

Treasury ALM: hedge stability, longevity and liquidity

Prasanna K | Research completed 19 October 2026 | Nominal USD

Finding and decision

A development-selected, constrained covariance hedge reduced the final-period surplus RMSE from 15.87 bp to 6.50 bp of initial liability present value relative to a 5/20-year duration match. The reduction was 59.03%; the paired six-month block-bootstrap 95% interval for the RMSE difference was [-12.34, -6.02]bp. This supports the selected hedge against the declared duration benchmark in this historical experiment. It does not establish universal dominance, an executable trading return or a better outcome than every untested hedge.

The separate September 2026 synthetic benefit case shows why the rate result is insufficient for an ALM decision. A mortality-force multiplier of 0.8 raises the liability by USD 6.95 million (5.82%) against every fixed, equally funded asset portfolio. A 2% assumed liquidation haircut leaves the covariance portfolio with a USD 0.705 million year-ten funding deficit at base mortality and reinvestment. The rate hedge should therefore be reviewed alongside longevity exposure, cash-flow timing and the cost of forced sales.

1. Question, population and source controls

The historical question is whether a constrained multi-node hedge improves out-of-sample stability of assets less a defined liability, beyond a simple duration match. The current case asks which other balance-sheet risks remain. Neither case represents an actual employer, insurer, pension plan or bond book.

The original Federal Reserve GSW nominal yield-curve CSV is pinned at 16,657,704 bytes and SHA256 7286435090933f0df1dcaeb15629f3a9129c083baee4c42ae5df03d0bb89335c. GSW SVENY values are continuously compounded percent; SVENPY values use a different coupon-equivalent convention and are excluded. The Svensson beta parameters are percent and tau parameters years. For maturity t , a decimal zero rate $y(t)$ gives $D(t)=\exp(-t*y(t))$. All 30 published zero maturities must reconcile within a fixed 0.02 bp tolerance.

The input-window audit has 6,712 complete daily parameter/yield rows, of which 6,710 pass. The two quarantined dates are 16 July 2008 (0.029257354 bp maximum discrepancy) and 16 October 2013 (0.037964987 bp). Both are outside their month-end selection. The audit also retains 288 all-missing dates and one partial row. No tolerance was widened, no raw row deleted, and no exclusion used a hedge outcome. All 313 required month-end curves and the 30 September 2026 current curve pass the original rule and seven-calendar-day staleness limit.

This is a pinned 2026 downloaded vintage of revised historical curves. It is not immutable point-in-time data. Source-convention research exposed some historical curves before evaluation; complete blindness to public rates is not claimed. The separate SSA transcription includes published 2023 period qx for ages 0-119 and both sexes, SHA256 1cc6300b231972794d25dbf8ed53f559ffe2557c2082e52d9e8774b66e41ae35. It describes the US Social Security area population, not Indian annuitants or future cohort mortality. No Indian curve or mortality input enters this study.

2. Historical design and selection

The December 1999 anchor and January 2000-December 2025 month-end curves produce 312 one-month observations, labelled by ending month. There are 180 training months (2000-2014), 48 development months (2015-2018) and 84 final months (2019-2025). No missing month is interpolated.

Each start date defines a fresh synthetic liability paying 100 nominal units at annual maturities 1-30. Hypothetical zeros at years 1, 5, 10, 20 and 30 are fully funded to its initial PV. All maturities roll by actual calendar days divided by 365.25 and are repriced on the next selected curve. The interval is shorter than the first payment or maturity. The

outcome is (next asset PV - next liability PV) / initial liability PV. These reset experiments are not a continuously financed portfolio wealth path.

Four candidates were specified before outcomes:

Candidate	Construction
Five-year zero	All initial asset PV in the five-year zero
Duration match	Analytic nonnegative 5/20-year weights matching liability duration
Key-rate hedge	Minimise squared five-node key-duration mismatch, long-only, each weight at most 60%
Covariance hedge	Minimise exposure' Sigma exposure plus 1e-8 times squared weights, with the same bounds

Key-duration bump hats are piecewise linear at the five asset nodes and partition a parallel shift. Analytic derivatives are checked against central differences. Sigma uses only training-period monthly changes in decimal zero rates, with fixed 25% diagonal shrinkage. It is never updated after 2014. The covariance objective is a local risk approximation, not a promise of minimum future realised surplus variance.

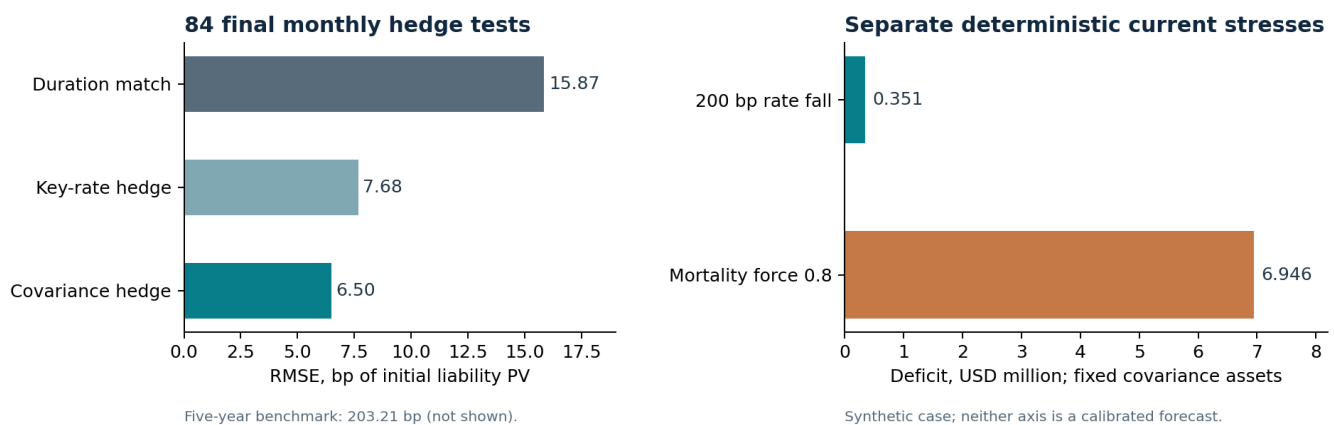
Selection uses development surplus RMSE. Within 1% of the minimum, prefer five-year, duration, key-rate, then covariance, in that order. Covariance was selected (development RMSE 4.50 bp versus 5.01 bp for key-rate and 13.04 bp for duration). The selected method, source, code and development results were frozen before final prediction generation; the final predictions were sealed before final scoring. No tuning used final outcomes.

3. Final historical evidence

All values in this table are basis points of each observation's initial liability PV, not yield basis points or percentage investment returns.

Method	RMSE	Mean surplus	Mean absolute surplus	Worst surplus	Empirical 5th percentile
Five-year zero	203.21	11.42	163.74	-551.68	-350.34
Duration match	15.87	-0.71	11.27	-58.10	-24.61
Key-rate hedge	7.68	0.55	6.14	-27.49	-10.05
Covariance hedge	6.50	0.49	5.01	-22.85	-9.89

The primary selected-minus-duration RMSE difference is -9.3656 bp. Its paired circular six-month block-bootstrap interval is [-12.3423, -6.0242]bp from 5,000 draws, seed 20261019. Pairing preserves common months across methods; blocks address short-run calendar dependence. The 84 months still cover few economic regimes, and the chosen block scheme does not remove all structural uncertainty. The empirical tail is descriptive, not a certified capital measure. No separate superiority test against key-rate is asserted.



4. Current liability, fixed portfolios and stresses

The base date is 30 September 2026. Four synthetic cohorts each contain 250 lives: male age 60, female 65, male 70 and female 75. Each survivor receives USD 12,000 annually in arrears. Survival to age $x+1$ uses q_x . A force multiplier m changes q to $1 - (1 - q)^m$; $m=0.8$ and 1.2 are imposed sensitivities, not forecasts. Payment at age 120 is allowed; later benefits are zero, equivalent to $q_{120}=1$. The alternative closure age is 110. Base curve extrapolation follows Svensson beyond 30 years; alternatives use a flat 29-30 forward or flat 30 year zero.

Base measure	Value
Liability PV, USD	119,380,495.42
Duration, years	8.473573
Convexity, years squared	113.758999
DV01 magnitude, USD per 1 bp	101,157.93
PV of benefits after 30 years, USD	583,401.06
Share of PV after 30 years	0.48869%

Each candidate is funded once against the base liability. Initial faces remain fixed in all stresses; there is no stress-specific rehedgeing. The 14 axes (including base) produce 56 method/scenario results. The largest absolute surplus among the eight declared rate stresses is USD 10.742 million for five-year, USD 0.371 million for duration, USD 0.688 million for key-rate and USD 0.351 million for covariance. The covariance portfolio's 200 bp parallel fall gives a USD 350,957.08 deficit. These deterministic axes do not create a new out-of-sample statistical test.

Force 0.8 raises liability PV to USD 126,326,153.13, a USD 6,945,657.71 deficit against equally funded fixed assets. Force 1.2 lowers PV to USD 113,491,792.81. Closing payments at 110 lowers PV by USD 49.79; the two alternative curve tails lower it by USD 7,930.27 and USD 21,482.99 respectively. These magnitudes are specific to this mortality table, population and portfolio.

5. Liquidity, reinvestment and event-ledger controls

The complete 108-case grid is four portfolios times three mortality forces, three assumed liquidation haircuts (0/1/2%) and three cash-rate shifts (-1/0/+1 percentage point). Each case runs for ten years. Cash accrues by $D(y-1)/D(y) \cdot \exp(\text{cash_shift})$. Zeros mature before benefits are paid. A cash shortage causes pro-rata sales of all remaining zeros at $D(T)/D(y)$, less the imposed haircut. These are deterministic initial-curve forwards, not a model of future market rates, executable prices or observed bid/ask spreads.

The ledger preserves maturities, cash interest, sales, haircut losses, paid and unpaid benefits, closing cash, remaining faces and residual liability. Unpaid benefits are forward-carried at base forwards and deducted from reported final surplus. Synthetic depletion tests verify that liabilities cannot disappear.

At base mortality and cash-rate shift zero with a 2% haircut:

Method	Gross sales, USD million	Nominal haircut losses, USD million	Year 10 net surplus, USD million
Five-year zero	46.7453	0.9349	-1.4066
Duration match	46.7453	0.9349	-1.4066
Key-rate hedge	24.6332	0.4927	-0.6411
Covariance hedge	26.3338	0.5267	-0.7050

The nominal flow sum of haircut losses and the year-ten funding stock have different valuation dates. Lost cash also loses forward accumulation; these quantities must not be treated as identical. No unpaid amount exceeds USD 0.01 anywhere in the actual 108-case grid. Raw tiny floating-point residuals are retained, with maximum cumulative unpaid USD 5.59e-9. No real default or calibrated shortfall probability is established.

6. Independent checks, workbook and reproduction

The 31 extension boundary checks cover source conventions, analytic/finite- difference risk, exact calendar roll, constrained objectives, annuity timing, depletion and wealth conservation. The separate 50-digit Decimal verifier rebuilds source discounts, cohort survival, all 56 instantaneous results and all 1,080 annual runoff records: 23,714 comparisons, USD 0.0001 financial tolerance, maximum financial discrepancy about USD 8.01e-8. Independent exhaustive enumeration of 243 bound-face combinations checks each current constrained optimum. Optimiser objective gaps are below 2 e-14.

The seven-sheet formula workbook independently calculates current survival, discounts, fixed-asset values, stresses and runoff. It reprices 12 historical portfolio-month combinations and aggregates all 84 final months into the four RMSEs and biases. All 30 current source yields reconcile. The 22,166 comparisons pass with maximum financial discrepancy USD 1.3411e-7, below USD 0.0001 tolerance; historical surplus fractions use 2 e-12. Historical weights and current initial faces are imported, not optimised inside Excel. Missing/invalid inputs fail visibly; zero benefits remain valid. All 7 sheets were rendered and reviewed. The actual exported XLSX was reopened and recalculated after a mortality-input change in a disposable copy. Native desktop Excel was not executed.

One clean integrated run copies the pinned source and required methods into an isolated directory, reruns all historical phases, regenerates the current case and independently rebuilds the workbook. All 5 historical and 3 current scientific files reproduce byte-for-byte. New historical seal timestamps are kept separately; current-case provenance retains the immutable original freeze. Workbook formulas, inputs, cached results, formatting, validations and panes match. Only ZIP/core date metadata and relationship identifier names may vary; every relationship must resolve to the same type/target. No numerical field is ignored. This is same-environment reproduction, not external replication.

7. Practical recommendation and limits

Use the covariance method as the documented rate-risk candidate supported against the duration benchmark in this setting. Retain the duration and key-rate comparators: key-rate has a slightly smaller deficit in the specific 2% haircut/base-cash runoff case. Do not substitute a rate-RMSE improvement for an overall liquidity or longevity mandate. A real decision would need actual liability data, executable asset choices, transaction costs, governance limits and mortality evidence appropriate to the population. This research supports model comparison and risk discussion, not a regulated valuation, solvency capital certification, investment recommendation or realised-profit claim.

Evidence map and sources

- Historical protocol and phases: `extensions/treasury_alm/PROTOCOL.md`, `run_study.py`; sealed outputs in `reports/treasury_20261019/`.
- Current implementation freeze: `f0168dd5a89af997e123ca2aabf06bf2e46d69e8`. Calculations: `CURRENT_CASE_METHODS.md`, `current_case.py`, `run_current_case.py`.
- Current results/ledger: `reports/current_case_20261019/`.
- Workbook: `artifacts/Treasury_ALM_Reconciliation.xlsx`; detailed guide and numeric/visual records: `reports/workbook_20261019/`.
- Integrated receipt: `reports/integrated_20261019/full_study_reproduction.json`.
- Environment: Python 3.12.14, NumPy 2.3.5, SciPy 1.17.0. Workbook builder requires the Artifact Tool JavaScript runtime; the saved XLSX exposes its formulas without that build dependency.
- Federal Reserve data and conventions: <https://www.federalreserve.gov/data/nominal-yield-curve.htm> and https://www.federalreserve.gov/data/yield-curve-tables/feds200628_1.html.
- Original CSV: <https://www.federalreserve.gov/data/yield-curve-tables/feds200628.csv>.
- GSW paper: <https://www.federalreserve.gov/pubs/feds/2006/200628/index.html>.
- SSA period table: <https://www.ssa.gov/oact/STATS/table4c6.html>.
- Source-use policies: <https://www.federalreserve.gov/disclaimer.htm> and <https://www.ssa.gov/policy/accessibility.html>.

No source-owner emails, paid data or paid inference were needed.