Measuring Equity Factor Uncertainty

Henri Huovinen*

March 18, 2026

Abstract

I construct a Factor Uncertainty Index (FUI), inspired by Jurado, Ludvigson, and Ng (2015), measuring time-varying conditional volatility of forecast errors across 13 equity factor themes from Jensen et al. (2023). The FUI predicts future average factor theme returns with in-sample R^2 of 19–23% at h = 6 months and 22–30% at h = 12 months, confirmed out of sample by the Clark and West (2007) test, and predicts cross-sectional factor return dispersion with R^2 up to 52%. Predictability is concentrated in value/fundamental factor themes while risk/defensive and momentum/technical factor themes show no statistical response, consistent with time-varying compensation for fundamental valuation risk rather than limits-to-arbitrage frictions. Subperiod analysis reveals that predictability is episodic rather than uniform: it is strongest during periods of elevated cross-sectional factor uncertainty, especially during the dot-com era and the COVID-19 shock. The FUI dominates VIX, Jurado, Ludvigson, and Ng (2015) macro uncertainty, and Baker-Bloom-Davis (2016) policy uncertainty in horse race regressions; VIX enters with a negative and highly significant coefficient once factor-specific volatility is controlled for, suggesting that the two measures capture opposing economic channels. A volatility-managed portfolio scaling inversely with the FUI generates a spanning alpha of 7.33 basis points per month (t = 5.2), improving the annualized Sharpe ratio from 1.25 to 1.53.

Keywords: Factor uncertainty; Equity factors; Asset pricing; Factor premia; Business cycles; Market regimes

JEL Codes: G11, G12, G17, E32, E37

Lappeenranta-Lahti University of Technology LUT, henri.huovinen@lut.fi

1 Introduction

Understanding how uncertainty shapes asset pricing remains a central question in financial economics. In recent decades, factor investing has become a dominant paradigm in investment management, and the stability of factor premia has attracted growing attention. Yet, while economic and financial uncertainty are well studied, uncertainty within the factor space remains largely unexplored.

Multiple economic and financial market–related uncertainty measures have been developed over the past decades (see, e.g., Bloom, 2009; Bekaert et al., 2013; Jurado et al., 2015; Baker et al., 2016). These studies link various forms of uncertainty, including stock market, macroeconomic, and policy uncertainty, to economic activity and asset prices.

According to Jurado, Ludvigson, and Ng (2015, henceforth JLN), uncertainty can be defined as the conditional volatility of a disturbance that is unpredictable from the perspective of economic agents. In the context of equity markets, which this study examines, uncertainty can generally be defined as the degree to which future returns or market conditions are unpredictable.

Most existing uncertainty-related measures that rely on stock market information focus on observable manifestations of uncertainty. Volatility-based measures, such as the CBOE VIX Index or realized volatility, proxy stock market risk by capturing implied or historical standard deviation. News-based indices, such as the Baker–Bloom–Davis Equity Market-related Economic Uncertainty Index, focus on uncertainty reflected in newspaper coverage of equity-market–related economic and financial conditions, capturing shifts in how the media portrays risks surrounding stock market activity and the broader economic outlook (Baker et al., 2016).

Despite existing research on macroeconomic and stock market–based uncertainty, little is known about uncertainty within the equity style factor space itself. Existing studies on factor-related risk focus on the returns and characteristics of individual equity factors and multifactor models, but they do not quantify the stability or predictability of the factor structure as a whole. This is surprising given the extensive literature on systematic factor investing and represents a limitation from both academic and industry perspectives, as economic agents have intellectual and financial incentives to analyze aggregate factor uncertainty. To address this gap, I construct a Factor Uncertainty Index (FUI) that captures the uncertainty of the factor space and can function as a risk proxy in long–short factor investing. My goal is to provide a measure of factor uncertainty that is as free as possible from dependencies on any single observable economic or financial indicator. Rather than conditioning on macroeconomic controls or financial market variables, the FUI is constructed purely from the unpredictable components of factor return series themselves, ensuring that the index captures genuine uncertainty about the factor structure rather than a projection of known macroeconomic conditions. My approach follows the spirit of JLN by first removing the forecastable component of each factor return series, then estimating the conditional volatility of the resulting forecast errors.

This study is motivated by the rich economic structure underlying the equity style factor universe. I employ 13 factor themes constructed by Jensen et al. (2023) from a set of 153 individual characteristics. These factors are based on purely financial market data (e.g., momentum), a combination of financial market and fundamental accounting data (e.g., book-to-market), or exclusively on fundamental accounting data (e.g., asset growth). This heterogeneous construction of the factor universe is a key strength for the measurement of factor uncertainty. Because the underlying factors originate from fundamentally different information sets, the resulting FUI aggregates uncertainty arising from multiple distinct sources. Uncertainty in purely fundamental factors is intuitively closely related to macroeconomic conditions and real economic risks, whereas uncertainty in financial market–based factors is more associated with investor sentiment and trading dynamics. This structure supports the interpretation of the FUI as a broad measure of systematic uncertainty in the long–short factor space rather than a narrowly defined factor-specific volatility proxy.

I make four main contributions. First, I construct a FUI using a three-layer procedure that extracts forecast errors from rolling AR(1) models for each factor and estimates their conditional volatility via four complementary methods: EWMA, GARCH(1,1), rolling volatility, and stochastic volatility following Kim, Shephard, and Chib (1998, henceforth KSC). The inclusion of stochastic volatility is motivated by the theoretical distinction between uncertainty and volatility: unlike deterministic models, the stochastic volatility model allows for an independent volatility innovation channel, and non-trivial volatility-of-volatility estimates across the 13 factor themes confirm this channel is empirically relevant.

Second, in-sample R^2 for average factor returns reaches 19–23% at h = 6 and 22–30% at h = 12, while dispersion R^2 reaches 17–38% at h = 6 and 21–51% at h = 12. These results are confirmed out of sample: Campbell and Thompson (2008) R^2_OOS of 6–12% at h = 6 and 5–13% at h = 12, with Clark and West (2007) statistics significant at conventional levels for most methods. Predictability is strongest in the post-2000 period, consistent with a higher density of large uncertainty episodes; the FUI spikes most dramatically during the dot-com period compared to the Global Financial Crisis, reflecting its sensitivity to cross-sectional factor divergence rather than aggregate market stress.

Subperiod analysis further reveals that the predictability is episodic rather than uniform, activating sharply during the dot-com era and the COVID-19 shock while lying dormant during the low-volatility post-GFC decade of 2010–2019. An investigation of the underlying economic mechanisms reveals that the predictive content of the FUI is concentrated in value/fundamental factors, while risk/defensive and momentum/technical factors show no significant response. This pattern is consistent with the interpretation that the results reflect time-varying compensation for fundamental valuation risk rather than limits-to-arbitrage frictions.

Third, I show that the FUI is moderately correlated with existing uncertainty measures, including JLN macro uncertainty ($\rho = 0.233$–0.329), Baker-Bloom-Davis (EPU) economic policy uncertainty ($\rho = 0.145$–0.273), and VIX ($\rho = 0.308$–0.401), but dominates all three in horse race regressions for factor returns. The FUI retains strong significance across all methods and horizons when any of the three benchmarks is included as a control, while JLN becomes insignificant and EPU and VIX enter with negative and significant coefficients, suggesting both capture opposing economic channels to the FUI. VIX exhibits a negative coefficient ($t = -2.82$ to -5.87) once factor-specific volatility is controlled for: conditional on factor-level uncertainty, higher market volatility predicts lower factor returns. This suggests that market-level volatility and factor-specific uncertainty act through distinct and partially offsetting channels, with the FUI capturing a risk compensation channel that operates independently of broad market stress conditions.

Fourth, I construct volatility-managed factor portfolios following Moreira and Muir (2017) that scale monthly factor returns inversely with the lagged FUI. The aggregate managed portfolio generates a spanning alpha of 7.33 basis points per month ($t = 5.2$), improving the annualized

Sharpe ratio from 1.25 to 1.53 at the same unconditional variance as the unmanaged benchmark, confirming that the gain reflects genuine timing ability rather than increased average leverage. This result is robust to weight capping and a binary regime strategy that halves exposure when the FUI exceeds its expanding historical median.

The remainder of this paper is organized as follows. Section 2 discusses the related literature. Section 3 presents the data, and Section 4 the construction of the FUI. Section 5 characterizes the statistical properties, cross-method robustness, and business cycle behavior of the FUI. Section 6 presents the predictive regressions, out-of-sample validation, volatility-managed portfolios, and economic mechanism tests. Section 7 concludes.

2 Related Literature

A large literature develops quantitative measures of uncertainty relevant for equity markets. The most widely used market-based indicator is implied volatility from index options, most prominently the CBOE Volatility Index (VIX), which represents the market's expectation of 30-day S&P 500 volatility and is often interpreted as a proxy for investor fear and near-term stock market uncertainty. Bloom (2009) uses stock market volatility and jump-based identification to isolate uncertainty shocks and study their macroeconomic effects. He shows that several cross-sectional measures of uncertainty, including dispersion in firm profit growth, firm stock returns, industry TFP growth, and GDP forecasts, are positively associated with stock market volatility. Baker et al. (2016) construct a monthly newspaper-based index of Economic Policy Uncertainty (EPU) and show that an auxiliary news-based equity market uncertainty index they develop is highly correlated with the monthly average of the daily VIX. Their framework has been extended to various national and thematic contexts.

JLN (2015) propose a data-rich approach that defines uncertainty as the conditional volatility of forecast errors across a large panel of macroeconomic and financial time series. Their macro and financial uncertainty indicators have become standard benchmarks for measuring pervasive uncertainty about the economic environment. Another influential line of work uses text-based methods and structural models to measure uncertainty arising from policy and institutional sources. In this spirit, Pástor and Veronesi (2013) show theoretically and empirically that political uncertainty affects equity risk premia, volatility, and correlations, implying that political risk is a priced source of market uncertainty.

These aggregate uncertainty indicators have been linked to stock market volatility, informational efficiency, and expected returns. For example, uncertainty shocks have been shown to generate large spikes in equity-market volatility and real activity (Bloom, 2009; JLN, 2015). Uncertainty also affects how financial markets process information. Brogaard and Detzel (2015) link EPU to the stock market, showing that an increase in EPU is followed by higher forecasted abnormal stock market returns, while innovations in EPU earn a significant negative risk premium in the Fama-French 25 size-momentum portfolios. Kelly et al. (2016) demonstrate that political uncertainty raises risk premia and implied volatility, reducing the clarity of valuations and increasing the value investors place on protection against adverse outcomes.

A body of evidence further demonstrates that uncertainty acts as a source of risk. Macroeconomic and policy uncertainty predict variation in expected returns, consistent with theoretical models of time-varying risk premia (see, e.g., Pástor and Veronesi, 2013; Bali et al., 2017). Macro-financial uncertainty is also linked to variation in expected returns and volatility. For example, Bekaert, Hoerova, and Lo Duca (2013) decompose the VIX into risk aversion and expected volatility, Conrad and Loch (2015) show that macro variables help forecast long-term stock market volatility, and Ludvigson, Ma, and Ng (2021) study whether uncertainty is an exogenous source of business-cycle fluctuations or an endogenous response to them.

However, existing uncertainty measures are typically defined at the aggregate financial market or macroeconomic level and are inherently silent on how uncertainty is distributed across the internal structure of equity factor returns, making them ill-suited to detect instability in the factor structure itself.

A growing literature highlights that factor premia themselves exhibit substantial time variation and structural instability. Research on factor timing and factor crowding documents that factor exposures, correlations, and performance vary over time, implying that factors evolve through distinct regimes rather than operating as stable, time-invariant constructs (see, e.g., Israel and Moskowitz, 2013; Asness et al., 2017; Baltas, 2019, Haddad et al., 2020; Blitz, 2021; Ilmanen et al., 2021). Related work on conditional factor models and instability in factor loadings (see, e.g., Ferson and Harvey, 1991; Lettau and Ludvigson, 2001; Ang and Chen, 2007; Harvey et al., 2016; Daniel et al., 2020) similarly shows that the explanatory power of traditional factors fluctuates across market environments.

These insights point to an important but underexplored dimension of risk: uncertainty about the factor structure itself. Such uncertainty reflects instability in the relationships that generate cross-sectional returns, including shifts in factor correlations, explanatory power, and the economic narratives driving factor premia. I address this research gap by constructing a FUI defined as the equal-weighted cross-sectional mean of factor-level conditional volatilities of the unforecastable components of factor returns, thereby providing a measure of structural uncertainty within the factor space.

3 Data

3.1 Data for FUI Construction

Data on U.S. long–short factor theme returns covering the period November 1951 to December 2024 are obtained from Global Factor Data (Jensen et al., 2023). The dataset consists of 13 factor themes, including Accruals, Debt Issuance, Investment, Low Leverage, Low Risk, Momentum, Profit Growth, Profitability, Quality, Seasonality, Short-Term Reversal, Size, and Value, which together comprise 153 individual factors. Each factor or characteristic $i \in \{1, \dots, 153\}$ belongs to one of $T = 13$ broad themes, following the categorization by Jensen et al. (2023). Following the recommendation by Jensen et al. (2023), I use capped value-weighted returns, in which stocks are weighted by market equity winsorized at the NYSE 80th percentile. In addition, U.S. capped value-weighted stock market returns are obtained from Global Factor Data.

I focus on 13 factor themes rather than 153 individual factors because Jensen et al. (2023) show that these factors naturally cluster into 13 economically meaningful themes using hierarchical agglomerative clustering. The resulting themes exhibit high within-cluster and low cross-cluster correlations, capturing distinct sources of systematic risk. In their approach, the distance between factors is defined as one minus their pairwise correlation, and clusters are formed using the Ward (1963) linkage criterion, following the implementation of hierarchical clustering described in Murtagh and Legendre (2014).

Table 1: Descriptive Statistics of 13 U.S. Factor Themes

Factor Theme	Mean (%)	Std (%)	Skewness	Kurtosis	SR
Accruals	0.235	1.107	0.27	1.26	0.734
Debt Issuance	0.196	0.699	0.86	9.56	0.969
Investment	0.237	1.973	0.79	8.98	0.416
Low Leverage	0.002	2.716	0.77	25.31	0.002
Low Risk	0.153	3.483	-0.17	9.04	0.152
Momentum	0.395	2.990	-0.69	12.76	0.458
Profit Growth	0.183	1.034	-0.62	5.54	0.614
Profitability	0.223	1.912	0.60	15.65	0.405
Quality	0.288	1.490	-0.09	1.44	0.670
Seasonality	0.140	0.528	0.22	2.94	0.917
Short-Term Reversal	0.181	1.142	0.73	17.57	0.550
Size	0.097	2.156	0.94	5.00	0.156
Value	0.311	3.010	-0.10	18.73	0.358
Average	***0.203***	***1.865***	***0.27***	***10.29***	***0.492***

Notes: Mean and Std are expressed as monthly percentages. SR is the annualized Sharpe ratio, computed as (Mean / Std) × √12. Skewness and Kurtosis are the third and fourth standardized central moments (excess kurtosis). The average row reports the cross-sectional mean across all 13 factors; Sample: November 1951 – December 2024 (878 monthly observations).

Table 1 reports summary statistics for the 13 factor themes. There is substantial heterogeneity across factor themes. Mean monthly returns range from effectively zero (low leverage, 0.002%) to approximately 0.395% (momentum). Monthly standard deviations range from 0.70% (debt issuance) to 3.48% (low risk). Annualized Sharpe ratios extend from near zero (low leverage, 0.002) to above 0.9 (debt issuance, seasonality). All 13 series are stationary (ADF $p < 0.001$ for all), and most exhibit mild positive first-order autocorrelation (median $\rho(1)$ = 0.0871), consistent with the literature on factor return momentum (see, e.g., Gupta and Kelly, 2019; Ehsani and Linnainmaa, 2022).

3.2 Data for Additional Variables

For the predictive regression analysis and comparison with the literature, I employ several additional data series.

CBOE Volatility Index (VIX). Monthly VIX data is available from January 1990 through December 2024 (434 observations). VIX represents the market's expectation of 30-day forward volatility, derived from S&P 500 index options, and serves as the standard measure of option-implied aggregate uncertainty.

The h = 1 macro uncertainty index of JLN (2015) is obtained from Sydney Ludvigson's website. This series measures the common variation in the conditional volatility of the unforecastable component of a large number of macroeconomic indicators. The data is available from July 1960, overlapping 772 monthly observations with the FUI.

The news-based policy uncertainty index of Baker, Bloom, and Davis (2016), obtained from policyuncertainty.com. This index quantifies newspaper coverage of policy-related economic uncertainty and is available from January 1900 through February 2026. The overlap with the FUI sample is the full sample of 878 monthly observations.

US recession dates, covering peak and trough dates for U.S. business cycles, used for event-study analysis of the FUI's behavior around economic downturns, are obtained from the website National Bureau of Economic Research.

4 Construction of the FUI

The Factor Uncertainty Index is constructed through a three-layer procedure inspired by the conceptual framework of JLN (2015). The motivating insight is that uncertainty about a variable should be measured as the conditional volatility of its unforecastable component, rather than the volatility of the raw series. By first removing the forecastable part of factor returns via a time-series model, the FUI captures conditional volatility of genuine forecast errors rather than total return variation. This section describes each of the three layers in detail.

4.1 Layer 1: Forecasting Model and Forecast Errors

For each factor theme i at time t, I estimate a one-step-ahead forecasting model to extract the forecast error. The baseline specification is rolling AR(1):

$$R_{i,t} = \alpha_i + \varphi_i * R_{i,t-1} + \varepsilon_{i,t},$$

where $R_{i,t}$ is the return on factor theme i in month t, α_i is an intercept, φ_i is the slope of AR(1), and $\varepsilon_{i,t}$ is the forecast error. The model is estimated by OLS on a rolling 120-month window, with a minimum of 60 months required to produce out-of-sample forecasts. For each month t, the one-step-ahead forecast is formed using parameters estimated on data up to and including $t - 1$, and the forecast error is the difference between the realized return and this forecast. This rolling-window approach allows the forecasting relationship to evolve over time, accommodating potential structural breaks in factor return dynamics.

The 120-month rolling window reflects a deliberate bias-variance tradeoff: long enough to estimate AR(1) parameters precisely, yet short enough to remain locally relevant. An expanding window would progressively incorporate data from structurally distinct periods — such as the high-inflation 1970s or the post-2008 low-rate environment — potentially distorting parameter estimates in ways unrelated to genuine factor uncertainty. Results are robust to using 96- and 144-month windows, with cross-window correlations exceeding 0.999 across all volatility specifications (RV12, EWMA, GARCH), confirming that the FUI signal is insensitive to moderate variation in estimation horizon.

The choice of AR(1) is deliberately parsimonious, reflecting the well-documented near-unpredictability of factor returns at the monthly frequency. In-sample R^2 values for AR(1) specification range from 0.24% (momentum theme) to 2.41% (profitability theme), with a median of 1.32% (Table A.1). These low R^2 values confirm that factor returns are dominated by their unpredictable component, which is precisely the component whose conditional volatility the FUI is designed to measure.

Extensive diagnostics indicate that richer mean specifications add negligible incremental information relative to AR(1): BIC selects AR(1) for all 13 factors, ARMA(1,1) is never preferred, PCA-based cross-factor augmentation is rarely selected, and the resulting FUI series from alternative specifications are nearly identical to the baseline (Table A.2). Even intercept-only forecasts yield an FUI correlated 0.9996 with the AR(1) version. I therefore retain AR(1) on both empirical and conceptual grounds: the conditional mean specification determines what variation is classified as uncertain by construction, a choice that matters regardless of the specific uncertainty framework employed, and the diagnostics above confirm the empirical adequacy of this choice. My results imply that short-horizon mean predictability in these factor returns is economically small in this sample, while the estimated uncertainty series remains highly persistent, consistent with volatility clustering in financial data.

4.2 Layer 2: Conditional Volatility Estimation

The second layer estimates the time-varying conditional volatility of the Layer 1 forecast errors. I employ four complementary methods, ranging from the simplest non-parametric approach to a full Bayesian stochastic volatility model. The use of multiple methods ensures that results do not depend on any particular volatility specification and allows a direct assessment of whether the stochastic volatility innovation channel matters empirically.

4.2.1 Exponentially Weighted Moving Average (EWMA)

The primary deterministic method is an EWMA with decay parameter of $\lambda = 0.85$:

$$\sigma_{i,t}^2 = \lambda \cdot \sigma_{i,t-1}^2 + (1-\lambda) \cdot \varepsilon_{i,t}^2,$$

where $\sigma_{i,t}^2$ is the EWMA conditional variance of the forecast error for factor theme i at time t, $\sigma_{i,t-1}^2$ is the previous period's EWMA estimate of that conditional variance, and $\varepsilon_{i,t}^2$ is the period's squared forecast error. The EWMA is non-parametric, adaptive to changing volatility regimes, and computationally efficient. The effective span of the EWMA is roughly 12 months, meaning that approximately 86% of the total weight is placed within that period. The decay parameter $\lambda = 0.85$ is selected by minimizing the QLIKE loss function across all 13 factors.

The QLIKE criterion, defined as $QLIKE = \frac{1}{T}\sum_{t=1}^{T}[\log(\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_t^2}]$ is the standard loss function for evaluating volatility forecasts (Patton 2011), as it is robust to noise in the volatility proxy. The optimization yields a sharp minimum at $\lambda = 0.85$, with QLIKE increasing monotonically for $\lambda > 0.85$, confirming that the optimum is well-identified and not on a knife-edge. The selected value is lower than the RiskMetrics default of $\lambda = 0.94$, reflecting the higher noise-to-signal ratio in monthly factor returns relative to the daily asset returns for which RiskMetrics was designed.

4.2.2 Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

The second method estimates a GARCH(1,1) with a zero-mean specification for the forecast errors of each factor theme:

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \cdot \varepsilon_{i,t-1}^2 + \beta_i \cdot \sigma_{i,t-1}^2,$$

where $\sigma_{i,t}^2$ denotes the conditional variance of the forecast error for factor theme i at time t, ω_i is the constant term, α_i is the ARCH coefficient, and β_i is the GARCH coefficient. Parameters are estimated by quasi-maximum likelihood on a 120-month rolling window and re-estimated every 12 months. Between re-estimations, the one-step-ahead conditional variance is propagated using the most recently estimated parameters. This approach limits estimation noise from frequent re-optimization while capturing the essential volatility dynamics. The zero-mean specification is appropriate because the inputs are already forecast errors from Layer 1 and should have mean zero by construction.

4.2.3 Rolling Volatility (RV12)

The third method computes a simple 12-month rolling root mean square of forecast errors:

$$\sigma_{i,t} = \sqrt{\frac{1}{12}\sum_{k=0}^{11} \varepsilon_{(i,t-k)}^2},$$

where $\sigma_{i,t}$ denotes the RV12 volatility measure, computed as the 12-month rolling root mean square of forecast errors for factor theme i at time t, and $\varepsilon_{(i,t-k)}^2$ is the squared error of factor theme i observed k periods earlier.

This measure serves as a useful benchmark against more sophisticated volatility models. The 12-month window balances responsiveness to volatility changes against estimation noise.

4.2.4 Stochastic Volatility (SV)

The fourth model in Layer 2 is stochastic volatility by KSC (1998), implementing the mixture sampler for each factor's forecast errors. The stochastic volatility model is:

$$y_t = \exp\left(\frac{h_t}{2}\right) \cdot \varepsilon_t, \varepsilon_t \sim N(0,1)$$

$$h_t = \mu + \varphi(h_{t-1} - \mu) + \sigma\eta \cdot \eta_t, \eta_t \sim N(0,1)$$

where y_t is the factor theme return forecast error $\varepsilon_{i,t}$ from Layer 1, h_t is the log-volatility, μ is the unconditional mean of log-volatility, φ governs persistence, and $\sigma\eta$ is the volatility-of-volatility, i.e., the standard deviation of the independent volatility innovation η_t. The key feature is that η_t is independent of ε_t: shocks to the variance process are separate from shocks to the level of the series. This is the theoretical mechanism that distinguishes "uncertainty" from "volatility" in JLN's framework, and which deterministic models like GARCH and EWMA cannot capture by construction.

Estimation proceeds as follows. Log-linearization transforms the observation equation into $y_t^* = \log(y_t^2 + c) = h_t + z_t$, where $c = 10^{-6}$ is a small offset to handle near-zero observations, and $z_t = \log(\varepsilon_t^2)$ where $\varepsilon_t^2 \sim \chi^2(1)$. KSC approximate this non-Gaussian distribution with a 7-component Gaussian mixture, using the exact mixture parameters. Conditional on the mixture indicators, the state-space model is linear and Gaussian, enabling efficient Kalman filtering. Because the backward-sampling step conditions on the full path $h(1{:}\mathrm{T})$, the SV-based FUI is an in-sample estimate. All real-time implementable results use the EWMA and GARCH estimates, which are strictly causal; the SV model serves as an in-sample benchmark confirming that the deterministic variance models adequately capture the volatility dynamics relevant for factor uncertainty.

The MCMC algorithm iterates over five blocks. First, mixture indicators s_t are sampled from their full conditional distribution. Second, the log-volatility path $h(1{:}T)$ is sampled via forward-filtering backward-sampling (de Jong and Shephard, 1995). Third, the unconditional mean is sampled from its Gaussian full conditional under a $N(0,100)$ prior. Fourth, the persistence parameter is sampled via a Metropolis-Hastings step with a $\text{beta}(20,1.5)$ prior on $\frac{\varphi+1}{2}$, which concentrates mass near $\varphi = 1$ while ensuring stationarity. Fifth, the volatility-of-volatility is sampled from its inverse-gamma full conditional under an $\text{IG}(2.5,0.025)$ prior. I run 5,000 iterations with 2,000 burn-in and thinning of 2, yielding 1,500 posterior draws per factor. Convergence is assessed via trace plots and the Geweke (1992) diagnostic.

Table 2: Stochastic Volatility Parameter Estimates for 13 U.S. Factor Themes

Factor Theme	φ	ση	μ
Accruals	0.934	0.234	-10.47
Debt Issuance	0.937	0.229	-11.48
Investment	0.968	0.236	-9.63
Low Leverage	0.986	0.205	-9.24
Low Risk	0.966	0.239	-8.45
Momentum	0.901	0.505	-9.00
Profit Growth	0.928	0.321	-10.79
Profitability	0.967	0.279	-9.95
Quality	0.940	0.196	-9.75
Seasonality	0.968	0.142	-11.78
Short-Term Reversal	0.926	0.329	-10.67
Size	0.945	0.236	-9.14
Value	0.965	0.298	-9.09

Notes: *Parameters are posterior means from the KSC (1998) MCMC mixture sampler with 5,000 iterations, 2,000 burn-in, and thinning factor of 2 (1,500 posterior draws). Estimation uses 818 monthly observations per factor. φ is the log-variance persistence parameter; ση is the volatility of log-variance innovations; μ is the unconditional mean of log-variance.*

Table 2 reports the posterior mean estimates, which reveal economically interpretable SV parameters across the 13 factors. Average log-volatility persistence is φ = 0.949, confirming slow-moving volatility dynamics. Low leverage has the highest persistence (φ = 0.986) and seasonality the lowest ση = 0.142, both reflecting stable, slowly-evolving volatility processes. Momentum stands out on both dimensions: the lowest persistence (φ = 0.901) and by far the largest volatility-of-volatility (ση = 0.505), driven by the sudden regime shifts and crash risk well-documented in momentum strategies (Daniel and Moskowitz, 2016; Barroso and Santa-Clara, 2015), which generate volatility jumps that deterministic models would attribute entirely to past squared errors. Average volatility-of-volatility is ση = 0.265, which is meaningfully positive: the independent volatility innovation is not negligible for factor returns.

The SV-based FUI, which is constructed as $FUI_t^{SV} = \frac{1}{13}\sum_{i=1}^{13} \exp\left(\frac{\hat{h}_{i,t}}{2}\right)$, where $\hat{h}_{i,t}$ is the posterior mean log-volatility for factor i at time t, has a correlation of 0.94 with the EWMA-based FUI. This is high but meaningfully below the near-unity correlations among the three deterministic methods, indicating that the SV innovation channel contributes material variation to the index. The SV-FUI has higher persistence (AR(1) = 0.996 vs. 0.983 for EWMA) and a substantially smoother path, consistent with the Bayesian shrinkage inherent in the MCMC estimator. The mean level differs in scale (0.0083 vs. 0.0161) because $\exp\left(\frac{h}{2}\right)$ and the EWMA standard deviation are not on identical scales, but rank correlations confirm that the two series identify the same high- and low-uncertainty periods.

4.3 Layer 3: Cross-Sectional Aggregation

The aggregate Factor Uncertainty Index is defined as the equal-weighted cross-sectional mean of the factor-level conditional volatilities:

$$FUI_t = \frac{1}{N}\sum_{i=1}^{N} \sigma_{i,t},$$

where N =13 and $\sigma_{i,t}$ is the conditional volatility of factor i at time t from any of the four Layer 2 methods. The equal-weighted aggregation follows the baseline specification in JLN. Alternative weighting schemes, such as value-weighted by factor AUM or eigenvector-weighted as in JLN's robustness checks could be explored but are unlikely to matter substantively for 13 relatively homogeneous factor theme series.

5 Properties of the FUI

5.1 Summary Statistics and Time Series Properties

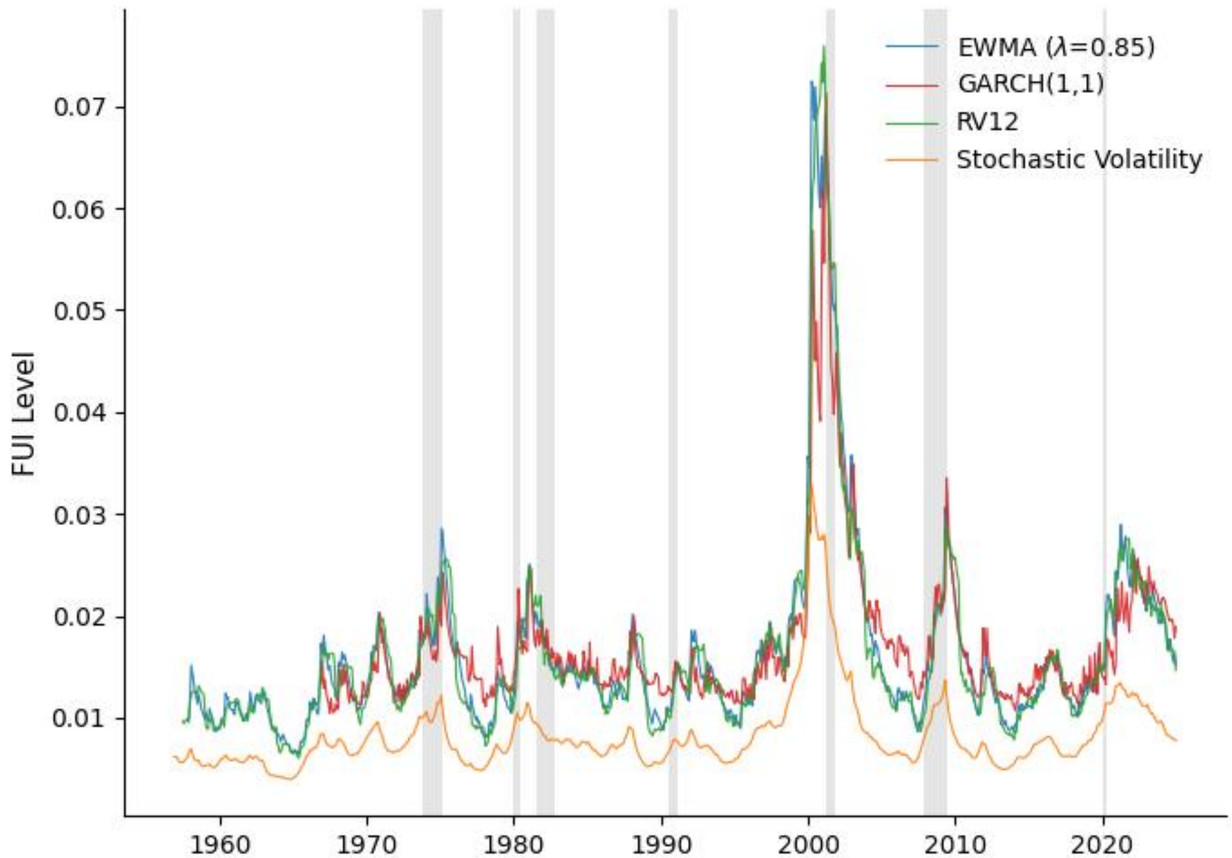

Figure 1: Factor Uncertainty Index Across Four Volatility Estimators.

Notes: *Figure 1 plots the monthly time series of the FUI constructed using four alternative volatility estimators: EWMA ($\lambda = 0.85$), GARCH(1,1), RV12, and stochastic volatility. EWMA, RV12, and SV are plotted from their individual start dates (October 1957, June 1957, and November 1956 respectively), while GARCH begins in November 1966. All series end in December 2024. NBER-dated recession periods are shaded in gray.*

Before turning to the FUI's predictive content in Section 6, I characterize its statistical properties, cross-method robustness, and relationship to the business cycle. Figure 1 reveals several notable patterns in the Factor Uncertainty Index across the full sample period. All four estimators co-move closely throughout, suggesting that the FUI is robust to the choice of volatility model. The indices display a clear cyclical pattern, with elevated levels coinciding with NBER-dated recessions. The most pronounced episode is the dot-com bubble around 2000, during which the FUI spikes before reverting sharply. A secondary spike is visible during the 2008–2009 Global Financial Crisis. Notably, the Stochastic Volatility estimator (orange) produces systematically lower and smoother estimates throughout the sample. In each case, spikes in factor uncertainty are temporary and tend to recede after periods of market stress.

Table 3: Summary Statistics of the FUI

Statistic	FUI(EWMA)	FUI(GARCH)	FUI(RV12)	FUI(SV)
Mean	0.0172	0.0172	0.0171	0.0087
Median	0.0146	0.0149	0.0144	0.0076
Std. Dev.	0.0098	0.0075	0.0100	0.0041
Minimum	0.0081	0.0104	0.0072	0.0049
Maximum	0.0724	0.0712	0.0759	0.0329
Skewness	3.3670	3.5413	3.4391	3.1810
Kurtosis (excess)	13.2790	15.6863	14.2069	12.2820
AR(1)	0.9847	0.9636	0.9894	0.9954

***Notes**: The table reports descriptive statistics for the four variants of the FUI, each constructed using an alternative volatility estimator: exponentially weighted moving average (EWMA, λ = 0.85), rolling GARCH(1,1) estimated on a 120-month window, 12-month rolling forecast-error volatility (RV12), and the KSC (1998) stochastic volatility model (SV). All statistics are computed over the common sample period November 1966 – December 2024 (N = 698 months), determined by the minimum start date of the GARCH(1,1) estimator. Individual series are available prior to the common sample: EWMA from October 1957 (N = 807), RV12 from June 1957 (N = 811), and SV from November 1956 (N = 818). Kurtosis is reported as excess kurtosis (Fisher's definition), normalized so that the normal distribution has a value of zero. AR(1) denotes the first-order autocorrelation coefficient.*

The three frequentist estimators, EWMA (λ = 0.85), GARCH(1,1), and RV12, exhibit near-identical dynamics throughout the sample, with pairwise correlations ranging from 0.928 to 0.982 (Table 4). All three share an unconditional mean of approximately 1.71–1.72% and exhibit pronounced right skewness (3.37–3.54) and excess kurtosis (13.28–15.69), reflecting the well-documented non-normality of financial volatility proxies. Persistence is uniformly high, with first-order autocorrelations between 0.964 and 0.989, confirming the slow-moving nature of factor uncertainty documented in JLN (2015).

The stochastic volatility (SV) estimator co-moves closely with the frequentist variants, with full-sample correlations between 0.840 and 0.930, but operates at a uniformly lower unconditional mean of 0.87%. This level difference reflects the Bayesian shrinkage embedded in the KSC (1998) MCMC posterior, which pulls extreme volatility draws toward the unconditional log-volatility mean. Importantly, the SV estimator remains the most persistent of the four, with AR(1) = 0.995, suggesting it captures the long-run component of uncertainty more cleanly by filtering out high-frequency noise. Despite this near-unit-root persistence, formal stationarity tests confirm that all four FUI variants are stationary: ADF tests reject the unit root null at the 1% level for all variants, while KPSS tests fail to reject the stationarity null throughout, with full results reported in Table A.3.

5.2 Cross-Method Robustness

Table 4: Pairwise Correlations Between FUI Variants

This table characterizes the agreement and divergence among the four FUI variants. Panel A reports full-sample pairwise Pearson correlations. Panel B reports statistics for 60-month rolling correlations, including the fraction of windows in which the rolling correlation falls below 0.90.

Panel A: Full-Sample Pairwise Correlations

	EWMA	**GARCH**	**RV12**	**SV**
EWMA	1.000	0.943***	0.982***	0.930***
GARCH	0.943***	1.000	0.928***	0.840***
RV12	0.982***	0.928***	1.000	0.910***
SV	0.930***	0.840***	0.910***	1.000

Notes: *Full-sample Pearson correlations between the four FUI variants, computed over the common sample shared by all four series ($N = 698$ monthly observations). ***, **, * denote significance at the 1%, 5%, and 10% levels.*

Panel B: 60-Month Rolling Correlations

Pair	**Mean**	**Min**	**Max**	**<0.90 (%)**
EWMA–GARCH	0.842	0.431	0.982	69.0%
EWMA–RV12	0.936	0.834	0.993	21.8%
EWMA–SV	0.797	0.430	0.993	80.3%
GARCH–RV12	0.772	0.244	0.971	80.4%
GARCH–SV	0.661	0.037	0.969	95.1%
RV12–SV	0.718	0.025	0.987	86.2%

Notes*: Statistics computed from 60-month rolling window Pearson correlations. "<0.90 (%)" reports the percentage of rolling windows in which the pairwise correlation falls below 0.90, indicating periods of meaningful divergence between variants*

A central question for the FUI is whether the results depend on the specific volatility estimation method. Panel A of Table 4 reports pairwise full-sample correlations across the four FUI variants. All pairs are highly correlated in the full sample, with correlations ranging from 0.840 (GARCH–SV) to 0.982 (EWMA–RV12), and all are significant at the 1% level.

However, rolling 60-month correlations reveal considerably more time-variation than the full-sample figures suggest. Panel B of Table 4 reports the mean, minimum, and maximum of the rolling correlations, along with the fraction of months where the correlation falls below 0.90. The EWMA–RV12 pair is the most stable, remaining above 0.90 in 78% of 60-month rolling windows (mean = 0.936). By contrast, pairs involving GARCH or SV show substantially lower and more volatile co-movement: the GARCH–SV rolling correlation averages only 0.661 and falls below 0.90 in 95% of rolling windows, while the RV12–SV and GARCH–RV12 pairs are below 0.90 in 86% and 80% of 60-month windows, respectively. This suggests that while the methods agree on the broad level of uncertainty over the full sample, they can diverge meaningfully at shorter horizons and in specific subperiods.

The SV-based FUI warrants particular attention. Although its full-sample correlation with EWMA is high at 0.930, this is meaningfully below the near-unity co-movement observed among the three deterministic methods, and its rolling correlation with all deterministic variants is considerably more volatile, falling below 0.90 in 80–95% of 60-month windows depending on the pair. The lower co-movement reflects the contribution of the independent volatility innovation that is unique to the SV specification. The SV-FUI is smoother, more persistent, and operates on a different scale (mean = 0.0083 vs. 0.0172 for EWMA), but rank correlations confirm that both series identify the same episodes of elevated uncertainty (Table A.4). All four methods agree on the relative ordering of major crisis periods, with the 2000–2002 dot-com period as the dominant event, followed by the 2007–2009 GFC, and the 2020 COVID onset.

Given these patterns, the EWMA variant serves as the primary specification in the economic recession analysis that follows in Section 5.4. Its high and relatively stable co-movement with RV12 suggests it captures the common signal in deterministic volatility-based uncertainty, and its longer available sample makes it the natural baseline.

5.3 Temporal Stability

Table 5: FUI Summary Statistics by Subperiod

Method	Period	N	Mean	Std	AR(1)
EWMA	Pre-1990	387	0.0134	0.0039	0.9586
	1990–1999	120	0.0150	0.0043	0.9489
	Post-2000	300	0.0205	0.0135	0.9873
GARCH	Pre-1990	278	0.0150	0.0026	0.8479
	1990–1999	120	0.0144	0.0025	0.8656
	Post-2000	300	0.0203	0.0103	0.9673
RV12	Pre-1990	391	0.0132	0.0041	0.9756
	1990–1999	120	0.0149	0.0042	0.9730
	Post-2000	300	0.0202	0.0138	0.9906
SV	Pre-1990	398	0.0070	0.0016	0.9899
	1990–1999	120	0.0087	0.0034	0.9967
	Post-2000	300	0.0099	0.0054	0.9969

Notes: *AR(1) is the first-order autocorrelation of the FUI level. Subperiods: Pre-1990 (full history to December 1989), 1990–1999, Post-2000 (January 2000 onward). Subperiod predictive regression results are reported in Table 8 (Section 6.1).*

A natural concern with any long-sample predictor is whether its properties are stable over time or concentrated in a particular historical episode. Table 5 examines the FUI across three subperiods: pre-1990, 1990–1999, and post-2000. The 1990–1999 decade is isolated separately as it represents the transitional period preceding the sharp rise in factor uncertainty documented post-2000, allowing the analysis to distinguish between the low-uncertainty regime of the 1990s bull market and the structurally elevated uncertainty environment that follows.

Table 5 reveals two clear patterns. First, persistence is high and stable throughout: first-order autocorrelations range from 0.85 to 0.99 across methods and subperiods, and this is not a post-2000 artifact — EWMA and RV12 exceed 0.95 even pre-1990, while SV exceeds 0.98 throughout. Second, both the level and standard deviation of the FUI rise sharply post-2000. The post-2000 standard deviation of the EWMA-based FUI is 0.0135, 3.5 times its pre-1990 level of 0.0039. The increase in variation is economically important: it implies that the post-2000 period contains far more signal for predictive purposes than the earlier subperiods.

5.4 Economic Recession Behavior

Having established the properties of the FUI, I examine how factor uncertainty behaves around NBER-dated recessions and whether its cyclical pattern reflects systematic co-movement with the business cycle.

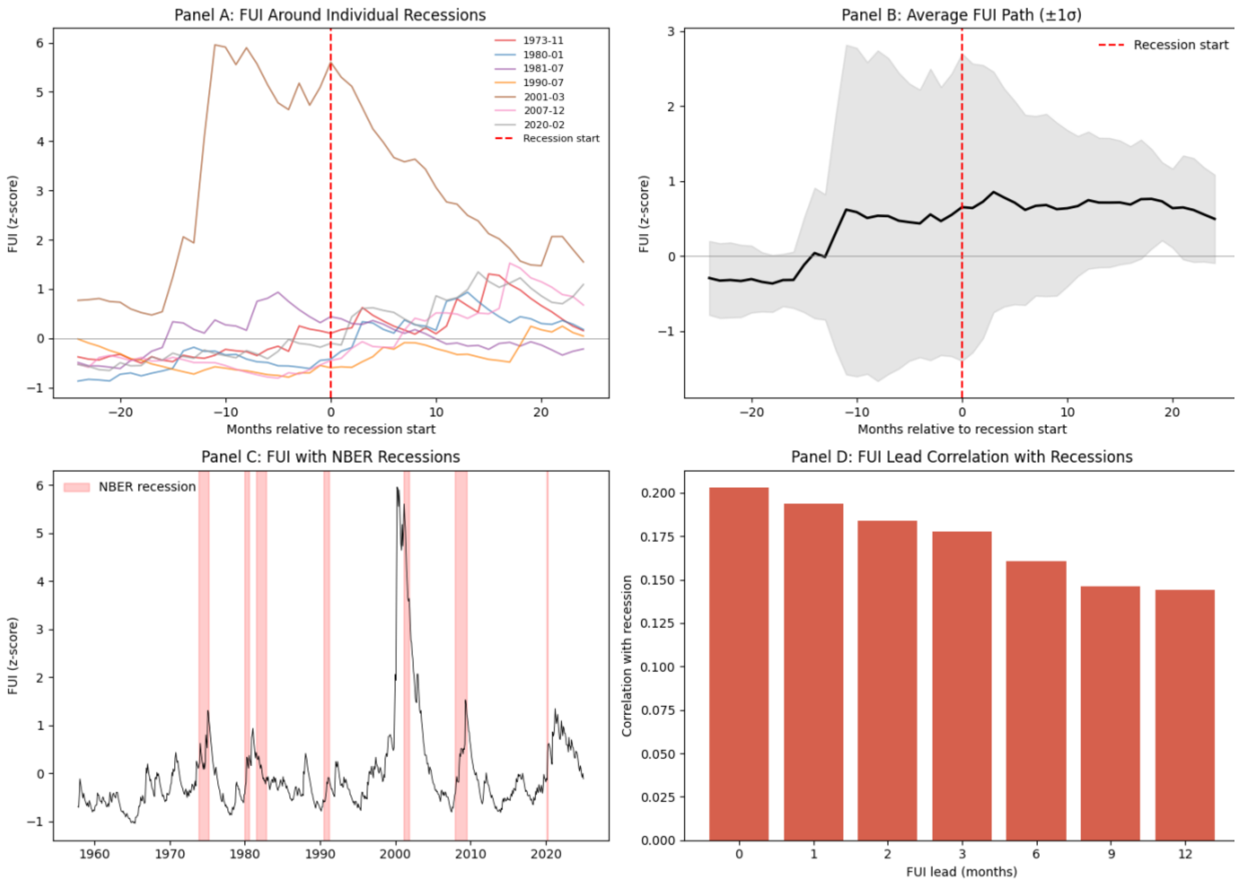

Figure 2: FUI Behavior Around NBER Recessions

Notes: Panel A shows the EWMA-based FUI (standardized) around each of the seven NBER recession start dates over the sample period. Panel B shows the cross-episode average path with a ±1σ band. Panel C plots the EWMA-based FUI time series with NBER recession periods highlighted in pink. Panel D reports Pearson correlations between the FUI and the NBER recession indicator at lead horizons of 0 to 12 months. All FUI values are expressed as z-scores relative to the full-sample mean and standard deviation.

Table 6: FUI Behavior Around NBER Recessions (EWMA Method, Standardized Units)

Panel A: Episode-Level FUI Around Recession Start Date

Episode	t−12m	t−6m	t−3m	t = 0 (Start)	t+3m	t+6m	Peak (month)
1973–11	−0.41σ	−0.23σ	0.25σ	0.10σ	0.62σ	0.24σ	1.30σ (m+15)
1980–01	−0.27σ	−0.49σ	−0.59σ	−0.43σ	0.33σ	0.10σ	0.93σ (m+13)
1981–07	0.10σ	0.81σ	0.58σ	0.44σ	0.28σ	0.17σ	0.93σ (m−5)
1990–07	−0.65σ	−0.74σ	−0.70σ	−0.60σ	−0.48σ	−0.22σ	0.24σ (m+22)
2001–03	**4.09σ**	**5.15σ**	**5.18σ**	**5.60σ**	**4.68σ**	**3.67σ**	**5.96σ** (m−11)
2007–12	−0.49σ	−0.79σ	−0.73σ	−0.45σ	−0.07σ	−0.19σ	1.52σ (m+17)
2020–02	−0.24σ	−0.42σ	−0.11σ	−0.10σ	0.60σ	0.52σ	1.34σ (m+14)

Panel B: Average FUI Path Around Recession Onset

t−24	t−18	t−12	t−9	t−6	t−3	t+0	t+3	t+6	t+12
−0.29σ	−0.37σ	0.31σ	0.51σ	0.47σ	0.55σ	**0.65σ**	0.85σ	0.61σ	0.74σ

Panel C: Lead Correlations Between FUI and NBER Recession Indicator

Lead 0m	Lead 1m	Lead 2m	Lead 3m	Lead 6m	Lead 9m	Lead 12m
ρ = 0.203	ρ = 0.194	ρ = 0.184	ρ = 0.178	ρ = 0.161	ρ = 0.146	ρ = 0.144

Notes: *FUI values are expressed in standardized units (σ), computed as deviations from the full-sample mean divided by the full-sample standard deviation. Panel A reports FUI levels at selected months relative to each NBER recession start date. The dot-com era (highlighted) is a notable outlier. Panel B reports the cross-episode average FUI path; t+0 (highlighted) denotes recession onset. Panel C reports Pearson correlations between the EWMA-based FUI and the NBER recession dummy at various lead lengths; a positive lead means FUI precedes the recession indicator.*

Figure 2, Panel C and Table 6, Panel A show that the largest spike in the standardized FUI occurs during the 2000–2002 dot-com period, when it reaches 5.96σ in March 2000. This is approximately 11 months before the official NBER recession start date, with the index still at 5.60σ at recession onset. The Global Financial Crisis of 2007–2009 produces a strikingly smaller response: the FUI stood at −0.45σ at recession onset and reached its peak of 1.52σ seventeen months later. This asymmetry may surprise readers who recall the GFC as the defining event of modern financial history. It reflects the fact that the GFC was predominantly a market-level event, while the dot-com era generated extreme cross-sectional divergence in factor performance, precisely the phenomenon the FUI is designed to capture.

Table 6, Panel B reports the average standardized FUI path around NBER recessions. The FUI is below zero at t−24 and t−18 (−0.29σ and −0.37σ), before rising on average from t−12 onward, and reaching 0.85σ three months after recession onset. This suggests that factor uncertainty begins building roughly one year before recessions begin on average, consistent with the leading indicator evidence in Panel A. However, the mean results are heavily influenced by the dot-com episode, which dominates the average path given its outsized spike relative to other recessions.

Table 6, Panel C reports lead-lag correlations between the FUI and the NBER recession indicator. The contemporaneous correlation is 0.203 and remains meaningful at a 12-month lead ($\rho = 0.144$), suggesting the FUI carries a modest but persistent signal about future recession risk across all horizons examined. The slow decay in correlation with lead length indicates that elevated factor uncertainty tends to precede recessions on average. This leading indicator property is partly driven by the 2001 episode, where the FUI led the recession start by approximately 11 months, though the positive correlations across all lead horizons suggest the pattern is not entirely episode-specific.

In conclusion, the results Figure 2 and Table 6 suggest that the FUI's relationship with the business cycle is more episodic rather than systematic. It spikes dramatically during periods of broad cross-sectional factor disruption, most clearly the dot-com episode, but provides little consistent signal ahead of recessions driven primarily by aggregate demand or financial sector stress. This is consistent with the FUI's design: it captures divergence across factor strategies rather than the level of macroeconomic uncertainty, and factor divergence is not a general feature of all recessions.

6 Predictive Power of the FUI

This section presents the paper's main empirical results. All predictive regressions use the beginning-of-month-timed FUI to ensure strict ex ante predictability. The FUI is a generated regressor, as the conditional volatility estimates in Layer 2 depend on estimated forecast errors from Layer 1. Block bootstrap with full pipeline re-estimation (2,000 replications, block length 12 months) confirms that generated-regressor bias is negligible: the bootstrap standard error is 1.03 times the Newey-West standard error for the dispersion prediction at $h = 1$.

6.1 Return Prediction: Baseline Results

I estimate the predictive regression:

$$r_{t\to t+h} = a + b \cdot FUI_{t,BOM} + \varepsilon_{t+h},$$

where $r_{t\to t+h}$ is the equal-weighted cumulative average return across the 13 factor themes from month t to $t+h$.

Table 7: Predictive Regressions of FUI on Factor Market Outcomes

This table reports results from predictive regressions of the form: $Y_{t+}^{h} = \alpha + \beta \cdot FUI_t + \varepsilon_{t+}^{h}$, where FUI_t is measured at the beginning of month t and Y_{t+}^{h} is the h-month cumulative outcome. Panel A: equal-weighted average factor theme return. Panel B: cross-sectional standard deviation of cumulative factor theme returns. Panel C: future rolling volatility of the average factor theme return. FUI variants: EWMA, GARCH, RV12, and SV. Horizons h ∈ {1, 3, 6, 12} months. Standard errors are Newey-West HAC-adjusted with horizon-matched lags. Sample sizes reflect each method's full available sample: EWMA (N = 800), GARCH (N = 691), RV12 (N = 804), and SV (N = 811).

Panel A: FUI → Average Factor Return

	EWMA			GARCH			RV12			SV		
h	**β**	**t**	R^2%	**β**	**t**	R^2%	**β**	**t**	R^2%	**β**	**t**	R^2%
h=1	0.122	2.27**	4.05	0.161	2.18**	4.13	0.124	2.24**	4.33	0.276	2.05**	3.64
h=3	0.382	3.79***	11.80	0.504	3.90***	11.88	0.385	4.15***	12.42	0.846	3.27***	10.10
h=6	0.770	4.35***	21.34	0.976	3.89***	19.67	0.778	4.72***	22.50	1.751	3.81***	19.23
h=12	1.441	3.87***	28.25	1.687	3.22***	21.90	1.416	4.16***	28.35	3.542	3.90***	29.96

Notes: *Newey-West HAC t-statistics with horizon-matched lags. Significance: *** p<0.01, ** p<0.05, * p<0.10. The FUI signal is measured at the beginning of month t (BOM timing) and used to forecast h-month-ahead outcomes. Sample sizes are similar for Panels B and C; h = 12 reduces N by approximately 11 observations per method due to forward return construction.*

Panel B: FUI → Cross-Sectional Factor Return Dispersion

	EWMA			GARCH			RV12			SV		
h	**β**	**t**	R^2%	**β**	**t**	R^2%	**β**	**t**	R^2%	**β**	**t**	R^2%
h=1	0.850	7.87***	33.07	0.969	5.79***	25.00	0.826	8.48***	32.48	2.537	9.82***	51.73
h=3	1.262	6.41***	23.14	1.395	5.01***	16.46	1.257	6.39***	23.82	3.961	6.55***	40.05
h=6	1.835	4.82***	25.05	1.960	3.24***	16.66	1.840	4.68***	26.14	5.371	6.35***	37.69
h=12	3.018	4.59***	32.74	3.135	3.44***	20.55	2.881	4.38***	30.89	8.953	6.98***	50.61

Notes: *Newey-West HAC t-statistics with horizon-matched lags. Significance: *** p<0.01, ** p<0.05, * p<0.10. The FUI signal is measured at the beginning of month t (BOM timing) and used to forecast h-month-ahead outcomes.*

Panel C: FUI → Future Volatility of Average Factor Return

	EWMA			GARCH			RV12			SV		
h	**β**	**t**	R^2%	**β**	**t**	R^2%	**β**	**t**	R^2%	**β**	**t**	R^2%
h=3	0.259	5.86***	39.54	0.295	4.31***	29.64	0.259	6.12***	41.17	0.685	7.59***	48.67
h=6	0.263	8.11***	49.74	0.299	5.28***	36.96	0.259	8.02***	50.10	0.718	13.26***	65.29
h=12	0.241	6.49***	48.62	0.264	4.40***	33.32	0.233	6.24***	47.12	0.692	11.80***	70.36

Notes: *h=1 omitted (rolling standard deviation requires ≥2 observations). Newey-West HAC t-statistics with horizon-matched lags. Significance: *** p<0.01, ** p<0.05, * p<0.10. The FUI signal is measured at the beginning of month t (BOM timing) and used to forecast h-month-ahead outcomes.*

Panel A of Table 7 shows that the FUI is a robust positive predictor of average factor theme returns at all horizons examined. The relationship strengthens markedly with the forecast horizon: the EWMA-based FUI explains only 4% of return variation one month ahead but accounts for over 28% at the 12-month horizon, with t-statistics ranging from 2.27 to 4.35 across horizons. This pattern is consistent with a factor uncertainty premium: periods of elevated cross-sectional factor uncertainty are followed by higher average factor returns, as investors demand compensation for bearing that uncertainty.

The predictability is not an artifact of a particular volatility estimator. At h = 6 months, t-statistics range from 3.81 to 4.72 across all four FUI variants, and R^2 values range from 19% to 23%. The SV-based FUI produces broadly similar results, with R^2 reaching 30% at h = 12 months, marginally above the EWMA equivalent of 28%, though the difference is economically modest. The consistency across methods with very different assumptions about the volatility process, covering deterministic smoothing, GARCH dynamics, realized volatility, and latent stochastic volatility, substantially strengthens the case that the predictability reflects a genuine feature of the data rather than a specification-specific artifact.

Table A.5 reports Stambaugh (1999) bias-corrected results, applied to account for the high persistence of the FUI. The correction has limited impact at longer horizons, where the predictability results are strongest: at h = 12 months, corrected t-statistics are 5.03 (EWMA), 3.07 (GARCH), 8.01 (RV12), and 2.02 (SV), all remaining statistically significant, though SV is only marginally significant at the 5% level. At shorter horizons, the correction is more consequential. Corrected betas turn negative for EWMA and RV12 at h = 1 and h = 3, and for SV at h = 1 through h = 6. GARCH is more robust, remaining positive from h = 3 to h = 12.

The sign reversal at the shortest horizons in the Stambaugh-corrected estimates warrants additional comment. The correction is most consequential when the predictor is highly persistent and its innovations are correlated with the regression disturbance, and the correction is well documented that it can behave unreliably in highly persistent finite samples (Campbell and Yogo, 2006). All four FUI predictors are persistent, but the degree differs materially across methods, which helps explain why the finite-sample adjustment is mild for GARCH and substantially stronger for EWMA, RV12, and especially SV. Accordingly, the short-horizon predictive results should be interpreted cautiously. By contrast, the h = 12 result remains positive and statistically meaningful across all four specifications after correction, indicating that the most robust evidence is concentrated at the 12-month horizon.

The individual factor regressions involve 13 simultaneous hypothesis tests at each horizon, raising the concern that some significant results reflect false discoveries rather than genuine predictability. I assess this using the Benjamini-Hochberg (1995) false discovery rate procedure with a 10% FDR threshold (Table A.6). The correction has negligible impact at h = 3, 6, and 12, where all 12, 12, and 10 factors significant at the raw 10% level survive BH adjustment. The correction is most consequential at h = 1, where 7 factors are significant at the raw 10% level but only 2 survive BH adjustment (accruals and debt issuance), consistent with the generally weaker and noisier short-horizon predictability documented throughout. Notably, momentum is the one factor that is never significant at any horizon under either the raw or adjusted criterion. The aggregate equal-weighted portfolio result, which constitutes a single test per horizon, is not subject to this concern.

Table 8: Subperiod Predictive Regressions of FUI on Factor Market Outcomes

Method	Period	N	β	t-stat	R^2%
EWMA	Pre-1990	386	0.2492	0.77	0.80%
	1990–1999	120	−0.4074	−1.04	2.38%
	Post-2000	294	**1.0584**	**7.04*** **	**44.54%**
GARCH	Pre-1990	277	0.2499	0.43	0.33%
	1990–1999	120	−0.8996	−1.39	4.17%
	Post-2000	294	**1.2948**	**5.31*** **	**38.22%**
RV12	Pre-1990	390	0.2784	0.91	1.12%
	1990–1999	120	−0.3436	−0.83	1.78%
	Post-2000	294	**1.0575**	**7.68*** **	**46.07%**
SV	Pre-1990	397	0.6636	0.83	1.01%
	1990–1999	120	−0.2602	−0.38	0.60%
	Post-2000	294	**2.6422**	**7.45*** **	**45.02%**

Notes: *Predictive regressions of the form: $y(t+6) = \alpha + \beta \cdot FUI(t-1) + \varepsilon$ with Newey–West standard errors (6 lags). y(t+6) is the 6-month cumulative equal-weighted average factor theme return. FUI(t−1) is the beginning-of-month FUI. Subperiods: Pre-1990 (full history to December 1989), 1990–1999, Post-2000 (January 2000 onward). Positive and statistically significant (≥ 1.96) betas, t-statistics and R^2 are bolded. *, **, *** denote 10%, 5%, 1% significance.*

While the full-sample results in Table 7 establish a robust predictive relationship, a natural question is whether this relationship is stable over time or concentrated in a particular historical episode. Table 8 presents subperiod predictive regressions at h = 6 for all four FUI variants. The subperiod results reveal substantial heterogeneity that the full-sample estimates conceal.

In the pre-1990 subsample, all four methods produce positive but insignificant slopes ($t = 0.43$–0.91) with R^2 below 2%. The 1990–1999 decade yields negative point estimates across all methods ($\beta = -0.26$ to -0.90), though none reaches conventional significance ($t = -0.38$ to -1.39). The post-2000 subsample tells a sharply different story. All four methods produce uniformly large and highly significant slopes: t-statistics range from 5.31 (GARCH) to 7.68 (RV12), with R^2 of 38–46%, roughly double the full-sample figures.

One potential explanation for the higher post-2000 betas and explanatory power is that the period contains a higher concentration of major uncertainty episodes, including the dot-com era, the Global Financial Crisis, and the COVID-19 shock. The evidence suggests that the full-sample estimates reflect an underlying average effect that is attenuated when calmer periods are pooled with high-uncertainty episodes, and that the predictive relationship is strongest when uncertainty is economically most salient.

The stronger post-2000 predictability likely reflects at least two forces. The post-2000 FUI standard deviation is more than three times its pre-1990 level (Table 5), suggesting that part of the improvement reflects greater variation in the predictor itself rather than a structural shift in the underlying relationship; the pre-2000 FUI simply did not move enough to generate detectable return predictability in finite samples. Beyond this mechanical explanation, shifts in how factor-related information is incorporated into prices may also play a role, though the data do not permit clean separation among these channels. What the subperiod evidence does rule out is that the full-sample result is driven by a single extreme episode: post-2000 predictability persists across a period containing three distinct uncertainty shocks of different economic origins, and the full-sample estimates therefore reflect an underlying average effect that is attenuated when calmer periods are pooled with high-uncertainty episodes.

Table A.7 decomposes the post-2000 period further to examine whether the predictive relationship persists after the dot-com era or whether the strong post-2000 results documented in Table 8 are entirely driven by the 2000–2002 spike. The results show that predictability does not disappear after the dot-com episode, but it is highly uneven across subperiods. The post-GFC decade of 2010–2019 produces near-zero and negative coefficients across all horizons, with the EWMA slope reaching -0.55 at $h = 6$ ($t = -1.93$) and -1.41 at $h = 12$ ($t = -2.27$). This aligns with the post-GFC period having the lowest average FUI level and the narrowest range in the entire sample (mean z-score = -0.36, maximum $z = 0.45$), reflecting a prolonged low-volatility regime where cross-sectional factor divergence was muted. When the predictor barely moves, regression-based predictability is mechanically attenuated regardless of the true underlying relationship. Beyond this mechanical explanation, the risk premium interpretation provides a complementary economic rationale: when uncertainty is low and stable, the compensation investors demand for bearing cross-sectional factor risk is small and difficult to detect in finite samples, whereas elevated uncertainty generates a larger and more detectable premium. By contrast, the COVID+ period generates the strongest predictability after the dot-com era, with slopes of 3.41 at h = 6 (t = 4.72) and 7.54 at h = 12 (t = 8.25), consistent with both elevated FUI levels and a large cross-sectional uncertainty premium. Overall, the within-post-2000 decomposition reinforces the episodic interpretation: the predictive relationship activates sharply during periods of genuine cross-sectional factor disruption and lies dormant during calmer regimes, confirming that the post-2000 result in Table 8 is not solely a dot-com artifact.

6.2 Asymmetric and Regime-Conditional Results

In Section 6.1, I find that factor uncertainty is not equally informative across all subperiods. If the FUI captures a risk premium that only activates during periods of genuine cross-sectional stress, one would expect its predictive power to be concentrated in high-uncertainty regimes rather than distributed uniformly across the full range of FUI values.

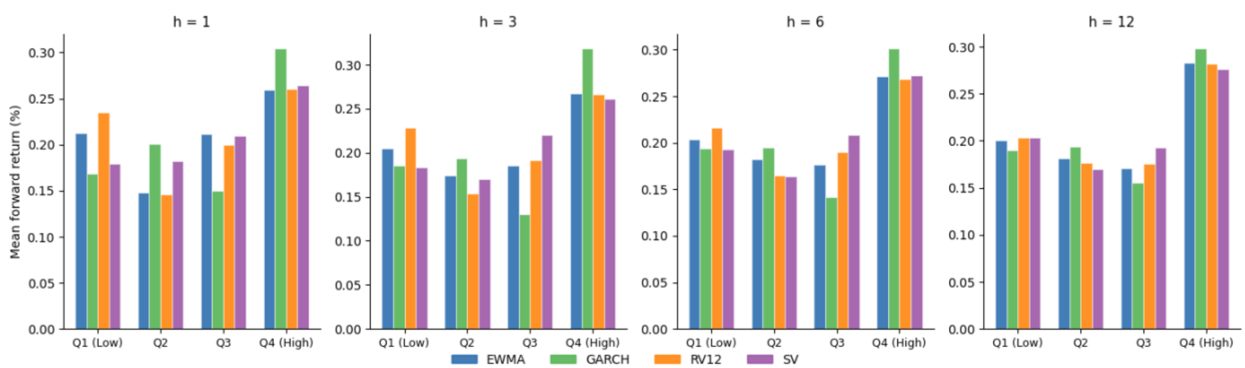

Figure 3: Equal-Weighted 13-Factor Theme Returns Across FUI Quartiles (h = 1, 3, 6, 12)

Notes: Each panel sorts months into quartiles based on the beginning-of-month FUI level and reports the mean monthly equal-weighted factor theme return within each quartile. Q1 (Low) contains the 25% of months with the lowest FUI values and Q4 (High) contains the 25% with the highest. Four FUI variants are shown: EWMA (λ = 0.85), GARCH(1,1), RV12, and Stochastic Volatility (SV). Horizons $h \in \{1, 3, 6, 12\}$ months. Returns are expressed as monthly percentages. The sample covers the full available period for each method. The Q4 jump visible across all methods and horizons is consistent with a threshold effect in the factor uncertainty premium, whereby elevated cross-sectional volatility is associated with meaningfully higher subsequent factor returns while low-to-moderate uncertainty levels carry little predictive content.

The relationship between FUI and subsequent factor returns is not monotonically linear. Figure 3 illustrates this nonlinearity. Quartile analysis, which sorts months by FUI level and computes average forward returns within each quartile, reveals that returns are broadly flat across Q1 through Q3 before jumping sharply in Q4 (high FUI). This pattern suggests a threshold effect concentrated in the highest FUI quartile rather than a smooth linear relationship.

Table 9: Q4–Q1 Return Spreads by Volatility Method and Horizon

Horizon	FUI(EWMA)	FUI(GARCH)	FUI(RV12)	FUI(SV)
h = 1				
Spread (%)	**0.046**	**0.136***	**0.026**	**0.139***
t-statistic	0.68	1.80	0.37	1.78
p-value	0.497	0.073	0.710	0.076
h = 3				
Spread (%)	**0.063**	**0.134*****	**0.037**	**0.131*****
t-statistic	1.47	2.90	0.88	2.72
p-value	0.143	0.004	0.381	0.007
h = 6				
Spread (%)	**0.068****	**0.107*****	**0.052**	**0.123*****
t-statistic	2.08	3.00	1.63	3.33
p-value	0.038	0.003	0.104	0.001
h = 12				
Spread (%)	**0.082*****	**0.109*****	**0.078*****	**0.123*****
t-statistic	3.08	3.63	2.95	4.07
p-value	0.002	<0.001	0.003	<0.001

Notes: The table reports the Q4–Q1 return spread (% per month), HAC t-statistic, and p-value for each volatility estimation method across investment horizons h = 1, 3, 6, 12 months. Spreads are computed as the difference in mean forward factor returns between the highest (Q4) and lowest (Q1) FUI quartile. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table 9 reports Q4-Q1 return spreads across all four volatility estimation methods and horizons, confirming that the nonlinear relationship between factor uncertainty and returns is both economically meaningful and statistically robust. At short horizons the spreads are modest and statistically mixed, with only GARCH and SV reaching significance at h = 1, but predictability builds steadily with horizon. At h = 12, all four methods yield uniformly large and highly significant spreads ranging from 0.078% per month (RV12) to 0.123% per month (SV), with t-statistics between 2.95 and 4.07. The consistency across methods is notable: despite operating at different levels and using different estimation approaches, all four FUI variants identify the same high-uncertainty quartile as the one associated with elevated subsequent returns. This pattern is difficult to reconcile with a smooth linear risk-return relationship and instead points to a threshold effect in which the volatility premium becomes economically relevant only when factor uncertainty exceeds a critical level, consistent with a time-varying uncertainty premium that activates when cross-sectional volatility is sufficiently elevated.

Table 10: Regime-Conditional Predictability — R², β, and t-statistic by Method and Horizon

Method	Regime	h = 1	h = 3	h = 6	h = 12
		R^2 / β / t-stat	*R^2 / β / t-stat*	*R^2 / β / t-stat*	*R^2 / β / t-stat*
EWMA	High FUI	6.92% / 0.174 / 2.53	17.51% / 0.172 / 4.51	32.28% / 0.175 / 6.95	43.79% / 0.166 / 6.75
	Low FUI	0.45% / -0.137 / -1.51	0.19% / -0.051 / -0.67	0.22% / -0.042 / -0.55	0.05% / -0.015 / -0.22
GARCH	High FUI	5.42% / 0.197 / 2.15	17.49% / 0.218 / 4.47	29.20% / 0.217 / 5.99	33.58% / 0.188 / 5.12
	Low FUI	0.05% / 0.079 / 0.38	0.06% / 0.054 / 0.38	0.00% / 0.002 / 0.01	0.17% / -0.054 / -0.36
RV12	High FUI	7.76% / 0.181 / 2.59	18.80% / 0.173 / 5.21	33.99% / 0.175 / 8.04	43.65% / 0.162 / 7.14
	Low FUI	0.51% / -0.136 / -1.63	0.96% / -0.114 / -1.46	0.93% / -0.083 / -1.20	0.66% / -0.054 / -0.91
SV	High FUI	4.58% / 0.357 / 1.93	11.93% / 0.355 / 3.22	24.33% / 0.384 / 4.55	42.84% / 0.414 / 7.16
	Low FUI	0.26% / 0.209 / 0.97	0.04% / 0.048 / 0.27	0.22% / 0.080 / 0.43	0.29% / 0.080 / 0.43

Notes: *The table reports R^2 (%), slope coefficient β, and HAC t-statistic from predictive regressions of forward factor returns on FUI. 'High FUI' ('Low FUI') restricts to months with FUI above (below) the sample median. Horizons h = 1, 3, 6, 12 are h-month ahead return windows.*

To examine whether this threshold effect is reflected in the predictive regression itself rather than just in unconditional return differences across quartiles, Table 10 splits the sample into High and Low FUI regimes and re-estimates the predictive regressions within each. In High FUI regimes, all four methods produce substantially larger betas and R^2 values than the Low FUI estimates at every horizon. At h = 6, High FUI R^2 values range from 24% (SV) to 34% (RV12), compared to Low FUI figures of 0% to 1%, and t-statistics range from 4.55 to 8.04 in High FUI regimes. At h = 12, High FUI R^2 values reach 34% to 44% for EWMA and RV12. The R^2 values are negligible throughout, falling below 1% in most cases. This sharp contrast between regimes, with predictability concentrated in the high uncertainty half of the sample, supports the interpretation that the factor uncertainty premium activates only when cross-sectional volatility is sufficiently elevated.

6.3 Horse Race with Market Volatility and Uncertainty Measures

To assess the incremental information content of the FUI, I compute correlations between FUI variants and uncertainty benchmarks along with the bivariate regressions that pair each of the four FUI variants with three comparison variables: VIX, JLN macro uncertainty, and Baker-Bloom-Davis economic policy uncertainty (EPU).

Table 11: Pearson Correlations Between FUI Variants and Uncertainty Benchmarks

This table reports pairwise correlations between the four FUI variants and three standard uncertainty benchmarks: VIX (CBOE implied volatility index), EPU (Baker, Bloom, and Davis policy uncertainty index), and JLN (Jurado, Ludvigson, and Ng macroeconomic uncertainty index).

	VIX	EPU	JLN
EWMA	0.335*** (N=420)	0.240*** (N=806)	0.291*** (N=774)
GARCH	0.308*** (N=420)	0.145*** (N=697)	0.233*** (N=697)
RV12	0.325*** (N=420)	0.246*** (N=810)	0.286*** (N=774)
SV	0.401*** (N=420)	0.273*** (N=817)	0.329*** (N=774)

Notes: *Standard Pearson correlations. All are significant at the 1% level (Pearson t-test). Sample sizes vary across benchmarks due to data availability. VIX sample is shorter (N=420) reflecting its later start date. *** p<0.01.*

Table 11 reports pairwise correlations between the four FUI variants and the three benchmark uncertainty indices. All correlations are positive and statistically significant, yet uniformly modest in magnitude, with no correlation exceeding 0.41. SV produces the strongest comovement across all three benchmarks, while GARCH is consistently the weakest, particularly against EPU (0.145). Across benchmarks, VIX exhibits the highest correlations for all four methods, followed by JLN and then EPU, reflecting the fact that VIX responds sharply to equity market stress, which is the primary driver of cross-sectional factor return dispersion, while EPU captures policy-specific uncertainty that need not coincide with elevated factor volatility. The modest correlations overall indicate that the FUI captures dimensions of factor-level uncertainty that are distinct from these aggregate indices.

Table 12: Horse Race Regressions: FUI vs Alternative Uncertainty Measures

Each cell reports results from OLS regressions of h-month forward average factor returns on FUI and the comparison variable, estimated with HAC standard errors. R^2 values in percent.

Panel A: FUI vs JLN (JLN macro uncertainty, Jurado, Ludvigson & Ng 2015).

Method	h	FUI R^2	JLN R^2	Joint R^2	FUI t	JLN t	N
EWMA	1	4.37%	0.01%	4.90%	**2.40****	-1.41	772
	6	22.55%	0.14%	26.01%	**5.38*****	-1.71*	767
	12	29.39%	0.00%	32.11%	**4.55*****	-1.47	761
GARCH	1	4.13%	0.01%	4.52%	**2.24****	-1.23	696
	6	19.67%	0.25%	22.38%	**4.28*****	-1.49	691
	12	21.90%	0.02%	23.69%	**3.48*****	-1.14	685
RV12	1	4.79%	0.01%	5.37%	**2.42****	-1.47	772
	6	24.14%	0.14%	27.84%	**6.07*****	-1.81*	767
	12	29.54%	0.00%	32.33%	**5.00*****	-1.53	761
SV	1	3.99%	0.01%	4.59%	**2.19****	-1.41	772
	6	21.09%	0.14%	25.04%	**4.86*****	-1.74*	767
	12	32.21%	0.00%	35.97%	**5.34*****	-1.69*	761

Panel B: FUI vs EPU (Baker-Bloom-Davis economic policy uncertainty).

Method	h	FUI R^2	EPU R^2	Joint R^2	FUI t	EPU t	N
EWMA	1	4.05%	1.76%	7.66%	**2.74****	**-4.69*****	805
	6	21.34%	4.14%	32.36%	**6.34*****	**-5.47*****	800
	12	28.25%	4.31%	40.80%	**5.83*****	**-5.09*****	794
GARCH	1	4.13%	2.11%	7.28%	**2.45****	**-4.35*****	696
	6	19.67%	5.31%	28.61%	**4.65*****	**-4.52*****	691
	12	21.90%	5.55%	31.36%	**3.93*****	**-3.51*****	685
RV12	1	4.33%	1.81%	8.06%	**2.75****	**-4.85*****	809
	6	22.50%	4.33%	33.96%	**7.03*****	**-5.77*****	804
	12	28.35%	4.32%	40.81%	**6.18*****	**-5.12*****	798
SV	1	3.64%	1.71%	7.27%	**2.52****	**-4.47*****	816
	6	19.23%	4.37%	30.96%	**5.39*****	**-5.33*****	811
	12	29.96%	4.54%	44.23%	**6.64*****	**-5.90*****	805

Panel C: FUI vs VIX (CBOE VIX). Post-1990 subsample.

Method	h	FUI R^2	VIX R^2	Joint R^2	FUI t	VIX t	N
EWMA	1	7.66%	0.60%	11.17%	**3.11****	**-3.07****	418
	6	36.63%	0.78%	47.08%	**8.43*****	**-4.48*****	413
	12	44.99%	0.02%	52.29%	**5.57*****	**-4.50*****	407
GARCH	1	6.80%	0.60%	9.77%	**2.89****	**-2.96****	418
	6	30.68%	0.78%	38.90%	**5.59*****	**-3.44*****	413
	12	32.87%	0.02%	37.62%	**4.98*****	**-2.82****	407
RV12	1	8.21%	0.60%	11.66%	**3.10****	**-3.07****	418
	6	38.46%	0.78%	48.56%	**9.04*****	**-4.34*****	413
	12	45.13%	0.02%	51.83%	**7.80*****	**-4.19*****	407
SV	1	7.06%	0.60%	11.21%	**2.96****	**-3.18****	418
	6	35.09%	0.78%	48.20%	**8.12*****	**-5.08*****	413
	12	50.61%	0.02%	61.79%	**10.98*****	**-5.87*****	407

Notes: The dependent variable is the h-month forward equal-weighted average factor return (h = 1, 6, 12). All regressions use HAC standard errors with maxlags = h. FUI is the Factor Uncertainty Index (beginning-of-month, lagged by one period). Panel A comparison variable is the JLN macro uncertainty index (h=1 horizon). Panel B comparison variable is the Baker-Bloom-Davis economic policy uncertainty index, normalized to the same scale as FUI. Panel C comparison variable is the CBOE VIX, available post-1990. Significant t-statistics ($|t| \geq 1.96$) are shown in bold. *, **, *** denote significance at the 10%, 5%, and 1% levels respectively.

Table 12 reports the results of horse race regressions on FUI to JLN, EPU, and VIX. The FUI–VIX horse race, estimated over the post-1990 subsample ($N \approx 418$), produces the most interesting result. Across all four methods, the FUI retains strong significance at all horizons ($t = 2.89$ to 10.98), while VIX enters with a negative and significant coefficient ($t = -2.82$ to -5.87) despite having near-zero standalone predictive power for returns ($R^2 < 1\%$ at all horizons). The joint R^2 substantially exceeds the FUI-only R^2 at all horizons. At $h = 12$, for example, the joint R^2 ranges from 38% (GARCH) to 62% (SV), compared to standalone FUI R^2 of 33% and 51% respectively, indicating that VIX provides complementary information with opposite sign rather than simply proxying for the FUI. The negative VIX coefficient implies that, conditional on factor-specific uncertainty, higher market volatility predicts lower factor returns, consistent with a risk-off channel that compresses factor premia during broad market stress episodes.

The FUI–EPU horse race produces equally striking complementarity. Despite modest full-sample correlations between the FUI and EPU ($\rho = 0.14$–0.27 depending on method), both variables are highly significant in the joint regression across all methods and horizons: FUI t-statistics range from 2.45 to 7.03 (positive), while EPU t-statistics range from −3.51 to −5.90 (negative). The opposite signs have a natural economic interpretation. The FUI captures risk compensation, whereby high factor-level volatility raises required returns, while EPU captures a demand effect where elevated policy uncertainty reduces investor risk appetite and thereby compresses realized factor returns. The two measures are therefore complementary rather than substitutes, each capturing a distinct channel through which uncertainty affects factor performance. At $h = 12$, the joint R^2 reaches 31–44% across methods, compared to EPU-only R^2 of 4–6%, confirming that EPU has modest but genuine incremental content beyond the FUI rather than the reverse.

The FUI–JLN horse race confirms that JLN macro uncertainty adds nothing to the FUI for factor return prediction, and this result is uniform across all four methods. In every combination of method and horizon, JLN is statistically insignificant ($t = -1.14$ to -1.81), while the FUI retains significance throughout ($t = 2.19$ to 6.07). The joint R^2 exceeds the FUI-only R^2 by at most 3 percentage points, indicating negligible incremental content. FUI–JLN correlations are low across all methods ($\rho = 0.23$–0.33), yet JLN still carries no predictive power once the FUI is included. The FUI fully subsumes whatever signal JLN might contain for this application, which is consistent with the interpretation that factor-level uncertainty is empirically distinct from, and more relevant than, aggregate macroeconomic uncertainty for predicting factor returns.

Overall, the horse race results establish that the FUI captures a distinct dimension of uncertainty that is not subsumed by existing market-level or macroeconomic uncertainty measures. Its predictive content for factor returns survives the inclusion of VIX, EPU, and JLN as competing variables, with the FUI remaining significant in every joint regression across all four methods and horizons.

6.4 Out-of-Sample Predictability

The in-sample results in Section 5.1 establish that the FUI has strong predictive content for average factor returns, but in-sample fit is well-known to overstate the practical value of a predictor due to parameter estimation error and look-ahead bias. I therefore assess out-of-sample predictability using the Campbell and Thompson (2008) R^2_OS statistic and the Clark and West (2007) MSFE-adjusted test. At each month t, I estimate the predictive regression using only data available up to t (expanding window) and generate a genuine out-of-sample forecast for the h-month ahead average factor return. The benchmark is the prevailing historical mean — the simplest possible competing predictor.

Table 13: Out-of-Sample Predictability of the FUI

Panel A reports Campbell & Thompson (2008) out-of-sample R^2 and Clark & West (2007) MSFE-adjusted statistics for all four FUI methods at horizons h = 1, 3, 6, 12 months. Panel B reports subperiod OOS R^2 for EWMA.

Panel A: Out-of-Sample R^2 and Clark-West Statistics — Full Sample

Method	h	R²_OS	CW stat	p-value	N_OOS
EWMA	1	**1.65%**	1.413*	**0.0788***	686
	3	**6.96%**	1.982**	**0.0238****	682
	6	**10.69%**	1.662*	**0.0482****	676
	12	**10.59%**	1.373*	**0.0849***	664
GARCH	1	**-0.49%**	1.027	0.1523	577
	3	**8.59%**	2.112**	**0.0173****	573
	6	**10.92%**	1.648**	**0.0496****	567
	12	**6.34%**	1.284*	**0.0996***	555
RV12	1	**1.83%**	1.383*	**0.0833***	690
	3	**7.76%**	2.007**	**0.0224****	686
	6	**12.48%**	1.719**	**0.0428****	680
	12	**12.53%**	1.401*	**0.0806***	668
SV	1	**0.71%**	0.951	0.1707	697
	3	**3.17%**	1.357*	**0.0874***	693
	6	**6.41%**	1.399*	**0.0810***	687
	12	**5.36%**	1.182	0.1185	675

Panel B: Subperiod OOS R^2 — EWMA method only

h	Pre-1990	1990–1999	Post-2000
1	-1.23%	**0.63%**	**2.71%**
6	-6.63%	-2.84%	**16.31%**
12	-5.08%	**1.53%**	**14.47%**

Notes: *Panel A: At each month t the FUI model is estimated on all targets fully realized by t (expanding window, no look-ahead) and used to forecast the h-month cumulative average factor return. The benchmark is the expanding historical mean.* $R^2_{as} = 1 - MSFE_{mo}{}^{de}{}_{l} / MSFE^{ie}{}_{n}{}^{ch}$ *(Campbell & Thompson 2008). CW stat is the Clark & West (2007) MSFE-adjusted t-statistic with HAC standard errors (Newey-West, maxlags = h−1); p-values are one-sided. Initial training window: 120 months. Panel B reports sub-sample OOS R² for the EWMA variant; negative values indicate the model underperforms the historical mean in that subperiod. Positive R²_OS values shown in blue, negative in red. *, **, *** denote 10%, 5%, 1% significance.*

Table 13 Panel A reports out-of-sample results for all four FUI methods across horizons h = 1, 3, 6, and 12. The evidence for genuine out-of-sample predictability is most robust at the three- and six-month horizon.

At the one-month horizon, R²_OS is positive but small, around 1–2% for EWMA, RV12, and SV, and marginally negative for GARCH. However, EWMA and RV12 clear the Clark-West test at the 10% level (p = 0.079 and p = 0.083 respectively), while GARCH and SV do not. This is consistent with the short-horizon results: idiosyncratic noise dominates and the signal in FUI is insufficient to reliably beat the historical mean benchmark on a month-by-month basis.

The picture strengthens at longer horizons. From h = 3 onward, all four methods produce positive R^2_OS, with CW statistics significant at least at the 10% level for most methods. At h = 3, EWMA, GARCH, and RV12 all achieve CW p-values below 5%, with R^2_OS ranging from 7% (EWMA) to 9% (GARCH). At h = 6, R^2_OS ranges from 6% (SV) to 12% (RV12), though CW significance is marginal for SV (p = 0.081). At h = 12, the out-of-sample fit remains economically meaningful. R^2_OS reaches 13% for RV12 and 11% for EWMA, but statistical evidence is weaker, with CW p-values in the 0.08–0.12 range and SV falling short of conventional thresholds entirely. This attenuation at long horizons is expected: HAC standard errors with h−1 lags are conservative on samples of 660–690 observations, and the overlapping-return construction mechanically inflates the R^2_OS metric relative to its sampling uncertainty. The consistency across volatility estimators nonetheless remains notable: despite meaningful differences in model complexity, the four methods deliver qualitatively identical out-of-sample conclusions, reinforcing that the predictability reflects a property of the FUI concept rather than any specific implementation.

The SV-based FUI, while an in-sample full-information estimate by construction, delivers lower R^2_OS than the causal EWMA and RV12 estimators at all horizons. One plausible explanation is that the SV smoother is substantially more persistent (AR(1) = 0.995), which may cause the resulting signal to adjust too slowly to emerging uncertainty episodes for real-time forecasting purposes. The weaker out-of-sample performance suggests that any advantage from backward-looking full-sample smoothing is not sufficient to offset the timing cost of this greater persistence. If look-ahead effects were the dominant force, SV would be expected to outperform the causal estimators in Table 13, which is not observed. Accordingly, SV is retained as an in-sample benchmark, while the practically relevant out-of-sample comparison centers on EWMA, GARCH, and RV12.

Panel B decomposes the EWMA out-of-sample performance by subperiod and reveals a concentration of predictability in the post-2000 period. Pre-1990 and 1990–1999 R^2_OS values are negligible or negative across all horizons, while the post-2000 subperiod drives essentially all of the full-sample result: R^2_OS reaches 3% at h = 1, 16% at h = 6, and 14% at h = 12. This pattern is consistent with the post-2000 period exhibiting more severe uncertainty episodes, including the dot-com era, the Global Financial Crisis, and the COVID-19 shock, during which factor-level uncertainty was both more volatile and more informative about subsequent compensation. The weak pre-2000 performance does not undermine the main finding. Rather, it suggests that the FUI's predictive content is most pronounced precisely when uncertainty variation is largest and economically most meaningful.

6.5 Volatility-Managed Factor Portfolios

To assess whether volatility management adds value beyond the unmanaged factor theme returns, I follow Moreira and Muir (2017) and construct volatility-managed versions of each factor theme by scaling monthly returns by the inverse of the lagged conditional variance: $R_{i,t}^{m} = (c_t/\hat{\sigma}_{i,t-1}^{2}) \cdot R_{i,t}$, where $R_{i,t}^{m}$ is the volatility-managed return on factor theme i in month t, $\hat{\sigma}_{i,t-1}^{2}$ is the EWMA conditional variance from the FUI pipeline, and c_t is a time-varying constant estimated recursively on an expanding window through $t-1$, chosen to equate the unconditional variance of the managed and unmanaged portfolios using only information available at the start of each month.

Table 14: Volatility-Managed Factor Theme Portfolios

Panel A reports individual factor theme results following Moreira & Muir (2017). Panel B reports aggregate FUI-managed portfolio results. Spanning α is from regressing managed on unmanaged returns (HAC, 6 lags).

Factor	Unmanaged SR	Managed SR	Δ SR	α (bps/m)	α t-stat
Panel A: Individual Factor Volatility-Managed Portfolios					
Accruals	0.772	0.617	−0.154	1.57	0.62
Debt Issuance	1.068	1.051	−0.017	4.44	2.76***
Investment	0.452	0.182	−0.270	−3.73	−0.99
Low Leverage	−0.005	0.194	+0.199	4.28	1.80*
Low Risk	0.156	0.287	+0.131	14.01	1.96**
Momentum	0.427	0.735	+0.309	30.97	4.60***
Profit Growth	0.612	0.699	+0.087	7.15	2.42**
Profitability	0.388	0.414	+0.026	7.72	1.86*
Quality	0.648	0.654	+0.006	4.03	1.78*
Seasonality	0.867	0.706	−0.160	0.15	0.19
Short-Term Reversal	0.479	0.419	−0.059	2.21	1.01
Size	0.157	−0.100	−0.258	−0.25	−2.72***
Value	0.367	0.134	−0.233	−0.01	−0.31
Panel B: Aggregate FUI-Managed Portfolio ($w_t = c_t / FUI^2_{t-1}$)					
Strategy	**Unmanaged SR**	**Managed SR**	**Δ SR vs raw**	**α (bps/m)**	**α t-stat**
Unmanaged (equal-weight)	1.245	—	—	—	—
Uncapped	—	1.525	+0.280	7.33	5.18***
Cap [5, 95]	—	1.548	+0.303	5.74	6.11***
Cap [10, 90]	—	1.528	+0.283	4.84	6.25***
Cap [1, 99]	—	1.530	+0.285	7.07	5.13***
Binary regime (50%)	—	1.461	+0.216	2.48	6.06***

*Notes: Panel A: Managed return $R^m_{it} = (c / \sigma^2_{i,t-1}) \cdot R_{it}$, where c is estimated recursively using an expanding window lagged one month so that only past information is used. Sharpe ratios (SR) are annualized. Panel B: Aggregate portfolio uses $w_t = c_t / FUI_{t-1}$, where FUI_{t-1} is the cross-sectional average EWMA conditional volatility across 13 factor themes at the start of month t, and c_t is estimated recursively on an expanding window through $t - 1$ to ensure no forward-looking information is used. Binary regime reduces weight to 50% when FUI exceeds its expanding median. *, **, *** denote 10%, 5%, 1% significance.*

Table 14 Panel A reports the results of individual factor theme performance. Volatility management improves the Sharpe ratio for 6 out of 13 factor themes, with the largest gains in momentum (+0.31, SR from 0.43 to 0.74), low leverage (+0.20, SR from −0.01 to 0.19), and low risk (+0.13, SR from 0.16 to 0.29). For some factor themes, however, volatility management actively hurts performance rather than helping it. Size is the clearest case, where the managed Sharpe ratio falls from 0.16 to −0.10 and the spanning alpha is significantly negative (t = −2.72), indicating that scaling back exposure during high-volatility periods removes precisely the periods when the size premium is earned.

Panel B shows that aggregating across factors substantially amplifies the net benefit of volatility management relative to the unmanaged equal-weighted benchmark (SR = 1.25). The uncapped aggregate FUI-managed portfolio achieves an annualized Sharpe ratio of 1.53, an improvement of 0.28, with a spanning alpha of 7.33 basis points per month (t = 5.18). This result is robust to weight capping: constraining weights to the 5th–95th percentile yields SR = 1.55 and alpha = 5.74 bps (t = 6.11), while the 10th–90th cap gives SR = 1.53 and alpha = 4.84 bps (t = 6.25). The stability of the alpha across capping schemes confirms that the improvement is not driven by extreme leverage during low-FUI periods. A simpler binary regime strategy, reducing exposure to 50% when the FUI exceeds its expanding median, also generates a significant spanning alpha of 2.48 bps per month (t = 6.06), with SR = 1.46. Together, these results suggest that the FUI is practically useful as a market-timing signal for scaling aggregate factor exposure, with gains that are robust to the specific implementation choice.

6.6 Economic Mechanism

The preceding evidence collectively points toward a specific economic interpretation. The temporal stability analysis shows FUI's predictive content is concentrated post-2000, coinciding with the period of highest cross-sectional factor uncertainty. The recession analysis confirms FUI captures episodic factor divergence rather than macroeconomic distress. The horse race establishes that FUI is empirically distinct from stock market volatility, policy uncertainty, and macro uncertainty. Overall, these results suggest FUI measures a time-varying factor-specific risk premium rather than a mispricing channel.

I consider three complementary tests: a limits-to-arbitrage channel test, a time-varying risk price test, and a factor characteristic decomposition. If the FUI's predictive power operates through limits-to-arbitrage, predictability should be concentrated in periods of high market stress when arbitrage constraints are tightest. I test this by augmenting the baseline predictive regression with a high-VIX indicator (equal to one when VIX exceeds its expanding median) and its interaction with the FUI.

Table 15: Economic Mechanism Tests

Panel A tests the limits-to-arbitrage channel via FUI×VIX interactions. Panel B reports Sharpe ratios across FUI quartiles. Panel C reports factor group decomposition results.

Panel A: Limits-to-Arbitrage Test — FUI × High-VIX Interaction

h	FUI t	High-VIX t	FUI×VIX t	R²
1	**4.08*** **	0.20	-0.81	8.3%
6	**3.30*** **	-1.10	0.09	38.3%
12	**3.66*** **	-0.98	0.35	45.8%

Cross-sectional test (h=6): β(FUI) regressed on unconditional factor volatility σi: b = 65.35, t = 1.24, R^2 = 12.2% (N = 13 factors). No significant cross-sectional pattern — limits-to-arbitrage channel not supported.

Panel B: Factor Sharpe Ratios Across FUI Quartiles

FUI Quartile	Mean (%/m)	Std (%/m)	SR (ann.)	N
Q1 (low FUI)	0.213	0.288	**2.56**	202
Q2	0.147	0.375	**1.36**	201
Q3	0.212	0.464	**1.58**	201
Q4 (high FUI)	0.259	0.927	**0.97**	202

Panel C: Factor Characteristic Decomposition

Factor Group	β (h=6)	t (h=6)	R² (h=6)	β (h=12)	t (h=12)	R² (h=12)
Value / Fundamental value, investment, profitability, accruals, debt issuance	1.642	**4.17*** **	20.82%	2.918	**4.07*** **	27.10%
Risk / Defensive low risk, low leverage, quality	0.154	0.79	0.29%	0.292	0.65	0.51%
Momentum / Technical momentum, short-term reversal, seasonality	0.347	1.48	1.81%	0.729	1.62	3.76%

Notes: *Panel A: Regression* $y_{t+h} = \alpha + \beta_1 \cdot FUI_t + \beta_2 \cdot HighVIX_t + \beta_3 \cdot (FUI_t \times HighVIX_t) + \varepsilon_{t+h}$. *HighVIX is an indicator for VIX above its expanding median. HAC standard errors (maxlags = h). Cross-sectional test: β(FUI)_i from factor-level regressions at h=6 regressed on unconditional factor volatility. Panel B: Equal-weighted average factor theme return within each FUI quartile. SR is annualized. Panel C: Factor groups averaged within each economic theme; FUI regressed on h-period forward group return. HAC standard errors. *, **, *** denote 10%, 5%, 1% significance.*

Panel A of Table 15 shows that the FUI×VIX interaction is statistically insignificant at all horizons (t = −0.81 to 0.35), while the FUI itself remains significant (t = 3.30 to 4.08). The limits-to-arbitrage channel is not supported by the data. Similarly, the cross-sectional test in Panel A shows no significant relationship between a factor's unconditional volatility and its FUI beta (slope coefficient b = 65.35, t = 1.24, R^2 = 12%), providing further evidence against the limits-to-arbitrage interpretation.

Panel B of Table 15 reports factor Sharpe ratios across FUI quartiles. The pattern is non-monotonic: the annualized Sharpe ratio is 2.56 in Q1 (lowest FUI), falls to 1.36 in Q2, partially recovers to 1.58 in Q3, and falls to 0.97 in Q4 (highest FUI). This decline is notable: Q4 delivers the highest mean return across all quartiles (0.26% per month vs 0.21% in Q1), yet also the highest standard deviation by a wide margin (0.93% per month vs 0.29% in Q1), with volatility more than tripling while the return increases only modestly. Conditional volatility rises faster than expected returns, compressing realized Sharpe ratios. This rules out a simple risk-price story in which investors demand higher compensation per unit of risk when uncertainty is elevated. Instead, the FUI appears to capture episodes of heightened factor volatility in which the additional expected return is more than offset by the increase in realized variance, precisely the environment where the volatility-managed portfolio construction of Section 6.5 adds the most value.

Panel C of Table 15 decomposes the FUI premium across three factor groups: Value/Fundamental (value, investment, profitability, accruals, debt issuance), Risk/Defensive (low risk, low leverage, quality), and Momentum/Technical (momentum, short-term reversal, seasonality). The FUI premium is almost entirely concentrated in Value/Fundamental factors, with t-statistics of 4.17 and 4.07 at $h = 6$ and $h = 12$ respectively and R^2 values of 21% and 27%. Risk/Defensive factors show no significant FUI predictability at either horizon ($t = 0.79$ and 0.65), and Momentum/Technical factors are also insignificant ($t = 1.48$ and 1.62). This decomposition is consistent with the interpretation that the FUI captures uncertainty about fundamental valuation, specifically the required return premium that uncertainty-averse investors demand for holding factors whose cash flow and discount rate dynamics are harder to forecast.

The three tests point toward a risk compensation interpretation rather than a limits-to-arbitrage story. The FUI reflects genuine time variation in the required return for bearing factor-level uncertainty, concentrated in fundamentally exposed factor styles, and operates independently of market-wide stress conditions, though the negative VIX coefficient in Section 6.3 suggests that market-level volatility and factor-specific uncertainty act through distinct and partially offsetting channels.

7 Conclusion

This paper introduces the Factor Uncertainty Index (FUI), a measure of aggregate uncertainty in the equity factor theme space constructed by applying conditional volatility estimation to the unpredictable components of 13 factor theme return series. The FUI is conceptually distinct from existing uncertainty measures: while VIX, JLN, and EPU capture uncertainty in aggregate markets, macroeconomic fundamentals, or policy environments respectively, the FUI measures how uncertain the factor structure itself is, specifically the degree to which the return-generating process of systematic equity strategies is unpredictable at a given point in time.

The FUI predicts future equal-weighted average factor theme returns with in-sample R^2 of 19–23% at h = 6 and 22–30% at h = 12, consistent across all four estimation methods. Out-of-sample R^2 ranges from 6% to 12% at h = 6 and 5% to 13% at h = 12, statistically significant at conventional levels by the Clark and West (2007) test. Subperiod analysis reveals that predictability is episodic rather than uniform, concentrated in periods of elevated cross-sectional factor uncertainty such as the dot-com era and the COVID-19 shock, and effectively dormant during the low-volatility periods, such as the post-GFC decade of 2010–2019.

Volatility-managed portfolios scaled inversely with the FUI generate a spanning alpha of 7.33 basis points per month (t = 5.18), improving the annualized Sharpe ratio from 1.25 to 1.53. Mechanism tests reveal that predictability is concentrated in value and fundamental factors, consistent with time-varying compensation for fundamental valuation risk rather than limits-to-arbitrage frictions.

The broader implication is that the factor space has its own uncertainty cycle, distinct from the macroeconomic and market uncertainty cycles studied in the existing literature. Periods of high factor uncertainty represent fundamentally different investment environments in which realized variance rises faster than expected compensation for bearing systematic factor risk, compressing risk-adjusted returns. Investors who allocate across factor strategies and condition on the FUI can improve risk-adjusted performance not by selecting better factors but by timing their aggregate exposure to the factor space. The negative VIX coefficient in joint regressions further suggests that market-level volatility and factor-specific uncertainty operate through distinct and partially offsetting channels. When broad market stress is controlled for, higher aggregate volatility is associated with lower subsequent factor returns, while the FUI captures a risk compensation channel that operates independently of market-wide conditions.

Several directions for future research emerge naturally. The relationship between FUI and uncertainty-about-uncertainty represents a promising theoretical avenue, as the stochastic volatility results suggest that volatility-of-volatility in the factor space is economically meaningful. A formal asset pricing model that incorporates factor uncertainty as a priced state variable could rationalize the empirical patterns documented here and generate testable cross-sectional predictions about which assets should load most heavily on factor uncertainty risk. The concentration of predictability in value and fundamental factors suggests a link to long-run risk models and cash flow uncertainty, where factor uncertainty may proxy for time-varying discount rates rather than short-run sentiment. Applying the FUI framework to international equity markets would test whether factor uncertainty premia are global phenomena or specific to the U.S. factor theme structure. Incorporating high-frequency data could allow for a real-time FUI with short-horizon tactical allocation applications. Finally, the weak pre-1990 predictability and the dominance of the post-2000 period suggest that a regime-switching model of factor uncertainty pricing may uncover important nonlinearities in the relationship between uncertainty and expected factor returns.

Appendix

Table A.1: In-Sample R^2 by Factor Theme

This table reports in-sample R^2 values (in percent) from rolling AR(1) forecasting models estimated for each of the 13 factor themes (Jensen et al., 2023). Models are estimated on a rolling 120-month window with a minimum of 60 months. Sample: 878 monthly observations.

Factor	**In-sample R^2 (%)**
Accruals	2.21
Debt Issuance	1.22
Investment	1.98
Low Leverage	0.84
Low Risk	0.90
Momentum	0.24
Profit Growth	0.70
Profitability	2.41
Quality	1.47
Seasonality	0.37
Short-Term Reversal	1.32
Size	1.86
Value	2.10
Median	**1.32**

Notes: The Layer 1 forecasting specification is AR(1) throughout, selected by BIC across all 13 factor themes.

Table A.2: AR Lag Selection Robustness

This table documents the robustness of the AR(1) specification used in the FUI pipeline. Panel A reports BIC values for AR(p) models with p = 1,...,6 estimated for each of the 13 factors, with the BIC-optimal lag highlighted in bold. BIC is defined such that lower values indicate a preferred model. Panel B reports pairwise correlations between FUI series constructed under alternative AR specifications. Sample: 878 monthly observations.

Panel A: BIC Lag Selection — AR(p) Models, p = 1,...,6

Factor	p = 1	p = 2	p = 3	p = 4	p = 5	p = 6	Optimal p
Accruals	-5415.46	-5401.70	-5389.78	-5377.19	-5364.45	-5359.15	**1**
Debt Issuance	-6231.71	-6221.26	-6214.00	-6211.06	-6197.16	-6184.50	**1**
Investment	-4394.54	-4383.51	-4371.21	-4359.81	-4352.21	-4349.26	**1**
Low Leverage	-3823.26	-3812.19	-3803.79	-3803.06	-3794.35	-3793.49	**1**
Low Risk	-3387.47	-3375.83	-3365.61	-3356.79	-3346.68	-3340.45	**1**
Momentum	-3649.31	-3639.88	-3627.95	-3616.46	-3605.23	-3595.82	**1**
Profit Growth	-5519.39	-5507.66	-5495.62	-5486.63	-5476.62	-5463.37	**1**
Profitability	-4452.99	-4442.57	-4441.22	-4441.66	-4429.00	-4421.80	**1**
Quality	-4882.43	-4870.63	-4857.89	-4844.65	-4834.64	-4821.52	**1**
Seasonality	-6691.31	-6676.07	-6662.23	-6647.88	-6633.75	-6620.09	**1**
Short-Term Reversal	-5347.60	-5333.73	-5320.57	-5307.29	-5293.47	-5280.47	**1**
Size	-4237.36	-4226.20	-4223.14	-4212.15	-4202.84	-4190.66	**1**
Value	-3653.94	-3643.10	-3635.09	-3637.01	-3630.15	-3631.05	**1**

Notes: *BIC values are reported to two decimal places. Bold entries in the Optimal p column indicate the BIC-selected lag order. AR(1) is selected unanimously for all 13 factors (lag distribution: {1: 13}).*

Panel B: FUI Proxy Correlations Across AR Specifications

	AR(1)	AR(2)	AR(3)	BIC-optimal
AR(1)	1.0000	1.0000	0.9999	1.0000
AR(2)	1.0000	1.0000	0.9999	1.0000
AR(3)	0.9999	0.9999	1.0000	0.9999
BIC-optimal	1.0000	1.0000	0.9999	1.0000

Notes: *Pairwise Pearson correlations between FUI series constructed using AR(1), AR(2), AR(3), and BIC-optimal lag specifications. All correlations exceed 0.9999, confirming that the FUI is effectively invariant to the choice of lag order. The BIC-optimal series is identical to AR(1) since BIC selects p = 1 for all 13 factors.*

Table A.3: Unit Root and Stationarity Tests for FUI Variants

	ADF Test		KPSS Test		
Series	**Statistic**	**p-value**	**Statistic**	**p-value**	**Conclusion**
EWMA	-3.742	0.004	0.230	>0.10	Stationary
GARCH	-4.177	0.001	0.347	>0.10	Stationary
RV12	-3.870	0.002	0.211	>0.10	Stationary
SV	-4.043	0.001	0.243	>0.10	Stationary

Notes: *ADF denotes the Augmented Dickey-Fuller test with lag length selected by AIC; the null hypothesis is a unit root. KPSS denotes the Kwiatkowski Phillips-Schmidt-Shin test; the null hypothesis is stationarity. All tests include a constant. The KPSS p-value is reported as >0.10 when it exceeds the upper bound of the standard critical value table. Sample period: November 1966 - December 2024 (N = 698 monthly observations).*

Table A.4: Spearman Rank Correlations Between FUI Variants

	EWMA	**GARCH**	**RV12**	**SV**
EWMA	1.0000	0.8392	0.9576	0.8590
GARCH	0.8392	1.0000	0.7807	0.7015
RV12	0.9576	0.7807	1.0000	0.8220
SV	0.8590	0.7015	0.8220	1.0000

Notes: *Spearman rank correlations computed over the common sample period November 1966 - December 2024 (N = 698 monthly observations). All four FUI variants are constructed as the cross-sectional mean of factor-level conditional volatilities across 13 equity factor portfolios. Rank correlations are computed on the aligned (non-missing) sample.*

Table A.5: Stambaugh (1999) Bias-Corrected Predictive Regressions

Panel A: FUI → Average Factor Theme Return

	EWMA				**GARCH**				**RV12**				**SV**			
h	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***
h=1	0.122	-0.085	-1.58	4.05%	0.161	-0.204	**-2.76*****	4.13%	0.124	-0.067	-1.22	4.33%	0.275	-0.313	**-2.33****	3.64%
h=3	0.382	-0.078	-0.78	11.80%	0.504	0.365	**2.83*****	11.88%	0.385	-0.146	-1.58	12.42%	0.843	-2.117	**-8.21*****	10.10%
h=6	0.770	0.093	0.53	21.34%	0.976	0.910	**3.62*****	19.67%	0.778	0.270	1.64	22.50%	1.748	-3.310	**-7.26*****	19.28%
h=12	1.441	1.874	**5.03*****	28.25%	1.687	1.607	**3.07*****	21.90%	1.416	2.729	**8.01*****	28.35%	3.540	1.813	**2.02****	30.10%

Panel B: FUI → Cross-Sectional Factor Return Dispersion

	EWMA				**GARCH**				**RV12**				**SV**			
h	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***
h=1	0.850	1.783	**16.50*****	33.07%	0.969	1.840	**11.00*****	25.00%	0.826	2.197	**22.54*****	32.48%	2.529	11.919	**46.31*****	51.75%
h=3	1.262	0.349	**1.77***	23.14%	1.395	1.932	**6.94*****	16.46%	1.257	1.219	**6.19*****	23.82%	3.944	15.053	**24.83*****	39.95%
h=6	1.835	0.254	0.67	25.05%	1.960	1.711	**2.83*****	16.66%	1.840	2.409	**6.12*****	26.14%	5.348	10.342	**12.25*****	37.60%
h=12	3.018	5.909	**8.99*****	32.74%	3.135	3.390	**3.72*****	20.55%	2.881	9.676	**14.71*****	30.89%	8.925	30.860	**24.21*****	50.60%

Panel C: FUI → Future Realized Volatility of Average Factor Return

	EWMA				**GARCH**				**RV12**				**SV**			
h	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***	***β_raw***	***β_corr***	***t***	***R²%***
h=3	0.259	0.586	**13.27*****	39.54%	0.295	0.427	**6.23*****	29.64%	0.259	0.738	**17.44*****	41.17%	0.683	2.054	**22.83*****	48.72%
h=6	0.263	0.369	**11.37*****	49.74%	0.299	0.347	**6.12*****	36.96%	0.259	0.587	**18.15*****	50.10%	0.716	1.687	**31.32*****	65.32%
h=12	0.241	0.469	**12.62*****	48.62%	0.264	0.369	**6.16*****	33.32%	0.233	0.667	**17.84*****	47.12%	0.690	2.224	**38.45*****	70.45%

Notes: β_raw = OLS estimate. β_corr = Stambaugh (1999) bias-corrected estimate. t = Newey–West HAC t-statistic on β_corr with maxlags = h. Bold t-statistics are significant at the 10% level or better. *, **, *** denote significance at the 10%, 5%, 1% level. Panel C omits h = 1 (rolling standard deviation requires ≥2 observations).

Table A.6: Individual Factor Predictive Regressions with Benjamini–Hochberg FDR Correction

	h = 1					h = 3					h = 6					h = 12				
Factor	**β**	**t**	**p_raw**	**q (BH)**	**BH**	**β**	**t**	**p_raw**	**q (BH)**	**BH**	**β**	**t**	**p_raw**	**q (BH)**	**BH**	**β**	**t**	**p_raw**	**q (BH)**	**BH**
Accruals	-0.141	-3.19***	0.0014	0.0182	✓	-0.441	-4.22***	<.0001	0.0003	✓	-0.808	-3.73***	0.0002	0.0008	✓	-1.466	-3.14***	0.0017	0.0028	✓
Debt Issuance	0.090	2.54**	0.0113	0.0732	✓	0.258	3.64***	0.0003	0.0018	✓	0.508	3.49***	0.0005	0.0015	✓	0.980	4.85***	<.0001	<.0001	✓
Investment	0.430	2.09**	0.0366	0.1190	✗	1.326	3.37***	0.0007	0.0028	✓	2.719	4.45***	<.0001	0.0001	✓	4.836	4.02***	0.0001	0.0003	✓
Low Leverage	-0.446	-1.43	0.1538	0.1817	✗	-1.455	-2.40**	0.0164	0.0305	✓	-3.035	-3.12***	0.0018	0.0039	✓	-5.425	-3.62***	0.0003	0.0008	✓
Low Risk	0.445	1.33	0.1849	0.2003	✗	1.437	2.35**	0.0189	0.0307	✓	2.927	2.87***	0.0041	0.0066	✓	5.214	2.74***	0.0061	0.0088	✓
Momentum	-0.017	-0.07	0.9434	0.9434	✗	0.079	0.22	0.8262	0.8262	✗	0.245	0.51	0.6085	0.6085	✗	0.724	0.97	0.3335	0.3335	✗
Profit Growth	-0.139	-1.60	0.1099	0.1761	✗	-0.413	-1.71*	0.0874	0.0947	✓	-0.835	-1.83*	0.0674	0.0731	✓	-1.174	-1.40	0.1610	0.1744	✗
Profitability	0.294	1.52	0.1285	0.1761	✗	0.917	2.45**	0.0144	0.0305	✓	1.752	2.43**	0.0150	0.0195	✓	2.895	2.26**	0.0241	0.0313	✓
Quality	0.099	1.49	0.1355	0.1761	✗	0.294	2.07**	0.0388	0.0502	✓	0.555	2.08**	0.0380	0.0449	✓	1.030	1.63	0.1032	0.1219	✗
Seasonality	0.069	1.72*	0.0863	0.1603	✗	0.224	2.03**	0.0425	0.0502	✓	0.515	3.01***	0.0026	0.0049	✓	1.044	3.48***	0.0005	0.0011	✓
Short-Term Reversal	0.124	1.78*	0.0754	0.1603	✗	0.396	3.07***	0.0021	0.0055	✓	0.674	3.44***	0.0006	0.0015	✓	1.195	3.13***	0.0017	0.0028	✓
Size	0.232	1.89*	0.0592	0.1539	✗	0.644	2.15**	0.0315	0.0455	✓	1.316	2.73***	0.0062	0.0090	✓	2.695	3.74***	0.0002	0.0006	✓
Value	0.680	2.14**	0.0324	0.1190	✗	2.095	3.34***	0.0008	0.0028	✓	4.249	4.13***	<.0001	0.0002	✓	7.503	4.39***	<.0001	0.0001	✓

Notes: Predictive regressions of FUI (EWMA, beginning-of-month) on h-month cumulative individual factor theme returns. Standard errors are Newey–West HAC with maxlags = h. p_raw = raw p-value. q (BH) = Benjamini–Hochberg (1995) adjusted q-value at a 10% FDR threshold. ✓ = factor survives BH adjustment; ✗ = factor dropped. Positive β shown in blue, negative in red. *, **, *** denote 10%, 5%, 1% significance on raw t-statistic.

Table A.7: Post-Dot-Com Subperiod Predictive Regressions (EWMA)

Subperiod	N	h = 1	h = 3	h = 6	h = 12
		β (t-stat)	β (t-stat)	β (t-stat)	β (t-stat)
Post-dotcom (full)	264	0.001 (0.02) R^2 = 0.00%	-0.004 (-0.02) R^2 = 0.00%	0.014 (0.04) R^2 = 0.00%	0.225 (0.30) R^2 = 0.40%
Pre-GFC (2003–2006)†	48	-0.039 (-0.53) R^2 = 1.36%	-0.155 (-0.79) R^2 = 5.95%	-0.018 (-0.09) R^2 = 0.08%	0.137 (0.79) R^2 = 3.24%
GFC (2007–2009)†	36	-0.157 (-1.10) R^2 = 2.22%	-0.600 (-1.98)* R^2 = 7.92%	-1.361 (-2.35)** R^2 = 15.56%	—
Post-GFC (2010–2019)	120	-0.063 (-0.64) R^2 = 0.24%	-0.218 (-1.06) R^2 = 1.06%	-0.552 (-1.93)* R^2 = 4.34%	-1.408 (-2.27)** R^2 = 9.75%
COVID+ (2020–)	60	0.481 (2.16)** R^2 = 5.35%	1.930 (3.26)*** R^2 = 20.86%	3.412 (4.72)*** R^2 = 27.52%	7.539 (8.25)*** R^2 = 55.62%†

*Notes: Predictive regressions of the form $y_{t+}h = \alpha + \beta \cdot FUI_{t-1} + \varepsilon$ using the EWMA-based FUI (beginning-of-month timing). $y_{t+}h$ is the h-month cumulative equal-weighted average factor theme return. Standard errors are Newey–West HAC with maxlags = h. t-statistics reported in parentheses. R^2 is the OLS coefficient of determination. Subperiods: Pre-GFC = January 2003–December 2006; GFC = January 2007–December 2009; Post-GFC = January 2010–December 2019; COVID+ = January 2020 onwards. Full-sample post-2000 results are reported in Table 8. † N < 50; results are illustrative only and should be interpreted with caution given the short sample. The COVID+ h = 12 result is also flagged given N = 48 at that horizon. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (based on Newey–West t-statistics). — = insufficient observations for regression at this horizon.*

References

Ang, A., & Chen, J. (2007). CAPM over the long run: 1926-2001. *Journal of Empirical Finance*, 14(1), 1–40.

Asness, C., Ilmanen, A., & Maloney, T. (2017). Market Timing: Sin a Little Resolving the Valuation Timing Puzzle. *The Journal of Investment Management*, 15(3), 23–40.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Bali, T. G., Brown, S. J., & Tang, Y. (2017). Is economic uncertainty priced in the cross-section of stock returns? *Journal of Financial Economics*, 126(3), 471–489.

Baltas, N. (2019). The Impact of Crowding in Alternative Risk Premia Investing. *Financial Analysts Journal*, 75(3), 89–104.

Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, 116(1), 111–120.

Bekaert, G., Hoerova, M., & Lo Duca, M. (2013). Risk, uncertainty, and monetary policy. *Journal of Monetary Economics,* 60(7), 771–788.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society.* Series B (Methodological), 57(1), 289–300.

Blitz, D. (2021). T*he Quant Cycle*. SSRN Working Paper No. 3930006.

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica,* 77(3), 623–685.

Brogaard, J., & Detzel, A. (2015). The Asset-Pricing Implications of Government Economic Policy Uncertainty. *Management Science*, 61(1), 3–18.

Campbell, J. Y, & Yogo, M. (2006). Efficient tests of stock return predictability. *Journal of Financial Economics*, 81(1), 27-60.

Campbell, J. Y, & Thompson, S. B. (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *The Review of Financial Studies*, 21(4), 1509–1531.

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1), 291–333.

Conrad, C., & Loch, K. (2015). Anticipating Long-Term Stock Market Volatility. *Journal of Applied Econometrics*, 30(7), 1090–1114.

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes, *Journal of Financial Economics*, 122(2), 221–247.

Daniel, K., Hirshleifer, D., & Sun, L. (2020). Short- and Long-Horizon Behavioral Factors. *The Review of Financial Studies*, 33(4), 1673–1736.

de Jong, P., & Shephard, N. (1995). The simulation smoother for time series models. *Biometrika*, 82(2), 339–350.

Ehsani, S., & Linnainmaa, J. T. (2022). Factor momentum and the momentum factor. *The Journal of Finance*, 77(3), 1877–1919.

Ferson, W. E., & Harvey, C. R. (1991). The Variation of Economic Risk Premiums. *The Journal of Political Economy*, 99(2), 385–415.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian Statistics 4* (pp. 169–193). Oxford University Press.

Gupta, T., & Kelly, B. (2019). Factor momentum everywhere. *Journal of Portfolio Management*, 45(3), 13–36.

Haddad, V., Kozak, S., & Santosh, S. (2020). Factor timing. *Review of Financial Studies*, 33(5), 1980–2018.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). …and the Cross-Section of Expected Returns. *The Review of Financial Studies*, 29(1), 5–68.

Ilmanen, A., Israel, R., Lee, R., Moskowitz, T. J., Thapar, A. (2021). How Do Factor Premia Vary Over Time? A Century of Evidence. *Journal of Investment Management*, 19(4), 15–57.

Israel, R., & Moskowitz, T. J. (2013). The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics*, 108(2), 275–301.

Jensen, T. I., Kelly, B., & Pedersen, L. H. (2023) Is There a Replication Crisis in Finance? *The Journal of Finance*, 78(5), 2465–2518.

Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring Uncertainty. *American Economic Review*, 105(3), 1177–1216.

Kelly, B., Pástor, L., & Veronesi, P. (2016). The Price of Political Uncertainty: Theory and Evidence from the Option Market. *The Journal of Finance*, 71(5), 2417–2480.

Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. The *Review of Economic Studies,* 65(3), 361–393.

Lettau, M., & Ludvigson, S. (2001). Consumption, Aggregate Wealth, and Expected Stock Returns. *The Journal of Finance*, 56(3), 815–849.

Ludvigson, S. C., Ma, S., & Ng, S. (2021). Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response? *American Economic Journal: Macroeconomics*, 13(4), 369–410.

Moreira, A., & Muir, T. (2017). Volatility-Managed Portfolios. *The Journal of Finance*, 72(4), 1611-1643.

Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, 31, 274–295.

Pástor, L., & Veronesi, P. (2013). Political uncertainty and risk premia. *Journal of Financial Economics*, 110(3), 520–545.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246–256.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, 54(3), 375–421.

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), 236–244.