# x-forecast Methodology • v1.0

A paper portfolio framework for translating macro views into a verifiable allocation

x-forecast Research

2026-07-15 (DRAFT)

**Abstract**

x-forecast is a publicly maintained paper portfolio that operationalises a single monthly macro view into three parallel allocation tilts —Base / Bull / Bear — across 10 ETFs spanning A-shares, US equities, duration, credit, gold, and cash. This whitepaper documents the asset universe, rebalancing discipline, attribution framework, and known v1 limitations. The goal is auditability: every NAV number, weight, and view is reproducible from the public repository.

## Contents

# 1 1 • Motivation

Macro research is typically published as a chain of four artefacts: a *forecast* of key indicators, a *recommendation* on portfolio tilts, an *allocation* implementing the recommendation, and an *attribution* of the realised return back to the original view. Most published research stops at the second link. The third and fourth links —turning a view into a sized position and then auditing whether the view actually produced excess return —are usually opaque, either because the underlying portfolio is proprietary or because no live record exists at all.

x-forecast addresses this gap by maintaining a paper portfolio whose entire decision history is public, version-controlled, and reproducible. Every monthly allocation is committed to a public Git repository at the time of decision; daily NAV is recomputed from market data on a fixed schedule; monthly attribution decomposes excess return into asset-class allocation effects and five named macro factor exposures (§5). No real capital is at risk, no subscriptions are accepted, and no investment advice is provided.

The site complements two sister sites maintained by the same author: economic-forecasting.com publishes daily 1-month forecasts of seven US macro indicators; macro-forecast.com publishes monthly cross-asset allocation views. x-forecast closes the forecast → recommendation → allocation → attribution loop by binding the recommendation to an actual paper portfolio whose performance can be audited end-to-end.

*In practice, this means we treat the site as a public commitment device: every view we publish on the first trading day of the month is paired with explicit falsification conditions (§4) and a verdict at month-end (§5), so that both the framework and individual views can be falsified rather than merely debated.*

# 2 2 • Three-group structure

The portfolio runs three parallel groups, denoted Base, Bull, and Bear. The three groups share the same underlying asset universe (§3) and the same monthly rebalancing cadence (§4); they differ only in how a single set of macro views is mapped to weights.

**Base** is a risk-parity-neutral allocation with a light view overlay. Initial target volatility is 8% annualised, and the view overlay is constrained to move any single asset weight by at most 5 percentage points relative to the risk-parity reference. Base is the framework's default expression of "no view"; when views are weak or contradictory, Base

converges toward equal risk-contribution weights.

**Bull** overweights the two or three highest-conviction views relative to Base. Single-position cap is 30%; target volatility is 12% annualised. Bull is not a separate strategy —it is Base with the conviction dial turned up. If Bull persistently outperforms Base, the framework attributes excess return to view skill; if Bull persistently underperforms Base, the views being expressed are subtracting rather than adding value.

**Bear** is a defensive book designed for the regime in which Base is wrong. Bear is not a short bet and does not attempt to profit from market declines directly. Instead, Bear holds high cash, long duration, and gold, with target volatility $\leq 6\%$ annualised. Bear's role is structural: by publishing its NAV alongside Base and Bull, the framework forces a comparison of "what we would have held if our central case were wrong."

The simultaneous publication of three NAVs converts an otherwise private question —"do my macro views add value?"—into a public test. Two observations are diagnostic:

1. If Bull − Base is persistently positive, views are creating value relative to the framework's view-free baseline.
2. If all three groups underperform a 60/40 stock-bond benchmark over a long horizon, the framework itself —not any individual view —requires revision.

*In practice, this means we publish three weight vectors each month rather than one, and we accept that one of the three (Bear) is expected to underperform in most months; the value of Bear is realised only in the small number of months when Base is materially wrong, which is precisely the value of a defensive book.*

## 3 3 • Asset universe

The portfolio is constructed from ten exchange-traded funds spanning five core asset classes. The choice of ETFs as building blocks —rather than individual securities, futures, or derivatives —is deliberate: ETFs are liquid, free-of-charge to track, and the resulting NAV calculation is fully reproducible from public price data. Tracking error against the underlying asset class is treated as a known limitation rather than an object of optimisation (§9).

The asset classes and their proxy instruments are:

| Class | Instruments | Role |
|---|---|---|
| Equity (China) | CSI 300, CSI 500 | onshore beta + size factor |
| Equity (Hong Kong) | Hang Seng Index ETF | offshore China beta with USD funding |
| Equity (US) | SPY, QQQ | developed-market beta + tech factor |
| Rates (Duration) | China 10Y CGB ETF, TLT (US 20Y+) | duration carry + risk-off hedge |

| Class | Instruments | Role |
|---|---|---|
| Credit | China LGFV bond ETF | spread carry; offshore HY excluded in v1 |
| Commodities | Gold ETF | inflation hedge + USD hedge |
| Cash | Money-market proxy | nominal preservation + dry powder |

Five additional **reference variables** are tracked but not directly traded. These variables enter the framework as inputs to the macro view (§4) and as proxies for the five factor exposures used in attribution (§5):

- USD/CNY exchange rate
- DXY (US dollar broad index)
- 10-year and 2-year US Treasury yields (and the 10Y−2Y term-structure spread)
- VIX (S&P 500 implied volatility)

Reference variables are pulled from FRED (for US series) and Tushare (for onshore China series), with yfinance as a fallback for cases where FRED does not publish a particular instrument. The list of indicators and their source mappings is committed at `data/macro/_series.yaml` and is itself a versioned artefact.

The decision to exclude offshore credit, emerging-market debt, foreign currencies as direct holdings, and any single-stock or derivative exposure reflects an early framework choice: the v1 universe is deliberately small enough that every weight change has a traceable rationale, and broad enough that the framework can express any of the four macro regimes used in the quadrant overlay (§6) without resorting to derivatives.

*In practice, this means we rebalance ten ETFs each month, we look at five additional indicators when forming the view, and we accept that any exposure we cannot express through these fifteen tickers is —by construction —outside the framework. Future universe expansion (Japan equity, EM debt, crypto) is deferred to v3 (§9) and gated on demonstrating that the existing universe is fully exploited first.*

# 4 4 • Monthly discipline

Three operating rules govern the portfolio. They are written here in the declarative form they are intended to be applied:

**Rule 4.1 • Cadence.** The portfolio is rebalanced on the first trading day of each calendar month. Intra-month adjustments are prohibited unless one of two hard stops is triggered: the cash floor (5%) is breached, or a single-position cap (30%) is exceeded by drift. Both hard stops are deterministic and require no judgement at the moment of trigger.

**Rule 4.2 • Position constraints.** A single instrument may not exceed 30% of any group’s gross weight. A single asset class may not exceed 50%. Cash holds a floor

of 5%. Each rebalance may change at most three positions relative to the prior month —a deliberate constraint to prevent overfitting to recent data, in the spirit of Asness, Frazzini & Pedersen (2012)' s caution against high-turnover risk-parity overlays.

**Rule 4.3 • Documentation.** Each rebalance must be accompanied by a record in the decision log (§5; schema at `data/decisions/_schema.yaml`). The record commits, in advance, to a one-line decision summary, a 3 – 5 line rationale, a confidence level (1 – 5), an expected outcome path with explicit return ranges, and 3 – 6 falsification conditions. Each falsification condition must be an observable event with a numeric threshold; vague or non-falsifiable statements ( "if the view turns out wrong" ) are not admitted. The record is committed to the public Git repository at the time of decision, providing an immutable timestamp.

At month-end, the record is updated with the realised outcome, the verdict (hit / partial / miss / unknown), and a 1 – 3 line lessons paragraph. The verdict criteria are pre-specified: a *hit* requires at least 80% of the expected-path range to be captured with zero falsifications triggered; a *partial* corresponds to 50 – 80% range capture or one non-fatal falsification; anything below is recorded as *miss*. The verdict is computed mechanically from the documented criteria, not reasoned post hoc.

*In practice, this means we cannot change weights between rebalance dates except under the two hard stops, we cap any one decision at three weight changes, and we will not let a view enter the portfolio until its 3 – 6 falsification conditions are written down with thresholds. The monthly discipline is the principal mechanism by which view skill, framework, and luck are separable in the resulting NAV record.*

# 5 5 • Attribution framework

Excess return is decomposed along two largely orthogonal axes. The first axis follows Brinson, Hood & Beebower (1986) and decomposes monthly performance versus a benchmark into allocation and selection effects. The second axis projects the NAV return time series onto a set of five macro factor proxies via a rolling regression and isolates the residual as view-specific alpha.

## 5.1 5.1 Brinson decomposition

For each month $t$, group $g$ □ {Base, Bull, Bear}, and asset class $c$, the attribution computes:

- *Allocation effect*: (w_g,c − w_bench,c) × (R_bench,c − R_bench), the portion of excess return attributable to over- or under-weighting an asset class relative to the benchmark.
- *Selection effect*: w_g,c × (R_g,c − R_bench,c), the portion attributable to holding different securities within an asset class. Because the v1 universe is composed entirely of broad ETFs, the selection effect is structurally small and is reported chiefly for completeness.
- *Interaction*: (w_g,c − w_bench,c) × (R_g,c − R_bench,c), the residual cross term.

The benchmark is the 60/40 portfolio (60% global equity, 40% global duration) unless an alternative benchmark is explicitly chosen for a given month. Two secondary benchmarks —equal weight and pure risk-parity —are reported alongside to expose framework-vs-view contributions separately.

## 5.2 5.2 Macro factor regression

Five factor proxies are pre-specified for the regression. Each is a public, daily time series with a transparent mapping to a macroeconomic state variable:

| Factor (β) | Proxy series | State variable represented |
|---|---|---|
| Rates duration | 10Y UST yield (FRED DGS10) | risk-free rate level |
| USD direction | USD broad index (FRED DTWEXBGS) | USD strength vs trade partners |
| Credit spread | BofA HY OAS (FRED BAMLH0A0HYM2) | corporate credit conditions |
| Equity volatility | VIX (FRED VIXCLS) | risk appetite |
| Commodity | Bloomberg Commodity Index (BCOM) | global goods inflation pressure |

The five factors were chosen to be (a) economically distinct, (b) tradable in principle, and (c) historically uncorrelated enough to admit a meaningful joint regression. Daily group return r_g,t is regressed on daily changes in the five factor proxies over a rolling 60-day window using ordinary least squares:

$$r\_g,t = \alpha\_g,t + \beta\_rates \cdot \Delta rates\_t + \beta\_USD \cdot \Delta USD\_t + \beta\_credit \cdot \Delta credit\_t + \beta\_vol \cdot \Delta VIX\_t + \beta\_commo \cdot \Delta BCOM\_t + \varepsilon\_g,t$$

The five β coefficients are interpreted as the portfolio's *implicit* exposures, irrespective of how the position book was constructed. A discrepancy between the implicit β and the explicit (weight-implied) β is itself a diagnostic signal: it indicates that an asset is acting as a proxy for a factor not captured in its weight description.

## 5.3 5.3 Residual interpretation

The intercept term α_g,t and the regression residual ε_g,t are jointly the portion of return not explained by the five factor exposures over the window. For a paper portfolio targeted at validating macro views, the residual is the closest available measure of "view-specific" return —i.e. return that cannot be attributed to mechanical factor beta and is therefore plausibly traceable to the active view.

The residual is reported monthly but **not** interpreted as alpha in any inferential sense until the closed-decision sample size in the decision log (§4.3) exceeds six months. Below six closed decisions, no inferential statistic on residual return is published. This threshold is intentionally conservative: the cost of a false alpha claim in the public

record is materially higher than the cost of withholding a true claim that may later be confirmed.

*In practice, this means each month we publish three numbers per group —a Brinson decomposition versus 60/40, a 5-β factor decomposition with implicit exposures, and a residual α —and we make no causal claim about the residual until the decision log contains at least six closed records. The residual serves as a target for future inferential work; in v1 it is a disclosure artefact, not a performance claim.*

# 6 6 · Regime overlay

A four-quadrant regime classification is overlaid on the portfolio's monthly allocation process. The two axes are growth momentum and inflation momentum, each constructed as a standardised z-score across a curated set of underlying indicators (§6.1). The resulting position on the (growth, inflation) plane maps to one of four regimes, each with a stylised cross-asset preference borrowed from Dalio (1996):

| Quadrant | Conditions | Preferred exposures |
|---|---|---|
| Q1 | Growth ↑ · Inflation ↑ | Equity + commodities |
| Q2 | Growth ↑ · Inflation ↓ | Equity + long duration (the Goldilocks case) |
| Q3 | Growth ↓ · Inflation ↓ | Long duration + cash |
| Q4 | Growth ↓ · Inflation ↑ | Gold + cash (stagflation hedge) |

The framework borrows the architecture —but not the parameter values —from the Bridgewater All Weather framework. Hedgeye Risk Management's GIP model (Growth / Inflation / Policy) and BlackRock Investment Institute's quarterly tactical matrix are also closely related antecedents.

## 6.1 6.1 Construction of the two axes

Growth momentum is computed as the cross-sectional mean of three standardised series: year-over-year industrial production, year-over-year nonfarm payrolls, and the inverse of the unemployment rate. Each series is standardised over a five-year rolling window before averaging. Inflation momentum is constructed analogously from core CPI, core PCE, and the 10-year breakeven inflation rate, all sourced from FRED. China-specific analogues (CPI, PPI, GDP year-over-year) are pulled from Tushare and used as supplementary signals in months where the US and China cycles diverge.

The choice to standardise rather than use raw levels is deliberate: regime membership is a *relative* statement about the current state of the economy versus its own recent history, not an absolute judgement about whether growth is "high" or "low". A z-score of +1.0 means one standard deviation above the trailing five-year mean —not a comment on whether 4% growth is fast.

### 6.2 6.2 Tactical translation

The current quadrant produces a tactical tilt across five asset classes (equity, rates, credit, commodities, cash), each tagged as overweight (OW), neutral (N), or underweight (UW) over three horizons (6 months, 12 months, 24 months). The tilt is *not* a target weight; it is an input to the view-formation process for the next monthly rebalance. The Base group's view overlay is most directly informed by the 12-month tilt; the Bull group's overweights are drawn from the 6-month tilt where conviction is highest; the Bear group's defensive book is calibrated against the 24-month tilt for the opposite-of-central case.

*In practice, this means each month we mark a single point on the (growth, inflation) plane, read off the implied quadrant, and accept the tilt matrix as one input —not the only input —to the rebalance decision. The mechanical mapping prevents idiosyncratic narrative drift, while the discretionary view layer (§5) preserves the ability to override the mechanical signal when explicit reasons exist.*

## 7 7 • Stress testing

Five historical crisis episodes are replayed monthly through the current weight book to surface non-obvious vulnerabilities before they manifest in live performance. The scenarios are chosen to span distinct failure modes of a diversified cross-asset portfolio:

| Scenario | Period | Dominant failure mode |
|---|---|---|
| 2008 Global Financial Crisis | Sep 2008 –Mar 2009 | Equity drawdown + credit spread blowout |
| 2020 COVID liquidity crisis | Feb –Mar 2020 | Correlations spike to one; even gold sells off |
| 2015 CNY devaluation | Jun 2015 –Jan 2016 | Direct hit to A-share book; offshore decoupling |
| 2022 inflation shock | Jan –Oct 2022 | Stock-bond correlation turns positive (risk parity's death case) |
| 1998 LTCM / Russia default | Aug –Oct 1998 | Emerging-market and credit-spread tail event |

### 7.1 7.1 Methodology

For each scenario *s* and group *g*, the simulated return r_g,s is computed as the inner product of current weights w_g and historical local-currency total returns r_i,s for each instrument *i* over the scenario period:

$$r_{g,s} = \Sigma_i \, w_{g,i} \cdot r_{i,s}$$

The simulation makes three deliberate simplifications, each consistent with the v1 NAV methodology to keep the comparison meaningful (§9):

1. Local-currency returns are used; no USD/CNY conversion is applied.
2. No transaction cost is deducted; the simulation assumes the weights are held statically through the scenario without intra-scenario rebalancing.
3. Dividends and coupons are assumed automatically reinvested at the scenario start.

The output of the stress test is not a probabilistic claim about the portfolio's future drawdown. It is a *conditional* claim: if a scenario exactly resembling the historical episode were to recur from today's weight starting point, the indicated loss is the loss the v1 framework would have produced. The value of the exercise lies in identifying scenarios where any of the three groups (Base, Bull, Bear) breaches its risk budget — and triggering a documented framework revision when that occurs.

### 7.2 7.2 Interpretation rule

A scenario is flagged for framework revision if any of the following holds for the current month's weight book:

- Base's simulated loss exceeds Base's target volatility × 2 (i.e. Base loses more than two annualised standard deviations in the simulated episode).
- Bear's simulated loss exceeds 8% in any scenario (Bear is designed as defensive; a deeper loss indicates the defensive book is miscalibrated).
- Bull and Base diverge by less than 2 percentage points in any scenario (indicating Bull's conviction overlay is contributing negligible distinguishing exposure).

*In practice, this means each month we re-run five historical replays against the freshly committed weight book and accept that the v1 simplifications (no FX adjustment, no transaction cost) make the absolute numbers approximate. The signal we act on is structural: which group breaks first, in which scenario, and whether that pattern is consistent with the published view for the month.*

## 8 8 · Reproducibility

The framework treats reproducibility as a brand commitment rather than a nice-to-have property. Three operational rules give this commitment concrete form:

**Rule 8.1 · Public source.** The complete repository —Astro source, Python data scripts, configuration YAML, monthly rebalance records, and decision logs —is published to a public Git repository. No part of the NAV computation chain depends on private code, private data, or non-deterministic inputs.

**Rule 8.2 · One-line reconstruction.** Any reviewer with Node 18+ and Python 3.9+ installed can reproduce the entire NAV history from a fresh clone in three commands:

```
git clone <repo> && cd x-forecast
npm install && pip install -r scripts/requirements.txt
npm run data:all
```

The fetched price data, the rebalance YAML files, and the computed NAV JSON are written to deterministic file locations. The resulting NAV time series is identical to that served on the public site to four decimal places, modulo upstream data revisions in the same period.

**Rule 8.3 • Stable JSON API.** A small static JSON API (§8 of the research hub) exposes the current weight vector, the NAV time series, and the current macro view at predictable URLs. The v1 endpoints are frozen: fields may be added but never removed or renamed. Breaking changes are deferred to a parallel /api/v2/ namespace while v1 continues to serve.

*In practice, this means a reviewer can verify a claim made on the site —for example, that the Base group' s NAV on a specific date is some specific value —by re-running the data pipeline and comparing JSON outputs locally. The reproducibility kit is also designed to be runnable during an interview, where a screen-share of the running pipeline substitutes for several pages of slides.*

# 9 9 • Known limitations (v1)

The v1 implementation makes a number of simplifications that produce systematic deviations from the returns that would be realised by an actual portfolio implementing the same allocation. Each is disclosed here explicitly rather than discovered by readers from the data:

**9.1 • No transaction cost or slippage.** Real implementations of this strategy incur execution costs of approximately 10 – 20 basis points per year for the ETF universe described in §3, including bid-ask spread and broker commission. The v1 NAV is gross of these costs. The omission biases reported returns upward by the same magnitude.

**9.2 • No FX conversion.** Offshore assets (US equity, US Treasuries, gold ETF) are valued in their native currencies (USD), and the NAV is computed as a weighted sum of native-currency returns. A CNY-domiciled investor would experience returns adjusted by USD/CNY moves; the v1 methodology omits this adjustment. The bias is bi-directional and typically smaller than 200 basis points per year, but can be larger in quarters with sharp FX moves.

**9.3 • No reinvestment modelling.** Dividends and coupons are assumed to be automatically reinvested at the relevant ETF price. This matches the assumption commonly embedded in adjusted-close price series from yfinance and akshare but does not match cash flows of a real account.

**9.4 • ETFs as asset-class proxies.** Tracking error between the constituent ETFs and the asset classes they proxy is ignored. Over a one-year horizon, tracking error for the universe described in §3 is typically 30–80 basis points.

**9.5 • Sample size insufficient for inference.** The v1 dataset begins 12 months before inception. Inferential statistics on residual return (§5) require at least six closed monthly

decisions, and even at that sample size the statistical power is low. Any claim about "view alpha" prior to that threshold is withheld by convention.

**9.6 · No tax treatment.** All returns are pre-tax. A tax-resident investor in any jurisdiction would face additional cash drag.

The v2 roadmap, published at https://x-forecast.com/roadmap, addresses items 9.1, 9.2, 9.4, and 9.6 as fixed milestones with target ship dates.

*In practice, this means readers should subtract a conservative estimate of roughly 50 basis points per year from any reported NAV when comparing against benchmarks they themselves price gross of fees, and should treat the offshore portion of the portfolio as accounting in its native currency until v2's FX adjustment ships.*

# 10 10 · References


Asness, C., Frazzini, A. & Pedersen, L. H. (2012). Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1), 47–59.

Bridgewater Associates. (1996). *The all weather story*. Internal memorandum, reproduced in Dalio, R. (2020). *Principles for navigating big debt crises*. Bridgewater.

Brinson, G. P., Hood, L. R., & Beebower, G. L. (1986). Determinants of portfolio performance. *Financial Analysts Journal*, 42(4), 39–44.

Damodaran, A. (continuously updated). *Implied equity risk premium update*. NYU Stern School of Business. Retrieved from https://pages.stern.nyu.edu/~adamodar/

GMO LLC. (continuously updated). *Seven-year asset class real return forecasts*. Retrieved from https://www.gmo.com/

Hoffstein, C. (continuously updated). *Flirting with models* [Blog]. Newfound Research. Retrieved from https://blog.thinknewfound.com/

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.


---

## 10.1 Appendix A · Group weights (2026-06-01 inception)

待粘贴—见 /api/v1/weights.json.

## 10.2 Appendix B · Reproducibility hash

- Repo commit at publication: `<git rev-parse HEAD>`
- Python packages: `pip freeze > scripts/requirements.lock`
- Node packages: `npm shrinkwrap`
- Built with Astro `<version>`, Node `<version>`

## 10.3 Appendix C ・Reviewer credits

上线前的 reviewer 列表 (2 位金融博士 + 1 位 sell-side). 留空待填.

---

**Compilation:**

```
# Markdown → PDF (任选其一)
pandoc whitepaper/v1.0.md -o whitepaper/v1.0.pdf --pdf-engine=xelatex \
       -V mainfont="Times New Roman" -V CJKmainfont="SimSun" \
       --toc --number-sections

# 或用 mdbook / tectonic / typst, 同效果.
```

**Versioning rule**: 每次更新版本号 (v1.0 → v1.1), 旧版本保留在 `public/whitepaper/v1.0.md`, 不覆盖. 简历附录的链接永远指向最新版.