Regime-dependent dynamics of the brazilian yield curve

Milene Maiser Moraes, [a]
João F. Caldeira[†], [b]

[a]Programa de Pós-Graduação em Economia da Universidade de Brasília
[b]Universidade Federal de Santa Catarina

Abstract This article investigates the dynamics of the Brazilian term structure of yields using a factor model with smooth transition. Latent yield curve factors are extracted through Principal Component Analysis from DI futures rates obtained from B3, covering the period from December 2005 to July 2024. Their temporal dynamics are modeled within a multivariate Smooth Transition Autoregressive framework, allowing both the conditional mean and the autoregressive coefficients to vary smoothly as a function of an observable macroeconomic variable. Among the candidate transition variables, the Selic rate provides evidence of nonlinearity and is selected as the transition variable. Results indicate that factor dynamics differ across regimes associated with lower and higher levels of the policy interest rate, with relevant changes in persistence and cross-factor interactions. The estimated LSTAR specification exhibits strong in-sample fit across maturities, including longer-term ones, suggesting that incorporating smooth transitions improves the empirical representation of the Brazilian term structure.

Keywords: Term structure; Yield curve; STAR models; Smooth transition.

JEL codes: C53, E43, G17.

1. Introduction

The dynamics of the term structure of interest rates contain information about macroeconomic expectations, monetary policy, and cyclical economic conditions, playing a central role in asset pricing, risk management, and monetary policy design. It is of central importance to policymakers, who rely on the yield curve as a signal of financial conditions, as well as to private-sector agents, including financial institutions, firms, and investors (Bie et al., 2024; Kim and Orphanides, 2007). An extensive literature shows that most of the variation in the yield curve can be summarized by a small number of latent factors (interpreted as level, slope, and curvature). Nevertheless, traditional linear models for the dynamics of these factors may be overly restrictive, as they impose constant parameters and disregard potential behavioral changes associated with the business cycle.

milenemoraes2000@hotmail.com
[†]emaildocaldeira@gmail.com

Revista Brasileira de Finanças (Online) Rio de Janeiro, Vol. none, 2025. ISSN 1679-0731, ISSN online 1984-5146

Within the class of linear factor models, the Nelson–Siegel model (Nelson and Siegel (1987)) and its dynamic version proposed by Diebold and Li (2006) (hereafter DNS) constitute a fundamental benchmark in the study of the term structure of interest rates and are widely used in the literature. The DNS model assumes that three latent factors evolve smoothly over time, however, empirical evidence from different economies suggests that the evolution of these factors is not always smooth, but instead subject to structural breaks and regime changes potentially associated with macroeconomic conditions such as inflationary shocks, shifts in monetary policy, and business cycle fluctuations. This motivates the relevance of regime-switching models in the analysis of the yield curve.

Regime-switching models have been applied to yield curves in studies such as Dai et al. (2007) and Tavanielli and Laurini (2023), the latter with Brazilian data. However, these approaches typically rely on Markov-switching frameworks, such as the algorithm proposed by Kim (1994), among others, these models present interpretational limitations, as the regimes are latent and therefore difficult to associate directly with observable economic conditions. In this context, Bie et al. (2024) propose a more economically interpretable methodology by incorporating machine learning techniques into the DNS framework; specifically, they employ decision trees to identify regimes explicitly, conditioning them on observable macroeconomic variables.

Given these considerations, this article proposes a nonlinear time series model for the dynamics of the Brazilian yield curve factors, in which both the conditional mean and the persistence of the factors vary smoothly as a function of an observable macroeconomic variable (several candidates are tested). Specifically, a multivariate Smooth Transition Autoregressive (STAR) model is employed, where the transition between regimes is governed by the Selic rate, interpreted as an indicator of the state of the economic cycle. The use of smooth transitions, rather than discrete regime shifts, allows for gradual changes in factor dynamics, avoiding potentially unrealistic abrupt breaks.

The latent factors of the yield curve are extracted using Principal Component Analysis (PCA), a widely adopted strategy in the empirical literature due to its simplicity and effectiveness in summarizing the information contained in the yields of different maturities. The threshold governing regime transitions is not imposed *a priori* but is instead selected empirically from the observed values of the transition variable, excluding extreme quantiles of the sample to ensure proper regime identification. The model is estimated via maximum likelihood, and model selection is based on information criteria, specifically the Akaike Information Criterion (AIC). In addition, formal tests of linearity and model specification within the STAR framework are conducted to assess the adequacy of the nonlinear specification.

The main contribution of this article is to demonstrate how multivariate STAR models can be consistently applied to the dynamics of factors extracted from the yield curve, preserving coherence between dimensionality reduction and temporal modeling. In addition, from an empirical perspective, the results provide evidence that the dynamics of Brazilian term structure factors differ systematically over the economic cycle, reinforcing the relevance of nonlinear approaches in yield curve modeling.

Besides this introduction, the paper is organized as follows. Section 2 presents a concise literature review, introducing the key concepts underlying term structure decomposition. Sections 3 and 4 present, respectively, the econometric model and tests employed in this study, and the data description. Section 5 discusses the empirical results, and Section 6 concludes.

2. Literature Review

This section presents the fundamental concepts underlying the study. In addition, it provides a concise review of the literature on yield curve modeling and regime-switching frameworks.

The main financial instruments in markets where the time value of money is traded are bonds, which are debt securities issued by governments or corporations to raise funds, in essence, the purchaser lends money to the issuer for a specified period (Caldeira, 2011; Filipovic, 2009). Bonds may pay periodic coupons (fixed cash amounts at predetermined dates) or be issued at a discount and pay a single amount at maturity (the latter are called zero-coupon bonds). However, any coupon-bearing bond can be decomposed into a portfolio of zero-coupon bonds, with each coupon payment represented by a zero-coupon bond maturing on the corresponding payment date (Smith, 2014).

The yield of a zero-coupon bond represents the return obtained by holding the bond until maturity. The price of a zero-coupon bond is expected to increase gradually over time until reaching maturity, at which point it attains its maximum value. However, the actual increase observed prior to maturity may deviate from this expected path and fluctuate, as it is determined by market forces and by the risk that changes in short-term interest rates affect the bond's resale value before maturity. The yield to maturity, in turn, is the rate that equates the current bond price to the discounted face value, it corresponds to the bond's internal rate of return and represents the rate at which the bond's price must grow in order to equal its face value at maturity (Caldeira, 2011).

Another important concept is the risk-free interest rate. Strictly speaking, there is no completely risk-free rate; however, in the literature, the shortest-maturity bond (e.g., one month) is typically used as a proxy. According to Saarinen (2012), the liquidity preference hypothesis (and the risk aversion) justifies the use of short-term securities as a risk-free benchmark, since investors demand an additional premium (the term premium) to hold longer-maturity bonds.

The term structure is the function that relates a financial variable to its maturities. The yield curve represents this structure for spot interest rates over time. It is worth noting that, because not all maturities are always available in the market at a given point in time, parametric and nonparametric techniques, such as interpolation, are employed to complete the missing information.

Parametric models, such as that of Nelson and Siegel (1987), rely on a single functional form and determine its parameters by minimizing the deviations between theoretical and observed prices. In contrast, nonparametric models, including those based on splines and exponential interpolation, fit the curve using segments or specific functional forms, offering greater flexibility at the expense of economic interpretability (BIS, 2005). Nelson and Siegel (1987) proposed a parsimonious functional specification capable of representing the yield curve through three components: level, slope, and curvature.

Diebold and Li (2006) developed a dynamic version of the Nelson and Siegel (1987) model by proposing that the factors evolve according to an autoregressive process, thereby enabling forecasting and time series analysis. The empirical robustness of the model has made it a central reference for central banks and financial institutions, including those in Brazil. Several studies have extended the DNS framework by incorporating no-arbitrage restrictions and explicit links with macroeconomic variables. For instance, Diebold et al. (2006) explore the interaction between yield curve factors and variables such as inflation and economic activity.

Although the DNS model provides a satisfactory description of the average evolution of the yield curve, empirical evidence suggests that latent factors do not follow purely smooth dynamics but are instead subject to structural breaks and regime changes. In this context, regime-switching models, particularly Markov-switching specifications, have become popular tools for capturing such nonlinearities. Examples applied to the yield curve include Dai et al. (2007), Tavanielli and Laurini (2023), and Hevia et al. (2015). Markov-type regime-switching models allow for the estimation of probabilities of being in distinct regimes (such as expansion or recession); however, they present interpretational limitations, as the regimes are latent and often lack a direct association with observable macroeconomic variables. Some studies relate regimes to business cycles or market volatility (e.g., Bansal and Zhou (2002)), but such interpretations are typically made *ex post*, limiting their predictive value.

Bie et al. (2024) advance the literature by proposing a new approach that replaces the latent structure of Markov-switching models with endogenous sample partitions obtained through decision trees. Their method uses macroeconomic variables as candidate splitting variables and selects partitions that maximize the marginal likelihood of the DNS model. In this way, regimes cease to be purely statistical constructs and become directly interpretable in terms of observable economic conditions, such as periods of high inflation or economic slowdown, depending on the variables chosen to define the splits. The use of decision trees provides not only interpretative clarity but also greater flexibility, as different combinations of macroeconomic variables may endogenously determine distinct regimes. This allows the model to better adapt to the specific characteristics of each economy, as indicated by the empirical data. This contribution represents a turning point in the regime-switching literature applied to the yield curve, as it combines the robustness of the Bayesian framework with the explanatory power of machine learning techniques.

Another possible response to these limitations is the use of Smooth Transition Autoregressive (STAR) models, in which the model parameters vary continuously as a function of an observable transition variable. These models allow for gradual changes in the dynamics of the process, avoiding the abrupt shifts imposed by discrete regime models, such as Threshold Autoregressive (TAR) models (Enders, 2015; van Dijk et al., 2002).

Two main versions of the STAR model stand out, both allowing for varying degrees of autoregressive decay and differing in the functional form governing the transition between regimes: the Logistic Smooth Transition Autoregressive (LSTAR) model and the Exponential Smooth Transition Autoregressive (ESTAR) model. In the LSTAR specification, the transition is governed by a logistic function, implying asymmetry between regimes and enabling the capture of distinct dynamics when the transition variable lies above or below a given threshold. In contrast, the ESTAR model employs an exponential transition function that is symmetric around the threshold, making it particularly suitable for describing dynamics that deviate from a "central" regime and revert to it as shocks dissipate (Enders, 2015; van Dijk et al., 2002).

When implementing this class of models, linearity must be formally tested. In this regard, Teräsvirta (1994) provides a suitable approach, as it offers a relatively simple testing framework that is often capable of detecting nonlinear behavior and assisting in the choice between ESTAR and LSTAR specifications. The selection between these specifications is not merely technical but carries relevant economic implications, as it reflects different ways in which the dynamics of the variable of interest respond to changes in the state of the economy. Another formal procedure, proposed by Lin and Teräsvirta (1994), allows for the joint assessment of the presence of nonlinearities and the relative adequacy of LSTAR versus ESTAR models, providing a statistical criterion for selecting the most appropriate functional form.

In this sense, the present article shares the motivation of Bie et al. (2024) but adopts an alternative and more parsimonious empirical strategy based on directly modeling the dynamics of yield curve factors using STAR models. By allowing both the conditional mean and the persistence of the factors to vary smoothly as a function of the transition variable, this study contributes to a deeper understanding of the nonlinearities embedded in the term structure of interest rates, particularly in the context of the Brazilian economy.

3. The Model

Building on the literature discussed in the previous section, this section describes the econometric framework adopted to model the dynamics of yield curve factors. The proposed model relies on a factor representation of the yields and incorporates a smooth transition mechanism into the factor dynamics, allowing both the conditional means and the autoregressive parameters to vary over the business cycle. This specification aims to capture relevant nonlinearities while preserving parsimony and interpretability.

Let $y_t \in \mathbb{R}^N$ denote the vector of yields observed at N maturities for periods $t = 1, \ldots, T$. It is assumed that the cross-sectional variation of these interest rates can be adequately represented by a small number of latent factors $f_t \in \mathbb{R}^K$. In line with the empirical literature, we set $K = 3$, corresponding to the usual interpretations of level, slope, and curvature of the yield curve. The factors f_t are estimated using Principal Component Analysis applied to the observed yields. This approach reduces the dimensionality of the problem without imposing additional parametric structure on the shape of the yield curve, thereby focusing the analysis on the temporal dynamics of the extracted factors.

Since the objective of this study is to capture potential nonlinearities in factor dynamics, we follow the proposal of Bie et al. (2024), with the necessary adaptations to the context of STAR models featuring two regimes. To this end, we define the indicator variable z_{t-1} as

$$z_{t-1} = \begin{cases} 1, & \text{if } s_{t-1} > c, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

where s_t denotes the observable transition variable and c is the threshold parameter.

Thus, the regime-dependent conditional mean is given by

$$\mu_{z_{t-1}} = (1 - z_{t-1})\mu_0 + z_{t-1}\mu_1, \tag{2}$$

where μ_0 and μ_1 denote the mean vectors associated with the regime in which the transition variable is below and above the threshold, respectively. This specification allows changes in the state of the economy to affect not only the persistence of the factors but also their conditional mean levels. The use of z_{t-1} ensures that regime classification depends solely on information available up to the previous period, thereby avoiding simultaneity issues.

Accordingly, the vector of regime-centered factors is defined as

$$F_t = f_t - \mu_{z_{t-1}}. \tag{3}$$

This transformation makes it possible to model factor dynamics around regime-dependent means, isolating the regime component from persistent fluctuations.

To allow factor dynamics to vary nonlinearly over time, a STAR model is adopted, in which the parameters of the process depend on an observable transition variable s_t. The specification used is given by

$$y_t = \alpha + \Lambda\mu_{z_{t-1}} + \Lambda F_t + \varepsilon_t \tag{4}$$

and

$$F_t = A_1 F_{t-1} G(s_{t-1};\gamma,c) + A_2 F_{t-1}\left[1 - G(s_{t-1};\gamma,c)\right] + \eta_t, \tag{5}$$

where Λ is the $N \times K$ matrix of factor loadings, ε_t is the vector of measurement errors, A_1 and A_2 are $K \times K$ matrices of autoregressive coefficients, c is the threshold parameter, $\gamma > 0$ governs the smoothness of the transition between regimes and $\eta_t \sim \mathcal{N}(0,\Sigma)$ represents the innovations vector. In addition, α enotes the intercept in Equation (4), included to allow the model to capture the average level of the yield curve not explained by the factors.

Formal linearity tests are conducted to assess whether factor dynamics can be adequately described by a linear model or whether there is statistical evidence in favor of a smooth transition specification. To this end, the test proposed by Teräsvirta (1994) is employed. The test relies on a Taylor expansion of the transition function around the linear case ($\gamma = 0$), allowing the null hypothesis of linearity to be evaluated against the STAR alternative.

In addition, the procedure proposed by Lin and Teräsvirta (1994) is employed, which extends the linearity test by allowing discrimination between different functional forms of the transition. This test makes it possible to assess whether the data dynamics are better described by an LSTAR model, characterized by asymmetry between regimes, or by an ESTAR model, in which the transition is symmetric around the threshold. If the LSTAR specification is preferred, the transition function takes the following form:

$$G(s_{t-1};\gamma,c) = [1+\exp\left(-\gamma(s_{t-1}-c)\right)]^{-1}, \tag{6}$$

whereas if the ESTAR specification is selected, the transition function is given by:

$$G(s_{t-1};\gamma,c) = 1-\exp\left[-\gamma(s_{t-1}-c)^2\right]. \tag{7}$$

The threshold parameter c is selected through a grid search over the interval between the 15th and 85th percentiles of the transition variable. The value that minimizes the Akaike Information Criterion (AIC) is chosen, based on the estimation of Equations (5) combined with either (6) or (7), according to the outcome of the Lin and Teräsvirta (1994) test.

In this way, the adopted specification combines dimensionality reduction via PCA with a flexible modeling of the temporal dynamics of the factors, allowing both the conditional mean and the persistence of the process to vary smoothly over the business cycle. This approach preserves parsimony and transparency while incorporating nonlinearities that are relevant for the analysis of the term structure of interest rates.

4. Base de Dados

The Interbank Deposit Futures Contract traded on the Brazilian exchange is one of the main instruments used to construct the Brazilian yield curve, as it represents one of the largest fixed-income markets among emerging economies. This contract trades the effective Interbank Deposit Certificate (CDI) rate accumulated daily from the current date until the contract's maturity and exhibits a high correlation with the Selic rate (*Sistema Especial de Liquidação e Custódia*), the benchmark interest rate of the Brazilian economy (Caldeira, 2020).

To construct the proposed model, data from Interbank Deposit Futures contracts are used as a proxy for nominal securities. These data, obtained from information provided by B3 - Brasil, Bolsa, Balcão (2025), consist of the yields of DI Futures contracts traded on the exchange, covering maturities ranging from one month to ten years (120 months). The sample uses end-of-month observations from December 2005 (with January 2 used as an approximation) through July 2024.

Table 1
Descriptive statistics of yields

Maturities (years)	Minimum	Median	Mean	Maximum	Standard Deviation	Skewness	Kurtosis
1	2.32	11.04	10.35	16.60	3.24	-0.58	-0.31
2	3.38	11.41	10.73	16.63	2.87	-0.51	-0.12
3	4.34	11.59	11.03	17.53	2.61	-0.42	0.03
4	5.07	11.65	11.24	18.10	2.45	-0.32	0.15
5	5.62	11.75	11.38	18.42	2.34	-0.24	0.26
6	6.03	11.81	11.48	18.53	2.26	-0.18	0.34
7	6.35	11.85	11.56	18.51	2.20	-0.14	0.39
8	6.55	11.86	11.62	18.40	2.15	-0.10	0.41
9	6.67	11.88	11.66	18.23	2.11	-0.08	0.42
10	6.76	11.91	11.70	18.03	2.08	-0.06	0.41

Note: This table reports descriptive statistics of interest rates (in percentage terms) for DI Futures contracts at selected maturities ranging from 1 to 10 years. For each maturity, the minimum, median, mean, maximum, standard deviation, skewness, and kurtosis are reported. The sample period spans from December 2005 to July 2024.

The Table 1 reports the descriptive statistics of annualized yields (in percentage terms) for DI Futures contracts traded on B3, with maturities ranging from 1 to 10 years, using end-of-month data from December 2005 to July 2024. It can be observed that both the mean and the median of the rates increase with maturity, while the standard deviation decreases, indicating lower dispersion at longer horizons. Skewness is negative across all maturities but tends toward zero, suggesting more symmetric distributions at longer maturities. Excess kurtosis shifts from negative to positive, indicating that although short-term rate distributions exhibit lighter tails, long-term rates display greater concentration around the mean and heavier tails.

In order to identify the most appropriate transition variable and determine a potential threshold value consistent with the dynamics of the term structure, several potentially relevant macroeconomic variables were collected and organized. These series were tabulated to match the sample period of the interest rate data, additionally including observations from November 2005, given the need for a one-period lag of the transition variable under the dynamic specification of the model. The following section discusses the sources of each dataset, as well as any data treatments applied when necessary.

The database on Brazilian unemployment rates was obtained from Instituto de Pesquisa Econômica Aplicada (IPEA) (2025), requiring the consolidation of two distinct series due to temporal discontinuities. For data from March 2012 onwards, the unemployment rate series from Instituto Brasileiro de Geografia e Estatística (IBGE) (2025c) is used. For earlier periods, data are drawn from the unemployment rate series of Instituto Brasileiro de Geografia e Estatística (IBGE) (2025b), although this dataset covered only the metropolitan regions of Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, and Porto Alegre. During the period in which both series overlap — March 2012 to December 2015 — the discontinued series consistently reports lower unemployment rates than the current one, this discrepancy likely reflects the regional scope of the earlier dataset, which was restricted to selected metropolitan areas. To ensure consistency and enable the joint use of both series over the entire sample period, a splicing procedure was implemented, the earlier series was adjusted upward by adding 1.84 percentage points, corresponding to the average difference observed during the overlapping period.

The policy interest rate was represented by the Selic rate accumulated over the month and annualized based on 252 business days, as provided by Banco Central do Brasil (2025b). The Selic rate constitutes the main instrument of monetary policy in Brazil. The use of the monthly accumulated rate annualized ensures compatibility with the monthly frequency of the other variables in the model, as well as with the annualized yield values.

For the industrial production series, the data obtained from Instituto de Pesquisa Econômica Aplicada (IPEA) (2025) are used. The series represents the annual percentage change in the physical production of the overall industry in Brazil, measured as an index (% p.a.), and is produced by Instituto Brasileiro de Geografia e Estatística (IBGE) (2025a). It is employed to capture cyclical dynamics in industrial activity and overall economic activity throughout the period of analysis.

The inflation was represented by the Índice de Preços ao Consumidor Amplo (IPCA), Brazil's official inflation index, calculated by Instituto Brasileiro de Geografia e Estatistica (IBGE) (2025) and obtained from Instituto de Pesquisa Econômica Aplicada (IPEA) (2025). The 12-month accumulated rate (% p.a.) was used. The choice of the twelve-month accumulated measure allows for capturing the inflation dynamics relevant to monetary policy decisions, reducing the short-term volatility present in monthly variations and aligning the variable with the temporal frequency of the other series in the model.

The M2 aggregate was used as the monetary variable, corresponding to broad money, obtained from Banco Central do Brasil (2025a). M2 comprises currency in circulation held by the public, demand deposits, savings deposits, and time deposits, reflecting the broader liquidity of the economy. The end-of-period balance, expressed in the current monetary unit, was used in order to capture the evolution of the money stock over time. This variable was considered a potential candidate for the transition function due to its role in determining the economy's monetary availability conditions.

The international oil price is proxied by the Brent (FOB) quotation, expressed in U.S. dollars per barrel, as reported by U.S. Energy Information Administration (2025) and obtained from Instituto de Pesquisa Econômica Aplicada (IPEA) (2025). The series reflects the daily price of Brent crude oil in the international market and is widely used as a benchmark for global energy prices. The inclusion of this variable as a potential candidate for the transition function is justified by the significant impact of commodity price shocks on inflation, monetary policy, and domestic macroeconomic conditions, particularly in emerging economies. Consistent with the treatment of interest rates, the value corresponding to the last available day of each month was used.

Table 2
Descriptive statistics of macroeconomic variables

Variable	Minimum	Median	Mean	Maximum	Standard Deviation	Skewness	Kurtosis
Unemployment Rate (%)	6.20	9.82	10.02	14.90	2.33	0.22	-1.13
Selic Rate(%)	1.19	10.90	10.18	18.87	3.56	-0.48	-0.21
Industrial Production (%)	-27.67	0.31	0.19	34.84	7.12	0.34	3.57
IPCA (%)	1.88	5.23	5.62	12.13	2.19	0.89	0.36
M2 (thousand)	548097119	2184788158	2451517882	6361258911	1495887201	0.87	-0.07
Oil Price	14.85	75.44	77.28	136.82	24.75	0.16	-0.75

Note: This table reports descriptive statistics for the unemployment rate (in percentage terms), the monthly accumulated Selic rate annualized (% p.a.), industrial production (% p.a.), IPCA inflation accumulated over 12 months (% p.a.), broad money aggregate M2 (in thousands of current monetary units), and the international oil price (in U.S. dollars per barrel). For each series, the minimum, median, mean, maximum, standard deviation, skewness, and kurtosis are reported. The sample period spans from November 2005 to July 2024.

Table 2 presents the descriptive statistics of the macroeconomic variables over the period from November 2005 to July 2024, allowing for an assessment of both their average behavior and the dispersion and shape of their distributions. The unemployment rate exhibits a mean of 10.02% and a median of

9.82%, with a relatively moderate standard deviation (2.33 p.p.), indicating relevant cyclical fluctuations throughout the sample. The slightly positive skewness (0.22) suggests a higher frequency of observations below the mean, with more pronounced increases concentrated in specific episodes. The Selic rate recorded an average of 10.18% per year, with a substantial range (minimum of 1.19% and maximum of 18.87%), reflecting different monetary policy regimes over the period. The negative skewness (-0.48) indicates a predominance of observations above the mean during the sample, while kurtosis close to zero suggests a distribution approximately consistent with normality.

Industrial production displays considerably higher relative volatility (standard deviation of 7.12) and pronounced extremes (minimum of -27.67% and maximum of 34.84%). The elevated kurtosis (3.57) indicates heavy tails, consistent with the presence of large shocks—such as economic crises or periods of strong recovery. The 12-month accumulated IPCA shows a mean of 5.62% and a standard deviation of 2.19, with positive skewness (0.89), indicating more pronounced inflationary episodes above the mean, though without significant excess kurtosis.

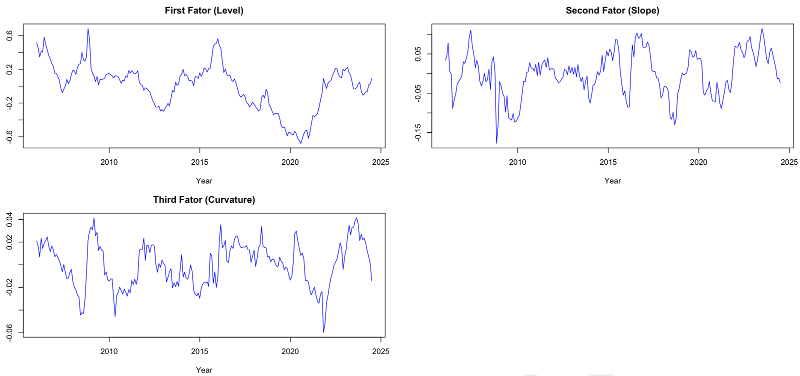
The monetary aggregate (M2) display high dispersion in absolute terms, consistent with their growth trajectory over time, as well as positive skewness (0.87), reflecting the monetary expansion observed during the period. And the international price of oil shows an average of US$ 77.28 per barrel and a standard deviation of US$ 24.75, with slight positive skewness (0.16) and negative kurtosis (-0.75), suggesting a relatively flatter distribution compared to the normal distribution. The results indicate that monetary and price variables exhibit more stable behavior, although they are still influenced by specific episodes of macroeconomic instability.

5. Empirical Results

The estimated model incorporates a smooth transition structure in order to capture potential nonlinearities in the dynamics of yields over time. Initially, the factors extracted from the term structure of interest rates are presented. Next, the linearity tests applied to the factor dynamics are reported, with the aim of assessing the adequacy of a nonlinear STAR-type specification. Finally, the estimation results of the selected model are presented, along with an analysis of its performance in fitting the yield curve across maturities.

In the Figure 1, the time series of the three factors used are plotted. These correspond to the first three principal components extracted from the cross-section of end-of-month yield observations. The factors display periods of higher and lower instability, suggesting the possibility of nonlinear dynamics

Figure 1
Pricing Factors



Note: This figure displays the time series of the model factors. The first three principal components extracted from the cross-section of interest rates of DI futures contracts with maturities ranging from 1 to 120 months are presented, covering the period from December 2005 to July 2024.

and gradual regime changes. This feature motivates the adoption of a smooth transition model capable of capturing structural variations in factor dynamics over time, rather than imposing an invariant linear structure over the entire sample period.

In estimating the model, several macroeconomic variables were considered as potential transition variables. The linearity tests of Teräsvirta (1994) did not provide statistically significant evidence of nonlinearity for most of the variables analyzed, specifically the unemployment rate (p-value of 0.4217), industrial production (p-value of 0.8935), and IPCA inflation (p-value of 0.6571). In some cases, the auxiliary regression exhibited problems of high collinearity, reflecting numerical limitations associated with low variability or high persistence of the series (as observed for oil prices and M2).

The only variable among those tested that provided evidence of nonlinearity as a transition variable was the Selic rate, with a p-value of 0.0649, suggesting rejection of the null hypothesis of linearity at the 10% significance level, although not at the conventional 5% level. Given the central role of monetary policy in shaping the dynamics of the term structure of interest rates, the STAR specification using the Selic rate as the transition variable is adopted.

The sequential procedure proposed by Lin and Teräsvirta (1994) was then applied to investigate the functional form of the nonlinearity. In this test, the lowest p-value was obtained for the hypothesis that all coefficients associated with cubic terms are equal to zero (p-value of 0.1039), indicating that an LSTAR specification is more appropriate for the data.

O modelo LSTAR estimado apresentou limiar de 0,064 (taxa Selic de 6,4%) e parâmetro de suavidade γ igual a 0,616. O valor relativamente moderado de γ indica que a transição entre regimes ocorre de forma gradual, e não abrupta, sugerindo ajustes progressivos na dinâmica do sistema quando a variável de transição cruza o limiar estimado. O comparativo entre diferentes limiares e γ foi feita pela comparação do AIC de cada modelo, assim os critérios de informação (AIC = -3082,874 e BIC = -3014,911) indicam melhor ajuste relativo em comparação às demais especificações.

The estimated LSTAR model yielded a threshold value of 0.064 (Selic rate of 6.4%) and a smoothness parameter γ equal to 0.616. The relatively moderate value of γ indicates that the transition between regimes occurs gradually rather than abruptly, suggesting progressive adjustments in the system's dynamics when the transition variable crosses the estimated threshold. The comparison across different threshold and γ values was conducted using the AIC of each specification, in a way that the information criteria (AIC = -3082.874 and BIC = -3014.911) indicate a superior relative fit compared to the alternative specifications.

The comparison between A_1 and A_2[1] reveals relevant structural changes in factor dynamics across monetary regimes. Substantial changes in both magnitude and sign are observed in several parameters, indicating that the joint dynamics of the factors vary significantly depending on the prevailing regime. In the regime associated with higher levels of the Selic rate, the cross-coefficients exhibit relatively larger magnitudes, indicating stronger interdependence among the yield curve factors. In contrast, under the lower interest rate regime, the dynamics become relatively more self-centered, with greater own persistence in the second and third factors and a reduction in the intensity of cross-factor interactions. For the level factor, however, the opposite pattern is observed: interdependence increases under the lower Selic regime, while its own persistence declines.

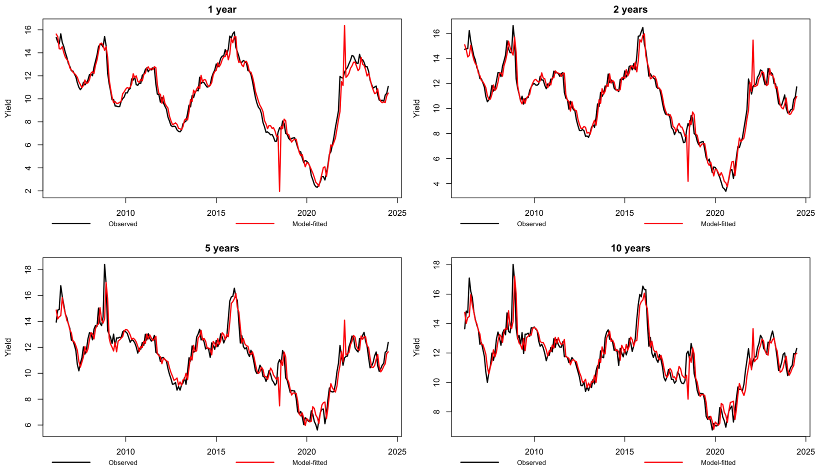
Figure 2 shows that the model reproduces the dynamics of the observed

[1]The following coefficient matrices were obtained:

$$A_1 = \begin{pmatrix} 1.5944 & 10.0996 & -13.7575 \\ -0.4490 & 0.1471 & 4.3150 \\ 0.1605 & -0.2058 & -0.1395 \end{pmatrix} \text{ and } A_2 = \begin{pmatrix} 0.2089 & -10.9199 & 13.0578 \\ 0.5258 & 1.6226 & -4.4730 \\ -0.1585 & 0.2977 & 1.9888 \end{pmatrix}.$$

.

Figure 2
Observed and model-fitted yields



Note: This figure displays the time series of observed yields and those fitted by the model for selected maturities. The sample period spans from December 2005 to July 2024.

interest rates quite satisfactorily across different maturities (1, 2, 5, and 10 years). There is a high degree of adherence between the estimated and observed series, both in terms of trend and cyclical behavior, with the model adequately capturing the main tightening and easing episodes throughout the sample. The most noticeable deviations occur during episodes of abrupt changes, in which the estimated curve tends to partially smooth the most extreme movements and/or prolong previous dynamics (due to the regime being determined based on the previous month). Nevertheless, the overall fit remains consistent across maturities, preserving the temporal pattern of the term structure. It is worth noting that the intercept α is fundamental to this result; when it is removed, a systematic level bias emerges, with the estimated curve persistently shifting downward relative to the observed one.

It is noteworthy that the model's strong performance is maintained even at longer maturities, such as 5 and 10 years. In general, factor models tend to exhibit a more pronounced deterioration in fit as maturity increases, reflecting greater difficulty in capturing the long-term components of the yield curve. In the present case, although the root mean squared error (RMSE) increases with maturity—as expected and reported in Table 3—this increase is relatively moderate; this suggests that the regime-dependent estimated factors are able to capture not only short-term dynamics but also a relevant portion of the variation in long-term rates, indicating that the model provides a good representation of the term structure across the full maturity spectrum, at least for Brazilian data over this period.

Table 3
Root mean squared error of the model

Maturity (years)	RMSE of the LSTAR Model
1	0.00704
2	0.00708
3	0.00721
4	0.00732
5	0.00736
6	0.00736
7	0.00733
8	0.00729
9	0.00727
10	0.00721

O período da amostra é de dezembro de 2005 até julho de 2024.Note: This table reports the root mean squared error of the three-factor LSTAR model estimated using DI futures contracts. The sample period spans from December 2005 to July 2024.

6. Final Remarks

This study investigated the dynamics of the Brazilian term structure of interest rates using a factor-based approach combined with a Smooth Transition Autoregressive (STAR) model. Starting from the extraction of the first three principal components of the yield curve via PCA, the temporal dynamics of the factors were modeled allowing both their mean levels and persistence to vary continuously as a function of the Selic rate, used as the transition variable. This strategy sought to reconcile parsimony, economic interpretability, and flexibility in capturing potential nonlinearities associated with the monetary cycle.

Formal tests were conducted using different variables; however, evidence in favor of nonlinearity was found only when the Selic rate was considered as the transition variable, and the selection procedure pointed to an LSTAR-type specification. The empirical results show that factor dynamics differ across low- and high-interest-rate regimes, with relevant changes in the magnitude and sign of both autoregressive and cross coefficients. In particular, the regime associated with higher interest rates exhibits greater interdependence among the factors, whereas the low-interest-rate regime tends to display relatively more self-centered dynamics, albeit with distinct specificities across factors.

From a goodness-of-fit perspective, the model demonstrated a high capacity to replicate the observed dynamics of yields across the maturities analyzed, including long maturities, for which linear models often exhibit greater performance deterioration. The analysis of the root mean squared error reinforces this evidence, indicating only a moderate increase in error as maturity rises. These results suggest that incorporating smooth transitions contributes to a more accurate representation of the term structure across different horizons.

In economic terms, the findings reinforce the central role of monetary policy in shaping the dynamics of the Brazilian yield curve. By allowing the Selic rate to smoothly govern the transition between regimes, the model captures gradual changes in factor behavior, consistent with how monetary policy is conducted in practice. Thus, the adopted approach provides an interpretable alternative to regime-switching models with latent states, preserving transparency and maintaining a direct connection with observable variables.

A suggestion for future research is to estimate nonlinearity using alternative methodologies, including approaches closer to that proposed by Bie et al. (2024), applied to Brazilian data. Other possible extensions include incorporating additional macroeconomic variables directly into the model specification (rather than only in the transition variable), expanding the set of candidate transition variables, and conducting comparisons with linear models. Nevertheless, the results indicate that STAR models applied to factor dynamics constitute a promising tool for analyzing the term structure of interest rates in Brazil.

References

B3 - Brasil, Bolsa, Balcão (2025). Séries históricas de mercado, https://www.b3.com.br/pt_br/institucional.

Banco Central do Brasil (2025a). Série 27810: Meios de pagamento amplos - m2 (saldo no final do periodo) - novo - u.m.c. (mil), Sistema Gerenciador de Series Temporais (SGS).

URL: https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries

Banco Central do Brasil (2025b). Série 4189: Taxa de juros - selic acumulada no mês anualizada base 252 - % a.a., Sistema Gerenciador de Séries Temporais (SGS).
URL: https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries

Bansal, R. and Zhou, H. (2002). Term structure of interest rates with regime shifts, *The Journal of Finance*, *57*(5), 1997–2043.
URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/0022-1082.00487

Bie, S., Diebold, F. X., He, J. and Li, J. (2024). Machine learning and the yield curve: Tree-based macroeconomic regime switching, *SSRN*, pp. 1–40.
URL: https://ssrn.com/abstract=4934442

BIS (2005). *Zero-coupon yield curves: Technical documentation*.

Caldeira, J. F. (2011). EstimaÇÃo da estrutura a termo da curva de juros no brasil atravÉs de modelos paramÉtricos e nÃo paramÉtricos, *Análise Econômica*, *29*(55).
URL: https://seer.ufrgs.br/index.php/AnaliseEconomica/article/view/13198

Caldeira, J. F. (2020). Investigating the expectation hypothesis and the risk premium dynamics: new evidence for Brazil, *Empirical Economics*, *59*(1), 395–412.
URL: https://ideas.repec.org/a/spr/empeco/v59y2020i1d10.1007_s00181-019-01629-0.html

Dai, Q., Singleton, K. J. and Yang, W. (2007). Regime shifts in a dynamic term structure model of u.s. treasury bond yields, *The Review of Financial Studies*, *20*(5), 1669–1706.
URL: https://academic.oup.com/rfs/article-abstract/20/5/1669/1592122?redirectedFrom=fulltext

Diebold, F. X. and Li, C. (2006). Forecasting the term structure of government bond yields, *Journal of Econometrics*, *130*(2), 337–364.
URL: https://www.sciencedirect.com/science/article/pii/S0304407605000795

Diebold, F. X., Rudebusch, G. D. and Aruoba, S. B. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach, *Journal of*

Econometrics, *131*(1-2), 309–338.
URL: https://ssrn.com/abstract=565168

Enders, W. (2015). *Applied Econometric Time Series*, 4 ed., John Wiley & Sons, Hoboken.

Filipovic, D. (2009). *Term-Structure Models: A Graduate Course*, Springer.

Hevia, C., Gonzalez-Rozada, M., Sola, M. and Spagnolo, F. (2015). Estimating and forecasting the yield curve using a markov switching dynamic nelson and siegel model, *Journal of Applied Econometrics*, *30*(6), 987–1009.
URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2399

Instituto Brasileiro de Geografia e Estatistica (IBGE) (2025). Ipca - indice geral - taxa de variacao em 12 meses (% a.a.), Sistema Nacional de Indices de Precos ao Consumidor (SNIPC). Taxa acumulada em 12 meses do Indice de Precos ao Consumidor Amplo.

Instituto Brasileiro de Geografia e Estatística (IBGE) (2025a). Produção industrial (pan12_qiigg12) variação percentual anual (% a.a.). Série mensal de produção industrial geral, variação percentual em relação a igual mês do ano anterior.
URL: https://www.ipeadata.gov.br/ExibeSerie.aspx?serid=38397

Instituto Brasileiro de Geografia e Estatística (IBGE) (2025b). Taxa de desemprego - inativa. Série PAN12 TD12. Acess with IPEADATA.

Instituto Brasileiro de Geografia e Estatística (IBGE) (2025c). Taxa de desocupação - pnad contínua. Série PNADC12 TDESOC12. Acess with IPEADATA.

Instituto de Pesquisa Econômica Aplicada (IPEA) (2025). Ipeadata: Base de dados macroeconômicos, http://www.ipeadata.gov.br.

Kim, C.-J. (1994). Dynamic linear models with markov-switching, *Journal of Econometrics*, *60*(1), 1–22.
URL: https://www.sciencedirect.com/science/article/pii/0304407694900361

Kim, D. H. and Orphanides, A. (2007). The bond market term premium: what is it, and how can we measure it?, *BIS Quarterly Review*, .
URL: https://ideas.repec.org/a/bis/bisqtr/0706e.html

Lin, C.-F. J. and Teräsvirta, T. (1994). Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics*, *62*(2), 211–228.

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves, *The Journal of Business*, *60*(4), 473–489.

Saarinen, J. A. (2012). *Using the yield curve in predicting real economic growth: Application to finland*, Master's thesis, University of Helsinki, Department of Political and Economic Studies, Helsinki.
URL: https://helda.helsinki.fi/server/api/core/bitstreams/8d7b1dc2-c43b-4be0-88b1-29b5bebe4dbb/content

Smith, D. J. (2014). *Bond Math: The Theory Behind the Formulas*, Wiley Finance.

Tavanielli, R. and Laurini, M. (2023). Yield curve models with regime changes: An analysis for the brazilian interest rate market, *Mathematics*, *11*(11).
URL: https://www.mdpi.com/2227-7390/11/11/2549

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association*, *89*(425), 208–218.

U.S. Energy Information Administration (2025). Preço - petróleo bruto - brent (fob) - us, $EnergyInformationAdministration(EIA).SrieobtidanoIPEADATA.$

van Dijk, D., Teräsvirta, T. and Franses, P. H. (2002). Smooth transition autoregressive models — a survey of recent developments, *Econometric Reviews*, *21*(1), 1–47.