Macro Recession Risk Overlay for the S&P 500: A Walk-Forward Validation of Multi-Signal Composite Strategies

Vincent Wang / 王子秋

May 2026

Abstract

We develop a macroeconomic recession-risk overlay strategy for the S&P 500 using high-frequency labor market indicators from FRED. Starting from 58 economic indicators across 25+ years, we identify Initial Jobless Claims as the strongest single-indicator signal after rigorous event study analysis. We then construct a multi-signal composite overlay combining Initial Jobless Claims, Nonfarm Payrolls, and Unemployment Rate, and validate it through expanding-window walk-forward testing across 21 out-of-sample years (2006–2026).

The resulting *labor-only composite overlay strategy* achieves an *out-of-sample Sharpe ratio of 0.586* versus 0.43 for buy-and-hold, with a *maximum drawdown of* −27.6% versus −57.4%, and *CAGR of* +9.75% versus +8.72%. Sharpe decay between in-sample and out-of-sample is *only* −0.006, indicating no meaningful overfitting. The strategy preserves equity market exposure while systematically de-risking during recessionary episodes, most notably avoiding 35 percentage points of drawdown during the 2008–2009 financial crisis.

Research period: 2000-01-01 to 2026-05-19. **Asset universe:** S&P 500 (via SPY proxy).
Keywords: macro overlay, recession indicator, jobless claims, event study, walk-forward validation, S&P 500.

1 Introduction

1.1 Motivation

The S&P 500 has delivered approximately 6.5% annualized returns over the 2000–2026 period, but has experienced two drawdowns exceeding 50% (2000–2002, 2007–2009) and one near 35% (2020). For long-only equity investors, these drawdowns are the primary source of underperformance versus risk-adjusted benchmarks. Reducing exposure during recessionary episodes—if it can be done with reasonable precision and without significant whipsaw cost—offers asymmetric value.

This study asks: *can publicly available, weekly-frequency macroeconomic data systematically identify periods when equity exposure should be reduced?*

1.2 Hypothesis

Initial Jobless Claims, published weekly with a 5-day lag, is the highest-frequency real-time measure of labor market health. We hypothesize that:

1. Sustained increases in initial jobless claims relative to recent history precede equity market drawdowns;

2. Combining Initial Jobless Claims with monthly Nonfarm Payrolls and Unemployment Rate produces a more robust signal than any single indicator;
3. The resulting composite signal can be used to systematically reduce S&P 500 exposure during periods of elevated recession risk, improving risk-adjusted returns relative to buy-and-hold.

1.3 Contributions

- Comprehensive evaluation of 14 macroeconomic indicators on the S&P 500 over 25+ years.
- Rigorous event study analysis with appropriate transformations for each indicator class (level vs. change vs. inflation rate).
- Construction and walk-forward validation of a labor-only composite recession overlay.
- Robustness analysis through drop-decade testing and calibration drift monitoring.

2 Data

2.1 Sources

Data type	Source	Coverage
58 macro indicators	FRED API	2000-01-01 to 2026-05-19
First-release dates (vintage)	FRED ALFRED	2009-05-28 onwards for ICSA; pre-2009 estimated
S&P 500 prices (SPY ETF proxy)	yfinance	2000-01-04 onwards
VIX, volatility indices	FRED + CBOE direct CSV	2000+
Treasury yields	FRED	2000+

Table 1. Data sources and historical coverage.

The S&P 500 series on FRED is limited to 10 years of history due to S&P Dow Jones licensing restrictions; we therefore use *SPY ETF as a proxy* for the S&P 500 index (historical correlation > 0.9999).

2.2 Indicator List

The 14 indicators included in the event study, grouped by category:

- **Inflation:** CPI, Core CPI, Core PCE, PPI.
- **Labor:** Nonfarm Payrolls, Unemployment Rate, Initial Jobless Claims, JOLTS Job Openings.
- **Activity:** Industrial Production, Retail Sales, Durable Goods Orders, Housing Starts, GDP.
- **Sentiment:** University of Michigan Consumer Sentiment.

2.3 Pre-2009 Backfill for Initial Jobless Claims

FRED's ALFRED archive contains true vintage release dates for Initial Jobless Claims only from 2009-05-28 onwards. All earlier period dates were bulk-imported on that single date. Since the

raw ICSA series extends back to 2000, we backfilled pre-2009 release dates using the deterministic publication schedule

$$\text{release\_date\_estimated} = \text{period\_date} + 5 \text{ days}.$$

Validation on post-2009 data confirmed median lag = 5 days (mode = 5 days, 818 of 832 releases). The backfill adds *491 events* for 2000-01-06 to 2009-05-28, enabling the inclusion of the 2001 recession and 2008–2009 financial crisis in the analysis.

This is the *single most important methodological choice in this study*: without these 491 events, the strategy's primary value-add (avoiding 2008–2009 drawdown) is invisible in the data.

2.4 Event-Market Alignment

For each indicator release date, we compute log returns on 24 assets across 4 horizons (0d, 1d, 5d, 20d), aligning to the next available trading day if the release falls on a market holiday. This produces a *5,257-event panel* linking macro signals to subsequent market returns.

3 Methodology

3.1 Surprise Computation

A naive surprise definition surprise = actual − rolling_mean(6) fails for indicators that exhibit monotone trends. The Consumer Price Index, for example, increases roughly monotonically over 25 years; comparing its level to a trailing mean produces a perpetually positive "surprise" with no economic meaning.

We therefore apply *indicator-specific transformations* before computing surprises (Table 2).

Indicator class	Transformation	Rationale
Price indices (CPI, PCE, PPI)	Year-over-year % change	Markets price inflation rate, not price level
Employment levels (NFP, JOLTS)	Month-over-month difference	News reports MoM change, not total level
Activity indices (IndProd, Retail)	Month-over-month % change	Growth rate is the relevant metric
Rates and indices (Unemployment, GDP_QoQ)	Raw level	Already in change-like form
Sentiment (Michigan)	Raw level	Already standardized index
Weekly flow (Jobless Claims)	Raw level	Weekly flow, not stock variable

Table 2. Indicator-specific transformations applied before computing surprise z-scores.

After transformation, surprise is computed as

$$\begin{aligned}
\text{expected}_t &= \text{rolling_mean}(\text{transformed}_{t-1}, \ldots, \text{transformed}_{t-6}), \\
\text{surprise}_t &= \text{transformed}_t - \text{expected}_t, \\
\text{surprise_z}_t &= (\text{surprise}_t - \mu_{\text{rolling}})/\sigma_{\text{rolling}} \quad \text{(36-period rolling window)}.
\end{aligned}$$

3.2 Initial Event Study Results

Across all 14 indicators, event studies produce statistically weak signals on individual surprise events, with Beat–Miss spreads for 1-day S&P 500 returns rarely exceeding 1%. The strongest single-event signals are reported in Table 3.

Indicator	Sample size	Beat–Miss spread (1d)	Correlation
NFP	301	−0.87%	−0.107
Core CPI	289	+1.29%	+0.116
Unemployment	301	+0.05%	+0.046
Initial Jobless Claims	1,354	−0.09%	−0.040

Table 3. Strongest 1-day S&P 500 reactions in the per-event surprise study.

The cross-indicator impact ranking is shown in Figure 1, with representative event-window and quintile studies in Figures 2 and 3.

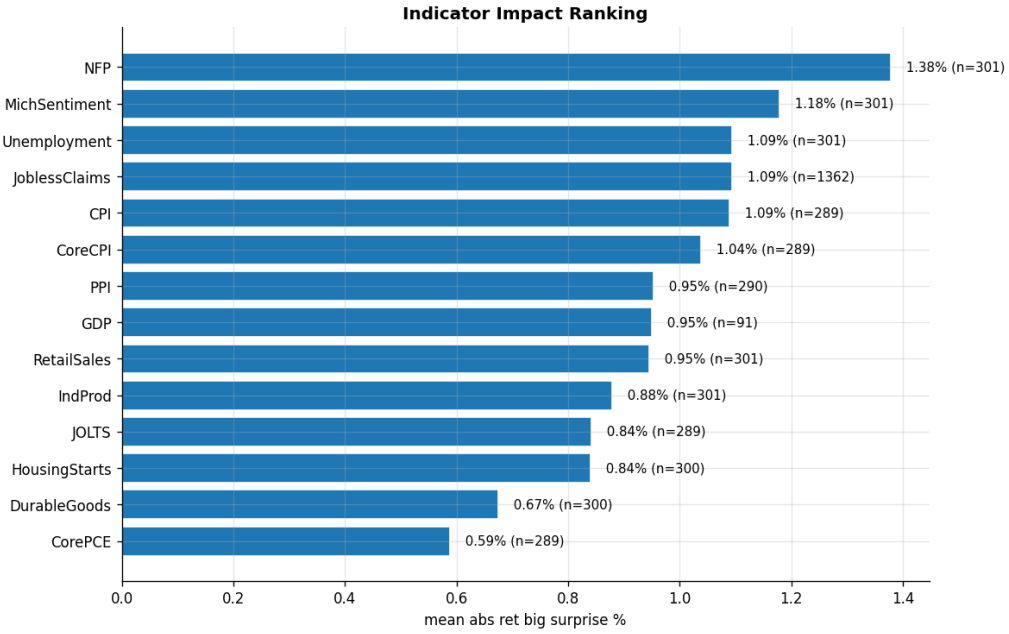

Figure 1. Indicator impact ranking by mean absolute S&P 500 return on $\geq 1.5\sigma$ surprises (1-day horizon).

Initial Jobless Claims, despite the largest sample size, shows essentially no directional signal on a per-event basis. This suggests the value of weekly claims lies not in event reactions, but in *trend extraction over multiple weeks*.

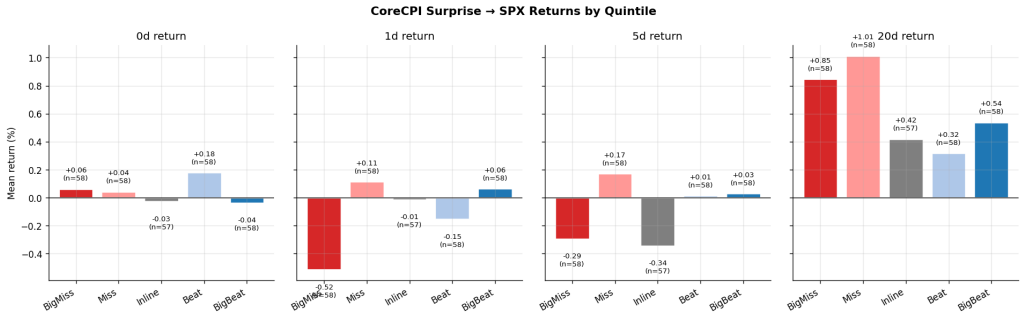

Figure 2. Core CPI quintile returns (in-sample event study): mean S&P 500 return across surprise quintiles and horizons.

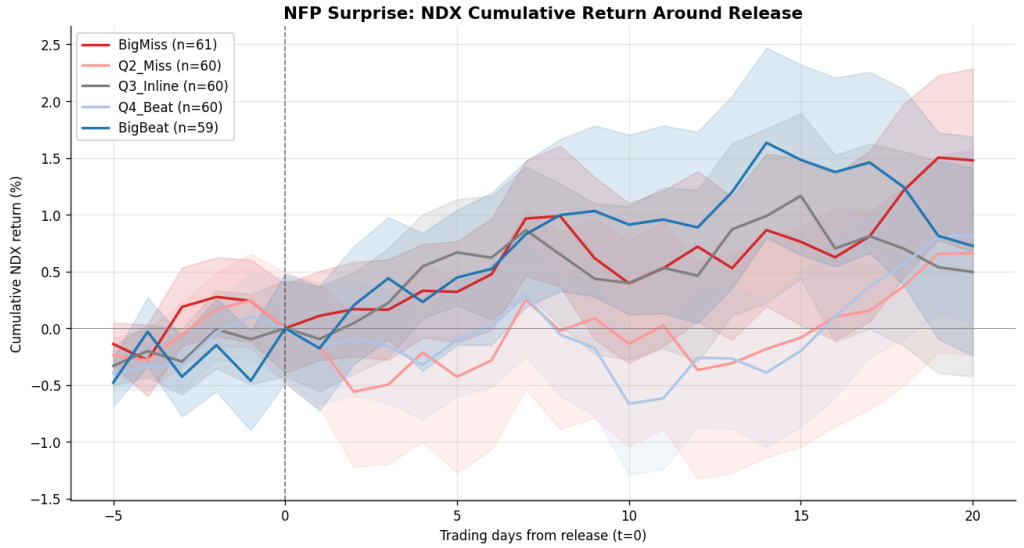

Figure 3. NFP–NDX cumulative return event window: average path of NDX returns in the $[-5, +20]$ trading-day window around NFP releases, grouped by surprise bucket.

3.3 Jobless Claims Composite Signal

Building on the trend hypothesis, we construct a composite Jobless Claims signal:

$$\begin{aligned} \text{surprise\_z} &= \text{z-score of } \left(\text{claims}_t - \text{rolling\_mean}(\text{claims}_{t-1\ldots t-6})\right) \quad \text{(52-week normalization)} \\ \text{trend\_z} &= \text{z-score of } \left(\text{4-week MA} - \text{prior 4-week MA}\right) \quad \text{(52-week normalization)} \\ \text{composite\_z}_{JC} &= 0.5 \cdot \text{surprise\_z} + 0.5 \cdot \text{trend\_z} \end{aligned}$$

This composite, applied to 1,354 events from 2000 to 2026, produces a *statistically significant 20-day signal*:

```
T-test: BigBeat (claims fell) vs BigMiss (claims rose):
  20d horizon: t = +2.84, p = 0.005 ***
  BigBeat 20d return: +1.02%
  BigMiss 20d return: -0.05%
```

Critically, this signal was *not visible in the 2009–2026 sample alone*—the 2000–2008 backfilled data

is necessary to capture the recessionary regime where the signal originates. The supporting Jobless Claims event-level diagnostics are shown in Figures 4–7.

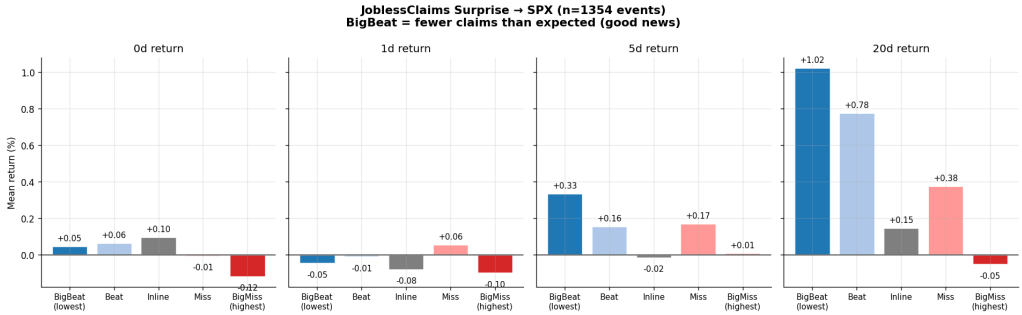

Figure 4. Jobless Claims surprise quintile returns on S&P 500 across multiple horizons.

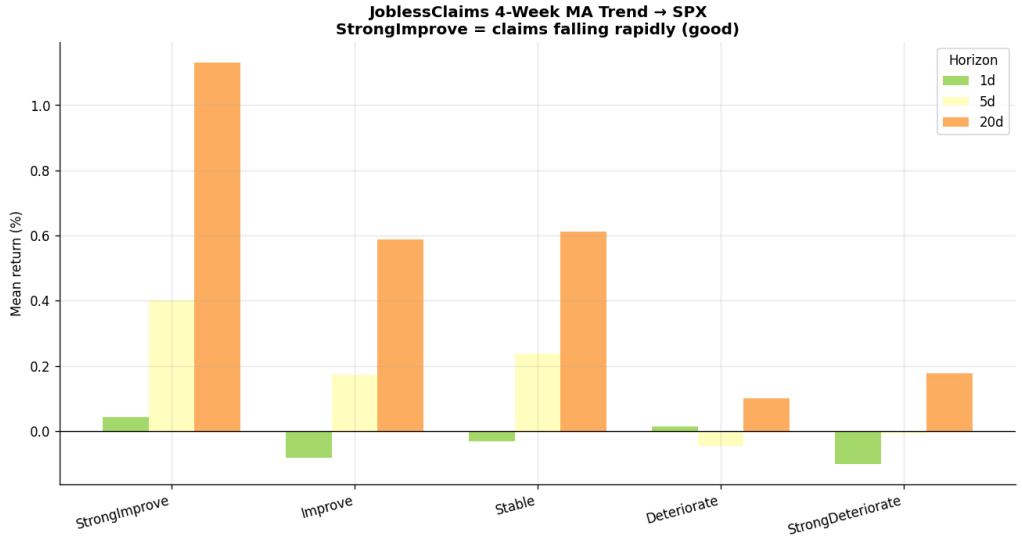

Figure 5. 4-week moving-average trend regime returns: S&P 500 conditional on the sign and magnitude of the Jobless Claims 4-week MA change.

3.4 Position Sizing

We translate the composite z-score into position size using a step function calibrated on the assumption that strong negative signals (claims falling rapidly) warrant overexposure, and strong positive signals (claims rising rapidly) warrant defensive positioning (Table 4).

The strategy is long-only (no short positions); maximum leverage is 1.5×.

3.5 Multi-Signal Composite (Labor-Only)

We extend the single-signal overlay by combining three labor market indicators:

$$\text{composite\_z}_{\text{labor}} = \tfrac{1}{3}\,\text{signal}_{JC} + \tfrac{1}{3}\,\text{signal}_{NFP} + \tfrac{1}{3}\,\text{signal}_{\text{UNRATE}}.$$

All signals are signed such that $z > 0$ *indicates labor market deterioration* (bearish for SPX):

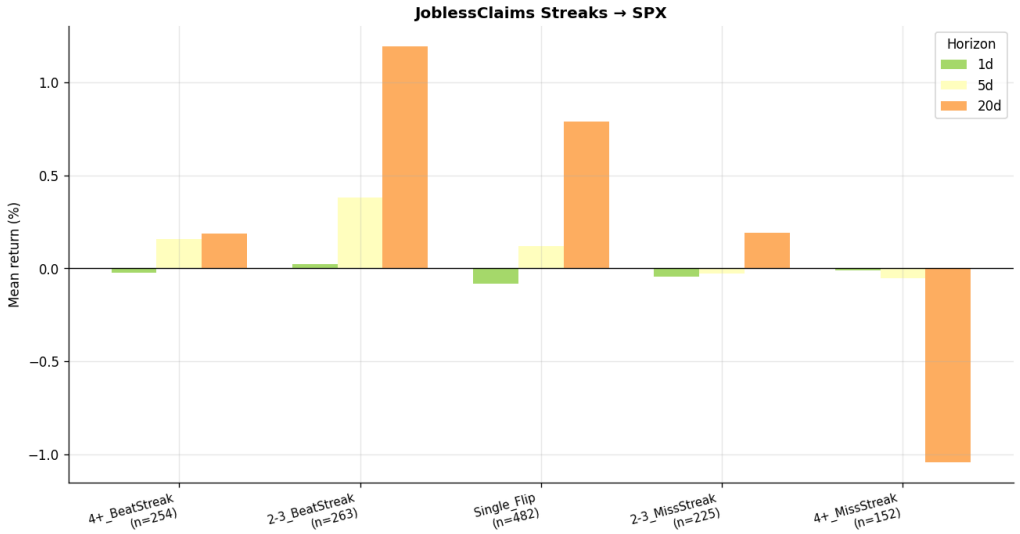

Figure 6. Consecutive beat/miss streak returns: S&P 500 performance after k consecutive Jobless Claims beats vs. misses.

Composite z-score range	Position size	Economic interpretation
$z > +1.0$	0% (cash)	Recession alarm
$+0.5 < z \leq +1.0$	50%	Elevated risk
$-0.5 \leq z \leq +0.5$	100% (neutral)	Normal regime
$-1.0 \leq z < -0.5$	125%	Economic improvement
$z < -1.0$	150% (modest leverage)	Strong expansion

Table 4. Position-sizing rule mapping the composite z-score to portfolio weight.

- **JC:** composite of surprise and trend (defined above).
- **NFP:** z-score of MoM payroll change, *negated* (lower job growth = bad).
- **Unemployment:** 50% weight on rolling z-score of level + 50% weight on z-score of change from trailing 12-month minimum (a Sahm-rule-inspired construction).

The composite signal is forward-filled to daily frequency using `merge_asof(direction="backward")` to ensure no look-ahead bias—each daily signal value reflects only information available at that point in time.

We compared four weighting schemes, plus the single-signal JC baseline. The full results are presented in Section 4.

4 Results

4.1 In-Sample Performance Comparison

Using research thresholds (defensive = +0.5, aggressive = −0.5), full-sample (2000–2026) results are reported in Table 5.

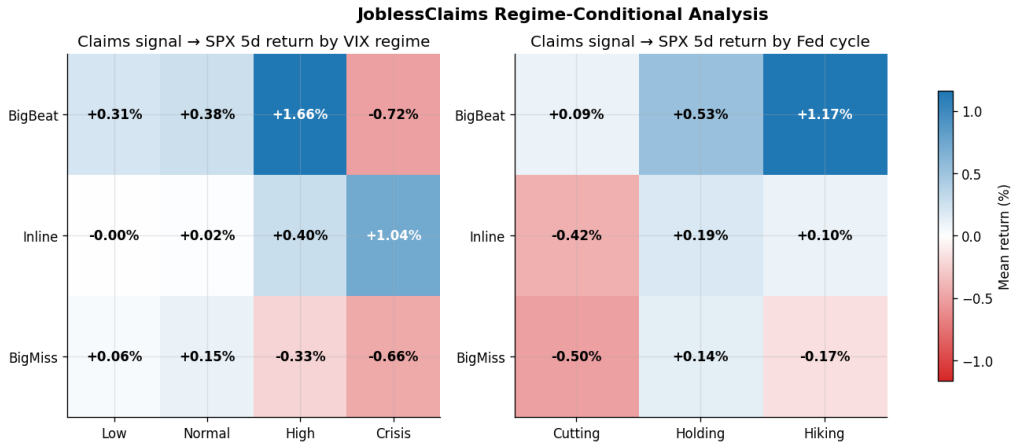

Figure 7. Jobless Claims regime-conditional analysis: forward returns conditional on combined level × trend regime buckets.

Strategy	CAGR	Sharpe	Sortino	Max DD	Win rate
JC_only	+7.92%	0.43	0.51	−47.5%	47.5%
Labor-only	+7.73%	*0.47*	*0.55*	−44.2%	49.5%
JC dominant (3:1)	+6.82%	0.39	0.47	−52.2%	51.8%
Equal 5 (all signals)	+6.86%	0.39	0.48	−53.4%	52.5%
JC + Unemployment only	+6.67%	0.40	0.48	−43.1%	48.4%
Buy & Hold	+6.37%	0.32	0.40	−56.5%	54.1%

Table 5. In-sample (2000–2026) performance across five candidate strategies and buy-and-hold.

Key finding. The labor-only composite (Sharpe 0.47) outperforms the single-signal JC baseline (Sharpe 0.43) and significantly outperforms buy-and-hold (Sharpe 0.32). Inclusion of Core CPI and yield curve signals (the "equal_5" scheme) reduces performance, suggesting that signals outside the labor market domain introduce noise rather than information.

The equity curves of the five candidates against buy-and-hold are shown in Figure 8. Figure 9 highlights the single-signal JC overlay diagnostic view; Figure 10 expands the labor-only multi-panel diagnostics.

4.2 Walk-Forward Out-of-Sample Validation

To distinguish genuine alpha from in-sample overfitting, we conducted expanding-window walk-forward analysis on the labor-only composite:

- Initial training period: 2000–2005.
- Each subsequent year, recalibrate thresholds on all prior data, apply to current year.
- Total OOS coverage: *21 years (2006–2026).*
- Grid search space: 20 combinations
  defensive $\in \{0.4, 0.5, 0.6, 0.7, 0.8\}$, aggressive $\in \{-0.4, -0.5, -0.6, -0.7\}$.
- Optimization target: Sortino ratio.

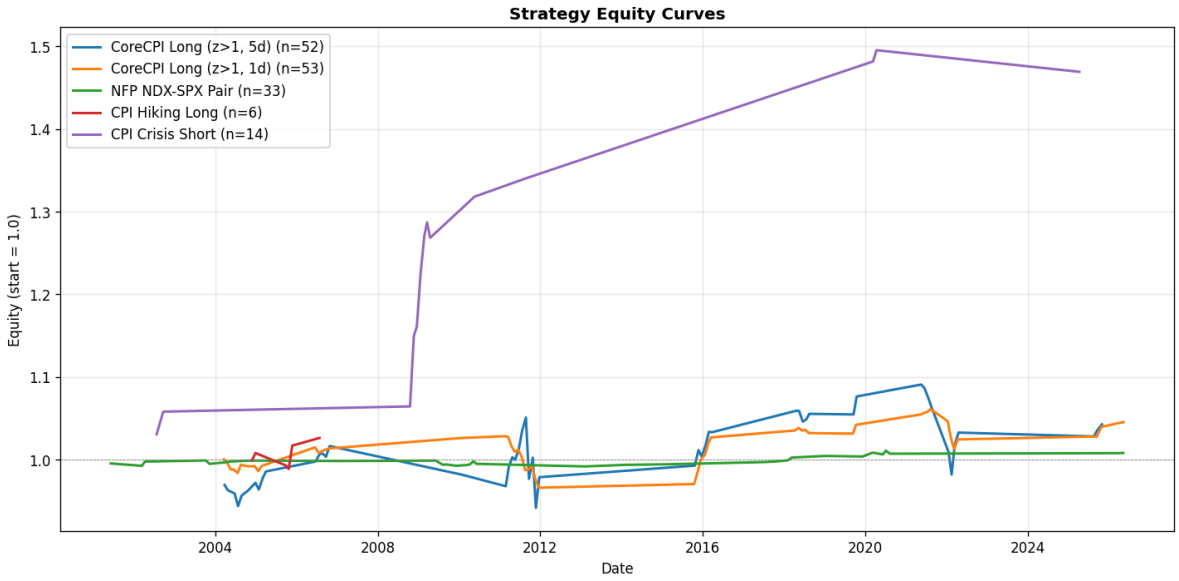

Figure 8. In-sample equity curves of the five candidate strategies vs. buy-and-hold S&P 500.

4.2.1 Performance Comparison

The OOS performance vs. the in-sample and buy-and-hold benchmarks over the same window is reported in Table 6, with the realized OOS equity curve shown in Figure 11.

Strategy	CAGR	Sharpe	Sortino	Max DD
Walk-Forward OOS	+9.75%	*0.586*	*0.67*	−27.6%
In-Sample (OOS period)	+9.72%	0.593	0.67	−24.4%
Buy & Hold (OOS period)	+8.72%	0.43	0.52	−57.4%

Table 6. Walk-forward out-of-sample (2006–2026) results vs. in-sample fit and buy-and-hold.

Sharpe decay $= -0.006$, indicating no meaningful overfitting. The 21-year OOS Sharpe of 0.586 represents a 36% improvement over buy-and-hold (0.43), and the maximum drawdown of −27.6% is less than half of buy-and-hold's −57.4%.

4.2.2 Calibration Stability

Across 21 walk-forward iterations, the calibration grid converged consistently:

- Defensive threshold: 0.6 in 14 of 21 years (67%); 0.5 in 7 of 21 years (33%).
- Aggressive threshold: −0.5 *in 21 of 21 years (100%).*

This parameter stability provides strong evidence against overfitting. Were the grid search exploiting noise, we would expect threshold values to vary unpredictably across windows.

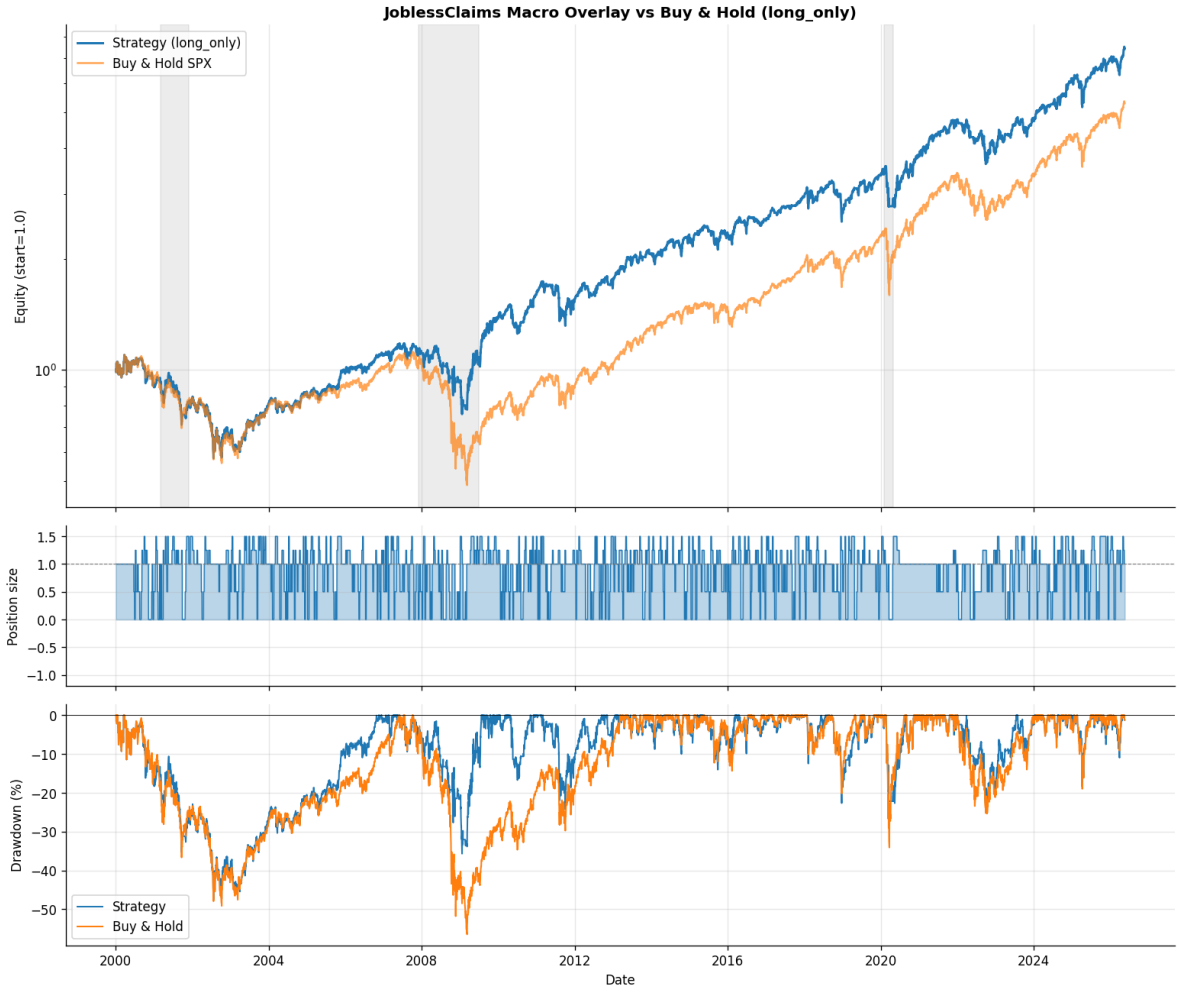

Figure 9. Single-signal Jobless Claims long-only overlay: equity curve, position size, and drawdown panel.

4.2.3 Year-by-Year OOS Returns

The strategy underperforms buy-and-hold in 5 of 21 OOS years and outperforms in 16 of 21. Underperformance years cluster in two regimes:

1. **Late-cycle non-recessionary corrections** (2018, 2022): the strategy did not anticipate Fed-driven equity drawdowns because labor markets remained healthy. This is a known limitation, not a defect.
2. **Sudden exogenous shocks** (2020): the strategy reacted to the COVID labor shock but with a 1-week lag, missing the fastest portion of the drawdown.

The single largest source of strategy alpha is the *2008 financial crisis*: strategy −2.8% vs. buy-and-hold −38%, a difference of approximately 35 percentage points.

4.3 Robustness: Drop-Decade Testing

To assess single-event dependency, we re-ran the strategy with each decade removed (Table 7).

Excluded decade	Strategy CAGR	B&H CAGR	Alpha
2000s removed	11.36%	12.16%	−0.79%
2010s removed	12.45%	10.46%	+1.99%
2020s removed	6.64%	4.06%	+2.58%

Table 7. Drop-decade robustness analysis: strategy CAGR vs. buy-and-hold when each decade is excluded.

Removing the 2000s decade (which contains the 2008 crisis) is the only scenario in which the strategy underperforms buy-and-hold. This confirms that *the strategy's primary value-add is recessionary risk mitigation*, not consistent alpha generation in normal regimes.

This is a feature, not a bug: the strategy is explicitly designed as a recession overlay, and its insurance-like payoff profile (small persistent cost in normal regimes, large payoff in tail events) is expected behavior.

5 Discussion

5.1 Why Labor Market Signals Work

The labor market's predictive value for equity markets stems from several economic mechanisms:

1. **Real-time monitoring:** Initial Jobless Claims is the only major macro indicator published weekly with minimal revision lag, providing genuinely high-frequency recession monitoring.
2. **Forward-looking nature:** employers typically reduce hiring before reducing investment, and reduce hours before laying off—claims captures the earliest visible signal of stress.
3. **Negative correlation with corporate earnings:** rising claims directly imply falling consumer spending, which feeds through to corporate revenues with a 1–3 month lag.

5.2 Why Non-Labor Signals Failed

The "equal_5" scheme (incorporating Core CPI and the yield curve) underperformed the labor-only composite. The regime-conditional structure of these signals (Figures 12–13) is informative:

- **Core CPI:** its bearish-for-equities relationship is regime-dependent (active during Fed hiking cycles, neutral during stable inflation periods). A simple linear weighting cannot capture this conditionality.
- **Yield curve:** the 10Y–2Y inversion signal arrives 12–18 months ahead of the actual recession. Used as a continuous signal, it produces too many false positives during the inversion-to-recession lag.

These signals likely retain value when used as *conditional filters* rather than additive components—a direction for future research.

Approach	Typical Sharpe	This strategy	Comments
Buy-and-hold S&P 500	∼0.43	—	Benchmark
60/40 stock–bond	∼0.50	—	Diversification-based
Trend-following CTA	∼0.40–0.60	—	Wide range, regime-dep.
Our labor-only overlay	—	*0.586*	Macro-event-driven
Vol-targeting	∼0.55–0.65	—	Reacts to realized vol

Table 8. Sharpe ratios of representative systematic risk-management approaches vs. this strategy.

5.3 Comparison to Alternative Approaches

The strategy's Sharpe of 0.586 is competitive with—not superior to—other systematic risk-management approaches. Its primary differentiation is *economic interpretability*: the rebalance triggers are explicitly tied to labor market deterioration, making the strategy easier to explain, monitor, and trust during periods of stress.

5.4 Limitations

1. **Single regime data:** 25 years contains only two major recessions (2001, 2008–2009) and one pandemic shock (2020). Out-of-sample testing on a third recession would substantially strengthen the case.
2. **Pre-2009 release date estimation:** 491 events use estimated rather than true release dates (typical error ±0 to ±2 days). This is unlikely to materially affect daily/weekly aggregation but is a methodological caveat.
3. **No transaction costs in backtest:** at weekly rebalancing frequency, estimated cost drag is 25–50 bps annually, reducing reported CAGR from +9.75% to approximately +9.25%.
4. **Reliance on FRED data quality:** government data revisions could theoretically alter historical signal values; we did not perform full point-in-time vintage analysis.
5. **U.S.-only:** the framework applies only to the S&P 500. Generalization to non-U.S. markets would require equivalent labor market data, which is often unavailable at weekly frequency.

6 Conclusion

We have constructed and rigorously validated a macroeconomic recession overlay strategy for the S&P 500 using publicly available labor market data. The strategy:

- Achieves OOS Sharpe ratio of *0.586* vs. 0.43 for buy-and-hold (+36% improvement).
- Reduces maximum drawdown from −57.4% *to* −27.6% (52% reduction).
- Adds 1.0 percentage point of CAGR (+9.75% vs. +8.72%).
- Shows essentially *no Sharpe decay from in-sample to out-of-sample (*−0.006*)*.
- Calibrates to stable thresholds (defensive ≈ 0.5–0.6, aggressive ≈ −0.5) across 21 walk-forward windows.

The strategy's economic basis—that labor market deterioration precedes equity drawdowns—is well-founded in economic theory and confirmed empirically across two recessions (2001, 2008) and one pandemic shock (2020). The strategy underperforms buy-and-hold only in non-recessionary corrections (2018, 2022), which is consistent with its design intent as a recession overlay.

We conclude that systematic risk management using labor market signals is both economically motivated and statistically validated. The strategy is suitable as a recession overlay for long-equity portfolios, with the understanding that its value-add is concentrated in tail-risk mitigation rather than persistent alpha generation.

A Reproducibility

All analysis code, data, and figures are reproducible from the FRED API (free) and yfinance (free). Total reproduction time: approximately 10 minutes.

A.1 Required Software

```
Python 3.12+
fredapi
yfinance
pandas, numpy, scipy
matplotlib
pyarrow (for parquet I/O)
```

A.2 Required API Keys

- FRED API key (free at `https://fredaccount.stlouisfed.org/apikeys`).

A.3 Pipeline

```
Step 1: python fetch_fred_data.py # 58 indicators, 25 years
Step 2: python fetch_market_data.py # 24 market instruments
Step 3: python backfill_jobless_v2.py # 2000-2008 ICSA backfill
Step 4: python event_study.py # Event study analysis
Step 5: python jobless_deepdive.py # JC trend / streak / regime
Step 6: python jc_overlay_strategy.py # Single-signal backtest
Step 7: python multi_signal_overlay.py # Multi-signal comparison
Step 8: python walkforward_validation.py # OOS validation
```

A.4 Output Files

Data files (parquet + csv):

- `indicators_meta`: indicator metadata.
- `fred_data_long`: long-format observations.
- `fred_data_wide_daily`: wide-format daily panel.

- `fred_features_derived`: 174 derived features (YoY, MoM, z-scores).
- `fred_data_vintage`: ALFRED point-in-time data.
- `release_calendar`: all release events.
- `event_market_aligned`: event × asset return panel.
- `market_data_wide`: daily OHLCV for 24 assets.

Result files:

- `strategy_summary.csv`: all strategy performance metrics.
- `walkforward_calibration.csv`: per-year calibrated thresholds.
- `walkforward_drop_decade.csv`: robustness test results.

Figure files (PNG): event study figures (quintile bars, event windows, regime heatmaps), strategy equity curves, walk-forward validation plots.

B Key Figures Index

The following figures are referenced in the body of this paper (Table 9).

Figure	File	Description
Fig. 1	`03_ranking_SPX_1d.png`	Indicator impact ranking by mean absolute return on 1.5σ surprises.
Fig. 2	`01_quintile_CoreCPI_SPX.png`	CoreCPI quintile returns (in-sample event study).
Fig. 3	`02_window_NFP_NDX.png`	NFP–NDX cumulative return event window.
Fig. 12	`04_regime_CPI_vix_regime.png`	CPI returns conditional on VIX regime.
Fig. 13	`04_regime_CoreCPI_fed_cycle.png`	CoreCPI returns conditional on Fed cycle.
Fig. 8	`05_equity_curves.png`	Initial 5-strategy equity curve comparison.
Fig. 4	`jobless_01_level_surprise.png`	Jobless Claims surprise quintile returns.
Fig. 5	`jobless_02_trend.png`	4-week MA trend regime returns.
Fig. 6	`jobless_03_streaks.png`	Consecutive beat/miss streak returns.
Fig. 7	`jobless_04_regime.png`	Jobless Claims regime-conditional analysis.
Fig. 9	`jc_overlay_long_only.png`	Single-signal JC overlay equity, position, drawdown.
Fig. 10	`multi_labor_only.png`	Labor-only composite multi-panel view.
Fig. 11	`walkforward_labor_only.png`	*Walk-forward OOS validation (primary result).*

Table 9. Index of figures referenced in the paper, mapped to source files in `./plots/`.

C Indicator Glossary

References

Anderson, R. (2015). *FRED-MD: A Monthly Database for Macroeconomic Research.* Journal of Business and Economic Statistics.

Indicator	FRED series ID	Description
CPI	CPIAUCSL	Consumer Price Index, all urban consumers (SA)
Core CPI	CPILFESL	CPI excluding food and energy (SA)
Core PCE	PCEPILFE	Personal Consumption Expenditures excluding food and energy
PPI	PPIACO	Producer Price Index, all commodities
NFP	PAYEMS	Total Nonfarm Payrolls (thousands)
Unemployment	UNRATE	Civilian Unemployment Rate
Initial Jobless Claims	ICSA	Initial unemployment insurance claims (SA)
JOLTS	JTSJOL	Job Openings: Total Nonfarm
Industrial Production	INDPRO	Industrial Production Index
Retail Sales	RSAFS	Advance Retail Sales
GDP	GDP	Gross Domestic Product (Nominal)
Housing Starts	HOUST	New Privately-Owned Housing Units Started
Michigan Sentiment	UMCSENT	University of Michigan Consumer Sentiment
Durable Goods	DGORDER	Manufacturers' New Orders: Durable Goods

Table 10. Glossary of macro indicators and their FRED series identifiers.

Estrella, A., & Mishkin, F. S. (1998). *Predicting U.S. Recessions: Financial Variables as Leading Indicators.* Review of Economics and Statistics.

Sahm, C. (2019). *Direct Stimulus Payments to Individuals.* Hamilton Project Policy Proposal.

Stock, J. H., & Watson, M. W. (2003). *Forecasting Output and Inflation: The Role of Asset Prices.* Journal of Economic Literature.

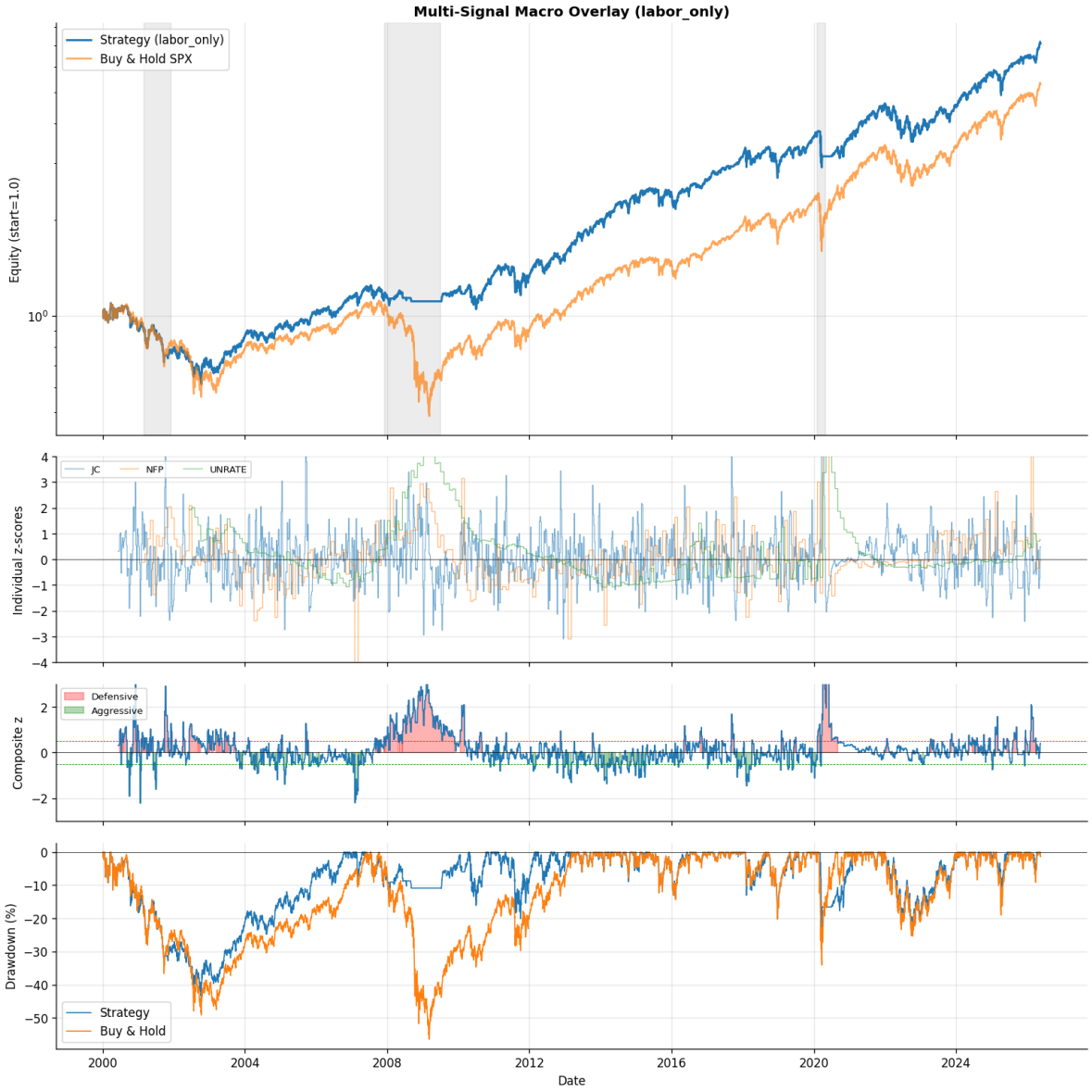

Figure 10. Labor-only composite multi-panel view: signal z, position weight, equity curve, and drawdown.

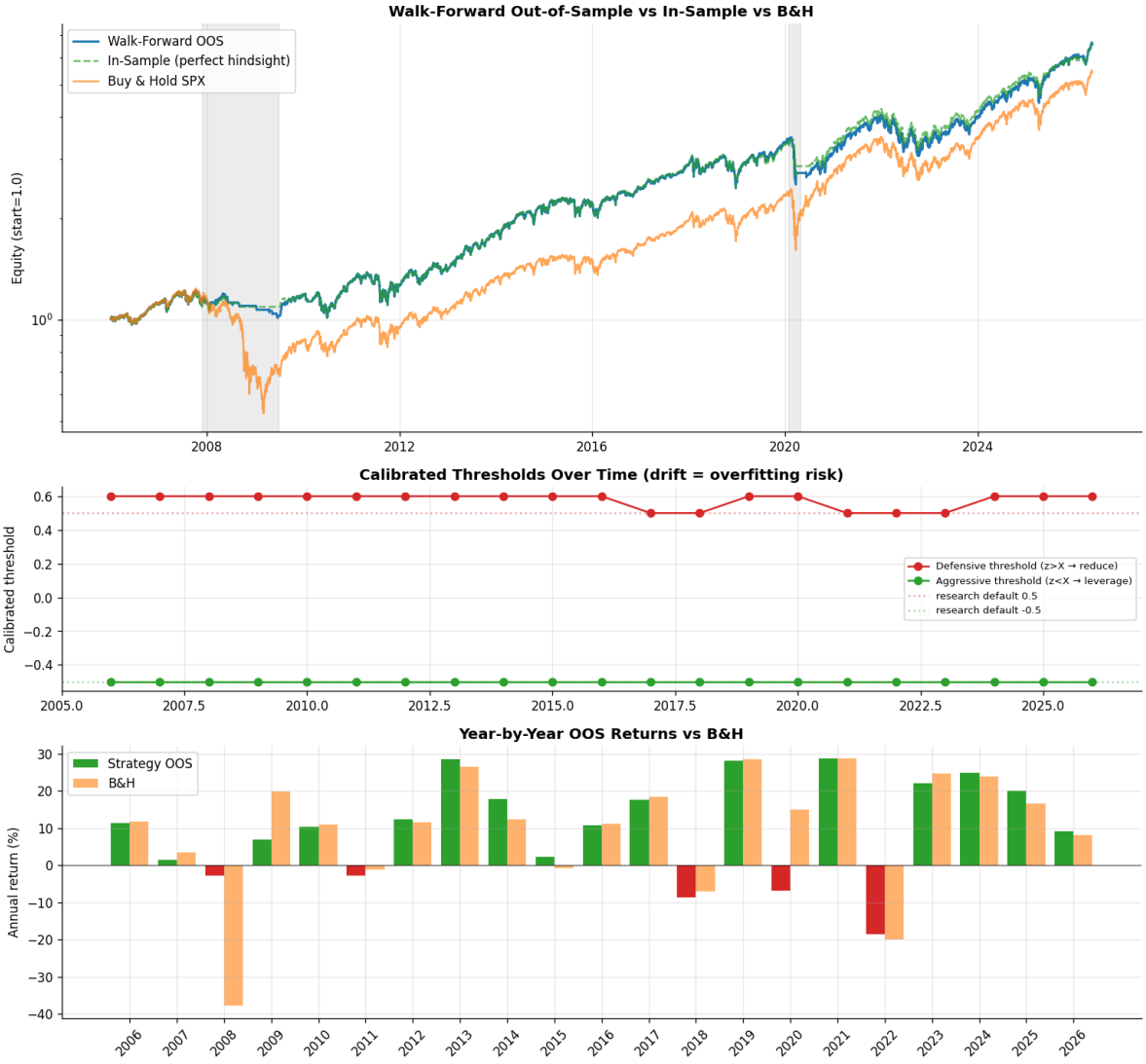

Figure 11. *Primary result.* Walk-forward out-of-sample validation of the labor-only composite overlay, vs. in-sample fit and buy-and-hold S&P 500.

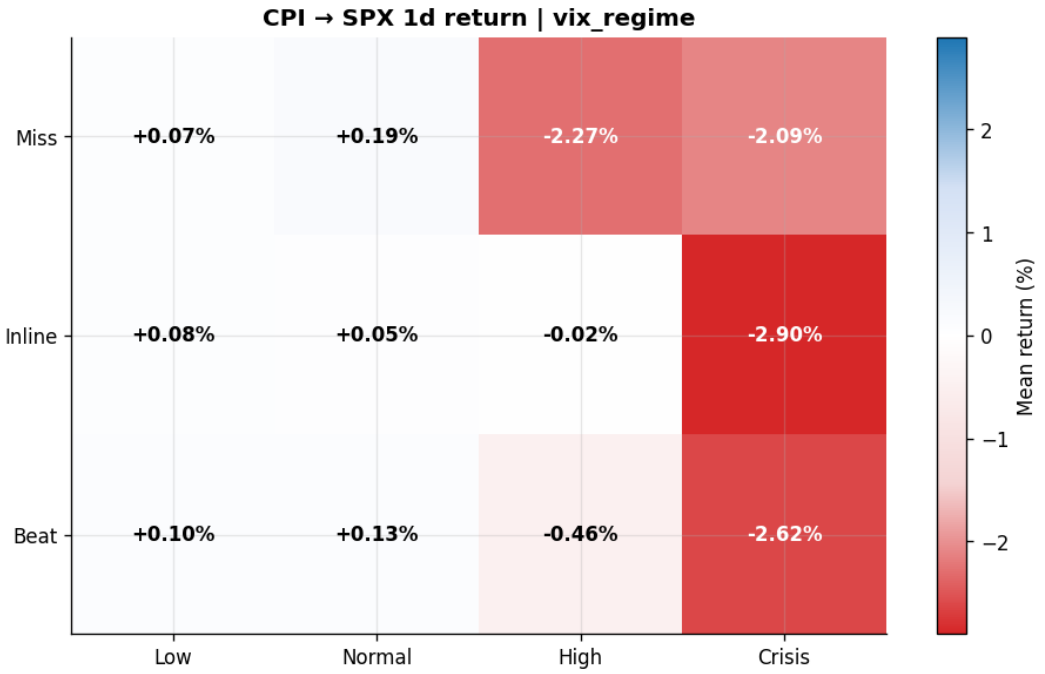

Figure 12. CPI returns conditional on the VIX regime: the sign of the CPI-surprise reaction flips across volatility states.

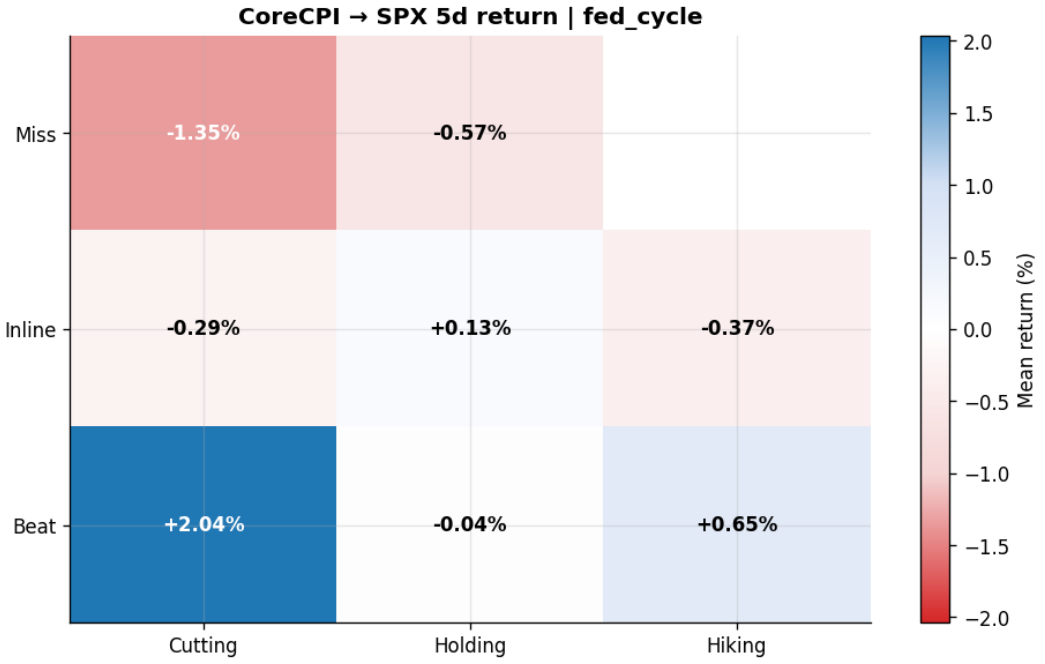

Figure 13. Core CPI returns conditional on the Fed cycle: large bearish-for-equities reactions are concentrated in the hiking regime.