Forecast Alpha and Residual Surprise: A Real-Time Framework for Evaluating Inflation Forecastability Relative to Market Expectations

John E. Kaprich, CFA

June 15, 2026

Working Paper V1.0.0

Abstract

This paper develops a real-time process for evaluating whether inflation models deliver information beyond market-based expectations. The framework combines real-time forecasting, regime classification, and a decomposition of market-relative inflation surprises into Forecast Alpha (the forecastable component) and Residual Surprise (the residual component).

The results indicate that inflation forecastability is economically meaningful but varies across forecast horizons and macroeconomic regimes. The model forecast contains information beyond that embedded in a market-based inflation benchmark, calibration improves the separation between forecastable and residual variation, and Forecast Alpha translates into positive breakeven-market performance. Together, the findings suggest that inflation information is not always fully reflected in contemporaneous market pricing.

1 Introduction

Markets respond most strongly when economic outcomes depart from prior expectations. In his Nobel Prize Lecture, Milton Friedman argued that inflation's economic effects are heavily influenced by deviations from what households, businesses, and investors previously expected. His contribution helped place expectations at the center of macroeconomic analysis and highlighted the importance of information surprises. If expectations matter, then evaluating economic outcomes relative to those expectations becomes equally important.

Market expectations provide a useful benchmark because they aggregate the inflation expectations of many participants. However, market expectations need not fully incorporate all available information at every point in time. Information may be dispersed across participants, difficult to process, or only partially reflected in prices. As a result, realized inflation may differ from market expectations for reasons that are either forecastable ex ante or genuinely unpredictable. Distinguishing between these two sources of market-relative inflation surprise is central to determining whether inflation contains incremental information beyond what is already embedded in market prices.

Subsequent research furthered the notion that inflation movements are not governed by a stable process over time. As expectations came to be viewed increasingly as adaptive and endogenous, economic outcomes became dependent in part on the expectations that preceded them. Once expectations influence the evolution of the process itself, inflation need not follow a single stable path. The forecastability of inflation may therefore become state-dependent, creating a natural role for regime identification when evaluating inflation expectations and information surprises.

Inflation forecasting literature continues to evaluate models primarily through forecast-error metrics such as root mean squared error (RMSE) and mean absolute error (MAE). While these measures assess average predictive accuracy, they provide limited insight into whether forecastability varies across states or whether a model contains information beyond what is already embedded in market expectations. A model may reduce forecast error yet contribute little incremental information if its forecasts merely replicate what markets already anticipate. This distinction between forecasting precision and incremental market information motivates the framework developed in this paper.

Rather than relying exclusively on market expectations and conventional forecast-error metrics, this paper evaluates inflation forecastability relative to market expectations. Specifically, market-relative inflation surprise is decomposed into Forecast Alpha (a forecastable component) and Residual Surprise (an orthogonal unforecastable component). The decomposition allows forecastability, regime dependence, and incremental information content to be evaluated within a unified framework.

Ultimately, the framework evaluates whether incremental information is available relative to market expectations, whether that information is regime-dependent, and whether it can serve as a real-time value-add signal for trading break-evens. The result is a unified framework that integrates macroeconomic expectations theory, nonlinear forecasting, and information transmission within markets.

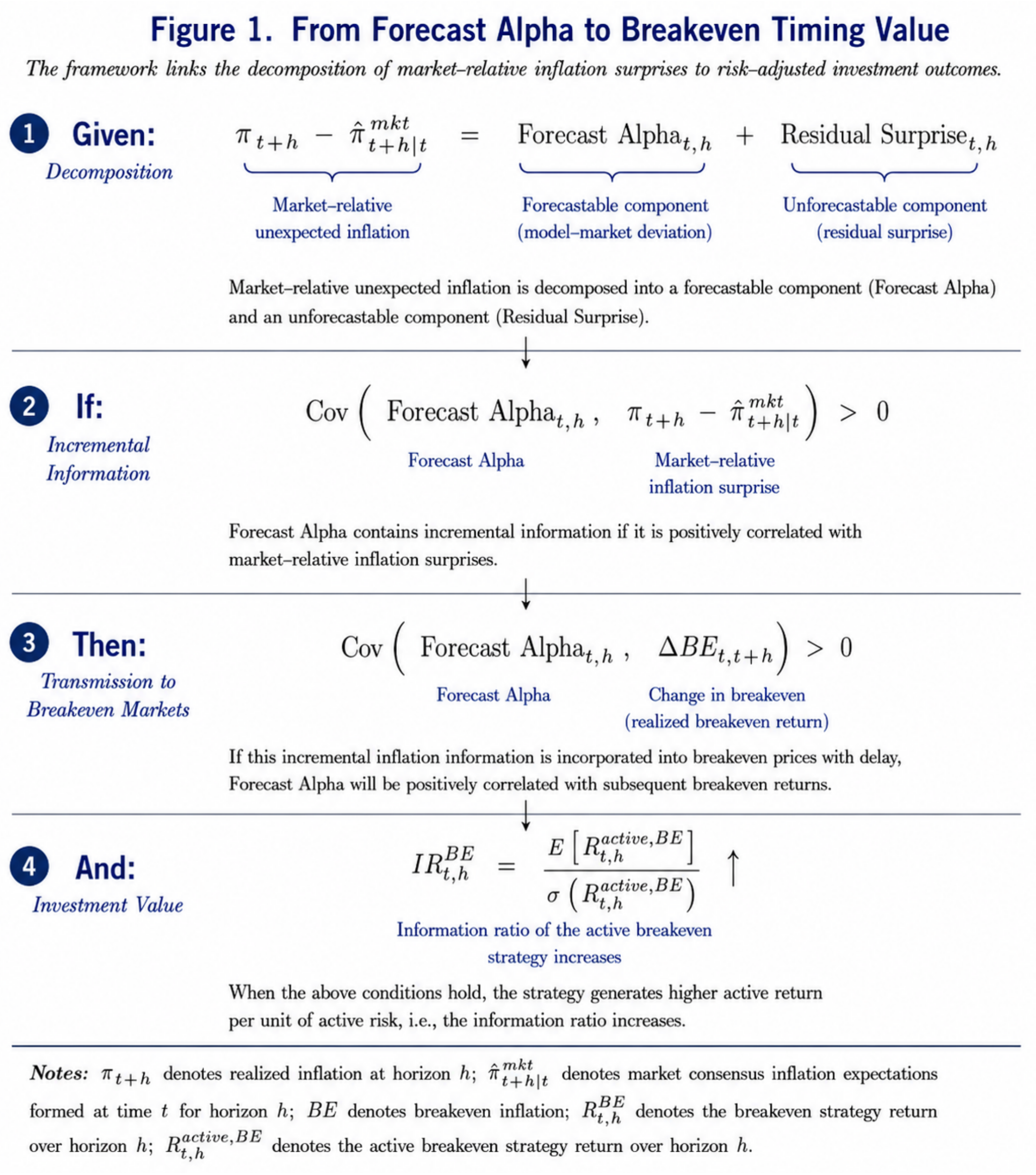

Figure 1. From Forecast Alpha to Breakeven Timing Value

The framework links the decomposition of market–relative inflation surprises to risk–adjusted investment outcomes.

① **Given:**
Decomposition

$$\underbrace{\pi_{t+h} - \hat{\pi}^{mkt}_{t+h|t}}_{\text{Market–relative unexpected inflation}} = \underbrace{\text{Forecast Alpha}_{t,h}}_{\text{Forecastable component (model–market deviation)}} + \underbrace{\text{Residual Surprise}_{t,h}}_{\text{Unforecastable component (residual surprise)}}$$

Market–relative unexpected inflation is decomposed into a forecastable component (Forecast Alpha) and an unforecastable component (Residual Surprise).

② **If:**
Incremental Information

$$\text{Cov}\left(\underbrace{\text{Forecast Alpha}_{t,h}}_{\text{Forecast Alpha}}, \; \underbrace{\pi_{t+h} - \hat{\pi}^{mkt}_{t+h|t}}_{\text{Market–relative inflation surprise}} \right) > 0$$

Forecast Alpha contains incremental information if it is positively correlated with market–relative inflation surprises.

③ **Then:**
Transmission to Breakeven Markets

$$\text{Cov}\left(\underbrace{\text{Forecast Alpha}_{t,h}}_{\text{Forecast Alpha}}, \; \underbrace{\Delta BE_{t,t+h}}_{\text{Change in breakeven (realized breakeven return)}} \right) > 0$$

If this incremental inflation information is incorporated into breakeven prices with delay, Forecast Alpha will be positively correlated with subsequent breakeven returns.

④ **And:**
Investment Value

$$IR^{BE}_{t,h} = \frac{E\left[R^{active,BE}_{t,h}\right]}{\sigma\left(R^{active,BE}_{t,h}\right)} \uparrow$$

Information ratio of the active breakeven strategy increases

When the above conditions hold, the strategy generates higher active return per unit of active risk, i.e., the information ratio increases.

Notes: π_{t+h} denotes realized inflation at horizon h; $\hat{\pi}^{mkt}_{t+h|t}$ denotes market consensus inflation expectations formed at time t for horizon h; BE denotes breakeven inflation; $R^{BE}_{t,h}$ denotes the breakeven strategy return over horizon h; $R^{active,BE}_{t,h}$ denotes the active breakeven strategy return over horizon h.

Figure 1: Forecast Alpha and Residual Surprise Decomposition

The rest of the paper is organized to guide the reader through this process. Section 2 discusses the previous literature that guides this process. Section 3 provides an overview of the data used in this process and Section 4 offers an overview of the methodologies used. Results are provided in Section 5, with a detailed evaluation in Section 6. The paper concludes with Section 7, which discusses how this process could be incorporated into the investment process and what additional research could be conducted to complement this paper's findings.

2 Literature Review

2.1 Expectations, Inflation Surprises, and Market-Based Expectations

Expectations are central to modern macroeconomics and inflation analysis. Milton Friedman argued that the economic consequences of inflation depend less on the level of inflation than on deviations from previous expectations [Friedman, 1977]. His insight motivated a deeper understanding of how unexpected inflation is measured, where expectations originate, and what information is embedded in market prices. Such questions are central when considering inflation expectations' influence on asset prices, real returns, and the valuation of inflation-linked securities. Market-based measures, including Treasury breakeven inflation rates, aggregate market participants' beliefs and provide a natural benchmark for evaluating whether forecasts contain information beyond what is reflected in prices [Ang et al., 2007, Haubrich et al., 2012, Chernov and Mueller, 2012].

Market prices are not pure measures of expected inflation. Treasury breakeven rates embed inflation risk premia, liquidity premia, and other market-specific distortions [D'Amico et al., 2018, Kajuth and Watzka, 2011, Andreasen et al., 2021, Bauer and Rudebusch, 2017]. Consequently, realized inflation may differ from market-implied expectations for reasons that are both forecastable and fundamentally unpredictable. This distinction motivates the decomposition framework shown in Figure 1 and raises the broader question of whether inflation contains incremental information beyond what is already embedded in market pricing.

Realized inflation can differ from market expectations, even if expectations are informative. Some deviations may reflect information that was potentially identifiable ex ante but not fully incorporated into market prices, while others represent genuinely unpredictable shocks. Distinguishing between these two sources of market-relative inflation surprises is central to understanding whether inflation contains incremental information beyond market expectations. Recent evidence further distinguishes between perceived, expected, and realized inflation, highlighting that inflation-related information may differ across agents and information sets [Weber et al., 2023]. More broadly, these findings suggest that realized inflation outcomes need not be interpreted as a single undifferentiated surprise, providing additional motivation for separating forecastable and unforecastable components of market-relative inflation outcomes.

2.2 Forecastable and Unforecastable Inflation Components

Early empirical approaches modeled inflation using relatively stable statistical relationships, including autoregressive specifications and Phillips-curve frameworks [Phillips, 1958, Nelson and Schwert, 1977, Stock and Watson, 1999, Atkeson and Ohanian, 2001, Robinson et al., 2003, Gabrielyan, 2016, Martins and Gabriel, 2020, Jørgensen and Lansing, 2019]. Subsequent research demonstrated that inflation forecasting performance varies substantially across periods and that no single specification consistently dominates across all environments [Stock and Watson, 2007, 2010, 2016, Faust and Wright, 2013].

Real-time forecasting studies further emphasize that forecast performance depends on the information set available at the forecast origin [Groen et al., 2012, Liebermann, 2012, Croushore, 2011]. Time-varying parameter models were developed in part to accommodate this instability by allowing coefficients and shock processes to evolve through time [Primiceri, 2005, Del Negro and Schorfheide, 2013]. These findings suggest that inflation forecastability is time varying, with the predictive content of economic information changing across periods and environments. Simple forecasting specifications may omit time-varying predictive relationships that emerge as economic conditions evolve.

Collectively, these studies show that some portion of inflation is predictable, but they generally stop short of distinguishing between forecastable inflation information and residual inflation surprises. As a result, the literature provides limited guidance regarding how much of realized inflation surprise is economically forecastable versus intrinsically unpredictable.

Machine-learning methods have expanded the forecasting toolkit by allowing high-dimensional predictors, nonlinear interactions, and more flexible function approximation [Medeiros et al., 2021, Goulet Coulombe et al., 2022, Kock and Medeiros, 2020, Medeiros and Mendes, 2016, Li and Chen, 2014, Ozgur and Aydin, 2022, Huang et al., 2024]. Recent evidence suggests that these approaches can improve inflation forecasting performance relative to traditional benchmark models by exploiting nonlinearities and information embedded in large predictor sets [Naghi et al., 2024]. In macro-finance more broadly, machine learning has also proved useful in extracting predictive structure from large information sets [Kelly et al., 2019]. Related model-selection approaches such as LASSO and elastic net provide disciplined methods for controlling predictor proliferation in high-dimensional settings [Tibshirani, 1996, Zou and Hastie, 2005]. However, improvements in forecast accuracy do not necessarily imply that a model contains information beyond what is already incorporated into market expectations. Yet these literatures remain primarily focused on predictive accuracy rather than the separation of forecastable information from residual surprise.

2.3 Regime Dependence and State-Dependent Forecastability

The inflation surge of 2021–22 renewed interest in the possibility that inflation dynamics and forecastability may vary across macroeconomic environments, particularly during periods of rapid structural change. More broadly, a growing literature suggests that inflation is not governed by a single stable process. Inflation dynamics vary over time with changes in monetary policy, macroeconomic volatility, and broader structural conditions [Clarida et al., 2000, Sims and Zha, 2006, Cogley and Sargent, 2005, Bai and Perron, 2003]. This instability motivated a shift toward state-dependent models of inflation.

A major step in that direction was the introduction of regime-switching models by Hamilton [1989]. By allowing the data-generating process to transition across latent states, these models capture nonlinear dynamics, structural breaks, and changes in persistence. In the inflation context, regime-switching frameworks have improved empirical fit and, in some cases, forecasting performance relative to constant-parameter alternatives [Ang et al., 2007, Bianchi, 2013]. More broadly, regime-dependent macro-finance models have helped explain time variation in inflation expectations, bond risk premia, and yield-curve behavior [Ang and Timmermann, 2012, D'Amico et al., 2018].

Subsequent research relaxed the assumption of constant transition probabilities by allowing regime dynamics to respond to observable economic conditions [Filardo, 1994, Diebold et al., 1994, Kim and Nelson, 1999]. These extensions, including time-varying Hidden Markov Models, provide a more flexible representation of macroeconomic state dynamics [Koop and Potter, 2007, Koop and Korobilis, 2013, Shi, 2016]. For inflation analysis, the implication is straightforward: forecastability itself may be state dependent, implying that both the relevance of predictors and the amount of forecastable inflation information may vary across macroeconomic environments.

2.4 Information Transmission and Breakeven Markets

If a forecast contains incremental information beyond market pricing, that information should eventually be reflected in market prices. Treasury breakeven inflation rates provide a natural setting in which to evaluate this proposition because they incorporate both inflation expectations and compensation for inflation risk. The literature has extensively examined the informational content of breakeven inflation rates and their relationship to subsequent realized inflation [Ang et al., 2007, Haubrich et al., 2012, Chernov and Mueller, 2012, D'Amico et al., 2018, Andreasen et al., 2021, Bauer and Rudebusch, 2017]. Accordingly, breakeven repricing provides an observable market-based test of whether forecast disagreement reflects genuine incremental information or estimation error.

Most studies focus either on the forecasting properties of breakevens or on the decomposition of breakeven inflation into expectations and risk-premium components. Relatively little attention has been given to whether model-derived information that is not yet reflected in breakeven prices predicts subsequent breakeven adjustment. This distinction is central to the transmission mechanism illustrated in Figure 1.

2.5 Present Study

The preceding literature leaves three related questions unresolved. First, while inflation surprises matter because expectations matter, it remains unclear how much of market-relative inflation surprise is forecastable in real time. Second, while inflation forecastability may vary across macroeconomic regimes, relatively little evidence exists regarding how the amount of forecastable inflation information changes across states. Third, while market prices provide a useful benchmark for evaluating forecast information, it remains unclear whether model-derived information not yet reflected in prices possesses economic value.

This study develops a framework for identifying, evaluating, and assessing the economic value of incremental inflation information in real time. Rather than focusing exclusively on forecast accuracy, it decomposes market-relative unexpected inflation into a forecastable component (Forecast Alpha) and an orthogonal unforecastable component (Residual Surprise). It then evaluates whether Forecast Alpha contains incremental information relative to market pricing, whether that information is regime dependent, and whether it is ultimately reflected in breakeven market performance.

The contribution is therefore not simply a new forecasting specification. Instead, the paper links inflation forecasting, regime dependence, and market-based information transmission within a unified framework that evaluates both statistical forecastability and economic relevance. By separating market-relative inflation surprises into forecastable and unforecastable components, the framework provides a direct method for evaluating whether model-market disagreement reflects genuine incremental information or merely forecast error.

3 Data

3.1 Predictor Variables

The forecasting environment incorporates macroeconomic and financial variables obtained from the Federal Reserve Economic Data (FRED) database. The broader FRED universe documents the information set considered in the research design and is intended to capture the primary drivers of inflation dynamics, including labor-market conditions, economic activity, monetary policy, financial conditions, commodity prices, and inflation expectations.

The candidate predictor universe spans seven broad categories: inflation, labor-market, monetary-policy, interest-rate, producer-price, commodity-price, and survey-based indicators obtained from FRED. Appendix Table 11 provides the complete list of candidate variables considered during model development.

The empirical results reported in the main paper are based on the Compact-Direct specification described in Section 4. This specification uses a fixed compact predictor set consisting of inflation-persistence variables and a small macro-financial information block. The broader candidate universe is therefore used to document the research design and motivate the compact information set, while the reported baseline forecasts do not rely on a separate broad recursive feature-selection procedure.

3.2 Market-Based Inflation Benchmark

To evaluate inflation relative to market pricing, the paper uses the 5-Year Treasury Inflation Breakeven Rate (FRED series `T5YIE`) as a market-based inflation benchmark. The breakeven rate is derived from the difference between nominal Treasury yields and Treasury Inflation-Protected Securities (TIPS) yields and reflects inflation compensation embedded in financial markets.

Although breakeven inflation is linked to CPI-based securities rather than PCE inflation directly and may reflect risk premia, liquidity effects, and other market-specific distortions, it is widely used as a forward-looking measure of market inflation expectations and exhibits substantial comovement with realized inflation over time. In this study, the series is employed as a common market-based inflation anchor against which model-generated information can be evaluated.

Core PCE is used because it captures more persistent inflation dynamics that are likely to influence medium-term inflation expectations and the pricing of inflation-sensitive assets. Core inflation is also less affected by short-term fluctuations in food and energy prices, making it a useful measure of the underlying inflation trend. To the extent that temporary food and energy shocks influence realized inflation, their effects are expected to be reflected primarily in the Residual Surprise component rather than Forecast Alpha.

The objective is not to match the inflation measure embedded in market pricing exactly, but rather to determine whether information contained in the underlying inflation trend is fully incorporated into market-based inflation expectations.

Figure 2 illustrates the relationship between realized Core PCE inflation and the market-based inflation anchor used throughout the analysis.

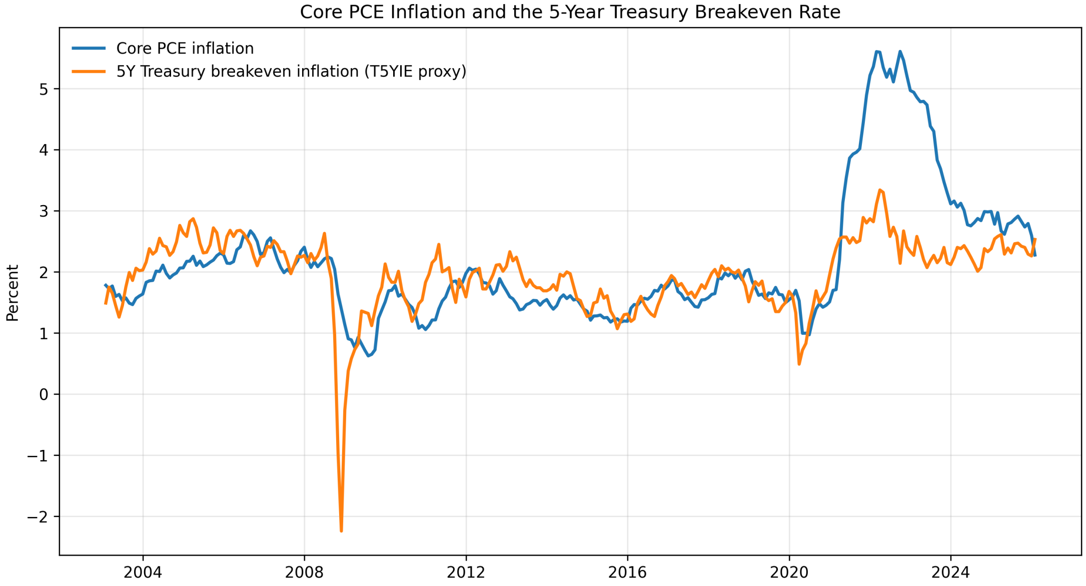

Figure 2: Core PCE inflation and the 5-year Treasury Inflation Breakeven Rate used as the market-based inflation anchor.

3.3 Data Frequency and Real-Time Alignment

The empirical implementation is monthly and uses month-end timestamps throughout. All variables are aligned such that only information available at time t is used to forecast inflation at time $t + h$. This structure is designed to approximate the real-time information set at each forecast origin and to prevent look-ahead bias in forecast construction and evaluation. The same timing convention is applied to forecast generation, regime classification, and model evaluation.

Where necessary, lower-frequency macroeconomic series are aligned to the monthly forecast calendar using the most recently available observation at the forecast origin. Publication-lag discipline is imposed where possible to ensure that the information set reflects real-time availability rather than revised data.

The empirical sample begins in January 1978, reflecting the start of the unified monthly panel, and extends through December 31, 2025, the formal endpoint of the empirical study. This period encompasses multiple inflation regimes, including the high-inflation environment of the late 1970s and early 1980s, the subsequent disinflation period, the low and stable inflation environment of the 2010s, and the inflation surge following 2020.

3.4 Data Integrity and Reproducibility

All series are obtained directly from FRED and cached locally to support reproducibility. The data pipeline is designed to preserve real-time information availability by preventing the inclusion of future-dated observations and supporting recursive model estimation. Forecasts, filtered regime labels, and evaluation statistics are merged and evaluated using the forecast-origin information set.

4 Methodology

This study develops a real-time framework for evaluating whether inflation contains forecastable information beyond that already embedded in market pricing. The framework combines six elements: (1) a direct horizon-specific compact inflation forecasting model, (2) a market-based inflation benchmark, (3) a market-relative decomposition that separates inflation surprise into Forecast Alpha and Residual Surprise, (4) an empirically estimated inflation-band structure that organizes latent macroeconomic regimes, (5) a recursive pass-through calibration that estimates the historically realizable portion of Forecast Alpha, and (6) a regime-conditioning and market-relevance framework linking incremental inflation information to subsequent breakeven-market repricing.

The empirical implementation is organized around the Compact-Direct forecasting system. The canonical forecast engine is a direct, horizon-specific Elastic Net model for year-over-year Core PCE inflation. Regime information is not used to define a separate regime-specific forecast equation and is not used as an alternative forecasting model. Instead, filtered TV-HMM regime probabilities are merged into the Compact-Direct Forecast Alpha panel as a conditioning layer for evaluating whether forecastability and market relevance vary across macroeconomic states.

The methodology operationalizes the framework illustrated in Figure 1 through six stages. First, direct inflation forecasts are generated using a compact real-time forecasting model. Second, model forecasts are compared with a Treasury breakeven-based market benchmark to construct raw Forecast Alpha and raw Residual Surprise. Third, an empirical inflation-band procedure is used to define the economic structure of the regime layer. Fourth, raw Forecast Alpha is calibrated recursively to estimate the portion of model–market disagreement that has historically passed through into realized market-relative inflation outcomes. Fifth, filtered TV-HMM regime labels are used to evaluate state dependence in Forecast Alpha and Residual Surprise. Finally, the economic relevance of Forecast Alpha is assessed through subsequent breakeven inflation repricing and related market outcomes.

The objective is not simply to minimize forecast error in an unconditional sense. Instead, the framework asks whether inflation contains forecastable information beyond that embedded in a market-based inflation benchmark, whether that information survives recursive calibration, and whether its relevance varies across macroeconomic states.

4.1 Forecasting Objective

Let π_t denote year-over-year Core PCE inflation at time t. In the empirical implementation, π_t is constructed from the Core PCE price index as

$$\pi_t = 100\left(\frac{P_t}{P_{t-12}} - 1\right),$$

where P_t denotes the level of the Core PCE price index. For each forecast origin t and horizon $h \in \{1, 3, 6, 12\}$ months, the forecasting objective is to form a direct forecast of the future level of year-over-year Core PCE inflation:

$$\hat{\pi}^{\text{model}}_{t+h|t} = E(\pi_{t+h} \mid \mathcal{I}_t),$$

where $\mathcal{I}_t$ denotes the information set available at the forecast origin.

The empirical implementation uses a direct horizon-specific forecasting design rather than an iterated multi-step design. A separate forecast equation is estimated for each horizon of interest. This choice is motivated by two considerations. First, direct forecasting avoids compounding specification error across intermediate steps. Second, it aligns naturally with the paper's decomposition framework, which evaluates the model's incremental information relative to a common market-based inflation benchmark at each horizon. This design is consistent with the broader real-time forecasting literature, which emphasizes that forecast evaluation should respect the information set available at the forecast origin [Croushore, 2011, Faust and Wright, 2013, Groen et al., 2012].

4.2 Market-Based Inflation Benchmark

Let $\hat{\pi}^{\text{mkt}}_{t+h|t}$ denote the market-based inflation benchmark associated with forecast origin t and horizon h. The empirical implementation uses the 5-year Treasury Inflation Breakeven Rate (T5YIE) as a common market-based inflation-pricing anchor. Because a clean horizon-matched market-implied Core PCE inflation series is not available for each horizon considered in the paper, the benchmark is not interpreted as a literal horizon-specific forecast of Core PCE inflation. Instead, it serves as a common market inflation anchor against which the model's incremental information is evaluated.

This interpretation is consistent with the literature showing that breakeven inflation rates contain market information about inflation expectations but may also embed inflation risk premia, liquidity premia, and other market-specific distortions [Ang et al., 2007, Haubrich et al., 2012, Chernov and Mueller, 2012, D'Amico et al., 2018, Kajuth and Watzka, 2011, Andreasen et al., 2021, Bauer and Rudebusch, 2017]. The paper therefore evaluates whether the Compact-Direct model contains information relative to market inflation pricing, rather than claiming that T5YIE is a pure measure of expected Core PCE inflation at each forecast horizon.

4.3 Market-Relative Inflation Decomposition

The first object in the decomposition is raw Forecast Alpha:

$$\alpha_{t,h}^{raw} \equiv \hat{\pi}_{t+h|t}^{\text{model}} - \hat{\pi}_{t+h|t}^{\text{mkt}}.$$

Raw Forecast Alpha measures the model's ex ante deviation from the market-based inflation benchmark. Market-relative unexpected inflation is defined as

$$s_{t,h} \equiv \pi_{t+h} - \hat{\pi}_{t+h|t}^{\text{mkt}}.$$

Raw Residual Surprise is defined as the realized model forecast error:

$$\varepsilon_{t,h}^{raw} \equiv \pi_{t+h} - \hat{\pi}_{t+h|t}^{\text{model}}.$$

The raw accounting decomposition is therefore

$$s_{t,h} = \alpha_{t,h}^{raw} + \varepsilon_{t,h}^{raw}.$$

This identity separates model–market disagreement observed at the forecast origin from the subsequent residual inflation outcome not captured by the model forecast. The identity does not imply that raw Forecast Alpha passes through one-for-one into realized market-relative inflation surprise, nor does it impose orthogonality between $\alpha_{t,h}^{raw}$ and $\varepsilon_{t,h}^{raw}$.

To estimate the portion of raw Forecast Alpha that has historically been realized in market-relative inflation outcomes, the paper introduces a recursive calibration step. For each horizon h, the following pass-through regression is estimated using only prior observations whose realized outcomes were available before the current forecast origin:

$$s_{\tau,h} = a_{t,h} + \lambda_{t,h}\alpha_{\tau,h}^{raw} + u_{\tau,h}, \qquad \tau \in \mathcal{H}_{t,h},$$

where $\mathcal{H}_{t,h}$ denotes the historical calibration sample available at forecast origin t for horizon h. The coefficient $\lambda_{t,h}$ measures the historical pass-through from raw model–market disagreement into realized market-relative inflation surprise.

The calibrated Forecast Alpha signal is defined as

$$\alpha_{t,h}^{cal} \equiv \hat{\lambda}_{t,h} \alpha_{t,h}^{raw}.$$

The calibration regression estimates an intercept, but the intercept is not included in calibrated Forecast Alpha. This convention keeps Forecast Alpha centered on directional model–market disagreement rather than treating a constant historical bias as part of the alpha signal. The intercept is retained as a calibration diagnostic.

Calibrated Residual Surprise is defined as

$$\eta_{t,h} \equiv s_{t,h} - \alpha_{t,h}^{cal}.$$

The calibrated decomposition is therefore

$$s_{t,h} = \alpha_{t,h}^{cal} + \eta_{t,h}.$$

The decomposition framework implies a testable condition for incremental information. Forecast Alpha contains economically relevant information if it is positively associated with subsequent market-relative inflation surprises:

$$\mathrm{Cov}\left(\alpha_{t,h}^{cal}, s_{t,h}\right) > 0.$$

A positive covariance indicates that the calibrated forecastable component captures variation in realized market-relative inflation outcomes that was not fully reflected in market pricing at the forecast origin.

4.4 Real-Time Information Set and Predictor Construction

The real-time information set combines an inflation-persistence backbone with a compact set of exogenous macro-financial predictors. The purpose of this design is to retain the discipline and interpretability of a persistence-based forecasting model while allowing a small number of economically motivated variables to improve forecast performance. The benchmark role of persistence and Phillips-curve information in inflation forecasting is well established in the inflation-forecasting literature [Stock and Watson, 1999, Atkeson and Ohanian, 2001, Robinson et al., 2003, Martins and Gabriel, 2020].

For each forecast origin, the predictor set is constructed from variables observable at, or lagged to, time t. The current Compact-Direct implementation uses the following feature blocks:

1. *Inflation persistence block:* lagged year-over-year Core PCE inflation, recent inflation changes, a short rolling inflation mean, and short rolling inflation volatility;

2. *Labor-market block:* unemployment level, an unemployment gap measure, and recent changes in unemployment;

3. *Policy and rates block:* the effective federal funds rate, the 2-year Treasury yield, changes in these rates, and a 2-year Treasury less federal funds rate spread;
4. *Price-pressure block:* producer-price and oil-price changes;
5. *Expectations block:* survey-based inflation-expectation information and recent changes in that measure.

The resulting design is intentionally compact. The goal is not to approximate a large machine-learning feature universe or to rely on regime-aware feature selection. Instead, the model estimates a disciplined inflation forecast that complements persistence with a limited amount of additional economic information. The final compact predictor set is constructed inside the Compact-Direct forecasting runner, while the shared data-loading and release-lag infrastructure is used to align the macro-financial panel.

All predictors are aligned to month-end timestamps. Publication-lag discipline is imposed where available so that the forecasting system approximates the information set observable at the forecast origin. The final predictor matrix therefore consists of month-end macro-financial variables transformed into a compact horizon-specific design matrix suitable for direct forecasting.

Because the dependent variable is year-over-year inflation and the predictor matrix is constructed from lags, changes, spreads, rolling moments, and rate variables rather than price levels, the empirical design limits reliance on nonstationary level relationships. The paper does not require a full-sample cointegrating relationship between macroeconomic price levels and inflation outcomes. Instead, identification comes from recursive out-of-sample forecast performance, market-relative decomposition, and forecast-origin information discipline.

4.5 Direct Compact Forecast Model

Given the real-time predictor set, inflation is forecast using a direct compact model estimated separately for each horizon. Let $X_{t,h}$ denote the horizon-specific predictor vector available at time t. The forecast equation is

$$\hat{\pi}^{\text{model}}_{t+h|t} = f_h(X_{t,h}),$$

where $f_h(\cdot)$ denotes the horizon-specific forecasting rule.

The canonical forecast engine is an Elastic Net regression estimated recursively using an expanding window. For each horizon h, the model solves

$$\hat{\beta}_{t,h} = \arg\min_{\beta_0,\beta} \left\{ \frac{1}{N_{t,h}} \sum_{\tau \in \mathcal{T}_{t,h}} \left(\pi_{\tau+h} - \beta_0 - X'_{\tau,h}\beta\right)^2 + \lambda^{EN}_{t,h} \left[\frac{1-\rho_{t,h}}{2} \|\beta\|_2^2 + \rho_{t,h}\|\beta\|_1 \right] \right\},$$

where $\mathcal{T}_{t,h}$ denotes the expanding-window training sample available at forecast origin t, $N_{t,h}$ is the number of usable training observations, $\lambda^{EN}_{t,h}$ is the Elastic Net penalty strength, and $\rho_{t,h}$ is the mixing parameter between the ridge and LASSO penalty components.

Elastic Net regularization is appropriate in this setting because the compact macro-financial predictor set contains correlated variables, including inflation lags, rate variables, labor-market indicators, and price-pressure measures. The Elastic Net penalty combines the shrinkage properties of ridge regression with the sparse-selection properties associated with the LASSO penalty [Tibshirani, 1996, Zou and Hastie, 2005]. The implementation does not estimate a separate pure LASSO model. References to LASSO in this paper therefore refer to the ℓ_1 component of the Elastic Net penalty rather than to a distinct LASSO forecast engine.

Predictors are standardized within each expanding-window training sample, and the same transformation is applied to the predictor vector observed at the forecast origin. Elastic Net tuning parameters are selected within the available training sample using a recent validation block. The selected specification is then refit on the full available expanding-window sample before generating the forecast $\hat{\pi}^{\text{model}}_{t+h|t}$.

A key methodological distinction is that the forecast engine is not a weighted mixture of regime-specific forecasting models. The Compact-Direct model is the paper's canonical forecast equation. The TV-HMM operates as a separate classification layer used to organize conditional analysis after the forecast and Forecast Alpha decomposition have been constructed.

4.6 Inflation-Band Structure for Regime Classification

The regime layer is anchored by an empirical inflation-band structure. This step is separate from the Compact-Direct forecast equation, but it is part of the regime methodology because it determines the economic organization of the latent states used in the TV-HMM. The purpose is to avoid imposing arbitrary regime cutoffs and instead let the historical distribution of inflation determine the boundaries between low-, moderate-, and high-inflation environments.

Let P_t denote the relevant inflation price index and define year-over-year inflation as

$$\pi_t = 100\left(\frac{P_t}{P_{t-12}} - 1\right).$$

The historical distribution of π_t is modeled using one-dimensional finite mixtures. Candidate Gaussian mixture and Student-t mixture specifications are estimated across alternative numbers of components. For a given number of components K, the generic mixture density is

$$f(\pi_t) = \sum_{k=1}^{K} \omega_k f_k(\pi_t;\theta_k), \qquad \sum_{k=1}^{K} \omega_k = 1, \qquad \omega_k > 0,$$

where ω_k is the component weight and $f_k(\cdot)$ is either a Gaussian or Student-t component density. Candidate specifications are compared using the Bayesian Information Criterion,

$$BIC_K = -2\ell_K + q_K \log(N),$$

where ℓ_K is the maximized log likelihood, q_K is the number of estimated parameters, and N is the number of inflation observations. In the empirical implementation, the lowest-BIC inflation-band specification produces a three-band structure, corresponding to low-, moderate-, and high-inflation environments. The TV-HMM is therefore estimated with $K = 3$ latent states, so the number of regimes is disciplined by the inflation distribution rather than imposed solely as a modeling convention.

The boundaries between inflation bands are defined by posterior component assignment. Let

$$\Pr(B_t = k \mid \pi_t) = \frac{\omega_k f_k(\pi_t; \theta_k)}{\sum_{\ell=1}^{K} \omega_\ell f_\ell(\pi_t; \theta_\ell)}$$

denote the posterior probability that observation π_t belongs to mixture component k. Inflation-band boundaries occur at values of π_t where the most likely posterior component changes. The resulting bands provide economically interpretable inflation states and define the structure used to organize the latent regimes.

These bands do not generate the inflation forecast and do not replace the Compact-Direct Elastic Net model. Their role is to discipline the regime architecture. The TV-HMM then estimates filtered probabilities for the latent regime states within this empirically motivated structure, allowing the paper to evaluate whether Forecast Alpha, Residual Surprise, and market relevance vary across inflation environments.

4.7 Regime Identification with a Time-Varying Hidden Markov Model

Macroeconomic regimes are represented using filtered posterior probabilities from a time-varying Hidden Markov Model. Let $S_t \in \{1, \ldots, K\}$ denote the latent regime at time t. Consistent with the three-band inflation structure selected above, the current implementation uses $K = 3$ latent states. The number and interpretation of these states are therefore tied to the empirical inflation distribution rather than to alternative forecasting models. The regime-classification layer is distinct from the inflation forecast engine. Its role is to characterize the prevailing macroeconomic environment and to support state-dependent analysis of Forecast Alpha, Residual Surprise, and market relevance.

The TV-HMM is based on a compact macro-financial state vector rather than on the full forecasting matrix. The state vector includes inflation conditions, recent inflation changes, rolling inflation moments, selected labor-market indicators, short-rate and yield-curve variables, and commodity-price changes. The latent state process follows a first-order Markov structure with time-varying transition probabilities:

$$P(S_t = j \mid S_{t-1} = i, Z_t) = p_{ij}(t),$$

where Z_t denotes the regime-detection input vector available at time t.

In the implemented TV-HMM, transition probabilities are allowed to vary with observable macro-financial information through a normalized probability mapping. Equivalently, for each prior state i, the transition probability to state j can be written in multinomial-logit form:

$$p_{ij}(t) = \frac{\exp(\gamma_{ij}' Z_t)}{\sum_{\ell=1}^{K} \exp(\gamma_{i\ell}' Z_t)}.$$

This specification builds on the regime-switching tradition introduced by Hamilton [1989] and on later time-varying transition-probability models in which transition dynamics respond to observable economic information [Filardo, 1994, Diebold et al., 1994, Kim and Nelson, 1999]. In macroeconomic applications, such models are useful because regime persistence and transition risk need not be constant across monetary-policy, inflation, and business-cycle environments.

For each forecast origin, the regime layer uses filtered posterior probabilities

$$P(S_t = k \mid \mathcal{I}_t),$$

which represent the probability assigned to regime k using information through time t. Filtered rather than smoothed probabilities are used in the Compact-Direct panel so that regime inference is aligned with the forecast-origin information set. The dominant regime is defined as the state with the largest filtered posterior probability:

$$\hat{S}_t = \arg\max_k P(S_t = k \mid \mathcal{I}_t).$$

In the present paper, the regime layer is used for conditioning and interpretation. Regime estimates are used to evaluate whether Forecast Alpha and Residual Surprise vary across macroeconomic states, rather than to define the main forecast equation itself.

4.8 Alpha Share

To summarize the relative importance of the forecastable and residual components, the paper defines Alpha Share using absolute-value magnitudes. For the raw decomposition, Alpha Share is defined as

$$AS_h^{raw} = \frac{E[|\alpha_{t,h}^{raw}|]}{E[|\alpha_{t,h}^{raw}|] + E[|\varepsilon_{t,h}^{raw}|]}.$$

This statistic measures the relative magnitude of the raw model–market signal compared with the residual component.

For the calibrated decomposition, the corresponding statistic is

$$AS_h^{cal} = \frac{E[|\alpha_{t,h}^{cal}|]}{E[|\alpha_{t,h}^{cal}|] + E[|\eta_{t,h}|]}.$$

Calibrated Alpha Share measures the relative magnitude of calibrated Forecast Alpha compared with calibrated Residual Surprise after accounting for the estimated pass-through from raw model–market disagreement into realized inflation outcomes. Higher values indicate that a larger portion of realized market-relative inflation variation is captured by the calibrated forecastable component.

Because the framework uses a regime-conditioning layer, Alpha Share can also be evaluated conditionally on the inferred regime. The calibrated regime-conditional measure is

$$AS_{h,k}^{cal} = \frac{E[|\alpha_{t,h}^{cal}| \mid \hat{S}_t = k]}{E[|\alpha_{t,h}^{cal}| \mid \hat{S}_t = k] + E[|\eta_{t,h}| \mid \hat{S}_t = k]}.$$

This conditional decomposition allows the paper to test whether inflation predictability varies systematically across macroeconomic states. In the empirical sections, raw Alpha Share is useful for describing the original model–market signal, while calibrated Alpha Share is the preferred statistic for interpreting the forecastable component after pass-through adjustment.

4.9 Recursive Real-Time Estimation Procedure

The empirical procedure is implemented as a sequence of forecast-origin calculations. The inflation-band structure is estimated as part of the regime-design step and is used to organize the latent state space. Conditional on that regime structure, the out-of-sample Compact-Direct forecast and decomposition panel is then generated recursively. For each forecast origin t and horizon h, the following steps are performed:

1. Construct the month-end macro-financial information set using data available through, or lagged to, the forecast origin;
2. Build the compact predictor matrix for the direct forecast model;
3. Estimate the horizon-specific Elastic Net forecast model on an expanding window ending at t;
4. Generate the direct inflation forecast $\hat{\pi}_{t+h|t}^{\text{model}}$;
5. Align the model forecast with the T5YIE-based market benchmark $\hat{\pi}_{t+h|t}^{\text{mkt}}$;
6. Compute raw Forecast Alpha, market-relative unexpected inflation, and raw Residual Surprise;
7. Estimate the horizon-specific pass-through coefficient using only prior observations whose realized outcomes were available before the current forecast origin;
8. Compute calibrated Forecast Alpha and calibrated Residual Surprise;
9. Merge filtered TV-HMM regime probabilities and dominant regime labels for regime-based evaluation;
10. Store the out-of-sample forecast, decomposition terms, regime labels, calibration diagnostics, and market-relevance variables for evaluation.

This recursive design preserves the real-time structure of the forecast and calibration exercises. At each forecast origin, the Compact-Direct forecast is estimated using only the expanding-window training sample available at that date. For multi-month horizons, the calibration step uses only historical observations whose target outcomes were already realized by the current forecast origin, rather than all prior forecast-origin rows. This timing restriction prevents overlapping-horizon look-ahead bias in the pass-through estimates and keeps the calibration exercise aligned with real-time forecast-evaluation practice [Croushore, 2011, Faust and Wright, 2013].

4.10 Market-Relevance and Breakeven Repricing

The final stage evaluates whether Forecast Alpha is related to subsequent breakeven-market repricing. The exercise is not intended to represent a fully optimized investment strategy. Instead, it provides a disciplined implementation test of whether the forecastable component of market-relative inflation surprise survives translation into inflation-sensitive market instruments.

For each signal date, Forecast Alpha is compared with subsequent changes in 5-year and 10-year breakeven inflation. The implementation evaluates simple threshold-based rules in which the position is determined by the sign and magnitude of Forecast Alpha:

$$q_{t,h} = \operatorname{sign}(\alpha_{t,h}) \cdot \mathbf{1}\{|\alpha_{t,h}| \geq c_h\},$$

where $q_{t,h}$ is the signal-implied position and c_h is the horizon-specific activation threshold. The realized payoff is measured using the sign convention implemented in the market-relevance tests:

$$r_{t,h}^{BE} = -q_{t,h}\Delta BE_{t+1} - TC_{t,h},$$

where ΔBE_{t+1} is the subsequent breakeven move and $TC_{t,h}$ denotes transaction-cost adjustment. Under this convention, the payoff is the negative of the signal-implied position multiplied by the realized breakeven move. The sign convention is therefore part of the implementation and is held fixed when evaluating hit rates, mean payoff, and information ratios.

The market-relevance tests are evaluated in full sample and by regime. The purpose is to determine whether Forecast Alpha corresponds to subsequent breakeven repricing and whether that relationship varies across the filtered TV-HMM regimes.

4.11 Interpretation

The methodology is designed to answer two related questions. These questions correspond directly to the framework presented in Figure 1: whether market-relative inflation surprises contain a forecastable component and whether that component ultimately possesses economic value.

The empirical implementation answers these questions by combining a parsimonious direct forecast model with a market-relative decomposition, a recursive calibration step, and a separate regime-conditioning layer. The Compact-Direct Elastic Net model provides the paper's canonical inflation forecast. The raw decomposition translates model–market disagreement into an ex ante signal. The calibrated decomposition estimates the portion of that signal that has historically passed through into realized market-relative inflation outcomes. The TV-HMM regime labels provide filtered state probabilities used to organize conditional analysis.

The contribution of the framework is therefore not simply a new point forecast of inflation. Rather, it is a real-time, market-relative, and regime-conditioned structure for distinguishing between raw model–market disagreement, the calibrated forecastable component of that disagreement, and the calibrated Residual Surprise that remains after the pass-through adjustment. This distinction is important because the raw model–market gap may contain useful directional information even when its magnitude does not pass through one-for-one into realized inflation outcomes. The framework therefore provides the bridge between inflation forecasting, market expectations, and the portfolio-relevance tests developed in the evaluation section.

5 Results

This section presents the empirical evidence associated with the paper's market-relative inflation framework. The analysis evaluates the existence of a forecastable component beyond market pricing, the dependence of forecastability on macroeconomic regimes, the credibility of the forecasting engine, the presence of information beyond market expectations, and the behavior of the calibrated residual component.

5.1 Market-Relative Inflation Decomposition

The objective of this subsection is to quantify the relative magnitudes of Forecast Alpha and Residual Surprise across forecast horizons.

Raw Forecast Alpha represents the ex ante model–market disagreement, calibrated Forecast Alpha represents the pass-through-adjusted forecastable component of that signal, and calibrated Residual Surprise represents the remaining market-relative inflation surprise after calibration.

Table 1 reports mean absolute raw Forecast Alpha, mean absolute calibrated Forecast Alpha, mean absolute calibrated Residual Surprise, and calibrated Alpha Share:

$$AS_h^{cal} = \frac{E[|\alpha_{t,h}^{cal}|]}{E[|\alpha_{t,h}^{cal}|] + E[|\eta_{t,h}|]}, \qquad h \in \{1, 3, 6, 12\}.$$

Table 1 reports the no-look-ahead calibrated decomposition by horizon. The mean pass-through coefficient declines monotonically with the forecast horizon, from 0.941 at the 1-month horizon to 0.499 at the 12-month horizon.

Calibrated Alpha Share equals 0.849, 0.716, 0.590, and 0.372 at the 1-, 3-, 6-, and 12-month horizons, respectively. Forecast Alpha exceeds calibrated Residual Surprise through the 6-month horizon, while calibrated Residual Surprise becomes the larger component at the 12-month horizon. The pass-through adjustment reduces the magnitude of Forecast Alpha at all horizons, with the

Table 1: Calibrated Forecast Alpha and Residual Surprise by Horizon

Horizon	N	Mean $\hat{\lambda}$	Mean $\lvert\alpha^{raw}\rvert$	Mean $\lvert\alpha^{cal}\rvert$	Mean $\lvert\eta\rvert$	Cal. Alpha Share
H1	217	0.941	0.602	0.575	0.103	0.849
H3	215	0.807	0.633	0.525	0.208	0.716
H6	212	0.684	0.675	0.465	0.323	0.590
H12	206	0.499	0.625	0.287	0.484	0.372

The table reports the no-look-ahead calibrated decomposition of market-relative inflation surprise by horizon. Raw Forecast Alpha, α^{raw}, is the model forecast minus the market-based inflation benchmark. Calibrated Forecast Alpha, α^{cal}, rescales raw Forecast Alpha using the recursively estimated pass-through coefficient $\hat{\lambda}$, estimated only from prior observations whose realized outcomes were available before the forecast origin. Calibrated Residual Surprise, η, is market-relative inflation surprise minus calibrated Forecast Alpha. Calibrated Alpha Share is computed as $E[|\alpha^{cal}|]/(E[|\alpha^{cal}|] + E[|\eta|])$. Market benchmark source: T5YIE proxy.

largest reduction occurring at the 12-month horizon where the estimated pass-through coefficient is lowest.

5.2 Regime Dependence of Forecastability

The objective of this subsection is to evaluate whether the calibrated decomposition varies across macroeconomic regimes. The calibrated decomposition is evaluated conditional on the dominant filtered TV-HMM regime at the forecast origin. For each horizon–regime cell, Table 2 reports mean absolute raw Forecast Alpha, mean absolute calibrated Forecast Alpha, mean absolute calibrated Residual Surprise, and calibrated Alpha Share.

Table 2 reports regime-conditional decomposition results by horizon. At the 1-month horizon, calibrated Alpha Share equals 0.888, 0.870, and 0.770 in regimes S0, S1, and S2, respectively. At the 3-month horizon, calibrated Alpha Share equals 0.802, 0.757, and 0.564 in regimes S0, S1, and S2, respectively.

At the 6-month horizon, calibrated Alpha Share equals 0.647, 0.640, and 0.460 in regimes S0, S1, and S2, respectively. At the 12-month horizon, calibrated Alpha Share equals 0.361, 0.434, and 0.302 in regimes S0, S1, and S2, respectively.

Across horizons and regimes, the magnitudes of raw Forecast Alpha, calibrated Forecast Alpha, calibrated Residual Surprise, and calibrated Alpha Share vary meaningfully, indicating that inflation forecastability is state dependent rather than constant across macroeconomic environments.

5.3 Forecast Model Validation

This subsection evaluates the forecasting performance of the Compact Inflation Model relative to benchmark alternatives. The objective is to verify that the forecast engine satisfies conventional

Table 2: Regime-Conditional Calibrated Forecast Alpha and Residual Surprise

Horizon	Regime	N	Mean $\|\alpha^{raw}\|$	Mean $\|\alpha^{cal}\|$	Mean $\|\eta\|$	Cal. Alpha Share
H1	S0	54	0.739	0.711	0.090	0.888
	S1	84	0.722	0.693	0.104	0.870
	S2	78	0.386	0.360	0.108	0.770
H3	S0	52	0.800	0.683	0.168	0.802
	S1	84	0.759	0.637	0.205	0.757
	S2	78	0.392	0.303	0.234	0.564
H6	S0	50	0.855	0.605	0.331	0.647
	S1	83	0.798	0.557	0.314	0.640
	S2	78	0.432	0.278	0.326	0.460
H12	S0	46	0.916	0.363	0.642	0.361
	S1	81	0.691	0.343	0.448	0.434
	S2	78	0.393	0.187	0.433	0.302

The table reports the no-look-ahead calibrated decomposition of market-relative inflation surprise by dominant filtered TV-HMM regime at the forecast origin. Raw Forecast Alpha, α^{raw}, is the model forecast minus the market-based inflation benchmark. Calibrated Forecast Alpha, α^{cal}, rescales raw Forecast Alpha using the recursively estimated pass-through coefficient. Calibrated Residual Surprise, η, is market-relative inflation surprise minus calibrated Forecast Alpha. Calibrated Alpha Share is computed as $E[|\alpha^{cal}|]/(E[|\alpha^{cal}|] + E[|\eta|])$ within each horizon-regime cell.

out-of-sample benchmark hurdles before its forecasts are used in the market-relative decomposition framework.

Table 3 reports out-of-sample forecast accuracy relative to several benchmark alternatives. The benchmark set includes the T5YIE market proxy, a naive no-change forecast, an autoregressive persistence benchmark, and a Phillips-curve benchmark. This benchmark set is designed to evaluate the Compact Inflation Model against market-implied inflation pricing, simple inflation persistence, and a conventional macroeconomic inflation specification. Survey-based forecasts and broader machine-learning benchmark ensembles are not included in the main comparison because their forecast horizons, target definitions, and real-time availability conventions are not directly aligned with the market-relative Core PCE panel used in this study.

Across all reported horizons, the Compact Inflation Model produces the lowest RMSE among the benchmark alternatives. The Compact Inflation Model also produces the lowest MAE at each reported horizon.

5.4 Incremental Information Beyond Market Expectations

This subsection evaluates whether the model forecast contains information beyond that embedded in the market-based inflation benchmark. Rather than relying solely on the decomposition framework, incremental information is evaluated using forecast-encompassing regressions:

Table 3: Out-of-Sample Forecast Accuracy Relative to Basic Benchmarks

Horizon	Model	RMSE	MAE	Bias
1	T5YIE market proxy	0.8696	0.5326	-0.1909
1	Naive no-change	0.1656	0.1182	0.0069
1	AR(p)	0.2089	0.1502	0.0069
1	Phillips curve	0.2101	0.1511	0.0062
1	Compact Inflation Model	**0.1437**	0.1043	0.0173
3	T5YIE market proxy	0.8696	0.5326	-0.1909
3	Naive no-change	0.3385	0.2416	0.0210
3	AR(p)	0.3176	0.2226	0.0181
3	Phillips curve	0.3147	0.2247	0.0097
3	Compact Inflation Model	**0.2771**	0.1951	0.0453
6	T5YIE market proxy	0.8696	0.5326	-0.1909
6	Naive no-change	0.5325	0.3769	0.0459
6	AR(p)	0.4738	0.3271	0.0330
6	Phillips curve	0.4624	0.3207	0.0241
6	Compact Inflation Model	**0.4046**	0.2853	0.0733
12	T5YIE market proxy	0.8696	0.5326	-0.1909
12	Naive no-change	0.8976	0.6200	0.0974
12	AR(p)	0.7415	0.4751	0.0480
12	Phillips curve	0.7107	0.4785	0.0735
12	Compact Inflation Model	**0.6692**	0.4616	0.1348

Notes: Forecasts are evaluated out of sample. The recommended model is the paper's Compact Inflation Model, estimated recursively in real time. T5YIE is used as a market-based inflation-compensation proxy rather than a horizon-matched short-horizon expectation measure. Bold indicates the lowest RMSE within each horizon.

$$\pi_{t+h} = a_h + b_h \hat{\pi}^{mkt}_{t+h|t} + c_h \hat{\pi}^{model}_{t+h|t} + u_{t,h},$$

where π_{t+h} is realized inflation, $\hat{\pi}^{mkt}_{t+h|t}$ is the market-based inflation benchmark observed at the forecast origin, and $\hat{\pi}^{model}_{t+h|t}$ is the model forecast formed at the same date.

Table 4: Forecast-Encompassing Regressions Relative to Market Expectations

Horizon	Market Coef.	Model Coef.	Adj. R^2	N	Interpretation
1-month	0.0566*** (3.22)	0.9780*** (71.77)	0.9837	277	Both contain distinct information
3-month	0.2238*** (4.59)	0.9117*** (25.70)	0.9335	277	Both contain distinct information
6-month	0.3946*** (3.75)	0.8344*** (15.85)	0.8728	277	Both contain distinct information
12-month	0.7016*** (2.70)	0.6792*** (6.33)	0.6561	277	Both contain distinct information

Notes: The table reports forecast-encompassing regressions of realized inflation on the market-based inflation benchmark and the model-based forecast, estimated separately by horizon:

$$\pi_{t+h} = a_h + b_h \hat{\pi}^{mkt}_{t+h|t} + c_h \hat{\pi}^{model}_{t+h|t} + u_{t,h}.$$

T-statistics based on heteroskedasticity- and autocorrelation-consistent standard errors are reported in parentheses, with Newey–West lag length matched to the forecast horizon. A significant model coefficient indicates incremental predictive information beyond market expectations. A significant market coefficient indicates that the market benchmark retains predictive content beyond the model. If both coefficients are significant, neither forecast fully encompasses the other. If only the model coefficient is significant, the results are consistent with model encompassing of the market benchmark. The market benchmark source used here is `T5YIE_proxy`.

Table 4 reports the horizon-by-horizon forecast-encompassing regression results. The coefficient on the model forecast is positive at all horizons and statistically significant at the 1-, 3-, and 6-month horizons. The estimated coefficient declines with the forecast horizon, reaching its lowest value at the 12-month horizon. The model coefficient remains positive in all specifications and statistically significant at the 1-, 3-, and 6-month horizons.

Table 5 reports the standardized forecast-encompassing results. Across all horizons, the standardized model coefficient exceeds the standardized market coefficient. The increase in adjusted R^2 associated with adding the model forecast to a market-only specification is larger than the increase associated with adding the market benchmark to a model-only specification.

5.5 Residual Surprise Diagnostics

This subsection evaluates whether the calibrated residual component behaves as a residual surprise after the forecastable portion of market-relative inflation variation has been removed. The diagnostic analysis is performed using both the raw and calibrated decompositions.

In the raw decomposition, Residual Surprise is defined as the model forecast error, $\varepsilon^{raw}_{t,h}$. In the calibrated decomposition, Residual Surprise is defined as the residual component,

Table 5: Standardized Incremental-Information Comparisons Relative to Market Expectations

Horizon	Std. Market Coef.	Std. Model Coef.	Delta Adj. R^2 (Model |Market)	Delta Adj. R^2 (Market |Model)	N
1-month	0.0338*** (3.22)	1.0259*** (71.77)	0.6328	0.0006	277
3-month	0.1337*** (4.59)	0.9426*** (25.70)	0.5825	0.0115	277
6-month	0.2356*** (3.75)	0.8552*** (15.85)	0.5218	0.0392	277
12-month	0.4190*** (2.70)	0.6184*** (6.33)	0.3051	0.1394	277

Notes: The table reports horizon-by-horizon incremental-information comparisons using standardized forecasts.

Within each horizon, the market forecast and the model forecast are standardized to mean zero and unit variance before estimation.

Reported coefficients measure the marginal association of a one-standard-deviation change in each forecast, conditional on the other forecast.

T-statistics in parentheses use Newey–West standard errors, with lag length matched to the forecast horizon.

Delta Adj. R^2 (Model |Market) reports the increase in adjusted R-squared from adding the model forecast to a market-only regression.

Delta Adj. R^2 (Market |Model) reports the increase in adjusted R-squared from adding the market forecast to a model-only regression.

Market benchmark source: `T5YIE_proxy`.

$$\eta_{t,h} = s_{t,h} - \alpha^{cal}_{t,h}.$$

The diagnostic test regresses the relevant residual component on the corresponding Forecast Alpha measure. In the raw specification, raw Residual Surprise is regressed on raw Forecast Alpha. In the calibrated specification, calibrated Residual Surprise is regressed on calibrated Forecast Alpha. If the calibrated residual behaves as a residual component, the coefficient on calibrated Forecast Alpha should not be statistically distinguishable from zero. Inference is evaluated using robust forecast-evaluation procedures that accommodate overlapping horizons and serially correlated forecast errors [Diebold and Mariano, 1995, West, 1996, Clark and West, 2007].

These diagnostics are interpreted as forecast-evaluation and decomposition tests rather than causal tests. The objective is to determine whether Forecast Alpha contains incremental predictive information relative to the market benchmark and whether the calibrated residual component behaves as an unexplained residual surprise. The paper therefore does not interpret the decomposition as establishing Granger causality or structural causation among inflation, market expectations, and breakeven repricing.

Table 6 reports the orthogonality diagnostics by horizon. The raw decomposition rejects orthogonality at the 3-, 6-, and 12-month horizons. The calibrated decomposition does not reject orthogonality at conventional significance levels at any reported horizon.

The calibrated HAC p-values equal 0.219, 0.126, 0.106, and 0.742 at the 1-, 3-, 6-, and 12-month

Table 6: Residual-Surprise Orthogonality Diagnostics by Horizon

Horizon	N	Raw β	Raw HAC p	Cal. β	Cal. HAC p	Interpretation
H1	217	-0.021	0.165	0.020	0.219	Do not reject calibrated orthogonality
H3	215	-0.095	0.025	0.077	0.126	Calibration removes 5% rejection
H6	212	-0.188	0.005	0.170	0.106	Calibration removes 5% rejection
H12	206	-0.441	0.014	0.147	0.742	Calibration removes 5% rejection

The table reports residual-surprise orthogonality diagnostics by horizon. The raw regression relates raw Residual Surprise, $\varepsilon_{t,h}^{raw}$, to raw Forecast Alpha, $\alpha_{t,h}^{raw}$. The calibrated regression relates calibrated Residual Surprise, $\eta_{t,h}$, to calibrated Forecast Alpha, $\alpha_{t,h}^{cal}$. Regressions include an intercept and use heteroskedasticity- and autocorrelation-consistent standard errors with horizon-specific lag lengths. Failure to reject in the calibrated regression indicates that the calibrated residual is not systematically related to the calibrated forecastable component at conventional significance levels.

horizons, respectively.

5.6 Regime Transitions and Signal Stability

This subsection examines whether Forecast Alpha is concentrated in periods of regime transition or remains present across broader macroeconomic environments. Transition periods are identified using filtered TV-HMM posterior transition intensity, with high-intensity episodes interpreted as periods of elevated regime change.

Table 7: Transition vs. Stable Forecast Alpha Diagnostics

Horizon	$\lvert\alpha\rvert$ Trans.	$\lvert\alpha\rvert$ Stable	Ratio	Diff.	N_T	N_S
1-month	0.4024	0.5861	0.6864	-0.1838	69	207
3-month	0.4049	0.6177	0.6555	-0.2128	69	207
6-month	0.4413	0.6500	0.6790	-0.2087	69	207
12-month	0.3802	0.6694	0.5680	-0.2892	69	207

Notes: The table compares the magnitude of Forecast Alpha between transition episodes and stable episodes by forecast horizon. Forecast Alpha is defined as model forecast minus market expectation. Transition classification is based on derived transition intensity, defined as one minus the maximum filtered posterior state probability. Transition episodes are defined as observations with transition intensity at or above the 0.75 quantile of the full matched sample; all remaining observations are classified as stable. $|\alpha|$ denotes the mean absolute magnitude of Forecast Alpha within each bucket. Ratio is the transition-to-stable ratio of mean $|\alpha|$, and Diff. is the transition-minus-stable difference in mean $|\alpha|$. Market benchmark source: `T5YIE_proxy`.

Table 7 reports the transition diagnostics for raw Forecast Alpha. Across all four horizons, mean absolute raw Forecast Alpha is larger in stable periods than in transition periods. The transition-to-stable ratio remains below one at every horizon.

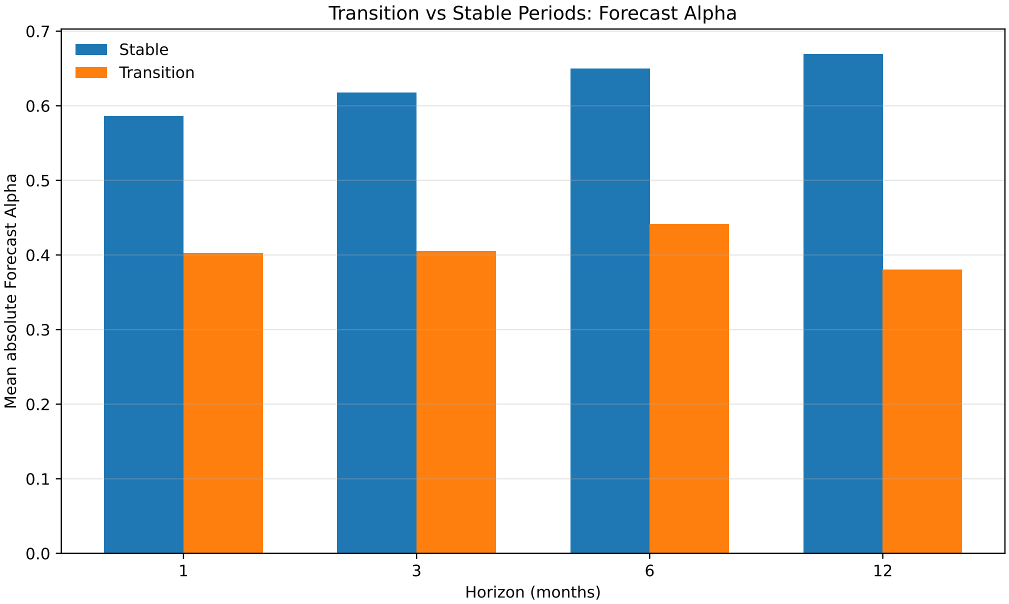

Figure 3: Raw Forecast Alpha in transition and stable episodes. Transition episodes are defined using filtered TV-HMM posterior transition intensity. Under the refreshed direct compact specification, mean absolute raw Forecast Alpha is consistently larger in stable periods than in transition periods across horizons.

6 Evaluation

This section evaluates the economic relevance of Forecast Alpha. Section 5 established that market-relative inflation surprises contain a forecastable component, that forecastability varies across macroeconomic regimes, and that the model forecast contains information beyond the market-based inflation benchmark. The remaining question is whether that information survives translation into inflation-sensitive market instruments.

The analysis focuses on breakeven inflation markets because the market benchmark used throughout the paper is derived from Treasury breakeven inflation. The objective is not to evaluate a fully optimized trading strategy. Instead, the goal is to determine whether Forecast Alpha corresponds to subsequent breakeven repricing and whether that relationship varies across regimes and implementation choices.

6.1 Interpretation of Results

The results reported in Section 5 can be summarized in five broad findings. First, a substantial portion of market-relative inflation surprise is associated with Forecast Alpha at short and intermediate horizons, although its relative importance declines as the forecast horizon increases. Second, forecastability varies across macroeconomic regimes, with meaningful differences in calibrated Alpha Share across states. Third, the Compact Inflation Model satisfies conventional out-of-sample benchmark hurdles relative to alternative forecasting approaches. Fourth, the model forecast contains information beyond that embedded in the market-based inflation benchmark. Finally, the calibrated decomposition produces a residual component that is substantially more independent of Forecast Alpha than the corresponding raw decomposition.

Taken together, these findings suggest that realized inflation outcomes contain information that is not fully reflected in contemporaneous market pricing. The remaining question is whether that information possesses economic relevance once translated into inflation-sensitive market instruments.

6.2 From Forecast Alpha to Market Relevance

The baseline implementation is evaluated through changes in 5-year breakeven inflation. This is the natural maturity-matched translation because the market benchmark in the forecasting system is the T5YIE-based proxy. The objective is narrow but economically meaningful: to determine whether the forecastable component isolated by the model corresponds to subsequent repricing in inflation-sensitive market instruments. This exercise is related to work linking inflation forecasts, inflation compensation, and asset-pricing outcomes, but the objective here is narrower: to test whether model–market disagreement has directional content for breakeven repricing rather than to estimate a complete asset-allocation strategy [Ang et al., 2007, Haubrich et al., 2012, Chernov and Mueller, 2012, D'Amico et al., 2018, Pesci et al., 2022].

The trading exercise uses the sign of raw Forecast Alpha as the ex ante directional signal. Because the estimated pass-through coefficient is positive throughout the calibration sample, calibration affects the magnitude of Forecast Alpha but generally does not alter its directional implication. The market-relevance exercise therefore evaluates the raw model–market disagreement directly.

The breakeven implementation uses the economically consistent sign convention for translating inflation forecast disagreement into breakeven exposure. Under that translation, raw Forecast Alpha from the direct compact model produces positive realized gross performance in simple breakeven-based rules. The purpose of the exercise is not to claim a fully optimized production strategy, but to show that the model's market-relative component contains information that survives translation into inflation-sensitive market prices.

6.3 Baseline Trading Results and 10-Year Robustness

Table 8 reports the preferred market-relevance rule in two implementations. The 5-year breakeven rule is the benchmark-consistent baseline because it is maturity-matched to the paper's T5YIE-based market anchor. The preferred baseline is the unthresholded 3-month raw Forecast Alpha rule. In that setting, the rule generates positive mean gross PnL, a positive ex post information ratio, and a hit rate above 60 percent. These results indicate that raw Forecast Alpha can be translated into economically usable breakeven signals in the maturity that most closely matches the market benchmark used throughout the paper.

Table 8: Forecast Alpha Signal Stability and Breakeven Market Transmission

Metric	2008–2022	2023–2025
Signal observations	180	36
Active trading observations	178	36
Corr(Forecast Alpha, Unexpected Inflation)	0.934	0.952
Mean |Forecast Alpha|	0.537	1.093
Corr(Forecast Alpha, Breakeven Move)	-0.307	-0.101
Hit Rate	64.6%	52.8%
Profit Factor	2.24	1.03

Notes: The table separates inflation signal quality from breakeven-market monetization for the preferred 3-month Forecast Alpha implementation. Corr(Forecast Alpha, Unexpected Inflation) measures the association between the model–market forecast differential and subsequent market-relative unexpected inflation. Corr(Forecast Alpha, Breakeven Move) measures the association between Forecast Alpha and subsequent realized 5Y breakeven repricing; the expected sign is negative because the trading payoff is defined as $-\text{Position} \times \Delta BE$. Hit Rate and Profit Factor are calculated from active 5Y breakeven trading observations. The 2026 observation is excluded because the year is incomplete. The post-2022 comparison is designed to distinguish preservation of the inflation signal from impairment in the breakeven-market transmission mechanism.

The same preferred signal is also evaluated using 10-year breakeven inflation. This longer-duration version is positive and produces stronger risk-adjusted performance with smaller drawdowns in the preferred specification. This matters because it shows that the signal's market relevance is not limited to the exact maturity of the T5YIE benchmark. Instead, the signal appears to carry information for a broader segment of breakeven repricing.

Taken together, Table 8 supports two conclusions. First, the 5-year implementation provides the natural benchmark-consistent test of whether raw Forecast Alpha survives into market prices. Second, the 10-year implementation shows that the signal's usefulness is not confined to that exact tenor. The economic value of Forecast Alpha is therefore best interpreted as maturity-robust market relevance rather than as a mechanical consequence of the T5YIE mapping.

6.4 State-Dependent Monetization

Table 9 reports regime-specific performance for the preferred trading rule in both the 5-year baseline and the 10-year robustness implementation. The regime breakdown shows that monetization is state-varying, but not state-exclusive. Gross performance remains positive across Regimes 0, 1, and 2 in the preferred baseline rule, which implies that the signal is not confined to a single narrow environment. At the same time, implementation quality differs meaningfully across states.

In the 5-year baseline, Regimes 0 and 1 generate the strongest average gross PnL and ex post information ratios, while Regime 2 remains positive but weaker on a risk-adjusted basis. The 10-year panel provides a useful cross-check by showing that the signal also remains positive across regimes in a longer-duration instrument. The main point is not that one regime alone determines the signal's value. Rather, the evidence indicates that the efficiency of implementation varies across macro

Table 9: Regime-Conditional Signal Monetization

Panel A. 5Y Breakeven Baseline

Regime	Active N	Hit Rate	Profit Factor	Reward/Risk	CVaR(95%)	Max DD
Regime 0	52	48.1%	2.39	0.99	-27.33	-50.00
Regime 1	84	67.9%	2.71	0.92	-38.20	-84.00
Regime 2	78	66.7%	1.39	0.33	-78.50	-67.00

Panel B. 10Y Breakeven Robustness

Regime	Active N	Hit Rate	Profit Factor	Reward/Risk	CVaR(95%)	Max DD
Regime 0	52	59.6%	2.25	1.03	-23.67	-45.00
Regime 1	84	65.5%	2.09	0.82	-40.80	-103.00
Regime 2	78	67.9%	2.10	1.06	-24.25	-47.00

Notes: The table reports regime-conditional signal monetization for the preferred 3-month Forecast Alpha trading rule. Regimes are assigned using the dominant filtered TV-HMM state at the forecast origin. Profit Factor is total gross gains divided by the absolute value of total gross losses within each regime. Reward/Risk is the annualized mean realized trading payoff divided by the standard deviation of realized trading payoffs. CVaR(95%) is the average payoff in the worst 5% of active observations. Payoff statistics are measured in basis points before transaction-cost adjustments. The 2026 observation is excluded because the year is incomplete.

environments.

This interpretation is consistent with the earlier results. Forecast Alpha is strongest at short and intermediate horizons, and the decomposition evidence shows that its relative importance varies across states. The monetization exercise extends those findings by showing that state dependence also appears in breakeven repricing. However, the relationship is not mechanical. Regimes in which Alpha Share is relatively larger are not necessarily the same regimes in which trading implementation is most efficient, because market translation also depends on instrument mapping, timing, turnover, and realized repricing dynamics.

The evidence therefore supports using Forecast Alpha as a portfolio-tilt variable rather than as a standalone active-risk engine. Positive gross implementation statistics indicate that the signal survives contact with market prices, but cross-regime variation shows that deployment quality is not uniform across states.

Table 10 evaluates whether Forecast Alpha remains informative during crisis and inflation-stress episodes identified in the editorial review. The table considers three windows: the 2008–09 global financial crisis and commodity-price shock, the 2020 COVID shock, and the 2021–22 post-COVID inflation surge. These windows are economically relevant because they correspond to episodes in which market-based inflation compensation and realized inflation dynamics were affected by severe financial stress, abrupt changes in demand and supply conditions, or unusually large inflation shocks. Panel A evaluates forecast accuracy relative to the market benchmark using matched observations of model forecasts, market-implied inflation expectations, and realized inflation. Panel B evaluates whether the preferred Forecast Alpha trading rule retains directional market content when translated into realized breakeven-market moves.

Table 10: Forecast and Market-Relevance Performance Across Inflation Stress Periods

Panel A. Forecast Accuracy Relative to the Market Benchmark

Stress Period	Horizon	N	Model RMSE	Market RMSE	RMSE Diff.	Mean Raw Alpha
2008–09 GFC / commodity shock	1	24	0.15	1.05	-0.89	0.23
2008–09 GFC / commodity shock	3	24	0.36	1.05	-0.69	0.29
2008–09 GFC / commodity shock	6	24	0.54	1.05	-0.51	0.42
2008–09 GFC / commodity shock	12	24	0.93	1.05	-0.12	0.95
2020 COVID shock	1	10	0.16	0.40	-0.24	0.02
2020 COVID shock	3	10	0.33	0.40	-0.07	0.07
2020 COVID shock	6	10	0.41	0.40	0.01	0.09
2020 COVID shock	12	10	0.42	0.40	0.02	0.33
2021–22 inflation surge	1	24	0.29	2.09	-1.80	1.67
2021–22 inflation surge	3	24	0.71	2.09	-1.39	1.40
2021–22 inflation surge	6	24	1.02	2.09	-1.07	1.21
2021–22 inflation surge	12	24	1.96	2.09	-0.13	0.27

Panel B. Preferred Forecast Alpha Trading-Rule Performance

Stress Period	Instrument	N	Gross PnL	Mean Monthly PnL	Stress IR	Hit Rate	Max DD
2008–09 GFC / commodity shock	10Y breakeven	24	234.00	9.75	1.00	70.8%	-136.00
2008–09 GFC / commodity shock	5Y breakeven	24	353.00	14.71	0.73	54.2%	-306.00
2020 COVID shock	10Y breakeven	10	72.00	7.20	1.04	60.0%	-47.00
2020 COVID shock	5Y breakeven	10	96.00	9.60	0.97	60.0%	-67.00
2021–22 inflation surge	10Y breakeven	24	35.00	1.46	0.29	58.3%	-66.00
2021–22 inflation surge	5Y breakeven	24	66.00	2.75	0.40	58.3%	-97.00

Notes: The table evaluates performance during editor-identified stress windows: the 2008–09 global financial crisis/commodity shock, the 2020 COVID shock, and the 2021–22 post-COVID inflation surge. Panel A reports forecast accuracy against realized inflation and the market-implied benchmark. RMSE Diff. is Model RMSE minus Market RMSE, so negative values indicate lower model forecast error than the market benchmark. Mean Raw Alpha is the average model–market forecast disagreement. Panel B reports the preferred unthresholded 3-month raw Forecast Alpha trading rule using realized breakeven-market moves. Gross PnL and Max DD are reported in basis points before transaction-cost adjustments. Stress IR is annualized from monthly gross PnL within each window and should be interpreted as descriptive stress-period evidence rather than a formal profitability test, particularly in short windows.

Panel A shows that the model generally remains informative during stress periods. During 2008–09, the model produces lower RMSE than the market benchmark at each reported horizon. During the 2020 COVID shock, the model improves on the market benchmark at the 1- and 3-month horizons and performs roughly in line with the benchmark at longer horizons. During the 2021–22 inflation surge, the model again produces lower RMSE than the market benchmark across all reported horizons. These results indicate that the model's incremental information is not confined to normal inflation environments; it remains visible during periods when inflation dynamics depart sharply from typical conditions.

Panel B provides complementary market-relevance evidence. The preferred raw Forecast Alpha trading rule generates positive gross PnL across the stress windows in both the 5-year and 10-year breakeven implementations. This does not establish a complete crisis-period trading strategy, since the results are gross of transaction costs and the stress windows are short. Instead, the evidence shows that Forecast Alpha retains economically meaningful directional content when mapped into inflation-sensitive market prices.

Taken together, the stress-period evidence supports a qualified interpretation. Forecast Alpha is not uniformly strong across every horizon and episode, but it remains informative in the periods most relevant to the editor's concern. The model improves on the market benchmark during the 2008–09 and 2021–22 inflation-stress windows and retains directional market relevance in the associated breakeven implementation. This reinforces the paper's interpretation of Forecast Alpha as a portfolio-tilt signal rather than a fully specified standalone trading strategy.

6.5 Candidate Specifications and Confirmation Effects

The broader candidate-rule comparison supports the same conclusion: the strongest overall implementation is the simple unthresholded 3-month raw Forecast Alpha rule. Confirmation rules that combine the 3-month and 6-month signals do not dominate the simpler baseline, even though some remain positive. Thresholding increases selectivity but does not improve on the preferred full-sample rule. The economic value of the framework therefore lies in the fact that a simple market translation of raw Forecast Alpha already survives into realized breakeven repricing.

This also sharpens the interpretation of the state results. Regime-filtered and posterior-weighted variants are useful diagnostically, but they do not dominate the unconstrained preferred rule in the current implementation. This suggests that the framework's economic value does not depend on isolating one single best state. Instead, the signal appears broadly usable in breakeven-market implementation, with the strongest average implementation arising from the simple full-sample 3-month specification.

6.6 Directional Rules and Position Sizing

Simple directional rules outperform the current variance-scaled specification. This does not mean that risk management is unimportant. Rather, it suggests that the current sizing rule is not yet rich enough to improve on the directional signal. The forecasting results indicate that the decomposition is strongest at shorter horizons and weaker at 12 months; the trading results add that even when direction is informative, a simple variance denominator may be too blunt to convert that information into better position sizes.

Accordingly, the framework's strongest current economic value lies in trade selection and conditional directional exposure rather than precise volatility-based scaling. In portfolio terms, the model appears more useful for deciding when to lean into or away from inflation-sensitive exposure than for determining an exact continuous position size from a variance-only rule. This points to a natural next step for implementation: improve the sizing rule without replacing the underlying directional signal.

6.7 Interpretation

Taken together, the economic evaluation supports three conclusions. First, the model's incremental information transfers to the investment domain under the breakeven mapping. Second, the strongest baseline implementation is the simple unthresholded 3-month raw Forecast Alpha rule applied to 5-year breakevens, which is the maturity most closely aligned with the paper's market expectation proxy. Third, monetization is state-varying but not state-exclusive: the signal remains positive across regimes on a gross basis, although its efficiency differs across macro environments.

This interpretation aligns closely with the framework introduced in Figure 1. Market-relative inflation surprises can be decomposed into a forecastable component and a residual component. The forecastable component varies across regimes, survives translation into breakeven-market signals, and exhibits positive gross implementation characteristics across multiple specifications and stress periods. The evidence is therefore most consistent with interpreting Forecast Alpha as a regime-aware portfolio-tilt variable whose usefulness depends on both horizon and state rather than as a uniformly scaled active-risk engine.

7 Conclusion

This study develops a real-time framework for evaluating whether inflation contains forecastable information beyond that embedded in market pricing. The framework combines a direct horizon-specific inflation forecasting model, a filtered time-varying Hidden Markov Model (TV-HMM) used for regime classification, and a market-relative framework that links model–market disagreement to realized market-relative inflation outcomes through a decomposition into Forecast Alpha and Residual Surprise. In this structure, raw Forecast Alpha measures the model's ex ante deviation from the market-based inflation benchmark, calibrated Forecast Alpha measures the historically realized component of that deviation, and calibrated Residual Surprise captures the remaining market-relative inflation surprise after the pass-through adjustment.

Several conclusions emerge.

First, inflation contains a non-trivial forecastable component relative to market pricing, especially at short and intermediate horizons. Calibrated Forecast Alpha accounts for a substantial share of market-relative unexpected inflation at shorter horizons, while Residual Surprise becomes increasingly important as the horizon lengthens. This pattern implies that the model captures structured variation in market-relative inflation surprise most clearly at tactical horizons, while residual uncertainty becomes more important over longer horizons.

Second, inflation forecastability is state dependent. The regime-conditioned decomposition shows that both the magnitude of Forecast Alpha and the share of market-relative inflation surprise attributable to the calibrated component vary across regimes. This state dependence is economically important. It implies that inflation dynamics are not governed by a single stable forecasting environment, and that the value of inflation forecasts depends on prevailing macroeconomic conditions.

Third, the forecast engine satisfies conventional out-of-sample validation hurdles. The direct compact model improves upon benchmark alternatives on a common recursive real-time sample, and formal predictive-accuracy tests indicate that these gains are not solely artifacts of sample construction.

Fourth, the model contains information beyond market pricing. Ex ante deviations between the model forecast and the market-based inflation benchmark are associated with subsequent realized market-relative inflation outcomes. This result supports the interpretation of raw Forecast Alpha as a meaningful measure of incremental information rather than a mechanical by-product of the decomposition identity.

Fifth, the market-relevance exercises indicate that raw Forecast Alpha survives translation into inflation-sensitive market signals. The baseline 5-year implementation, which is the natural maturity match to the paper's T5YIE-based market anchor, produces positive gross implementation statistics under the preferred unthresholded 3-month rule. A parallel 10-year implementation is also positive and indicates that the signal's usefulness is not limited to the exact tenor of the market benchmark. The appropriate interpretation is therefore not that the paper has produced a fully optimized trading strategy, but that the forecastable component identified by the model retains measurable economic relevance once translated into breakeven-market signals.

Several clarifications are important.

The decomposition of market-relative unexpected inflation into Forecast Alpha and Residual Surprise is an accounting structure whose economic interpretation depends on the forecast model, the recursive calibration procedure, and the market benchmark. Forecast Alpha should therefore be interpreted as incremental information relative to the forecasting system and market anchor used in this study, not as a definitive partition of all inflation uncertainty.

The use of Treasury breakeven inflation as the market benchmark introduces additional qualification. Breakeven rates embed inflation risk premia, liquidity effects, and other pricing distortions. Moreover, because a clean horizon-matched market-implied inflation series is not available for all horizons considered here, the paper uses a common market-based inflation anchor rather than a fully horizon-specific market forecast curve. Accordingly, Forecast Alpha should be interpreted as a model-based deviation from market pricing, not as a pure estimate of disagreement with true expected inflation.

Horizon heterogeneity remains central to interpretation. The strongest evidence concerns short- and intermediate-horizon predictability, while longer-horizon results are weaker and less stable. This implies that the framework is best viewed as a tool for tactical inflation monitoring, market interpretation, and conditional portfolio tilting rather than as evidence of broad long-horizon forecast dominance.

The trading evidence should also be interpreted with appropriate discipline. The market-relevance exercises are intended to test whether raw Forecast Alpha maps into subsequent breakeven repricing in a directionally useful way, not to deliver a full excess-return backtest with carry, roll, funding, transaction costs, and portfolio-construction frictions that would be required in a fully implemented investment strategy. For that reason, the economic contribution of the paper lies less in claiming a finished trading strategy than in showing that the model's market-relative component is usable for conditional positioning and market interpretation.

Taken together, the findings support a qualified but meaningful conclusion: inflation contains a regime-dependent forecastable component beyond what is embedded in market pricing, that component is strongest at shorter decision horizons, and it has economically relevant implications for inflation-sensitive markets. More broadly, the paper's contribution is not a particular forecasting model or trading rule, but a real-time framework for identifying, measuring, and evaluating forecastable information relative to market expectations. By separating forecastable and non-forecastable components of realized outcomes, the framework provides a mechanism for linking forecasting, market expectations, and portfolio implementation within a unified empirical structure.

7.1 Limitations and Future Research

The analysis has several limitations. The market benchmark is based on Treasury breakeven inflation rather than a clean horizon-specific Core PCE expectations curve. The market-relevance exercises are not intended to represent a fully optimized excess-return strategy, and the regime-conditioning layer should be evaluated further in future real-time extensions.

Future research may extend this framework by decomposing breakeven inflation into expectation and risk-premium components, constructing alternative market-based inflation anchors, refining the mapping from Forecast Alpha into signal sizing, and examining whether the same regime-dependent predictability structure appears across alternative inflation measures and international settings. A natural next step is to extend the framework beyond directional signals toward risk-aware allocation schemes that translate Forecast Alpha into portfolio weights while preserving the real-time discipline that underlies its identification and evaluation. More generally, future work may explore how forecastable information extracted from market-relative outcomes can be incorporated systematically into portfolio construction and risk management decisions.

Although the framework is developed in the context of inflation, the underlying logic is not inflation-specific. Any economic or financial variable for which market expectations can be observed or approximated may be evaluated through a similar decomposition. The broader contribution is therefore methodological: a general approach for determining whether model-based information contains economically meaningful content beyond market expectations, how that information varies across regimes, and whether it survives translation into real-world decision environments.

8 Acknowledgements

Acknowledgments. The author used ChatGPT (OpenAI) for coding assistance. The author specified the intended functionality, modeling objectives, and analytical requirements, and ChatGPT was used to draft portions of the implementation. All code was reviewed, tested, revised, and validated by the author. The author was solely responsible for the study design, variable construction, model specification, backtesting protocol, empirical verification, interpretation of findings, and final manuscript content.

9 Disclosure Statement

The author reports no conflicts of interest.

10 Funding

No external funding was received for this research.

References

Martin M. Andreasen, Jens H. E. Christensen, and Simon Riddell. The TIPS liquidity premium. *Review of Finance*, 25(6):1639–1675, 2021. doi: 10.1093/rof/rfab025. Identifies and quantifies the liquidity premium embedded in TIPS yields using no-arbitrage term structure models.

Andrew Ang and Allan Timmermann. Regime changes and financial markets. *Annual Review of Financial Economics*, 4:313–337, 2012. doi: 10.1146/annurev-financial-110311-101808.

Andrew Ang, Geert Bekaert, and Min Wei. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, 54(4):1163–1212, 2007. doi: 10.1016/j.jmoneco.2006.04.006.

Andrew Atkeson and Lee E. Ohanian. Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(1):2–11, 2001. Influential paper showing that naive random-walk models often outperform Phillips curve forecasts.

Jushan Bai and Pierre Perron. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22, 2003. doi: 10.1002/jae.659. Provides methods for detecting and dating structural breaks, relevant to regime identification.

Michael D. Bauer and Glenn D. Rudebusch. Resolving the spanning puzzle in macro-finance term structure models. *Review of Finance*, 21(2):511–553, 2017. doi: 10.1093/rof/rfw044. Examines the reliability of market-based inflation expectations and their informational content.

Francesco Bianchi. Regime switches, agents' beliefs, and post-world war ii u.s. macroeconomic dynamics. *The Review of Economic Studies*, 80(2):463–490, 2013. doi: 10.1093/restud/rds032.

Mikhail Chernov and Philippe Mueller. The term structure of inflation expectations. *Journal of Financial Economics*, 106(2):367–394, 2012. doi: 10.1016/j.jfineco.2012.06.004.

Richard Clarida, Jordi Galí, and Mark Gertler. Monetary policy rules and macroeconomic stability: Evidence and some theory. *The Quarterly Journal of Economics*, 115(1):147–180, 2000. doi: 10.1162/003355300554692.

Todd E. Clark and Kenneth D. West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311, 2007. doi: 10.1016/j.jeconom.2006.05.023. Develops tests for forecast accuracy comparison in nested model settings.

Timothy Cogley and Thomas J. Sargent. Drifts and volatilities: Monetary policies and outcomes in the post wwii u.s. *Review of Economic Dynamics*, 8(2):262–302, 2005. doi: 10.1016/j.red.2004.10.009.

Dean Croushore. Forecasting with real-time data: What has changed since the early 1990s? *Annual Review of Economics*, 3:267–286, 2011. doi: 10.1146/annurev.economics.050708.142927. Reviews developments in real-time data forecasting over two decades.

Stefania D'Amico, Don H. Kim, and Min Wei. Tips from TIPS: The informational content of treasury inflation-protected security prices. *Journal of Financial and Quantitative Analysis*, pages 1–35, 2018.

Marco Del Negro and Frank Schorfheide. DSGE model-based forecasting. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2, pages 57–140. Elsevier, 2013.

Francis X. Diebold and Roberto S. Mariano. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263, 1995. doi: 10.1080/07350015.1995.10524599. Foundational test for comparing predictive accuracy of competing forecasts.

Francis X. Diebold, Joon-Haeng Lee, and Gretchen C. Weinbach. Regime switching with time-varying transition probabilities. In Colin P. Hargreaves, editor, *Nonstationary Time Series Analysis and Cointegration*, pages 283–302. Oxford University Press, Oxford, 1994.

Jon Faust and Jonathan H. Wright. Forecasting inflation. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2, pages 2–56. Elsevier, 2013. doi: 10.1016/B978-0-444-53683-9.00001-3.

Andrew J. Filardo. Business-cycle phases and their transitional dynamics. *Journal of Business & Economic Statistics*, 12(3):299–308, 1994.

Milton Friedman. Nobel lecture: Inflation and unemployment. *Journal of Political Economy*, 85(3): 451–472, 1977. doi: 10.1086/260579. Lecture to the memory of Alfred Nobel, delivered December 13, 1976. Awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, 1976.

Lilit Gabrielyan. Forecasting inflation using Phillips curve. *Social Science Research Network*, 2016. Reviews Phillips curve specifications for inflation forecasting.

Philippe Goulet Coulombe, Maxime Leroux, Dalibor Stevanovic, and Stéphane Surprenant. How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5): 920–964, 2022. doi: 10.1002/jae.2910.

Jan J. J. Groen, Richard Paap, and Francesco Ravazzolo. Real-time inflation forecasting in a changing world. *Social Science Research Network*, 2012. doi: 10.2139/ssrn.1465985. Develops real-time inflation forecasting methods robust to structural change.

James D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.

Joseph G. Haubrich, George Pennacchi, and Peter Ritchken. Inflation expectations, real rates, and risk premia: Evidence from inflation swaps. *The Review of Financial Studies*, 25(5):1588–1629, 2012. doi: 10.1093/rfs/hhs003.

Yingying Huang, Paul Luk, and Jinsong Zhao. China's inflation forecasting in a data-rich environment: Based on machine learning algorithms. *Applied Economics*, 2024. doi: 10.1080/00036846.2024.2322572. Evaluates machine learning algorithms for inflation forecasting in a data-rich environment.

Peter Lund Jørgensen and Kevin J. Lansing. Anchored inflation expectations and the slope of the Phillips curve. *Federal Reserve Bank of San Francisco Working Paper*, 2019. doi: 10.24148/WP2019-27. Examines how anchored inflation expectations affect the slope of the Phillips curve.

Florian Kajuth and Sebastian Watzka. Inflation expectations from index-linked bonds: Correcting for liquidity and inflation risk premia. *The Quarterly Review of Economics and Finance*, 51(3): 225–235, 2011. doi: 10.1016/j.qref.2011.03.004. Corrects breakeven inflation for liquidity and inflation risk premia using a term structure model.

Bryan T. Kelly, Seth Pruitt, and Yinan Su. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524, 2019.

Chang-Jin Kim and Charles R. Nelson. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press, Cambridge, MA, 1999.

Anders Bredahl Kock and Marcelo C. Medeiros. Penalized time series regression. In *Handbook of Statistics*. 2020. doi: 10.1007/978-3-030-31150-6_7. Provides theoretical foundations for penalized (LASSO/Elastic Net) regression in time-series settings.

Gary Koop and Dimitris Korobilis. Large time-varying parameter VARs. *Journal of Econometrics*, 177(2):185–198, 2013. doi: 10.1016/j.jeconom.2013.04.007.

Gary Koop and Simon Potter. Estimation and forecasting in models with multiple breaks. *The Review of Economic Studies*, 74(3):763–789, 2007.

Jiahan Li and Weiye Chen. Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models. *Social Science Research Network*, 2014. doi: 10.2139/ssrn.2410214. Compares LASSO-based approaches with dynamic factor models for macroeconomic time series forecasting.

Joelle Liebermann. Real-time forecasting in a data-rich environment. 2012. Examines the use of large datasets for real-time macroeconomic forecasting.

Luis F. Martins and Vasco J. Gabriel. Forecasting inflation with the new Keynesian Phillips curve: Frequency matters. *Social Science Research Network*, 2020. doi: 10.2139/ssrn.3594397. Shows that data frequency affects the forecasting performance of New Keynesian Phillips curve models.

Marcelo C. Medeiros and Eduardo F. Mendes. ℓ_1-regularization of high-dimensional time-series models with flexible innovations. *Journal of Econometrics*, 191(1):255–271, 2016. doi: 10.1016/

j.jeconom.2015.10.011. Establishes consistency and model selection properties of LASSO for high-dimensional time-series models.

Marcelo C. Medeiros, Gabriel F. R. Vasconcelos, Álvaro Veiga, and Eduardo Zilberman. Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics*, 39(1):98–119, 2021.

Andrea A. Naghi, Eoghan O'Neill, and Martina Danielova Zaharieva. The benefits of forecasting inflation with machine learning: New evidence. *Journal of Applied Econometrics*, 39(7):1321–1331, aug 2024. doi: 10.1002/jae.3088. URL `https://doi.org/10.1002/jae.3088`.

Charles R. Nelson and G. William Schwert. Short-term interest rates as predictors of inflation: On testing the hypothesis that the real rate of interest is constant. *American Economic Review*, 67 (3):478–486, 1977.

Onder Ozgur and Levent Aydin. Inflation forecasting in an emerging economy: Selecting variables with machine learning algorithms. *International Journal of Emerging Markets*, 17(3):741–764, 2022. doi: 10.1108/IJOEM-05-2020-0577. Uses machine learning for variable selection in inflation forecasting in an emerging market context.

Valentina Pesci, Sandra Paterlini, and Emanuele Taufer. Inflation forecasts and European asset returns: A regime-switching approach. *Journal of Risk and Financial Management*, 15(10):475, 2022. doi: 10.3390/jrfm15100475. Applies a regime-switching approach to link inflation forecasts with asset return dynamics.

A. W. Phillips. The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861–1957. *Economica*, 25(100):283–299, 1958. doi: 10.1111/j.1468-0335.1958.tb00003.x.

Giorgio E. Primiceri. Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72(3):821–852, 2005.

Tim Robinson, Alan Stone, and Marileze van Zyl. The real-time forecasting performance of Phillips curves. 2003. Evaluates the real-time forecasting performance of various Phillips curve specifications.

Shuping Shi. Bayesian model averaging under regime switching with application to cyclical macro variable forecasting. *Journal of Forecasting*, 35(8):739–750, 2016. doi: 10.1002/for.2375. Combines Bayesian model averaging with regime-switching for macroeconomic forecasting.

Christopher A. Sims and Tao Zha. Were there regime switches in u.s. monetary policy? *American Economic Review*, 96(1):54–81, 2006. doi: 10.1257/000282806776157678.

James H. Stock and Mark W. Watson. Forecasting inflation. *Journal of Monetary Economics*, 44 (2):293–335, 1999.

James H. Stock and Mark W. Watson. Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking*, 39(s1):3–33, 2007.

James H. Stock and Mark W. Watson. Modeling inflation after the crisis. *Jackson Hole Economic Policy Symposium Proceedings*, pages 173–220, 2010.

James H. Stock and Mark W. Watson. Core inflation and trend inflation. *The Review of Economics and Statistics*, 98(4):770–784, 2016. doi: 10.1162/REST_a_00608.

Robert Tibshirani. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996. doi: 10.1111/j.2517-6161.1996.tb02080.x. Foundational LASSO paper; provides theoretical basis for the Elastic Net used in the paper.

Michael Weber, Yuriy Gorodnichenko, and Olivier Coibion. The expected, perceived, and realized inflation of us households before and during the covid19 pandemic. *IMF Economic Review*, 71(1):326–368, aug 2023. doi: 10.1057/s41308-022-00175-7. URL `https://doi.org/10.1057/s41308-022-00175-7`.

Kenneth D. West. Asymptotic inference about predictive ability. *Econometrica*, 64(5):1067–1084, 1996. doi: 10.2307/2171956. Provides asymptotic theory for evaluating forecast accuracy; complements Diebold-Mariano testing.

Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005. doi: 10.1111/j.1467-9868.2005.00503.x. Original Elastic Net paper introducing the combined L1/L2 penalty for regularization and variable selection.

Appendix

A Forecast Model Validation and Robustness

This appendix reports validation results for the paper's forecasting engine. The objective is to verify that the selected model is competitive relative to a persistence benchmark, stable across samples, economically interpretable, and robust across regimes.

A.1 Candidate Predictor Universe

Table 11 reports the complete set of candidate variables available to the direct compact forecasting model. The specification combines inflation persistence measures with a limited set of labor-market, monetary-policy, interest-rate, producer-price, commodity-price, and survey-based indicators. Derived features used in estimation include lags, changes, rolling statistics, and selected spread variables constructed from the underlying series listed below.

Table 11: Variables Used in the Direct Compact Forecasting Model

Category	Variable	FRED Series ID
Inflation	Core PCE Price Index	PCEPILFE
Labor Market	Unemployment Rate	UNRATE
Monetary Policy	Effective Federal Funds Rate	DFF
Interest Rates	2-Year Treasury Yield	DGS2
Producer Prices	Producer Price Index Commodities	PPIACO
Commodity Prices	WTI Crude Oil Spot Price	WTISPLC
Survey-Based	University of Michigan Consumer Sentiment	MICH

Notes: The forecasting model combines inflation persistence measures derived from Core PCE inflation with a compact set of exogenous macroeconomic and financial indicators. Engineered model inputs include lagged inflation terms, rolling statistics, changes in predictor variables, and selected spread measures constructed from the series listed above.

A.2 Common-Sample Comparison

Tables 12A and 12B compare the direct compact model with the ARIMA benchmark on the exact overlapping forecast dates at each horizon.

Table 12A: Common-Sample Forecast Comparison with ARIMA: RMSE and MAE

Horizon	N	RMSE ARIMA	RMSE Direct	Δ RMSE	MAE ARIMA	MAE Direct	Δ MAE
1	433	0.1539	0.1437	-0.0102	0.1077	0.1043	-0.0034
3	431	0.2937	0.2771	-0.0166	0.2034	0.1951	-0.0084
6	428	0.4543	0.4046	-0.0497	0.3150	0.2853	-0.0296
12	422	0.7667	0.6692	-0.0974	0.5310	0.4616	-0.0694

The table reports common-sample forecast performance for the ARIMA benchmark and the direct compact model. Δ RMSE and Δ MAE are computed as Direct minus ARIMA, so negative values indicate superior performance by the direct compact model.

The common-sample results confirm that the direct compact model's gains are not driven by differences in forecast coverage. Relative to ARIMA, the direct compact model reduces RMSE and MAE at every horizon on the overlapping forecast sample.

Table 12B: Common-Sample Forecast Comparison with ARIMA: Bias

Horizon	N	Bias ARIMA	Bias Direct	Δ Bias
1	433	0.0094	0.0173	0.0079
3	431	0.0280	0.0453	0.0172
6	428	0.0586	0.0733	0.0147
12	422	0.1278	0.1348	0.0070

The table reports common-sample forecast bias for the ARIMA benchmark and the direct compact model on the exact overlapping forecast dates. Δ Bias is computed as Direct minus ARIMA, so positive values indicate greater upward bias in the direct compact model relative to ARIMA.

Bias differences are modest relative to the error-level improvements, although the direct compact model exhibits somewhat higher positive bias than ARIMA at several horizons.

A.3 Predictive-Accuracy Tests

Table 13 reports Diebold–Mariano tests of equal predictive accuracy between the ARIMA + Exog. specification and ARIMA.

Table 13: Diebold–Mariano Tests versus ARIMA

Horizon	Loss	N	Mean loss ARIMA	Mean loss ARIMA + Exog.	Δ loss	DM stat	p-value
1	AE	433	0.1077	0.1043	-0.0034	-1.3134	0.1891
1	SE	433	0.0237	0.0207	-0.0030	-2.3677	0.0179
3	AE	431	0.2034	0.1951	-0.0084	-1.1767	0.2393
3	SE	431	0.0863	0.0768	-0.0095	-1.7765	0.0756
6	AE	428	0.3150	0.2853	-0.0296	-2.1386	0.0325
6	SE	428	0.2064	0.1637	-0.0427	-2.4849	0.0130
12	AE	422	0.5310	0.4616	-0.0694	-2.7218	0.0065
12	SE	422	0.5878	0.4479	-0.1399	-2.3184	0.0204

Diebold–Mariano tests of equal predictive accuracy are reported on the common sample. AE denotes absolute error and SE denotes squared error. Δ loss is computed as ARIMA+Exog. minus ARIMA, so negative values indicate lower loss for the ARIMA+Exog. specification.

These tests assess whether the forecast-error improvements are statistically distinguishable from zero.

A.4 Feature Stability and Interpretability

Table 14 summarizes the frequency with which predictors enter the model across recursive refits.

Table 14: Feature Stability Summary

Horizon	Feature	Nonzero freq.	Dominant sign	Mean $\|\beta\|$
1	infl_lag1	1.00	+	0.5809
1	infl_lag3	1.00	+	0.5303
1	infl_roll3_mean	1.00	+	0.4761
1	infl_chg3	1.00	+	0.1312
1	mich	1.00	+	0.0968
1	infl_chg1	1.00	+	0.0921
1	infl_lag2	0.93	+	0.2045
1	oil_chg3	0.87	+	0.0147
3	infl_lag1	1.00	+	0.5547
3	infl_lag3	1.00	+	0.4847
3	infl_roll3_mean	1.00	+	0.4390
3	mich	1.00	+	0.1770
3	infl_chg3	1.00	+	0.1106
3	infl_chg1	1.00	+	0.0841
3	infl_roll6_std	0.97	+	0.0515
3	infl_lag6	0.82	-	0.1501
6	infl_lag1	1.00	+	0.5719
6	infl_lag3	1.00	+	0.4933
6	infl_roll3_mean	1.00	+	0.4623
6	mich	1.00	+	0.2336
6	infl_chg3	1.00	+	0.1629
6	infl_roll6_std	1.00	+	0.1068
6	infl_chg1	1.00	+	0.0612
6	infl_lag2	0.86	+	0.2236
12	infl_lag1	1.00	+	0.4176
12	infl_lag3	1.00	+	0.3290
12	infl_roll6_std	1.00	+	0.2591
12	infl_chg3	1.00	+	0.1478
12	mich	1.00	+	0.2413
12	infl_chg1	1.00	+	0.0494
12	infl_roll3_mean	0.98	+	0.3338
12	infl_lag2	0.88	+	0.2003

Top retained predictors by recursive nonzero frequency. The table reports the most stable features in the direct compact model by horizon.

Table 15 reports whether the dominant signs of the retained predictors align with standard economic intuition.

Table 15: Economic Interpretability Summary

Horizon	Feature	Expected sign	Dominant sign	Alignment
1	infl_lag1	positive	positive	aligned
1	infl_lag3	positive	positive	aligned
1	infl_roll3_mean	positive	positive	aligned
1	infl_chg3	positive	positive	aligned
1	mich	positive	positive	aligned
1	infl_chg1	positive	positive	aligned
1	infl_lag2	positive	positive	aligned
1	oil_chg3	positive	positive	aligned
3	infl_lag1	positive	positive	aligned
3	infl_lag3	positive	positive	aligned
3	infl_roll3_mean	positive	positive	aligned
3	mich	positive	positive	aligned
3	infl_chg3	positive	positive	aligned
3	infl_chg1	positive	positive	aligned
3	infl_roll6_std	ambiguous	positive	not_scored
3	infl_lag6	positive	positive	aligned
6	infl_lag1	positive	positive	aligned
6	infl_lag3	positive	positive	aligned
6	infl_roll3_mean	positive	positive	aligned
6	mich	positive	positive	aligned
6	infl_chg3	positive	positive	aligned
6	infl_roll6_std	ambiguous	positive	not_scored
6	infl_chg1	positive	positive	aligned
6	infl_lag2	positive	positive	aligned
12	infl_lag1	positive	positive	aligned
12	infl_lag3	positive	positive	aligned
12	infl_roll6_std	ambiguous	positive	not_scored
12	infl_chg3	positive	positive	aligned
12	mich	positive	positive	aligned
12	infl_chg1	positive	positive	aligned
12	infl_roll3_mean	positive	positive	aligned
12	infl_lag2	positive	positive	aligned

Economic interpretability audit for the most stable retained features. Alignment indicates whether the dominant estimated sign accords with standard economic intuition.

The feature-stability results indicate that the forecasting specification is not being driven by a rapidly changing or opaque predictor set. Across horizons, the most persistent inputs are inflation-persistence variables, including lagged inflation and rolling inflation summaries, while a smaller group of exogenous variables—most notably survey expectations and selected macro-financial controls—enters more selectively. The interpretability audit in Table 15 is broadly consistent with that pattern: the dominant signs for the stable retained predictors generally accord with standard economic intuition, while the few variables with less clear directional priors are explicitly treated as ambiguous rather than forced into a stronger interpretation. Taken together, Tables 14 and 15 support the view that the selected specification combines a stable persistence block with a limited and economically interpretable set of exogenous predictors.

A.5 Temporal Robustness

Table 16 reports subsample forecast performance.

Table 16: Subsample Forecast Robustness

Horizon	Subsample	N	RMSE ARIMA	RMSE Direct	Δ RMSE	MAE ARIMA	MAE Direct
1	Pre-2000	120	0.1332	0.1307	-0.0025	0.1091	0.1077
1	2000–2019	240	0.1410	0.1243	-0.0167	0.0936	0.0873
1	Post-2020	73	0.2143	0.2089	-0.0053	0.1517	0.1546
3	Pre-2000	118	0.2074	0.2344	0.0270	0.1650	0.1915
3	2000–2019	240	0.2445	0.2091	-0.0354	0.1777	0.1534
3	Post-2020	73	0.4933	0.4699	-0.0234	0.3502	0.3377
6	Pre-2000	115	0.3368	0.3525	0.0157	0.2737	0.2879
6	2000–2019	240	0.3534	0.3169	-0.0364	0.2576	0.2363
6	Post-2020	73	0.7879	0.6586	-0.1294	0.5685	0.4423
12	Pre-2000	109	0.5774	0.5152	-0.0622	0.5013	0.4646
12	2000–2019	240	0.4980	0.4748	-0.0231	0.3864	0.3572
12	Post-2020	73	1.4439	1.2048	-0.2391	1.0507	0.8001

Subsample robustness for the common forecast sample. Δ RMSE is computed as Direct minus ARIMA; negative values indicate superior performance by the direct compact model.

Figure 4 plots rolling RMSE differences between the selected specification and ARIMA. Negative values indicate windows in which the selected specification performs better.

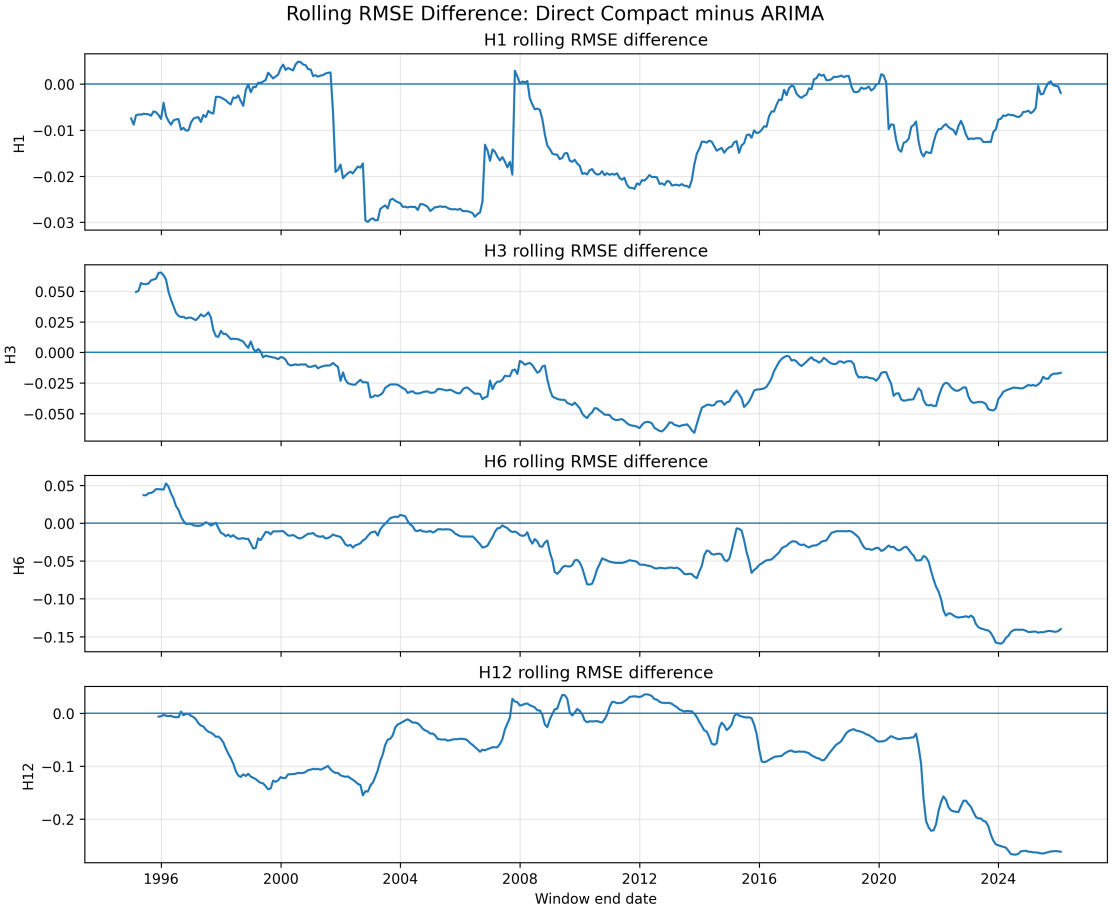

Figure 4: Rolling RMSE Difference: Direct Compact minus ARIMA

Negative values indicate rolling windows in which the direct compact model outperforms the ARIMA benchmark.

The temporal-robustness evidence is favorable but not uniform. Table 16 shows that the model's gains relative to ARIMA are strongest in the 2000–2019 and post-2020 subsamples, particularly at intermediate and longer horizons, while the pre-2000 results are more mixed. Figure 4 reinforces that interpretation: rolling RMSE differences are not negative in every window, but they are predominantly negative over substantial portions of the evaluation sample, with the clearest and most sustained improvements appearing at the 6- and 12-month horizons. The results therefore indicate that the model's advantage is not confined to a single historical episode, even though the magnitude of the gains varies over time.

A.6 Regime Robustness

Table 17 reports forecast performance by horizon and regime.

Table 17: Regime-Specific Forecast Robustness

Horizon	Regime	N	RMSE ARIMA	RMSE Direct	Δ RMSE	MAE ARIMA	MAE Direct
1	0	126	0.1419	0.1300	-0.0118	0.1029	0.0967
1	1	193	0.1541	0.1418	-0.0122	0.1079	0.1064
1	2	113	0.1643	0.1587	-0.0056	0.1109	0.1076
3	0	126	0.2405	0.2264	-0.0141	0.1746	0.1645
3	1	192	0.2629	0.2410	-0.0219	0.2001	0.1913
3	2	112	0.3833	0.3695	-0.0138	0.2397	0.2336
6	0	126	0.3902	0.3022	-0.0880	0.2745	0.2268
6	1	189	0.4045	0.3611	-0.0434	0.3005	0.2827
6	2	112	0.5817	0.5471	-0.0347	0.3826	0.3525
12	0	124	0.8075	0.6851	-0.1224	0.5579	0.4473
12	1	189	0.6524	0.5805	-0.0719	0.4918	0.4469
12	2	108	0.8956	0.7881	-0.1075	0.5706	0.5067

Forecast robustness by TV-HMM regime on the common sample. Δ RMSE is computed as Direct minus ARIMA; negative values indicate superior performance by the direct compact model.

The regime-specific results show that the selected specification remains competitive across the regime distribution rather than benefiting from a single narrow state. The RMSE differences are favorable across the reported regime-horizon cells, with the improvement generally becoming more pronounced at the 6- and 12-month horizons than at the 1- and 3-month horizons. This pattern is consistent with the broader validation evidence: the model's gains are broad-based, but they are economically more meaningful at intermediate and longer horizons than at the very shortest horizon.

A.7 Summary

Taken together, the validation results support the use of the selected forecasting specification in the main analysis. The common-sample comparisons show that the model's gains relative to ARIMA are not an artifact of forecast coverage. The Diebold–Mariano tests indicate that those improvements are statistically meaningful in several horizon-loss combinations, though not uniformly in every case. The feature-stability and interpretability tables show that the model relies on a stable and economically intelligible predictor set. The subsample, rolling-window, and regime-specific evidence further indicate that the model's relative performance is reasonably robust across time and across states, with the strongest gains appearing at intermediate and longer horizons.

B Formal Hypotheses

This appendix states the formal hypotheses corresponding to the empirical tests reported in the main text. The hypotheses distinguish between raw Forecast Alpha, calibrated Forecast Alpha, and calibrated Residual Surprise. Raw Forecast Alpha measures the model's ex ante deviation from the market-based inflation benchmark. Calibrated Forecast Alpha measures the recursively estimated pass-through component of that deviation into realized market-relative inflation surprise. Calibrated Residual Surprise is the remaining market-relative inflation surprise after the pass-through adjustment.

B.1 Existence of a Forecastable Component

The first empirical question is whether market-relative unexpected inflation contains a non-trivial forecastable component. Let calibrated Forecast Alpha at horizon h be denoted by $\alpha_{t,h}^{cal}$ and calibrated Residual Surprise by $\eta_{t,h}$. The calibrated Alpha Share is defined as

$$AS_h^{cal} = \frac{E[|\alpha_{t,h}^{cal}|]}{E[|\alpha_{t,h}^{cal}|] + E[|\eta_{t,h}|]}.$$

The corresponding formal hypothesis is

$$H_0 : E[|\alpha_{t,h}^{cal}|] = 0 \quad \text{vs.} \quad H_1 : E[|\alpha_{t,h}^{cal}|] > 0.$$

Evidence that AS_h^{cal} is bounded away from zero implies that a meaningful portion of market-relative inflation surprise is forecastable after recursive calibration.

B.2 State Dependence of Forecastability

The second empirical question is whether the forecastable component varies across regimes. Regime-conditioned calibrated Alpha Share is defined as

$$AS_{h,k}^{cal} = \frac{E[|\alpha_{t,h}^{cal}| \mid S_t = k]}{E[|\alpha_{t,h}^{cal}| \mid S_t = k] + E[|\eta_{t,h}| \mid S_t = k]}.$$

The formal null is that calibrated Alpha Share does not differ across regimes:

$$H_0 : AS_{h,k}^{cal} = AS_{h,j}^{cal} \quad \forall k, j,$$

against the alternative that at least one regime pair differs:

$$H_1 : AS^{cal}_{h,k} \neq AS^{cal}_{h,j} \quad \text{for some } k, j.$$

This hypothesis corresponds to the regime-conditioned decomposition reported in the main text.

B.3 Relative Forecast Performance versus the Market Benchmark

The third empirical question is whether the model improves upon a market benchmark in conventional forecast-error terms. Let $L(e_{t,h})$ denote forecast loss at horizon h, with emphasis on squared-error loss.

The formal null for equal predictive accuracy is

$$H_0 : \mathbb{E}\Big[L(e^{model}_{t,h}) - L(e^{mkt}_{t,h})\Big] = 0,$$

against the alternative

$$H_1 : \mathbb{E}\Big[L(e^{model}_{t,h}) - L(e^{mkt}_{t,h})\Big] < 0,$$

when lower loss indicates superior model performance.

This hypothesis is evaluated using forecast-error metrics and Diebold–Mariano tests by horizon.

B.4 Incremental Information beyond Market-Implied Inflation Compensation

The fourth empirical question is whether model-based forecasts contain information beyond that embedded in market pricing. The encompassing regression is

$$\pi_{t+h} = a_h + b_h \hat{\pi}^{mkt}_{t+h|t} + c_h \hat{\pi}^{model}_{t+h|t} + u_{t,h}.$$

The coefficient of primary interest is c_h. The formal null is

$$H_0 : c_h = 0,$$

against the alternative

$$H_1 : c_h \neq 0.$$

Failure to reject $b_h = 0$ jointly with rejection of $c_h = 0$ is consistent with model encompassing. Rejection of both restrictions indicates that the market benchmark and model forecast each contain distinct predictive information.

This hypothesis is closely related to raw Forecast Alpha, because raw Forecast Alpha is the model's ex ante deviation from the market-based inflation benchmark.

B.5 Calibrated Residual Surprise as an Approximate Innovation

The fifth empirical question is whether calibrated Residual Surprise behaves like an approximately unpredictable residual after recursive pass-through adjustment. Formally,

$$H_0 : E[\eta_{t,h}] = 0 \quad \text{and} \quad \mathrm{Cov}(\eta_{t,h}, \alpha^{cal}_{t,h}) = 0,$$

against the alternative that calibrated Residual Surprise retains systematic association with calibrated Forecast Alpha or has a non-zero mean.

This hypothesis is evaluated using residual-surprise diagnostics by horizon. The key diagnostic is whether calibrated Residual Surprise remains systematically related to calibrated Forecast Alpha after the recursive pass-through adjustment. Failure to reject orthogonality supports the interpretation that the calibration separates the historically realized component of model–market disagreement from the remaining residual surprise.

B.6 Market Relevance of Raw Forecast Alpha

The sixth empirical question is whether raw Forecast Alpha maps into subsequent repricing in inflation-sensitive market instruments. Let r^{BE}_{t+1} denote the realized breakeven-market payoff or breakeven-rate change used in the market-relevance exercise, and let $d_{t,h}$ denote the directional signal implied by raw Forecast Alpha. The formal null is

$$H_0 : E[d_{t,h} r^{BE}_{t+1}] = 0,$$

against the alternative

$$H_1 : E[d_{t,h} r^{BE}_{t+1}] > 0.$$

This hypothesis is evaluated using gross market-relevance statistics, including mean gross PnL, ex post information ratios, hit rates, and regime-conditioned performance. The test does not claim to establish a fully optimized production trading strategy. Instead, it evaluates whether raw Forecast Alpha contains directional market information before imposing transaction-cost, funding, carry, roll, or vehicle-specific implementation assumptions.

B.7 Distributional Diagnostics

This subsection reports distributional-shape and serial-dependence diagnostics for Forecast Alpha and Residual Surprise by horizon. The objective is to assess whether Gaussian iid assumptions provide a reasonable approximation for the decomposition objects used in the paper and, in turn, to motivate the use of heteroskedasticity- and autocorrelation-robust inference in the main tests.

The diagnostics indicate that neither component is well described by a Gaussian iid benchmark. Across horizons, both Forecast Alpha and Residual Surprise exhibit positive skewness, substantial excess kurtosis, and strong evidence of serial dependence under Ljung–Box tests. Jarque–Bera

Table 18: Distributional and Serial-Dependence Diagnostics for Forecast Alpha and Residual Surprise

Horizon	Series	Mean	Std. Dev.	Skewness	Ex. Kurt.	JB p-value	LB p-value
1-month	Forecast Alpha	0.1954	0.8576	1.8549	3.1611	0.0000	0.0000
1-month	Residual Surprise	-0.0045	0.1370	1.1226	6.1382	0.0000	0.0000
3-month	Forecast Alpha	0.1991	0.8846	1.8006	3.3931	0.0000	0.0000
3-month	Residual Surprise	-0.0082	0.2952	1.7579	7.9352	0.0000	0.0000
6-month	Forecast Alpha	0.1948	0.9221	1.7570	3.4186	0.0000	0.0000
6-month	Residual Surprise	-0.0039	0.4345	1.6054	7.2543	0.0000	0.0000
12-month	Forecast Alpha	0.1950	0.9074	1.6562	3.3020	0.0000	0.0000
12-month	Residual Surprise	-0.0041	0.7484	1.9572	7.7630	0.0000	0.0000

Notes: Forecast Alpha is defined as the model forecast minus the market expectation. Residual Surprise is defined as realized inflation minus the model forecast. Ex. Kurt. denotes excess kurtosis. JB p-value is the p-value from the Jarque–Bera normality test. LB p-value is the p-value from the Ljung–Box serial-correlation test using horizon-matched lag length. These diagnostics are reported to assess whether Gaussian iid assumptions are plausible; the paper's main inference relies on heteroskedasticity- and autocorrelation-consistent procedures rather than on exact normality. Market benchmark source: `T5YIE_proxy`.

statistics also strongly reject normality at every horizon. These results support the paper's use of robust inference procedures rather than reliance on homoskedastic Gaussian error assumptions.

C Supplementary Decomposition and Transition Diagnostics

Table 19: Forecast Alpha and Realized Market-Relative Inflation Surprise

Horizon	N	Intercept	t-stat	Alpha Coef.	t-stat	R^2
1	277	-0.000	-0.035	0.978	78.244	0.974
3	277	0.011	0.411	0.906	23.953	0.889
6	277	0.032	0.634	0.814	13.162	0.779
12	277	0.074	0.625	0.598	4.385	0.408

This table reports horizon-by-horizon regressions of market-relative unexpected inflation (actual − market) on Forecast Alpha (model − market). Reported t-statistics use Newey–West heteroskedasticity- and autocorrelation-robust standard errors with horizon-adjusted lag length $h - 1$. A positive and significant Alpha coefficient indicates that ex ante model-market gaps are associated with realized market-relative inflation differences.

C.1 Inflation-Band Identification and Final State Mapping

The regime structure used in the paper is supported by a preliminary inflation-band identification exercise. This exercise provides a data-driven basis for discrete inflation environments before estimating the time-varying Hidden Markov Model used in the main analysis. The objective is not to maximize the number of latent states, but to identify economically distinct inflation environments that can be used reliably in recursive, real-time forecasting and decomposition exercises.

I first estimate one-dimensional inflation-band specifications using year-over-year core PCE

Table 20: Legacy Regime-Conditional Decomposition Summary

Horizon	Regime	N	Mean |Alpha|	Mean |Residual|	Mean Alpha Share	Mean Residual Share
1	0	80	0.602	0.080	0.771	0.229
1	1	109	0.631	0.103	0.769	0.231
1	2	87	0.371	0.103	0.682	0.318
3	0	80	0.616	0.159	0.674	0.326
3	1	109	0.677	0.211	0.681	0.319
3	2	87	0.375	0.222	0.595	0.405
6	0	80	0.648	0.216	0.666	0.334
6	1	109	0.711	0.309	0.608	0.392
6	2	87	0.411	0.342	0.524	0.476
12	0	80	0.705	0.433	0.572	0.428
12	1	109	0.685	0.481	0.514	0.486
12	2	87	0.388	0.521	0.453	0.547

Regime-conditional decomposition of market-relative unexpected inflation. Forecast Alpha is defined as the model forecast minus the market-based inflation expectation benchmark, and Residual Surprise is defined as realized inflation minus the model forecast. Shares are computed in absolute-value terms.

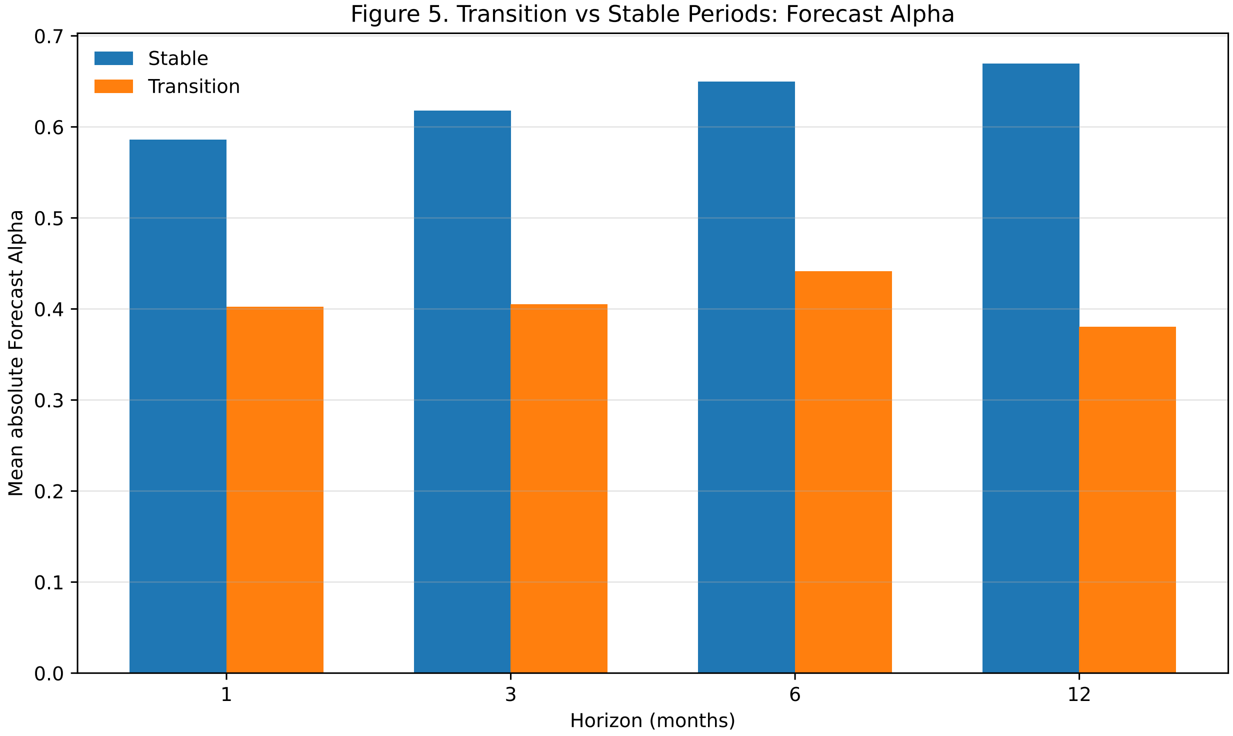

Figure 5: Transition versus stable periods. Bars report mean absolute Forecast Alpha by horizon, comparing observations classified as transition periods with those classified as stable periods using filtered TV-HMM posterior transition intensity.

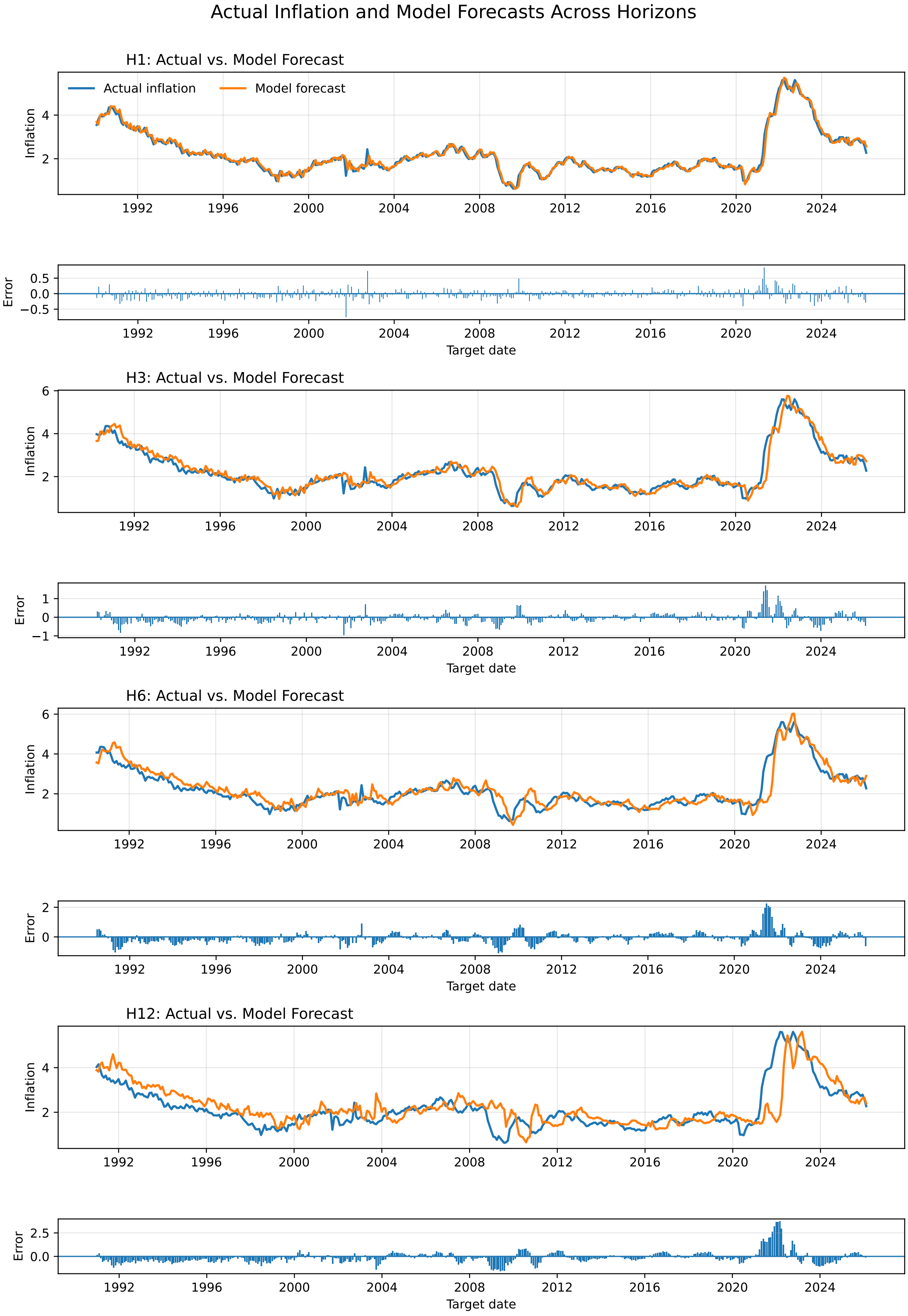

Figure 6: Actual inflation, horizon-specific forecasts, and forecast errors through time. The figure reports four vertically stacked panels corresponding to the 1-, 3-, 6-, and 12-month forecast horizons. In each panel, the upper segment plots realized inflation and the model-implied forecast over time, while the lower segment plots the forecast error, defined as realized inflation minus the model forecast, with a horizontal zero line.

inflation. Candidate Gaussian and Student-t mixture models are estimated over alternative values of K, and the preferred preliminary band structure is selected using the Bayesian information criterion,

$$\text{BIC} = -2\ell(\hat{\theta}) + p\log(n),$$

where $\ell(\hat{\theta})$ is the maximized log likelihood, p is the number of estimated parameters, and n is the number of observations. This procedure selects a four-component Gaussian mixture for year-over-year core PCE inflation. The resulting estimated decision boundaries are 2.439, 5.619, and 8.015 percent.

Table 21 reports the empirical inflation bands implied by these boundaries. The first two bands contain the majority of observations and correspond to low- and moderate-inflation environments. The third and fourth bands correspond to elevated- and very-high-inflation environments, respectively. However, the highest-inflation band contains only 35 observations, or 4.40 percent of the sample. This limited sample support makes it unsuitable as a separate state for recursive regime-specific forecasting, decomposition, and trading-rule evaluation. I therefore collapse the highest-inflation band into the adjacent elevated-inflation band.

The final regime layer used in the paper is therefore a parsimonious three-state specification. State 0 corresponds to low inflation, State 1 to moderate inflation, and State 2 to high inflation. This mapping preserves the BIC-selected evidence that inflation environments are discrete and nonlinear, while avoiding an over-fragmented state space with insufficient observations in the upper tail. The final TV-HMM estimates filtered posterior probabilities over these three states recursively using only information available at each forecast origin.

Table 21: BIC-Guided Inflation-Band Selection and Final State Mapping

Preliminary Band	Inflation Range	Mean Inflation	N	Sample Share	Final State
Band 0	$\pi \leq 2.439$	1.657	400	50.31%	State 0
Band 1	$2.439 < \pi \leq 5.619$	3.833	286	35.97%	State 1
Band 2	$5.619 < \pi \leq 8.015$	6.637	74	9.31%	State 2
Band 3	$\pi > 8.015$	9.178	35	4.40%	State 2

Notes: Preliminary inflation bands are based on year-over-year core PCE inflation. Candidate mixture specifications are compared using BIC, and the selected preliminary specification is a four-component Gaussian mixture. The highest-inflation band contains only 35 observations and is collapsed into the adjacent high-inflation band for the final three-state TV-HMM implementation. Inflation values are reported in percentage points.

C.2 Economic Characterization of the Latent States

This appendix provides a descriptive economic characterization of the latent states used in the main analysis. The purpose is interpretive rather than structural: the states are intended to summarize distinct environments for market-relative inflation predictability, not to identify uniquely causal macroeconomic regimes.

This appendix figure provides an economic visualization of the latent states in unemployment–

Table 22: Macro-Augmented Profile of the Latent Regimes

Regime	N	Sample Share	Mean $\|\alpha\|$	Mean α	Core PCE YoY	UNRATE	DFF
Unknown	4	0.4%	0.183	0.114	2.276		3.640
State 0	320	28.9%	0.643	0.245	2.269	5.486	2.389
State 1	436	39.4%	0.676	0.360	2.291	5.451	2.108
State 2	348	31.4%	0.386	-0.054	1.842	6.367	0.745

Notes: The table summarizes the latent regimes using both model-based and macroeconomic descriptors. Forecast Alpha is defined as the direct compact model forecast minus the market expectation benchmark. Core PCE inflation, unemployment, and the effective federal funds rate are regime-average macro characteristics included for descriptive interpretation only. Regime labels should be interpreted descriptively rather than structurally.

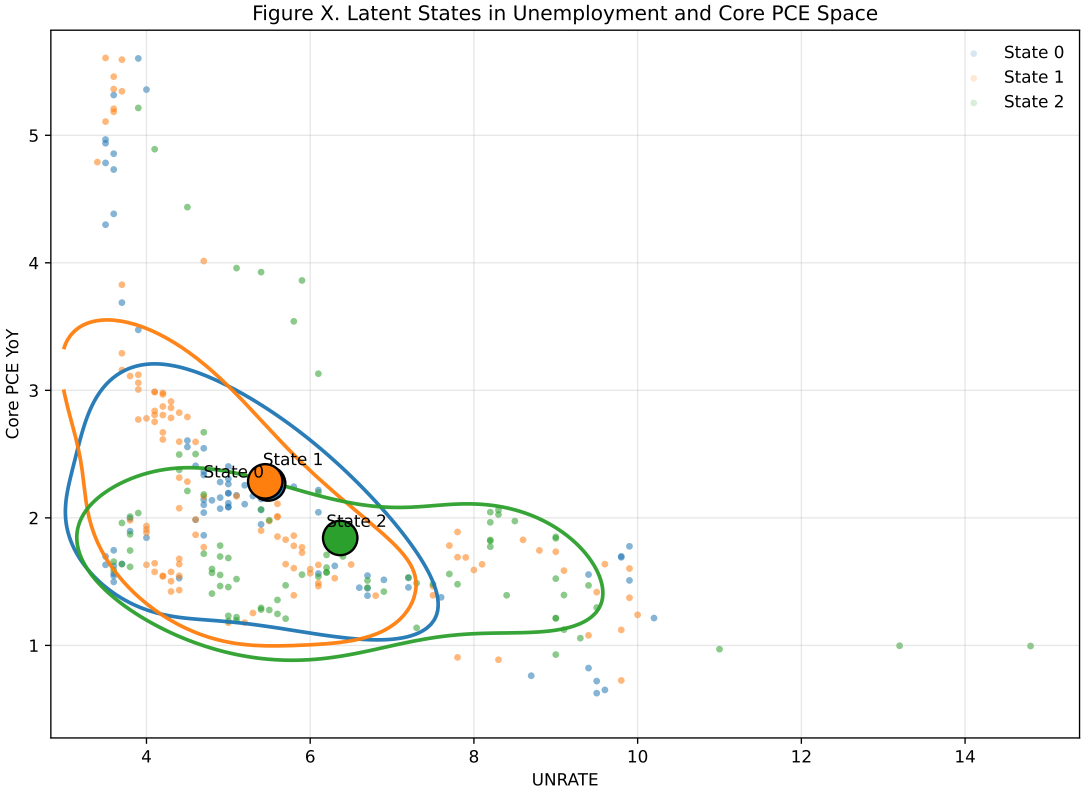

Figure 7: Latent states in unemployment and core PCE space. Each point represents a matched monthly observation, colored by dominant latent regime. Large black-edged circles denote regime centroids. Colored contours trace state-specific kernel-density regions and are intended as descriptive summaries of where each regime is most concentrated rather than as hard classification boundaries. State 2 is centered in a softer macroeconomic region and extends further into higher-unemployment, lower-core-inflation space, while States 0 and 1 overlap more substantially. The figure therefore supports an interpretation in which the latent states are economically meaningful but not perfectly separated macroeconomic clusters.

core-inflation space. The figure should be interpreted as a descriptive map of regime concentration rather than as evidence of sharply separated macroeconomic partitions. State 2 is the most clearly distinct regime, with density concentrated in a softer macroeconomic region marked by higher unemployment and lower core inflation. States 0 and 1 overlap more substantially, indicating that their separation is not driven primarily by large differences in macroeconomic levels alone. Instead, the overlap is consistent with the interpretation that States 0 and 1 represent nearby macroeconomic environments that differ more in the strength of market-relative inflation predictability than in raw unemployment or inflation levels. This reading aligns with the main results of the paper: State 1 is the strongest positive-alpha regime, State 0 is a more moderate positive-alpha regime, and State 2 is the weakest regime for incremental information beyond market pricing.

Table 23: Quantitative Characterization of TV-HMM States

State	N	Core PCE YoY	UNRATE	DFF	Mean $\lvert\alpha\rvert$	Mean α
State 0	320	2.269	5.486	2.389	0.643	0.245
State 1	436	2.291	5.451	2.108	0.676	0.360
State 2	348	1.842	6.367	0.745	0.386	-0.054

State	Persist.	Accel.	Vol.	Exp.	Front-End	Upstream	Labor	Policy
State 0	77.2	5.9	4.6	5.2	1.6	2.1	2.4	1.1
State 1	78.6	5.8	4.2	4.5	1.7	1.8	2.4	1.1
State 2	70.4	8.8	5.2	5.9	3.7	2.8	1.9	1.3

Notes: The top panel reports regime-level macroeconomic and Forecast Alpha descriptors. Core PCE YoY, UNRATE, and DFF are regime averages. Forecast Alpha is defined as the direct compact model forecast minus the market-based inflation benchmark. The bottom panel reports average factor-contribution shares, in percent, across the 1-, 3-, 6-, and 12-month forecast horizons. Factor shares are calculated from absolute forecast contributions, where each feature-level contribution equals the standardized predictor value multiplied by its recursive Elastic Net coefficient. "Persist." denotes inflation level and persistence; "Accel." denotes inflation acceleration/deceleration; "Vol." denotes inflation volatility; "Exp." denotes expectation conditions; "Front-End" denotes front-end market pricing; "Upstream" denotes upstream price pressure; "Labor" denotes labor slack; and "Policy" denotes policy stance. State labels are descriptive rather than structural.

D Risk-Adjusted Forecast Alpha Robustness

The results indicate that RAFA-CVaR does not improve the benchmark-consistent 5-year implementation, where the unscaled Forecast Alpha rule remains stronger. However, RAFA-CVaR improves the 10-year robustness implementation on an information-ratio and hit-rate basis while reducing reliance on the unfiltered signal. The evidence therefore supports interpreting RAFA-CVaR as a portfolio-implementation overlay rather than as the paper's primary market-relevance test.

Table 24: Forecast Alpha versus RAFA-CVaR Trading Rule Robustness

Instrument	Rule	CVaR Half-Life	Mean Net PnL	Ex Post IR	Hit Rate	Max DD
5Y	Forecast Alpha	–	4.3478	0.5865	60.51%	-306.0000
5Y	RAFA-CVaR	36 months	4.3478	0.5865	60.51%	-306.0000
5Y	RAFA-CVaR	60 months	4.3478	0.5865	60.51%	-306.0000
10Y	Forecast Alpha	–	3.2572	0.6999	60.87%	-136.0000
10Y	RAFA-CVaR	36 months	3.1014	0.7137	64.35%	-136.0000
10Y	RAFA-CVaR	60 months	3.1014	0.7137	64.35%	-136.0000

Notes: The table compares the baseline 3-month Forecast Alpha rule with a secondary risk-adjusted Forecast Alpha rule. RAFA-CVaR scales Forecast Alpha by an ex ante CVaR estimate of residual-surprise risk. Forecast Alpha determines the long/short direction; the CVaR adjustment affects signal filtering and confidence. CVaR is estimated using prior residual surprises only, with exponentially weighted histories using 36- and 60-month half-lives. Mean Net PnL and Max DD are reported in basis points after transaction-cost adjustments. The exercise is interpreted as an implementation robustness check rather than as a replacement for the baseline market-relevance test.