

THE ROLE OF FISCAL SENTIMENT IN BRAZIL'S YIELD CURVE

Tatiana Pinheiro Marcelo Kfoury Muinhos Marcelo Fernandes

Abstract

This paper constructs a novel Brazilian Fiscal Sentiment Index (FSI-BR) from daily newspaper and newswire articles using natural language processing and studies its relation to the term structure of interest rates. The index tracks major macroeconomic events and co-moves with standard risk indicators and asset prices. Embedding the FSI in a dynamic Nelson–Siegel model with forward-looking macroeconomic variables, we show that fiscal sentiment significantly affects the slope and curvature of the yield curve, with limited impact on the level. Deteriorations in fiscal sentiment are associated with a flattening of the yield curve, a result consistent with simultaneous increases in short- and long-term risk premia under Brazil's inflation-targeting regime. These findings identify fiscal perceptions as a key driver of yield curve dynamics.

Keywords: natural language processing, sentiment index, term structure of interest rates, dynamic Nelson-Siegel model and macroeconomic expectations.

1 Introduction

While a growing literature uses textual data to construct measures of economic uncertainty and sentiment, there is no high-frequency fiscal-specific sentiment index for Brazil. This paper fills this gap by constructing a Fiscal Sentiment Index (FSI-BR) from domestic news sources using natural language processing (NLP) techniques and integrating it into a macro-financial term structure framework of Diebold, Rudebusch, and Aruoba (2006).

The FSI-BR captures agents' perceptions of fiscal conditions and provides a novel high-frequency measure of fiscal expectations in an economy where fiscal dynamics have persistently shaped macroeconomic outcomes and financial markets.

The use of textual data in economics has expanded significantly with advances in machine learning and computational linguistics, allowing researchers to extract structured information from unstructured sources. Seminal contributions such as Baker, Bloom, and Davis (2015) introduced the Economic Policy Uncertainty (EPU) index based on newspaper coverage, followed by a broad literature applying similar approaches across countries. In Brazil, related measures include the EPU-Brazil index developed by Baker and the Economic Uncertainty Indicator (IIE) developed by Ferreira et al. (2019), which combines news-based metrics with forecast dispersion from the Central Bank's Focus survey. More broadly, textual data have been shown to contain valuable information for financial markets and macroeconomic dynamics, including volatility forecasting (Manela and Moreira, 2017), asset pricing and investment strategies (Santos et al., 2023), and real-time activity tracking (Martins, 2022).

Literature has paid limited attention to fiscal-specific sentiment, particularly in emerging markets. Avila (2023) constructs a fiscal risk index using news data and shows that increases in perceived fiscal risk affect long-term interest rates and the exchange rate. Nevertheless, this measure focuses on negative sentiment and does not capture the full distribution of fiscal perceptions.

This paper addresses these gaps by constructing the FSI-BR from daily news coverage in major Brazilian newspapers: *O Estado de S. Paulo*, *Folha de S. Paulo*, and *Valor Econômico*, combined with financial newswire data (Broadcast), covering the period

from 2008 to 2022. Sentiment is extracted using FinBERT-PT-BR (Santos et al., 2023), a transformer-based model tailored to financial texts in Portuguese, allowing classification into positive, neutral, and negative categories. Unlike previous approaches, our index incorporates the full distribution of sentiment, providing a comprehensive measure of fiscal perceptions over time.

To assess its macro-financial relevance, the FSI-BR is incorporated into a dynamic latent factor model of the yield curve following Diebold, Rudebusch, and Aruoba (2006), augmented with forward-looking macroeconomic variables, following the evidence of Vieira et al. (2017) and Fernandes and Vieira (2019) that such variables improve the modeling of the Brazilian yield curve. The model is estimated at weekly frequency to better capture the high volatility of the Brazilian term structure and focuses on explaining the in-sample dynamics of the level, slope, and curvature factors. This framework allows fiscal sentiment to act as an additional state variable influencing the joint evolution of yield curve factors and macroeconomic conditions.

The results show that fiscal sentiment is an economically relevant driver of the yield curve. The model closely tracks observed yields, and fiscal sentiment shocks generate meaningful responses primarily in the slope and curvature factors, with more limited effects on the level. Deteriorations in fiscal sentiment are associated with a flattening of the yield curve, at first glance appears counterintuitive, but this result is consistent with increases in both short- and long-term risk premia under Brazil's inflation-targeting regime. These findings suggest that fiscal perceptions play a central role in shaping financial conditions in Brazil.

This paper contributes to literature in three dimensions. First, it constructs a novel high-frequency fiscal sentiment index for Brazil based on domestic news sources. Second, it provides new evidence on the role of fiscal sentiment in the dynamics of the term structure within a macro-financial latent factor framework. Third, it integrates advances in NLP into a structural model of the yield curve, bridging the gap between textual analysis and macro-financial modeling.

2 Building the Fiscal Sentiment Index

The objective is to develop an instrument that captures the effects of fiscal policy conduct on market expectations, going beyond the information contained in the Focus survey forecasts for the primary balance, nominal balance, net debt, and gross debt. To this end, the paper uses NLP techniques to extract fiscal sentiment from qualified media coverage of fiscal news. This approach helps overcome key limitations of analysts' forecasts in the Focus survey, which are reported at annual frequency and incorporate fiscal policy developments only with a lag, generating a timing mismatch between expectations and asset prices that adjust at the time of policy announcements. More importantly, this methodology allows for the construction of an index that captures market agents' perceptions of fiscal conditions.

Brazil Fiscal Sentiment Index recovers an unobservable variable - fiscal sentiment - from textual data in a setting where no labeled benchmark exists and where raw documents are highly dimensional. The database combines two complementary sources. The first consists of daily articles from three major Brazilian newspapers: *O Estado de S. Paulo*, *Folha de S. Paulo*, and *Valor Econômico*. The second consists of daily reports from Agência Estado's Broadcast platform -*Abertura*, *Cenário 1*, and *Cenário 2*. The data range is from January 2008 to December 2022. The resulting corpus spans multiple presidential administrations, fiscal regimes, sovereign rating changes, and major reform episodes, making it particularly suitable for tracking shifts in fiscal narratives over time.

Standard preprocessing procedures are applied, following Gentzkow et al. (2019) to reduce dimensionality and transform the corpus into a format suitable for sentiment analysis. Sentiment classification is performed using FinBERT-PT-BR (Santos et al., 2023), a transformer-based language model tailored to Portuguese financial texts. For consistency of interpretation, the original model outputs are remapped from {negative, neutral, positive} = {1, 2, 0} to {-1, 0, +1}. At the sentence level, the model thus assigns negative, neutral, or positive sentiment scores that can be directly aggregated.

Additionally, a dictionary-based strategy is necessary because the target variable is not observable in the data and no training labels exist for supervised estimation of fiscal sentiment. As emphasized by Gentzkow et al. (2019), dictionary methods are particularly appropriate in exactly this type of setting: when the latent concept of interest cannot be

directly observed and no benchmark dataset is available to train or validate a supervised classifier. In this paper, the dictionary is to identify the subset of sentences that is relevant for fiscal interpretation. Sentiment is therefore evaluated only for sentences containing at least one fiscal-related dictionary term; sentences without any dictionary term are treated as neutral for the purpose of fiscal sentiment extraction. This filtering step is important because it constrains the model to operate on the fiscal content of the corpus rather than on the overall tone of the news, thereby reducing noise and improving interpretability.

Two dictionaries¹ strategies are implemented. The first is an ad hoc dictionary composed of unigrams and bigrams commonly associated with fiscal policy, such as “fiscal,” “debt,” “deficit,” “public debt,” “fiscal target,” “primary result,” and “public expenditures.” Its main advantage is transparency. However, an ad hoc list is inevitably subject to researcher discretion and may fail to capture changes in the vocabulary of fiscal debate across time. For that reason, a second dictionary is constructed - an endogenous dictionary directly from the corpus. The endogenous list is generated separately for each administration period, allowing the fiscal vocabulary to adapt to shifting policy agendas and media narratives. The corpus is divided into the Lula 2, Dilma 1, Dilma 2, Temer, and Bolsonaro periods.

The endogenous dictionary is built using the K-means++ clustering algorithm² of Arthur and Vassilvitskii (2006) applied to word embeddings generated by FinBERT-PT-BR. Because embeddings encode contextual similarity, this procedure identifies words that tend to appear in the fiscal-policy discourse of each period even when they are not obvious candidates from an ex-ante perspective. This feature is especially valuable in the Brazilian case, where the content of fiscal debate changes across administrations - from debt sustainability and exchange-rate concerns to spending caps, pension reform, primary balance target, and other institutional themes. The endogenous dictionary therefore reduces subjective interference, captures the time-varying semantics of fiscal debate, and provides a measure more closely aligned with the actual discourse observed in the data.

Once sentence-level sentiment is obtained, document-level sentiment is computed as the average sentiment across the sentences in each text. The next step is to build daily

¹ Full table of dictionary lists is in Appendix A

² The most frequent nouns in the corpus are represented as dense vectors, and clustering is performed around the semantic neighborhood of the term “fiscal policy.”

sentiment series. The average sentiment within each source is normalized using z-scores, following Baker, Bloom, and Davis (2015), as raw sentiment measures are not directly comparable across sources or over time. Specifically, they exhibit a strong negative bias whose magnitude varies across sources, and document volumes differ substantially both across sources and over time.

This normalization serves three purposes. First, it mitigates the strong negative bias present in the raw sentiment series, particularly in Broadcast. Second, it makes the newspaper and Broadcast series comparable despite differences in scale and sentiment dispersion. Third, it attenuates distortions arising from unequal news volumes across sources and across administrations. In this setting, normalization is not merely a statistical convenience; it is required to ensure that the final index reflects shifts in fiscal tone rather than mechanical differences in publication intensity or source composition.

After normalization, the daily sentiment index is constructed from the counts of positive, neutral, and negative texts. Following Hiew et al. (2019), the fiscal sentiment measure for day t is defined as the difference between the number of positive and negative texts divided by the total number of texts:

$$FSI_t = \frac{Pos_t - Neg_t}{Pos_t + Neu_t + Neg_t}. \quad (1)$$

This formulation preserves the sign of sentiment while scaling it by the number of available texts on each day. For newspapers, daily raw sentiment is aggregated across *Folha de S. Paulo*, *O Estado de S. Paulo*, and *Valor Econômico*, weighting each paper by its share of total articles on that day. For Broadcast, the daily index is the average across the three summary reports—*Abertura*, *Cenário 1*, and *Cenário 2*. All series are ultimately constructed at working-day frequency to ensure consistency.

Given the high volatility of daily news-based sentiment, the normalized series is then smoothed using an exponential moving average (EMA), which preserves the high-frequency nature of the data, enhances the readability of the series around major fiscal events, and reduces the risk that the final measure is dominated by idiosyncratic daily shocks. Thereby extracting a more stable signal that is better suited for interpretation and

for use in downstream macro-financial models. The index is integrated in section 3 into a term structure framework estimated at weekly frequency.

The newspaper-only index benefits from broader coverage and longer-form reporting, while the Broadcast-only index provides timelier and market-oriented fiscal information. The combined index exploits both dimensions, summarizing the common fiscal signal across sources. Accordingly, we construct seven fiscal sentiment indices: three based on an ad hoc dictionary, three based on an endogenous dictionary, and an alternative GPT-based index built from the same corpus for robustness analysis. Because no benchmark series for fiscal sentiment exists, validation relies on the standard approach in the textual-index literature: comparing the series with historically relevant events and with variables that should co-move with fiscal perceptions. Accordingly, we examine the correlation of the alternative FSI measures with Brazil's five-year sovereign CDS, the BRL/USD exchange rate, the Economic Policy Uncertainty index for Brazil (Baker et al., 2015), and FGV's Economic Uncertainty Index (Ferreira et al., 2019).

Based on the comparison results reported in Appendix B, we adopt the FSI Total Endogenous (hereafter FSI-BR) as the benchmark measure of fiscal sentiment. The index is defined as the simple average of the normalized and EMA-smoothed newspaper and Broadcast indices constructed using the endogenous dictionary over the common sample (January 2008–December 2022). This index combines broader media coverage with real-time financial reporting, adapts to shifts in fiscal discourse across administrations, aligns more closely with major fiscal events, and outperforms the main alternatives in the robustness exercises. It is therefore the most suitable summary measure of fiscal sentiment for the macro-financial analysis that follows.

3 The Yield Curve and the Fiscal Sentiment Index (FSI-BR)

Understanding the determinants of the yield curve is central to evaluating the transmission of monetary and fiscal policies, forecasting interest rates, and pricing fixed-income securities. This issue is particularly relevant in Brazil, where fiscal constraints have historically shaped macroeconomic dynamics and risk premia. In this context, this section examines the interaction between the Brazilian yield curve and fiscal sentiment by introducing the Fiscal Sentiment Index (FSI-BR) into a macro-financial term structure framework.

3.1 Data

The empirical analysis is based on the Brazilian term structure of interest rates (ETTJ), constructed by ANBIMA using a parametric Svensson (1994) methodology (ANBIMA, 2021). This approach produces zero-coupon equivalent yields across maturities by discounting future cash flows, ensuring consistency with standard fixed-income theory (Duffie and Singleton, 2003; Diebold and Rudebusch, 2013).

The dataset includes seven maturities, $\tau = \{1,3,12,24,60,84,120\}$ months, observed weekly (Fridays) from January 18, 2008, to December 30, 2022 (746 observations). Since ANBIMA reports annualized yields, rates are converted to weekly equivalents.

The choice of weekly frequency is motivated by the historically elevated volatility of Brazilian interest rates. Monthly data, as in Diebold, Rudebusch, and Aruoba (2006), hereafter DRA (2006), would result in a loss of information, particularly during periods of rapid macroeconomic and fiscal adjustment. This choice is supported by Vieira et al. (2017) and Fernandes and Vieira (2019), who show that higher-frequency data improves model performance in the Brazilian context.

Descriptive statistics confirm that the Brazilian yield curve is typically upward sloping but highly volatile, especially at short maturities. This motivates the use of latent factor models to capture time-varying dynamics.

Empirical proxies for the latent factors are defined as:

$$\text{Level}_t = \frac{y_t(1) + y_t(24) + y_t(120)}{3} \quad (2)$$

$$\text{Slope}_t = y_t(1) - y_t(120) \quad (3)$$

$$\text{Curvature}_t = 2y_t(24) - y_t(1) - y_t(120) \quad (4)$$

Table 1 – Descriptive Statistics of the Brazilian Yield Curve (annualized rates).

	maturities (months)							Level_	Slope_	Curvature_
	1	3	12	24	60	84	120	proxy	proxy	proxy
Mean	9.50	9.52	9.88	10.39	11.09	11.28	11.45	10.44	-1.95	-0.17
Median	10.16	10.10	10.44	11.10	11.64	11.73	11.79	11.17	-1.89	-0.26
Maximum	14.72	14.77	16.38	17.41	18.00	18.04	18.05	16.53	2.69	5.16
Minimum	1.95	1.90	2.30	3.38	5.62	6.35	6.69	4.12	-6.78	-4.02
Std. Dev	3.36	3.42	3.31	2.93	2.31	2.16	2.03	2.67	2.11	1.61
Skewness	-0.50	-0.46	-0.43	-0.43	-0.31	-0.25	-0.17	-0.48	-0.19	0.27
Kurtosis	-0.60	-0.65	-0.67	-0.45	-0.04	0.08	0.22	-0.43	-0.42	-0.61
Observations	746	746	746	746	746	746	746	746	746	746

Source: Prepared by the author (2025) with data from ANBIMA (2021). From 2008 to 2022. Note: Slope values are negative, consistent with the notation adopted by the model, explained below.

The Brazilian data show a positive correlation between level and slope, indicating that increases in long-term rates are often associated with a flattening of the curve. It is important to note that, under the Diebold and Li (2003) parameterization, a positively sloped curve (steepening) implies a negative slope factor, which affects the interpretation of results.

Table 2 – Correlation Matrix of Level, Slope, and Curvature

	Level_	Slope_	Curvature_
	proxy	proxy	proxy
Level_proxy	1.000	0.618	0.414
Slope_proxy	0.618	1.000	-0.168
Curvature_proxy	0.414	-0.168	1.000

Source: Prepared by the author (2025) with data from 2008 - 2022.

3.2 Model Framework

To model the yield curve, I adopt the Dynamic Nelson–Siegel (DNS) framework extended by DRA (2006)³. For a given maturity τ , the yield is defined as:

³ The DRA (2006) model provides a natural starting point because it combines the empirical tractability of the dynamic Nelson–Siegel approach with a macroeconomic state–space representation.

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t} \left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_{3,t} \left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) + \varepsilon_t(\tau) \quad (5)$$

where $\beta_{1,t} \equiv L_t$ (level), $\beta_{2,t} \equiv S_t$ (slope), $\beta_{3,t} \equiv C_t$ (curvature), λ is the decay parameter (fixed at 0.069159⁴), $\varepsilon_t(\tau) \sim \mathcal{N}(0, H)$ is the measurement error.

The model is cast in state-space form. Let s_t denote the state vector.

Yields-Only Model:

$$s_t = f_t = \begin{bmatrix} L_t \\ S_t \\ C_t \end{bmatrix} \quad (6)$$

Yields-Macro Model:

$$s_t = \begin{bmatrix} L_t \\ S_t \\ C_t \\ \text{GDP}_t^{12m} \\ \text{Selic}_t^{12m} \\ \text{IPCA}_t^{12m} \end{bmatrix} \text{ or } \begin{bmatrix} L_t \\ S_t \\ C_t \\ \text{GDP}_t^{12m} \\ \text{Selic}_t^{12m} \\ \text{FSI}_t \end{bmatrix} \quad (7)$$

Transition Equation

$$s_t = \mu + A s_{t-1} + \eta_t, \eta_t \sim \mathcal{N}(0, Q) \quad (8)$$

where μ is the intercept vector, A is the transition matrix (VAR(1)), Q is the covariance matrix of state innovations.

⁴ The decay parameter λ was previously estimated in order to find a parameter that represents the optimal solution where latent factors are stable and the state-space converges. λ of 0,069159 implies that the loading on the curvature factor is maximized at a maturity of 25,93 months; maximum curvature loading value of 0,298. Diebold and Li (2003) found that it is commonly observed that the curvature achieves its maximum between 24 and 36 months.

Measurement Equation

$$y_t = \Lambda(\tau)f_t + \varepsilon_t, \varepsilon_t \sim \mathcal{N}(0, H) \quad (9)$$

where:

$$\Lambda(\tau) = \begin{bmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_n}}{\lambda\tau_n} & \frac{1-e^{-\lambda\tau_n}}{\lambda\tau_n} - e^{-\lambda\tau_n} \end{bmatrix} \quad (10)$$

and H is assumed diagonal.

3.3 Macroeconomic Variables and Fiscal Sentiment

The macroeconomic block includes forward-looking variables⁵: $IPCA_t^{12m}$ (inflation expectations), GDP_t^{12m} (growth expectations), $Selic_t^{12m}$ (monetary policy expectations), and FSI_t (Brazil Fiscal Sentiment Index)

All series are weekly and aligned with the yield curve data. The use of expectations (instead of contemporaneous variables) reflects the forward-looking nature of financial markets and is supported by Vieira et al. (2017) and Fernandes and Vieira (2019), who show that expectations significantly improve model performance.

3.4 Identification and IRFs

Impulse-response analysis is based on the covariance matrix Q , using a Cholesky decomposition. The ordering follows DRA (2006):

$$(L_t, S_t, C_t, GDP_t^{12m}, Selic_t^{12m}, IPCA_t^{12m})$$

or alternatively:

$$(L_t, S_t, C_t, GDP_t^{12m}, Selic_t^{12m}, FSI_t)$$

⁵ All expectations of 12 months ahead

This ordering assumes that yield curve factors contemporaneously affect macro variables, but not vice versa.

3.5 Estimation

The model is estimated via maximum likelihood using the Kalman filter. Key assumptions include: H diagonal, Q unrestricted, λ fixed, Stability of A (eigenvalues inside unit circle), Positive definiteness of Q . The Kalman smoother is used to extract latent factors.

This framework allows for bidirectional interactions between macroeconomic expectations, fiscal sentiment, and the yield curve. The inclusion of the FSI-BR extends the DRA (2006) model by incorporating a real-time measure of fiscal perceptions, which is particularly relevant in Brazil, where fiscal risk is a key determinant of interest rates.

By combining high-frequency data, forward-looking macroeconomic variables, and a novel fiscal sentiment measure, this framework provides a more informative representation of the Brazilian term structure, improving both empirical fit and the economic interpretation of yield curve dynamics.

4 Results

This section presents the main empirical results. It first discusses the latent factors extracted from the Kalman filter and compares the alternative term-structure specifications. It then reports the impulse-response evidence and its macroeconomic interpretation. Finally, it summarizes the robustness exercises.

4.1 Yields-Only models

As a first step, the latent factors—level, slope, and curvature—are estimated using the standard two-step Nelson–Siegel procedure, with the decay parameter fixed at ($\lambda = 0.069159$). This preliminary decomposition provides economically meaningful factors and supplies initial values for the state-space estimation. Consistent with the stylized behavior of the Brazilian yield curve, the estimated level factor is always positive and less volatile, whereas slope and curvature fluctuate around zero. Relative to the U.S. evidence in DRA (2006), the Brazilian factors are highly persistent, with autoregressive

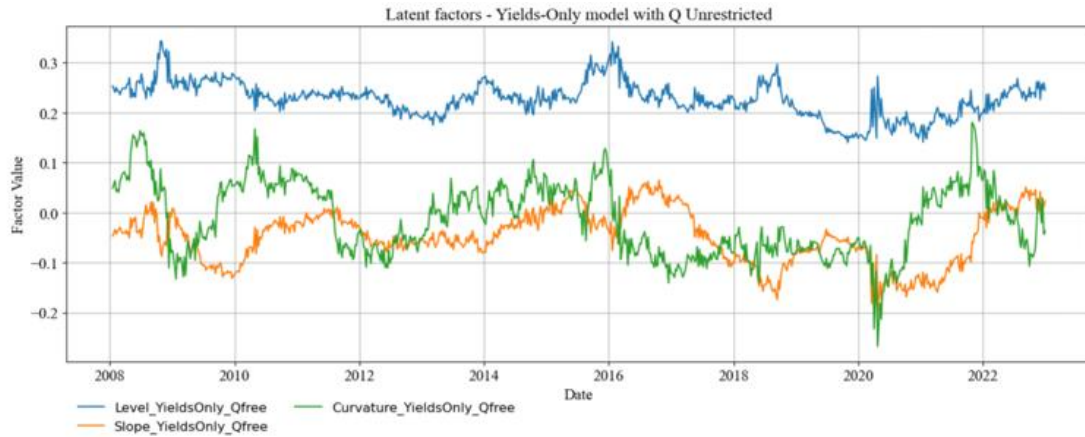
coefficients of 0.96, 0.98, and 0.98 for level, slope, and curvature, respectively. This persistence reflects the inertial nature of Brazilian interest rates and the slow dissipation of shocks to the term structure.

The Yields-Only model is then estimated by maximum likelihood via the Kalman filter. When initialized with the two-step factors and allowing for an unrestricted state covariance matrix (Q), the model converges successfully and yields a log-likelihood of 21,784.11. The estimated transition matrix is stable, with all eigenvalues inside the unit circle, and the covariance matrix (Q) is positive definite. The filtered latent factors preserve the expected Nelson–Siegel interpretation: the level factor remains highly persistent and relatively smooth, while slope and curvature are more volatile and mean reverting.

A useful benchmark is obtained by comparing unrestricted- (Q) and diagonal- (Q) versions of the model. The diagonal- (Q) specification produces a lower log-likelihood (20,822), but slightly lower root mean squared errors (RMSE) at some maturities. This difference is not contradictory. The log-likelihood evaluates the model's overall statistical fit, whereas RMSE focuses directly on pricing errors along the yield curve. Since the primary objective of the empirical exercise is to reproduce the observed term structure as closely as possible, RMSE is an especially relevant metric.

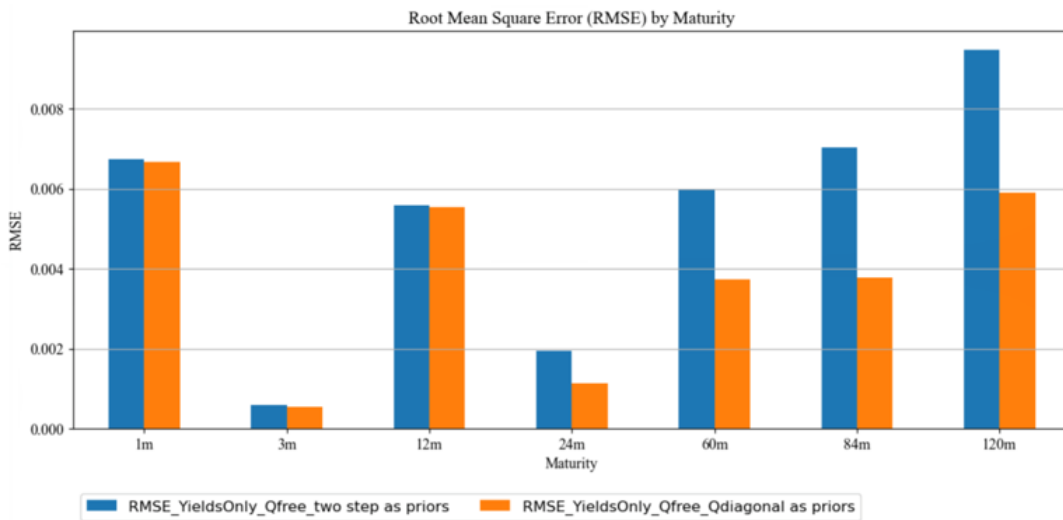
To improve the fit further, I re-estimate the unrestricted- (Q) model using the parameters from the diagonal- (Q) specification as priors. This version achieves the best performance among the Yields-Only models, with a higher log-likelihood of 22,059.91 and lower RMSE across most maturities.

Chart 1 - Latent factors estimated of Yields-Only model with Q Unrestricted (using Q-diagonal as priors)



Source: Prepared by the author (2025).

Chart 2 – RMSE of Yields-Only models using the parameters estimated by two-step and Q-diagonal as priors



Source: Prepared by the author (2025).

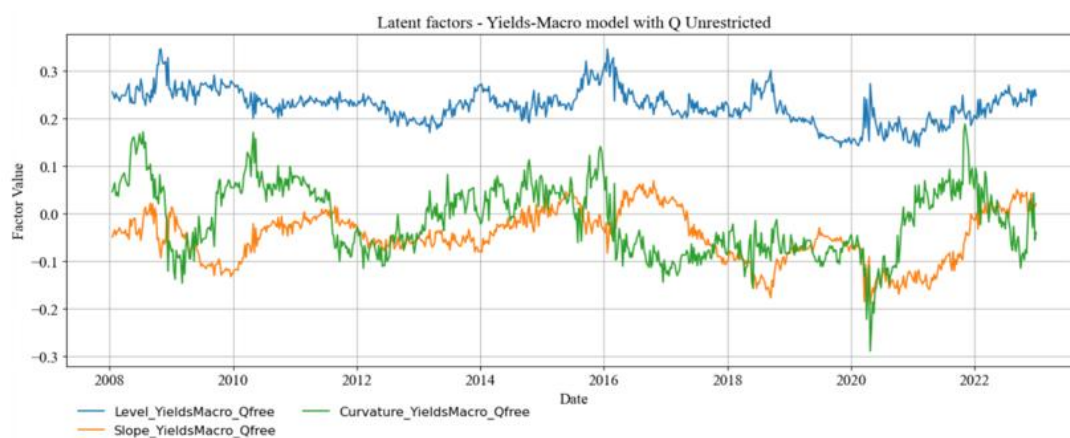
The estimated transition matrix confirms that the latent factors are highly persistent, with significant diagonal elements and only limited cross-factor spillovers, presented in Appendix C. The strongest off-diagonal effects are modest influences of slope on level and curvature. The covariance matrix (Q) indicates that curvature is subject to relatively larger shocks, consistent with the higher variability of the medium segment of the Brazilian curve. The intercept vector suggests a small positive drift in level and negative drifts in slope and curvature, consistent with stationary mean-reverting dynamics.

The tests of (Q)-diagonality deliver mixed evidence. The likelihood ratio test rejects the null of a diagonal (Q), whereas the Wald test does not. This result mirrors DRA (2006), who also find that unrestricted covariance structures in yields-only models often approximate diagonal matrices. Following Stock and Watson (2016), this divergence can be interpreted as reflecting the greater sensitivity of the likelihood ratio test to local alternatives. On balance, the unrestricted-(Q) specification is retained because it provides a superior overall fit and greater flexibility in capturing the in-sample dynamics of the Brazilian term structure. Accordingly, this Yields-Only specification serves as the benchmark against which the Yields-Macro models are evaluated.

4.2 Yields-Macro models

The next step augments the latent-factor system with forward-looking macroeconomic variables. In the benchmark macro-financial specification, hereafter DRABZ, the state vector includes the three Nelson–Siegel factors together with twelve-month-ahead expectations for GDP growth, the Selic rate, and IPCA inflation. This model converges successfully, with a log-likelihood of 23,498.66, outperforming the Yields-Only specifications. The eigenvalues of the transition matrix remain inside the unit circle, confirming stability, and the covariance matrix (Q) is positive definite. The latent factors continue to exhibit the expected behavior: level is smoother and always positive, while slope and curvature fluctuate around zero.

Chart 3 - Latent factors estimated by DRABZ

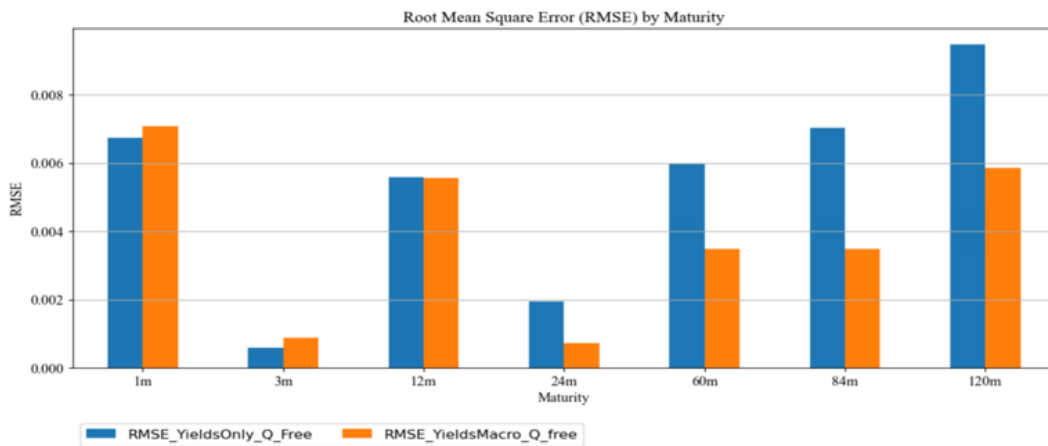


Source: Prepared by the author (2025).

The estimated state dynamics are economically informative. Presented in Appendix C, the three latent factors remain highly persistent, with autoregressive coefficients of approximately 0.92–0.95, while inflation expectations are highly persistent, GDP expectations moderately persistent, and Selic expectations comparatively less persistent. The off-diagonal coefficients indicate strong macro-financial linkages. Higher Selic expectations raise the level factor and reduce both slope and curvature. Given the sign convention of the Diebold and Li (2003) slope, this corresponds to a steepening of the yield curve: long-term rates rise more than short-term rates. The same qualitative effect arises from stronger GDP growth expectations and, to a lesser extent, from higher inflation expectations. These responses are consistent with the historical Brazilian pattern in which confidence shocks and rising macroeconomic risk tend to increase long-term yields disproportionately.

The macro-financial model also performs better in terms of pricing accuracy. RMSE comparisons show that DRABZ dominates the Yields-Only specification, especially at medium and long maturities. This result is important because it indicates that forward-looking macro variables contain economically relevant information for the long end of the Brazilian yield curve.

Chart 4 - RMSE of DRABZ and Yields-Only model



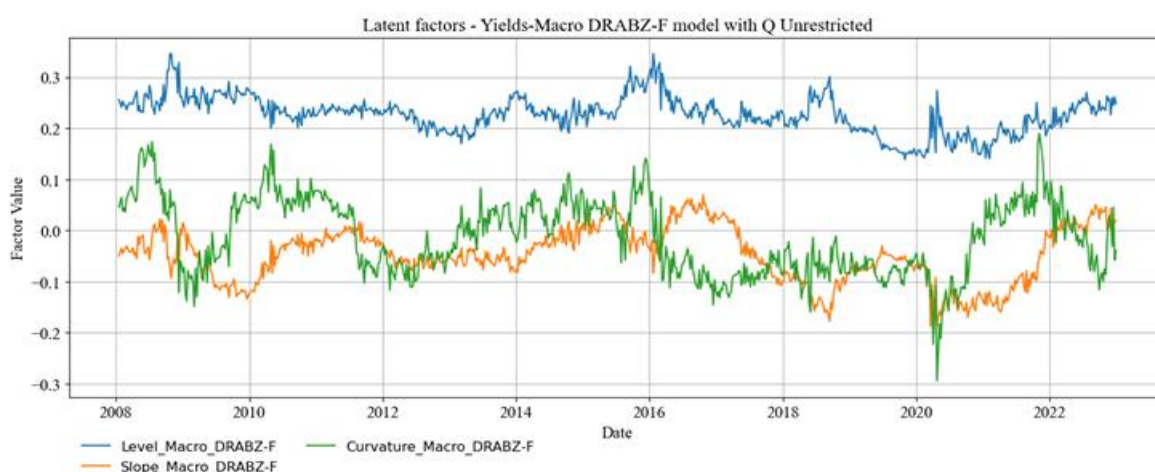
Source: Prepared by the author (2025).

4.2.1 The fiscal-sentiment specification

The central specification of the paper, hereafter DRABZ-F, replaces inflation expectations with the Fiscal Sentiment Index (FSI-BR). This substitution is motivated by the DRA interpretation of the factors and by the institutional relevance of fiscal risk in Brazil. In the DRA framework, inflation is associated primarily with long-term yield movements, the policy rate with the short end, and real activity with the belly of the curve. Because fiscal sentiment is expected to affect long-term risk premia and sovereign credibility, it is natural to allow it to enter the state vector in place of inflation expectations.

The DRABZ-F model converges successfully and delivers a log-likelihood of 23,468.79, only slightly below the benchmark DRABZ specification and well above the Yields-Only models. Stability is again confirmed by eigenvalues within the unit circle, and the covariance matrix remains positive definite. Presented in Appendix C, the three latent factors retain their standard interpretation and high persistence, with autoregressive coefficients of 0.93, 0.95, and 0.92 for level, slope, and curvature, respectively. Fiscal sentiment itself is also highly persistent, whereas GDP expectations remain moderately persistent and Selic expectations less so.

Chart 5 - Latent factors estimated by DRABZ-F



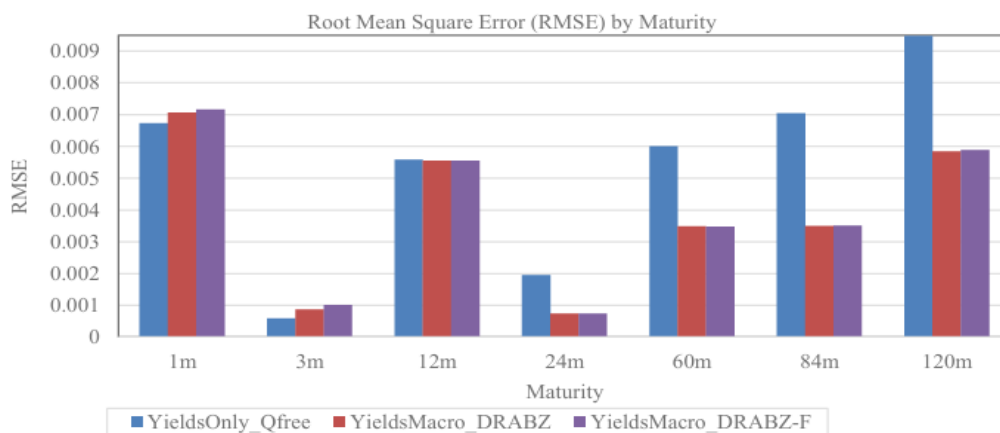
Source: Prepared by the author (2025).

The fiscal block yields the main result of the paper. Because the FSI-BR is constructed such that lower values indicate more negative sentiment, a decline in fiscal sentiment is interpreted as a worsening of fiscal perceptions. The estimated transition coefficients

show that a deterioration in fiscal sentiment has only a small effect on the level factor but significantly affects slope and curvature. Specifically, a lower FSI-BR raises the slope factor and reduces curvature. Given the sign convention of the Diebold and Li (2003) slope, this implies a flattening of the yield curve led by a stronger increase in short-term rates relative to long-term rates. At first glance, this may appear counterintuitive, since fiscal deterioration is often associated with long-end stress. However, the result is consistent with Brazil’s experience under inflation targeting: fiscal laxity increases long-run risk premia, but it also triggers short-run monetary tightening expectations, producing a front-end-led flattening. This interpretation is reinforced by the negative lagged effect of fiscal sentiment on Selic expectations: a worsening in fiscal sentiment leads the market to revise upward the expected policy rate.

The macroeconomic interpretation of the remaining coefficients is also instructive. GDP and Selic expectations continue to affect all three factors in ways similar to the benchmark DRABZ model. In contrast, fiscal sentiment has negligible effects on expected GDP growth, suggesting that fiscal perceptions in this framework operate mainly through financial pricing rather than through direct growth expectations. By comparison, Selic expectations exert a much stronger negative influence on GDP expectations, highlighting the asymmetric roles of monetary policy and fiscal sentiment: the former is the main instrument affecting real activity, while the latter primarily shapes the term structure and risk premia.

Chart 6 – RMSE of Yields-Only, Yields-macro DRABZ and DRABZ-F



Source: Prepared by the author (2025).

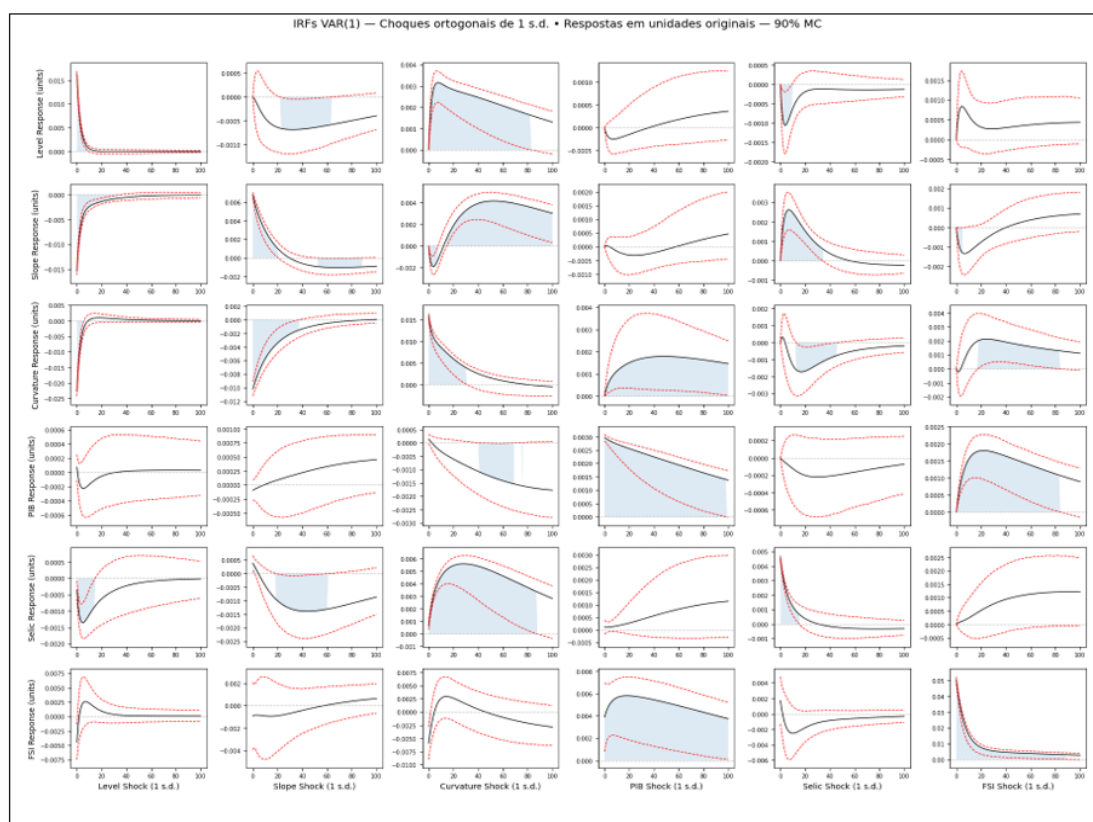
RMSE comparisons confirm that DRABZ-F performs similarly to DRABZ and better than the Yields-Only specification, particularly from the 12-month to the 120-month maturities. Thus, including fiscal sentiment does not materially improve short-end pricing, but it adds substantial economic content and maintains strong performance at the medium and long end of the curve. Because its empirical fit is close to that of DRABZ and because it directly isolates the fiscal channel of interest, DRABZ-F is chosen as the preferred specification for the impulse-response analysis.

4.3 Impulse-response analysis

Impulse-response functions are used to trace the propagation of shocks through the system. Following DRA (2006), the responses are computed for one-unit shocks rather than one-standard-deviation shocks, and identification is achieved through recursive ordering using a Cholesky decomposition. Confidence intervals are constructed using a 90% parametric bootstrap, which is preferable here to asymptotic delta-method approximations given the complexity of the system.

The responses of the latent factors confirm the intuitive interpretation of the DNS decomposition. A level shock produces a short-lived upward shift in the curve, accompanied by a temporary decline in slope and curvature. In Brazil, such shocks are consistent with episodes in which increases in long-term yields reflect rising inflation or risk premia rather than conventional monetary tightening. A slope shock generates temporary flattening, driven by relatively larger increases in short-term rates, and is accompanied by a decline in curvature. This is consistent with a monetary policy surprise that affects the front end most strongly but is expected to dissipate over time. Curvature shocks are the most persistent and influence all three factors, initially steepening the curve before eventually producing a longer-lasting flattening as long-end yields adjust downward.

Figure 1 – Impulse response of DRABZ-F



Source: Prepared by the author (2025).

Macroeconomic shocks also display sensible propagation patterns. A positive shock to GDP expectations raises curvature strongly, underscoring the sensitivity of the belly of the curve to business-cycle conditions. A positive shock to Selic expectations generates a short-lived flattening: short-term rates rise immediately, while long-term rates react more modestly, and curvature declines as tighter expected policy dampens medium-term inflation premia. These responses are fully consistent with the role of policy expectations in shaping the short end of the curve.

The main contribution comes from the fiscal-sentiment shock. Deterioration in the FSI-BR has no statistically meaningful effect on the level factor, but it raises the slope factor and persistently increases in curvature. Under the DNS sign convention, this implies a flattening of the curve. Although counterintuitive relative to the standard narrative of fiscal stress steepening the curve, the result is plausible in the Brazilian context. Negative fiscal news simultaneously worsens long-run risk perceptions and increases the expected probability of near-term monetary tightening. The latter effect dominates in the short run, pushing the front end up more strongly than the long end.

Notably, negative fiscal sentiment reduces growth expectations, while fiscal austerity raises them, consistent with the parameter estimates of the DRABZ-F specification. And, as expected, growth shock improves fiscal sentiment. Interestingly, however, the system does not identify a significant response of growth expectations to monetary policy expectation shocks, suggesting a weaker direct transmission from policy rate expectations to activity forecasts, which is consistent with the market's anticipation that higher policy rates will not persist in time.

Taken together, the responses of the yield curve to Level, Slope, and Curvature shocks, the responses of growth expectations to fiscal sentiment, and the response of fiscal sentiment to growth expectations indicate that the model is well specified. Accordingly, the evidence supports the finding of a yield curve flattening in response to a deterioration in fiscal assessments, rather than a steepening of the curve - a result that, while counterintuitive, proves robust within the model's framework.

4.4 Interaction tests

To complement the impulse responses, we conduct interaction tests based on the transition matrix (A) and the covariance matrix (Q).

Table 3 shows the interaction tests on matrix A reveal strong evidence that GDP expectations, fiscal sentiment (FSI-BR), and monetary policy expectations (Selic) exert significant effects on the yield curve, whereas the reverse causality from yields to these macroeconomic variables is not supported. Complementary tests on matrix Q indicate that the structural shocks to GDP, FSI-BR, and Selic are statistically independent from shocks to the yield curve. Together, these results suggest a clear direction of causality from macroeconomic fundamentals and policy expectations to the term structure, rather than the opposite, reinforcing the interpretation of the yield curve as a channel of transmission rather than an autonomous driver of macroeconomic expectations. as the findings of Yields-Macro DRA (2006) for the US term-structure interest rate. In this sense, the Brazilian evidence differs from the bidirectional interactions emphasized in DRA (2006) for the U.S. economy and instead points to a more one-sided macro-to-finance transmission mechanism.

Table 3 - Interaction test

Wald Tests on Transition Matrix A			
Hypothesis	Wald statistic	df	p-value
GDP → Yields (A[:3,3]=0)	3604.39	3	0.000
FSI → Yields (A[:3,4]=0)	954.85	3	0.000
SELIC → Yields (A[:3,5]=0)	2598.69	3	0.000
Yields → GDP (A[3,:3]=0)	0.06	3	0.996
Yields → FSI (A[4,:3]=0)	2.88	3	0.411
Yields → SELIC (A[5,:3]=0)	0.01	3	1.000

Shock Independence Tests on State Covariance Q			
Hypothesis	Wald		
	statistic	df	p-value
Shocks GDP–Yields independent (Q[:3,3]=0)	7.80145E-07	3	1.000
Shocks FSI–Yields independent (Q[:3,4]=0)	6.60373E-07	3	1.000
Shocks SELIC–Yields independent (Q[:3,5]=0)	1.70317E-09	3	1.000

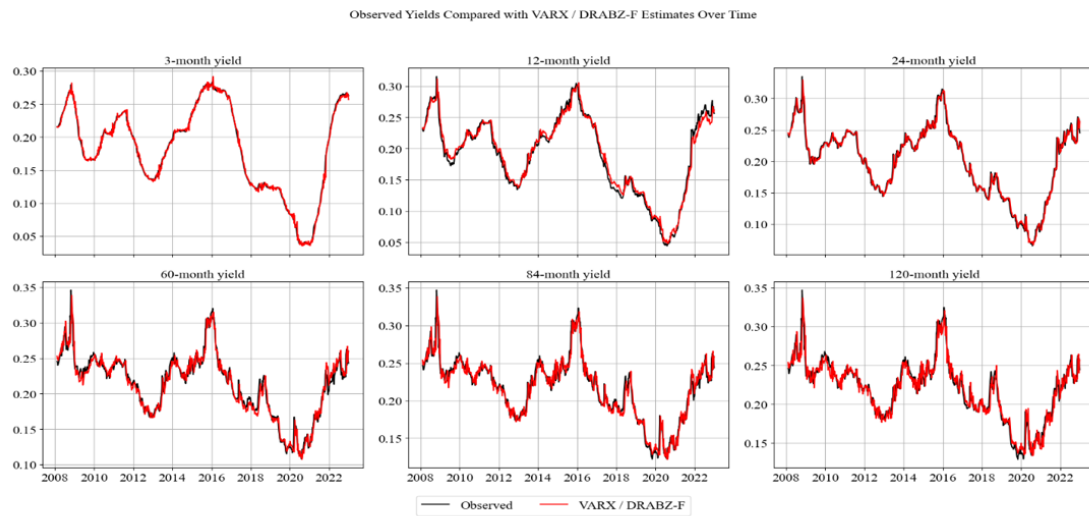
Source: Prepared by the author (2025).

4.5 Robustness

The final step is to assess whether the strong in-sample fit of DRABZ-F and the associated impulse responses are statistically credible rather than sample-specific. The robustness analysis focuses on stability diagnostics because the main concern is whether the model's fit deteriorates during episodes of extreme market stress, especially during the COVID-19 shock.

The first diagnostic consists of visual inspection of residuals by maturity. These visuals provide a quick, model-agnostic check on adequacy and the reliability of inference by highlighting potential nonlinearity, heteroskedasticity, serial dependence/structural breaks, and outliers.

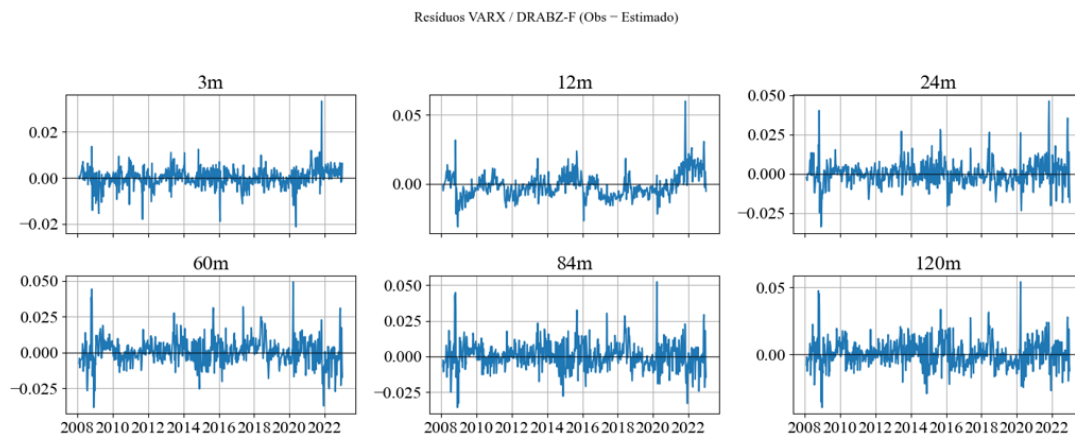
Figure 2 – Observed yields compared with yields estimated by DRABZ-F model



Source: Prepared by the author (2025).

Residuals are tightly centered around zero for most of the sample, with relatively stable variance, but display pronounced spikes during 2020–2021. This suggests that any instability is concentrated in the pandemic period rather than spread evenly over the sample.

Figure 3 – Residuals of DRABZ-F model (observed yields – estimated)



Source: Prepared by the author (2025).

To formalize this evidence, we compute residual-based stability tests for each maturity. Specifically, we run CUSUM and CUSUMSQ tests with wild-bootstrap p-values using Rademacher weights, as well as a univariate Sup-Wald test for an endogenous mean break

over a trimmed search grid. These tests serve distinct purposes. CUSUM assesses parameter constancy over time, CUSUMSQ tests for variance instability, and the Sup-Wald test identifies the most likely structural break date in the mean of the residual process. Together, they provide a comprehensive picture of model stability.

The results show a clear pattern. CUSUMSQ never rejects the null of stability, indicating no material change in residual variance across maturities. By contrast, CUSUM rejects parameter constancy at the short end—particularly for the 1-, 3-, and 12-month maturities—but not at medium and long maturities at conventional significance levels. The Sup-Wald statistics identify economically meaningful breaks concentrated in the COVID period: 3- to 12-month yields break between February and July 2021, the 24-month tenor breaks in August 2020, and the 60-month tenor in May 2021. The long–short spread also exhibits a sharp break in July 2021. Additional pre-COVID breaks are detected at the 84-month tenor in August 2018 and marginally at the 120-month tenor in February 2016, but these are more limited in scope.

Table 4 – CUSUM, CUSUMSQ and Sup-Wald tests results

Maturity	CUSUM_ stat	CUSUM_ p	CUSUMSQ_ stat	CUSUMSQ_ p	SUPW_ stat	SUPW_ p	break_date
1m	111.035	0	156.240	1	174.434	0	7/30/2021
3m	77.651	0.004	93.727	1	70.875	0	2/12/2021
12m	125.838	0	193.249	1	207.078	0	6/4/2021
24m	34.890	0.344	110.881	1	12.358	0.034	8/7/2020
60m	51.599	0.112	93.169	1	28.828	0	5/21/2021
84m	44.686	0.17	94.837	1	13.664	0.014	8/31/2018
120m	40.482	0.214	90.545	1	9.238	0.064	2/26/2016

	CUSUM_ stat	CUSUM_ p	CUSUMSQ_ stat	CUSUMSQ_ p	SUPW_ stat	SUPW_ p	break_ date
Spread 120m-1m	67.016	0.03	74.444	1	64.600	0	7/2/2021

Source: Prepared by the author (2025).

The key point is that the instability is concentrated in the short end of the curve and in the pandemic episode. Since CUSUMSQ does not reject, the evidence points to a level shift rather than a change in the variance of innovations. This distinction is important. The model does not fail because of heteroskedasticity or explosive residual variance; rather,

it faces a temporary mean shift during an extraordinary macro-financial event. Medium and long maturities -the segment most relevant for the fiscal-sentiment results -remain comparatively stable.

This finding provides the basis for maintaining the main conclusions of the paper. The COVID episode was an exceptional period marked by abrupt changes in monetary policy, liquidity conditions, and risk aversion. Treating it as an atypical shock and assuming that the underlying macro-financial relationships revert toward their pre-COVID regime, the DRABZ-F specification remains a reliable basis for inference, especially for medium- and long-term maturities. Since the main contribution of the paper concerns the role of fiscal sentiment in shaping the curve beyond the short end, and since the long maturities are precisely where the model performs best, the presence of pandemic-related instability does not invalidate the central results.

4.6 Summary of findings

Several conclusions emerge from the empirical analysis. First, all specifications recover economically meaningful latent factors, but the Yields-Macro models outperform the Yields-Only models in terms of fit, particularly at longer maturities. Second, replacing inflation expectations with the FSI-BR yields a model that fits nearly as well as the benchmark macro specification while adding substantial economic content. Third, the main result is robust across model comparisons and impulse responses: a deterioration in fiscal sentiment flattens the Brazilian yield curve, primarily through effects on slope and curvature rather than on the level. Fourth, interaction tests indicate that fiscal sentiment, GDP expectations, and Selic expectations drive the yield curve, not vice versa – different than DRA found in the US. Finally, the robustness analysis shows that residual instability is largely confined to the COVID period and the short end of the curve, leaving the medium- and long-term results intact.

Taken together, these findings suggest that fiscal perceptions are a relevant and distinct determinant of the Brazilian term structure. The FSI-BR improves the economic interpretation of yield-curve dynamics and provides evidence that, in Brazil's inflation-targeting regime, fiscal deterioration affects the curve mainly through the front end and the belly, rather than through a simple upward shift in long-term rates.

5 Conclusion

The Brazilian yield curve differs markedly from those of advanced economies. Its historically high volatility, sensitivity to fiscal conditions, and strong macro-financial interactions require a framework that is both flexible and economically interpretable. The DRA (2006) model provides a natural starting point because it combines the empirical tractability of the dynamic Nelson–Siegel approach with a macroeconomic state-space representation.

Building on this framework, this study introduces three innovations. First, it uses weekly rather than monthly data, which is more appropriate for the Brazilian environment. Second, it replaces contemporaneous macroeconomic indicators with forward-looking expectations, which are more relevant for financial pricing. Third, and most importantly, it incorporates a novel Fiscal Sentiment Index (FSI-BR), capturing real-time perceptions of fiscal conditions.

The contribution of this paper is therefore to provide a coherent macro-financial framework in which fiscal sentiment can be analyzed as a determinant of the Brazilian yield curve. By embedding the FSI-BR into a dynamic Nelson–Siegel model of the DRA type, the analysis makes it possible to examine how fiscal perceptions interact with macroeconomic expectations and affect the level, slope, and curvature of the term structure. This approach improves both the empirical fit and the economic interpretation of yield curve dynamics in a setting where fiscal risk is central to the pricing of fixed-income assets.

References

Alves, Cássio RA, and Márcio P. Laurini. The effects of Brazilian Central Bank communication on the yield curve. **Macroeconomic Dynamics** 29 (2025): e83.2025.

ANBIMA. Estrutura A Termo Das Taxas De Juros Estimada E Inflação Implícita – Metodologia. 2021.

Andrews, Donald W K,. Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, **Econometric Society**, vol. 61(4), pages 821-856, July. 1993.

Ang, A.; Piazzesi, M. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. **Journal of Monetary Economics**, p. 745-787, 2003.

Arthur, David; Vassilvitskii, Sergei. **k-means++: The advantages of careful seeding**. Stanford, 2006.

Baker, S.; *et al.* Measuring economic policy uncertainty. **NBER Working Paper Series**, n. 21633, <http://www.nber.org/papers/w21633>, 2015.

Banco Central do Brasil. **Focus - Relatório de Mercado**. 2025. Available at: <https://www.bcb.gov.br/publicacoes/focus>

Bloomberg. Bloomberg Terminal. Available at: <https://www.bloomberg.com/professional/products/bloomberg-terminal/>.

Cañete, J. *et al.* Spanish pre-trained bert model and evaluation data. arXiv preprint arXiv:2308.02976, 2023.

Caldeira, J.F.; *et al.* Efficient Yield Curve Estimation and Forecasting in Brazil. **EconomiA**, v.11, n.1, p.27–51, 2010.

Caldeira, J.F.; *et al.* Forecasting the yield curve with the arbitrage-free dynamic Nelson–Siegel model: Brazilian evidence. **EconomiA**, v.17, p.221–237, 2016.

Campbell, J. Y.; Shiller, R. J. Cointegration and tests of present value models. **NBER Working Paper Series**, n. 1885, 1986.

Avila, R.C. Fiscal Policy Risk and the Yield Curve: an Alternative Measure. Advisor: Carlos Viana de Carvalho; co-advisor: Marcelo Cunha Medeiros. **Dissertação (mestrado)** - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, 2023.

Carriero, A. Forecasting the yield curve using priors from no-arbitrage affine term structure models. **International Economic Review**, v. 52, n. 4, p. 425-459, 2011.

Carriero, A.; Giacomini, R., How useful are no-arbitrage restrictions for forecasting the term structure of interest rates?, **Journal of Econometrics**, Elsevier, vol. 164(1), pages 21-34, September. 2011.

Cho, K. *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: **Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)**, pages 1724–1734, Doha, Qatar. Association for Computational Linguistics. 2014.

Christensen, J. H. E.; Diebold, F.X; Rudebusch, G.D.. The affine arbitrage-free class of Nelson-Siegel term structure models. **NBER Working Paper Series**, n. 13611, <http://www.nber.org/papers/w13611>, 2007.

Cowles, A. Can stock market forecasters forecast? **Econometrica**, v. 1, n. 3, p. 309–324, 1933.

Dai, A.; Le, Q. V. Semi-supervised sequence learning. **Advances in neural information processing systems**, v. 28, 2015.

Delobelle, M.; Gautier, J.; Van Elsen, L. Robbert: a Dutch RoBERTa-based language model. **arXiv preprint arXiv:2001.06286**, 2020.

Devlin, Jacob, *et al.* Bert: Pre-training of deep bidirectional transformers for language understanding. **Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)**, <https://arxiv.org/abs/1810.04805>, 2019.

Diebold, *et al.* (2006). The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach. **Journal of Econometrics**, 2006, vol. 131, issue 1-2, 309-338

Diebold, F. X.; Li, C. Forecasting the term structure of government bond yields. **NBER Working Paper Series**, n. 10048, <http://www.nber.org/papers/w10048>, 2003.

_____. Forecasting the term structure of government bond yields. **Journal of Econometrics**, Volume 130, Issue 2, 2006, Pages 337-364, ISSN 0304-4076, <https://doi.org/10.1016/j.jeconom.2005.03.005>.

DIEBOLD, F. X.; RUDÉBUSCH, G. D. Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach. **Princeton: Princeton University Press**, 2013.

Duffie, D; Singleton, K.J. Credit Risk: Pricing, Measurement, and Management. **Princeton: Princeton University Press**, 2003.

Fernandes, M., & Vieira, F. A dynamic Nelson-Siegel model with forward-looking macroeconomic factors for the yield curve in the US. **Journal of Economic Dynamics and Control**, 106, 103720, 2019

Fernandes, B.; *et al.* Explaining the Brazilian Yield Curve with Macroeconomic Variables and Latent Factors. **Working Paper**, n. 523, Banco Central do Brasil, 2020.

Ferreira, P.C.; Vieira, R.M; Silva, F.B; Oliveira, I.C.L. Measuring Brazilian Economic Uncertainty. **Journal of Business Cycle Research**, v. 15, n. 2, p. 25–40, 2019.

Filho, J. A. W. *et al.* The brwac corpus: a new open resource for brazilian portuguese. In: Proceedings of the 11th International conference on language resources and evaluation. [s.n.], 2018. p. 4339–4344.

Froot, Kenneth A. New hope for the expectations hypothesis of the term structure of interest rates. **The Journal of Finance**, v. 44, n. 2, p. 283-305, 1989.

Gentzkow, M.; Kelly, B.; Taddy, M. Text as data. **Journal of Economic Literature**, v. 57, n. 3, p. 535-574, 2019.

Hiew, J. Z. G. *et al.* BERT-based financial sentiment index and LSTM-based stock return predictability. arXiv:1906.09024v2, 2022.

Hinton, G. E.; Osindero, S.; Teh, Y. W. A fast learning algorithm for deep belief nets. **Neural Computation**, v. 18, n. 7, p. 1523–1554, 2006.

Hochreiter, S.; Schmidhuber, J. Long short-term memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997.

Liu, Y. *et al.* RoBERTa: a robustly optimized BERT pretraining approach. arXiv:1907.11692v1, 2019.

Martins, Leonardo Caio. From Micro to Macro: Essays in Textual Analysis / Chapter 2. Nowcasting GDP with Unstructured Data. Dissertação de Mestrado. PUC Rio. 2022.

Manela, A.; Moreira, A. News implied volatility and disaster concerns. **Journal of Financial Economics**, v. 123, n. 1, p. 137-162, 2017.

Martin, L. *et al.* CamemBERT: a tasty French language model. **arXiv preprint arXiv:1911.03894**, 2019.

Marzi, G. *et al.* K-Alpha Calculator–Krippendorff's Alpha Calculator: A user-friendly tool for computing Krippendorff's Alpha inter-rater reliability coefficient, **MethodsX**, Volume 12, 2024, 102545, ISSN 2215-0161, <https://doi.org/10.1016/j.mex.2023.102545>. 2024.

McCulloch, W. S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, v. 5, n. 4, p. 115–133, 1943.

Nelson, C. R.; Siegel, A. F. Parsimonious Modeling of Yield Curves for U.S. Treasuries Bills. **NBER Working Paper Series**, n. 1594, 1985.

Poder 360. Evolução da Circulação total dos Jornais no Brasil. Instituto Verificador de Comunicação. 2023. Available at: <https://static.poder360.com.br/2023/12/circulacao-jornais-IVC-nov2023.pdf>. Access: 09 set. 2025.

Polignano, Marco *et al.* Alberto: Italian BERT language understanding model for NLP challenging tasks based on tweets. In: **CEUR workshop proceedings**. CEUR, 2019. p. 1-6.

Radford, A. *et al.* Improving language understanding by generative pre-training. 2018. Available at: <https://www.cs.unc.edu/research/gemini/papers/10.1.1.84.4578.pdf>. Access: 9 set. 2025.

Rossi, J.W. A Estrutura a Termo da Taxa de Juros: Uma Síntese. Texto para Discussão nº 0447, IPEA, Rio de Janeiro, dezembro de 1996.

Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533–536, 1986.

Santos, L. *et al.* FinBERT-PT-BR: Análise de Sentimentos de Textos em Português do Mercado Financeiro. In: **BRAZILIAN WORKSHOP ON ARTIFICIAL INTELLIGENCE IN FINANCE (BWAIF)**, 2., 2023, João Pessoa/PB. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2023. p. 144-155. DOI: <https://doi.org/10.5753/bwaif.2023.231151>.

Shannon, C. E. Prediction and entropy of printed English. **Bell System Technical Journal**, v. 30, n. 1, p. 50–64, 1951.

Souza, F. *et al.* BERTimbau: Pretrained BERT Models for Brazilian Portuguese. In **Brazilian Conference on Intelligent Systems**, pages 403–417, 2020. DOI: 10.1007/978-3-030-61377-8_28.

Stock, J. H.; Watson, M. W. **Introduction to Econometrics**. 3. ed. Pearson Education, 2016.

Svensson, L. E. O. Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994. **NBER Working Paper Series**, n. 4871, 1994.

Vaswani, A. *et al.* Attention is all you need. In: **31st Conference on Neural Information Processing Systems (NIPS 2017)**, Long Beach, CA, USA, p. 5998–6008, 2017.

Vieira, F., Fernandes, M., & Chague, F. Forecasting the Brazilian yield curve using forward-looking variable. **International Journal of Forecasting**, 33(1), 121-131, 2017.

Vries, W. *et al.* BERTje: a Dutch BERT language model. arXiv:1912.10301, 2019.

Yang, Z. *et al.* Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, v. 32, 2019.

Appendix A

Table 5A - Comparison between the Endogenous dictionary list obtained in each administration by fiscal policy terms cluster from Newspapers and *Ad Hoc* dictionary list

Lula 2	Dilma 1	Dilma 2	Temer	Bolsonaro	Ad Hoc
economy	economy	economy	pension	pension	fiscal
dollar	dollar	dollar	economy	economy	debt
politics	politics	politics	politics	dollar	deficit
investment	investment	investment	dollar	politics	politics
government	government	government	reforms	stock market	congress
agreement	agreement	adjustment	government	spending cap	public debt
recovery	labor	agreement	spending cap	investments	fiscal policy
debt	analysts	fiscal	agreement	government	fiscal target
bank	debt	debt	labor	agreement	primary target
bonds	bank	budget	bank	labor	primary result
target	target	bank	recovery	recovery	fiscal package
exchange rate	exchange rate	target	target	debt	fiscal measures
fund	fund	credit	debt	budget	public spending
credit	credit	company	company	bank	federal revenue
company	company	project	reform	reform	national congress

Source: Prepared by the author (2025) with data from 2007 - 2024.

Table 5B - Endogenous dictionary lists obtained in each administration by fiscal policy terms cluster from Broadcast

Abertura

Lula 2	Dilma 1	Dilma 2	Temer	Bolsonaro	Lula 3
economy	economy	economy	economy	pension	economy
interest rate	politics	politics	pension	economy	politics
dollar	dollar	dollar	politics	politics	dollar
government	government	government	dollar	dollar	government
earnings	earnings	agreement	government	earnings	earnings
commodities	commodities	retail	agreement	government	agreement
president	agreement	target	retail	retail	retail
recovery	retail	president	reform	bank	bank
analysts	president	minister	president	agreement	president
bank	debt	euro	target	companies	companies
companies	bonds	project	contracts	analysts	target
company	companies	oil	minister	recovery	analysts
contracts	banks	stock markets	euro	company	debt
banks	currency	inflation	project	project	reform
balance	euro	market	oil	proposal	company

Source: Prepared by the author (2025) with data from 2007 - 2024.

Cenário 1

Lula 2	Dilma 1	Dilma 2	Temer	Bolsonaro	Lula 3
economy	economy	economy	pension	economy	economy
interest rate	politics	politics	economy	pension	politics
dollar	interest rate	dollar	politics	interest rate	interest rate
government	dollar	government	dollar	dollar	dollar
commodities	swap	commodities	government	politics	government
adjustment	government	agreement	commodities	government	commodities
exchange rate	commodities	recovery	agreement	spending cap	adjustment
recovery	exchange rate	adjustment	exchange rate	commodities	labor
currency	adjustment	president	adjustment	adjustment	exchange rate
contracts	president	contracts	reform	agreement	payroll
banks	debt	target	target	exchange rate	target
credit	analysts	exchange rate	president	recovery	president
euro	currency	euro	analysts	reform	analysts
oil	contracts	oil	currency	president	bank
president	bonds	currency	contracts	oil	banks

Source: Prepared by the author (2025) with data from 2007 - 2024.

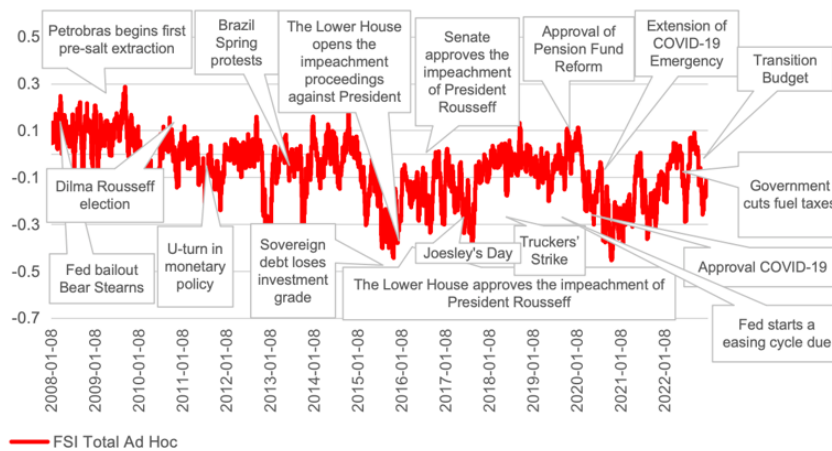
Cenário 2

Lula 2	Dilma 1	Dilma 2	Temer	Bolsonaro	Lula 3
economy	economy	economy	pension	economy	politics
interest rate	politics	politics	economy	pension	economy
dollar	interest rate	interest rate	politics	dollar	dollar
government	dollar	dollar	dollar	politics	government
turnover	government	government	government	government	commodities
commodities	agreement	commodities	commodities	spending cap	labor
president	turnover	agreement	adjustment	commodities	exchange rate
recovery	exchange rate	exchange rate	agreement	exchange rate	president
currency	adjustment	adjustment	exchange rate	bank	banks
contracts	president	target	reform	reform	target
banks	currency	president	president	agreement	minister
exchange rate	bonds	currency	currency	president	bets
credit	over-the-counter	over-the-counter	contracts	currency	euro
euro	contracts	contracts	minister	banks	oil
oil	banks	minister	bets	recovery	currency

Source: Prepared by the author (2025) with data from 2007 - 2024.
 Note: The tables contain the English free translation of Portuguese terms used in the Endogenous dictionary. The original versions in Portuguese are available in the Appendix A.

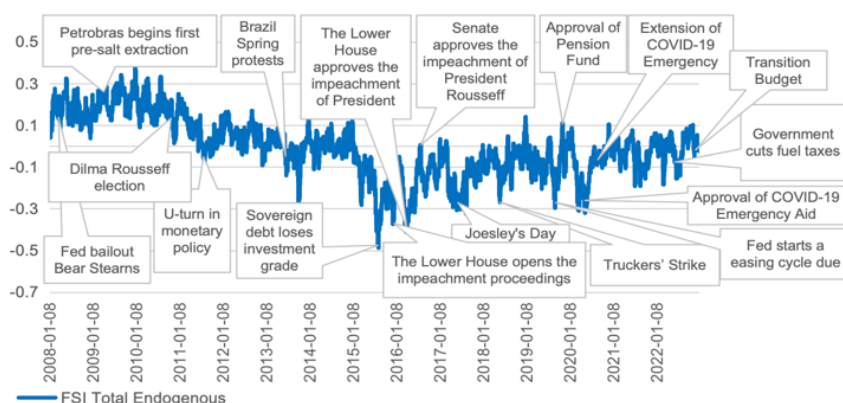
Appendix B

Figure 4A - Overall FSI Ad Hoc



Source: prepared by the author (2025) with data from AE Broadcast 2008-2022.

Figure 4B - Overall FSI Endogenous



Source: prepared by the author (2025) with data from AE Broadcast 2008-2022.

Table 6 - Correlation between IIE, EPU, BRL Brazil CDS 5y and FSI indices.

	EPU		Brazil's	FSI Total	FSI Total	FSI Total	
	IIE	Brazil	BRL/USD	CDS 5-year	Ad Hoc	Endogenous	GPT
FGV's Economic							
Uncertainty Index (IIE)	1.000	0.463	0.636	0.521	0.438	0.437	0.544
Economic Policy							
Uncertainty Index (EPU							
Brazil)	0.463	1.000	0.337	0.581	0.456	0.552	0.277
BRL/USD	0.636	0.337	1.000	0.365	0.516	0.539	0.544
Brazil's CDS 5-year	0.521	0.581	0.365	1.000	0.380	0.514	0.331
FSI Total Ad Hoc	0.438	0.463	0.516	0.380	1.000	0.683	0.595
FSI Total Endogenous	0.437	0.559	0.539	0.514	0.683	1.000	0.321
FSI Total GPT	0.544	0.279	0.544	0.331	0.595	0.321	1.000

Source: Prepared by the author (2025) with data from *AE Broadcast*, *Folha de S. Paulo*, *Estadão and Valor Econômico* (Martins, 2022), Bloomberg and *ANBIMA* data series from 2008-2022 (ANBIMA, 2021). FSI GPT Total is a simple average of FSI GPT newspaper and FSI GPT Broadcast.

Appendix C

Figure 5 – Yields-Only model parameters estimate

	Yields-only estimates	model	parameter	
	L_{t-1}	S_{t-1}	C_{t-1}	μ
L_t	0.938 (0.014)	0.064 (0.009)	0.042 (0.006)	0.017 (0.003)
S_t	0.049 (0.014)	0.934 (0.010)	-0.020 (0.006)	-0.014 (0.003)
C_t	0.038 (0.025)	-0.073 (0.017)	0.938 (0.013)	-0.013 (0.006)

	Estimated Q matrix		
	L_t	S_t	C_t
L_t	0.00013039 (0.000)	-0.00012957 (0.000)	-0.00013372 (0.000)
S_t	- 0.00012957 (0.000)	0.00013550 (0.000)	0.00011874 (0.000)
C_t	- 0.00013372 (0.000)	0.00011874 (0.000)	0.00042491 (0.000)

Test	Test Statistic	df	p-value
Likelihood			
Ratio	2473.921095	6	0.000
Wald	4.87681E-06	6	1.000

Source: Prepared by the author (2025).

Figure 6 – DRABZ model parameters estimate

Yields-Macro model parameter estimates							
	L_{t-1}	S_{t-1}	C_{t-1}	GDP 12m ahead $t-1$	Selic 12m ahead $t-1$	IPCA 12m ahead $t-1$	μ
L_t	0.930 (0.000)	0.039 (0.000)	-0.006 (0.000)	1.668 (0.000)	0.668 (0.000)	0.079 (0.000)	0.001 (0.000)
S_t	0.069 (0.000)	0.951 (0.000)	0.034 (0.000)	-1.673 (0.000)	-0.586 (0.000)	-0.355 (0.000)	-0.000 (0.000)
C_t	-0.071 (0.000)	-0.004 (0.000)	0.925 (0.000)	-1.891 (0.000)	-1.231 (0.000)	-2.344 (0.000)	-0.001 (0.000)
GDP 12m ahead t	-0.003 (0.000)	0.000 (0.000)	-0.012 (0.000)	0.537 (0.000)	-0.330 (0.000)	-0.097 (0.000)	-0.001 (0.000)
Selic 12m ahead t	-0.051 (0.000)	0.033 (0.000)	-0.062 (0.000)	-0.388 (0.000)	0.275 (0.000)	-1.052 (0.000)	-0.014 (0.000)
IPCA 12m ahead t	0.004 (0.000)	-0.002 (0.000)	0.003 (0.000)	0.014 (0.000)	0.052 (0.000)	0.924 (0.000)	0.000 (0.000)

Estimated Q matrix						
	L_t	S_t	C_t	GDP 12m ahead t	Selic 12m ahead t	IPCA 12m ahead t
L_t	0.00003675 (0.000)	-0.00003917 (0.000)	-0.00003642 (0.000)	-0.00001590 (0.000)	-0.00002991 (0.000)	0.00000055 (0.000)
S_t	-0.00003917 (0.000)	0.00005531 (0.000)	0.00001236 (0.000)	0.00001444 (0.000)	0.00002793 (0.000)	-0.00000034 (0.000)
C_t	-0.00003642 (0.000)	0.00001236 (0.000)	0.00033033 (0.000)	0.00003759 (0.000)	0.00001134 (0.000)	0.00000154 (0.000)
GDP 12m ahead t	-0.00001590 (0.000)	0.00001444 (0.000)	0.00003759 (0.000)	0.00001350 (0.000)	0.00001179 (0.000)	-0.00000203 (0.000)
Selic 12m ahead t	-0.00002991 (0.000)	0.00002793 (0.000)	0.00001134 (0.000)	0.00001179 (0.000)	0.00003549 (0.000)	-0.00000105 (0.000)
IPCA 12m ahead t	0.00000055 (0.000)	-0.00000034 (0.000)	0.00000154 (0.000)	-0.00000203 (0.000)	-0.00000105 (0.000)	0.00000090 (0.000)

Test	Test Statistic	df	p-value
Likelihood			
Ratio	221.6711103	30	4.05E-31
Wald	6.83734E-07	30	1.000

Source: Prepared by the author (2025).

Figure 7 – DRABZ-F model parameters estimate

Yields-Macro DRABZ-F model parameter estimates

	L _{t-1}	S _{t-1}	C _{t-1}	GDP 12m ahead t-1	Selic 12m ahead t-1	FSI _{t-1}	μ
L_t	0.932	0.038	-0.001	1.659	0.674	0.058	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
S_t	0.064	0.950	0.031	-1.661	-0.581	-0.372	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
C_t	-0.077	-0.003	0.926	-1.872	-1.263	-2.521	-0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GDP 12m ahead t	-0.003	-0.000	-0.012	0.542	-0.329	-0.062	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Selic 12m ahead t	-0.053	0.031	-0.058	-0.398	0.255	-1.062	-0.013
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FSI_t	0.003	-0.002	0.003	0.015	0.052	0.921	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Estimated Q matrix

	L	S	C	GDP 12m ahead	Seli 12m ahead	FSI
L	0.00003579	-0.00003883	-0.00003663	-0.00001564	-0.00002892	0.00000056
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
S	-0.00003883	0.00005834	0.00000933	0.00001478	0.00002610	-0.00000058
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
C	-0.00003663	0.00000933	0.00034117	0.00003856	0.00001156	0.00000190
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GDP 12m ahead	-0.00001564	0.00001478	0.00003856	0.00001398	0.00001237	-0.00000200
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Selic 12m ahead	-0.00002892	0.00002610	0.00001156	0.00001237	0.00003468	-0.00000148
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FSI	0.00000056	-0.00000058	0.00000190	-0.00000200	-0.00000148	0.00000078
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Test	Test Statistic	df	p-value
Likelihood			
Ratio	83.24176214	30	6.69E-07
Wald	6.70156E-07	30	1.000

Source: Prepared by the author (2025).