

When a simple inflation forecast survives stricter evidence

A release-aware, frozen evaluation of Indian headline CPI inflation. Old base 2012; July 2024-December 2025 final test.

Prasanna K | Economic research portfolio | Prepared 17 October 2026

Decision

Retain development-selected persistence for this bounded old-base experiment. Its held-out RMSE is **0.819 pp**, versus **1.690 pp** for the fixed Ridge AR(1). The complex challenger does not earn adoption.

Method	MAE (pp)	RMSE (pp)	Bias (pp)
persistence	0.6828	0.8185	+0.2083
ridge_ar1	1.5397	1.6901	+1.4053
seasonal_12month	2.3356	2.8280	+2.1344

Evidence strength

Primary paired MSE contrast, AR(1) minus persistence: **2.186524 pp squared**. Conditional 95% moving-block interval: **[0.613693, 3.816555]**. All 18 release dates are held out from fitting and development selection. Two thousand circular draws use three-month blocks and seed 20261017.

This is uncertainty about average error in a short dependent panel, not a prediction interval or a guarantee about new-base/current forecasting. Publication-dated documents were retrieved retrospectively and may have been edited.

Research contribution

A reusable, audited vintage panel; explicit information timing; numerical measurement of hindsight contamination; a rejected challenger; and a stopping rule at the 2026 basket/base break.

1. Measurement and frozen design

48 official monthly releases yield 144 old-base observations: current provisional, previous-month final and prior-year reference. Every value carries a source hash, exact HTML cell and publication timestamp. Seven of 47 revision pairs change, by at most 0.10 pp.

What the forecast may know

Cutoff is one minute after the previous reference month's printed PIB publication timestamp, with IST offset. An as-of query must match reference month, base 2012, headline YoY and Combined sector. It cannot use a final revision first published with the target release. Missing or conflicting evidence fails rather than becomes zero.

Training and selection

Training: 23 lag/target pairs, February 2022-December 2023. The feature standardizer is fitted only on training observations. Challenger: Ridge AR(1), alpha 1, intercept, no parameter search. Coefficients remain fixed through the entire test; observed lags update when available.

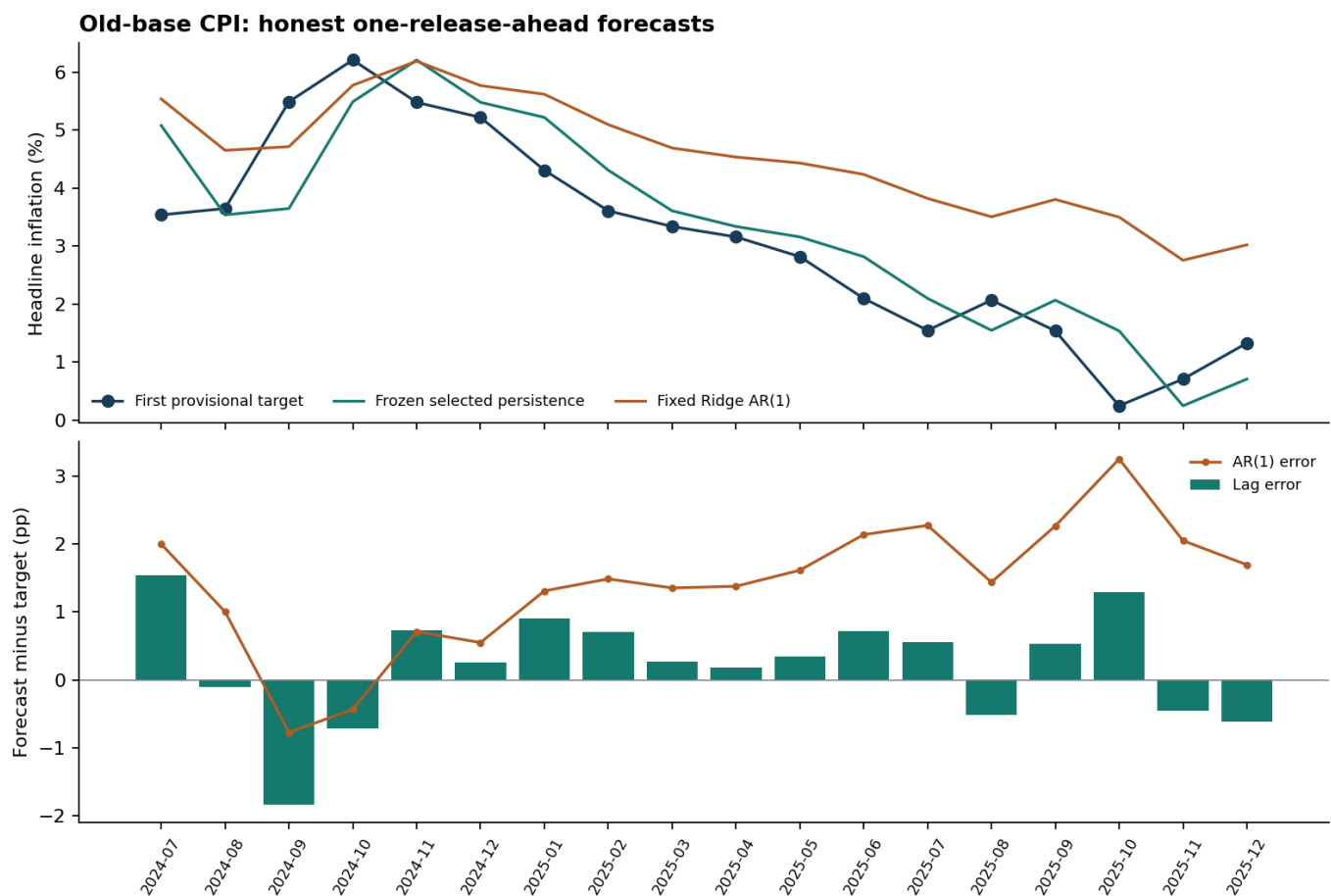
Development: six releases, January-June 2024. Persistence RMSE 0.2948, seasonal-12-month 0.8897, AR(1) 0.5969. The predeclared development rule selects persistence. Final: July 2024-December 2025, 18 releases. No test-based reselection, refit or recalibration occurs.

Audit contract

Eleven protected hashes seal protocol, source manifest, data, scientific code, fitted coefficients, development results and final predictions. The evaluation plan is fixed before first final scoring. A separate closed form reconciles the one-variable penalized regression. Five attached official PDFs reconcile with the HTML and were visually checked.

These checks establish source consistency and local reproducibility, not external independent review. They do not authenticate web content as it existed on each original release date.

2. Results and error mechanics



Arithmetic decomposition

AR(1) error = last-release lag error + adjustment away from that lag. Mean lag error is +0.2083 pp; mean model adjustment +1.1969; total mean error +1.4053. The fixed training environment was substantially above subsequent low observations, so shrinkage can move a forecast in the wrong direction.

MSE decomposition: lag 0.669972 + adjustment 1.948181 + cross term +0.238343 = 2.856496 pp squared. This is an exact identity, not causal inflation attribution.

No regime was chosen after seeing the test. The full 18-date prediction/source ledger, yearly descriptive slices and decomposition are in the accompanying JSON and Markdown paper.

3. Hindsight, revisions and the base break

Unavailable revised lag

For each final forecast, replace only the previous-month lag with the final value disclosed in the target release; hold the regression coefficients fixed. Four predictions change. Maximum absolute forecast change is 0.057558 pp. Contaminated AR(1) RMSE is 1.685431 versus honest 1.690117; MSE falls by 0.015820 pp squared. This small apparent improvement does not validate using unavailable information or rescue the challenger.

A different target is a separate question

Four first-provisional targets differ from their next-release final values. Persistence RMSE on those later-final targets is 0.817360 pp versus 0.818518 on the primary targets. Preserve both estimands. December 2025 final old-base 1.33% is independently traced to the pinned February 2026 source row as auxiliary evidence, outside the frozen 144-row panel.

Do not splice measurement systems

The 12 February 2026 release starts base 2024=100 and reports January headline 2.75%. The old-base December final 1.33% appears separately. Basket, weight and classification changes make a cosmetic rescaling inappropriate. Backcast 2025 new-base history first published in 2026 cannot enter a 2025 information set.

Stopping rule

The primary study ends in December 2025. A new-base study needs separately justified history and future evaluation. No current October 2026 forecast, monetary-policy effect, investment recommendation or ongoing automation is claimed.

Limits

Only 23 training pairs, six development releases and 18 dependent test dates; conditional bootstrap uncertainty; publication-dated reconstruction; no official archived-as-of attestation; no external independent replication. Simple forecasts can still have economically material errors.

4. Reproduction and source trail

Use pinned cached source bytes and pinned requirements. Restore data/raw_sources.tar.xz with scripts/restore_sources.py if required. Run the seven boundary tests, then scripts/evaluate_final.py and scripts/report_final.py. Do not overwrite the original freeze to make an edited model look predeclared.

Checked failure modes

Timezone-required timestamps; equality at publication; one-minute-late revision excluded; unavailable month/base rejected; conflicting values rejected; exact duplicates deduplicated; future data unable to alter earlier queries; source-cell equality; independent baseline/regression arithmetic; no training labels from the final period.

Evidence files

data/source_manifest.json: all 49 document URLs, hashes and retrieval dates. data/processed/vintages.json: 144 cell-level observations. reports/model_freeze.json: 11 scientific guards. reports/final.json: metrics, paired contrast, every final prediction and source. reports/error_decomposition.json: exact additive identities. reports/source_spot_checks.json: five PDF/HTML checks. reports/final_verification.json: evaluation assertions.

Primary sources

PIB / Ministry of Statistics and Programme Implementation monthly CPI releases. Publisher statistical content remains separately attributed; no broader redistribution license is asserted. Representative first source and final source are below; the complete ledger is supplied with the project.

<https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=1798312&lang=2®=3>

<https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=2213736&lang=2®=3>

<https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=2227012&lang=2®=3>

The accompanying two-page brief explicitly separates the reconstructed historical information set from the later outcome audit. It is prepared in October 2026 and is never backdated as a publication made in December 2025.