Time-Varying Factor-Augmented Models for Volatility Forecasting

Duo Zhang
dz2349@nyu.edu
New York University
New York, New York, USA

Jiayu Li
jl15681@stern.nyu.edu
New York University
New York, New York, USA

Junyi Mo
junyi.mo@stern.nyu.edu
New York University
New York, New York, USA

Elynn Chen*
elynn.chen@stern.nyu.edu
New York University
New York, New York, USA

Abstract

Accurate volatility forecasts are vital in modern finance for risk management, portfolio allocation, and strategic decision-making. However, existing methods face key limitations. Fully multivariate models, while comprehensive, are computationally infeasible for realistic portfolios. Factor models, though efficient, primarily use static factor loadings, failing to capture evolving volatility co-movements when they are most critical. To address these limitations, we propose a novel, model-agnostic *Factor-Augmented Volatility Forecast framework*. Our approach employs a time-varying factor model to extract a compact set of dynamic, cross-sectional factors from realized volatilities with minimal computational cost. These factors are then integrated into both statistical and AI-based forecasting models, enabling a unified system that jointly models asset-specific dynamics and evolving market-wide co-movements. Our framework demonstrates strong performance across two prominent asset classes—large-cap U.S. technology equities and major cryptocurrencies—over both short-term (1-day) and medium-term (7-day) horizons. Using a suite of linear and non-linear AI-driven models, we consistently observe substantial improvements in predictive accuracy and economic value. Notably, a practical pairs-trading strategy built on our forecasts delivers superior risk-adjusted returns and profitability, particularly under adverse market conditions.

CCS Concepts

• **Applied computing** → **Operations research**; • **Computing methodologies** → **Machine learning**.

Keywords

Volatility Forecasting, Factor Models, Long Short-Term Memory (LSTM), HAR, MIDAS, Volatility-Based Pairs Trading

---

*Corresponding author

1 Introduction

From tick-by-tick hedging to long-term investment strategy, reliable volatility forecasts are indispensable for guiding financial decisions. A vast literature—ranging from linear ARCH [24], GARCH [6], and HAR [19] to nonlinear neural networks such as LSTMs [35]—is typically applied on an asset-by-asset basis, modeling each asset in isolation. This univariate approach overlooks market-wide shocks that propagate rapidly across assets [33], often biasing forecasts and understating systemic risk. Although multivariate extensions such as BEKK-GARCH [25], DCC-GARCH [23], Conditional Autoregressive Wishart [31], and deep multivariate LSTM architectures capture shared dynamics, they are rarely used in real-world settings due to parameter explosion, fragile estimation, and heavy tuning burdens [5, 26].

Factor models offer a natural remedy by efficiently summarizing co-movements through a low-dimensional set of latent drivers. Yet, two key limitations hinder their utility in volatility forecasting. First, most factor models are designed for returns, not volatility. Canonical linear frameworks such as Fama-French and Carhart [9, 27], and their modern machine-learning successors [32], all distill factors from returns data. Consequently, the rich patterns of commonality in volatility—which often exhibit their own distinct dynamics—remain largely unexplored. Second, although a few recent pioneering studies do extend factor models to realized volatility [3, 22], they still impose static factor loadings, which fail to capture time-varying interdependencies—an especially critical flaw given that volatility co-movement evolves over time [20]. Although recent econometric advances do provide time-varying loadings estimators [41], these techniques have so far been confined exclusively to asset returns.

To bridge these gaps, we propose a model-agnostic *Factor-Augmented Volatility Forecast framework* that captures evolving market-wide volatility shocks with computational tractability. Unlike traditional factor approaches with static loadings, our method employs a time-varying factor model to extract latent factors directly from realized volatilities. This dynamic approach ensures that both the factors and their loadings adapt in real time to shifting

market regimes. Compared to conventional multivariate techniques, our framework achieves this enhanced market awareness while preserving the parsimony of the base model, furthermore, introducing only a minimal number of extra parameters. The result is a robust, accurate, and interpretable framework that delivers computationally efficient volatility forecasts, suitable for practical deployment in dynamic financial environments.

We evaluate the framework on two representative markets—U.S. technology equities and cryptocurrencies—for both 1-day and 7-day forecast horizons. Across all model-market-horizon combinations, factor augmentation consistently improves performance. Out-of-sample coefficient of determination (R^2) increases by up to 12.8% for equities and 22.8% for cryptocurrencies, with corresponding reductions in Mean Squared Error (MSE) and Quasi-likelihood loss (QLIKE). Furthermore, a volatility-scaled pairs-trading backtest confirms the framework's tangible economic value. Most strikingly, in a challenging market period, augmentation reversed an unprofitable strategy from an annualized loss of -5.5% into a +7.3% gain, flipping its Sharpe ratio from negative to positive. In brief, our key contributions are as follows:

- We introduce the first volatility forecasting framework that systematically integrates dynamic, cross-sectional information through factor augmentation. Our approach combines three unique innovations: (i) direct factor extraction from realized volatilities, distilling interpretable factors that capture volatility-specific commonalities; (ii) time-varying loadings that adapt to evolving market interdependencies; and (iii) a computationally efficient augmentation mechanism that ensures both robustness and scalability.

- Our framework achives superior statistical accuracy across all tested scenarios, and a practical pairs-trading backtest confirms that these accuracy gains translate directly into higher risk-adjusted returns, generating real-world economic value.

- We provide a practical, end-to-end guide to factor extraction, selection, and integration, offering a clear roadmap for tailoring the framework to any forecasting model, horizon, or asset class.

2 Related Work

Research in discrete-time volatility forecasting has evolved significantly. However, existing models present common trade-offs, which our framework is designed to address.

Univariate Models. Frameworks from early AR models to more sophisticated ARCH [24], GARCH [6, 7], and HAR [19] have long been pillars of the field due to their interpretability. Extensions have incorporated features like mixed-frequency data [18, 44] and long-memory processes [4]. However, by modeling each asset in isolation, these models suffer from a "univariate blind spot," ignoring the systemic co-movements that are crucial drivers of market-wide risk [33, 34].

Multivariate Models. Multivariate models are designed to capture asset interplay directly. Within the GARCH family, the BEKK specification [25] guarantees positive-definite covariance matrices, while the Dynamic Conditional Correlation (DCC) model [23] offers a more feasible estimation process. AR-style alternatives

like the Conditional Autoregressive Wishart (CAW) [31] provide further flexibility. While theoretically comprehensive, these models are often impractical for large systems due to the explosion of parameter count [5] and numerically unstable estimation [26].

Machine Learning Models. A parallel track has leveraged machine learning to capture the complex, nonlinear volatility patterns. Deep learning models like LSTM [35] have shown superior performance over classical models by modeling long-range dependencies in financial time series [39, 40]. However, these models are not immune to the "curse of dimensionality." When applied in a fully multivariate setting, they face excessive parameterization, high overfitting risk, and prohibitive computational costs [29], inheriting the same fundamental limitations as their statistical counterparts.

Factor Models. Factor methods are the backbone of modern portfolio theory, distilling hundreds of correlated returns into a handful of common drivers and powering scalable asset allocation, hedging, and attribution. Widely used examples range from statistical factor-GARCH variants such as OGARCH [1] and GO-GARCH [45] to recent factor-structured deep networks [28]. While these models achieve parsimony by using latent factors, most of the existing work targets returns [10, 12–17, 32, 37, 38]; factor models for volatility remain comparatively sparse and always assume time-invariant loadings [3, 22].

Our research directly confronts these trade-offs by proposing a factor augmentation framework. This approach enhances existing volatility models with a set of dynamic, cross-sectional factors, enabling the framework to capture systemic, time-varying market interdependence without adding additional computational overhead, thereby balancing predictive accuracy with practical feasibility.

3 Data and Volatility Measurement

3.1 Data

The empirical analysis is conducted on two high-frequency datasets, encompassing U.S. equities and major cryptocurrencies.

The equity dataset focuses on five prominent firms within the S&P 500 Information Technology Sector, selected for their large market capitalization and historical sectoral comovements: Microsoft (MSFT), Advanced Micro Devices (AMD), Intel (INTC), Oracle (ORCL), and Cisco (CSCO). For these firms, we obtained high-frequency Consolidated Trade and Quote records from Wharton Research Data Service, covering the period from January 2018 to December 2020 at millisecond precision.

The cryptocurrency dataset, sourced from Kaiko, covers January 2018 to April 2021. This dataset comprises detailed, high-frequency trading data for five leading cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Cardano (ADA), and Litecoin (LTC). Kaiko compiles transaction-level records from major digital asset exchanges, providing comprehensive data including timestamps, trade volumes, and direction.

3.2 Volatility Measurement

3.2.1 Daily Realized Volatility. Following [2], we estimate daily realized volatility (RV) nonparametrically from high-frequency midpoint prices. After filtering out non-positive, crossed, and spurious quotes, midpoint prices for asset *i* on day *t* are sampled every five

minutes:

$$\text{MID}_{i,t} = \frac{\text{BID}_{i,t} + \text{ASK}_{i,t}}{2}. \tag{1}$$

The daily RV is then calculated as the square root of the sum of squared intraday returns:

$$RV_{i,t} = \sqrt{\sum_{s=1}^{N} \left[\ln \left(\frac{\text{MID}_{i,t_s}}{\text{MID}_{i,t_{s-1}}} \right) \right]^2}, \tag{2}$$

where MID_{i,t_s} is the mid-price at the sth five-minute interval, $N = 78$ for a standard equity trading day and $N = 288$ for 24-hour cryptocurrency markets.

3.2.2 Aggregated Realized Volatility. To capture medium-term risk dynamics, we construct weekly realized volatility by averaging daily RVs over a 7-day window ($h = 7$) following [8, 19]:

$$RV_t^7 = \frac{1}{7} \sum_{j=0}^{6} RV_{t-j}. \tag{3}$$

This 7-day moving average smooths daily noise, yielding a stable volatility gauge that matches common weekly rebalancing and reporting cycles.

4 Methodology

4.1 Volatility Factor Analysis Across Assets

4.1.1 Time-Varying Factor Model. Let $\mathbf{y}_t = (RV_{1,t}, \ldots, RV_{p,t})$ denote the p-dimensional vector of realized volatilities at time $t = 1, \ldots, T$. Following the time-varying matrix factor model of [11], we posit the locally-smooth factor representation:

$$\mathbf{y}_t = \mathbf{\Lambda}_t \mathbf{f}_t + \varepsilon_t, \tag{4}$$

where $\mathbf{f}_t \in \mathbb{R}^k$ is the k-dimensional latent factors, and $\varepsilon_t \in \mathbb{R}^p$ is the idiosyncratic errors, assumed to be uncorrelated with $\mathbf{f}_t$ but allowed to exhibit weak serial dependence. To accommodate a wide range of potential smooth temporal variation, we specify $\mathbf{\Lambda}_t \in \mathbb{R}^{p \times k}$ to be the time-varying loading matrix. Specifically, we model $\mathbf{\Lambda}_{t,i} = \mathbf{\Lambda}_{i \cdot}(t/T)$, where $\mathbf{\Lambda}_{i \cdot}(\cdot)$ is an unknown smooth function of t/T on $[0, 1]$ for each $i \in [p]$.

Let $\Sigma_f = \mathbb{E}(\mathbf{f}_t \mathbf{f}_t^\top)$ and $\Sigma_\varepsilon = \mathbb{E}(\varepsilon_t \varepsilon_t^\top)$. The second uncentered moment satisfies

$$\Sigma_{y,t} := \mathbb{E}(\mathbf{y}_t \mathbf{y}_t^\top) = \mathbf{\Lambda}_t \Sigma_f \mathbf{\Lambda}_t^\top + \Sigma_\varepsilon. \tag{5}$$

Under mild regularity conditions, the rank-k signal component $\mathbf{\Lambda}_t \Sigma_f \mathbf{\Lambda}_t^\top$ implies that the top k eigenvectors of $\Sigma_{y,t}$ span the column space of $\mathbf{\Lambda}_t$ up to rotation, as established in [11].

In principle, one could estimate $\Sigma_{y,t}$ using kernel smoothing in t/T as in [11]. For practical implementation, we instead adopt a *rolling-window* (uniform-kernel) approximation. Fix a window length n and define $W_t = \{t - n + 1, \ldots, t\}$. We estimate the local covariance matrix by

$$\widehat{\Sigma}_{y,t} = \frac{1}{n} \sum_{s \in W_t} \mathbf{y}_s \mathbf{y}_s^\top = \sum_{s=1}^{T} K_n(s,t) \, \mathbf{y}_s \mathbf{y}_s^\top, \ K_n(s,t) = \frac{1}{n} \mathbf{1}\{s \in W_t\}.$$

We extract the top-k eigenvectors $\hat{\mathbf{u}}_{1,t}, \ldots, \hat{\mathbf{u}}_{k,t}$ of $\widehat{\Sigma}_{y,t}$. The estimated loading space and the contemporaneous factors are then defined by

$$\widehat{\mathbf{\Lambda}}_t = \sqrt{p} \, [\hat{\mathbf{u}}_{1,t}, \ldots, \hat{\mathbf{u}}_{k,t}], \qquad \widehat{\mathbf{f}}_t = \frac{1}{p} \widehat{\mathbf{\Lambda}}_t^\top \mathbf{y}_t.$$

We apply this time-varying factor model separately to equities and cryptocurrencies, producing two distinct sets of $\widehat{\mathbf{\Lambda}}_t$ and $\widehat{\mathbf{f}}_t$.

4.1.2 Divergent Factor Structures: Concentration in Crypto vs. Diffusion in Equity. Both the equity and cryptocurrency markets exhibit hierarchical factor structures, yet they diverge significantly in risk concentration. This structural difference enables clear economic interpretations of the factors and provides a direct rationale for our horizon-dependent factor selection.

The equity market's risk is diffuse. Its leading factor, $\hat{\mathbf{f}}_t^{(1)}$, represents broad technology sector risk, explaining 45.2% of the total variance, with firms like Oracle loading most heavily at 0.96. However, as the gradual decay in the scree plot Figure 1b illustrates, a single factor is insufficient. Subsequent components, from $\hat{\mathbf{f}}_t^{(2)}$ to $\hat{\mathbf{f}}_t^{(5)}$, capture firm-specific risks for companies like Microsoft and Intel, meaning a broader set of factors is needed to comprehensively describe market dynamics.

In stark contrast, cryptocurrency's risk is highly concentrated. A single, dominant systemic factor $\hat{\mathbf{f}}_t^{(1)}$ serves as the primary market-wide driver, accounting for a commanding 79.1% of total variance. This factor is closely linked to the Bitcoin cycle, with its overwhelming influence confirmed by the sharp "elbow" in Figure 1d. While subsequent factors isolate the dynamics of major altcoins like Ethereum and Ripple, their marginal contributions are more modest, highlighting a market driven by one principal force.

These distinct factor structures provide a clear rationale for our horizon-dependent factor selection. For short-term (1-day) forecasts, we select only the single, dominant factor for each asset class to maximize the signal-to-noise ratio. For medium-term (7-day) forecasts, we employ an explained-variance threshold [36, 43]. This involves selecting the minimum number of factors required to meet a predetermined cumulative variance target, which we tailor to each market's unique structure (e.g., 85% for equities and 90% for crypto). This data-driven approach ensures we capture significant secondary dynamics without overfitting on noise.

4.2 Factor-Augmented Forecasting Models

To demonstrate the flexibility and model-agnostic nature of our framework, we apply it uniformly to four benchmark models. The models forecast a common target variable, Y_{t+h}, across two horizons $h \in \{1, 7\}$: the next day's daily volatility $Y_{t+1} = RV_{t+1}$ and the aggregated weekly volatility for the upcoming week $Y_{t+7} = RV_{t+7}^7$.

The forecasts are based on two estimation approaches: an expanding window for the three statistical models and a fixed 80:20 train-test split for the more computationally intensive LSTM.

4.2.1 Autoregressive (AR). The AR model regress the target volatility Y_{t+h} on the five most recent daily RVs to capture short-memory dynamics over a full trading week without over-parameterization:

$$Y_{t+h} = \beta_0 + \sum_{i=0}^{4} \beta_i RV_{t-i} + \epsilon_t. \tag{6}$$

(a) Factor Loadings for Equity (b) Explained Variance for Equity (c) Factor Loadings for Crypto (d) Explained Variance for Crypto
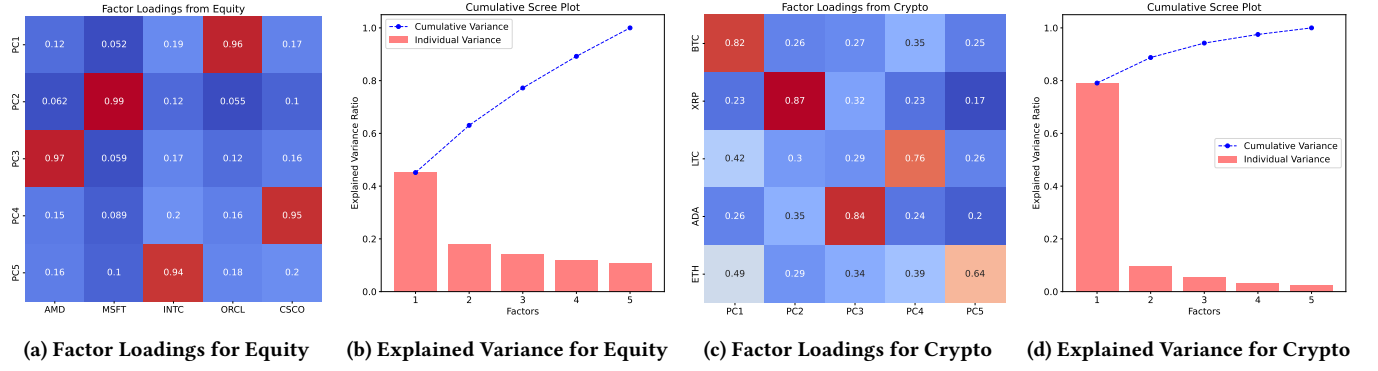
Figure 1: Factor Loadings and Scree Plots for Equities and Crypto

Given the AR model's simple linear structure, the factor-augmented version extends the baseline by directly incorporating the top-S factors as additional predictors:

$$Y_{t+h} = \beta_0 + \sum_{i=0}^{4} \beta_i RV_{t-i} + \sum_{j \in [S]} \gamma_j \widehat{f}_{j,t} + \epsilon_t. \tag{7}$$

4.2.2 Heterogeneous Autoregressive (HAR). The HAR model captures multi-scale volatility persistence using daily, weekly, and monthly components [19]:

$$Y_{t+h} = \beta_0 + \beta^d RV_t + \beta^w RV_t^7 + \beta^m RV_t^{30} + \epsilon_t, \tag{8}$$

where RV_t^d, RV_t^w, and RV_t^m represent daily, weekly, and monthly volatilities, respectively.

To align with the HAR's multi-scale structure, our augmentation introduces factors constructed over corresponding daily and weekly horizons:

$$Y_{t+h} = \beta_0 + \beta^d RV_t + \beta^w RV_t^7 + \beta^m RV_t^{30} \\ + \sum_{j \in [S^d]} \gamma_j^d \widehat{f}_{j,t}^d + \sum_{r \in [S^w]} \gamma_r^w \widehat{f}_{r,t}^w + \epsilon_t, \tag{9}$$

where $\widehat{f}_{j,t}^d$ and $\widehat{f}_{r,t}^w$ denote the top-S^d daily factors and the top-S^w weekly factors, respectively.

4.2.3 Mixed Data Sampling (MIDAS). The MIDAS model captures long-memory dynamics by incorporating many past observations through a parsimonious weighting scheme, specified as a Beta-polynomial distributed lag [30]:

$$Y_{t+h} = \beta_0 + \beta_1 \left[a(1)^{-1} a(L) RV_t \right] + \epsilon_t, \tag{10}$$

where lag weights $a(L)$ follow:

$$a_i = \left(\frac{i}{k} \right)^{\theta_1 - 1} \left(1 - \frac{i}{k} \right)^{\theta_2 - 1} \frac{\Gamma(\theta_1 + \theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)}, \quad i = 1, \ldots, k, \tag{11}$$

with normalization $a(1) = \sum_{i=1}^{k} a_i$. $\theta_1 = 1$ and θ_2 are optimized via in-sample grid search. Default $k = 30$ is extended with $k \in \{30, 50, 80\}$ for longer horizons.

To align with the MIDAS's core principle of parsimoniously weighting long-lag structures, we augment each factor through its

own analogous Beta-polynomial $b(L)$:

$$Y_{t+h} = \beta_0 + \beta_1 \left[a(1)^{-1} a(L) RV_t \right] + \sum_{j \in [S]} \gamma_j \left[b(1)^{-1} b(L) \widehat{f}_{j,t} \right] + \epsilon_t. \tag{12}$$

4.2.4 Long Short-Term Memory (LSTM). Long Short-Term Memory network is a gated recurrent architecture designed to capture long-range dependencies in time-series data. Each input vector concatenates a 7-day window of daily realized volatility with the contemporaneous factor. This sequence passes through a stack of three LSTM layers, which learn the joint temporal patterns of volatility and factors. The final hidden state feeds a dense output layer that forecasts the target volatility Y_{t+h}.

4.3 Evaluation Metrics

Model performance is examined through three lenses: (i) statistical accuracy, (ii) statistical significance, and (iii) economic value.

4.3.1 Prediction Accuracy. Prediction accuracy is assessed with three widely used metrics: R^2, MSE, and QLIKE, which respectively capture explained variance, average prediction error, and distributional accuracy respectively [42, 46]:

$$R^2 = 1 - \sum_t \left(\frac{RV_{t+h}^h - \widehat{RV}_{t+h}^h}{RV_{t+h}^h - \overline{RV}_{t+h}^h} \right)^2, \tag{13}$$

$$\text{MSE} = \frac{1}{T} \sum_t \left(RV_{t+h}^h - \widehat{RV}_{t+h}^h \right)^2, \tag{14}$$

$$\text{QLIKE} = \frac{1}{T} \sum_t \left[\frac{RV_{t+h}^h}{\widehat{RV}_{t+h}^h} - \ln \left(\frac{RV_{t+h}^h}{\widehat{RV}_{t+h}^h} \right) - 1 \right], \tag{15}$$

4.3.2 Economic Implication. We assess practical relevance using the out-of-sample Utility-of-Wealth (UoW) metric:

$$\widehat{\text{UoW}} = \frac{1}{T} \sum_{t=1}^{T} \left[\frac{SR^2}{\gamma} \frac{RV_{t+7}^7}{\widehat{RV}_{t+7}^7} - \frac{SR^2}{2\gamma} \left(\frac{RV_{t+7}^7}{\widehat{RV}_{t+7}^7} \right)^2 \right], \tag{16}$$

where we follow [8] in setting the target Sharpe ratio to SR = 0.4, a typical long-run estimate for volatility-timing strategies, and choose risk aversion $\gamma = 2$ to represent moderate investor preferences.

Table 1: Cryptocurrency Volatility Forecast Accuracy: 1- and 7-Day Out-of-Sample Results

	BTC			ETH			XRP			ADA			LTC		
	R^2	MSE	QLIKE	R^2	MSE	QLIKE	R^2	MSE	QLIKE	R^2	MSE	QLIKE	R^2	MSE	QLIKE
Panel A: Crypto in 1-Day Horizon															
AR	43.60	3.27	0.0802	40.34	3.50	0.0687	45.06	7.69	0.0842	41.33	4.16	0.0499	39.95	4.13	0.0658
AR-Aug	47.57	2.99	0.0770	42.34	3.38	0.0679	48.58	7.15	0.0789	42.01	4.13	0.0498	40.02	4.12	0.0654
HAR	39.52	3.12	0.0831	35.93	3.41	0.0691	47.58	6.98	0.0779	40.51	4.16	0.0497	34.08	4.09	0.0640
HAR-Aug	43.93	0.88	0.0797	40.43	3.19	0.0681	49.73	6.63	0.0778	41.09	4.14	0.0497	37.41	3.85	0.0640
MIDAS(k=30)	42.20	2.96	0.0828	39.07	3.29	0.0691	45.95	7.21	0.0843	40.94	4.19	0.0508	33.58	4.03	0.0683
MIDAS-Aug(k=30)	42.32	2.96	0.0827	38.93	3.30	0.0693	46.00	7.21	0.0844	40.94	4.19	0.0508	34.12	4.01	0.0679
LSTM	53.88	2.13	0.0603	51.24	2.75	0.0551	50.73	13.20	0.0824	45.12	6.12	0.0550	51.64	3.66	0.0530
LSTM-Aug	57.70	1.99	0.0564	55.72	2.64	0.0506	53.57	12.66	0.0800	47.34	6.08	0.0529	53.64	3.65	0.0519
Panel B: Crypto in 7-Day Horizon															
AR	48.63	1.87	0.0677	39.98	2.08	0.0552	41.78	5.53	0.0801	38.37	2.89	0.0426	39.21	2.24	0.0475
AR-Aug	50.89	1.78	0.0648	41.85	2.03	0.0544	43.07	5.49	0.0792	40.70	2.78	0.0414	42.38	2.14	0.0453
HAR	44.96	2.00	0.0793	38.25	2.15	0.0571	45.63	5.24	0.0782	36.71	2.96	0.0442	39.14	2.26	0.0492
HAR-Aug	47.38	1.94	0.0737	37.90	2.17	0.0577	46.58	5.29	0.0782	38.03	2.90	0.0435	40.36	2.22	0.0474
MIDAS(k=30)	49.40	1.84	0.0670	40.94	2.06	0.0546	43.26	5.47	0.0796	38.92	2.86	0.0424	40.82	2.20	0.0469
MIDAS-Aug(k=30)	53.26	1.70	0.0634	44.04	1.95	0.0516	47.23	5.09	0.0724	47.01	2.47	0.0390	45.08	2.07	0.0438
MIDAS(k=50)	49.64	1.83	0.0670	40.90	2.06	0.0546	43.26	5.52	0.0804	39.00	2.86	0.0423	40.69	2.20	0.0470
MIDAS-Aug(k=50)	52.95	1.71	0.0621	43.80	1.95	0.0514	47.34	5.05	0.0727	47.29	2.47	0.0389	44.30	2.07	0.0443
MIDAS(k=80)	49.58	1.83	0.0673	40.96	1.95	0.0516	42.73	5.09	0.0724	38.92	2.48	0.0390	40.80	2.04	0.0438
MIDAS-Aug(k=80)	52.95	1.71	0.0628	43.72	1.96	0.0518	47.33	5.08	0.0736	47.79	2.45	0.0384	44.31	2.07	0.0447
LSTM	59.31	1.37	0.0419	46.25	2.28	0.0441	42.81	1.43	0.1222	37.32	6.28	0.0629	52.39	2.38	0.0401
LSTM-Aug	61.66	1.30	0.0374	51.89	2.03	0.0385	44.25	1.41	0.1186	41.16	6.19	0.0616	59.83	2.38	0.0340

Note: All MSE values are reported in units of 10^{-4}. Improvement: ▮ ≥ +15% ▮ ≥ +10% ▮ ≥ +5% ▮ > 0% ▮ ≤ −5%

Higher UoW denotes greater expected utility for a volatility-timing investor.

4.3.3 Statistical Significance. We assess systematic improvements with the Diebold–Mariano test[21]. For each horizon h, we compute the one-period loss differential:

$$d_t = L\big(RV^h_{t+h}, \widehat{RV}^h_{t+h,\text{aug}}\big) - L\big(RV^h_{t+h}, \widehat{RV}^h_{t+h,\text{Baseline}}\big), \quad (17)$$

where

$$L(RV, \hat{RV}) = \begin{cases} (RV - \hat{RV})^2, & \text{MSE-based loss,} \\ -\left[\frac{SR^2}{\gamma}\frac{RV}{\hat{RV}} - \frac{SR^2}{2\gamma}\left(\frac{RV}{\hat{RV}}\right)^2\right], & \text{Utility-based loss.} \end{cases}$$
$$(18)$$

The DM statistic:

$$\text{DM} = \frac{\bar{d}}{\sqrt{\widehat{\text{Var}}(\bar{d})}}, \quad \bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t, \quad (19)$$

is asymptotically $\mathcal{N}(0,1)$ under $H_0 : E[d_t] = 0$. A positive DM value indicates that the factor-augmented forecasts significantly outperform the unaugmented benchmarks.

5 Experiment Result

5.1 A Universal Performance Boost Across All Models and Horizons

Factor augmentation universally and significantly boosts forecast accuracy across every model, asset class, and horizon tested, proving its effectiveness for both short- and medium-term predictions, as shown in Tables 1 and 4.

Specifically, in 1-day forecasts, the benefits of factor augmentation are most pronounced for short-memory models, with HAR and LSTM boosting equity R^2 by up to 10% and crypto R^2 by a commanding 9–12%. The MIDAS model, as expected from its long-memory design, shows more modest improvement.

At the 7-day horizon, long-memory models exhibit more substantial gains. While AR and HAR still show solid improvements, the augmented LSTM achieves an accuracy boost of up to 14%. The MIDAS model now stands out, delivering remarkable gains: its R^2 in equities surges by 11–13%, and by as much as 23% in cryptocurrencies, with corresponding reductions in MSE and QLIKE.

These results demonstrate the versatile power of factor augmentation, which consistently improves forecast accuracy, especially when a model's architecture aligns with the task horizon.

5.2 Value Across Markets: Robust in Equities, Super-Charged in Crypto

Factor augmentation framework proves its power and versatility across fundamentally different market structures, delivering robust

gains in the diffuse equity market and a super-charged performance boost in the highly concentrated crypto market.

In the challenging equity market, where volatility drivers are highly dispersed, our framework delivers substantial value. For instance, as shown in Table 4, the augmented 7-day MIDAS model boosts its R^2 by an average of 4.7%, while the 1-day LSTM improves its accuracy by 4.1%, proving the framework's reliability even in such diffuse regimes.

In cryptocurrency markets, the framework's competitive edge is even sharper. Here, where the top three factors capture nearly 95% of market variance, the performance gains are dramatic: as detailed in Table 1, the 7-day MIDAS model's R^2 soars by an average of 11.1% and the 1-day HAR model's by 7.9%—both more than doubling the improvements seen in equities.

These results confirm that factor augmentation is not only a reliable enhancer in different markets, but also becomes especially powerful when risk drivers are concentrated.

5.3 Statistically and Economically Significant Gains

Factor augmentation delivers gains that are both substantial and statistically significant in both predictive accuracy and economic value. We demonstrate this on the crypto 7-day horizon task, where our framework shows the largest improvements. Three unaugmented baselines are employed: the Random Walk, an "Individual" model with asset-specific coefficients, and an "Panel" model that share coefficients across assets.

The framework's edge in predictive accuracy is both substantial and statistically unequivocal. As shown in Table 5a, while the augmented models' consistently higher R^2 offer strong initial evidence, the positive statistics from an MSE-based DM test provide formal proof. This confirms that the forecasting improvements are not due to random chance but are a robust gain of our methodology.

Crucially, these statistical superiority translate directly into significant economic value. As detailed in Table 5b, the augmented models consistently deliver unambiguously higher UoW outcomes, and a corresponding UoW-based DM test formally validates that these gains are statistically significant, providing clear evidence of the framework's real-world benefits for investors.

6 Portfolio Backtest

To rigorously assess the framework's financial value, we conduct a volatility-scaled pair-trading backtest.

6.1 Cointegration Test

To identify a stationary, mean-reverting spread suitable for our pairs-trading strategy, we perform a two-stage cointegration analysis on the price series within each market.

We apply the Johansen trace test to each five-asset group. While it finds no evidence of cointegration at any rank for the equity group, the crypto sector indicates a cointegration rank of $r = 2$.

Table 2: Cointegrating Vector Coefficients for Crypto

	ADA	BTC	ETH	LTC	XRP
CEV_1	+2.62	+3.69	−4.60	−0.99	+2.23
CEV_2	+5.15	−0.19	−5.64	+1.82	−3.51

Table 2 reports the first two cointegrating vectors ($CEV_{1,2}$), both of which load ADA and ETH with opposite signs, indicating that the dominant mean-reverting combination is the ADA–ETH spread. To further validate this pairing, we conduct pairwise Engle–Granger tests on each residual and find only ADA–ETH stationary at the 5% level ($p = 0.045$). Consequently, our backtest focuses exclusively on the ADA–ETH spread, excluding all other pairs.

6.2 Volatility-Based Pairs Trading

Our backtest framework simulates a pairs trading strategy on the cointegrated spread between ADA-ETH. The log-price spread S_t is defined as:
$$S_t = \ln P_{\text{ADA},t} - \beta \ln P_{\text{ETH},t}, \tag{20}$$
where the hedge ratio β is estimated via rolling OLS regression. The performance is evaluated for all four benchmark models and is measured by annualized return and Sharpe ratio.

The performance is evaluated by final equity, cumulative return, and annualized Sharpe ratio.

6.2.1 Trading Signal Generation. Trading signals are generated from a rolling z-score of the spread, where $Z_t = \frac{S_t - \mu_{S_t}}{\sigma_{S_t}}$. We initiate a long position when $Z_t < $ -1.5 and a short position when $Z_t > $ +1.5, closing the position when the z-score reverts to zero. To limit turnover, each trade is held for at least one day. The rolling window length (W) is set to 70 days for the statistical models to balance responsiveness to regime changes with statistical stability, while a shorter window of 30 days is used for the LSTM to align with its input structure.

6.2.2 Volatility-based Position Sizing. Our strategy employs a volatility-targeting approach, where daily position sizes are scaled to achieve a 25% annualized volatility target. Each day, we estimate the spread volatility using our model's forecasts for the component assets' volatility $\widehat{\sigma}_{\text{ADA},t}$, $\widehat{\sigma}_{\text{ETH},t}$ and their return covariance:

$$\widehat{\sigma}_{\text{spread},t} = \sqrt{252} \cdot \sqrt{\widehat{\sigma}_{\text{ADA},t}^2 + \beta^2 \widehat{\sigma}_{\text{ETH},t}^2 - 2\beta \widehat{\text{Cov}}(r_{\text{ADA},t}, r_{\text{ETH},t})}. \tag{21}$$

The position size is then set based on this forecast, subject to a 5:1 maximum leverage. The backtest is initialized with an equity of $50,000 and incorporates a 5 basis point transaction cost per leg.

6.3 Backtest Results

To ensure a fair comparison, we evaluate statistical models on the full sample with an expanding window and the LSTM on a fixed 80:20 split. Each approach is benchmarked against a corresponding Random Walk baseline, with full results presented in Table 3.

6.3.1 Dual Gains: Amplifying Return and Sharpening Risk. Factor augmentation decisively boosts both profitability and risk-adjusted performance across all statistical models we examine.

The short-memory models realize the most impressive gains in raw profitability. Relative to its unaugmented version, the AR

Table 3: Out-of-Sample Backtest for ADA-ETH Pair

Model	Metric	Unaugmented	Factor-Augmented	Improvement
Panel A: Expanding Window Models				
Random Walk (Full Sample)	Portfolio Value ($)	59,161.34	–	–
	Ann. Return	5.32%	–	–
	Ann. Sharpe Ratio	1.402	–	–
AR	Portfolio Value	58,711.22	60,898.40	3.73%
	Ann. Return	5.35%	6.61%	23.55%
	Ann. Sharpe Ratio	1.437	1.443	0.42%
HAR	Portfolio Value	56,311.89	57,298.08	1.75%
	Ann. Return	3.86%	4.52%	17.10%
	Ann. Sharpe Ratio	1.457	1.475	1.24%
MIDAS (k=30)	Portfolio Value	58,339.32	58,121.55	-0.37%
	Ann. Return	5.31%	5.04%	-5.08%
	Ann. Sharpe Ratio	1.460	1.485	1.71%
Panel B: Late Sample Models				
Random Walk (Last 20% Window)	Portfolio Value	49,125.74	–	–
	Ann. Return	-2.67%	–	–
	Ann. Sharpe Ratio	-0.18	–	–
LSTM	Portfolio Value	47,427.84	53,306.58	12.40%
	Ann. Return	-5.48%	7.27%	232.66%
	Ann. Sharpe Ratio	-0.467	0.787	268.52%

specification boosts its annualized return by a remarkable 23.6%. The HAR model shows a parallel improvement, with its return climbing by 17.1% and its Sharpe ratio rising from 1.457 to 1.475.

The long-memory MIDAS model, in contrast, demonstrates its value through a different channel: superior risk calibration. Although its raw return sees only modest gains—a result of its long-term design—the augmented strategy's precise calibration of volatility drives its Sharpe ratio from 1.460 to a higher 1.485, signaling tighter risk management.

Collectively, this evidence confirms that factor augmentation is not merely an incremental tweak but a fundamental advance in volatility forecasting. It consistently amplifies reward, sharpens risk-adjusted performance, and sets a new benchmark for the field.

6.3.2 **Robustness in All Market Conditions: Enhancing Gains, Reversing Losses.** Our framework proves its value in all market conditions, systematically enhancing gains in favorable periods while acting as a powerful shield during downturns.

During the full-sample "bull-market" window, when even a naïve RW was profitable, factor augmentation still delivered a distinct competitive edge. Every statistical model, when augmented, posted a higher Sharpe ratio, sharpening performance even when market conditions are already favorable.

The framework's power becomes even more pronounced in adversity. During the challenging late-sample period, when the market turned negative and the RW benchmark lost 2.7% annually, the unaugmented LSTM suffered a -5.5% annualized loss, likely overfitting on patterns that no longer held. In a stunning reversal, our factor-augmented LSTM transformed this failing strategy into a resounding success, flipping the annualized loss into a +7.3% gain. This phenomenal turnaround propelled its Sharpe ratio from a negative -0.467 to a positive 0.787, unequivocally converting a losing proposition into a clear winner.

This remarkable dual capability—to amplify gains in good times and decisively reverse losses in bad—undeniably proves our framework's practical utility for navigating the complexities of any market conditions.

7 Conclusion

This paper proposes a novel, model-agnostic Factor-Augmented Volatility Forecast framework, resolving the trade-off between computationally infeasible multivariate models, overly simplistic univariate models, and rigid static-factor models. By extracting a compact set of dynamic factors and their time-varying loadings directly from realized volatilities, our frasmework allows any baseline model to efficiently capture evolving market-wide co-movements. The framework's robust performance—evidenced by substantial gains in forecast accuracy and superior risk-adjusted returns—combined with its computational efficiency and interpretability, establishes it as a valuable and practical solution for dynamic risk management.

Acknowledgments

References

[1] Carol Alexander. 2000. A primer on the orthogonal GARCH model. *manuscript ISMA Centre, University of Reading, UK* 2 (2000).

[2] Torben G. Andersen, Tim Bollerslev, Francis X. Diebold, and Heiko Ebens. 2001. The Distribution of Realized Stock Return Volatility. *Journal of Financial Economics* 61, 1 (July 2001), 43–76. doi:10.1016/S0304-405X(01)00055-1

[3] Alev Atak and George Kapetanios. 2013. A factor approach to realized volatility forecasting in the presence of finite jumps and cross-sectional correlation in pricing errors. *Economics Letters* 120, 2 (2013), 224–228.

[4] Richard T Baillie. 1996. Long memory processes and fractional integration in econometrics. *Journal of econometrics* 73, 1 (1996), 5–59.

[5] Luc Bauwens, Sébastien Laurent, and Jeroen VK Rombouts. 2006. Multivariate GARCH models: a survey. *Journal of applied econometrics* 21, 1 (2006), 79–109.

[6] Tim Bollerslev. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics* 31, 3 (1986), 307–327.

[7] Tim Bollerslev. 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics* (1987), 542–547.

[8] Tim Bollerslev, Benjamin Hood, John Huss, and Lasse Heje Pedersen. 2018. Risk everywhere: Modeling and managing volatility. *The Review of Financial Studies* 31, 7 (2018), 2729–2773.

[9] Mark M Carhart. 1997. On persistence in mutual fund performance. *The Journal of finance* 52, 1 (1997), 57–82.

[10] Bin Chen, Elynn Chen, Stevenson Bolivar, and Rong Chen. 2024. Time-varying Matrix Factor Models.

[11] Bin Chen, Elynn Y. Chen, Stevenson Bolivar, and Rong Chen. 2024. Time-Varying Matrix Factor Models. arXiv:2404.01546 [stat.ME] https://arxiv.org/abs/2404.01546

[12] Elynn Chen and Rong Chen. 2022. Modeling dynamic transport network with matrix factor models: with an application to international trade flow. *Journal of Data Science* (2022).

[13] Elynn Chen and Jianqing Fan. 2023. Statistical inference for high-dimensional matrix-variate factor models. *J. Amer. Statist. Assoc.* 118, 542 (2023), 1038–1055.

[14] Elynn Chen, Jianqing Fan, and Xiaonan Zhu. 2024. Factor Augmented Matrix Regression. *arXiv preprint arXiv:2405.17744* (2024).

[15] Elynn Chen, Ruey S Tsay, and Rong Chen. 2019. Constrained factor models for high-dimensional matrix-variate time series. *J. Amer. Statist. Assoc.* (2019).

[16] Elynn Chen, Dong Xia, Chencheng Cai, and Jianqing Fan. 2024. Semi-parametric tensor factor analysis by iteratively projected singular value decomposition. *Journal of the Royal Statistical Society Series B: Statistical Methodology* (2024), qkae001.

[17] ELYNN Y CHEN, DONG XIA, CHENCHENG CAI, and JIANQING FAN. 2020. SEMIPARAMETRIC TENSOR FACTOR ANALYSIS BY ITERATIVELY PROJECTED SVD BY ELYNN Y. CHEN, DONG XIA, CHENCHENG CAI, AND JIANQING FAN. *arXiv preprint arXiv:2007.02404* (2020).

[18] Christian Conrad and Onno Kleen. 2020. Two are better than one: volatility forecasting using multiplicative component GARCH-MIDAS models. *Journal of Applied Econometrics* 35, 1 (2020), 19–45.

[19] Fulvio Corsi. 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7, 2 (2009), 174–196.

[20] Marco Del Negro and Chris Otrok. 2008. Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *FRB of New York Staff Report* 326 (2008).

[21] Francis X Diebold and Robert S Mariano. 2002. Comparing predictive accuracy. *Journal of Business & economic statistics* 20, 1 (2002), 134–144.

[22] Yi Ding, Robert Engle, Yingying Li, and Xinghua Zheng. 2025. Multiplicative factor model for volatility. *Journal of Econometrics* 249 (2025), 105959.

[23] Robert Engle. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of business & economic statistics* 20, 3 (2002), 339–350.

[24] Robert F Engle. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society* (1982), 987–1007.

[25] Robert F Engle and Kenneth F Kroner. 1995. Multivariate simultaneous generalized ARCH. *Econometric theory* 11, 1 (1995), 122–150.

[26] Robert F Engle, Olivier Ledoit, and Michael Wolf. 2019. Large dynamic covariance matrices. *Journal of Business & Economic Statistics* 37, 2 (2019), 363–375.

[27] Eugene F Fama and Kenneth R French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of financial economics* 33, 1 (1993), 3–56.

[28] Guanhao Feng, Jingyu He, Nicholas G Polson, and Jianeng Xu. 2024. Deep learning in characteristics-sorted factor models. *Journal of Financial and Quantitative Analysis* 59, 7 (2024), 3001–3036.

[29] Wenbo Ge, Pooia Lalbakhsh, Leigh Isai, Artem Lenskiy, and Hanna Suominen. 2023. Neural Network–Based Financial Volatility Forecasting: A Systematic Review. *Comput. Surveys* 55, 2 (2023), 14:1–14:30. doi:10.1145/3483596

[30] Eric Ghysels, Arthur Sinko, and Rossen Valkanov. 2007. MIDAS Regressions: Further Results and New Directions. *Econometric Reviews* 26, 1 (Feb. 2007), 53–90. doi:10.1080/07474930600972467

[31] Vasyl Golosnoy, Bastian Gribisch, and Roman Liesenfeld. 2012. The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics* 167, 1 (2012), 211–223.

[32] Shihao Gu, Bryan Kelly, and Dacheng Xiu. 2020. Empirical asset pricing via machine learning. *The Review of Financial Studies* 33, 5 (2020), 2223–2273.

[33] Yasushi Hamao, Ronald W Masulis, and Victor Ng. 1990. Correlations in price changes and volatility across international stock markets. *The review of financial studies* 3, 2 (1990), 281–307.

[34] Bernard Herskovic, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh. 2016. The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics* 119, 2 (2016), 249–283.

[35] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.

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

[37] Linghang Kong, Elynn Chen, Yuzhou Chen, and Yuefeng Han. 2024. TEAFormers: TEnsor-Augmented Transformers for Multi-Dimensional Time Series Forecasting. In *The 34th International Joint Conference on Artificial Intelligence, 2025, AI for Time Series Workshop; arXiv preprint arXiv:2410.20439*.

[38] Xialu Liu and Elynn Chen. 2022. Identification and estimation of threshold matrix-variate factor models. *Scandinavian Journal of Statistics* (2022).

[39] Yang Liu. 2019. Novel volatility forecasting using deep learning–long short term memory recurrent neural networks. *Expert Systems with Applications* 132 (2019), 99–109.

[40] Sean McNally, Jason Roche, and Simon Caton. 2018. Predicting the Price of Bitcoin Using Machine Learning. In *2018 26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*. IEEE, 339–343. doi:10.1109/PDP2018.2018.00060

[41] Jakob Guldbæk Mikkelsen, Eric Hillebrand, and Giovanni Urga. 2019. Consistent estimation of time-varying loadings in high-dimensional factor models. *Journal of Econometrics* 208, 2 (2019), 535–562.

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

[43] Kuntara Pukthuanthong and Richard Roll. 2009. Global market integration: An alternative measure and its application. *Journal of Financial Economics* 94, 2 (2009), 214–232.

[44] Douglas G Santos and Flavio A Ziegelmann. 2014. Volatility forecasting via MIDAS, HAR and their combination: An empirical comparative study for IBOVESPA. *Journal of Forecasting* 33, 4 (2014), 284–299.

[45] Roy Van der Weide. 2002. GO-GARCH: a multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics* 17, 5 (2002), 549–564.

[46] Chao Zhang, Yihuang Zhang, Mihai Cucuringu, and Zhongmin Qian. 2024. Volatility forecasting with machine learning and intraday commonality. *Journal of Financial Econometrics* 22, 2 (2024), 492–530.

A Equity Market: Forecast Performance

Table 4: Equity Volatility Forecast Accuracy: 1- and 7-Day Out-of-Sample Results

Panel A: Equity in 1-Day Horizon

Models	MSFT R^2	MSE	QLIKE	AMD R^2	MSE	QLIKE	INTC R^2	MSE	QLIKE	ORCL R^2	MSE	QLIKE	CSCO R^2	MSE	QLIKE
AR	33.75	7.30	0.1756	32.07	8.79	0.0570	33.64	13.79	0.1115	36.49	23.47	0.1486	34.58	12.72	0.1133
AR-Aug	34.26	7.30	0.1741	32.33	8.78	0.0566	34.01	13.57	0.1111	37.69	22.78	0.1443	34.78	12.55	0.1132
HAR	33.48	7.40	0.1740	32.87	8.72	0.0540	31.37	14.28	0.1139	37.11	23.37	0.1469	34.34	12.75	0.1087
HAR-Aug	33.78	7.37	0.1755	32.52	8.84	0.0548	31.96	14.20	0.1135	37.21	23.21	0.1453	37.79	12.08	0.1049
MIDAS (k=30)	35.58	7.16	0.1759	34.58	8.61	0.0535	33.89	14.04	0.1134	38.19	23.39	0.1504	36.84	12.53	0.1083
MIDAS-Aug	35.44	7.17	0.1771	34.49	8.62	0.0533	34.29	13.95	0.1129	38.60	23.46	0.1509	36.99	12.52	0.1083
LSTM	38.87	4.47	0.1237	27.51	13.22	0.0577	42.92	21.12	0.1103	27.10	54.25	0.1887	46.74	18.00	0.0887
LSTM-Aug	39.33	4.43	0.1223	28.44	13.05	0.0564	44.16	20.73	0.1089	29.84	52.21	0.1876	48.08	17.79	0.0882

Panel B: Equity in 7-Day Horizon

Models	MSFT R^2	MSE	QLIKE	AMD R^2	MSE	QLIKE	INTC R^2	MSE	QLIKE	ORCL R^2	MSE	QLIKE	CSCO R^2	MSE	QLIKE
AR	44.99	3.48	0.0946	47.20	4.12	0.0269	44.03	6.71	0.0499	45.67	12.87	0.0730	47.85	6.12	0.0514
AR-Aug	45.34	3.46	0.0964	47.43	4.10	0.0268	43.87	6.73	0.0503	46.81	12.59	0.0718	49.38	5.94	0.0504
HAR	44.39	3.86	0.1084	42.88	4.48	0.0257	37.83	8.36	0.0540	20.99	21.98	0.1154	40.32	8.14	0.0558
HAR-Aug	45.71	3.81	0.1096	43.24	4.39	0.0258	40.07	8.06	0.0542	21.57	21.82	0.1036	45.27	7.47	0.0555
MIDAS (k=30)	47.26	3.33	0.0875	48.09	3.73	0.0254	45.53	6.21	0.0477	47.82	11.46	0.0741	50.02	5.61	0.0515
MIDAS-Aug	49.61	3.18	0.0853	48.76	3.76	0.0251	46.29	6.12	0.0472	48.37	11.33	0.0711	55.53	4.99	0.0493
MIDAS(k=50)	46.35	3.41	0.0927	50.17	3.65	0.0251	45.36	6.27	0.0500	48.67	11.35	0.0842	55.30	5.05	0.0572
MIDAS-Aug(k=50)	48.38	3.28	0.0859	50.49	3.63	0.0247	46.52	6.13	0.0472	48.70	11.35	0.0793	55.92	4.98	0.0545
MIDAS(k=80)	49.05	3.15	0.0885	50.85	3.76	0.0249	45.13	6.42	0.0496	48.22	11.95	0.0750	50.07	5.72	0.0526
MIDAS-Aug(k=80)	51.36	3.01	0.0866	51.49	3.71	0.0239	45.68	6.35	0.0485	49.86	11.87	0.0728	56.50	4.98	0.0503
LSTM	46.38	3.55	0.1002	43.82	7.02	0.0337	42.98	15.13	0.0779	42.14	39.23	0.1491	49.47	13.01	0.0644
LSTM-Aug	47.36	3.51	0.0964	44.82	6.91	0.0326	45.15	14.54	0.0766	42.74	33.29	0.1416	50.55	12.74	0.0628

Note: All MSE values are reported in units of 10^{-4}. Improvement: ≥ +15% ≥ +10% ≥ +5% > 0% ≤ −5%

B Statistical Significance

(a) Out-of-Sample R^2 and MSE-based DM Test for 7-day Crypto.

Asset	Model	Random Walk	AR5	HAR	MIDAS (k=30)	MIDAS (k=50)	MIDAS (k=80)	LSTM
Panel A: Out-of-Sample R^2 (in %)								
BTC	Individual	32.48	48.63	44.96	49.40	49.64	49.58	59.31
	Panel	–	48.02	47.65	48.52	47.59	45.62	57.91
	Factor-Augmented	–	50.89	47.38	53.26	52.95	52.95	61.66
ETH	Individual	17.27	39.98	38.25	40.94	40.90	40.96	46.25
	Panel	–	40.83	39.54	41.06	41.13	39.37	48.65
	Factor-Augmented	–	41.85	37.90	44.04	43.80	43.72	51.89
XRP	Individual	22.94	41.78	45.63	43.26	43.26	42.73	42.81
	Panel	–	42.68	46.76	42.86	42.27	41.10	46.81
	Factor-Augmented	–	43.07	46.58	47.23	47.34	47.33	44.25
ADA	Individual	12.32	38.37	36.71	38.92	39.00	38.92	37.32
	Panel	–	40.15	42.36	40.51	40.76	39.48	48.96
	Factor-Augmented	–	40.70	38.03	47.01	47.29	47.79	41.16
LTC	Individual	20.23	39.21	39.14	40.82	40.69	40.80	52.39
	Panel	–	39.37	41.22	39.84	38.87	36.58	56.54
	Factor-Augmented	–	42.38	40.36	45.08	44.30	44.31	59.83
Panel B: MSE-based Diebold–Mariano Test Statistics								
BTC	Aug vs. RW	NA	2.46	1.25	3.34	3.36	3.32	3.71
	Aug vs. Ind	–	3.13	1.29	2.18	2.21	2.06	4.32
ETH	Aug vs. RW	NA	3.07	2.34	4.60	4.56	4.48	3.56
	Aug vs. Ind	–	1.65	1.61	2.10	2.19	1.95	2.71
XRP	Aug vs. RW	NA	2.60	3.09	3.50	3.57	3.55	2.94
	Aug vs. Ind	–	1.89	1.78	1.79	1.77	1.68	2.13
ADA	Aug vs. RW	NA	3.29	3.35	3.80	3.76	3.84	1.95
	Aug vs. Ind	–	2.14	1.14	2.11	1.79	1.70	1.61
LTC	Aug vs. RW	NA	2.18	1.94	3.78	3.78	3.79	3.12
	Aug vs. Ind	–	3.94	1.52	1.27	1.10	1.27	2.09

Note: Aug = Factor-Augmented, Ind = Unaugmented Individual, RW = Random Walk.

(b) Out-of-Sample UoW and UoW-based DM Test for 7-day Crypto

Asset	Model	Random Walk	AR5	HAR	MIDAS (k=30)	MIDAS (k=50)	MIDAS (k=80)	LSTM
Panel A: Out-of-Sample Utility of Wealth (in %)								
BTC	Individual	2.66	3.31	2.87	3.25	3.23	3.22	3.66
	Panel	-	3.32	3.22	3.27	3.27	3.27	3.65
	Factor-Augmented	-	3.35	3.10	3.34	3.33	3.34	3.68
ETH	Individual	3.13	3.51	3.46	3.49	3.48	3.48	3.58
	Panel	-	3.50	3.46	3.48	3.48	3.48	3.60
	Factor-Augmented	-	3.52	3.47	3.53	3.52	3.51	3.65
XRP	Individual	2.72	3.14	3.11	3.08	3.06	3.04	2.36
	Panel	-	3.10	3.13	3.03	3.03	3.03	2.48
	Factor-Augmented	-	3.24	3.23	3.20	3.20	3.19	2.41
ADA	Individual	3.39	3.62	3.57	3.59	3.59	3.59	3.31
	Panel	-	3.60	3.61	3.60	3.60	3.60	3.45
	Factor-Augmented	-	3.65	3.62	3.62	3.62	3.61	3.35
LTC	Individual	3.23	3.59	3.52	3.56	3.55	3.55	3.68
	Panel	-	3.56	3.56	3.56	3.56	3.56	3.69
	Factor-Augmented	-	3.60	3.57	3.59	3.58	3.58	3.71
Panel B: UoW-based Diebold-Mariano Test Statistics								
BTC	Aug vs. RW	NA	3.35	3.10	3.34	3.33	3.34	3.68
	Aug vs. Ind	–	1.46	1.43	1.96	2.17	2.06	2.93
ETH	Aug vs. RW	NA	3.52	3.47	3.53	3.52	3.51	3.65
	Aug vs. Ind	–	1.55	0.668	1.84	1.79	1.74	2.38
XRP	Aug vs. RW	NA	3.24	3.23	3.20	3.20	3.19	2.41
	Aug vs. Ind	–	1.82	1.20	1.52	1.62	1.70	1.63
ADA	Aug vs. RW	NA	3.65	3.62	3.62	3.62	3.61	3.35
	Aug vs. Ind	–	2.43	1.13	2.17	2.30	2.24	2.32
LTC	Aug vs. RW	NA	3.60	3.57	3.59	3.58	3.58	3.71
	Aug vs. Ind	–	1.67	1.72	1.89	2.24	1.89	2.44

Note: Definitions same as table (a).