Risk-controlled Bitcoin allocation with Machine-Learning views

PRELIMINARY DRAFT

Rosella Castellano[1] Federico Cini[2,*] Annalisa Ferrari[3]

July 2026

[1] Department of Law and Economics, University of Rome UnitelmaSapienza
[2] Department of Social Sciences and Economics, Sapienza University of Rome
[3] Department of Law and Digital Society, University of Rome UnitelmaSapienza
[*] Corresponding author, Email: federico.cini@uniroma1.it

Abstract

This study examines the economic conditions under which Bitcoin can improve a traditional portfolio without becoming a source of excessive risk. The central question is whether imperfect predictive information can be converted into economically useful allocation decisions when uncertainty and exposure are explicitly controlled. Using a portfolio composed of Bitcoin, the S&P 500, and gold, machine-learning probabilities of positive next-period Bitcoin returns are translated into expected-return views and incorporated as private information/beliefs into a Black–Litterman framework, which combines them with a strategic prior according to their estimated uncertainty. The analysis evaluates whether these probabilistic views improve risk-adjusted performance relative to an equity–gold benchmark, static Bitcoin allocations, and conventional rolling mean–variance portfolios. The results show that Bitcoin contributes most effectively as a limited complementary asset. Small allocation caps improve the portfolio's return–risk balance, whereas progressively larger admissible exposures increase volatility and drawdowns and reduce overall efficiency. The proposed framework also produces more stable portfolios and substantially lower turnover than rolling mean–variance optimization, although the latter achieves higher absolute returns. The evidence suggests that Bitcoin's economic value depends less on strong predictability than on disciplined belief regularization and exposure control.

Keywords: Bitcoin; Machine learning; Black–Litterman model; Portfolio allocation; Probabilistic forecasting

JEL classification: G11; G17; C11

1 Background, motivation and research questions

Portfolio allocation is fundamentally a decision problem under uncertainty. Within the mean–variance framework, portfolio weights depend on expected returns and the covariance structure of asset returns (Markowitz, 1952). Because these parameters are unobservable, investors must rely on estimates, and portfolio decisions therefore reflect beliefs about future returns and risks rather than their subsequent realizations. Both expected returns and covariance parameters are estimated with error. However, portfolio weights are especially sensitive to errors in expected returns, since relatively small differences in estimated means may be translated by the optimizer into large and concentrated positions (Best and Grauer, 1991). In this sense, portfolio optimization may interpret estimation noise as an investment opportunity. The relevant problem is therefore not merely that returns are difficult to forecast, but that uncertain forecasts can generate economically significant allocation errors once they are converted into portfolio exposures.

Bitcoin provides a particularly challenging environment for this problem. Originally conceived as a peer-to-peer electronic cash system (Nakamoto, 2008), it has evolved into a financial asset increasingly considered in portfolio-allocation decisions. Bitcoin combines technological, monetary, and financial characteristics that make it difficult to classify within conventional asset-pricing categories (Böhme et al., 2015; Biais et al., 2019; Chiu and Koeppl, 2022). Moreover, it is a non-cash-flow-generating asset whose valuation depends strongly on expectations, adoption, scarcity, liquidity, and investor attention (Ciaian et al., 2016; Hayes, 2017, 2019; Cong et al., 2021). These features make both its expected return and its appropriate portfolio weight particularly difficult to estimate.

The portfolio role of Bitcoin is further complicated by its unstable dependence with traditional assets. Earlier evidence emphasized phases of weak correlation and potential diversification benefits (Brière et al., 2015; Dyhrberg, 2016a,b; Corbet et al., 2019). Related studies have examined whether Bitcoin can act as a diversifier, hedge, or safe-haven asset, frequently comparing it with gold and conventional financial assets (Bouri et al., 2017; Smales, 2019; Selmi et al., 2018; Wen et al., 2022). More recent evidence, however, documents increasing exposure to conventional financial markets, macro-financial conditions, investor attention, and liquidity cycles (Zhang et al., 2021; Bouri et al., 2021; Adelopo and Luo, 2025). During periods of market stress, Bitcoin may behave less as an independent diversifier and more as a high-volatility risky asset (Conlon and McGee, 2020; Liu and Yuan, 2024; Gorman and Hughen, 2024). Under these conditions, historical estimates and model-based forecasts of Bitcoin should be treated as uncertain inputs to the portfolio decision process rather than as precise measures of future investment opportunities.

From this perspective, the portfolio value of a predictive signal cannot be assessed independently of the decision rule through which it is used. Its contribution depends on the information contained in the signal, the uncertainty attached to it, its interaction with the remaining portfolio assets, and the mechanism through which it is translated into portfolio exposure. This dependence is particularly important because the apparent predictability of Bitcoin may vary across samples, market regimes, liquidity conditions, transaction costs, and implementation choices (Noda, 2021; Sebastião and Godinho, 2021; Bellocca et al., 2022; Guo et al., 2025). Portfolio allocation can therefore be approached as a decision problem in which uncertain forecasts are transformed into positions. The economic value of predictive information is jointly determined by the information content of the signal, the uncertainty assigned to it, and the decision architecture through which it is translated into exposure.

The quality of the informational input is consequently a central element of the allocation problem. A large body of research has examined whether Bitcoin returns contain predictable components, linking cryptocurrency prices to investor attention, search intensity, sentiment, social-media information, news flows, market conditions, and cross-asset dynamics (Valencia et al., 2019; Subramaniam and Chakraborty, 2020; Zhu et al., 2021; Aslanidis et al., 2022; Yao et al., 2024). Other contributions investigate technical trading rules, market inefficiencies, bubble dynamics, and machine-learning-based prediction (Grobys et al., 2020; Fry and Cheah, 2016; Kyriazis et al., 2020; Svogun and Bazán-Palomino, 2022; Khedr et al., 2021; Akyildirim et al., 2021; Fang et al., 2025). This evidence suggests that potentially relevant information may be distributed across a broad and heterogeneous set of financial variables, although its stability and economic usefulness remain uncertain.

Machine-learning methods provide flexible tools for extracting such information because they can accommodate large predictor sets, nonlinearities, and complex interactions without imposing an overly restrictive functional form (Breiman, 2001; Chen and Guestrin, 2016; Gu et al., 2020; Alpaydin, 2020). This flexibility is relevant for Bitcoin, whose relationship with traditional financial variables may change across return horizons and market conditions. Interpretability methods such as SHAP values can also be used to investigate which variables and economic groups contribute to the fitted relationships (Lundberg and Lee, 2017). Model flexibility, however, does not eliminate predictive uncertainty (Alpaydin, 2020). Accordingly, the role of machine learning in the present framework is informational rather than decisional. It transforms observable financial variables into time-varying probabilistic assessments of future Bitcoin returns, but it neither determines portfolio weights directly nor determines how much influence those assessments should receive in portfolio construction.

The preceding discussion establishes two central points. First, portfolio decisions rely on imperfect beliefs rather than true expected returns. Second, machine-learning probabilities represent uncertain and model-dependent information. A suitable allocation framework must therefore regulate both how predictive information modifies expected-return estimates and how those estimates are translated into actual portfolio positions.

The Black–Litterman framework provides a natural mechanism for combining structured prior information with new evidence while explicitly accounting for the uncertainty associated with both sources. Bayesian updating produces posterior expected returns that reflect the relative precision assigned to the prior and the additional view (Black and Litterman, 1992; Idzorek, 2007; Meucci, 2008). It thereby regulates how strongly uncertain predictive information modifies the expected-return inputs used in portfolio construction. In the proposed framework, the predicted probability, the expected-return view, and the uncertainty assigned to that view are distinct modelling objects. The probability conveys directional information, the probability-to-view mapping translates that information into return units, and the uncertainty specification determines the influence of the resulting view relative to the prior. Neither the return mapping nor the uncertainty assigned to the view is mechanically implied by the predicted probability.

Bayesian belief regularization alone does not fully control portfolio risk. Exposure constraints provide a second and distinct form of discipline. Black–Litterman updating regulates how predictive beliefs affect posterior expected returns, whereas portfolio constraints regulate how those posterior returns affect actual positions. This distinction is especially important for Bitcoin. Because Bitcoin combines uncertain expected returns, high volatility, and unstable dependence with traditional assets, larger admissible allocations increase the portfolio consequences of both informative and erroneous beliefs. A Bitcoin allocation cap therefore acts as an additional regularization mechanism, limiting the extent to which a volatile and difficult-to-estimate asset can dominate the portfolio.

The proposed framework consequently combines three forms of discipline: probabilistic signal formation, because machine learning generates time-varying probability estimates rather than deterministic trading signals; Bayesian belief regularization, because the resulting view is combined with a structured prior according to its assigned uncertainty; and portfolio exposure control, because long-only constraints and explicit Bitcoin caps limit the translation of posterior beliefs into portfolio positions. The central empirical issue is therefore whether imperfect probabilistic information can contribute to portfolio efficiency when both its influence on expected returns and its translation into portfolio exposure are explicitly controlled.

Against this background, the study addresses the following research questions:

RQ1: Can portfolios that incorporate machine-learning-generated probabilistic views through a Bayesian allocation framework achieve higher risk-adjusted performance than passive and conventional active benchmark portfolios?

RQ2: Do the risk-adjusted benefits of including Bitcoin diminish as the maximum admissible Bitcoin exposure increases?

The empirical design consists of two complementary stages. The first is an exploratory machine-learning analysis that investigates the relationship between Bitcoin and a broad set of traditional financial variables at daily, weekly, and biweekly horizons. This stage characterizes the information environment from which the probabilistic beliefs are generated and examines whether nonlinear and horizon-dependent associations emerge across equity, currency, commodity, interest-rate, volatility, and investor-attention variables. The exploratory analysis provides an interpretable assessment of the information set subsequently used in portfolio construction.

The second stage evaluates the portfolio consequences of incorporating those beliefs into the allocation process. A weekly walk-forward design is used in which lag-augmented predictors are subjected to rolling standardization and principal component analysis before estimating the probability of a positive next-week Bitcoin return. The estimated probability is mapped into a volatility-scaled expected-return view and assigned an explicit degree of uncertainty. The resulting view is combined with a Black–Litterman prior and used in a constrained long-only optimization problem. All preprocessing, probability estimation, posterior updating, and portfolio formation steps are performed using information available at the corresponding decision date.

The investable universe is intentionally stylized and consists of Bitcoin, the S&P 500, and gold. Bitcoin represents the high-volatility digital asset whose controlled integration is the focus of the analysis, the S&P 500 represents the conventional risky component, and gold provides a defensive asset frequently considered in diversification and safe-haven studies (Barro and Misra, 2016; Baur et al., 2018; Dyhrberg, 2016a). The broader set of financial variables is used as predictive information rather than as part of the investable universe. This design isolates the contribution of dynamically managed Bitcoin exposure within a conventional equity–gold allocation.

The contribution of the paper is both methodological and empirical. Methodologically, the study develops a direct link between probabilistic machine-learning outputs and portfolio construction by treating predicted probabilities as uncertain Black–Litterman views rather than as binary buy or sell signals. The framework separates signal generation, return-view construction, uncertainty modelling, and exposure control, thereby allowing the economic value of predictive information to be evaluated within the complete allocation process. Empirically, the study examines whether this mechanism improves risk-adjusted portfolio performance relative to three benchmark families: a no-Bitcoin equity–gold portfolio, static Bitcoin allocations under equivalent exposure limits, and conventional rolling mean–variance optimization. The research questions are evaluated through out-of-sample benchmark comparisons and an extensive specification-based robustness analysis covering alternative probability models, view mappings, uncertainty specifications, covariance windows, risk-aversion parameters, and Bitcoin allocation caps.

The remainder of the paper is organized as follows. Section 2 presents the dataset, variable construction, and exploratory machine-learning analysis. Section 3 develops the ML-informed Black–Litterman framework, including the rolling preprocessing procedure, probability-to-view mapping, uncertainty specification, posterior-return construction, and constrained optimization. Section 4 reports the portfolio results, benchmark comparisons, and robustness analysis. Section 5 concludes.

2 Exploratory ML Analysis

2.1 Dataset and Variable Construction

The empirical analysis is based on a daily price-level dataset covering the period from January 2018 to February 2026. The sample is selected to include both mature cryptocurrency-market phases and recent episodes of stronger interaction between digital assets and traditional financial markets. Bitcoin is used as the target asset throughout the exploratory analysis, while the explanatory information set includes variables representing sets of different assets. This broad cross-market design is motivated by the evidence that Bitcoin valuation is affected not only by cryptocurrency-specific factors, but also by liquidity conditions, investor attention, macro-financial shocks, and time-varying spillovers from conventional markets (Ciaian et al., 2016; Hayes, 2019; Zhang et al., 2021; Aslanidis et al., 2022; Adelopo and Luo, 2025).

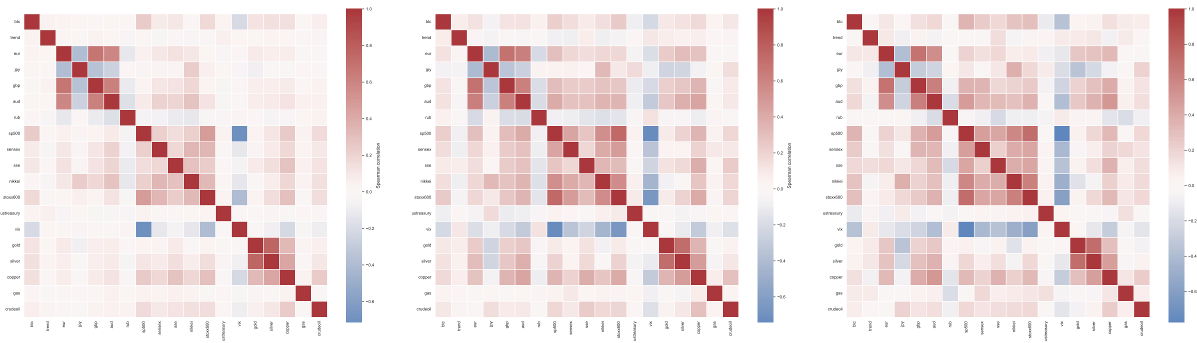
Table 1: Dataset, variable groups, and roles

Group	Representative variables	Role
Crypto	Bitcoin	Target variable
Attention	Google Trends	Investor attention proxy
Currencies	EUR, JPY, GBP, AUD, RUB	Exchange-rate conditions
Equity markets	S&P 500, Sensex, SSE, Nikkei, Stoxx 600	Global risk appetite
Interest rates	US Treasury	Rate and discount-rate conditions
Market volatility	VIX	Market uncertainty and risk aversion
Metals	Gold, Silver, Copper	Safe-haven and commodity channels
Energy	Gas, Crude oil	Energy and macro-cost shocks

The grouping in Table 1 is used for both empirical organization and economic interpretation. Equity indices capture global risk-on/risk-off conditions, while the VIX and Treasury-rate variables proxy for uncertainty and discount-rate effects. Gold is included because of its traditional role as a defensive asset and because its relationship with Bitcoin has been widely discussed in the literature on hedging and safe-haven properties (Dyhrberg, 2016a; Bouri et al., 2017; Baur et al., 2018; Wen et al., 2022). Metals and energy variables capture broader commodity-market conditions, whereas currencies reflect global macro-financial and exchange-rate movements. Finally, Google Trends is included as an attention-based variable, in line with the literature emphasizing investor attention, search intensity, and sentiment as relevant drivers of cryptocurrency price dynamics (Valencia et al., 2019; Subramaniam and Chakraborty, 2020; Zhu et al., 2021; Yao et al., 2024).

Logarithmic returns are computed at three horizons: daily, weekly, and biweekly. Daily returns are obtained directly from the aligned daily price series. Weekly and biweekly returns are instead computed by resampling prices using the last available Friday observation and then taking logarithmic differences. The use of multiple horizons is important because Bitcoin-market relationships may differ between high-frequency noise, weekly adjustment, and lower-frequency allocation-relevant dynamics. In particular, weekly and biweekly transformations reduce the influence of daily idiosyncratic fluctuations and are closer to the medium-frequency portfolio-allocation perspective adopted later in the paper (Dehouche, 2021; Petukhina et al., 2021). Extreme observations are treated using a 30-observation rolling z-score winsorization procedure to reduces the influence of extreme observations while preserving the temporal ordering of the data.

The feature-screening step combines pairwise rank correlations and variance-inflation diagnostics. Figure 1 show Spearman correlations computed separately for each horizon to examine the association between Bitcoin returns and traditional financial variables and to identify redun-



(a) Daily (b) Weekly (c) Biweekly

Figure 1: Spearman correlation matrices by return horizon

dant predictors before estimating the machine-learning models (Liu and Yuan, 2024). Several additional variables were initially considered in the broader dataset, but they were pruned before the main analysis when they exhibited excessive pairwise correlation with retained variables or contributed to multicollinearity according to VIF diagnostics. In particular, highly overlapping equity and currency variables were excluded: the Dow Jones and Nasdaq indices were removed because of their strong overlap with the S&P 500, the FTSE index because of its close relationship with Stoxx 600, and the New Zealand dollar because of its high correlation with the Australian dollar. After this screening step, the final selected variables do not exhibit severe multicollinearity: the maximum VIF is 2.437 for the daily dataset, 3.439 for the weekly dataset, and 4.039 for the biweekly dataset. The remaining correlation structure is economically informative rather than mechanically redundant. The strongest rank correlations with Bitcoin are presented in Table 2. Higher correlation values are concentrated among equity-market and market-volatility variables. At the daily horizon, the S&P 500 displays the largest positive correlation with Bitcoin returns, while the VIX shows the largest negative association. The same pattern persists at the weekly horizon, with Stoxx 600 and S&P 500 among the most correlated positive variables and the VIX retaining a negative relationship. At the biweekly horizon, these associations become stronger in absolute value: Bitcoin is positively associated with the S&P 500 and other equity indices, while its correlation with the VIX becomes more negative.

Table 2: Strongest Spearman correlations with Bitcoin returns

	Variable	Spearman correlation with Bitcoin
	S&P 500	0.238
	VIX	-0.206
Daily	Stoxx 600	0.164
	Copper	0.137
	Silver	0.134
	Stoxx 600	0.246
	S&P 500	0.224
Weekly	VIX	-0.211
	Copper	0.163
	Sensex	0.150
	VIX	-0.359
	S&P 500	0.342
Biweekly	Stoxx 600	0.297
	Nikkei	0.274
	Sensex	0.241

2.2 Machine-Learning Models

The analysis compares linear and nonlinear supervised-learning models estimated separately for different returns horizons. Let $r_{t,h}^{BTC}$ denote the Bitcoin log-return at horizon $h \in \{\text{daily, weekly, biweekly}\}$ and let $\mathbf{x}_{t,h}$ be the corresponding vector of screened predictors described in Section 2.1. For each horizon, the generic relation can be written as

$$r_{t,h}^{BTC} = f_h(\mathbf{x}_{t,h}) + \varepsilon_{t,h}, \tag{1}$$

where $f_h(\cdot)$ is allowed to vary across model classes. The objective is not to obtain an accurate forecasting model, but to assess whether the selected cross-market variables contain linear or nonlinear information about Bitcoin returns and whether such information can be interpreted in economically meaningful terms. In this sense, ML is widely recognised as a useful tool and is often employed in empirical asset pricing and financial forecasting to detect nonlinearities and interaction effects that are difficult to capture with standard linear specifications (Gu et al., 2020; Alpaydin, 2020; Khedr et al., 2021).

Linear Regression here is used as the benchmark model. It assumes that the conditional mean of Bitcoin returns is a linear function of the input vector,

$$\widehat{r}_{t,h}^{BTC} = \widehat{\alpha}_h + \mathbf{x}_{t,h}'\widehat{\boldsymbol{\beta}}_h, \tag{2}$$

with parameters estimated by minimizing the residual sum of squares. This specification provides a transparent reference point and allows the analysis to evaluate whether more flexible nonlinear models improve explanatory performance relative to a simple linear relationship (Hastie, 2009; James et al., 2013).

Support Vector Regression (SVR) is included as a nonlinear kernel-based benchmark. Using a radial basis function (RBF) kernel, SVR estimates a function that is flat while allowing deviations within an ε-insensitive tube around the fitted value (Vapnik, 2013; Cortes and Vapnik, 1995). In simplified form, the SVR problem can be written as

$$\min_{f} \ \frac{1}{2}\|f\|^2 + C\sum_{t}\left(\xi_t + \xi_t^*\right), \tag{3}$$

subject to deviations from the fitted function being no larger than ε up to slack variables ξ_t and ξ_t^*. In the empirical implementation, the SVR uses an RBF kernel, $C = 0.3$, and $\varepsilon = 0.1$. The relatively conservative value of C limits the penalty assigned to errors, reducing the risk that the model adapts too aggressively to noisy Bitcoin return observations.

The tree-based models are Random Forest and XGBoost. Random Forest (RF) estimates an ensemble of regression trees and averages their predictions,

$$\widehat{f}_{RF}(\mathbf{x}_{t,h}) = \frac{1}{B}\sum_{b=1}^{B} T_b(\mathbf{x}_{t,h}), \tag{4}$$

where $T_b(\cdot)$ denotes the b-th tree and B is the number of trees (Breiman, 2001). This structure allows the model to capture nonlinearities and interactions while reducing the variance of individual trees through bootstrap aggregation. The implementation uses 300 trees, a maximum depth of 6, a minimum leaf size of 3, and $\sqrt{p}$ candidate predictors at each split. These parameters impose moderate regularization: tree depth is restricted to avoid overly complex partitions, the minimum leaf size prevents splits based on very small samples, and random feature selection reduces dependence on highly correlated predictors. Out-of-bag scoring is retained as an internal diagnostic of generalization performance.

XGBoost is used as a gradient-boosted tree model. It approximates the prediction function through an additive expansion of regression trees,

$$\widehat{f}_{XGB}(\mathbf{x}_{t,h}) = \sum_{k=1}^{K} \eta\, T_k(\mathbf{x}_{t,h}), \tag{5}$$

where K is the number of boosting iterations and η is the learning rate (Chen and Guestrin, 2016). The model is estimated using the squared-error objective. The baseline specification uses 200 trees, maximum depth equal to 3, learning rate equal to 0.05, minimum child weight equal to 10, subsampling equal to 0.7, and column subsampling equal to 0.7. This configuration is deliberately conservative: shallow trees limit interaction complexity, a low learning rate reduces the contribution of each individual tree, the minimum child weight discourages splits with limited statistical support, and row and column subsampling mitigate overfitting.

Model performance is assessed using complementary error, explanatory, and directional metrics. Average forecast errors are evaluated through the mean absolute error (MAE), median absolute error ($MedAE$), and root mean squared error ($RMSE$), while explanatory performance is measured using the coefficient of determination (R^2), and explained variance. The reader is referred to Alpaydin (2020) for a detailed description of both models and metrics. Additionally, Directional accuracy is also reported as

$$DA = \frac{1}{T} \sum_{t=1}^{T} \mathbb{1} \left[\operatorname{sign}\left(r_{t,h}^{BTC} \right) = \operatorname{sign}\left(\widehat{r}_{t,h}^{BTC} \right) \right], \tag{6}$$

where $\mathbb{1}[\cdot]$ is an indicator function. DA is included as an exploratory diagnostic because the sign of Bitcoin returns is relevant for the later probability-based allocation framework, but it is not interpreted as sufficient evidence of a profitable trading rule.

Finally, model-interpretability tools are used to connect models outputs with economic interpretation. For tree-based models, feature-importance measures are computed and then interpreted both at the individual-variable level and by economic group. For XGBoost, SHAP values are used to decompose the fitted prediction into additive feature contributions (Lundberg and Lee, 2017). For observation t, the SHAP representation can be written as

$$\widehat{f}(\mathbf{x}_{t,h}) = \phi_0 + \sum_{j=1}^{p} \phi_{j,t}, \tag{7}$$

where ϕ_0 is the baseline prediction and $\phi_{j,t}$ is the contribution of predictor j for observation t. SHAP values provide a local and global interpretation of nonlinear models while preserving an additive attribution structure.

2.3 Explanatory evidence and economic interpretation

Table 3 reports the performance of the different models across different horizons. The ranking is stable across frequencies: XGBoost provides the strongest explanatory performance, followed by Random Forest, while SVR and Linear Regression generally produce weaker results. The improvement is particularly visible when moving from daily to weekly and biweekly returns. This pattern suggests that the relationship between Bitcoin and traditional financial variables becomes more structured at lower frequencies, where idiosyncratic daily noise is attenuated.

Table 3: Exploratory ML performance by return horizon

Horizon	Model	MAE	MedAE	RMSE	R^2	Dir. Acc.	Spearman corr.
Daily	Linear	0.0277	0.0187	0.0016	0.084	0.619	0.290
	SVR	0.0315	0.0243	0.0017	0.007	0.543	0.157
	RF	0.0251	0.0171	0.0013	0.265	0.705	0.575
	XGBoost	0.0228	0.0156	0.0010	0.396	0.729	0.648
Weekly	Linear	0.0601	0.0446	0.0066	0.101	0.610	0.313
	SVR	0.0546	0.0507	0.0043	0.410	0.600	0.479
	RF	0.0449	0.0345	0.0036	0.512	0.763	0.804
	XGBoost	0.0325	0.0242	0.0018	0.752	0.835	0.877
Biweekly	Linear	0.0822	0.0643	0.0122	0.181	0.632	0.411
	SVR	0.0693	0.0720	0.0069	0.535	0.703	0.631
	RF	0.0578	0.0401	0.0061	0.588	0.804	0.855
	XGBoost	0.0364	0.0271	0.0023	0.847	0.856	0.924

Residual diagnostics for the XGBoost specifications confirm that the fitted models do not eliminate the main distributional features of Bitcoin returns. Residual autocorrelation is not detected at the daily, weekly, or biweekly horizon. However, heteroskedasticity remains present, and residual normality is rejected at the daily and weekly frequencies. These diagnostics are consistent with the view that lower-frequency returns are smoother, but they also confirm that Bitcoin returns retain volatility clustering and tail-risk characteristics that are not fully captured by the predictor set.

The variable-level interpretation focuses on the XGBoost specification, which provides the strongest explanatory performance across horizons. Table 4 jointly reports the five most relevant variables according to standard XGBoost feature importance and mean absolute SHAP values. XGBoost feature importance reflects the relevance of a variable within the boosted-tree structure, whereas mean absolute SHAP values measure the average contribution of each variable to the fitted prediction. The comparison therefore separates split-based relevance from contribution-based relevance in the nonlinear fitted function.

Table 4: XGBoost feature importance and mean absolute SHAP top 5 rankings across horizons

	Feature importance		Mean absolute SHAP	
	Feature	Importance	Feature	Mean absolute SHAP
Daily	S&P 500	0.081	S&P 500	0.00337
	VIX	0.079	Google Trends	0.00318
	Google Trends	0.068	VIX	0.00310
	Silver	0.066	Silver	0.00218
	AUD	0.057	Nikkei	0.00168
Weekly	Google Trends	0.079	Stoxx 600	0.00998
	Stoxx 600	0.076	RUB	0.00740
	Copper	0.067	SSE	0.00722
	US Treasury	0.060	Copper	0.00719
	Crude oil	0.059	Google Trends	0.00710
Biweekly	VIX	0.103	VIX	0.02233
	US Treasury	0.075	US Treasury	0.02100
	Nikkei	0.072	Copper	0.01257
	Google Trends	0.071	Nikkei	0.01219
	Stoxx 600	0.068	Google Trends	0.01091

The results in Table 4 and the SHAP summary plots in Figure 2 indicate that the relevant information set changes substantially with the return horizon. At the daily frequency, the most relevant variables are concentrated around equity-market conditions, investor attention, and market stress. The S&P 500, Google Trends, and the VIX appear among the leading variables in the SHAP ranking, suggesting that short-run Bitcoin movements are associated with both global risk appetite and attention-driven demand. At the weekly horizon, the ranking becomes more geographically and economically diversified: Stoxx 600, RUB, SSE, copper, and Google Trends dominate the SHAP-based interpretation. This suggests that weekly Bitcoin returns reflect a broader cross-market adjustment process involving equity markets, currencies, industrial commodities, and attention. At the biweekly horizon, the VIX and US Treasury variables become the two most influential predictors according to both XGBoost importance and SHAP values, while copper, Nikkei, and Google Trends remain relevant. This longer-horizon pattern is consistent with the interpretation that Bitcoin becomes more closely connected to risk, rates, and macro-financial conditions as daily noise is reduced.

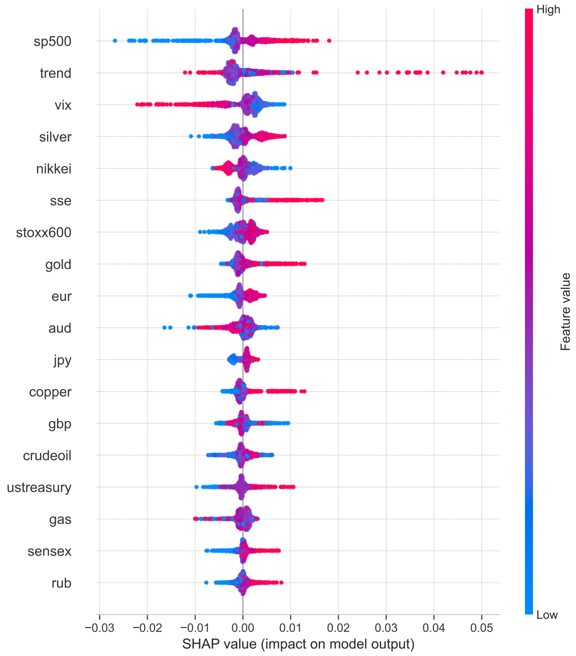
The SHAP summary plots provide additional information beyond the average rankings. They show not only which variables are important, but also how the sign and magnitude of their contributions vary across observations. This is particularly relevant in the present setting because Bitcoin's relationship with traditional financial variables is unlikely to be constant across market regimes.

To assess whether the interpretability evidence depends on a specific importance measure or on a particular subperiod, four robustness diagnostics are considered in Table 5. First, the importance rank correlation measures the average rank agreement across alternative importance diagnostics, including tree-based feature importance, SHAP-based importance, and permutation-based importance. Second, the top-5 overlap reports the average share of variables that appear simultaneously among the five most relevant predictors across alternative importance measures. Third, the temporal SHAP rank correlation evaluates whether the ranking of SHAP-based variable relevance remains stable across subperiods. Fourth, the dominant SHAP sign share measures whether the sign of each variable's SHAP contribution is stable over time. The results indicate that the leading predictors are reasonably robust, especially at the daily horizon. Temporal SHAP rank correlations remain high across all horizons, indicating that the relative importance of predictors is not driven by a single subperiod. However, the dominant SHAP sign share remains around 0.59–0.63, showing that the direction of individual feature contributions is not perfectly stable and that the strength of the relationships vary over time and should not be interpreted as fixed structural effects.

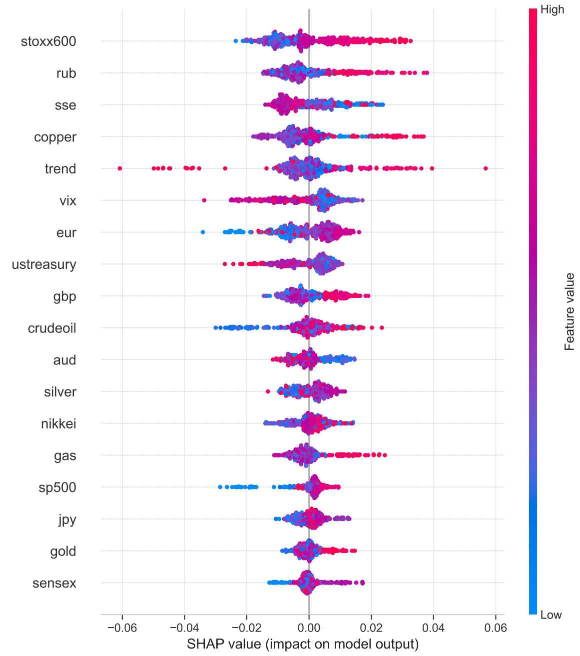
Table 5: Robustness diagnostics for interpretability results

Horizon	Importance rank corr.	Top-5 overlap	Temporal SHAP corr.	Dominant SHAP sign share
Daily	0.698	0.800	0.842	0.628
Weekly	0.613	0.600	0.877	0.592
Biweekly	0.635	0.600	0.801	0.624

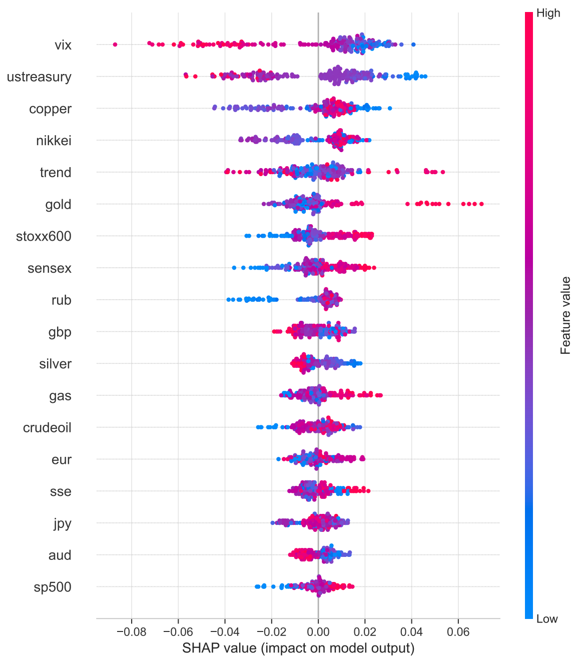
Finally, Figure 3 aggregates SHAP values by economic group. This aggregation is useful because individual predictor rankings may be affected by correlation among financial variables, whereas group-level interpretation better reflects the economic channels represented in the information set. Equity-market variables account for the largest share of SHAP importance at the daily horizon, followed by currencies, metals, attention, and market volatility. At the weekly horizon, currencies and equity markets become almost equally important, while metals, energy, attention, volatility, and interest rates contribute smaller but non-negligible shares. At the biweekly horizon, equity markets remain the largest group, but currencies, metals, market volatility, and interest rates jointly account for a substantial portion of the model explanation. The group-level evidence therefore reinforces the main message of the exploratory analysis: Bitcoin returns are linked to a broad cross-market information set whose relevance changes with the return horizon.



(a) Daily



(b) Weekly



(c) Biweekly

Figure 2: XGBoost SHAP summary plots by return horizon

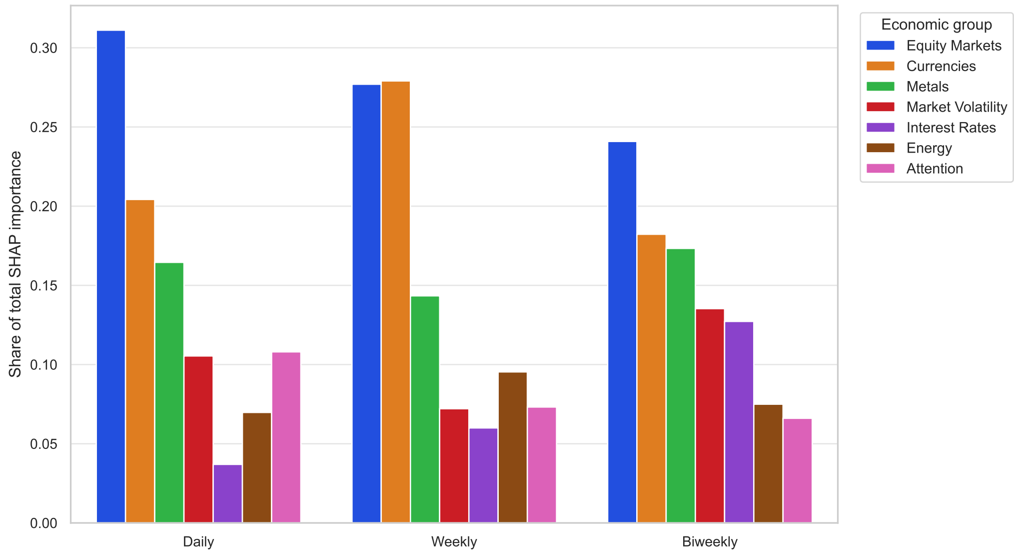


Figure 3: SHAP importance by economic group and return horizon

3 ML-Informed Black–Litterman Framework

3.1 Portfolio Dataset and Rolling Preprocessing

The portfolio component of the analysis transforms the cross-market information set examined in Section 2 into rolling probability signals that are subsequently incorporated into a Black–Litterman framework. While the exploratory analysis is designed to study the explanatory relationship between Bitcoin and traditional financial variables, the following allocation exercise requires a stricter timing structure. The empirical design is therefore based on a weekly strategy dataset in which all predictors are observed at time t, the Bitcoin directional target is defined at time $t+1$, and portfolio weights formed at time t are evaluated using next-period realized returns. This construction ensures that the information used to generate the allocation signal is available before the return realization used to assess portfolio performance.

The investable universe is intentionally restricted to three assets:

$$\mathcal{A} = \{BTC, S\&P500, Gold\}. \tag{8}$$

Bitcoin represents the high-volatility digital asset whose controlled integration is the object of the study, the S&P 500 represents the traditional risky equity component, and gold provides the defensive asset. The broader set of financial variables described in Section 2.1 is not treated as investable in the portfolio exercise. Instead, it defines the information set used by the machine-learning models to estimate the probability of a positive Bitcoin return.

For each weekly decision date t, the next-period Bitcoin return is denoted by r_{t+1}^{BTC}, and the directional classification target is defined as

$$y_{t+1} = \begin{cases} 1, & \text{if } r_{t+1}^{BTC} > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

The corresponding vector of next-period portfolio returns is

$$\mathbf{r}_{t+1} = \left(r_{t+1}^{BTC}, r_{t+1}^{SP500}, r_{t+1}^{Gold}\right)'. \tag{10}$$

The weekly dataset covers the period from January 2018 to February 2026 and initially contains 402 weekly observations The test period begins in January 2021, while the first valid walk-forward portfolio decision is obtained on 26 February 2021 because the initial rolling estimation window must be available before generating out-of-sample probabilities. The final walk-forward sample contains 246 weekly allocation decisions, ending on 30 January 2026. The resulting classification target is moderately balanced: positive Bitcoin weeks account for approximately 54.5% of the full strategy sample and 52.0% of the test sample.

The final predictor design is lag-augmented. Let $\mathbf{z}_t^{-BTC}$ denote the contemporaneous non-Bitcoin predictors observed at time t, and let $\mathbf{z}_{t-1}$ and $\mathbf{z}_{t-2}$ denote the first and second weekly lags of the available predictor set. The information vector used by the classifier can be written as

$$\mathbf{x}_t = \left[\mathbf{z}_t^{-BTC}, \mathbf{z}_{t-1}, \mathbf{z}_{t-2}\right]. \tag{11}$$

The contemporaneous Bitcoin return is excluded from $\mathbf{z}_t^{-BTC}$ to avoid using the target asset's current return directly in the same-period predictor block. However, lagged Bitcoin returns are retained through $\mathbf{z}_{t-1}$ and $\mathbf{z}_{t-2}$. This choice allows short-term Bitcoin momentum or reversal patterns to enter the model without violating the timing structure of the backtest. In the implementation, two weekly lags are added for the available variables. The final predictor matrix contains 56 predictors: 18 contemporaneous non-Bitcoin predictors and 38 lagged predictors, including two lagged Bitcoin-return variables.

The classification and portfolio construction are implemented in a rolling walk-forward setting. At each decision date t, the estimation sample is restricted to the most recent 156 weekly observations available before t:

$$\mathcal{T}_t = \{s < t : s \text{ belongs to the most recent 156 weekly observations}\}. \tag{12}$$

All preprocessing steps are estimated only on $\mathcal{T}_t$. In particular, each predictor is standardized using the mean and standard deviation computed from the rolling training window,

$$\widetilde{x}_{j,s}^{(t)} = \frac{x_{j,s} - \widehat{\mu}_{j,t}}{\widehat{\sigma}_{j,t}}, \qquad s \in \mathcal{T}_t, \tag{13}$$

where $\widehat{\mu}_{j,t}$ and $\widehat{\sigma}_{j,t}$ are estimated using only past observations. The same transformation is then applied to the test observation at date t. This prevents information from the evaluation period from entering the scaling step.

After standardization, Principal Component Analysis (PCA) is applied within each rolling training window. PCA is used because the lag-augmented predictor matrix is relatively high-dimensional and contains correlated financial variables. Rather than estimating the classifiers on the original predictor space, the standardized variables are projected onto orthogonal components that summarize the common variation in the information set (Jolliffe and Cadima, 2016). The PCA basis is fitted only on the rolling training sample and then used to transform the corresponding test observation. Formally, if $\widehat{\mathbf{V}}_t$ denotes the matrix of principal-component loadings estimated from $\mathcal{T}_t$, the transformed predictor vector is

$$\mathbf{p}_t = \widetilde{\mathbf{x}}_t \widehat{\mathbf{V}}_t. \tag{14}$$

The baseline specification retains the number of components required to explain 90% of the training-sample variance. With the lag-augmented predictor matrix, this typically requires approximately 30 components at the beginning of the walk-forward period and around 32–33 components toward the end of the sample.

3.2 Black–Litterman Allocation Framework

The allocation component is based on the Black–Litterman model, which provides a disciplined way to combine a strategic prior with additional information about expected returns. This choice is motivated by a well-known limitation of classical mean-variance optimization: optimal weights are highly sensitive to expected-return estimates, especially when these estimates are obtained from historical sample means (Markowitz, 1952; Best and Grauer, 1991). This problem is particularly relevant in the present setting because Bitcoin returns are volatile, non-Gaussian, and difficult to estimate precisely. The Black–Litterman framework mitigates this instability by shrinking portfolio inputs toward a prior return vector and by controlling the impact of investor views through an explicit view-uncertainty matrix (Black and Litterman, 1992; He and Litterman, 2002; Idzorek, 2007; Meucci, 2008). Here, the additional view is not imposed subjectively but it is generated from the machine-learning probability signal and is therefore treated as informative but noisy.

Let $\mathbf{r}_{t+1}$ denote the vector of next-period returns for the three investable assets,

$$\mathbf{r}_{t+1} = \left(r_{t+1}^{BTC}, r_{t+1}^{SP500}, r_{t+1}^{Gold}\right)', \tag{15}$$

and let Σ_t be the conditional covariance matrix estimated at time t using a rolling 52-week window. The Black–Litterman prior is obtained by reverse optimization:

$$\boldsymbol{\pi}_t = \delta \Sigma_t \mathbf{w}_m, \tag{16}$$

where $\boldsymbol{\pi}_t$ is the vector of implied prior excess returns, δ is the risk-aversion parameter, and $\mathbf{w}_m$ is the reference portfolio. The baseline specification sets $\delta = 3.0$ and the reference portfolio is defined as

$$\mathbf{w}_m = \begin{bmatrix} 0.05 \\ 0.75 \\ 0.20 \end{bmatrix}, \tag{17}$$

for the asset ordering $[BTC, S\&P500, Gold]$. This vector is not intended to reproduce market-capitalization weights. It is a stylized strategic allocation with limited Bitcoin exposure, dominant equity exposure, and a defensive gold component. The prior therefore provides a conservative starting point from which the machine-learning view can tilt the allocation without allowing the portfolio to be driven entirely by unstable return estimates.

Uncertainty around the prior is scaled by the scalar parameter τ, so that the prior covariance is proportional to $\tau\Sigma_t$. The baseline specification sets $\tau = 0.05$, implying that the prior is treated as relatively informative while still allowing the view to affect posterior expected returns. The investor view is represented through the standard Black–Litterman view equation,

$$P\boldsymbol{\mu}_t = Q_t + \eta_t, \qquad \eta_t \sim N(0, \Omega_t), \tag{18}$$

where $\boldsymbol{\mu}_t$ is the unknown expected-return vector, Q_t is the view value, and Ω_t is the view-uncertainty matrix. Since the framework imposes a single absolute view on Bitcoin, the view matrix is

$$P = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}. \tag{19}$$

Thus, Q_t is a scalar view on Bitcoin's expected return. The S&P 500 and gold do not receive direct views. Nevertheless, the posterior returns and final portfolio weights of all assets are affected through the covariance matrix and the optimization constraints.

Given the prior, the view, and their respective uncertainty matrices, the Black–Litterman posterior expected-return vector is

$$\boldsymbol{\mu}_t^{BL} = \left[(\tau\Sigma_t)^{-1} + P'\Omega_t^{-1}P\right]^{-1} \left[(\tau\Sigma_t)^{-1}\boldsymbol{\pi}_t + P'\Omega_t^{-1}Q_t\right]. \tag{20}$$

This expression shows that the posterior is a precision-weighted combination of the prior and the view. When the view uncertainty Ω_t is low, the ML-generated Bitcoin view has a stronger influence on the posterior. When Ω_t is high, posterior returns remain closer to the strategic prior. The posterior expected return therefore acts as a shrinkage estimator that regularizes the machine-learning signal through the equilibrium prior and the estimated covariance structure.

The posterior expected returns are then used in a constrained mean-variance allocation problem. At each decision date t (every weeek), portfolio weights are obtained by solving

$$\max_{\mathbf{w}_t} \quad \mathbf{w}_t'\boldsymbol{\mu}_t^{BL} - \frac{\lambda}{2}\mathbf{w}_t'\Sigma_t\mathbf{w}_t, \tag{21}$$

subject to

$$\mathbf{1}'\mathbf{w}_t = 1, \tag{22}$$

$$w_{i,t} \geq 0 \qquad \forall i \in \{BTC, S\&P500, Gold\}, \tag{23}$$

and

$$w_{BTC,t} \leq \bar{w}_{BTC}. \tag{24}$$

The parameter λ controls the risk-return trade-off in the optimization step. The no-short-selling and full-investment constraints make the portfolio comparable with implementable long-only asset-allocation strategies. The Bitcoin cap is a central part of the research design rather than a secondary technical restriction, since it prevents the optimizer from assigning excessive weight to Bitcoin when the posterior expected return is high and allows the analysis to study controlled Bitcoin integration. The tested caps are

$$\bar{w}_{BTC} \in \{0.02, 0.05, 0.10, 0.20\}. \tag{25}$$

The optimization is solved numerically at each decision date, and convergence is checked to avoid silent allocation failures.

3.3 ML-to-Black–Litterman Pipeline

The previous subsection introduced the theoretical foundations of the Black–Litterman model and showed that portfolio allocation ultimately depends on an expected-return view Q_t and its associated uncertainty Ω_t. The remaining methodological step consists of specifying how these quantities are generated from the machine-learning models. The proposed framework therefore transforms rolling probability estimates into Black–Litterman views through a sequential procedure that combines predictive modelling, probability-to-view mapping, Bayesian return updating, and constrained portfolio optimization. Unlike conventional quantitative trading strategies, the machine-learning models are not used to generate direct buy or sell decisions. Instead, they provide probabilistic information about the direction of future Bitcoin returns, which is subsequently incorporated into the Black–Litterman framework. This separation between prediction and portfolio construction allows predictive uncertainty to be explicitly quantified and propagated throughout the allocation process, reducing the sensitivity of the final portfolio to noisy return forecasts.

The first stage estimates the probability that Bitcoin will generate a positive return over the following week. Three alternative classifiers are considered: Logistic Regression, Random Forest, and XGBoost, all estimated using the rolling PCA-transformed predictors described above. At each portfolio formation date t, the models estimate

$$p_t = \Pr(y_{t+1} = 1 \mid \mathcal{F}_t), \tag{26}$$

where $\mathcal{F}_t$ denotes the information set available at time t. Rather than retaining the binary classification, the proposed framework exploits the estimated probability itself, since the Black–Litterman methodology requires a continuous measure of confidence that can be translated into an expected-return view rather than a deterministic trading signal.

The estimated probability is first converted into a centered directional signal,

$$s_t = p_t - 0.5, \tag{27}$$

whose sign identifies the expected market direction while its magnitude measures the strength of the prediction. Prediction confidence is then defined as

$$c_t = 2|p_t - 0.5|, \tag{28}$$

which ranges between zero, corresponding to maximum uncertainty, and one, corresponding to the highest confidence level. The directional signal is subsequently transformed into an absolute expected-return view,

$$Q_t = \bar{r}_t^{BTC} + \gamma s_t \sigma_t^{BTC}, \tag{29}$$

where $\bar{r}_t^{BTC}$ and σ_t^{BTC} denote the rolling mean and volatility of Bitcoin returns, respectively, while γ regulates the aggressiveness of the adjustment. Consequently, probabilities close to 0.5 generate views that remain close to the historical mean return, whereas increasingly confident predictions progressively shift the expected return upward or downward. Scaling the adjustment by rolling volatility expresses the view directly in return units while automatically adapting its magnitude to prevailing market conditions.

The influence of the machine-learning view on the posterior allocation is determined by the associated uncertainty matrix Ω_t. Three alternative specifications are considered. The baseline approach links uncertainty directly to the confidence implied by the estimated probability, assigning greater weight to highly confident predictions. Two additional specifications are examined as robustness checks: one based on the historical prediction error of the classifier and an Idzorek-style calibration that infers confidence from the implied portfolio allocation. During the initial observations, when the historical information required to estimate these measures is unavailable, missing uncertainty values are replaced using the rolling variance of Bitcoin returns. This initialization avoids assigning unrealistically precise views during the early stages of the backtest while preserving a coherent probabilistic interpretation.

The resulting Bitcoin view and its associated uncertainty are then combined with the equilibrium prior through the Black–Litterman updating equations introduced in Section 3.2. The posterior expected-return vector therefore incorporates both the long-run strategic allocation implied by the reference portfolio and the short-run information extracted from the machine-learning models. These posterior expected returns constitute the inputs of the constrained mean–variance optimization problem, which is solved sequentially at every portfolio formation date using the rolling covariance matrix and the Bitcoin exposure constraints described in the previous subsection. The resulting portfolio weights define a fully dynamic allocation strategy that continuously updates its asset allocation as new information becomes available.

Portfolio weights are obtained by solving the optimization problem presented in Section 3.2. At each decision date, the optimizer receives the posterior expected returns together with the rolling covariance matrix and determines the optimal long-only allocation subject to the Bitcoin exposure constraint. The optimization is repeated sequentially over the entire out-of-sample period, producing a fully dynamic asset-allocation strategy.

Portfolio performance is evaluated through a walk-forward backtesting procedure. At each decision date t, the optimized portfolio weights are applied to the realized asset returns observed during period $t+1$. Gross portfolio returns are adjusted for proportional transaction costs according to

$$r_t^{net} = r_t^{gross} - \kappa \cdot Turnover_t, \tag{30}$$

where $\kappa = 10$ basis points denotes the assumed transaction cost and

$$Turnover_t = \sum_{i=1}^{N} |w_{i,t} - w_{i,t-1}| \tag{31}$$

measures the portfolio rebalancing intensity between consecutive allocation dates. Incorporating turnover into the return calculation provides a more realistic assessment of the implementability of the proposed strategy by penalizing excessive portfolio rebalancing.

To isolate the contribution of the different components of the proposed methodology, the ML-informed Black–Litterman framework is compared with three benchmark families. The first benchmark consists of a traditional portfolio invested exclusively in the S&P 500 and gold, thereby measuring the incremental contribution of introducing Bitcoin into a conventional allocation. The second benchmark includes static portfolios with fixed Bitcoin allocations equal to 2%, 5%, 10%, and 20%, allowing the analysis to distinguish the value of dynamic allocation from the simple inclusion of the cryptocurrency. Finally, the third benchmark employs a classical rolling mean–variance optimization under the same Bitcoin exposure constraints (Markowitz, 1952; Best and Grauer, 1991). This comparison isolates the contribution of the Black–Litterman updating mechanism relative to a conventional optimization procedure based solely on rolling historical estimates.

The empirical comparison is conducted at two complementary levels. First, the probability models are evaluated using standard classification and probability-quality measures, including balanced accuracy, precision, recall, F_1 score, ROC-AUC, Brier score, log loss, and calibration diagnostics (Alpaydin, 2020). Second, portfolio performance is evaluated through a comprehensive set of return, risk, and implementation measures. Besides cumulative and annualized returns, the analysis considers annualized volatility together with the Sharpe ratio (Sharpe, 1966; Sharpe et al., 1998),

$$SR = \frac{\bar{r} - r_f}{\sigma}, \tag{32}$$

where $\bar{r}$ denotes the annualized portfolio return, r_f the annualized risk-free rate, and σ the annualized portfolio volatility. Downside performance is further assessed through the Sortino ratio (Sortino and Price, 1994), while drawdown characteristics are summarized using the maximum drawdown (Magdon-Ismail et al., 2004) and the Calmar ratio. Tail risk is evaluated through the 95% Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR), the latter measuring the expected loss beyond the VaR threshold (Rockafellar et al., 2000). Finally, portfolio turnover is reported to quantify trading intensity and implementation costs.

4 Results and Discussion

As discussed in Section 3.3, the objective of the machine-learning stage is to generate reliable probabilistic views that can be incorporated into the Black–Litterman allocation framework. Consequently, model selection is primarily based on probability quality rather than directional accuracy alone.

Table 6 reports the out-of-sample classification performance of the three competing probability models estimated on the rolling PCA-transformed predictors. The results indicate that predicting next-week Bitcoin direction remains a challenging task. Balanced accuracy is close to 0.50 for all specifications, while ROC–AUC values remain only moderately above the random benchmark. These findings are fully consistent with the growing literature documenting that medium-frequency Bitcoin returns exhibit only limited standalone predictability once realistic out-of-sample evaluation protocols are adopted (Sebastião and Godinho, 2021).

Table 6: Out-of-sample classification performance of the probability models.

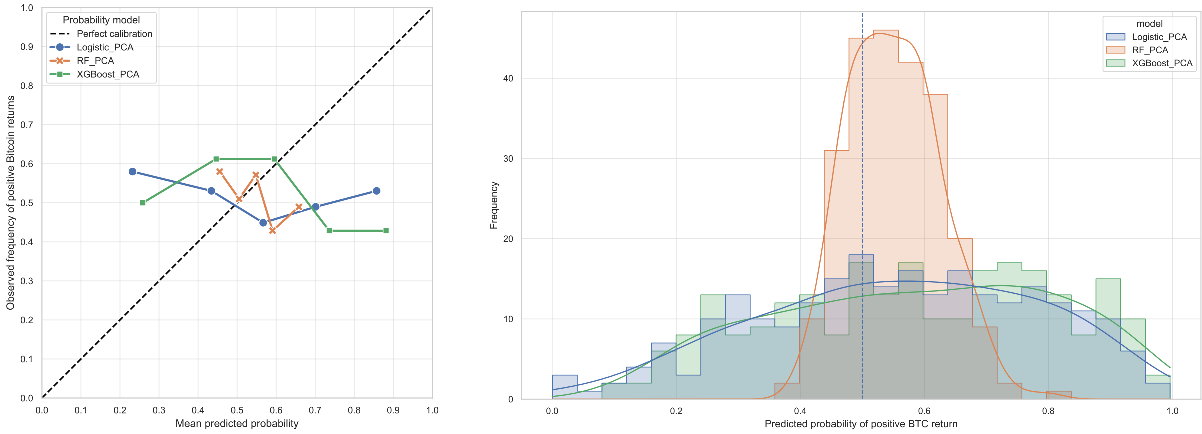
Model	Balanced accuracy	ROC–AUC	Brier score	Log-loss
Logistic_PCA	0.541	0.541	0.309	0.866
RF_PCA	0.520	0.550	**0.262**	**0.717**
XGBoost_PCA	0.496	0.475	0.325	0.905

Despite these modest classification results, clear differences emerge in the quality of the estimated probabilities. The Random Forest classifier achieves the lowest Brier score and log-loss, indicating better probabilistic calibration than other models. Logistic Regression produces slightly higher balanced accuracy, whereas XGBoost exhibits the weakest overall probabilistic performance. However, since the subsequent Black–Litterman framework exploits the complete probability distribution rather than binary class labels, it must be taken into account that probability quality represents the most relevant model-selection criterion. Consequently, Random Forest is adopted as the baseline probability engine throughout the portfolio analysis for constructing uncertainty-adjusted portfolio views.

Figure 4 provides complementary evidence on the quality of the estimated probabilities. Panel 4a reports the calibration curves of the three classifiers. Although none of the models achieves perfect calibration, reflecting the intrinsic difficulty of forecasting Bitcoin returns, the Random Forest classifier remains comparatively closer to the 45-degree reference line across most probability bins and exhibits the smallest overall calibration bias. Panel 4b reports the empirical distribution of the predicted probabilities, showing that Logistic Regression and XGBoost generate relatively dispersed forecasts, while Random Forest produces a more concentrated distribution centered around the decision threshold, indicating that the model avoids unwarranted overconfidence while still preserving discriminatory power.

Once Random Forest is selected, the central question becomes whether these probabilistic signals generate economically meaningful improvements if embedded within the proposed portfolio-allocation framework. Table 7 reports transaction cost adjusted performance over 246 weekly observations, comparing the performance of the proposed ML-informed Black–Litterman strategies with three benchmark families: a traditional equity–gold portfolio without Bitcoin exposure, static portfolios with fixed Bitcoin allocations, and classical rolling mean–variance optimization under identical allocation constraints. The risk-free rate used in the Sharpe and Sortino ratios is time-varying and based on the US Treasury series already discussed in Section 2.1.

Table 7 reveals a clear distinction between return maximization and portfolio efficiency. The



(a) Probability calibration curves.

(b) Distribution of predicted probabilities.

Figure 4: Probability quality of the competing classification models.

Table 7: Net performance of the ML-informed Black–Litterman strategies and benchmark portfolios

Family	Strategy	Cum. return	Ann. return	Ann. vol.	Sharpe	Sortino	Max DD	Calmar	Avg. turnover
Black–Litterman	BTC cap 2%	108.54%	16.81%	11.45%	**1.336**	**2.035**	-20.32%	**0.827**	0.000
	BTC cap 5%	106.06%	16.51%	11.88%	1.270	1.837	-22.55%	0.732	0.000
	BTC cap 10%	103.65%	16.22%	12.95%	1.157	1.587	-26.19%	0.619	0.013
	BTC cap 20%	84.03%	13.76%	12.37%	1.032	1.430	-24.37%	0.565	0.016
No Bitcoin	S&P 500–Gold	94.88%	15.15%	12.16%	1.147	1.735	-21.70%	0.698	0.000
Static Bitcoin	BTC 2%	93.68%	15.00%	12.36%	1.120	1.634	-22.96%	0.653	0.000
	BTC 5%	91.71%	14.75%	12.82%	1.067	1.489	-25.05%	0.589	0.000
	BTC 10%	87.98%	14.27%	13.96%	0.961	1.278	-28.73%	0.497	0.000
	BTC 20%	78.95%	13.09%	17.25%	0.748	0.993	-36.19%	0.362	0.000
Rolling MVO	BTC cap 2%	120.33%	18.17%	15.04%	1.126	1.749	-27.56%	0.659	0.124
	BTC cap 5%	119.06%	18.03%	15.03%	1.119	1.749	-27.86%	0.647	0.117
	BTC cap 10%	122.39%	18.41%	15.45%	1.113	1.772	-29.07%	0.633	0.123
	BTC cap 20%	**135.96%**	**19.90%**	16.90%	1.106	1.814	-31.42%	0.633	0.118

rolling mean–variance benchmark achieves the highest cumulative and annualized returns, but the high cost is a substantially higher volatility, deeper drawdowns, and considerably larger portfolio turnover. By contrast, the ML-informed Black–Litterman framework consistently delivers superior risk-adjusted performance under conservative Bitcoin allocation caps while maintaining remarkably stable allocations and minimal trading activity.

Several patterns emerge from the Black–Litterman strategies. Net annualized returns range between 13.8% and 16.8%, with annualized volatility remaining within a relatively narrow interval. More importantly, conservative Bitcoin allocations systematically dominate larger exposures across virtually all risk-adjusted performance measures. The 2% allocation simultaneously achieves the highest Sharpe, Sortino and Calmar ratio and exhibits the smallest maximum drawdown among the Black–Litterman specifications. Performance gradually deteriorates as the admissible Bitcoin allocation increases from 2% to 20%, indicating that the contribution of Bitcoin to portfolio efficiency is maximized under relatively small exposures. This result provides strong empirical support for the idea that Bitcoin should primarily be viewed as a complementary diversifier rather than as a dominant portfolio component (Petukhina et al., 2021; Wen et al., 2022).

The comparison with the traditional equity–gold benchmark further reinforces this conclusion. Introducing a limited Bitcoin allocation through the ML-informed Black–Litterman framework simultaneously improves cumulative return, annualized return, and all major risk-adjusted performance indicators without materially increasing downside risk. In particular, the 2% specification increases the Sharpe ratio and slightly reduces the maximum drawdown, supporting again the idea that a carefully controlled Bitcoin exposure can enhance portfolio efficiency without compromising portfolio resilience during adverse market conditions.

An equally important comparison concerns the static Bitcoin portfolios. Simply allocating a fixed fraction of the portfolio to Bitcoin already provides some diversification benefits, but ignores the time-varying information contained in the predictive variables. By contrast, the proposed framework continuously adjusts Bitcoin exposure according to the machine-learning probabilities after they have been regularized through the Black–Litterman updating mechanism. As a consequence, every dynamic specification outperforms its corresponding static allocation in terms of risk-adjusted performance and achieve a notable Sharpe-ratio improvement. These indicate that the economic value of the framework originates primarily from dynamically managing Bitcoin exposure rather than from simply introducing the cryptocurrency into the portfolio.

Lastly, the comparison with rolling mean–variance optimization highlights the role of the proposed framework. Although the MVO benchmark achieves higher realized cumulative and annualized returns, these gains are accompanied by substantially greater portfolio risk and considerably higher trading activity, which obviously implies greater transaction costs. In fact, average weekly turnover exceeds 11% across all MVO specifications, whereas turnover is essentially zero for the 2% and 5% Black–Litterman portfolios and remains below 2% even under the 20% allocation cap. Consequently, the proposed framework delivers higher Sharpe ratios for the 2%, 5%, and 10% allocation constraints despite producing lower absolute returns, demonstrating a positive effect on the efficiency and stability of the optimization process relative to classical mean–variance allocation.

The cumulative performance trajectories reported in Figure 5 further illustrate these differences. All strategies naturally experience the major market-wide shocks affecting both traditional financial assets and cryptocurrencies over the sample period that includes the aftermath of the COVID-19 pandemic and the macro-financial disruptions associated with the Russia–Ukraine conflict (Conlon and McGee, 2020; Goodell et al., 2023). Nevertheless, the ML-informed Black–

Litterman portfolios consistently outperform both the static Bitcoin allocations and the traditional equity–gold benchmark over most of the evaluation horizon.

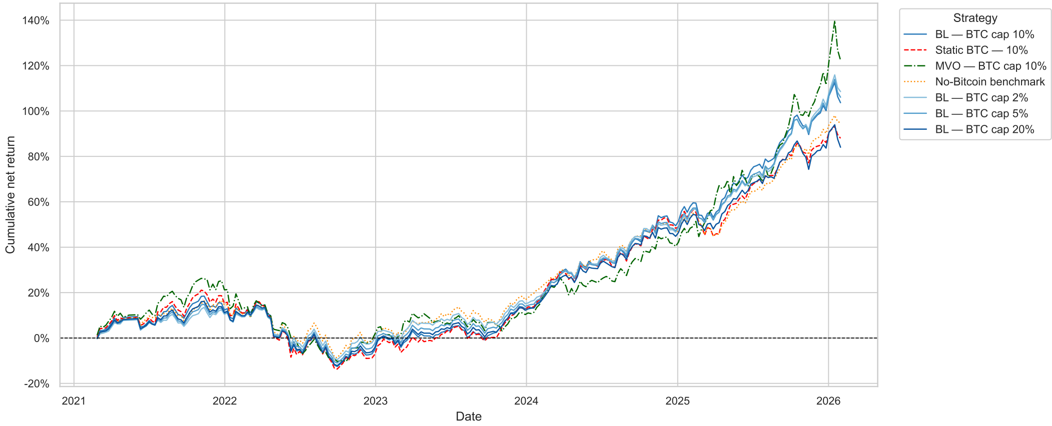


Figure 5: Cumulative net returns of the baseline ML-informed Black–Litterman strategies and benchmark portfolios.

The allocation dynamics reported in Figure 6 provide further insight into the mechanism generating these results. The 2% strategy remains almost permanently at its maximum admissible allocation, indicating that a small Bitcoin exposure consistently contributes to portfolio efficiency. The 5% specification displays a similar pattern, with only modest fluctuations around the allocation limit. In contrast, the 10% and especially the 20% portfolios exhibit substantially greater variation, as the optimizer actively reduces Bitcoin exposure whenever the estimated views become less favorable or the associated portfolio risk increases. Rather than mechanically exploiting the largest admissible allocation, the Black–Litterman framework dynamically adjusts exposure according to both expected returns and risk considerations. This behavior explains why increasing the Bitcoin allocation limit does not necessarily improve portfolio performance: the optimization itself recognizes that excessive cryptocurrency exposure often deteriorates the overall risk–return trade-off.

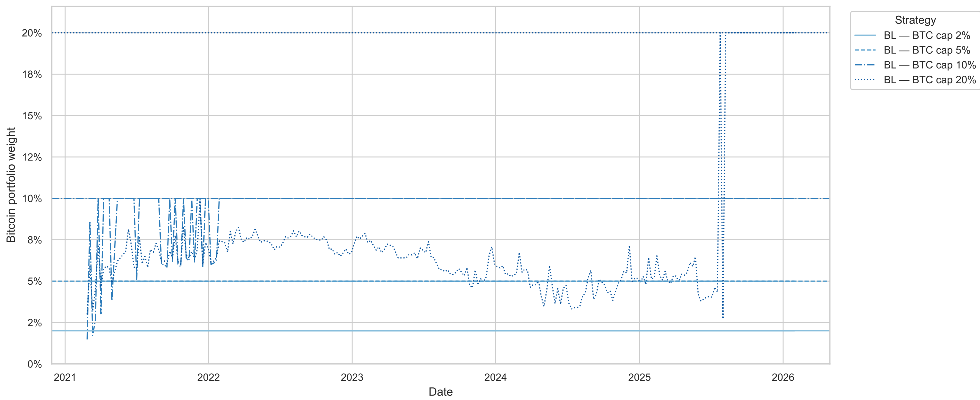


Figure 6: Evolution of Bitcoin portfolio weights across the four allocation constraints.

The analyses presented so far focused on the baseline specification of the proposed framework. Although these results already suggest that the ML-informed Black–Litterman strategy provides an effective mechanism for integrating Bitcoin into a traditional portfolio, they remain conditional on a particular set of modelling choices. To assess the sensitivity of the conclusions, an extensive robustness analysis is performed, spanning all the principal dimensions of the framework.

Specifically, the sensitivity analysis considers three probability models (Logistic Regression, Random Forest, and XGBoost), three view-uncertainty specifications (probability-confidence, prediction-error, and Idzorek-style uncertainty), three probability-to-view mappings (linear, directional, and hyperbolic tangent), three values of the view-aggressiveness parameter γ, three covariance estimation windows, three levels of investor risk aversion, and four different Bitcoin allocation caps. In total, 2,916 alternative portfolio specifications are evaluated in order to verify whether the principal empirical conclusions remain stable across a broad set of plausible methodological choices. Consequently, the discussion focuses primarily on average performance across specifications rather than on isolated best-performing configurations. Table 8 presents a summary of the whole exercise.

Table 8: Summary of the robustness analysis across 2,916 alternative specifications.

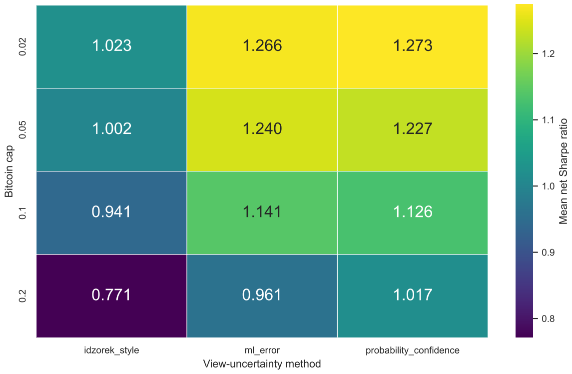
Dimension	Specification	Mean Sharpe	Median Sharpe	Best Sharpe	Mean Volatility	Mean Max DD
BTC cap	2%	**1.188**	1.203	1.417	11.63%	-21.8%
	5%	1.156	1.173	1.385	11.89%	-22.4%
	10%	1.069	1.087	1.331	12.64%	-24.9%
	20%	0.916	0.934	1.239	14.26%	-28.8%
Probability model	RF_PCA	**1.139**	1.160	1.385	12.39%	-23.9%
	Logistic_PCA	1.055	1.071	1.358	12.61%	-24.5%
	XGBoost_PCA	1.054	1.065	**1.417**	12.71%	-24.6%
View uncertainty	Probability confidence	**1.161**	1.181	1.385	12.27%	-23.6%
	ML prediction error	1.152	1.169	1.417	12.34%	-23.8%
	Idzorek	0.934	0.951	1.302	13.10%	-26.1%
View mapping	Linear	**1.123**	1.140	1.417	12.36%	-24.0%
	Directional	1.075	1.094	1.356	12.59%	-24.4%
	Tanh	1.049	1.064	1.334	12.77%	-24.9%
γ	0.5	**1.124**	1.145	1.417	12.31%	-23.8%
	1.0	1.077	1.094	1.385	12.57%	-24.5%
	1.5	1.047	1.060	1.352	12.83%	-25.0%
Risk aversion	2.0	**1.117**	1.133	1.417	12.34%	-23.9%
	3.0	1.100	1.118	1.385	12.48%	-24.2%
	5.0	1.030	1.048	1.328	12.89%	-25.1%

The robustness analysis reveals that the Bitcoin allocation cap is the primary driver of portfolio performance. Smaller allocation caps consistently deliver higher risk-adjusted performance, lower volatility, and shallower drawdowns than more aggressive Bitcoin exposures. This pattern remains remarkably stable regardless of the underlying machine-learning model or Black–Litterman parameterization, reinforcing the central conclusion that Bitcoin contributes most effectively as a complementary portfolio asset under disciplined exposure limits.

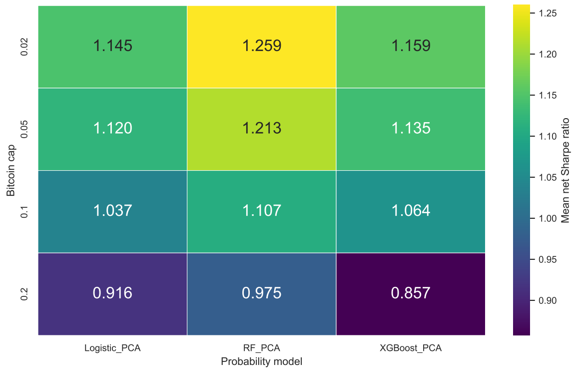
The remaining modelling choices exert a comparatively smaller influence on performance. Although the probability models exhibit some differences, their average results remain broadly comparable, with only isolated configurations substantially outperforming the others. Likewise, the probability-confidence uncertainty specification and the linear probability-to-view mapping generally provide the strongest results, while more conservative values of the view-aggressiveness parameter (γ) tend to outperform more aggressive alternatives. Differences across investor risk-aversion levels are relatively modest, indicating that the proposed framework remains stable under a wide range of reasonable portfolio preferences.

Hence, the dominant and most persistent result is that controlled Bitcoin exposure, combined with uncertainty-adjusted Black–Litterman views, consistently produces the most attractive risk-adjusted portfolio performance.

Figure 7 complements the numerical evidence reported in Table 8. Panel 7a shows that the superiority of the 2% and 5% Bitcoin allocation caps is remarkably stable across all view-uncertainty specifications, while Panel 7b demonstrates that the relative ranking of the probability models is also highly stable.



(a) Bitcoin allocation cap vs. view uncertainty.



(b) Probability model vs. view uncertainty.

Figure 7: Robustness heatmaps. Cell values report the mean transaction-cost-adjusted Sharpe ratio across the remaining parameter combinations.

The broader distribution of the robustness configurations is illustrated in Figures 8 and 9. The risk–return map shows that the highest-performing specifications are concentrated in the upper-left region of the performance space, combining relatively high annualized returns with comparatively low volatility. Likewise, the Sharpe–drawdown map reveals that these configurations are also associated with moderate drawdowns rather than with excessive downside risk. Hence, the proposed framework improves portfolio quality primarily through superior risk management rather than by pursuing increasingly aggressive return targets.

Figure 8: Risk–return distribution of the 2,916 robustness configurations.

The robustness exercise also helps explain the apparent discrepancy between the relatively modest predictive performance of the machine-learning classifiers and the strong portfolio results documented in the previous sections. As shown in Table 6, the probability models achieve only moderate directional accuracy, reflecting the inherent difficulty of forecasting medium-frequency Bitcoin returns. However, the proposed framework does not rely on binary class predictions, but exploits the full probability distribution to construct expected-return views and explicitly accounts for the associated uncertainty. Consequently, even modest but well-calibrated probability estimates can become economically valuable once combined with covariance information and Bayesian updating. The portfolio improvements therefore arise from the interaction between predictive information and the allocation framework, rather than from

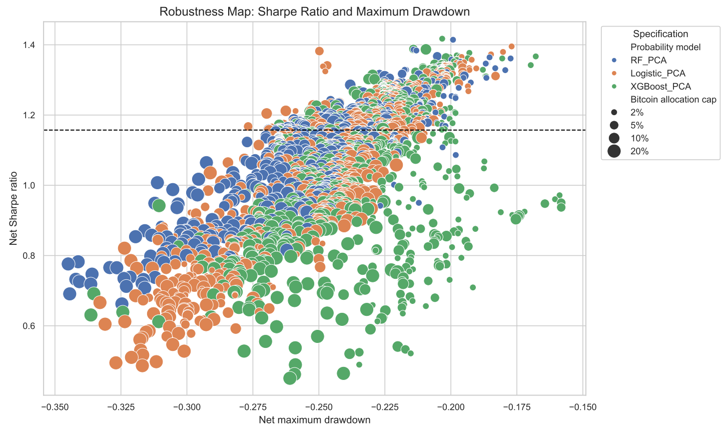


Figure 9: Sharpe ratio versus maximum drawdown across the robustness configurations.

strong standalone forecasting accuracy.

5 Conclusion

This study examined whether Bitcoin can be integrated into a traditional portfolio through a disciplined allocation mechanism that does not depend on strong standalone return predictability. The underlying problem is inherently different from identifying a profitable Bitcoin trading rule. In fact, the relevant question for portfolio managers is whether imperfect and time-varying information about Bitcoin can be converted into allocation decisions without allowing a volatile digital asset to destabilize the broader portfolio. To address this issue, this work combines an exploratory analysis of Bitcoin's cross-market relationships with a machine-learning-informed Black–Litterman framework for portfolio construction.

Building on this theoretical framework, the study formulates two related research questions. The first examines whether machine-learning-generated probabilistic views, when incorporated through a Bayesian allocation framework, can improve risk-adjusted performance relative to passive and conventional active benchmarks. The second investigates whether the portfolio benefits associated with Bitcoin diminish as the maximum admissible exposure increases. To address these questions, the framework treats directional probabilities as noisy beliefs rather than as direct buy or sell signals. The probabilities are converted into volatility-scaled views and incorporated into the Black–Litterman posterior according to their estimated uncertainty, while the strategic prior, covariance structure, and portfolio constraints regulate their effect on the final allocation. The methodological contribution therefore lies in separating signal formation from portfolio construction and in jointly controlling how predictive information affects expected returns and actual portfolio exposure.

With respect to RQ1, the ML-informed Black–Litterman portfolios achieve higher risk-adjusted performance than the equity–gold benchmark and the corresponding static Bitcoin allocations. Relative to rolling mean–variance optimization, the proposed framework produces lower absolute returns but generally combines lower volatility, shallower drawdowns, and substantially lower turnover with stronger risk-adjusted performance under conservative and moderate Bitcoin caps. Its economic value should therefore be interpreted primarily in terms of allocation stability, downside-risk control, and implementability rather than return maximization. Moreover, because the smallest-cap portfolios remain close to their exposure limits, the observed benefits cannot be attributed exclusively to frequent dynamic adjustments; they reflect the combined contribution of Bayesian regularization, covariance-based allocation, and disciplined exposure constraints.

Regarding RQ2, the results indicate that the risk-adjusted benefits associated with Bitcoin diminish as the maximum admissible exposure increases. Small allocation limits provide the most favorable balance between return and risk, whereas progressively larger caps increase volatility and drawdown exposure and weaken overall portfolio efficiency. This pattern remains stable across the alternative probability models, view mappings, uncertainty specifications, covariance windows, and risk-aversion parameters considered in the robustness analysis. The evidence therefore supports the interpretation of Bitcoin as a complementary portfolio asset whose contribution depends critically on controlled exposure. In this setting, the allocation cap is an economically meaningful regularization mechanism that limits the portfolio consequences of uncertain expected-return beliefs.

From a ML-related literature perspective, this study contributes by illustrating how machine-learning outputs can be incorporated directly into portfolio construction rather than evaluated solely as standalone forecasts. In particular, treating predicted probabilities as uncertain Black–Litterman views provides a practical link between predictive modelling and constrained asset allocation. This approach may offer a useful reference for future research examining how even modest predictive signals can be translated into economically interpretable portfolio decisions.

The analysis is subject to several limitations that also identify directions for future research. The investable universe is intentionally restricted to Bitcoin, the S&P 500, and gold, and the conclusions therefore refer to a stylized equity–gold portfolio rather than to a fully diversified institutional allocation. The weekly strategy is evaluated over a relatively limited out-of-sample period and relies on historical covariance estimates, PCA-based dimensionality reduction, one absolute view on Bitcoin, and proportional transaction costs. Future work could extend the framework to larger multi-asset portfolios, multiple cryptocurrencies, simultaneous relative and absolute views, and alternative covariance estimators based on shrinkage, factor structures, or dynamic volatility models. Additional research could also incorporate on-chain indicators, liquidity measures, macroeconomic announcements, alternative probability-calibration procedures, and time-varying transaction costs. These extensions would clarify whether the central conclusion generalizes beyond the present setting: machine-learning information is most valuable when it is used to support, rather than replace, disciplined and economically grounded portfolio construction.

References

Adelopo, I. and Luo, X. (2025). Interconnectedness among cryptocurrencies and financial markets: A systematic literature review. *Digital Finance*, 7(4):1119–1171.

Akyildirim, E., Goncu, A., and Sensoy, A. (2021). Prediction of cryptocurrency returns using machine learning. *Annals of Operations Research*, 297(1):3–36.

Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.

Aslanidis, N., Bariviera, A. F., and López, Ó. G. (2022). The link between cryptocurrencies and google trends attention. *Finance Research Letters*, 47:102654.

Barro, R. J. and Misra, S. (2016). Gold returns. *The Economic Journal*, 126(594):1293–1317.

Baur, D. G., Dimpfl, T., and Kuck, K. (2018). Bitcoin, gold and the us dollar–a replication and extension. *Finance research letters*, 25:103–110.

Bellocca, G. P., Attanasio, G., Cagliero, L., and Fior, J. (2022). Leveraging the momentum effect in machine learning-based cryptocurrency trading. *Machine Learning with Applications*, 8:100310.

Best, M. J. and Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The review of financial studies*, 4(2):315–342.

Biais, B., Bisiere, C., Bouvard, M., and Casamatta, C. (2019). The blockchain folk theorem. *The Review of Financial Studies*, 32(5):1662–1715.

Black, F. and Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43.

Böhme, R., Christin, N., Edelman, B., and Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of economic Perspectives*, 29(2):213–238.

Bouri, E., Gupta, R., Tiwari, A. K., and Roubaud, D. (2017). Does bitcoin hedge global uncertainty? evidence from wavelet-based quantile-in-quantile regressions. *Finance Research Letters*, 23:87–95.

Bouri, E., Saeed, T., Vo, X. V., and Roubaud, D. (2021). Quantile connectedness in the cryptocurrency market. *Journal of International Financial Markets, Institutions and Money*, 71:101302.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Brière, M., Oosterlinck, K., and Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with bitcoin. *Journal of Asset Management*, 16(6):365–373.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.

Chiu, J. and Koeppl, T. V. (2022). The economics of cryptocurrency: Bitcoin and beyond. *Canadian Journal of Economics/Revue canadienne d'économique*, 55(4):1762–1798.

Ciaian, P., Rajcaniova, M., and Kancs, d. (2016). The economics of bitcoin price formation. *Applied economics*, 48(19):1799–1815.

Cong, L. W., Li, Y., and Wang, N. (2021). Tokenomics: Dynamic adoption and valuation. *The Review of Financial Studies*, 34(3):1105–1155.

Conlon, T. and McGee, R. (2020). Safe haven or risky hazard? bitcoin during the covid-19 bear market. *Finance Research Letters*, 35:101607.

Corbet, S., Lucey, B., and Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, 62:182–199.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

Dehouche, N. (2021). Scale matters: The daily, weekly and monthly volatility and predictability of bitcoin, gold, and the s&p 500. *arXiv preprint arXiv:2103.00395*.

Dyhrberg, A. H. (2016a). Bitcoin, gold and the dollar – a garch volatility analysis. *Finance Research Letters*, 16:85–92.

Dyhrberg, A. H. (2016b). Hedging capabilities of bitcoin. is it the virtual gold? *Finance Research Letters*, 16:139–144.

Fang, F., Ventre, C., Basios, M., Kanthan, L., Martinez-Rego, D., Wu, F., and Li, L. (2025). Cryptocurrency trading: a comprehensive survey. *Blockchain, Crypto Assets, and Financial Innovation: A Decade of Insights and Advances*, pages 55–127.

Fry, J. and Cheah, E.-T. (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, 47:343–352.

Goodell, J. W., Yadav, M. P., Ruan, J., Abedin, M. Z., and Malhotra, N. (2023). Traditional assets, digital assets and renewable energy: Investigating connectedness during covid-19 and the russia-ukraine war. *Finance Research Letters*, 58:104323.

Gorman, M. and Hughen, W. K. (2024). Does bitcoin still enhance an investment portfolio in a post covid-19 world? *Finance Research Letters*, 62:105170.

Grobys, K., Ahmed, S., and Sapkota, N. (2020). Technical trading rules in the cryptocurrency market. *Finance Research Letters*, 32:101396.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273.

Guo, W., Intini, S., and Jahanshahloo, H. (2025). Bitcoin arbitrage and exchange default risk. *Finance Research Letters*, 71:106364.

Hastie, T. (2009). *The elements of statistical learning: data mining, inference, and prediction*. springer.

Hayes, A. S. (2017). Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telematics and informatics*, 34(7):1308–1321.

Hayes, A. S. (2019). Bitcoin price and its marginal cost of production: support for a fundamental value. *Applied economics letters*, 26(7):554–560.

He, G. and Litterman, R. (2002). The intuition behind black-litterman model portfolios. *Available at SSRN 334304*.

Idzorek, T. (2007). A step-by-step guide to the black-litterman model: Incorporating user-specified confidence levels. In Satchell, S., editor, *Forecasting Expected Returns in the Financial Markets*, pages 17–38. Academic Press, Oxford.

James, G., Witten, D., Hastie, T., Tibshirani, R., et al. (2013). *An introduction to statistical learning: with applications in R*, volume 103. Springer.

Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202.

Khedr, A. M., Arif, I., El-Bannany, M., Alhashmi, S. M., and Sreedharan, M. (2021). Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey. *Intelligent Systems in Accounting, Finance and Management*, 28(1):3–34.

Kyriazis, N., Papadamou, S., and Corbet, S. (2020). A systematic review of the bubble dynamics of cryptocurrency prices. *Research in International Business and Finance*, 54:101254.

Liu, P. and Yuan, Y. (2024). Is bitcoin a hedge or safe-haven asset during the period of turmoil? evidence from the currency, bond and stock markets. *International Review of Financial Analysis*, 96:103663.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.

Magdon-Ismail, M., Atiya, A. F., Pratap, A., and Abu-Mostafa, Y. S. (2004). On the maximum drawdown of a brownian motion. *Journal of applied probability*, 41(1):147–161.

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

Meucci, A. (2008). The black-litterman approach: Original model and extensions. *SSRN Electronic Journal*.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized business review*, page 21260.

Noda, A. (2021). On the evolution of cryptocurrency market efficiency. *Applied Economics Letters*, 28(6):433–439.

Petukhina, A., Trimborn, S., Härdle, W. K., and Elendner, H. (2021). Investing with cryptocurrencies–evaluating their potential for portfolio allocation strategies. *Quantitative Finance*, 21(11):1825–1853.

Rockafellar, R. T., Uryasev, S., et al. (2000). Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42.

Sebastião, H. and Godinho, P. (2021). Forecasting and trading cryptocurrencies with machine learning under changing market conditions. *Financial Innovation*, 7(1):3.

Selmi, R., Mensi, W., Hammoudeh, S., and Bouoiyour, J. (2018). Is bitcoin a hedge, a safe haven or a diversifier for oil price movements? a comparison with gold. *Energy Economics*, 74:787–801.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, 39(1):119–138.

Sharpe, W. F. et al. (1998). The sharpe ratio. *Streetwise–the Best of the Journal of Portfolio Management*, 3(3):169–85.

Smales, L. A. (2019). Bitcoin as a safe haven: Is it even worth considering? *Finance Research Letters*, 30:385–393.

Sortino, F. A. and Price, L. N. (1994). Performance measurement in a downside risk framework. *the Journal of Investing*, 3(3):59–64.

Subramaniam, S. and Chakraborty, M. (2020). Investor attention and cryptocurrency returns: Evidence from quantile causality approach. *Journal of Behavioral Finance*, 21(1):103–115.

Svogun, D. and Bazán-Palomino, W. (2022). Technical analysis in cryptocurrency markets: Do transaction costs and bubbles matter? *Journal of International Financial Markets, Institutions and Money*, 79:101601.

Valencia, F., Gómez-Espinosa, A., and Valdés-Aguirre, B. (2019). Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *entropy*, 21(6):589.

Vapnik, V. (2013). *The nature of statistical learning theory.* Springer science & business media.

Wen, F., Tong, X., and Ren, X. (2022). Gold or bitcoin, which is the safe haven during the covid-19 pandemic? *International Review of Financial Analysis*, 81:102121.

Yao, S., Sensoy, A., Nguyen, D. K., and Li, T. (2024). Investor attention and cryptocurrency market liquidity: a double-edged sword. *Annals of Operations Research*, 334(1):815–856.

Zhang, Y.-J., Bouri, E., Gupta, R., and Ma, S.-J. (2021). Risk spillover between bitcoin and conventional financial markets: An expectile-based approach. *The North American Journal of Economics and Finance*, 55:101296.

Zhu, P., Zhang, X., Wu, Y., Zheng, H., and Zhang, Y. (2021). Investor attention and cryptocurrency: Evidence from the bitcoin market. *PLoS One*, 16(2):e0246331.