Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Front-End Rates and Long-End Relative Value in the U.S. Treasury Curve

Fuli Yang

dr@garch.ac

GARCH Institute

Jing Huang

GARCH Institute

Research Article

Keywords: Treasury yield curve; conditional relative value; front-end rates; curve steepener; debt supply

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11586111/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Front-End Rates and Long-End Relative Value in the U.S. Treasury Curve

Abstract

Front-end rates provide a conditional reference for relative value between ten-year and thirty-year Treasuries. We evaluate the full link from conditional flatness to spread adjustment, portfolio expression, and supply sensitivity using public U.S. data. In a chronological evaluation beginning in 2015, a lagged reference gap lowers twenty-observation official-curve forecast mean squared error (MSE) by 3.48 percent. Removing front-end conditioning reverses that improvement. Frozen-reference accounting associates a one-basis-point initial zero-coupon gap with 0.35 basis points of spread adjustment and 0.31 basis points of gap closure. Under full cash-flow pricing, positive-gap states outperform nonpositive states for the same funded steepener by 1.33 basis points per unit of gross dollar value of a basis point (DV01), with positive intervals across four fixed dependence assumptions. The positive-state steepener has a mean net outcome of 0.72 and standard deviation of 3.68, compared with 19.89 for a two-year long. Its lower directional sensitivity provides relative protection during rising-rate episodes. Forecast strength varies across yield conventions and periods; longer dependence blocks and financing wedges reduce the precision of positive absolute means. Known auction quantities leave conditional adjustment essentially unchanged, while prospective supply changes can shift the valuation reference. The results establish front-end-conditioned curve selection and directional duration as distinct portfolio decisions.

Keywords: Treasury yield curve; conditional relative value; front-end rates; curve steepener; debt supply.
JEL: E43, G11, G12, G14.

1 Introduction

Treasury yields reflect macroeconomic news, the expected policy path, risk appetite and portfolio demand jointly. Strong inflation and labor-market data can therefore coincide with a bond rally. A directional position aggregates these forces. A relative-value position evaluates the long-end curve against prevailing front-end conditions and uses the resulting gap to choose an exposure. We ask whether conditional flatness between ten and thirty years predicts subsequent spread adjustment and supports a lower-beta expression of a rate view.

The relative-value proposition has three linked elements. Front-end rate conditions provide a benchmark for the ten-year/thirty-year slope; an unusually flat curve can offer scope for subsequent adjustment; and a balanced steepener can express that view with lower directional sensitivity than an outright duration position. Prospective changes in long-end issuance can alter the benchmark and the attractiveness of the expression.

We examine the complete chain using public Treasury data. First, we estimate the conditional relationship between front-end rates and the long-end slope. Second, we test whether a lagged deviation forecasts subsequent spread adjustment and survives simple forecast comparisons. Third, we compare a DV01-balanced steepener with outright front-end and ten-year duration expressions. Fourth, we examine known maturity-specific supply and distinguish announced quantities from expectation revisions. These components connect conditional pricing, forecasting, portfolio expression, and the conditions under which a relative-value position can fail.

We estimate the long-end spread against a one-year forward beginning one year ahead over the preceding 126 common observations. The primary common-date analysis uses the continuously compounded forward; a separate auxiliary analysis uses the published coupon-equivalent forward. On October 16, 2026 the official constant-maturity curve is 3.03 basis points flatter than the continuous-forward reference, whose trailing regression has an R^2 of 0.851. The fitted-par gap is 3.34 basis points. The same reference-window rule governs the dated illustration and the historical evaluation.

The longer evaluation supports conditional relative adjustment. The forward-conditioned official spread forecast improves twenty-observation MSE by 3.48 percent over no change. The corresponding improvement is 0.50 percent for fitted par yields and 0.99 percent for zero-coupon yields. Frozen-parameter accounting shows that daily gap closure is associated mainly with movement of the spread rather than movement of the front-end-implied reference. Performance is much stronger after 2021 than in 2015–2019.

The strategy comparison connects conditional adjustment to an explicit risk expression. On the 1,164 evaluation origins with a positive lagged official-curve gap, an equal-gross-DV01 steepener has a twenty-observation first-order payoff standard deviation of 3.51 basis points, versus 22.14 for an outright two-year long. Its slope on subsequent two-year yield changes is -0.098, versus the two-year long's algebraic sensitivity of -1. Full cash-flow pricing preserves the lower-risk expression after carry, roll, benchmark financing and assumed execution charges. A comparison of the same funded steepener across gap states gives a positive-minus-nonpositive mean difference of 1.327 basis points per gross DV01. Conditional selection and balanced exposure thus make distinct contributions to the strategy.

Our contribution is an integrated evaluation of conditional pricing, subsequent adjustment and exposure choice. The design combines lagged chronological forecasts, a front-conditioning ablation, frozen-reference accounting, same-steepener state comparisons and full cash-flow pricing at equal gross DV01. This sequence establishes where the signal adds information and how a curve position translates it into a lower directional-risk exposure. The return decomposition connects to Koijen et al. (2018), who distinguish carry from subsequent price appreciation across

asset classes. In our application, that distinction identifies the respective contributions of maturity aging, curve repricing, financing and execution charges.

The forecasting literature connects yield-curve shape to subsequent interest-rate movements and bond returns. Campbell and Shiller (1991) examine the information in yield spreads, while Cochrane and Piazzesi (2005) use forward rates to predict excess bond returns. Diebold and Li (2006) forecast the curve through level, slope and curvature factors. Our object is the ten-year/thirty-year slope relative to a recent front-end-conditioned mapping. We evaluate that deviation through subsequent adjustment, comparisons with direct yield forecasts, and the portfolio expression of its information.

The economic interpretation of a curve residual depends on its benchmark. Hu et al. (2013) construct a security-level curve-noise measure to study market illiquidity. Our gap measures the long-end slope relative to front-end conditions and can contain both relative adjustment and maturity-specific risk compensation. Term-structure decompositions in Adrian et al. (2013), Kim and Wright (2005) and Crump et al. (2016) provide the corresponding distinction between expected rates and term premia. In related Treasury research, Yang (2026) examines oil-price innovations and the composition of yield and term-premium movements. Here, consistent yield conventions and frozen-reference accounting locate the observed adjustment within the curve.

The reference can move as monetary news and maturity-specific demand change. Hanson and Stein (2015) link monetary news to long-term real rates; Greenwood and Vayanos (2014) study debt supply and excess bond returns; and Vayanos and Vila (2021) formalize the interaction of preferred-habitat investors with risk-bearing arbitrageurs. These mechanisms motivate our supply scenarios and stability comparisons. The empirical sequence runs from conditional pricing to forecast evaluation, funded holding outcomes, and the circumstances that change the reference.

2 A dated application of conditional pricing

An illustrative week highlights the distinction between yield direction and relative curve shape. Official H.15 yields declined across the two-, five-, ten-, and thirty-year maturities between October 9 and October 16, 2026. The larger declines at longer maturities coincided with a curve that remained flat relative to its recent front-end-conditioned reference.

The respective weekly declines were 6, 13, 15, and 17 basis points. They establish the dated market setting for a conditional-pricing calculation that combines the observed spread with prevailing front-end rates. The full-history tests below assess the information in that valuation gap.

Table 1. Conditional relative value on October 16, 2026

Spread convention	Reference minus observed (bp)	Trailing R^2
Official CMT	3.03	0.851
GSW fitted par	3.34	0.832
GSW zero coupon	6.42	0.786

Notes: October 16, 2026; each mapping is estimated on the previous 126 common observations and evaluated at that date's continuous one-year/one-year Treasury forward. The gap is positive when the observed curve is relatively flat.

The official and fitted-par estimates both indicate conditional flatness of approximately three basis points. The zero-coupon gap is larger, reflecting a different maturity-weighting convention. The fitted R^2 measures the reference relationship within the estimation window. Subsequent adjustment and holding outcomes supply the separate forecast and economic evaluations.

Reference strength changes over time. Across 2015 onward, the median trailing official-curve

R^2 is 0.602, with tenth and ninetieth percentiles of 0.038 and 0.872. The dated value of 0.851 lies toward the strong end of that distribution. Rolling estimation and calendar-period comparisons capture this variation in the reference relationship.

3 A framework for conditional relative value

3.1 Expectations and maturity-specific compensation

For a continuously compounded zero-coupon yield of maturity n, write the conventional decomposition as

$$z_t^{(n)} = \frac{1}{n}\int_0^n E_t[r_{t+u}]\,du + TP_t^{(n)}. \tag{1}$$

The thirty-year minus ten-year zero-coupon slope combines different averages of expected short rates and the difference between thirty-year and ten-year term premia. Near-term policy news, distant rate expectations and maturity-specific risk compensation can each move the slope. We evaluate official par yields alongside fitted par and zero-coupon outcomes to track the consequences of curve construction and maturity weighting.

Let p_t denote a front-end rate measure, and let q_t collect information about term premia, supply, and demand. Define a conditional reference spread

$$s_t^* = a_t + b_t p_t + c_t' q_t, \qquad s_t = y_t^{(30)} - y_t^{(10)}, \qquad g_t = s_t^* - s_t. \tag{2}$$

The baseline sets q_t aside and estimates a_t, b_t over the preceding 126 common observations. Thus s_t^* is an estimated conditional benchmark. Its role is a statistical valuation reference. A positive g_t indicates that the observed curve is flatter than its conditional reference.

3.2 Partial adjustment and testable implications

A reduced-form adjustment law is

$$s_t - s_{t-1} = \kappa_t(s_t^* - s_{t-1}) + u_t, \qquad 0 < \kappa_t \le 1. \tag{3}$$

For illustration, the same law follows from a quadratic adjustment device minimizing $\frac{1}{2}(s - s_t^*)^2 + \frac{\chi_t}{2}(s - s_{t-1})^2$, giving $\kappa_t = (1+\chi_t)^{-1}$. This partial-adjustment device supplies the empirical sign restriction.

If the reference spread is locally unchanged and κ is constant, expected cumulative adjustment satisfies

$$E_t[s_{t+h} - s_t] = [1 - (1-\kappa)^h]g_t. \tag{4}$$

The framework yields three propositions to test. First, an alignment gap predicts subsequent spread changes with the same sign. Second, the adjustment profile varies with the persistence of front-end news and the market's capacity to absorb duration risk. Third, a persistent shift in supply or risk compensation can move the reference itself, leaving an unaugmented gap misleading. The third proposition is essential for distinguishing delayed alignment from a changing equilibrium spread.

3.3 From conditional adjustment to a lower-beta expression

The connection between a curve gap and portfolio choice can be stated directly. For a common front-end repricing factor $\Delta_h p_t$, write the reduced-form yield responses as

$$\Delta_h y_t^{(n)} = \beta_{n,h}\Delta_h p_t + e_{n,t+h}, \qquad n \in \{2, 10, 30\}. \tag{5}$$

At equal initial gross DV01, the first-order steepener and two-year long have factor sensitivities

$$B_{S,h} = \tfrac{1}{2}(\beta_{30,h} - \beta_{10,h}), \qquad B_{L2,h} = -\beta_{2,h}. \tag{6}$$

If ten- and thirty-year yields have similar responses to front-end repricing, their difference carries less directional exposure than an outright front-end long. A modest negative $B_{S,h}$ retains the sign of duration exposure when front-end rates decline. Total risk also depends on the variance of the residual difference. We therefore measure both factor sensitivity and realized exposure dispersion, rather than infer low risk solely from net duration balance. The empirical beta comparisons use two-year yield changes as an observable front-end factor, while the conditional reference uses the Treasury forward rate.

The forecasting equation supplies a linear-projection forecast of the first-order gross curve payoff per unit DV01:

$$\widehat{m}^{S}_{t,h} = \tfrac{1}{2}\left(\alpha_h + \lambda_h g_{t-1} + \Gamma_h X_{t-1}\right). \tag{7}$$

For a positive adjustment coefficient, conditional flatness increases the forecast attractiveness of a steepener without requiring a decline in the overall level of yields. Full cash-flow holding outcomes add maturity aging, financing, execution and nonlinear repricing. In the same risk units, an economically attractive position requires the forecast curve payoff plus expected net carry and nonlinear effects to exceed expected implementation costs. A positive gap is therefore a transparent valuation signal, while the cost scenarios assess how much of that value remains under different assumptions. We retain the zero sign boundary as the empirical rule and do not optimize a trading hurdle from observed performance.

These relations explain why a lower-beta curve position can be preferable to an outright long without dominating it in every state. The relative expression gives up part of a directional rally's payoff. Its protection during a selloff concerns comparative losses, with curve direction determined by the realized relative adjustment.

4 Data and measurement

Daily two-, five-, ten-, and thirty-year constant-maturity yields come from the Federal Reserve H.15 series distributed through FRED. These official CMT rates are estimated par-curve rates. We use CMT for the official constant-maturity estimates and GSW for the fitted curves; the accompanying code retains the short label cash for CMT. Both the official CMT and GSW par series are curve estimates with different inputs and fitting procedures. The usable common sample begins on February 9, 2006, when the thirty-year series resumed, and ends on October 29, 2026. It contains 5,185 common observations. Downloaded records preceding the documented resumption are retained in a separate audit file and excluded from estimation.

The primary common-date front-end measure is the continuously compounded one-year forward beginning one year ahead, $p_t = 2z_t^{(2)} - z_t^{(1)}$, constructed from the Gürkaynak–Sack–Wright (GSW) zero-coupon curve. Section 6 also reports an auxiliary replication using the published coupon-equivalent forward SVEN1F01 and a daily H.15 two-year reference. We label these measures separately throughout. The GSW estimates use off-the-run coupon securities; the latest fitted date in the retrieved vintage is October 23, 2026. The common CMT/forward panel contains 5,181 observations. Both Treasury forward measures combine policy-path expectations with risk compensation.

GSW estimates are generally released weekly, and historical estimates can be revised. We use the retrieved historical vintage to evaluate dated curve states. A six-observation delay of the complete gap provides a publication-timing sensitivity. A reconstructed real-time exercise would additionally require the original release and vintage archive.

The primary outcome is the twenty-observation change in the thirty-year minus ten-year spread. Five observations provide an auxiliary horizon. Forecast calibration uses origins from 2007 with target dates ending by December 31, 2014; evaluation begins in 2015. We retain the project's calendar comparisons of 2015–2019, 2020–2021, and 2022–2026. The design is an exploratory chronological evaluation, with the same specification and comparison grid carried through the long-sample extension.

Table 2. Sample clocks and primary information sets

Panel	Observations	First date	Last date
Official CMT common rates	5,185	2006-02-09	2026-10-29
Common CMT/GSW states	5,181	2006-02-09	2026-10-23
Twenty-observation evaluation origins	2,934	2015-01-02	2026-09-24
Positive official-gap holding origins	1,164	2015-01-02	2026-09-24
Policy-announcement events	169	2006-09-20	2026-09-16

Notes: State dates, forecast origins, holding origins and announcement dates are distinct sample objects. Evaluation endpoints require complete future target windows; the last origin therefore precedes the last observed curve date. The positive-state rule uses a lagged gap. Release timing is assessed separately from the historical state dates.

5 Empirical design

Following the horizon-by-horizon regression approach of Jorda (2005), the baseline local projection is

$$s_{t+h} - s_t = \alpha_h + \lambda_h \widehat{g}_{t-1} + \Gamma_h X_{t-1} + \varepsilon_{t+h}. \tag{8}$$

The controlled specification includes the previous forward level and its previous five-observation change. A positive λ_h measures conditional alignment. Separate equations for two-, ten-, and thirty-year yield changes locate that adjustment within the curve. With the same information set, the spread coefficient equals the thirty-year coefficient minus the ten-year coefficient; the code verifies this identity. The daily regression estimates conditional predictability, while the supplementary event analysis addresses identified announcement shocks.

Forecast comparisons use the unchanged spread, a directly estimated AR1 spread forecast, a rolling 252-resolved-observation AR1, a fixed four-yield direct forecast, and a rolling standardized four-yield ridge regression. The ridge penalty is fixed at 10. The main alignment specification uses the original 126-observation conditional window. An expanding forecast-calibration sensitivity changes only the calibration sample and uses resolved targets available before each forecast origin. No window or threshold is selected from the long-sample results.

HAC covariance estimates following Newey and West (1987) address overlapping targets. Primary revised comparisons use 5,000 circular block draws over the complete calendar, drawing on the dependent-data bootstrap framework of Kunsch (1989). The baseline block contains forty observations; sensitivity blocks contain twenty, eighty or 126 observations. Original auxiliary analyses retain their documented 2,000-draw procedures. We evaluate paired forecast-loss differences in the spirit of Diebold and Mariano (1995), using the stated block intervals for inference. The loss intervals condition on the generated forecasts and quantify the fixed comparison's uncertainty. Uncertainty from re-estimating the complete reference-generation pipeline is a separate object.

5.1 Research propositions and empirical tests

The pricing proposition concerns g_t, the conditional flatness of the long-end curve. The adjustment proposition concerns the sign and size of subsequent spread changes conditional on a lagged g_t. The expression proposition concerns the sensitivity and downside behavior of a balanced steepener compared with an outright long. The supply proposition concerns whether known maturity-specific quantities absorb the relation and whether revised future supply can move the reference. The daily gap conditions on the full curve state; the event study conditions on a specific monetary-policy innovation. Each design tests its corresponding proposition.

The relative-value extension retains both positive and negative gaps in forecast estimation. Its long-steepener exposure comparison uses only a strictly positive lagged gap, with zero as the economically determined sign boundary. Five observations are auxiliary and twenty are primary. All measured conventions, calendar periods, and scenarios are retained. The extension is exploratory; no new rule is selected to maximize performance or target a particular dated valuation.

6 Daily forecasting and relative adjustment

The auxiliary daily-rate reference uses the official two-year yield. In its cash-rate panel, the fixed alignment model reduces twenty-observation mean squared forecast error by 4.36 percent over 2015–2026. Its RMSE is 7.340 basis points, compared with 7.505 for the unchanged spread and 7.961 for the rolling AR1. The loss-improvement interval relative to the rolling AR1 is positive in this long panel.

The auxiliary published-forward estimates use the coupon-equivalent SVEN1F01 on common CMT/GSW dates. Its RMSE is 7.379 basis points versus 7.510 for the unchanged spread, a 3.48 percent MSE improvement. The paired loss gain is 1.961 squared basis points, with the original 2,000-draw block interval [0.167,3.760]. These are the published-forward auxiliary estimates. The primary continuous-forward comparison in Table 5 uses the revised 5,000-draw procedure: its loss gain is 1.964 and its interval is [0.139,3.834]. Delaying the published-forward gap by six observations lowers its MSE improvement to 2.35 percent, with a loss interval spanning zero.

The auxiliary published-forward local projection associates a one-basis-point gap with a 0.352-basis-point spread change, with HAC standard error 0.090. Including lagged forward conditions gives a coefficient of 0.404 and standard error 0.091. In the same controlled equations, the ten-year coefficient is 0.011 and the thirty-year coefficient is 0.415. The point estimates place more adjustment in the thirty-year sector; the uncertainty around each individual leg is wider than that around their difference. The common-date continuous-forward accounting below evaluates the relative adjustment under consistent yield conventions.

Period differences are economically material. The forward model raises MSE by 1.94 percent in 2015–2019, lowers it by 1.85 percent in 2020–2021, and lowers it by 5.74 percent in 2022–2026. The corresponding cash-rate changes are -0.61, 1.65, and 6.71 percent. The conditional relationship is strongest in the later block; forecast usefulness is state dependent.

6.1 Common-date forecast benchmarks

For a direct comparison under the unified continuous-forward measure, we recalibrate all fixed models on exactly the same 2007–2014 resolved targets and evaluate the same 2,934 origins. The four-yield model uses lagged official two-, five-, ten-, and thirty-year levels. Adding the gap to this model tests incremental predictive information under a common calibration rather than comparing different dates.

The parsimonious gap model performs best among these fixed specifications for each spread

Table 3. Common-date fixed-model forecast RMSE at twenty observations

Model	Official CMT	GSW par	GSW zero
No change	7.510	8.355	10.646
Conditional gap	7.378	8.334	10.594
Lagged-spread AR1	8.319	8.496	10.657
Four lagged yields	8.046	8.739	11.025
Four yields plus gap	7.773	8.801	10.980

Notes: Basis points; identical origins and training targets. Each model predicts the cumulative change directly. The full specification and five-observation outputs are retained in the replication package.

convention. Adding the gap improves the four-yield forecast for official CMT and zero-coupon outcomes, but worsens it for fitted par yields; the combined models still underperform no change. The results favor the parsimonious conditional reference within this fixed comparison.

6.2 The incremental role of front-end conditioning

The reference gap contains the observed spread with a negative sign, so a successful gap forecast could partly reflect ordinary mean reversion. We therefore remove front-end conditioning while preserving the 126-observation reference window, information lag, calibration dates and targets. The resulting comparator is the preceding-window mean spread minus the current spread. A direct spread/front-end forecast provides a second benchmark, with fixed or 252-resolved-target rolling calibration. We also report the conditional-gap forecast under the same rolling calibration rule. All models use identical evaluation dates within a horizon; no benchmark is selected according to its performance.

Table 4. Conditioning ablation and calibration comparisons

Model	Official CMT	GSW par	GSW zero
Conditional gap, fixed calibration	3.48	0.50	0.99
No-front mean gap, fixed	-1.65	-0.61	-0.21
Direct spread/front, fixed	-68.92	-6.08	-6.52
Direct spread/front, rolling 252	-17.80	-17.81	-17.96
Conditional gap, rolling 252	-0.21	-5.82	-7.74
Conditional gap, lag6	2.35	-0.43	0.05
H.15 two-year reference, fixed	4.36	–	–

Notes: Percent MSE improvement over no change; $N = 2,934$. Rolling models use only targets resolved strictly before each origin. Lag 6 delays the entire generated gap. The two-year reference uses the official daily rate and is reported for the official-CMT outcome. Full five-observation results are retained.

For official CMT, removing front-end conditioning changes the twenty-observation MSE gain from 3.48 percent to -1.65 percent. The paired mean loss improvement of the front-conditioned model is 2.894 squared basis points, with a forty-observation block interval of [1.075,5.024]. Its interval remains positive under all four fixed block lengths. Fitted-par and zero-coupon point comparisons also favor front-end conditioning, but their paired intervals span zero.

Rolling recalibration reduces forecast performance in this comparison. The rolling conditional-gap model slightly underperforms no change in official CMT and more clearly underperforms in the fitted measures. Its comparison with the rolling direct forecast is imprecise. The positive forecast result belongs to the parsimonious conditional reference with fixed calibration; the original 126-observation baseline remains the primary specification.

7 Yield conventions and frozen-reference adjustment

7.1 Cash, fitted par, and zero-coupon spreads

To distinguish market-yield measurement from conditional adjustment, we construct three thirty-year minus ten-year outcomes on identical dates: Treasury constant-maturity par yields distributed through H.15, GSW fitted par yields, and GSW continuously compounded zero-coupon yields. All three conditional-reference regressions use the same continuously compounded one-year forward beginning one year ahead, $p_t = 2z_t^{(2)} - z_t^{(1)}$. This exercise differs from the preceding published-forward baseline only in its explicit convention-unification design. The common panel has 5,181 observations through October 23, 2026.

The reference is fitted on 126 preceding observations, and each gap enters forecasts with a one-observation lag. Calibration uses resolved 2007–2014 targets; evaluation begins in 2015. Table 5 locates the strongest fixed-calibration gain in H.15 CMT yields. The zero-coupon MSE gain is 0.99 percent, with a paired block interval spanning zero. Moving from CMT to GSW par yields changes the security universe and fitting method; moving from par to zero yields also changes maturity weighting and quotation convention. These measurement differences affect forecast strength.

Table 5. Identical-date twenty-observation forecast comparisons

Spread measure	No-change RMSE	Gap RMSE	MSE gain (%)	Gain interval
H.15 CMT	7.510	7.378	3.48	[0.14,3.83]
GSW fitted par	8.355	8.334	0.50	[-1.19,1.94]
GSW zero coupon	10.646	10.594	0.99	[-1.64,4.02]

Notes: $N = 2,934$ origins per outcome. RMSE in basis points; gain intervals in squared basis points. Paired full-calendar circular blocks contain forty observations, with 5,000 draws. Current-vintage historical inputs; timing sensitivities are reported separately.

7.2 A frozen-reference accounting test

A rolling reference can change because short rates move or because estimated parameters change. To isolate the first channel, freeze a_t, b_t at the origin and define

$$\widetilde{s}_{t+h|t} = a_t + b_t p_{t+h}, \qquad \widetilde{g}_{t+h|t} = \widetilde{s}_{t+h|t} - s_{t+h}. \tag{9}$$

The exact accounting restriction is

$$\widetilde{g}_{t+h|t} - g_t = b_t(p_{t+h} - p_t) - (s_{t+h} - s_t). \tag{10}$$

The first term on the right measures movement of the front-end-implied reference under the frozen mapping; the second measures observed spread adjustment. The identity separates these two sources of gap change while holding reference parameters fixed. Shifts in maturity-specific risk compensation can change the economic interpretation of the mapping.

Using the zero-coupon spread and continuous forward, we regress each component on the origin gap, controlling for the preceding forward level and its preceding five-observation change. This is a retrospective association test on dated market states, distinct from the lagged fixed-calibration forecast experiment. The common regression sample has 2,934 observations from 2015 onward. The origin-gap coefficients are 0.355 for spread adjustment, 0.044 for reference movement, and -0.311 for the remaining-gap change, with HAC standard errors 0.099, 0.053, and 0.080. The coefficients satisfy the accounting restriction exactly. Conditional gap closure thus does not rely on re-estimating the reference parameters after the origin.

The retained calendar blocks show heterogeneity. The gap-change coefficient is -0.079 (0.111) in 2015–2019, -0.474 (0.163) in 2020–2021, and -0.640 (0.103) in 2022–2026. These comparisons locate stronger conditional adjustment in the later periods. Treasury changed its CMT interpolation method on December 6, 2021; Appendix F repeats the fixed-forecast comparison on windows respecting that documented date. The comparison tracks variation across both economic conditions and measurement regimes.

8 Relative-value expression and directional risk

8.1 Exposure units and comparison rules

A ten-year/thirty-year steepener receives ten-year duration and pays thirty-year duration. Its first-order price exposure is positive when the thirty-year minus ten-year yield spread widens. Let d be total gross absolute initial DV01. Allocating $d/2$ to each leg gives the yield-risk payoff

$$\frac{\Pi^{S}_{t,h}}{d} = \frac{1}{2}\left(\Delta_h y_t^{(30)} - \Delta_h y_t^{(10)}\right), \qquad \frac{\Pi^{L2}_{t,h}}{d} = -\Delta_h y_t^{(2)}, \qquad \frac{\Pi^{L10}_{t,h}}{d} = -\Delta_h y_t^{(10)}. \tag{11}$$

Yield changes are measured in basis points. The two-year long represents front-end duration exposure; the ten-year long supplies an additional directional benchmark. Each comparison uses the same positive-gap origins and corresponding yield convention. Common gross DV01 holds initial exposure scale fixed; volatility, financing and capital requirements depend on the resulting positions.

The initial comparison holds first-order exposures fixed. Section 9 extends it to full cash-flow pricing, carry, roll, convexity, financing and transaction costs. The CMT and GSW inputs describe curve-priced exposures. We evaluate overlapping observations as horizon outcomes per gross DV01.

8.2 Average behavior and front-end sensitivity

Table 6 evaluates all 1,164 positive-gap official-CMT origins from 2015 onward with complete twenty-observation outcomes. The gap is lagged one observation before an origin enters the exposure comparison. The steepener mean is 0.65 basis points per unit gross DV01, with a full-calendar block interval of [0.02,1.33]. Its standard deviation is 3.51, versus 22.14 for the two-year long and 21.49 for the ten-year long. The steepener's tenth percentile is -3.50, compared with -20.00 and -27.00 for the outright expressions.

Table 6. Positive-gap exposure outcomes at equal gross DV01

Expression	Mean	Standard deviation	Tenth percentile	Front-end beta
Steepener	0.65	3.51	-3.50	-0.098
Two-year long	0.12	22.14	-20.00	-1.000
Ten-year long	-1.96	21.49	-27.00	-0.804

Notes: $N = 1,164$; twenty-observation first-order yield-risk outcomes, in basis points per common gross DV01. Beta is the descriptive OLS slope on the realized official two-year yield change, with an intercept. The two-year beta equals -1 algebraically. Outcomes overlap; bootstrap means resample forty-observation blocks of the entire evaluation calendar together with their selection indicators.

The steepener's modest negative beta retains the favorable sign of front-end duration exposure when yields decline. Doubling both steepener legs to one full DV01 each gives a beta of -0.196 at twice the gross DV01, still far below the two-year long's unit sensitivity. Thus the small beta reflects the similarity of long-end yield responses as well as the chosen exposure scale.

The sign and scale persist across yield measures. Fitted-par and zero-coupon positive-gap steepeners have mean exposure payoffs of 0.51 and 0.76 basis points, standard deviations of 3.94 and 4.83, and front-end betas of -0.124 and -0.138. Each convention selects its own positive-gap dates. Across the retained calendar blocks, official steepener means are 0.03, 0.44, and 2.09 basis points. The positive overall outcome is therefore concentrated in the later period.

8.3 When bearish rate pressure resumes

We classify future yield paths only to diagnose scenarios; their direction is not available when the exposure is initiated. On the 633 positive-gap origins followed by an increase in the official two-year yield, the steepener loses 0.74 basis points on average, compared with 12.72 for the two-year long and 14.05 for the ten-year long. On the remaining 531 origins it gains 2.31, compared with 15.43 and 12.45. The expression sacrifices much of outright duration's rally participation while reducing its adverse directional exposure.

Table 7. Ex-post rate scenarios for the same positive-gap signal

Subsequent scenario	Origins	Steepener	Two-year long	Ten-year long
Two-year yield rises	633	-0.74	-12.72	-14.05
Two-year yield does not rise	531	2.31	15.43	12.45
Ten-year yield rises	622	-0.18	-11.23	-16.95
Ten-year yield does not rise	542	1.61	13.15	15.25

Notes: Conditional mean first-order outcomes in the same units as Table 6. Two different partitions of the same 1,164 origins are shown; the rows are not four mutually exclusive states. Directional-long losses in their own rising-yield partition follow algebraically from the exposure definition. The empirical content is the accompanying curve movement and magnitude of the relative expression.

These results give a precise interpretation to relative protection. The steepener can deliver a smaller drawdown than an outright long without generating a positive payoff in the selloff. Conditional flatness, overall positive subsequent adjustment, and a lower-beta expression can coexist with further flattening in a particular rate regime. The scenario evidence locates relative protection in the size of losses.

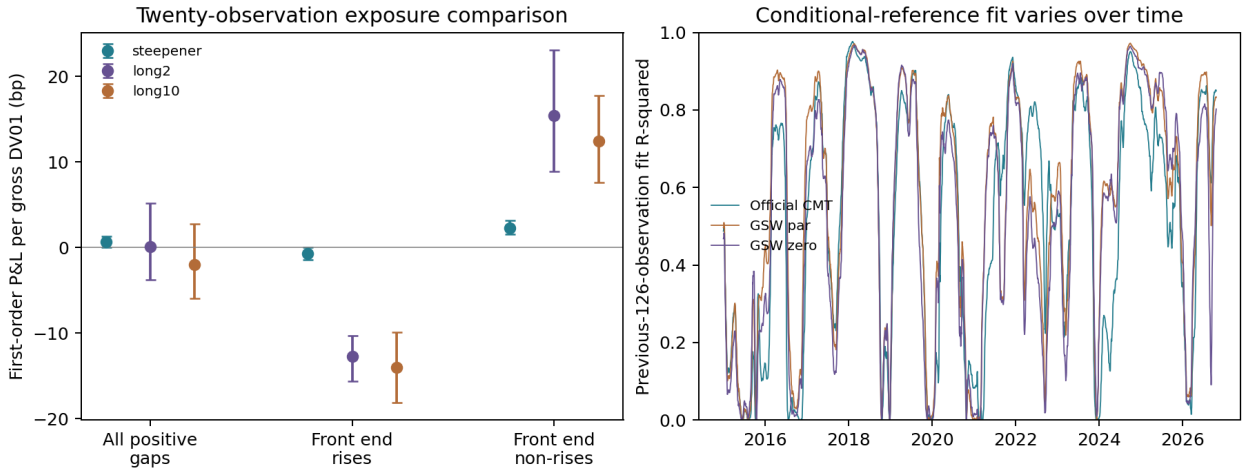


Figure 1. Relative exposure and changing reference strength. Left: mean twenty-observation official-CMT exposure outcomes and 95 percent full-calendar block intervals for positive lagged gaps. Right: preceding-window conditional-reference fit under the three yield measures. Future front-end direction is an ex-post scenario classification.

9 Coupon-bond holding outcomes, financing and costs

9.1 Cash-flow pricing and common risk normalization

We extend the first-order comparison to fixed-coupon synthetic bonds valued on the full GSW discount curve. This approach follows the curve-based synthetic-security pricing framework in Gurkaynak et al. (2007). At each origin we form two-, ten-, and thirty-year bonds with semiannual payments and an origin coupon selected to make their initial dirty value 100. Coupons and principal are then held fixed. Remaining maturities decline by the actual number of calendar days divided by 365.25. All five- and twenty-observation holding windows end before the first coupon payment; coupon accrual is captured in the terminal dirty price, with no intervening coupon reinvestment.

Let $C_{tj}^{(n)}$ be the fixed cash flows of an origin-t bond and τ_j its original payment times. With $D_t(u) = \exp[-z_t(u)u]$, terminal value after calendar aging δ is

$$P_{t+h}^{(n)} = \sum_j C_{tj}^{(n)} D_{t+h}(\tau_j - \delta), \qquad P_{t+h|t}^{(n)} = \sum_j C_{tj}^{(n)} D_t(\tau_j - \delta). \tag{12}$$

The frozen-curve aged value $P_{t+h|t}$ yields an exact decomposition of the price change into carry plus roll and subsequent curve repricing:

$$P_{t+h} - P_t = (P_{t+h|t} - P_t) + (P_{t+h} - P_{t+h|t}). \tag{13}$$

Full cash-flow repricing incorporates nonlinear price effects, including convexity. Coupons are selected using only origin curve parameters, not terminal prices.

Initial gross absolute parallel-zero-curve DV01 is normalized to one dollar per basis point. The steepener receives one-half of that exposure in the ten-year bond and pays one-half in the thirty-year bond; each outright long receives a full unit. Its per-100-face DV01 is $10^{-4} \sum_j \tau_j C_{tj} D_t(\tau_j)$. The cash-flow risk normalization is closely related to, but not identical to, the preceding yield-change normalization. Analytical DV01 agrees with finite-difference values within a maximum relative discrepancy of 1.37×10^{-6}. Across evaluation dates, parameter-implied par coupons agree with published par quotes within 0.006 basis points. A few larger pre-2015 quote discrepancies are archived and do not enter the holding-outcome evaluation.

9.2 Realized financing and assumed execution costs

The financing benchmark is the daily effective federal funds rate, DFF, from H.15 through FRED (Board of Governors, 2026). We compound each realized daily rate over actual calendar days using ACT/360. Positive bond market value is financed at DFF plus an annual wedge w; cash from the short is credited at DFF minus w. For initial long and short market values L_t and S_t, the financing charge is

$$F_{t,h}(w) = L_t \left[\prod_{d=t}^{t+h^-} \left(1 + \frac{r_d + w}{360} \right) - 1 \right] - S_t \left[\prod_{d=t}^{t+h^-} \left(1 + \frac{r_d - w}{360} \right) - 1 \right]. \tag{14}$$

Here d indexes calendar days rather than the observation count used in the holding horizon. We retain wedges of 0, 25, 50 and 100 annual basis points. These are transparent financing scenarios; DFF is not a measured security-specific repo rate.

For a one-way execution charge c in price basis points, both entry and exit incur $c \times 10^{-4}$ times gross dirty market value across the legs. We retain $c = 0, 0.5, 1, 2$ without selecting a cost level from the resulting performance. Net dollar P&L is the weighted full price change less financing and execution charges. Because initial gross DV01 is one, numerical P&L also has units

of basis points per common dollar DV01. It is not a percentage return on equity or an estimate of achievable execution profits.

9.3 Funded outcomes and the source of gains

The primary comparison uses exactly the same 1,164 positive lagged official-CMT gap origins as the preceding exposure analysis. At the DFF benchmark with a zero execution charge, the steepener earns a mean 0.804 basis points per gross DV01. Carry plus roll contributes 0.644, curve repricing contributes 0.574, and net financing costs 0.415. Thus the positive mean reflects both maturity aging and curve movement, rather than being a relabeling of the slope change.

Table 8. Twenty-observation synthetic holding outcomes with costs

Financing wedge	Expression	Net mean	Mean interval	Std. dev.	Break-even c
0 bp	Steepener	0.72	[0.05,1.45]	3.68	4.92
	Two-year long	0.93	[-2.57,5.09]	19.89	1.42
	Ten-year long	-0.84	[-4.98,3.48]	21.54	-3.24
25 bp	Steepener	0.56	[-0.11,1.29]	3.68	3.91
	Two-year long	-0.10	[-3.60,4.06]	19.89	0.41
	Ten-year long	-1.07	[-5.21,3.26]	21.54	-4.25

Notes: $N = 1,164$ per row. Net outcomes include a 0.5 price-basis-point one-way charge at entry and exit. Mean intervals use 5,000 forty-observation circular full-calendar block draws with signal indicators. Break-even c is the sample-mean-zero one-way price charge under the stated financing wedge; negative values indicate losses before execution costs.

At a 0.5 price-basis-point one-way charge, the steepener mean is 0.722 with a conditional block interval of [0.053,1.454]. The two-year long mean is 0.933 and the ten-year long mean is -0.840. The steepener combines substantially lower volatility with a positive mean under benchmark financing. Paired mean differences against both outright expressions have block intervals spanning zero. With the 25-basis-point financing wedge, its mean is 0.557 and its interval spans zero. These estimates establish relative risk reduction more precisely than mean-return superiority. Section 10 examines reference memories and holding horizons.

The transaction-cost boundary is economically informative. The official-signal steepener's sample-mean break-even one-way charge is 4.92 price basis points at benchmark funding, 3.91 with a 25-basis-point financing wedge, 2.89 with a 50-basis-point wedge, and 0.86 with a 100-basis-point wedge. These boundaries reflect model-priced outcomes under explicit assumptions. At the 25-basis-point wedge, the 2015–2019 steepener break-even charge is only 0.38, compared with 4.34 in 2020–2021 and 10.11 in 2022–2026. Cost resilience therefore varies across the original calendar blocks; a positive full-sample mean does not establish equal performance in every period.

9.4 Rising-rate scenarios and implementation interpretation

With benchmark funding and the 0.5 price-basis-point charge, the 633 positive-gap origins followed by a rising two-year yield have mean funded outcomes of -0.75 for the steepener, -10.36 for the two-year long, and -12.95 for the ten-year long. On the remaining 531 origins, the means are 2.48, 14.39 and 13.60. Adding the 25-basis-point financing wedge gives -0.92, -11.39 and -13.18 in the rising-front-end partition. Incorporating aging, nonlinear pricing and the stated costs therefore preserves the comparative protection found in the first-order exercise.

The synthetic design isolates the contributions of fixed cash flows, maturity aging, financing and curve repricing. It prices historical outcomes from the current-vintage fitted curve, with execution and financing wedges specified as scenarios. Observed CUSIP execution, issue-specific repo specialness, haircuts and margin requirements lie outside this pricing design. We retain the

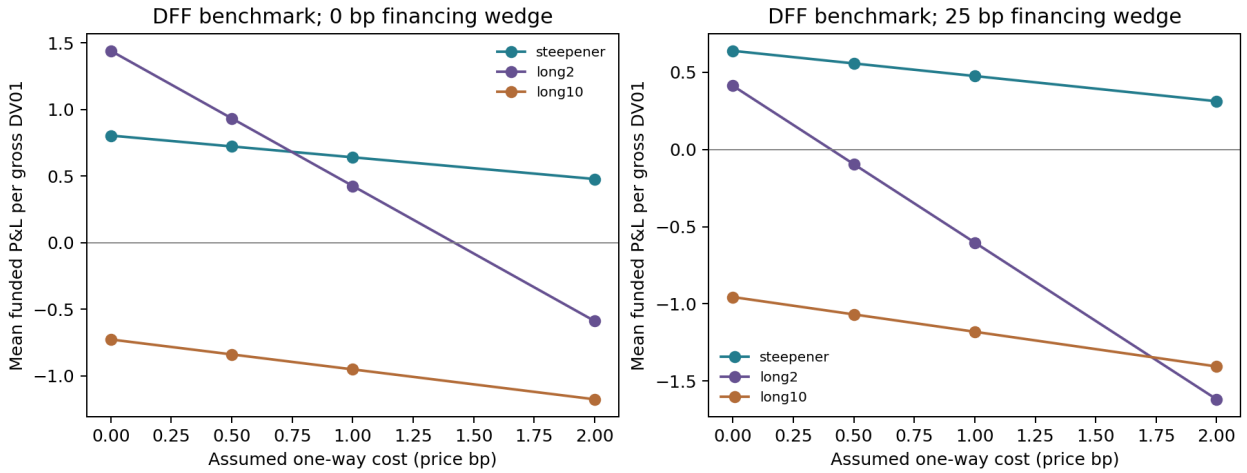


Figure 2. Funded synthetic outcomes across assumed execution charges. Each line uses the same positive official-CMT gap origins. The financing wedge applies above DFF for long positions and below DFF for short-cash credit. Charges are one-way price basis points, while the vertical axis measures P&L per initial gross DV01.

complete five-observation, alternative-gap, financing and cost grid. Overlapping holdings enter as conditional horizon outcomes, each measured per initial gross DV01.

10 Stability across reference windows, horizons and periods

10.1 A fixed comparison grid

We compare reference windows of 63, 126 and 252 observations, corresponding approximately to a quarter, half-year and year, at five- and twenty-observation horizons. The 126-observation window remains the baseline. The complete grid reports every window, outcome convention, horizon and retained calendar block.

For a fair forecast comparison, all windows use identical complete-feature calibration dates within a horizon: origins from 2007 onward with targets resolved by the end of 2014 and sufficient prior observations for the longest window. This leaves 1,969 calibration targets at five observations and 1,954 at twenty. Evaluation still uses the original 2,949 and 2,934 dates, respectively. Restricting the training sample to common dates makes the 126-window calibration slightly different from the original baseline; it does not change that baseline's reported results. Each window's positive lagged official-CMT gap selects its holding-outcome origins from the unchanged synthetic payoff panel, with DFF benchmark financing and the same 0.5 price-basis-point one-way cost.

10.2 Forecast stability

Table 9 reports the complete full-period forecast grid. All eighteen point estimates improve on no change, with most fitted-par and zero-coupon loss intervals spanning zero. Official-CMT twenty-observation MSE gains range from 3.50 to 3.81 percent. The aggregate gain therefore persists across the three reference memories, with precision varying across yield measures.

The chronological pattern also persists across windows. At twenty observations, official-CMT MSE gains in 2015–2019 are -2.82, -1.87 and -4.79 percent for the three windows. They are 1.51, 1.90 and 4.20 percent in 2020–2021, and 6.44, 5.74 and 6.74 percent in 2022–2026. Changing reference memory therefore does not eliminate the earlier period of weak performance. The later strengthening is a material feature of the evidence rather than an artifact of choosing only the

Table 9. Forecast MSE improvement under common calibration dates

	Five observations			Twenty observations		
Spread	63	126	252	63	126	252
Official CMT	0.89	1.37	1.25	3.68	3.50	3.81
GSW par	0.37	0.50	0.56	1.98	0.61	1.73
GSW zero	0.32	0.63	0.66	1.58	1.14	1.96

Notes: Percent MSE improvement over no change. Column headings are reference-window observation counts. All windows share calibration and evaluation dates within a horizon. Full loss intervals and the original calendar-period comparisons are retained; the intervals describe individual fixed comparisons.

126-observation window.

10.3 Funded expression stability

Table 10 uses each window's own positive-gap origins. Twenty-observation funded steepener means range from 0.57 to 0.72 basis points per common gross DV01, with standard deviations of 3.59–3.68. Two-year long standard deviations range from 18.77 to 19.89 and ten-year long standard deviations from 19.63 to 21.79. The steepener's descriptive beta on subsequent two-year yield changes is between -0.095 and -0.105, compared with approximately -0.87 to -0.88 for the funded two-year long. Lower directional sensitivity survives the alternative reference windows under full cash-flow pricing.

Table 10. Funded steepener outcomes across windows and horizons

Horizon	Window	Origins	Mean	Std. dev.	Front-end beta	Rising-front mean
5	63	1,227	0.05	2.10	-0.121	-0.81
5	126	1,169	0.15	2.01	-0.119	-0.70
5	252	1,034	0.13	2.07	-0.112	-0.66
20	63	1,223	0.57	3.59	-0.095	-0.86
20	126	1,164	0.72	3.68	-0.105	-0.75
20	252	1,027	0.65	3.65	-0.096	-0.66

Notes: Funded synthetic P&L per initial gross DV01, including benchmark DFF financing and the 0.5 price-basis-point one-way charge. Reference windows select different origin sets; the table compares rule implementations rather than matched-origin treatment effects. Rising-front mean conditions ex post on a positive official two-year yield change. No future direction enters signal selection.

The twenty-observation means exceed their five-observation counterparts under all three windows. Shorter-horizon means have block intervals spanning zero, as do twenty-observation means under the 63- and 252-observation reference windows. The comparisons favor a longer holding horizon for this expression. Overlapping horizons and different eligible origins make this a comparison of implementations, with the sign and risk reduction more stable than the precision of positive absolute means.

Across calendar blocks, funded steepener means are 0.10, 0.50 and 1.35 for the 63-window rule; 0.14, 0.73 and 1.91 for the baseline rule; and -0.04, 0.90 and 2.58 for the 252-window rule. Thus the early funded result is close to zero even where full-period means are positive. All three rules exhibit stronger later-period outcomes. The findings support conditional relative-value assessment and choice of exposure, while making period dependence part of the empirical conclusion.

Holding the original predictions fixed, excluding one origin year at a time leaves official-CMT MSE gains between 1.35 and 4.38 percent; excluding 2022 leaves 1.35 percent. This fixed-prediction

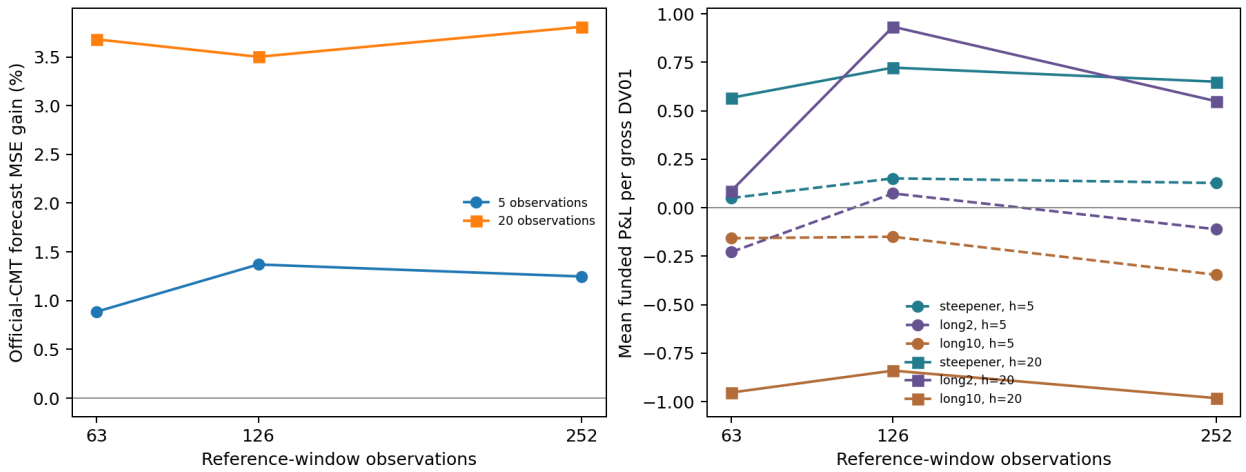


Figure 3. Fixed-grid stability without window selection. Left: official-CMT forecast MSE improvement. Right: funded synthetic outcomes using each reference window's positive-gap origins. Dashed lines denote five observations and solid lines twenty; all costs and financing assumptions are held fixed.

accounting attributes a material share of the aggregate improvement to 2022 and shows that a positive aggregate gain remains across every omitted-year comparison.

11 Conditional information, timing and inference

11.1 Signal selection versus conventional curve hedging

Balancing ten- and thirty-year DV01 removes much common duration exposure regardless of the signal. To isolate conditional information, we compare the same funded steepener at all origins, at positive lagged gaps and at nonpositive gaps. All three use identical cash-flow pricing, DFF financing and the 0.5 price-basis-point one-way cost. Lower beta is an exposure property; differences across gap states measure the incremental selection information.

Table 11. The same steepener under different dated information states

State rule	Origins	Net mean	Std. dev.	Mean interval
All origins	2,934	-0.08	4.10	[-0.63,0.47]
Positive lagged gap	1,164	0.72	3.68	[0.05,1.45]
Nonpositive lagged gap	1,770	-0.61	4.26	[-1.35,0.06]
Positive gap delayed six	1,164	0.66	3.74	[-0.04,1.44]

Notes: Twenty-observation outcomes per common initial gross DV01. Mean intervals use 5,000 forty-observation full-calendar block draws. The delayed rule has its own eligible positive states; equal counts do not imply identical origin sets. State contrasts measure conditional historical horizon outcomes.

The positive-minus-nonpositive mean difference is 1.327 basis points per gross DV01, with interval [0.475,2.224]. Positive states contribute 0.574 through curve repricing versus -0.706 in nonpositive states, while carry plus roll contributes 0.644 versus 0.779. The selection contrast therefore reflects more favorable curve movement rather than a higher mechanical carry contribution. It remains positive across every fixed block length.

11.2 Information timing and matched curve conventions

Delaying the entire conditional gap by six observations reduces the official-CMT forecast MSE improvement from 3.48 to 2.35 percent. The corresponding fitted-par change is from 0.50 to -0.43 percent and the zero-coupon change from 0.99 to 0.05 percent. The delayed official-signal funded mean is 0.663, compared with 0.722 under the original lag, and its forty-observation interval spans zero. The daily H.15 two-year reference improves official-spread MSE by 4.36 percent on the same dates. These results separate the role of a front-end rate condition from reliance on one specific fitted-forward representation.

The six-observation rule measures sensitivity to a publication delay. Initial curve estimates also determine synthetic coupons and exposure normalization, so interpretation remains at the level of historical curve-priced outcomes. Original-release curves and instrument-specific trading records would supply the information required for a real-time implementation test.

For convention matching, we retain the same GSW cash-flow pricing but select dates using GSW par and zero-coupon reference gaps as well as the primary official-CMT gap. Funded mean outcomes are 0.586 and 0.569 under the two GSW signals, compared with 0.722 under the official signal; their standard deviations are 3.85, 3.83 and 3.68. These comparisons use different signal-origin sets and do not impose equal precision across conventions. They show that the positive holding-outcome sign is not confined to pairing a CMT signal with GSW pricing.

11.3 Dependence, relative risk and nonoverlapping horizons

The primary twenty-observation outcomes overlap and the reference gap is persistent. Full-calendar inference therefore preserves the signal's actual timing, using fixed block lengths of 20, 40, 80 and 126 observations. Table 12 reports the complete grid. The absolute funded-mean interval crosses zero under the longest block, while the state-selection contrast and front-conditioning ablation retain positive intervals. The official forecast's loss gain over no change is also less precise under longer blocks. The most stable inferential result is the incremental information in front-end conditioning and gap-based selection.

Table 12. Fixed dependence-sensitivity grid

Block	Positive-gap mean	Positive minus nonpositive	Front versus no-front
20	[0.10,1.39]	[0.51,2.17]	[0.94,4.99]
40	[0.05,1.45]	[0.48,2.22]	[1.08,5.02]
80	[0.02,1.52]	[0.39,2.31]	[0.95,5.21]
126	[-0.05,1.61]	[0.36,2.43]	[0.75,5.53]

Notes: 95 percent full-calendar block intervals from 5,000 draws per block length. First two columns: funded P&L per gross DV01, with point estimates 0.722 and 1.327. Last column: paired forecast MSE improvement in squared basis points, point estimate 2.894. Blocks are fixed sensitivities, not independent replications or a significance-selection grid.

The funded steepener beta on subsequent two-year yield changes is -0.105 with a calendar-HAC interval [-0.137,-0.074]. The corresponding outright two-year beta is -0.881; their difference has interval [0.741,0.811]. The conditional variance ratio of steepener to two-year long is 0.034 with a block interval [0.024,0.059], and the ratio interval remains below one for every block length. The calendar-HAC scores are zero-filled at unselected dates, preserving actual spacing rather than treating signal observations as consecutive days.

Finally, we partition the complete origin calendar into all twenty offsets of nonoverlapping twenty-observation holding windows before applying the signal rule. The positive-state funded steepener mean is positive in every offset, ranging from 0.25 to 1.30 per gross DV01. The full

offset grid supplies a nonoverlap diagnostic using dependent partitions of the same historical sample.

12 Maturity-specific supply and valuation scenarios

12.1 Data and predetermined information

Long-end relative value depends on both front-end rates and the maturity composition of debt supply. We therefore construct a dated auction-information panel from TreasuryDirect's historical records. The source contains 2,371 nominal note and bond records. Selecting the ten-year note and thirty-year bond market segments over 2005–2026 leaves 474 auctions: 252 ten-year and 222 thirty-year auctions. The thirty-year group includes four early records labeled 29-year nine-month securities; the original labels are retained. Inflation-protected securities, floating-rate securities, twenty-year bonds, and other maturities are excluded. Four original offering announcements, covering both tenors in 2010 and 2026, reproduce the API offering amount, CUSIP, and auction date.

For an origin t, the known pending pipeline includes only auctions with announcement date strictly before t and auction date on or after t. Quantities announced subsequently are not assigned to the origin, even if the eventual auction falls inside its response horizon. Realized auction yields, bid-to-cover ratios, and bidder allocations are archived separately and never enter predetermined controls. The records form a current historical panel, not a complete vintage archive of every release and revision.

Let Q^{10}_{t-1} and Q^{30}_{t-1} denote these pending offering amounts in billions of dollars. Define

$$U_{t-1} = \frac{Q^{30}_{t-1} D^{30}_{t-1}/D^{10}_{t-1} - Q^{10}_{t-1}}{10}, \qquad V_{t-1} = \frac{Q^{30}_{t-1} D^{30}_{t-1}/D^{10}_{t-1} + Q^{10}_{t-1}}{10}. \tag{15}$$

The modified-duration proxies use semiannual par bonds and prior-close H.15 yields at ten and thirty years. They do not use coupons or yields determined at a future auction. U measures relative scheduled supply and V total scheduled supply in units of ten billion ten-year-equivalent dollars. This is a gross auction-pipeline proxy, rather than net debt supply or exact security-specific DV01. An unweighted dollar version is retained as a sensitivity.

12.2 Does the known pipeline absorb conditional adjustment?

On the same 2,934 daily origins beginning in 2015, we estimate

$$s_{t+20} - s_t = \alpha + \lambda g_{t-1} + \eta U_{t-1} + \xi V_{t-1} + \rho g_{t-1} U_{t-1} + \Gamma W_{t-1} + \delta_{\text{year}} + \delta_{\text{weekday}} + \epsilon_{t+20}, \tag{16}$$

where W contains the prior forward level and its prior five-observation change. We retain baseline, supply-level, interaction and unweighted-dollar specifications. Pending ten- or thirty-year auctions occur at 691 eligible origins. Calendar-HAC inference uses a bandwidth of forty common observations. These regressions evaluate conditional adjustment after controlling for the known auction pipeline.

Table 13 shows that adding the known pipeline scarcely changes the gap coefficient. The estimate is 0.528 for cash yields, 0.434 for fitted par yields, and 0.409 for zero-coupon yields after adding supply levels. The corresponding relative-supply level coefficients are -0.093 (0.097), -0.010 (0.121), and -0.034 (0.156). Supply interactions are also estimated imprecisely. Thus the conditional adjustment relation survives this concrete scheduled-supply control; the pipeline does not provide a clear explanation for its strength across calendar blocks.

Table 13. Conditional alignment with predetermined auction quantities

Spread	Calendar baseline gap	Supply-level gap	Gap × relative supply
H.15 CMT	0.528 (0.096)	0.528 (0.096)	0.023 (0.031)
GSW fitted par	0.434 (0.106)	0.434 (0.106)	0.012 (0.035)
GSW zero coupon	0.409 (0.095)	0.409 (0.095)	0.003 (0.030)

Notes: Calendar-HAC standard errors in parentheses; $N = 2,934$. All models include year and weekday fixed effects and lagged forward conditions. The last column comes from the interaction specification, not the supply-level specification. Units of relative supply are ten billion ten-year-equivalent dollars.

12.3 Supply expectations as a valuation scenario

The scheduled pipeline is one observable part of the information set. Prospective cuts to long-end issuance can change the relative scarcity and risk compensation of thirty-year exposure before any individual auction is announced. Holding other determinants fixed, a decline in expected long-end supply can support thirty-year prices and flatten the ten-year/thirty-year spread, offsetting a steepener thesis. An absence of the anticipated cut can remove that source of support. This supply scenario is therefore relevant to the conditional valuation and the choice of curve exposure.

The archived quarterly evidence illustrates why quantities and expectations must be distinguished. In November 2023, the announced three-month thirty-year total increased from 63 to 66 billion dollars while a dated MNI forecast had called for 69 billion. An increase relative to the prior quarter thus coexisted with a decrease relative to that named expectation. The pipeline regressions test whether known quantities absorb daily adjustment. The forecast audit clarifies how a supply-news interpretation would be measured. Neither is required to establish the conditional pricing and risk-expression results. Supplementary quantity-event estimates and dated guidance paths are retained in the appendices.

13 Portfolio interpretation and synthesis

The evidence links conditional pricing, subsequent adjustment, and portfolio expression. Uncertainty about yield direction motivates the comparison between outright duration and a relative expression. Front-end conditions then supply a valuation benchmark. A positive gap identifies relative flatness, historical forecasts test its subsequent information, and the DV01 comparison evaluates how that information can be expressed with lower directional sensitivity. Supply expectations belong in the scenario set because they can alter the long-end reference itself.

The portfolio implications follow from separating duration exposure from curve exposure. A curve position and an outright long can share a negative front-end beta while exhibiting very different sensitivities and scenario outcomes. An investor can therefore assess the conditional spread opportunity and choose directional duration separately. Our comparisons quantify this choice at common gross DV01.

The supplementary policy-event study shows immediate flattening after tightening surprises and an imprecisely estimated subsequent official-spread response over twenty observations. A ten-year decomposition associates the immediate repricing with both expected rates and term premia. These announcement-conditioned dynamics complement the daily valuation analysis by identifying circumstances in which the reference can change.

Taken together, the dated reconstruction, conditional adjustment, and lower-beta expression support the central relative-value narrative. The greatest empirical variation concerns the strength of forecast improvements across yield conventions and calendar periods. The full cash-flow extension adds carry, roll, nonlinear repricing, realized benchmark financing and execution-cost

scenarios. It preserves the low-risk expression result, while showing that mean-return rankings depend on financing and cost assumptions. Observed-security implementation remains a separate empirical layer.

14 Conclusion

The ten-year/thirty-year Treasury curve provides a conditional relative-value expression of front-end rate views. The October 16, 2026 application places the official curve 3.03 basis points flat against its trailing reference. Across the historical evaluation, lagged gaps predict subsequent spread adjustment, and frozen-reference accounting associates gap closure mainly with realized spread movement. The front-conditioning ablation establishes incremental information over a same-window mean reference. For the same funded steepener, positive-gap states outperform nonpositive states by 1.327 basis points per gross DV01, with positive intervals across the fixed block grid. Together, these results connect conditional pricing to adjustment and exposure selection.

A balanced steepener carries substantially less front-end sensitivity and volatility than outright duration and incurs much smaller average losses when front-end yields rise. Full cash-flow pricing preserves that relative protection after carry, roll, benchmark financing and assumed costs. Forecast strength and cost resilience vary across periods and yield conventions; longer blocks reduce the precision of positive absolute means. Known auction quantities leave the conditional adjustment relation essentially unchanged, while prospective issuance changes can shift the valuation reference. The resulting framework combines a conditional curve opportunity, an explicit choice of directional risk and supply-sensitive valuation scenarios.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

The authors used ChatGPT (OpenAI) for coding assistance, document preparation, and language editing. The authors take responsibility for the analysis, interpretation, and final text.

References

Acosta, M., A. Ajello, M. Bauer, F. Loria, and S. Miranda-Agrippino. 2026. Financial market effects of FOMC communication: Evidence from a new event-study database. Federal Reserve Bank of San Francisco Working Paper 2025-30, revised August 27. DOI: 10.24148/wp2025-30.

Adrian, T., R. K. Crump, and E. Moench. 2013. Pricing the term structure with linear regressions. *Journal of Financial Economics* 110(1): 110–138. DOI: 10.1016/j.jfineco.2013.04.009.

Board of Governors of the Federal Reserve System. 2026. Federal Funds Effective Rate [DFF]. H.15 Selected Interest Rates, distributed through FRED. https://fred.stlouisfed.org/series/DFF. Retrieved November 1, 2026.

Campbell, J. Y., and R. J. Shiller. 1991. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies* 58(3): 495–514. DOI: 10.2307/2298008.

Cochrane, J. H., and M. Piazzesi. 2005. Bond risk premia. *American Economic Review* 95(1): 138–160. DOI: 10.1257/0002828053828581.

Crump, R. K., S. Eusepi, and E. Moench. 2016. The term structure of expectations and bond yields. Federal Reserve Bank of New York Staff Report 775, revised April 2018. https://www.newyorkfed.org/research/staff_reports/sr775.html.

Diebold, F. X., and C. Li. 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* 130(2): 337–364. DOI: 10.1016/j.jeconom.2005.03.005.

Diebold, F. X., and R. S. Mariano. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13(3): 253–263. DOI: 10.1080/07350015.1995.10524599.

Greenwood, R., and D. Vayanos. 2014. Bond supply and excess bond returns. *Review of Financial Studies* 27(3): 663–713. DOI: 10.1093/rfs/hht133.

Gürkaynak, R. S., B. Sack, and J. H. Wright. 2007. The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics* 54(8): 2291–2304. DOI: 10.1016/j.jmoneco.2007.06.029.

Hanson, S. G., and J. C. Stein. 2015. Monetary policy and long-term real rates. *Journal of Financial Economics* 115(3): 429–448. DOI: 10.1016/j.jfineco.2014.11.001.

Hu, G. X., J. Pan, and J. Wang. 2013. Noise as information for illiquidity. *Journal of Finance* 68(6): 2341–2382. DOI: 10.1111/jofi.12083.

Jarociński, M., and P. Karadi. 2020. Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics* 12(2): 1–43. DOI: 10.1257/mac.20180090. Author-updated shocks retrieved October 31, 2026.

Jordà, Ò. 2005. Estimation and inference of impulse responses by local projections. *American Economic Review* 95(1): 161–182. DOI: 10.1257/0002828053828518.

Kim, D. H., and J. H. Wright. 2005. An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates. Finance and Economics Discussion Series 2005-33, Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/pubs/feds/2005/200533/200533abs.html.

Koijen, R. S. J., T. J. Moskowitz, L. H. Pedersen, and E. B. Vrugt. 2018. Carry. *Journal of Financial Economics* 127(2): 197–225. DOI: 10.1016/j.jfineco.2017.11.002.

Künsch, H. R. 1989. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* 17(3): 1217–1241. DOI: 10.1214/aos/1176347265.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3): 703–708. DOI: 10.2307/1913610.

Vayanos, D., and J.-L. Vila. 2021. A preferred-habitat model of the term structure of interest rates. *Econometrica* 89(1): 77–112. DOI: 10.3982/ECTA17440.

Yang, F. 2026. Oil-price innovations and the Treasury term premium: Evidence from 2010–2026. Working paper, GARCH Institute. DOI: 10.2139/ssrn.7497760.

A Research design and empirical counterparts

The analysis combines public yield curves, dated auction records, and monetary-policy measures. The pricing and exposure tests use dated curve states. Event studies address narrower conditioning questions.

Table 14. Research propositions and empirical counterparts

Research proposition	Empirical counterpart
Yield direction and curve shape	Dated official-yield changes and a simultaneous conditional-pricing application.
Front-end conditional reference	Prior-126-observation forward mappings and full-history fit diagnostics.
Illustrative relative flatness	Official-CMT gap 3.03 and fitted-par gap 3.34 on October 16, 2026.
Subsequent adjustment	Lagged chronological forecasts, simple benchmarks and frozen-reference accounting.
Lower-beta curve expression	Same-origin, equal-gross-DV01 steepener versus two-year and ten-year long exposures.
Protection during rising rates	Ex-post rate-direction partitions and comparative exposure outcomes.
Debt-management scenarios	Predetermined auction controls, quantity-event evidence and dated forecast audits.
Portfolio exposure choices	Separate curve and duration sensitivities under common gross DV01.

The dated calculations, reference-fit series, forecasts and exposure scenarios are retained with the supporting event analyses in the replication package.

B Supplementary policy-event evidence

B.1 Shocks, timing, and information sets

We use the announcement-level monetary policy and central-bank information shocks supplied by Jarocinski and Karadi (2020), using their median-rotation decomposition. The two components enter every primary event equation jointly. A unit regressor corresponds to a ten-basis-point contribution to the source policy-surprise factor, not to a ten-basis-point change in the policy target. This distinction matters because the factor summarizes changes in derivatives at several short horizons.

The source contains 334 announcements. Exact matching to the Treasury panel, a preceding 126-observation reference window, and a complete twenty-observation response leave 169 eligible events from September 20, 2006 to September 16, 2026. Five source records with missing equity surprises contain zero-valued decomposed components despite nonzero rate surprises; they are excluded as unavailable decompositions rather than treated as zero policy shocks. Another 156 announcements fall outside the common cash-yield panel and four lack the preceding reference window. No eligible observations are dropped according to their estimated effects. Thirteen included events have another source announcement within the following twenty observations.

Let e index announcements and t_e denote the matched common yield date. Immediate responses are $v_{t_e} - v_{t_e-1}$, post-event responses are $v_{t_e+h} - v_{t_e}$, and total responses are $v_{t_e+h} - v_{t_e-1}$. We estimate

$$\Delta_e^{(h)} v = \alpha_h + \beta_h^{MP} MP_e + \beta_h^{CBI} CBI_e + \Theta_h W_{e,-1} + \epsilon_{e,h}, \tag{17}$$

where $W_{e,-1}$ contains the prior alignment gap, the prior two-year yield, and its preceding

five-observation change. The reference intercept and slope are estimated strictly before each announcement and frozen over its response horizon. Freezing the reference prevents later re-estimation from mechanically erasing a displacement. This event specification uses the two-year cash rate so that dated cash-market conditions can be analyzed without assuming daily publication of fitted GSW forwards.

Outcomes use daily H.15 yields, whereas the shocks use intraday announcement windows. Consequently, the immediate estimate is a shock-linked daily response, not a thirty-minute yield response. Announcements after 15:30 New York time and dates without common yield observations are excluded by rule. Calendar-time Bartlett HAC covariance estimates use a 40-common-observation bandwidth; HC1 results are retained as a sensitivity. The intervals condition on the author's shock decomposition and the estimated reference curves.

B.2 Immediate displacement and the twenty-observation response

Table 15 reports the controlled median-decomposition results. A ten-basis-point tightening surprise raises the two-year yield by 8.64 basis points on the announcement date. The ten-year response is 5.43 basis points and the thirty-year response is 1.68, producing a 3.75-basis-point flattening. The estimated conditional-reference gap increases by 2.28 basis points. Thus the observed relative move is not fully summarized by the recent cash-rate reference relationship.

Table 15. Responses to a ten-basis-point monetary policy surprise

Outcome	Immediate	Subsequent 20	Total 20
Two-year yield	8.64 (1.68)	3.22 (4.43)	11.86 (5.01)
Ten-year yield	5.43 (1.98)	9.75 (5.95)	15.17 (6.49)
Thirty-year yield	1.68 (1.61)	8.58 (5.07)	10.26 (5.51)
Thirty-year minus ten-year	-3.75 (0.71)	-1.17 (1.78)	-4.92 (1.91)
Frozen-reference gap	2.28 (0.60)	1.03 (1.66)	3.31 (1.83)

Notes: Responses in basis points; calendar-HAC standard errors in parentheses. All equations include both median-decomposition shocks and the same predetermined controls; $N = 169$. Subsequent 20 excludes the announcement-day move. Total equals immediate plus subsequent response.

The subsequent twenty-observation spread coefficient is -1.17 basis points, with standard error 1.78 and a 95 percent interval of [-4.65,2.31]. The point estimate continues the initial flattening, with substantial uncertainty around its direction. The total twenty-observation response is -4.92 basis points. The announcement sample therefore shows a policy-linked displacement and an imprecise post-event path. Figure 4 reports the complete cumulative response profile at the specified horizons.

B.3 Alignment within the event sample

We also regress the post-event spread change on the event-close gap, controlling for the pre-event gap, both shocks and prior front-end conditions. At twenty observations, the event-close gap coefficient is 0.275 with standard error 0.219. Excluding the thirteen overlapping event horizons gives 0.216 with standard error 0.245. Adding the announcement-day spread change gives 0.072 with standard error 0.329. Announcement-sample gap closure is therefore imprecisely estimated. Leave-one-event-out coefficients for the baseline gap association range from 0.131 to 0.371, preserving a positive point estimate across omitted events.

The daily sample evaluates prediction across dated market states. The announcement sample isolates monetary-policy and central-bank-information innovations and follows their associated yield paths. Together, the designs distinguish a conditional valuation signal from one particular

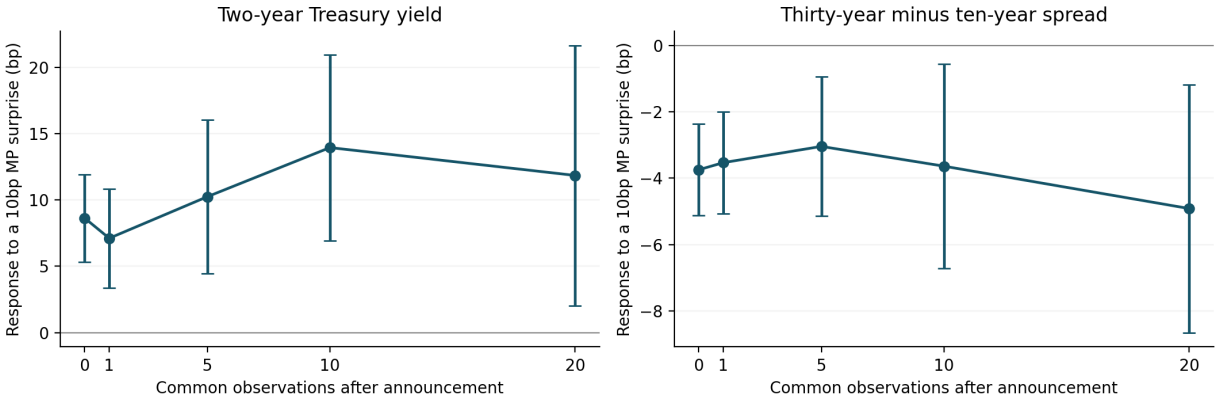


Figure 4. Cumulative daily responses around monetary policy announcements. Bars show 95 percent calendar-HAC intervals. Horizon zero is the announcement-day response.

source of reference displacement.

B.4 Alternative communication windows

As a sensitivity, we use the publicly supplied statement and combined monetary-event surprises from the San Francisco Fed USMPD (Acosta et al., 2026). The combined window includes the statement and post-meeting press conference. Matching to the eligible JK event sample leaves 166 events. These factors are not purged of central-bank information and are normalized differently from the JK components, so their response magnitudes should not be equated directly.

The combined-window surprise produces an immediate spread coefficient of -5.30 basis points and a subsequent twenty-observation coefficient of -3.18, with respective standard errors 0.65 and 2.21. Thus the broader window preserves immediate flattening but supplies no clear evidence of subsequent steepening. Both statement and combined-window estimates, as well as the poor-man sign decomposition, are retained in the replication output rather than selected according to their significance.

C Supplementary ten-year component evidence

We next use the published Kim–Wright fitted ten-year zero-coupon yield and ten-year term premium, distributed through FRED. The expected-average-short-rate component is their difference (Kim and Wright, 2005). These public estimates extend only to ten years, so they do not supply the thirty-year premium or the thirty-year minus ten-year premium differential. We retain the GSW ten-year minus KW fitted ten-year difference as a separate cross-model basis.

All 169 original event dates and their original cash-calendar endpoints can be matched exactly; no missing component is forward-filled. Both policy and information shocks, the prior cash gap, and prior front-end conditions enter every equation. Table 16 decomposes only the ten-year fitted yield. The announcement-day response comprises 2.60 basis points from the expected-rate component and 2.38 from the premium component. Their sum is the fitted-yield response of 4.98 basis points. Over the following twenty observations the expected component rises by 6.82 basis points and the premium component by 4.35, with the latter response estimated less precisely.

The zero-coupon spread falls by 5.43 basis points on the announcement day and a further 7.12 over the following twenty observations, with standard errors 0.90 and 3.51. The frozen-reference gap increases by 4.43 and 8.12 basis points. Adding the predetermined ten-year premium level and its preceding five-observation change leaves the post-event spread response at -6.35 (3.11). These results locate persistent announcement-linked dynamics alongside daily conditional adjustment.

Table 16. Ten-year components and zero-coupon curve responses

Outcome	Immediate	Subsequent 20
KW fitted ten-year yield	4.98 (1.45)	11.17 (6.28)
KW expected-rate component	2.60 (0.71)	6.82 (3.21)
KW term-premium component	2.38 (0.78)	4.35 (3.26)
GSW ten-year minus KW fitted	-0.09 (0.60)	0.93 (0.63)
GSW thirty-year minus ten-year	-5.43 (0.90)	-7.12 (3.51)
Frozen zero-spread reference gap	4.43 (1.07)	8.12 (3.48)

Notes: Responses in basis points per ten-basis-point MP surprise-factor contribution; calendar-HAC standard errors in parentheses. $N = 169$. Both source shocks enter jointly. The zero-spread event gap conditions on the two-year zero yield and freezes the pre-event 126-observation mapping. KW components are model estimates, not observed expectations or premiums.

The available KW decomposition measures the ten-year sector; a thirty-year decomposition would be needed to attribute the long-end premium differential.

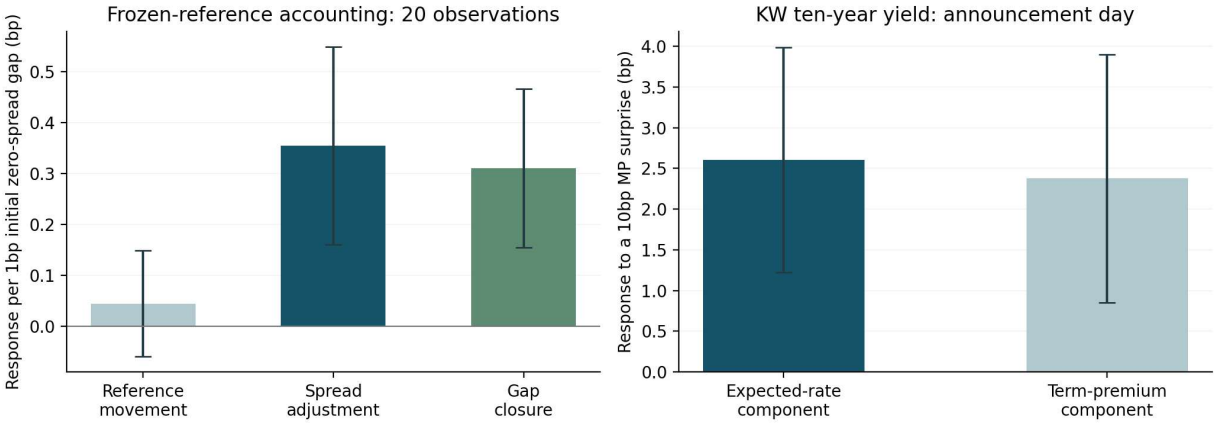


Figure 5. Conditional reference accounting and the ten-year response decomposition. Bars show 95 percent conditional HAC intervals. Left: twenty-observation daily associations per basis point of initial gap. Right: announcement-day component responses per ten-basis-point policy surprise.

D Supplementary supply-announcement evidence

D.1 Announcement-size changes and the auction cycle

We separately aggregate securities by announcement date and relate spread responses to ten- and thirty-year offering-size changes from the preceding auction in each market segment. These are recorded size changes, not innovations relative to dealer expectations. The event regressions include both tenor changes, the count of newly issued securities, the prior cash gap, prior forward conditions, and year fixed effects. Exact dates with a preceding reference window and twenty future observations leave 234 announcements.

The original size-change specification associates a ten-billion-dollar increase in thirty-year size with a 7.96-basis-point subsequent cash-spread increase, standard error 3.92. The corresponding zero-spread coefficient is 2.40, standard error 6.77. Because new and reopened auctions have systematically different sizes, we add a sensitivity comparing each auction with the previous observation in the same tenor and calendar-month-within-quarter group. Both change definitions are then evaluated on the same 231 announcement dates. Table 17 reports all tenor effects for this matched comparison. The thirty-year cash coefficient under the cycle-adjusted measure is

7.19 with standard error 7.13. The isolated positive association in the original cash specification is therefore not a stable cross-measurement explanation of curve adjustment.

Table 17. Size changes on matched announcement dates

Spread	Size change	Previous auction	Same cycle month
Cash	10-year	2.63 (1.89)	0.82 (1.36)
Cash	30-year	8.67 (3.96)	7.19 (7.13)
Zero coupon	10-year	5.13 (3.76)	1.75 (1.85)
Zero coupon	30-year	4.61 (6.47)	4.76 (9.70)

Notes: Subsequent twenty-observation spread responses in basis points per ten billion dollars of size change; calendar-HAC standard errors in parentheses. Both definitions use $N = 231$ dates. The same-cycle sensitivity was added after the initial size-change estimates; the original specifications are retained.

D.2 The relevant remaining supply margin

The auction pipeline measures quantities already announced for specific auctions. Prospective long-end issuance cuts concern a broader debt-management information set, including quarterly guidance, later auction paths, redemptions and Federal Reserve absorption. The quarterly extension distinguishes announced quantities from available forecast benchmarks. The daily regressions establish that routine known auction quantities leave the conditional alignment relation essentially unchanged.

D.3 Quarterly quantities and same-day monetary news

Quarterly refunding dates are identified from ten-year new issues auctioned in February, May, August, or November, paired with a thirty-year-segment security announced on the exact same date. Dates are not inferred from a first-Wednesday rule. Adjacent paired quarters and complete yield windows leave 78 events from May 2, 2007 to August 5, 2026. Thirty-two coincide with a source FOMC announcement. Excluding all such dates leaves 46 events, fifteen with a nonzero change in either initial auction size.

Regressions include initial ten- and thirty-year size changes, prior gap and forward conditions, a linear calendar trend, and auction-quarter fixed effects. The initial auction quantity is not the complete quarterly path, and its change is not a surprise measure. We report the non-FOMC sample and an all-event sensitivity controlling monetary-policy and information surprises. Calendar-HAC covariance estimates have finite-sample Student-t reference intervals.

Table 18. Quarterly size changes and subsequent twenty-observation responses

Spread	Size change	Non-FOMC ($N = 46$)	All with shocks ($N = 78$)
Official CMT	10-year	-0.45 (6.35)	1.17 (8.23)
Official CMT	30-year	16.74 (6.57)	9.73 (9.42)
GSW zero	10-year	-4.71 (15.34)	3.52 (13.76)
GSW zero	30-year	22.92 (14.30)	5.15 (13.82)

Notes: Basis points per ten billion dollars of initial size change; calendar-HAC standard errors in parentheses. The specifications differ in event coverage and controls. All immediate and subsequent five- and twenty-observation estimates are retained.

The non-FOMC thirty-year coefficient is positive in both conventions but more precise in official CMT yields: 16.74 (6.57), compared with 22.92 (14.30) for zero-coupon yields. Leave-one-event-out CMT coefficients range from 9.59 to 21.14; the zero-coupon range crosses zero.

The all-event estimates are 9.73 (9.42) and 5.15 (13.82). These are conditional quantity-change associations with material sample and measurement sensitivity, rather than identified exogenous supply-news effects.

The forecast audit in Appendix E clarifies the economic distinction. In November 2023 the announced three-month thirty-year total rose from 63 to 66 billion dollars, while an October 31 MNI forecast put the new total at 69 billion. A positive announced increase therefore coexisted with a negative deviation from this specified forecast. This distinction matters when evaluating prospective long-end issuance cuts.

E Dated forecasts and prospective quarterly paths

For a named pre-event benchmark $E^b_{t^-} Q_t$, an announced quantity change decomposes as

$$Q_t - Q_{\text{prev}} = (E^b_{t^-} Q_t - Q_{\text{prev}}) + (Q_t - E^b_{t^-} Q_t). \tag{18}$$

The second term is a deviation from the specified forecast. Calling it market-wide news requires representative expectations and a suitable identification design.

Four dated public pre-announcement reports are audited. The 2021 record is a TBAC-based schedule reported by MNI, kept distinct from MNI's own forecasts. Complete three-month paths are separately verified for August 2023, November 2023, and October 2024. Eighteen monthly-tenor entries are read from contemporary forecast and official guidance tables. Later realized auction records provide a consistency check only, not the source of ex-ante announced quantities.

Table 19. Matched first-auction forecasts and announcements

Event	Forecast 10/30	Announced 10/30	Error 10/30	FOMC
2021-11-03	38/25	39/25	+1/+0	Yes
2023-08-02	38/23	38/23	+0/+0	No
2023-11-01	41/25	40/24	-1/-1	Yes
2024-10-30	42/25	42/25	+0/+0	No

Notes: Billions of dollars. Forecast dates: November 2, 2021; August 1, 2023; October 31, 2023; October 28, 2024. The 2021 benchmark is a TBAC-based schedule, not a consensus estimate. Other entries are named MNI forecasts.

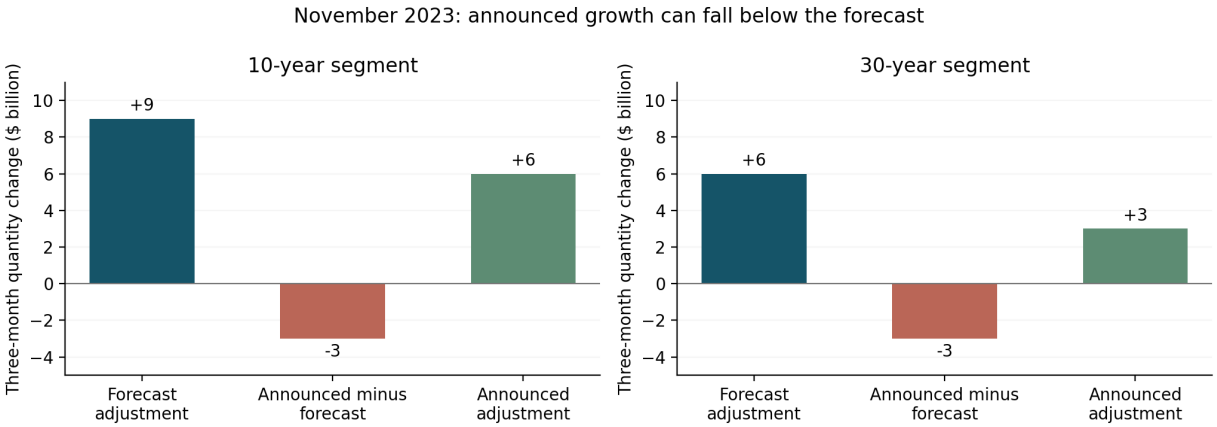


Figure 6. Positive quantity growth and a negative forecast deviation can coexist. November 2023 three-month quantities compared with the specified MNI forecast. The bars represent arithmetic identities, not estimated causal responses.

Full paths match the selected forecasts in August 2023 and October 2024. In November 2023, announced ten- and thirty-year totals are each three billion dollars below the selected forecast. Both nonzero first-auction benchmark deviations in the four-case audit occur on FOMC dates. The examples do not constitute a clean non-FOMC consensus-surprise regression sample. Missing expectations are not filled with prior quantities.

Treasury's official semiannual auction-size survey is also examined. Its October 2023 questions concern fiscal-year-end 2024–2026 auction sizes, which are not interchangeable with next-refunding forecasts. Source reports: November 2021 preview, August 2023 preview, November 2023 preview, October 2024 preview. Dates, hashes, and numerical facts are retained; copyrighted full reports are not redistributed.

F Dating sensitivity around the CMT method change

Treasury documents that monotone-convex interpolation replaced the earlier quasi-cubic Hermite spline method on December 6, 2021.[1] This date is close to the calendar division used to describe stronger later-period alignment. We reuse the original fixed forecasts, without refitting, and isolate origins whose targets end before the change or whose entire signal reference window begins after it.

The first set has 1,713 origins from January 2, 2015 to November 3, 2021; the second has 1,074 origins from June 8, 2022 to September 24, 2026. Every convention uses identical dates within each set. Intervening target and reference windows are excluded.

Table 20. Unchanged fixed forecasts on method-consistent dated windows

Window	Spread	MSE gain (%)	Paired gain interval
Pre-change	Official CMT	-0.87	[-2.34,1.65]
Pre-change	GSW par	-1.40	[-2.42,1.30]
Pre-change	GSW zero	-1.75	[-4.32,2.01]
Post full-window	Official CMT	5.46	[1.42,7.65]
Post full-window	GSW par	1.50	[-1.11,4.24]
Post full-window	GSW zero	2.07	[-1.26,8.66]

Notes: Gain intervals are 95 percent paired block intervals in squared basis points, not percentage points. Forty-observation blocks and 2,000 draws; original calibration and forecasts unchanged.

The later CMT gain remains positive; fitted-par and zero-coupon gains are smaller and less precise. The windows span macroeconomic changes and a documented interpolation-method date jointly. Their comparison supports evaluating the official curve alongside consistent fitted-yield conventions.

[1]The interpolation change is documented in the Treasury's yield curve methodology description.