What the DI Curve Is Actually Pricing

Separating Copom Expectations from Risk Premia in Brazil's Rate Curve

MARCOS MOLLICA

Monthly ACM estimation · 2010–2026 · Newey-West HAC

When the 3-year DI rate moves 20 basis points in a session, the headline explanation is almost always the same: the market repriced its Selic call. That explanation is frequently wrong, or at best incomplete. A rate move can reflect a genuine revision to the expected path of the policy rate – or it can reflect investors demanding more, or less, compensation to hold Brazilian duration for the same expected path. These are economically distinct events. One is about what Copom will do; the other is about what the market charges to bet on it. Most curve-watchers conflate them by construction, because the DI rate itself doesn't distinguish the two – it only reports their sum:

$$y_t^{(n)} = \underbrace{\frac{1}{n}\sum_{i=0}^{n-1} E_t[r_{t+i}]}_{\text{expected average Selic path}} + \underbrace{\phi_t^{(n)}}_{\text{term premium}}$$

where $y_t^{(n)}$ is the n-period DI rate observed at time t and r_t is the short (one-period) policy rate.

This note separates the two terms on the right, using sixteen years of Brazilian DI futures data (2010–2026) and two complementary tools: a statistical decomposition of what actually moves the curve, and a no-arbitrage model of why.

Methodology

Layer 1: What Moves the Curve

Principal component analysis on the DI curve's daily changes recovers a small number of factors that explain nearly all of its variation. Writing $\Delta y_t \in \mathbb{R}^N$ for the vector of daily rate changes across N curve vertices, PCA solves for the eigenvectors of the covariance matrix $\Sigma = \text{Cov}(\Delta y_t)$:

$$\Sigma w_k = \lambda_k w_k, \qquad k = 1, 2, 3, \ldots$$

ranked by eigenvalue λ_k. The first three loadings w_1, w_2, w_3 are, respectively, a level factor (the whole curve moving together, same sign at every maturity), a slope/twist factor (short and long ends moving in opposite directions, one sign change), and a curvature factor (the belly moving against the wings, two sign changes). This is standard technique in fixed income, but running it separately across four distinct Brazilian episodes – the calm 2010–13 period, the 2015–16 political and fiscal crisis, the aggressive 2021–22 post-pandemic hiking cycle, and today – lets us ask a sharper question: is the shape of w_1, w_2, w_3 stable across regimes, or does it change?

Layer 2: Why the Curve Moves

A statistical factor tells you how the curve moved; it says nothing about why. Splitting a DI yield into the two terms of the identity above requires a model of the pricing kernel. We use the regression-based no-arbitrage affine framework of Adrian, Crump and Moench (2013) – "ACM" below – estimated monthly on the curve from 1 to roughly 110 months.

State variables and dynamics. Let $X_t \in \mathbb{R}^3$ be the level/slope/curvature factors extracted by PCA on the level (not change) of yields, standardized before extraction. The factors are assumed to follow a Gaussian VAR(1) under the physical measure:

$$X_{t+1} = \mu + \Phi X_t + v_{t+1}, \qquad v_{t+1} \sim N(0, \Sigma_v)$$

Pricing kernel. Bond prices are priced by a stochastic discount factor with time-varying prices of risk $\lambda_t = \lambda_0 + \lambda_1 X_t$:

$$m_{t+1} = \exp\left(-y_t^{(1)} - \tfrac{1}{2}\lambda_t'\lambda_t - \lambda_t' v_{t+1}\right), \qquad \lambda_t = \lambda_0 + \lambda_1 X_t$$

The prices of risk $\lambda_0 \in \mathbb{R}^3$ and $\lambda_1 \in \mathbb{R}^{3\times3}$ are estimated in three regression steps rather than by maximum likelihood, following ACM: (i) the VAR above is estimated by OLS to obtain the innovations v_{t+1}; (ii) one-month holding-period excess returns $rx_{t+1}^{(n)}$ on each maturity n are regressed on $[1, v_{t+1}, X_t]$, for every n, giving loadings β_n on the shocks; (iii) a cross-sectional regression of the step-(ii) intercepts and slopes on β_n recovers λ_0, λ_1 without ever needing a survey of analysts' rate expectations. This is the appeal of the no-arbitrage approach: the curve's own cross-section and time series are enough to identify the price of risk, provided the model is well specified.

Risk-neutral dynamics and the affine recursion. Given λ_0, λ_1, the risk-neutral measure has drift $\mu^Q = \mu - \Sigma_v\lambda_0$ and persistence $\Phi^Q = \Phi - \Sigma_v\lambda_1$.

$$\mu^{\mathrm{Q}} = \mu - \Sigma_v \lambda_0, \qquad \Phi^{\mathrm{Q}} = \Phi - \Sigma_v \lambda_1$$

Log bond prices are affine in the state, $p_t^{(n)} = A_n + B_n'X_t$, with coefficients obtained by the standard recursion (short-rate equation $y_t^{(1)} = \delta_0 + \delta_1'X_t$):

$$\begin{aligned} A_n &= A_{n-1} + B'_{n-1}\mu^{\mathrm{Q}} + \tfrac{1}{2}B'_{n-1}\Sigma_v B_{n-1} - \delta_0, \\ B'_n &= B'_{n-1}\Phi^{\mathrm{Q}} - \delta'_1, \qquad A_1 = -\delta_0, \quad B_1 = -\delta_1 \end{aligned}$$

The model-fitted yield $y_t^{(n),\mathrm{model}}$ is then compared against the yield implied by the same short-rate equation evaluated under the physical measure with no risk pricing – i.e. the literal expected average short rate:

$$y_t^{(n),\mathrm{model}} = -\, p_t^{(n)}/n$$

$$y_t^{(n),\mathrm{EH}} = \frac{1}{n}\sum_{i=0}^{n-1} E_t\left[\delta_0 + \delta_1' X_{t+i}\right], \qquad E_t[X_{t+i}] = \Phi^i X_t + \sum_{j=0}^{i-1}\Phi^j \mu$$

The term premium is the gap between the two:

$$\phi_t^{(n)} = y_t^{(n),\mathrm{model}} - y_t^{(n),\mathrm{EH}}$$

Two corrections were necessary to get this to behave over a sample spanning Selic at both 2% and 15%. First, a Bauer–Rudebusch–Wu (2012) bootstrap correction for the well-known small-sample downward bias in Φ – near-unit-root persistence combined with a limited monthly sample (200 observations) otherwise produces an explosive risk-neutral matrix Φ^Q, and model yields diverge at long maturities. Second, and more consequentially, standardizing the factors before extraction rather than using raw yield levels: an unstandardized PC1 inherits the full historical range of Brazilian rates directly into Σ_v, and the convexity term compounds that inflated variance every month out to n periods, producing a long-end bias of several percentage points. Standardizing brought the model back into a Q-measure that was stationary without needing to impose a constraint at all, and cut the long-end fit error by roughly two-thirds.

Result 1: The Factors Are Remarkably Stable in Shape, Not in Scale

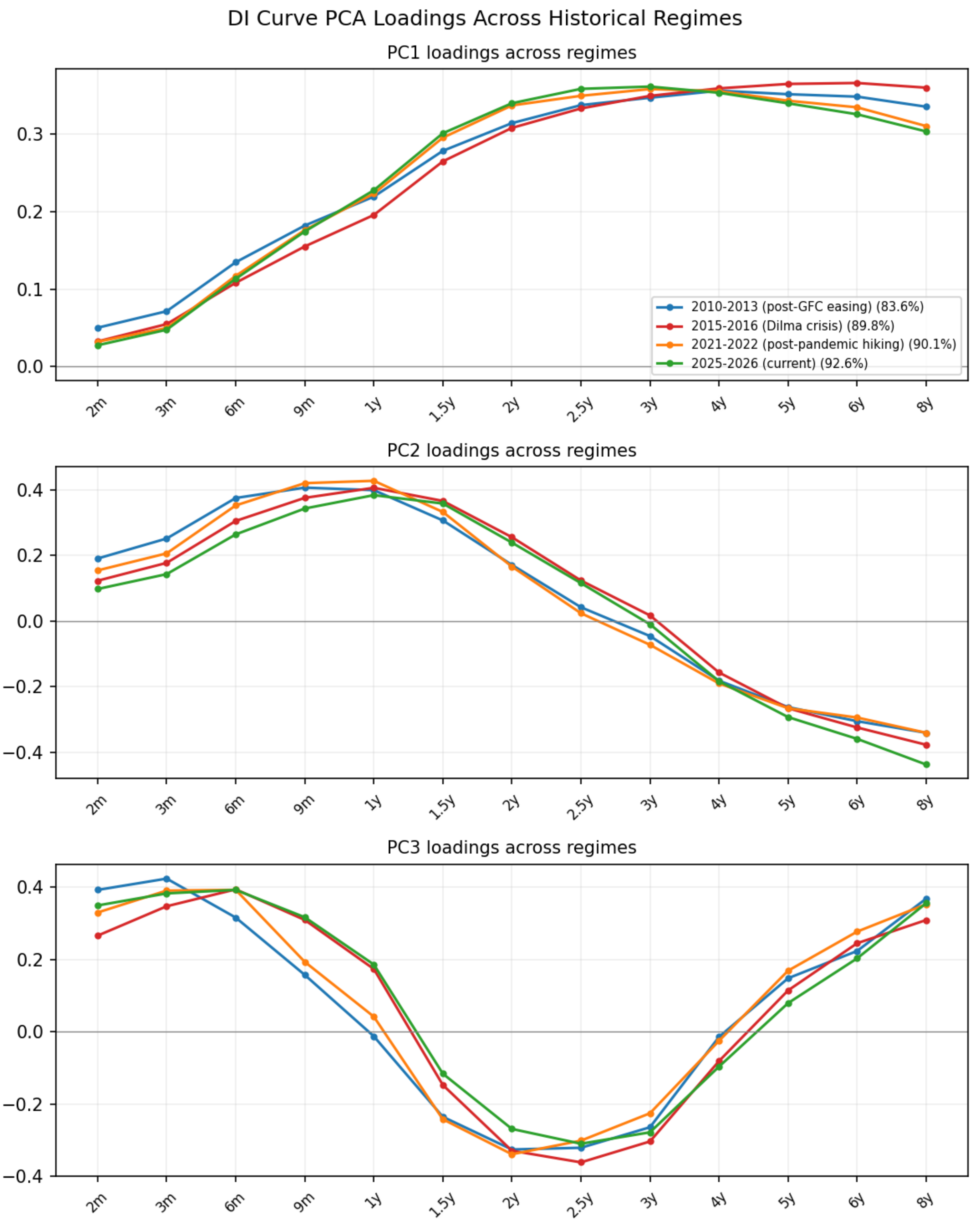

Figure 1: PCA factor loadings across four Brazilian rate regimes.

Across sixteen years, a Dilma-era political and fiscal crisis, a 1,175-basis-point hiking cycle, and whatever the current regime turns out to be, the shape of w_1, w_2, w_3 barely moves. The slope factor's pivot sits consistently around three years; the curvature factor's belly-trough sits consistently around two to two-and-a-half years. What moves with the regime is not the shape but the concentration: the level factor's share of variance rises from 83% in the calm 2010–13 period to 90–94% during the crisis, hiking, and current episodes. Quieter markets leave more room for idiosyncratic, vertex-specific noise; stressed or fast-moving ones reprice the whole curve on a single dominant theme. The structure of the Brazilian curve is not fragile – the volume of what moves it is.

The structure of the Brazilian curve is not fragile – the volume of what moves it is.

That stability is not just a descriptive curiosity. Because w_1, w_2, w_3 hold their shape across regimes, a vertex's residual against the factor model – what it does after the common level/slope/curvature moves are stripped out – is a comparatively clean read on that vertex's own idiosyncratic pricing, rather than an artifact of whichever regime happens to be in effect. That residual is exactly the raw material a relative-value screen across DI vertices needs: which points on the curve are cheap or rich net of what the common factors already explain. This is a direct, structural link between the two halves of this note – the same PCA that recovers w_1, w_2, w_3 recovers the object an RV signal is built from, developed with a worked example below.

A Relative-Value Application

Expressed pairwise and DV01-neutral across tradable DI contracts, the residual described above is the natural input to a relative-value screen: which points on the curve are cheap or rich net of what the common factors already explain, hedged against those same common factors so a position expresses the residual mispricing specifically rather than a directional or curve-shape view available more cheaply elsewhere.

Figure 2 shows this screen applied to twelve live, quarterly-cycle DI contracts (January, April, July, and October expiries, spanning the front three years of the curve) as of the most recent available date. Cells report the pairwise DV01-neutral spread residual in basis points, shaded by statistical conviction against each pair's own trailing 12-month distribution – a weak flag at 1.5 standard deviations, a strong flag at 2.5 – rather than by a fixed spread level, since what counts as an extreme differs by pair and by regime, consistent with the stability result above.

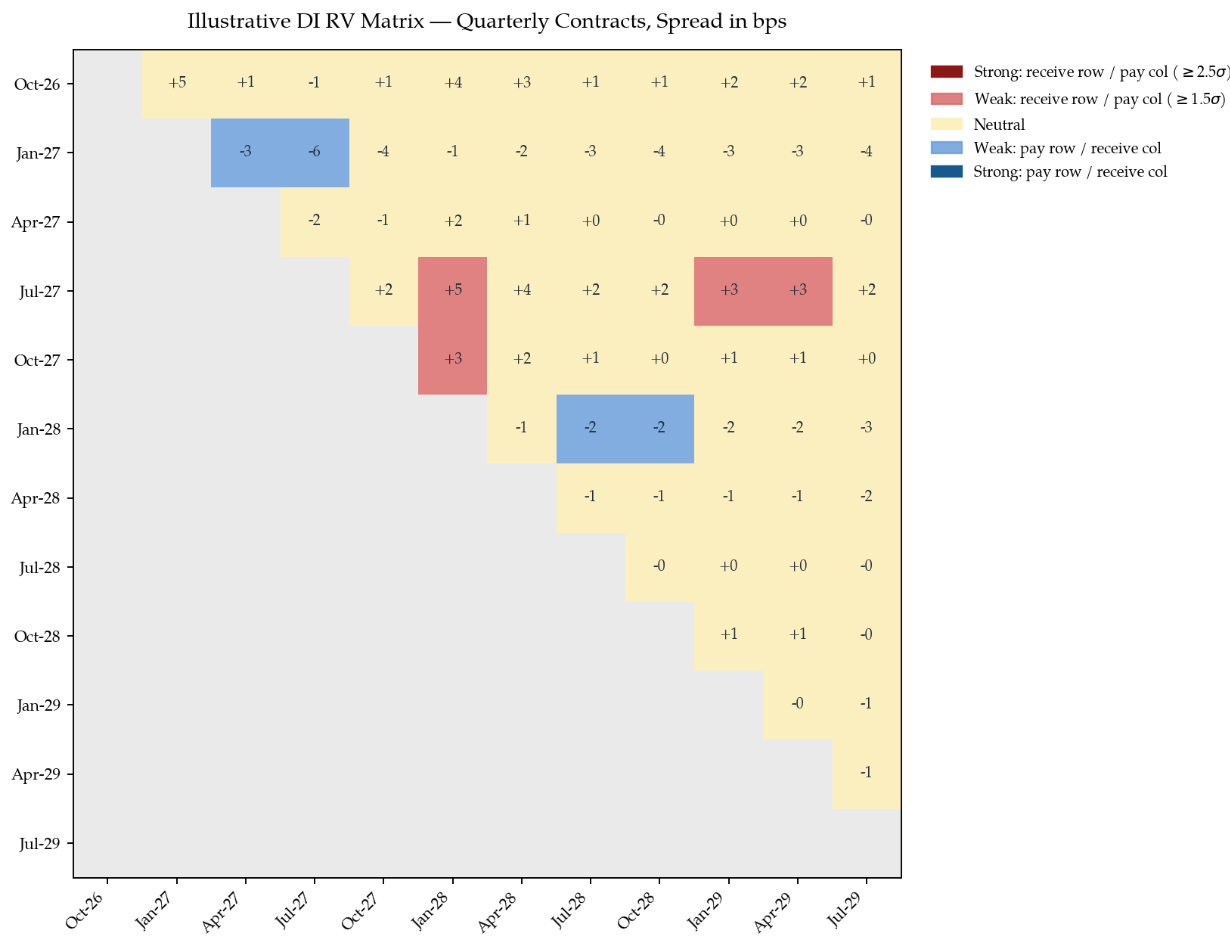

Figure 2: Illustrative DI relative-value matrix, quarterly contracts, DV01-neutral pairwise spreads in bps, conviction-shaded against each pair's trailing distribution.

This is deliberately a compact, illustrative example rather than the full production screen: the same construction extends directly to every live monthly DI contract, not only the quarterly subset shown here, and in practice is re-computed continuously as new curve prints arrive rather than evaluated at a single snapshot. What the example demonstrates is the methodology and the shape of its output – which pairs on the curve are statistically stretched right now, net of the common factors – not a trading strategy. The specific rules governing which flagged pairs are actually traded, at what size, and with what holding-period discipline are a separate, undisclosed layer built on top of this screen.

Result 2: Where the Risk Premium Actually Lives

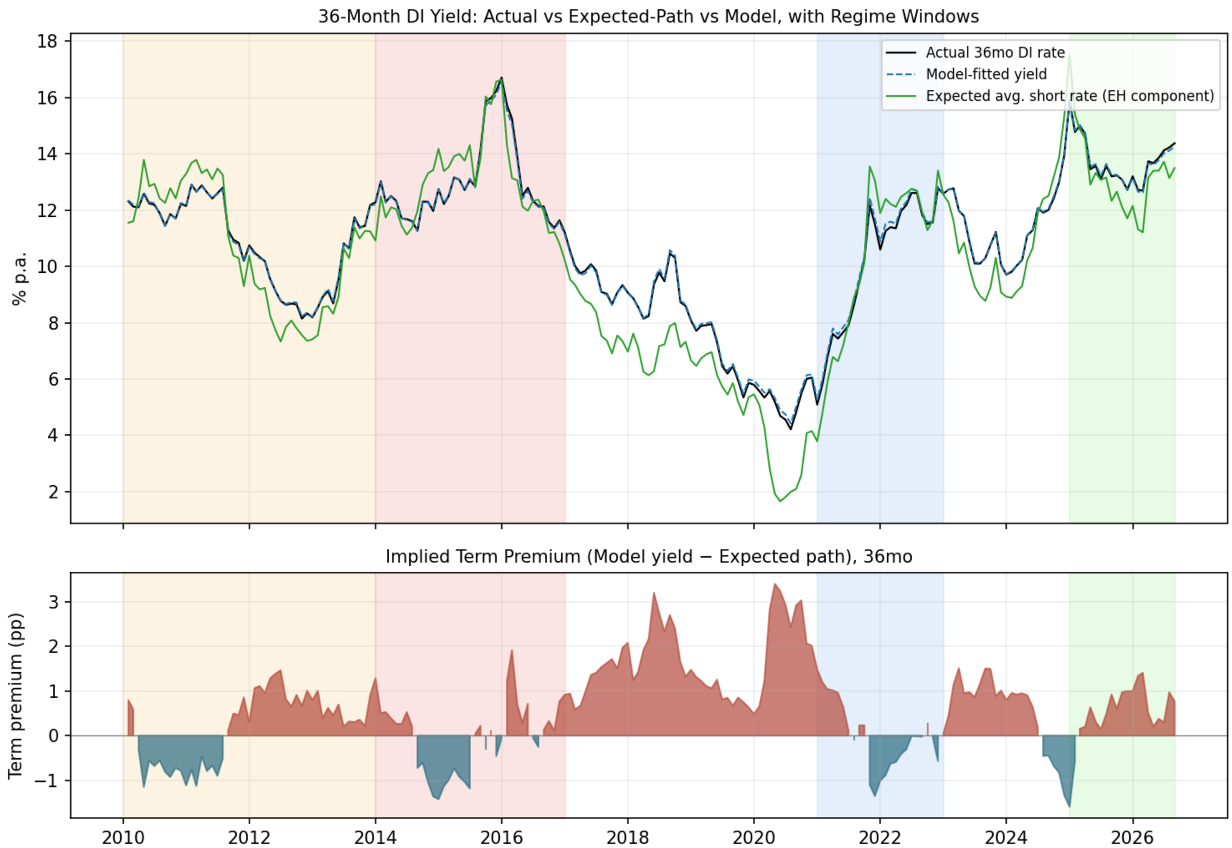

Figure 3: 36-month DI yield: actual vs. expected path vs. model, with implied term premium.

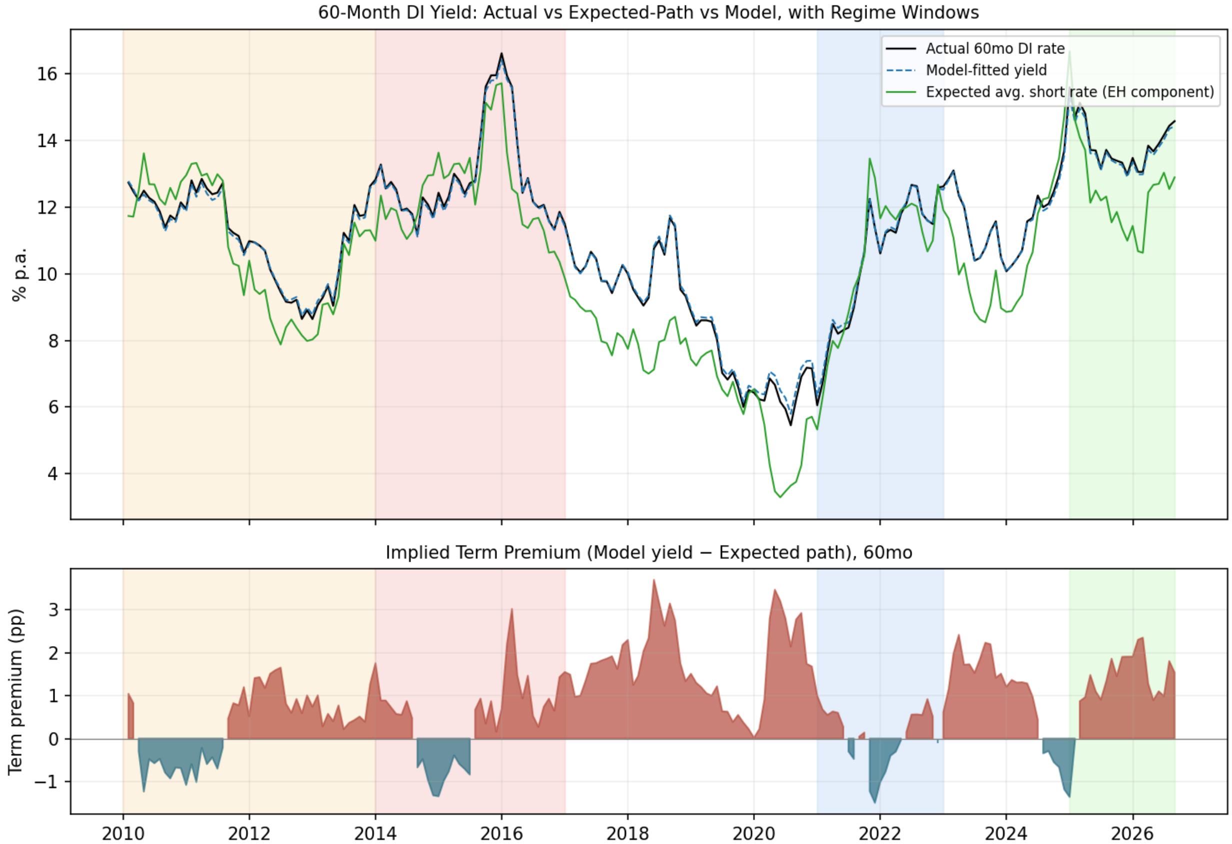

Figure 4: 60-month DI yield: actual vs. expected path vs. model, with implied term premium.

Three episodes stand out, and each tells a different story about what the market was actually pricing.

2015–16 produced the highest term premium of the entire sample – $\phi^{(36)} \approx +1.9$pp, $\phi^{(60)} \approx +2.3$pp, right at the peak of the political and fiscal crisis. This is direct empirical evidence that the market was not merely revising its Selic forecast during that period; it was charging real, measurable compensation to hold Brazilian duration on top of whatever it expected the policy rate to do. That finding sits comfortably alongside the fiscal dominance argument this author has made elsewhere: when the sustainability of the fiscal path itself is in question, the premium investors demand for duration is one of the first places it shows up, well before it shows up in the level of rates.

2021–22 is the counter-intuitive result. Despite one of the fastest hiking cycles in Brazilian history – Selic from 2% to 13.75% in under two years – the term premium went negative, troughing near $\phi^{(36)} \approx -1.3$pp in early 2022. The most economically coherent reading is that an aggressive, front-loaded, credible tightening cycle reduced rather than added to long-run uncertainty: the market seems to have priced high conviction in a subsequent easing cycle more strongly than the model's own physical-measure forecast captured. Whether that is a genuine feature of credible front-loaded hiking or an artifact of the specific episode is a question this note doesn't fully resolve.

Today, the term premium is positive and has been rising – roughly $\phi^{(36)} \approx +1.2$pp and $\phi^{(60)} \approx +1.6$pp as of the most recent observation, materially above its sixteen-year average of $\bar{\phi}^{(36)} = 0.38$pp ($\bar{\phi}^{(60)} = 0.61$pp). Some of that reading likely reflects the residual fitting bias the model still carries at longer maturities, so the precise magnitude should be treated with some caution. The direction – elevated and rising, and by a wide margin the highest reading since the 2015–16 crisis – is the more defensible part of the finding, and the timing is not subtle: Brazil is headed into one of its closest presidential races in a generation, with an October 4 first round and a near-certain October 25 runoff between President Lula, seeking a fourth term, and Senator Flávio Bolsonaro, recent polling putting the second-round scenario at a dead heat. Political risk is the most economically plausible single driver of a premium spike of this size at this particular moment, in the same way fiscal and political risk drove the 2015–16 spike – the market is not merely debating Copom's next move, it is pricing genuine uncertainty about who sets fiscal and monetary policy direction from 2027 onward.

Controlling for the US Term Premium

A natural objection to Result 2: some of the co-movement in $\phi_t^{(n)}$ could simply be global duration risk being repriced everywhere at once, rather than anything Brazil-specific. This objection has particular force right now, for a reason worth stating plainly – US Treasuries are not the clean risk-free benchmark they were for most of the post-GFC decade. The New York Fed's own published ACM decomposition of the 10-year Treasury (the identical methodology used in this note, applied to the US curve) shows the US term premium sitting near zero or negative for nearly all of 2016–2022, then turning decisively positive from 2024 onward – 0.43% in April 2026, and still running near 0.5–0.8% through mid-2026. Controlling for the level of US yields would be the wrong fix, since that level now embeds a growing premium of its own; the correct control is the US term premium specifically, sourced from the same no-arbitrage framework rather than a raw yield.

Before specifying that control, the same discipline this author has applied elsewhere to this question – test the choice of levels versus changes rather than assume it – applies here too, and here the test result and the specification actually chosen diverge, which deserves to be stated plainly rather than glossed over. Unit-root tests (ADF and KPSS, run jointly) on the two Brazilian term premium series show a consistent stationary signature: ADF rejects a unit root, and KPSS does not reject stationarity. The US 2-year and 10-year ACM term premia show the opposite signature – ADF fails to reject a unit root, and

KPSS rejects stationarity. Allowing for a single endogenous structural break (Zivot–Andrews) rather than a simple trend does not change this: the US series still fail to reject a unit root even with a break permitted, while Brazil's series look, if anything, more clearly stationary under the same test.

No unit-root test, however constructed, can distinguish a genuine unit root from a level shift that simply hasn't finished reverting yet when the entire post-shift sample is two years long.

Taken at face value, that evidence points toward differencing. This note uses levels anyway, and the reason is a judgment call, stated explicitly rather than left implicit. A term premium is risk compensation, not a persistently trending economic quantity; there is no structural reason for it to random-walk to infinity, and the ACM/Kim–Wright literature this note draws on treats it as mean-reverting by construction. The US series has been trending only since 2024 – roughly two years of monthly data – after sitting near zero or negative for most of the prior decade. That is a sample-size limitation on the test, not evidence for either hypothesis. Given that, treating the US premium as I(1) because it has not yet shown mean reversion risks mistaking a still-unresolved regime for a permanent trend – exactly the kind of judgment this note is generally skeptical of making without testing, made here on theoretical rather than statistical grounds because the statistics cannot currently settle it either way. A levels regression of $\phi_t^{(n),BR}$ on $\phi_t^{(2y),US}$ and $\phi_t^{(10y),US}$ is used on that basis.

Using the NY Fed's published monthly ACM series for the US 2-year and 10-year term premia, each Brazilian term premium is regressed contemporaneously on both, by OLS with Newey–West HAC standard errors:

$$\phi_t^{(n),\mathrm{BR}} = \alpha + \beta_1 \phi_t^{(2y),\mathrm{US}} + \beta_2 \phi_t^{(10y),\mathrm{US}} + \varepsilon_t^{(n)}$$

The residual $\varepsilon_t^{(n)}$ is, by construction, the Brazil-specific component of the term premium – whatever is left after netting out its co-movement with global duration compensation.

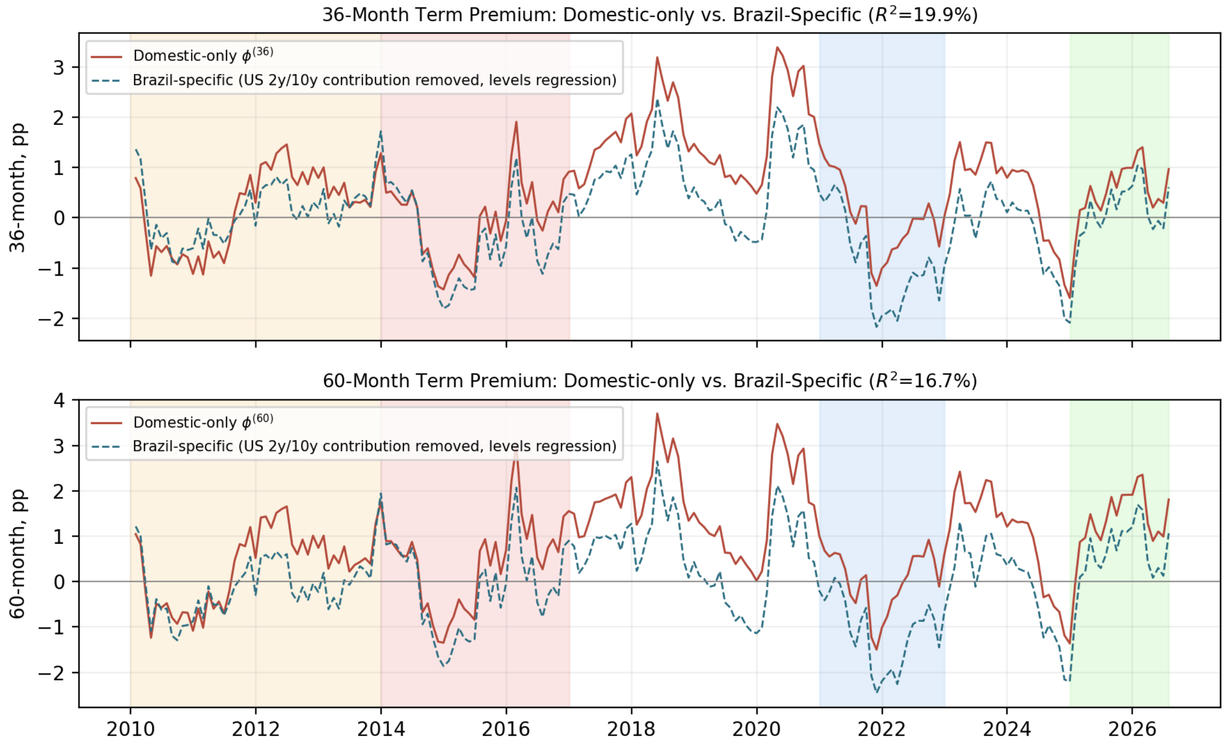

Figure 5: 36- and 60-month Brazilian term premium: domestic-only vs. Brazil-specific component, net of the US 2y/10y ACM term premium (levels regression).

The headline result is the low R^2: **19.9% at 36 months, 16.7% at 60 months**. Global duration risk explains a modest but non-trivial share of the variation in Brazil's own term premium; the majority of what moves $\phi_t^{(n),BR}$ is still not a beta to US Treasuries, but the fraction that is is larger than a naive reading of "the two are unrelated" would suggest. That is, on its own, a useful finding: it partially, not wholly, argues against the objection this section set out to address.

Even so, the residual is not identical to the raw series, and the differences are informative rather than cosmetic. In every regime, $\varepsilon_t^{(n)}$ sits below the raw $\phi_t^{(n)}$ – the calm 2010–13 period, which showed a small positive raw premium, is essentially flat to slightly negative once global duration compensation is removed, suggesting that period's modest premium was mostly imported rather than Brazil-specific. The 2021–22 episode, already the paper's most surprising result, becomes more pronounced once controlled – the residual troughs deeper than the raw series, because global term premia were also unusually compressed over that window (the Fed at the zero bound), so netting out the global component reveals an even larger Brazil-specific collapse in compensation than the headline number suggested. And today's reading survives the control, smaller but clearly intact: as of July 2026, the raw $\phi^{(36)}$ of 0.97pp nets to a residual of 0.61pp, and the raw $\phi^{(60)}$ of 1.81pp nets to a residual of 1.06pp – both comfortably above the residual series' own zero mean and roughly two-thirds to six-tenths of a standard deviation into positive territory. The election-risk reading from Result 2 is weakened in magnitude by this exercise, appropriately, but not overturned.

One coefficient is worth flagging rather than over-interpreting: β_2, the loading on the US 10-year premium, comes out negative and statistically significant in both regressions ($t = -2.40$ at 36 months, $t = -2.60$ at 60 months), while β_1, the loading on the US 2-year premium, is positive but insignificant in both. Given the low overall R^2 and the correlation between the US 2- and 10-year series themselves, the negative β_2 is more likely a symptom of limited identifying variation – the regression assigning a spurious sign to whichever of two correlated regressors happens to carry the residual explanatory power once the other is included – than a real structural relationship, and should not be read as "Brazilian term premium falls when US long-end compensation rises."

Why This Matters Beyond Description

A framework that can label a given curve move as "the market changed its Copom call" versus "the market got paid more to hold duration" does more than satisfy curiosity. Once the two are separated, a natural next question presents itself: which of the two – the expectations path, or the premium being charged for it – looks out of line with its own recent history, and at which point on the curve? That question is directly tradable on the term-premium side, and it sits on top of the complementary relative-value question already developed above: net of the common factors, which vertices are mispriced relative to each other, right now? Both are versions of the same underlying idea – a persistently elevated premium, or a persistently stretched factor residual, at a given maturity is, by construction, the market asking to be paid more than usual to hold that risk, or offering a mispriced spread against its neighbors.

This note is a first look at a line of work still in progress. A fuller treatment – with error bands on the premium estimates, a closer look at what is actually driving the unresolved 2021–22 result, and a more complete relative-value evaluation – is forthcoming.

APPENDIX · DETAILED ESTIMATION PROCEDURE

This appendix gives the complete, step-by-step estimation procedure behind Result 2 and the US-control exercise, in enough detail to reproduce every figure in this note from the underlying DI futures data alone. It is intended as a companion to the methodology section, not a restatement of it: the main text gives the model; this appendix gives every implementation choice, including two that materially affect the result and are not obvious from the model's equations alone.

A.1 From Bloomberg Contracts to a Fitted Curve

Raw data. Daily settlement prices for all listed DI1 futures contracts (Bloomberg tickers of the form ODmyy Comdty, e.g. ODF10 for the January 2010 contract), 2010–2026, plus the B3 trading holiday calendar.

Expiry and business-day counting. Each contract's ticker encodes its expiry month via a standard futures month code (F=Jan, G=Feb, H=Mar, ..., Z=Dec) and a two-digit year. DI1 contracts expire on the first business day of the expiry month. Days-to-expiry d_u for a given trade date is the count of B3 business days between that date and expiry, using the official holiday calendar (a fixed list of national and financial-sector holidays through 2099). Only observations with $d_u > 0$ are retained.

Curve interpolation. For each trade date, the set of (days-to-expiry, rate) pairs across all live contracts is interpolated to a target maturity via flat-forward interpolation on log discount factors:

$$P(d_u) = (1 + y(d_u))^{-d_u/252}, \qquad \log P(d_{uf}) = \text{interp}(d_{uf}, d_u, \log P(d_u))$$

with linear interpolation performed in (d_u, log P) space, then converted back to an annualized rate. This is standard flat-forward curve construction and preserves no-arbitrage consistency between adjacent points on the curve.

Two panels are built from this common interpolation routine:

• **Daily panel**, a fixed set of 13 standard tenors (2m, 3m, 6m, 9m, 1y, 1.5y, 2y, 2.5y, 3y, 4y, 5y, 6y, 8y in business-day terms), used for Layer 1 (Result 1, Figure 1).

• **Monthly panel**, a fine grid of 112 vertices at every 21-business-day multiple (approximately 1 to 110 months), resampled to month-end, used for Layer 2 (the ACM estimation). Only maturities with a fully populated history (no missing observations across the full 2010–2026 window) are retained, which leaves 112 usable monthly maturities.

A.2 Layer 1: Regime-Conditional PCA (Result 1)

For each of the four regime windows, principal components are extracted from the daily changes of the 13-tenor panel:

$$\Sigma w_k = \lambda_k w_k, \qquad k = 1, 2, 3, \ldots$$

with the sign of each loading vector fixed by convention (level loadings normalized positive; slope loadings normalized positive at the short end; curvature loadings normalized negative at the belly) purely for visual comparability across regimes — this has no effect on the variance decomposition itself. The variance share reported in Figure 1's legend is $\lambda_1/\Sigma_k\lambda_k$.

This layer is fully independent of Layer 2 below: it uses no VAR, no risk pricing, and no affine recursion. It is included here because the same PCA construction — applied to yield levels rather than changes, and standardized rather than raw — supplies the state variables for Layer 2.

A.3 Layer 2: The ACM Estimation, Step by Step

A.3.1 State Variables

The monthly panel of 112 maturities, converted to decimal, is standardized column-wise (zero mean, unit variance) before extraction:

$$Y_z = \frac{Y - \bar{Y}}{\sigma_Y}, \qquad \text{Cov}(Y_z) = \Sigma_Y w_k = \lambda_k w_k$$

The first three eigenvectors give raw factor scores $X_t^{raw} = Y_z w_{1:3}$. A second normalization step, not explicit in the main text's equations, is necessary here: X_t^{raw} inherits the eigenvalue scale of Σ_Y directly — in practice the level factor's eigenvalue is roughly two orders of magnitude larger than the curvature factor's, since a single dominant level factor absorbs most of the curve's variance. Left uncorrected, this scale disparity produces a poorly conditioned covariance matrix for the VAR innovations (see A.3.2) and an explosive risk-neutral transition matrix. The factor scores are therefore rescaled to unit variance component-wise before proceeding:

$$X_t = X_t^{\text{raw}} \,/\, \text{std}(X_t^{\text{raw}})$$

This rescaling is what the main text's "standardizing the factors before extraction… brought the model back into a Q-measure that was stationary without needing to impose a constraint" is describing operationally — standardizing the input yields alone is necessary but not sufficient; the extracted scores need the same treatment.

A.3.2 VAR(1) and the Bauer-Rudebusch-Wu Correction

$$X_{t+1} = \mu + \Phi X_t + v_{t+1}, \qquad v_{t+1} \sim N(0, \Sigma_v)$$

is first estimated by OLS, giving $\hat{\mu}_{OLS}$, $\hat{\Phi}_{OLS}$. Bauer–Rudebusch–Wu (2012) bootstrap bias correction proceeds as follows: 2,000 simulated paths are generated from $\hat{\mu}_{OLS}$, $\hat{\Phi}_{OLS}$ using residuals resampled with replacement from the OLS fit; each simulated path is itself refit by OLS; the average simulated $\hat{\Phi}$ across replications, less $\hat{\Phi}_{OLS}$, gives the estimated small-sample bias. The correction $\Phi = \hat{\Phi}_{OLS} - \delta \cdot \text{bias}$ is applied with $\delta = 1$, the standard, full correction — provided the internal-consistency requirement below is met; see the note at the end of this subsection for why this qualification matters.

A genuine implementation subtlety, worth stating explicitly because it materially affects the result: the innovations v_t used downstream, in the excess-return regressions of A. 3.4, must be computed using the same, bias-corrected Φ used for the risk-neutral dynamics in A.3.5 and the physical-measure forecast in A.3.6 — not the uncorrected $\hat{\Phi}_{OLS}$. That is,

$$v_t = X_t - (\mu + \Phi X_{t-1})$$

using the corrected Φ, not $\hat{\Phi}_{OLS}$. Using $\hat{\Phi}_{OLS}$-based residuals while using the corrected Φ elsewhere introduces an internal inconsistency between the shocks entering the risk-price estimation and the dynamics those risk prices are subsequently used to price — the shocks and the model they are priced against no longer describe the same process. This does not show up as an estimation error in any individual step; each regression still runs and reports a high R^2. It shows up only in the aggregate shape and magnitude of the recovered term premium, and specifically in the ability of the model to distinguish genuinely different episodes (a level-driven crisis from a slope-driven hiking cycle) rather than compressing all episodes toward a similar, muted average. With this consistency requirement satisfied, the standard $\delta = 1$ correction performs as the theory would suggest, without further adjustment.

A.3.3 Short-Rate Equation

$$y_t^{(1)} = \delta_0 + \delta_1' X_t + \varepsilon_t$$

estimated by OLS on the shortest available monthly maturity, using the same standardized, rescaled X_t. In-sample fit is very high ($R^2 > 0.98$ throughout the estimation window), as expected given the short end is close to a linear combination of the three dominant factors by construction.

A.3.4 Excess-Return Regressions

For each of 16 maturities $n \in \{2, 3, 6, 9, 12, 15, 18, 24, 30, 36, 48, 60, 72, 84, 96, 108\}$ months, the one-month holding-period log excess return is

$$rx_{t+1}^{(n)} = \log P_{t+1}^{(n-1)} - \log P_t^{(n)} - \tfrac{1}{12} \log(1 + y_t^{(1)})$$

where $P(\cdot)$ denotes the bond price implied by the corresponding decimal yield, $P(y,n) = (1+y)^{-n/12}$. The risk-free leg here is scaled by 1/12 because $y_t^{(1)}$ is quoted on an annualized basis (the DI market convention) while the holding period is one month — an easy point to get wrong, since omitting the 1/12 produces excess returns that are uniformly biased downward by a constant close to $\log(1+y^{(1)})$ regardless of maturity, a distinctive symptom (excess returns come out large, negative, and essentially maturity-independent) worth checking for directly if reproducing this step from scratch.

Each maturity's excess return is regressed on the contemporaneous VAR innovation and the lagged state:

$$rx_{t+1}^{(n)} = a_n + \beta_n' v_{t+1} + c_n' X_t + e_{t+1}^{(n)}$$

giving, for each of the 16 maturities, an intercept a_n, a (3×1) loading on the shocks β_n, and a (3×1) loading on the lagged state c_n. Stacking across the 16 maturities gives $a \in \mathbb{R}^{16}$, $B \in \mathbb{R}^{16\times3}$, $C \in \mathbb{R}^{16\times3}$.

A.3.5 Risk-Price Recovery

This is the step where the pricing kernel stated in the main text is actually converted into estimating equations. The no-arbitrage restriction implied by this kernel is

$$E_t[rx_{t+1}^{(n)}] + \tfrac{1}{2}\,\mathrm{Var}_t[rx_{t+1}^{(n)}] = \mathrm{Cov}_t(rx_{t+1}^{(n)}, v_{t+1}')\,\lambda_t$$

Given the regression of A.3.4, $\mathrm{Cov}_t(rx_{t+1}^{(n)}, v_{t+1}) = \Sigma_v\beta_n$ (since rx's only source of covariance with v_{t+1} runs through the $\beta_n' v_{t+1}$ term), so matching the constant and state-dependent parts of the no-arbitrage condition against the regression coefficients gives, for each maturity n:

$$a_n + \mathrm{conv}_n = \beta_n'\Sigma_v\lambda_0, \qquad c_n = \beta_n'\Sigma_v\lambda_1$$
$$\lambda_0 = (B\Sigma_v)^+(a + \mathrm{conv}), \qquad \lambda_1 = (B\Sigma_v)^+C$$

where $\mathrm{conv}_n = ½(\beta_n'\Sigma_v\beta_n + \sigma_n^2)$ is the convexity adjustment, σ_n^2 being the residual variance of the maturity-n regression. Stacking across all 16 maturities and solving by (pseudo-inverse) least squares gives λ_0, λ_1 as shown above.

This is the point most likely to be implemented incorrectly, because the natural-looking simplification $\lambda_0 = B^+(a+\mathrm{conv})$ — omitting the Σ_v factor — is not equivalent to the equation above except in the degenerate case $\Sigma_v = I$. The two formulas can differ by an order of magnitude in the recovered term premium; only the Σ_v-inclusive version is consistent with the stated pricing kernel.

A.3.6 Risk-Neutral Dynamics and the Affine Recursion

Given λ_0, λ_1, the risk-neutral drift and persistence are $\mu^Q = \mu - \Sigma_v\lambda_0$, $\Phi^Q = \Phi - \Sigma_v\lambda_1$. Log bond prices are affine in the state, $p_t^{(n)} = A_n + B_n'X_t$, with the recursion

$$A_{n+1} = A_n + B_n'\mu^Q + \tfrac{1}{2}B_n'\Sigma_v B_n - \delta_0, \qquad B_{n+1}' = B_n'\Phi^Q - \delta_1'$$
$$A_1 = -\delta_0/12, \qquad B_1 = -\delta_1/12$$

The 1/12 scaling in the initial condition is necessary because δ_0, δ_1 are estimated against the annualized short rate (A.3.3) while the recursion steps monthly — the same annualized-versus-monthly distinction as A.3.4's risk-free leg, appearing here in the recursion's own units. The model-implied yield is then $y_t^{(n),\mathrm{model}} = -(A_n + B_n'X_t)/(n/12)$.

No eigenvalue clamp is applied to Φ^Q: with the internal-consistency requirement of A.3.2 satisfied, Φ^Q's eigenvalues are stationary (modulus below 1) without any additional constraint, matching the claim in the main text. An eigenvalue clamp was tested as a fallback during development and found to distort the sign and timing of the recovered term premium in specific episodes, not merely rescale its magnitude — a consequence of the clamp interacting with the model's complex-conjugate eigenvalue pair. It is not used in the final specification.

A.3.7 Expectations-Hypothesis Yield and the Term Premium

$$y_t^{(n),\mathrm{EH}} = \frac{1}{n}\sum_{i=0}^{n-1} E_t\left[\delta_0 + \delta_1' X_{t+i}\right], \qquad E_t[X_{t+i}] = \Phi^i X_t + \sum_{j=0}^{i-1} \Phi^j \mu$$

computed by direct iteration of the (physical-measure, bias-corrected) VAR forward n steps and averaging, for each of $n \in \{24, 36, 60\}$ separately. The term premium is

$$\phi_t^{(n)} = y_t^{(n),\mathrm{model}} - y_t^{(n),\mathrm{EH}}$$

A.4 The US Term Premium Control (Figure 5)

The NY Fed's published monthly ACM decomposition of the US Treasury curve — the identical three-step methodology of A.3.4–A.3.6, applied by the Fed to the US curve — supplies $\phi_t^{(2y),\mathrm{US}}$ and $\phi_t^{(10y),\mathrm{US}}$. Each Brazilian maturity's term premium, and each US series, is first tested for the order of integration (ADF and KPSS, jointly, plus Zivot–Andrews allowing one endogenous structural break as a check against a break being mistaken for a unit root). The Brazilian series test as stationary on every version of this test; the US series test as I(1) even allowing for a break, but only around two years of data exist since the US premium began trending in 2024 – too short a post-break window for any unit-root test to distinguish a genuine unit root from a level shift still finding its new equilibrium. Given that a term premium is theoretically mean-reverting risk compensation rather than a persistently trending quantity, and that the statistical evidence for treating the US series as I(1) is a sample-size artifact rather than a settled result, the control regression is estimated in levels, by OLS with Newey–West HAC standard errors, on the overlapping monthly sample:

$$\phi_t^{(n),\mathrm{BR}} = \alpha + \beta_1 \phi_t^{(2y),\mathrm{US}} + \beta_2 \phi_t^{(10y),\mathrm{US}} + \varepsilon_t^{(n)}$$

The fitted residual $\varepsilon_t^{(n)}$ is the Brazil-specific component of the term premium, plotted directly (no cumulation or re-anchoring is needed here, unlike a differenced specification, since the regression is already estimated in levels).

A.5 A Validation Checklist

Given the number of implementation choices above that materially affect the result without producing an obvious error at the point they are made, the following checks are used to validate any re-estimation of this model before trusting its output:

1. Mean excess returns by maturity (A.3.4, before regression) should be small and positive, increasing with maturity. Large, negative, and roughly maturity-independent values indicate the annualized/monthly risk-free-leg error described in A.3.4.

2. Φ^Q's eigenvalue moduli should be under 1 without needing a clamp. If a clamp is required, first check the internal-consistency requirement of A.3.2 before falling back to a clamp, which was found to distort episode-level sign and timing.

3. The fitted model yield should track the actual observed curve closely (mean absolute error well under 0.5pp at the maturities of interest) — a large, systematic gap between model and actual yield, as distinct from the intended model-versus-EH gap that defines the term premium, indicates a specification error upstream of the term premium calculation itself, not a genuine finding.

4. The 2015–16 and 2021–22 episodes should show clearly opposite-signed term premium behavior (a sustained positive peak in the former, a trough in the latter) of comparable order of magnitude. A model that reproduces one episode but collapses the other toward zero is displaying a genuine, diagnosable asymmetry — in this implementation's development, tracing that asymmetry to its source (via a direct linear decomposition of the term premium into its state-variable-specific components, holding all else fixed) was what surfaced the internal-consistency issue of A.3.2 in the first place.

⁂