de; JOSEPH ABRAHAM, Kuruvilla; POLETTI LAURINI, Márcio. Can Brazilian Central Bank communication help to predict the yield curve? **Journal of Forecasting**, Wiley Online Library, v. 42, n. 6, p. 1429–1444, 2023.

CALDEIRA, João et al. Forecasting the yield curve: the role of additional and time-varying decay parameters, conditional heteroscedasticity, and macro-economic factors. **Forecasting the Yield Curve: The Role of Additional and Time-Varying Decay Parameters, Conditional Heteroscedasticity, and Macro-Economic Factors (July 31, 2023)**, 2023.

CALDEIRA, Joao F; LAURINI, Márcio P; PORTUGAL, Marcelo S. Bayesian inference applied to dynamic Nelson-Siegel model with stochastic volatility. **Brazilian Review of Econometrics**, v. 30, n. 1, p. 123–161, 2010.

CALDEIRA, João F; MOURA, Guilherme V; SANTOS, André AP. Bond portfolio optimization using dynamic factor models. **Journal of Empirical Finance**, Elsevier, v. 37, p. 128–158, 2016.

CALDEIRA, Joao F; MOURA, Guilherme V; SANTOS, André AP. Predicting the yield curve using forecast combinations. **Computational Statistics & Data Analysis**, Elsevier, v. 100, p. 79–98, 2016.

CORDEIRO, Werley da Costa et al. Classe de modelos Nelson-Siegel com parâmetros variando no tempo: ajuste e previsão da estrutura a termo da taxa de juros, 2019.

CREAL, Drew; KOOPMAN, Siem Jan; LUCAS, André. A general framework for observation driven time-varying parameter models. Tinbergen Institute Discussion paper, 2008.

________. Generalized autoregressive score models with applications. **Journal of Applied Econometrics**, Wiley Online Library, v. 28, n. 5, p. 777–795, 2013.

DIEBOLD, Francis X; LI, Canlin. Forecasting the term structure of government bond yields. **Journal of econometrics**, Elsevier, v. 130, n. 2, p. 337–364, 2006.

DIEBOLD, Francis X; RUDEBUSCH, Glen D. **The Dynamic Nelson-Siegel Approach to Yield Curve Modeling and Forecasting**. [S.l.]: Princeton, NJ: Princeton University Press, 2011.

DIEBOLD, Francis X; RUDEBUSCH, Glenn D; ARUOBA, S Boragan. The macroeconomy and the yield curve: a dynamic latent factor approach. **Journal of econometrics**, Elsevier, v. 131, n. 1-2, p. 309–338, 2006.

FERNANDES, Marcelo; NUNES, Clemens; REIS, Yuri. What drives the nominal yield curve in Brazil? **Brazilian Review of Econometrics**, v. 40, n. 2, p. 267–284, 2020.

FERNANDES, Marcelo; THIELE, Eduardo. The macroeconomic determinants of the term structure of inflation expectations in Brazil. Sociedade Brasileira de Econometria, 2015.

FRANCISCANGELO, João Gabriel Costa. **Vantagens e desvantagens do modelo dinâmico de Nelson-Siegel: aplicação ao mercado brasileiro**. 2015. PhD thesis.

HARVEY, Andrew C. **Dynamic models for volatility and heavy tails: with applications to financial and economic time series**. [S.l.]: Cambridge University Press, 2013. v. 52.

________. Forecasting, structural time series models and the Kalman filter. Cambridge university press, 1990.

HAUTSCH, Nikolaus; OU, Yangguoyi. Analyzing interest rate risk: Stochastic volatility in the term structure of government bond yields. **Journal of Banking & Finance**, Elsevier, v. 36, n. 11, p. 2988–3007, 2012.

HAUTSCH, Nikolaus; YANG, Fuyu. Bayesian inference in a stochastic volatility Nelson–Siegel model. **Computational Statistics & Data Analysis**, Elsevier, v. 56, n. 11, p. 3774–3792, 2012.

HO, Thomas SY. Key Rate Durations. **The Journal of Fixed Income**, Portfolio Management Research, v. 2, n. 2, p. 29–44, 1992.

KOOPMAN, Siem Jan; LUCAS, Andre; ZAMOJSKI, Marcin. **Dynamic term structure models with score-driven time-varying parameters: estimation and forecasting**. [S.l.], 2017.

KOOPMAN, Siem Jan; MALLEE, Max IP; WEL, Michel Van der. Analyzing the term structure of interest rates using the dynamic Nelson–Siegel model with time-varying parameters. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 28, n. 3, p. 329–343, 2010.

LAURINI, Márcio Poletti; HOTTA, Luiz Koodi. Bayesian extensions to diebold-li term structure model. **International Review of Financial Analysis**, Elsevier, v. 19, n. 5, p. 342–350, 2010.

________. Modelos de fatores latentes generalizados para curvas de juros em múltiplos mercados. Insper, 2009.

LITTERMAN, Robert B; SCHEINKMAN, Jose. Common factors affecting bond returns. **The journal of fixed income**, Institutional Investor Journals Umbrella, v. 1, n. 1, p. 54–61, 1991.

MEIRELLES, Sofia Kusiak; FERNANDES, Marcelo. Estratégias de Imunização de Carteiras de Renda Fixa no Brasil. **Brazilian Review of Finance**, v. 16, n. 2, p. 179–219, 2018.

MESTERS, Geert; SCHWAAB, Bernd; KOOPMAN, Siem Jan. A dynamic yield curve model with stochastic volatility and non-Gaussian interactions: an empirical study of non-standard monetary policy in the euro area. Tinbergen Institute Discussion Paper 14-071/III, 2014.

NELSON, Charles R; SIEGEL, Andrew F. Parsimonious modeling of yield curves. **Journal of business**, JSTOR, p. 473–489, 1987.

PINHEIRO, Felipe; ALMEIDA, Caio; VICENTE, José. Um modelo de fatores latentes com variáveis macroeconômicas para a curva de cupom cambial. **Revista Brasileira de Finanças**, Sociedade Brasileira de Finanças, v. 5, n. 1, p. 79–92, 2007.

QUAEDVLIEG, Rogier; SCHOTMAN, Peter. Hedging long-term liabilities. **Journal of Financial Econometrics**, Oxford University Press, v. 20, n. 3, p. 505–538, 2022.

SANTOS, Daiane Rodrigues dos; RIBEIRO, Tiago Costa; SANFINS, Marco Aurélio. Estimating the Level of the Brazilian Yield Curve Using the Time-Varying Coefficient Model GAS (2, 2) with Gamma Distribution. **International Journal of Economics and Finance**, Canadian Center of Science and Education, v. 11, n. 9, p. 1–1, 2019.

SVENSSON, Lars EO. **Estimating and interpreting forward interest rates: Sweden 1992-1994**. [S.l.]: National bureau of economic research Cambridge, Mass., USA, 1994.

VIEIRA, Fausto; FERNANDES, Marcelo; CHAGUE, Fernando. Forecasting the Brazilian yield curve using forward-looking variables. **International Journal of Forecasting**, Elsevier, v. 33, n. 1, p. 121–131, 2017.

VLEUTEN, Pieter van der; LANGE, Rutger-Jan; WEL, Michel van der. Proximal-Parameter Updates in a Score-Driven Dynamic Nelson-Siegel Framework, 2023.