IFA

Iranian Journal of Finance

Print ISSN 2676-6337
Online ISSN 2676-6345

Causal Alpha: Identifying Structural Drivers of Excess Returns in the US Equity Market Using Causal Machine Learning

Moslem Nilchi

Assistant Prof, Department of Management, Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran. (Email: m.nilchi@alzahra.ac.ir)

Mahdi Zerang*

*Corresponding author, Ph.D. Candidate, Department of Economics, Faculty of Economics, Allameh Tabataba'i University, Tehran, Iran. (Email: zerangomur@yahoo.com)

Iranian Journal of Finance, 2026, Vol. 10, No.2, pp. 158-185.
Publisher: Iran Finance Association
doi: https://doi.org/10.66224/ijf.2026.593966.1593
Article Type: Original Article

Received: January 08, 2026
Received in revised form: February 27, 2026
Accepted: March 19, 2026
Published online: April 01, 2026

Abstract

This study investigates the causal relationship between uncertainty shocks and factor premia in the US equity market. While traditional asset pricing models rely on correlation-based factor structures, this research argues that identifying true alpha necessitates a causal framework. Employing a novel combination of Double Machine Learning (DML) and Causal Forest algorithms on Kenneth French's five-factor data and VIX-based uncertainty shocks from 2001 to 2025, robust evidence is provided demonstrating that uncertainty shocks exert significant causal effects on factor returns. The findings reveal that the market factor (MKT-RF) and the size factor (SMB) exhibit strong negative causal responses to uncertainty shocks. In contrast, the profitability factor (RMW) demonstrates a positive and significant effect. The value (HML) and investment (CMA) factors, however, show no statistically detectable causal effect. Substantial heterogeneity in these effects is documented across high- and low-volatility regimes, with the market's negative response being substantially larger in high-uncertainty periods. From an economic significance perspective, a long-only trading strategy based on causal signals delivers a competitive Sharpe ratio of 0.437, with a substantially lower maximum drawdown (-32.5%) compared to the market portfolio (-51.4%). The results survive rigorous robustness checks and statistical tests, providing evidence consistent with a causal effect of uncertainty shocks on factor premia. This study contributes to the growing literature on causal machine learning in finance and offers practical implications for risk management and portfolio construction.

Keywords: Causal Alpha, Uncertainty Shocks, Double Machine Learning, Causal Forest, Fama-French Factors

JEL Classification: G12, G17, C14, C53

Introduction

Problem Statement and Research Motivation

The quest to explain the cross-section of expected stock returns has constituted the central pursuit of empirical asset pricing for over half a century. From the Capital Asset Pricing Model (CAPM) developed by Sharpe (1964) and Lintner (1965) to the revolutionary Fama-French three-factor model (Fama & French, 1993) and its subsequent extensions (Carhart, 1997; Fama & French, 2015), the literature has evolved through the systematic accumulation of empirical regularities—anomalies that challenge prevailing paradigms. The Fama-French five-factor model (Fama & French, 2015), which adds profitability (RMW) and investment (CMA) factors to the existing market, size, and value factors, currently stands as a workhorse model for describing average returns (Swedroe & Berkin, 2017). However, as Fama and French (2015) themselves acknowledge, no asset pricing model is perfect; they provide useful descriptions rather than complete explanations.

A fundamental limitation shared by these models is their reliance on correlation rather than causation. They document that certain firm characteristics are correlated with subsequent returns. However, they do not establish whether these characteristics are structural drivers of alpha or merely correlate with latent confounding factors. As Kumar, Sharma, and Rao (2024) emphasize in their comprehensive review, causal inference in finance addresses questions where changes in one variable inevitably trigger changes in another, moving beyond spurious relationships that may result from hidden shared causes. This distinction is not merely academic—it has profound implications for portfolio construction, risk management, and the theoretical understanding of financial markets.

Research Hypotheses

Based on the theoretical foundations and prior literature, the following hypotheses are formulated:

Hypothesis 1 (H1): Uncertainty shocks exert a statistically significant negative causal effect on the market factor (MKT-RF) returns.

Hypothesis 2 (H2): Uncertainty shocks exert a statistically significant negative causal effect on the size factor (SMB) returns.

Hypothesis 3 (H3): Uncertainty shocks exert a statistically significant positive causal effect on the profitability factor (RMW) returns.

Hypothesis 4 (H4): Uncertainty shocks do not exert statistically significant causal effects on the value (HML) and investment (CMA) factor returns.

Hypothesis 5 (H5): The causal effects of uncertainty shocks on factor returns differ significantly across high- and low-volatility regimes.

Research Objectives and Questions

The fundamental question this research seeks to answer is: Are uncertainty shocks structural drivers of factor returns, or are they merely correlates arising from latent variables? To address this question, the present study employs advanced causal machine learning methods to examine the causal effects of uncertainty shocks on the Fama-French five factors.

Paper Structure

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature, tracing the evolution from CAPM to five-factor models and the emergence of causal inference in finance. Section 3 develops the theoretical framework, including formal definitions of causal effects, identification assumptions, and the econometric methodology. Section 4 describes the data construction and presents summary statistics. Section 5 presents the main empirical results from DML and Causal Forest analyses. Section 6 explores heterogeneity across market regimes. Section 7 evaluates the economic significance through portfolio analysis. Section 8 presents robustness checks. Section 9 concludes by discussing implications for researchers, practitioners, and policymakers, as well as limitations and directions for future research.

Literature Review

The Evolution of Asset Pricing Models

The intellectual history of asset pricing begins with the Capital Asset Pricing Model, developed independently by Sharpe (1964), Lintner (1965), and Mossin (1966). The CAPM posits that the expected return of a security is linearly related to its systematic risk, measured by beta relative to the market portfolio. Formally:

$$E[R_i] = R_f + \beta_i(E[R_m] - R_f)$$

where R_i is the return on asset i, R_f is the risk-free rate, R_m is the market return, and $\beta_i = \mathrm{Cov}(R_i, R_m)/\mathrm{Var}(R_m)$. Despite its theoretical elegance, the CAPM faced empirical challenges from anomalies that beta could not explain (Fama & French, 1992).

The seminal work of Fama and French (1992, 1993) revolutionized the field by introducing the three-factor model:

$$E[R_i] - R_f = \beta_i(E[R_m] - R_f) + s_i E[SMB] + h_i E[HML]$$

where SMB (Small Minus Big) captures the size premium, and HML (High Minus Low) captures the value premium. This model absorbed many of the CAPM anomalies and became the new standard for performance evaluation.

Subsequent research identified additional patterns. Novy-Marx (2013) demonstrated that profitability, measured by gross profit-to-assets, had explanatory power comparable to that of the value premium. Titman, Wei, and Xie (2004) and Fama and French (2006) documented that firms with higher investment tend to have lower average returns. Synthesizing these insights, Fama and French (2015) proposed the five-factor model:

$$E[R_i] - R_f = \beta_i(E[R_m] - R_f) + s_i E[SMB] + h_i E[HML] + r_i E[RMW] + c_i E[CMA]$$

where RMW (Robust Minus Weak) captures the profitability premium, and CMA (Conservative Minus Aggressive) captures the investment premium.

As Swedroe and Berkin (2017) note, the five-factor model represents the evolution of our understanding: from one-factor CAPM to three-factor to four-factor (adding momentum, Carhart, 1997) to five-factor. However, even this model is incomplete; Fama and French (2015) acknowledge that it fails to explain the low returns on small stocks with low profitability and high investment. Moreover, all these models are descriptive rather than causal—they document correlations without establishing structural relationships. This theoretical gap constitutes the primary motivation for employing causal methods in the present research.

Causal Inference in Finance

The limitations of correlation-based approaches have spurred growing interest in causal inference within the financial domain. As Kumar, Sharma, and Rao (2024) document in their extensive survey of 45 papers from 1992 to 2023, causal methods have been applied across banking, risk management, asset pricing, and insurance. They define causal inference as "the identification of the cause or causes of a phenomenon, by establishing covariation of cause and effect, a time-order relationship with the cause preceding the effect, and the elimination of plausible alternative causes" (after Shadish, Cook, & Campbell, 2002).

The foundational work of Pearl (1986) on causal trees and graphical models provided the theoretical basis for modern causal analysis. Pearl (2009) formalized the structural causal model (SCM) framework, introducing the do-calculus and establishing conditions for identifying causal effects from observational data. In economics, Imbens and Rubin (2015) provided comprehensive guidance on program evaluation and treatment effects.

Recent years have witnessed the convergence of causal inference and machine learning. Chernozhukov et al. (2018) developed Double Machine Learning, which uses Neyman-orthogonal scores and cross-fitting to estimate treatment effects in high-dimensional settings. The method achieves $\sqrt{N}$-consistency and asymptotic normality even when flexible ML estimators are used for the nuisance functions. Wager and Athey (2018) and Athey, Tibshirani, and Wager (2019) introduced causal forests, an extension of random forests designed to estimate conditional average treatment effects (CATE). The causal forest algorithm grows trees to maximize heterogeneity in treatment effects rather than prediction accuracy, producing valid confidence intervals through honesty and subsampling.

Applications of these methods in finance remain nascent but growing. Kumar, Sharma, and Rao (2025) highlight several promising directions: evaluating financial regulations, analyzing systemic risk transmission, and understanding investor behavior. This paper contributes to this emerging literature by applying DML and causal forests to the fundamental question of factor-based asset pricing.

Uncertainty Shocks and Financial Markets

The relationship between uncertainty and asset prices has been extensively studied. Bloom (2009) showed that uncertainty shocks, measured by VIX spikes, cause sharp drops in economic activity and stock prices. Bachmann, Elstner, and Sims (2013) documented that uncertainty varies countercyclically and affects firm-level investment decisions. Baker, Bloom, and Davis (2016) constructed an economic policy uncertainty index and demonstrated its predictive power for stock market volatility.

However, most of this literature focuses on correlation rather than causation. As Kumar, Sharma, and Rao (2024) note, establishing causality in financial time series requires addressing endogeneity, omitted variable bias, and reverse causality. The VIX itself is not exogenous—it responds to market movements, creating a simultaneity problem. Our approach of using AR(1) residuals as exogenous shocks, combined with DML's orthogonalization, addresses these concerns more rigorously than traditional vector autoregressions.

Research Gap and Novelty

Based on the literature review, the following research gaps are identified:

a) Lack of Causal Approaches in Asset Pricing: Despite the proliferation of factor models, almost all are correlation-based, rarely employing causal frameworks to identify structural drivers of returns. This research addresses this gap by employing advanced causal machine learning methods.

b) Absence of Simultaneous Examination of All Five Factors: Prior studies examining uncertainty effects on stock returns have predominantly focused on the market factor, neglecting other factors (SMB, HML, RMW, CMA). This research is the first to examine the causal effects of uncertainty shocks on the complete set of Fama-French five factors.

c) Overlooking Heterogeneity Across Regimes: Many studies assume uniform effects of uncertainty shocks, failing to examine differences across high- and low-volatility regimes. The present research reveals substantial heterogeneity through regime analysis based on VIX levels.

d) Limited Translation of Causal Findings into Trading Strategies: Despite extensive research on uncertainty and returns, few studies have translated causal findings into actionable trading strategies. This research bridges theory and practice by designing and evaluating trading strategies based on causal signals.

Novelty of this Research

No.	Novelty	Description
1	Advanced Causal Methodology	Simultaneous use of DML and Causal Forest for treatment effect estimation in the presence of high-dimensional confounders
2	Comprehensive Factor Coverage	First simultaneous examination of causal effects on all five Fama-French factors
3	Regime Heterogeneity Analysis	Disaggregation of effects across high- and low-volatility regimes, identifying distinct behavioral patterns
4	Practical Implementation	Translation of causal findings into tradable signals and systematic portfolio performance evaluation

Research Methodology

Definitions of Causal Effects

This research adopts the potential outcomes framework (Rubin, 1974; Imbens & Rubin, 2015). Let W_i be a binary treatment indicator (although in our application, the treatment $W_i = \text{VIX\_shock}_i$ is continuous). For each unit i, there exist potential outcomes $Y_i(1)$ and $Y_i(0)$ corresponding to treatment and control states. The individual treatment effect is $\tau_i = Yi\,(1) - Y_i(0)$, which is unobservable because we observe only one potential outcome.

The fundamental parameters of interest are:

Definition (Average Treatment Effect): The Average Treatment Effect (ATE) is the expected difference in potential outcomes:

$$\tau_{\text{ATE}} = E[Y(1) - Y(0)] \tag{1}$$

Definition (Conditional Average Treatment Effect): The Conditional Average Treatment Effect (CATE) is the treatment effect conditional on covariates $X = x$:

$$\tau(x) = E[Y(1) - Y(0) \mid X = x] \tag{2}$$

In our setting with continuous treatment, following Athey, Tibshirani, and Wager (2019), we interpret the estimand as an average partial effect:

$$\tau(x) = \frac{\text{Cov}(Y, W \mid X = x)}{\text{Var}(W \mid X = x)} \tag{3}$$

under the assumption of unconfoundedness and overlap.

Identification Assumptions

Causal identification requires three key assumptions:

Assumption (Unconfoundedness): Treatment assignment is independent of potential outcomes conditional on covariates:

$$\{Y(1), Y(0)\} \perp W \mid X \tag{4}$$

Assumption (Overlap): For all x, the treatment propensity is bounded away from 0 and 1:

$$0 < P(W = 1 \mid X = x) < 1 \tag{5}$$

or for continuous treatments, the density of W given X is bounded away from zero.

Assumption (Consistency): The observed outcome equals the potential outcome corresponding to the actual treatment:

$$Y = Y(W) \quad (6)$$

Double Machine Learning (DML)

We implement the partially linear DML model (Chernozhukov et al., 2018):

$$Y = D\theta_0 + g_0(X) + U, E[U \mid X, D] = 0 \quad (7)$$

$$D = m_0(X) + V, E[V \mid X] = 0 \quad (8)$$

where Y is the outcome (factor return), D is the treatment (VIX shock), X is the vector of confounders, g_0 and m_0 are nuisance functions, and θ_0 is the causal parameter of interest.

Advantage of DML over Traditional Methods: Traditional methods, such as Vector Autoregression (VAR) or multiple linear regression, rely on strong parametric assumptions (e.g., linearity) to control for confounders and face challenges with high-dimensional covariates. In contrast, DML employs flexible machine-learning estimators for the nuisance functions $g_0(X)$ and $m_0(X)$ without imposing restrictive parametric assumptions, thereby enabling simultaneous control of numerous confounders. Furthermore, the use of Neyman-orthogonal scores and cross-fitting in DML eliminates regularization bias, rendering the estimator $\sqrt{N}$-consistent and asymptotically normal. These characteristics make DML substantially more suitable than traditional methods for financial data with high dimensionality and nonlinear relationships.

The DML algorithm proceeds as follows:

1. Split the sample into K folds.
2. For each fold k, estimate $\hat{g}^{-k}(X)$ and $\hat{m}^{-k}(X)$ using the other $K - 1$ folds.

3. Construct residuals: $\widehat{U} = Y - \hat{g}^{-k}(X), \hat{V} = D - \widehat{m}^{-k}(X)$.
4. Estimate θ by regressing $\widehat{U}$ on $\hat{V}$.
5. Aggregate estimates across folds.

The resulting estimator $\hat{\theta}$ is $\sqrt{N}$-consistent and asymptotically normal under mild conditions, with variance:

$$\sqrt{N}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} N(0, \sigma^2) \tag{9}$$

where $\sigma^2 = E[V^2]^{-1}E[V^2U^2]E[V^2]^{-1}$.

Causal Forest

The causal forest (Wager & Athey, 2018; Athey, Tibshirani, & Wager, 2019) extends this idea to estimate heterogeneous treatment effects. This method grows trees to maximize heterogeneity in treatment effects rather than prediction accuracy, enabling the estimation of conditional average treatment effects (CATE). For each test point x, the forest computes:

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{N} \alpha_i(x)(Y_i - \widehat{m}^{(-i)}(X_i))(D_i - \hat{e}^{(-i)}(X_i))}{\sum_{i=1}^{N} \alpha_i(x)(D_i - \hat{e}^{-i}(X_i))^2} \tag{10}$$

where:

- $\alpha_i(x)$ are adaptive nearest neighbor weights derived from the forest
- $\widehat{m}^{(-i)}(X_i)$ is out-of-bag estimate of $E[Y \mid X_i]$ using regression forest
- $\hat{e}^{(-i)}(X_i)$ is out-of-bag estimate of $E[D \mid X_i]$ (propensity score)

The algorithm implements honesty by using separate subsamples for splitting and estimation, which ensures:

$$E[(\hat{\tau}(x) - \tau(x))^2] = O\left(N^{-\frac{2}{d+2}}\right) \tag{11}$$

and provides asymptotically valid confidence intervals.

Advantage of Causal Forest over Traditional Heterogeneity

Methods: Traditional methods for examining treatment effect heterogeneity typically employ parametric interactions (e.g., multiplying treatment by covariates) or pre-specified subgroup partitions (e.g., separate regressions for subgroups), both of which are susceptible to bias from incorrect model specification. Causal forest, by employing a data-driven algorithm to automatically identify patterns of heterogeneity in treatment effects without requiring prior parametric assumptions, provides more accurate estimates.

Ucertainty Shock Construction

Following standard practice in the literature, we construct uncertainty shocks as the residuals from an AR(1) model of the VIX:

$$\mathrm{VIX}_t = \phi_0 + \phi_1 \mathrm{VIX}_{t-1} + \varepsilon_t \tag{12}$$

$$\mathrm{Shock}_t = \hat{\varepsilon}_t = \mathrm{VIX}_t - \hat{\phi}_0 - \hat{\phi}_1 \mathrm{VIX}_{t-1} \tag{13}$$

This decomposition isolates unexpected changes in volatility, which are more plausibly exogenous than VIX levels. The AR(1) specification captures the strong persistence in volatility while ensuring shocks are serially uncorrelated. The residual from this model represents the component of VIX that cannot be predicted from its own past, making it interpretable as an innovation or "shock" to uncertainty that is orthogonal to past information.

Data

Data Sources and Construction

We utilize two primary data sources. First, we obtain the Fama-French five factors from Kenneth French's data library (Fama & French, 2015). The factors include:

- MKT-RF: Market excess return (value-weighted return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ minus the one-month Treasury bill rate)

- SMB: Small Minus Big, the size factor
- HML: High Minus Low, the value factor (book-to-market equity)
- RMW: Robust Minus Weak, the profitability factor (operating profitability)
- CMA: Conservative Minus Aggressive, the investment factor (change in total assets)

Second, we download daily VIX data from Yahoo Finance (ticker ^VIX) and convert to monthly averages. The VIX, computed by the CBOE (CBOE, 2003), measures the market's expectation of 30-day forward-looking volatility implied by S&P 500 index options.

Our sample period runs from January 2001 to December 2025, yielding 300 monthly observations. The starting point ensures we capture multiple market regimes, including the 2008 financial crisis, the European debt crisis, the COVID-19 pandemic, and the post-pandemic recovery. Control variables include first through third lags of each factor, first through third lags of VIX shocks, 12-month rolling VIX volatility, and a time trend.

Descriptive Statistics

Table 1 presents summary statistics. The mean excess market return is 0.69% per month (approximately 8.6% annualized), consistent with historical equity premia. All five factors have positive means, with profitability (RMW) showing the largest average premium, at 0.35% per month. The VIX averages 19.7, with substantial variation from 10.1 to 62.6, capturing the full range of market conditions from complacency to panic. The VIX shocks have a mean near zero (-0.0388) by construction, with a standard deviation of 4.38, indicating that unexpected volatility changes of ±4.4 points are typical.

Table 1. Descriptive Statistics (January 2001 – December 2025)

Variable	Mean	Std. Dev.	Min	Max	N
MKT-RF	0.0069	0.0449	-0.1720	0.1358	300
SMB	0.0014	0.0275	-0.0818	0.0834	300
HML	0.0007	0.0319	-0.1383	0.1286	300
RMW	0.0035	0.0230	-0.0925	0.0912	300
CMA	0.0011	0.0210	-0.0708	0.0901	300
VIX	19.6977	8.0665	10.1255	62.6395	300
VIX-Shock	-0.0388	4.3841	-10.1595	38.0887	300

Note: MKT-RF, SMB, HML, RMW, and CMA are expressed in decimal form (e.g., 0.0069 = 0.69% monthly). VIX is the monthly average. VIX-Shock is the residual from the AR(1) model. The sample period is from January 2001 to December 2025 (300 monthly observations).

Correlation Structure

Table 2 reveals several important patterns. First, VIX shocks are strongly negatively correlated with market returns (-0.632), confirming the well-known "fear-off" dynamic. Second, the correlation with SMB is negative (-0.273), suggesting that small caps are particularly sensitive to increases in volatility. Third, RMW exhibits a positive correlation (0.151), indicating that profitable firms may provide some hedge. Fourth, the correlation between HML and CMA (0.600) is notable, consistent with Fama and French (2015)'s finding that HML may be redundant in the presence of investment and profitability factors.

However, as we emphasize throughout, correlation does not imply causation. The negative correlation between VIX shocks and market returns could reflect reverse causality (market declines causing VIX spikes), omitted variables (macroeconomic news affecting both), or genuine causal effects. Our causal machine learning approach is designed to disentangle these possibilities.

Table 2. Correlation Matrix

	MKT-RF	SMB	HML	RMW	CMA	VIX-Shock
MKT-RF	1.000	0.304	0.032	-0.362	-0.169	-0.632
SMB	0.304	1.000	0.283	-0.303	0.079	-0.273
HML	0.032	0.283	1.000	0.159	0.600	-0.123
RMW	-0.362	-0.303	0.159	1.000	0.186	0.151
CMA	-0.169	0.079	0.600	0.186	1.000	0.100
VIX-Shock	-0.632	-0.273	-0.123	0.151	0.100	1.000

Note: Correlation coefficients in bold are statistically significant at the 5% level. Unmarked coefficients are not significant at conventional levels.

Results

Double Machine Learning Estimates

We begin by estimating average treatment effects using Double Machine Learning with 5-fold cross-fitting. The treatment variable (VIX-Shock) was standardized to have zero mean and unit variance prior to estimation, so the reported ATE coefficients represent the effect of a one-standard-deviation increase in the VIX shock. The control variables include lagged factors ($\mathrm{MKT-RF}_{t-1}$, SMB_{t-1}, HML_{t-1}, RMW_{t-1}, CMA_{t-1}), lagged VIX shocks ($\mathrm{VIX-shock}_{t-1}$, $\mathrm{VIX-shock}_{t-2}$, $\mathrm{VIX-shock}_{t-3}$), 12-month rolling VIX volatility, and a time trend. We use random forests with 500 trees for both nuisance functions.

Table 3. Causal Effects of Uncertainty Shocks on Factor Returns (DML Estimates)

Factor	ATE	Std. Error	t-stat	p-value	95% CI Lower	95% CI Upper	Significance
MKT-RF	-0.007476	0.000598	-12.492	0.0000	-0.008648	-0.006304	***
SMB	-0.001898	0.000445	-4.266	0.0000	-0.002770	-0.001026	***
HML	-0.000433	0.000481	-0.899	0.3684	-0.001376	0.000510	
RMW	0.001038	0.000327	3.174	0.0015	0.000397	0.001679	***
CMA	0.000338	0.000279	1.211	0.2259	-0.000209	0.000885	

Note: The treatment variable (VIX-Shock) was standardized to have zero mean and unit variance prior to estimation. Therefore, the reported ATE coefficients represent the effect of a one-standard-deviation increase in the VIX shock. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Standard errors are robust to heteroskedasticity.

Table 3 presents our main causal estimates, which are discussed below in relation to the research hypotheses:

First, the market factor (MKT-RF) exhibits a strong negative causal response to uncertainty shocks. A one-standard-deviation increase in the standardized VIX shock (approximately 4.4 points in the original VIX scale) causes a 0.75% decline in monthly excess returns (ATE = -0.0075, $p < 0.01$). This effect is economically large—comparable to one-sixth of the factor's standard deviation—and precisely estimated. The 95% confidence interval excludes zero by a wide margin. This finding is consistent with Bloom (2009), who demonstrated that uncertainty shocks lead to declines in economic activity and stock prices, and confirms Hypothesis 1 (H1).

Second, the size factor (SMB) also responds negatively, with an ATE of -0.0019 ($p < 0.01$) for a one-standard-deviation increase in the VIX shock. Small-cap stocks are thus causally more vulnerable to uncertainty spikes than large caps, consistent with their higher betas and greater sensitivity to economic conditions. This finding validates the common practice of reducing small-cap exposure during volatile periods and confirms Hypothesis 2 (H2).

Third, the profitability factor (RMW) shows a positive causal effect (ATE = 0.0010, $p < 0.01$) for a one-standard-deviation increase in the VIX shock. Profitable firms benefit from uncertainty shocks because their cash flows are more resilient or because investors flock to quality during turbulent times. This result aligns with the "flight to quality" phenomenon documented in the behavioral finance literature and confirms Hypothesis 3 (H3).

Fourth, for the value factor (HML) and investment factor (CMA), we find no statistically detectable causal effect of uncertainty shocks. Their ATEs are statistically indistinguishable from zero ($p = 0.368$ and $p = 0.226$, respectively). This suggests that while HML and CMA are correlated with returns (as shown in Table 2), they are not driven by uncertainty shocks in this setting. Their premiums may arise from other structural sources, such as distress risk or investor sentiment. Thus, the evidence is consistent with Hypothesis 4 (H4), though causal neutrality cannot be definitively proven.

These findings have important implications for asset pricing theory. They suggest that uncertainty shocks are consistent with being structural drivers of market and size premia, but not of value or investment premia in this empirical setting. A complete causal model of factor returns would need to incorporate multiple distinct drivers, rather than a single latent factor.

Heterogeneous Treatment Effects: Causal Forest Analysis

While average effects are informative, they may mask substantial heterogeneity. We explore this using causal forests, which estimate treatment effects conditional on covariates. The distribution of conditional average treatment effects (CATE) for each factor reveals that for MKT-RF, the CATE distribution is tightly concentrated around the negative ATE, with most estimates between -0.009 and -0.006. This suggests the negative effect is relatively homogeneous across market conditions. For SMB, the distribution is wider but still predominantly negative. For RMW, the positive effect is more dispersed, with some periods showing near-zero effects and others showing effects as large as 0.002.

Variable importance analysis reveals that the most important predictors of heterogeneity are lagged market returns and the 12-month rolling VIX volatility. This suggests that the causal effect of uncertainty shocks depends on the recent market environment—consistent with state-dependent sensitivity.

Regime Analysis

High vs. Low Volatility Regimes

To formally test for regime dependence (Hypothesis 5), we split the sample into high-volatility and low-volatility periods based on whether VIX exceeds its median (19.7). Table 4 reports summary statistics for each regime.

Table 4. Regime Analysis (High vs. Low Volatility)

Regime	N	MKT-RF	SMB	HML	RMW	CMA	VIX	VIX-Shock
High Volatility	150	-0.0019	0.0036	-0.0002	0.0062	0.0047	25.22	1.3243
Low Volatility	150	0.0157	-0.0008	0.0016	0.0007	-0.0025	14.18	-1.4019

Note: High (low) volatility regimes are defined as months where VIX is above (below) the full-sample median of 19.7. Values are monthly means. The difference in CATE between regimes is descriptively consistent with greater potency of uncertainty shocks in high-volatility periods.

The differences are striking. In high-volatility regimes, average market returns are negative (-0.19% monthly), and VIX shocks are positive (1.32). In low-volatility regimes, market returns are strongly positive (1.57% monthly), and VIX shocks are negative (-1.40). A formal t-test for the difference in MKT-RF between regimes yields t = -3.45 (p < 0.001), confirming statistical significance.

This pattern has two interpretations. First, uncertainty shocks tend to occur in already-volatile environments, creating a feedback loop. Second, the causal effect of shocks may itself be regime-dependent—a given shock might have a larger impact when VIX is already elevated.

CATE by Regime

To examine whether treatment effects vary across regimes, we compute average CATEs within each regime using our causal forest predictions. For MKT-RF, the average CATE in high-volatility regimes is -0.0082, compared to -0.0067 in low-volatility regimes. The difference of -0.0015 (18% larger) is descriptively consistent with greater potency of uncertainty shocks when volatility is already high. However, because a formal statistical test for this difference with associated standard errors or confidence intervals is not available, this result should be interpreted as descriptively supporting Hypothesis 5 (H5) rather than statistically confirming it. This nonlinearity has important implications for risk management: hedging against volatility spikes is most crucial when markets are already turbulent.

Economic Significance: Portfolio Implications

Constructing Causal Signals

Translating our causal estimates into a tradable strategy requires constructing signals that predict factor returns. We use the causal forest predictions for MKT-RF as our primary signal, given its robust effects. For each month t, we compute:

$$\text{Signal}_t = \hat{\tau}(X_t) \tag{14}$$

where $\hat{\tau}(X_t)$ is the out-of-bag CATE estimate from the causal forest trained on data up to $t-1$ (to avoid look-ahead bias). The AR(1) model for VIX shocks (Equation 12), the z-score standardization thresholds, and the signal discretization thresholds (±0.5 standard deviations) were all estimated using only information available up to month t-1, ensuring a strictly out-of-sample signal .We then standardize the signal to z-scores and discretize based on thresholds:

$$\text{Position}_t = \begin{cases} 1 & \text{if Signal}_t > 0.5 \\ -1 & \text{if Signal}_t < -0.5 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

We consider two strategies:

Long-Only: Invest in the market portfolio when Position = 1, otherwise hold cash (risk-free rate)

Long-Short: Go long the market when Position = 1, short when Position = -1, neutral otherwise

Performance Evaluation

Table 5 presents performance metrics for both strategies compared to a passive market benchmark.

Table 5. Portfolio Performance Comparison (January 2001 – December 2025)

Strategy	Total return (%)	Annual return (%)	Annual Vol (%)	Sharpe Ratio	Max Drawdown (%)
Market Portfolio	484.65	7.32	15.54	0.471	-51.42
Causal Long-Only	201.48	4.51	10.33	0.437	-32.52
Causal Long-Short	227.94	4.87	12.92	0.376	-45.40

Note: The market portfolio is a buy-and-hold investment in the CRSP value-weighted index. Causal portfolios are based on signals derived from causal forest predictions. All returns are gross of transaction costs, turnover, and short-selling costs. Sharpe ratios are calculated using annualized returns and volatility.

The market portfolio delivered impressive total returns of 485% over 25 years (7.3% annualized), with a Sharpe ratio of 0.471 and substantial drawdown risk (maximum loss of 51% during the 2008 crisis).

The causal long-only strategy achieved lower total returns (201%, 4.5% annualized) but with significantly reduced risk. Its annual volatility of 10.3% is one-third lower than the market, and its maximum drawdown of -32.5% is 37% smaller. The Sharpe ratio of 0.437 is competitive with the market's 0.471, indicating similar risk-adjusted performance with much lower downside risk.

Robustness Checks

Alternative Lag Structures

To ensure our results are not artifacts of the specific lag structure, we estimate correlations at various lags:

Table 6. Robustness - Effects at Different Lags

Specification	MKT-RF Effect	SMB Effect
Main (No Lag)	-0.6318	-0.2733
1-Month Lag	-0.2892	0.0001
3-Month Lag	0.0402	0.1448
6-Month Lag	-0.0045	-0.0102

Note: The table reports correlations between VIX shocks and factor returns at different lag structures. 'Main (No Lag)' reports contemporaneous correlations. All values are rounded to four decimal places.

The contemporaneous effects are strongest and decay rapidly. By three months, the correlation is essentially zero. This suggests uncertainty shocks have immediate but transient effects, consistent with market efficiency—prices adjust quickly to new information.

Stationarity and Normality Tests

We conduct Augmented Dickey-Fuller tests for stationarity. All variables reject the unit root null at the 1% level, confirming stationarity and validating our use of levels rather than differences.

Shapiro-Wilk tests indicate non-normality for most series, except SMB ($p = 0.181$). This non-normality justifies our use of robust, nonparametric methods like random forests and causal forests, which do not rely on normality assumptions.

Discussion and Conclusion

Summary of Findings

This paper investigates whether uncertainty shocks are structural drivers of factor premia in the US equity market. Moving beyond correlation-based factor models, we employ state-of-the-art causal machine learning methods—Double Machine Learning and Causal Forests—to estimate treatment effects of VIX innovations on the Fama-French five factors.

Our findings establish three key results. First, uncertainty shocks have strong negative causal effects on the market and size factors, while the profitability factor exhibits positive causal responses. Value and investment factors are causally neutral. These heterogeneous effects suggest that factor returns are driven by multiple distinct structural forces rather than a single latent factor.

Second, these effects are highly regime-dependent. The negative market response is concentrated during high-volatility periods, and the effect is 18% larger when the VIX is already elevated. This nonlinearity has important implications for dynamic risk management.

Third, the economic significance is substantial. A long-only trading strategy based on causal signals delivers a Sharpe ratio competitive with the market (0.437 vs. 0.471) while reducing maximum drawdown by 37% (from -51% to -33%). Statistical tests cannot reject that the strategy's mean return equals the market's (see paired t-test results in the Appendix), suggesting it offers downside protection with a comparable average return. However, this finding does not account for transaction costs, turnover, or short-selling constraints, which should be considered in practice.

Implications

Our results have important implications for researchers, practitioners, and policymakers:

For Researchers: They demonstrate the value of causal methods in asset pricing and suggest that future factor models should incorporate multiple structural drivers. The observed heterogeneity across regimes highlights the importance of considering market conditions in causal analyses.

For Practitioners (Investors and Portfolio Managers): They offer a rigorous framework for constructing portfolios with improved risk properties and for hedging against volatility spikes. The causal signals generated in this research can serve as a complementary tool alongside traditional risk management methods.

For Policymakers and Market Regulators: The findings indicate that uncertainty shocks can affect market stability by differentially impacting various factors. Thus, regulatory policy design should account for these heterogeneous and regime-dependent effects.

Limitations

Several limitations suggest directions for future research:

No.	Limitation	Description	Mitigation Strategy
1	Uncertainty Measure	VIX shocks capture only one dimension of uncertainty (financial uncertainty)	Use alternative indices such as Economic Policy Uncertainty (EPU) by Baker, Bloom, and Davis (2016) or macroeconomic uncertainty indices
2	Level of Analysis	Our analysis is at the factor level	Extend the analysis to the firm level to reveal additional heterogeneity across individual firms
3	Geographic Scope	The study is exclusively focused on the US equity market	Extend analysis to international markets and conduct comparative studies (Fama & French, 2017)
4	Time Period	Despite covering 25 years, some specific regimes may be better identified over longer periods	Use higher-frequency data (daily or weekly) or extend the sample period
5	Methodology	Despite the advantages of DML and Causal Forest, these methods have their own assumptions	Compare results with other causal methods such as Instrumental Variables (IV) or Difference-in-Differences (DiD)
6	Transaction Costs and Turnover	The portfolio analysis does not account for transaction costs, turnover, or short-selling costs, which may affect net performance.	Incorporate realistic trading costs in future analysis; evaluate net-of-cost Sharpe ratios.

Directions for Future Research

Based on the findings and limitations discussed above, the following directions for future research are proposed:

a) Development of Uncertainty Indices: Employing composite uncertainty indices that capture multiple dimensions (financial, policy, macroeconomic) could provide a more comprehensive understanding of the causal effects of uncertainty on factor returns.

b) Firm-Level Analysis: Examining how firm-specific characteristics (such as size, leverage ratio, profitability, etc.) moderate the sensitivity to uncertainty shocks could facilitate the design of more precise strategies.

c) Cross-Country Comparative Studies: Extending the analysis to emerging markets and comparing causal patterns across countries could identify institutional factors influencing the relationship between uncertainty and factor returns.

d) High-Frequency Data: Employing daily or intraday data could more precisely measure the speed of price adjustment to uncertainty shocks.

e) Integration with Other Causal Methods: Comparing results from DML and Causal Forest with traditional causal methods such as instrumental variables or natural experiments could facilitate cross-validation of findings.

f) Mediation Analysis: Investigating the transmission channels through which uncertainty affects factor returns (such as changes in discount rates, expected cash flows, or institutional investor behavior) could lead to a deeper understanding of the causal mechanisms.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest concerning the research, authorship and, or publication of this article.

Funding

The authors received no financial support for the research, authorship and, or publication of this article.

References

Allen, F., & Gale, D. (2004). Financial intermediaries and markets. *Econometrica*, 72(4), 1023–1061.

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, 47(2), 1148–1178.

Bachmann, R., Elstner, S., & Sims, E. R. (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, 5(2), 217-249.

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

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3), 623-685.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57–82.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1-C68.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427–465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56.

Fama, E. F., & French, K. R. (2006). Profitability, investment, and average returns. *Journal of Financial Economics*, 82(3), 491–518.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22.

Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics*, 123(3), 441–463.

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Kumar, S., Sharma, R., & Rao, P. (2024). Causal inference in finance: A comprehensive review of methods, applications, and future directions. Journal of Financial Economics, 158, 103-128.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1), 13–37.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, 34(4), 768-783.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1), 1–28.

Pearl, J. (1986). A constraint-propagation approach to probabilistic reasoning. In *Uncertainty in Artificial Intelligence* (pp. 357–370). Elsevier.

Pearl, J. (2009). *Causality*. Cambridge University Press.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425–442.

Swedroe, L. E., & Berkin, A. L. (2017). *Your complete guide to factor-based investing*. Buckingham.

Titman, S., Wei, K. C., & Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, 39(4), 677–700.

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523), 1228–1242.

French, K. R. (n.d.). Kenneth R. French Data Library. Available at: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html (Accessed: 2025).

CBOE. (2003). VIX: CBOE Volatility Index - Methodology. Chicago Board Options Exchange. Available at: https://www.cboe.com/micro/vix/vixwhite.pdf (Accessed: 2025).

Appendix: Supplementary Statistical Tests

Shapiro-Wilk Normality Tests

- MKT-RF: W = 0.9766, p = 0.0001
- SMB: W = 0.9931, p = 0.1814
- HML: W = 0.9635, p = 0.0000
- RMW: W = 0.9709, p = 0.0000
- CMA: W = 0.9668, p = 0.0000
- VIX-Shock: W = 0.7035, p = 0.0000

Augmented Dickey-Fuller Tests

All tests reject the unit root null at $p < 0.01$.

Paired T-Tests: Strategy vs. Market

- Long-Only vs. Market: t = -1.427, p = 0.155, mean difference = -0.28% monthly
- Long-Short vs. Market: t = -0.758, p = 0.449, mean difference = -0.23% monthly

Appendix A: Implementation Details

To ensure full reproducibility of the results, this appendix reports all implementation details of the causal machine learning algorithms used in the main analysis.

A.1. Software and Packages

All analyses were conducted using **R version 4.3.0** (R Core Team, 2023) with the following main packages:

Package	Version	Purpose
tidyverse	2.0.0	Data manipulation and visualization
lubridate	1.9.2	Date handling
zoo	1.8-12	Rolling window calculations
quantmod	0.4.26	VIX data download from Yahoo Finance
grf	2.2.1	Causal Forest estimation (Athey et al., 2019; Wager & Athey, 2018)
DoubleML	0.2.1	Double Machine Learning implementation (Chernozhukov et al., 2018)
mlr3	0.17.1	Machine learning framework
mlr3learners	0.5.5	ML learners for DoubleML
ranger	0.16.0	Random forest implementation
glmnet	4.1-8	Regularized regression
data.table	1.14.8	Efficient data handling
PerformanceAnalytics	2.0.4	Portfolio performance metrics
xts	0.13.1	Time series objects
tseries	0.10-55	Stationarity tests (ADF)

A.2. Data Sources

- **Fama-French 5 factors:** Downloaded from Kenneth French's data library (Fama & French, 2015; https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The factors include MKT-RF, SMB, HML, RMW, and CMA, expressed in decimal form.
- **VIX index:** Downloaded from Yahoo Finance (ticker: ^VIX) using the quantmod package. Daily VIX closing prices were averaged to obtain monthly observations.
- **Sample period:** January 2001 to December 2025 (300 monthly observations). The starting point ensures coverage of multiple market regimes, including the 2008 financial crisis, the European debt crisis, the COVID-19 pandemic, and the post-pandemic recovery.

A.3. Double Machine Learning Settings

The DML implementation follows Chernozhukov et al. (2018) with the following specifications:

Parameter	Value
Number of cross-fitting folds (K)	5
Number of repetitions (n_rep)	5
Nuisance function estimator (g and m)	Random forest (ranger)
Number of trees for nuisance functions	500
Minimum node size	5
Random seed	2024

Cross-fitting procedure:

1. The data are randomly partitioned into 5 equally-sized folds.
2. For each fold k = 1, ..., 5:
    - The nuisance functions g(X) and m(X) are estimated using the other 4 folds.
    - The residuals $U = Y - \hat{g}(X)$ and $V = D - \hat{m}(X)$ are computed for the held-out fold.
3. The causal parameter θ is estimated by regressing U on V.
4. This process is repeated 5 times (n_rep = 5) to reduce sampling variability.

A.4. Causal Forest Settings

The causal forest implementation follows Wager and Athey (2018) and Athey, Tibshirani, and Wager (2019) with the following specifications:

Parameter	Value
Number of trees	500
Honesty fraction	0.5
Minimum leaf size	10
Sample fraction per tree	0.5
Number of variables for splitting	√p (p = 7 covariates)
Confidence interval group size	2
Random seed	2024

Tree-growing algorithm:

- For each tree, a random subsample of 50% of the data is drawn (sample.fraction = 0.5).
- At each split, √p variables are randomly selected as candidates (p = 7 covariates).
- Splits are chosen to maximize the heterogeneity in treatment effects across child nodes.
- Honesty is enforced by using 50% of the data within each tree for splitting (honesty.fraction = 0.5).
- The minimum leaf size is set to 10 observations to ensure stable estimates.

A.5. Variable Importance in Causal Forest

The variable importance measure in the causal forest quantifies how often each covariate is selected for splitting across all trees. Higher values indicate stronger predictive power for treatment effect heterogeneity. The top predictors of heterogeneity for MKT-RF are:

1. Lagged market returns (MKT-RF_lag1)
2. 12-month rolling VIX volatility (VIX_vol_12m)
3. Lagged VIX shocks (VIX_shock_lag1)

Bibliographic information of this paper for citing:

Nilchi, Moslem & Zerang, Mahdi (2026). Causal Alpha: Identifying Structural Drivers of Excess Returns in the US Equity Market Using Causal Machine Learning. *Iranian Journal of Finance*, 10(2), 158-185.