Crypto Factor Zoo (.zip)

Aleksander Mercik[†*], *Adam Zaremba*[‡], *and Ender Demir*[⁑]

Abstract

How many factors are genuinely needed to explain the cross-section of cryptocurrency returns? To answer this, we are the first to apply the alpha-based, iterative factor selection methodology of Swade et al. (2024)—originally developed for equities—to the cryptocurrency market. Using a comprehensive set of 36 return-predictive factors, we find that just two to three factors can eliminate all significant portfolio alphas. The most influential factors include turnover volatility, bid–ask spreads, and blockchain-native metrics such as the new-address-to-price ratio. Liquidity-related variables dominate the selection process, appearing consistently across weighting schemes, model specifications, and periods.

Keywords: cryptocurrency markets, factor zoo, asset pricing, GRS test, return predictability

JEL codes: G11, G12, G17

Declaration of Interest: The authors declare no conflicts of interest.

This version: July 2025

† Aleksander Mercik, Wroclaw University of Economics and Business, Department of Financial Investments and Risk Management, ul. Komandorska 118/120, 53-345 Wrocław, Poland; email: aleksander.mercik@ue.wroc.pl.

* Corresponding author.

‡ Adam Zaremba, 1) MBS School of Business, 2300, avenue des Moulins, 34185 Montpellier cedex 4, France, email: a.zaremba@montpellier-bs.com; 2) Department of Investment and Capital Markets, Institute of Finance, Poznan University of Economics and Business, al. Niepodległości 10, 61-875 Poznań, Poland, adam.zaremba@ue.poznan.pl; 3) Monash University, Melbourne, Australia.

⁑ Department of Business and Economics, Reykjavik University, Reykjavik, Iceland, enderd@ru.is.

1. Introduction

In recent decades, the explosion of research into cross-sectional return predictors has created what Cochrane (2011) famously termed a "zoo of factors," with hundreds of documented characteristics claiming to predict returns (Harvey et al., 2016; Mclean & Pontiff, 2016; Hou et al., 2020; Chen & Zimmermann, 2021; Jensen et al., 2023). In traditional asset markets, the increasing number of proposed anomalies has raised serious questions about the true dimensionality of returns. Researchers seek to determine the minimum number of factors needed to capture the full range of the "zoo" (Feng et al., 2020; Kozak et al., 2020; Anatolyev & Mikusheva, 2022; Chinco et al., 2022; Ahmed et al., 2023; De Nard & Zhao, 2023). A similar problem has emerged in the cryptocurrency market, where the number of proposed return drivers skyrockets (Liu et al., 2022; Zhang & Li, 2023; Cakici et al., 2024). Yet, the true dimensionality of risk in cryptocurrency markets remains an open question.

In this paper, we address this problem by applying Swade et al.'s novel iterative alpha-based factor compression methodology (2024). Using cryptocurrency data from January 2018 to July 2024, we construct 36 factors that encompass traditional market-based measures, microstructure variables, and blockchain-native characteristics. Our procedure begins with the standard market factor and sequentially incorporates the factor that most reduces the Gibbons et al. (1989) test statistic (GRS) until all statistically significant alphas are eliminated. We apply this method to equal- and value-weighted long-short portfolios and perform diverse robustness tests across weighting schemes, sorting depths, and subsample periods.

Our findings reveal a sparse yet consistent factor structure. First, only three factors beyond the market risk premium are needed to explain all abnormal returns on equal- and value-weighted portfolios. Specifically, the iterative GRS reduction procedure selects three factors before convergence for equal-weighted portfolios: turnover volatility, salience theory value, and the new-address-to-price ratio. A similarly small set dominated by bid-ask spread and 7-day momentum is required for value-weighted portfolios.

Second, the composition of the selected factors differs fundamentally from that of traditional asset markets. While equity markets emphasize size, value, profitability, and investment factors (Fama & French, 2015), cryptocurrency returns are primarily explained by liquidity-related variables and blockchain-native metrics. Microstructure factors, particularly turnover volatility and bid-ask spread, consistently rank among the most important. Blockchain-specific signals, such as the new-address-to-price ratio and network activity growth, provide additional explanatory power, even when controlling for traditional factors. This highlights how on-chain data captures unique information about adoption and usage. Conventional risk measures such as market beta, downside risk, and size effects dominate equity markets and play minimal roles in the cryptocurrency cross-section.

Third, although the overall sparsity and broad factor categories are stable, the specific factors selected have limited temporal persistence. When we divide our sample into two subperiods, only one factor, bid-ask spread, survives in both halves for value-weighted portfolios. Additionally, factor rankings shift substantially between periods. This lack of stability suggests that cryptocurrency factor premia may reflect evolving market inefficiencies rather than stable risk compensation. Early-period factors include idiosyncratic choices, such as minimum daily returns and kurtosis. Later periods emphasize more conventional liquidity and momentum measures, which could indicate market maturation.

Our findings relate to three strands of literature. First, we extend the factor zoo compression methodology to a new asset class, demonstrating that the principle of factor sparsity documented by Swade et al. (2024) in equity markets also holds in cryptocurrencies. Although previous studies have identified various factors, ranging from momentum and size (Liu et al., 2021) to blockchain-specific metrics (Cong et al., 2022; Bhambhwani et al., 2023; Sakkas & Urquhart, 2024; Lan & Frömmel, 2025), few have addressed the fundamental question of how many factors are necessary to capture the cross-sectional variation and remove unexplained abnormal returns fully.

Second, we add to the debate on cryptocurrency anomalies and their predictability of cross-sectional returns. The literature thus far has documented an extensive list of return-predicting signals, explore by, e.g. Jia et al. (2021), Li et al. (2021), Liu et al. (2021), Liu et al. (2022), Liu and Tsyvinski (2021), Zhang et al. (2021), Borri and Shakhnov (2022), Cong et al. (2022), Bianchi et al. (2022), Chen et al. (2022), Chang et al. (2023), Zhang and Zhao (2023), Cai and Zhao (2024), Fieberg et al. (2024), Dobrynskaya (2024), Younus and Naeem (2024), Zhao et al. (2024), Fieberg et al. (2024) and Lan and Frömmel (2025). By replicating a comprehensive set of 36 predictors, we shed additional light on their robustness and unique information content.

Third, we emphasize the role of liquidity as a driver of cryptocurrency returns, particularly in terms of turnover volatility and bid-ask spreads, matching earlier evidence from cryptocurrency markets (Brauneis et al., 2021; Zhang & Li, 2023; Liu et al., 2022; Han, 2023; Dong et al., 2022; Leirvik, 2022; and Garfinkel et al., 2023).

The structure of this article is organized as follows: Section 2 reviews the related literature and develops our hypotheses. Section 3 describes our data and methodology. Section 4 presents our empirical results, including robustness tests. Section 5 concludes.

2. Data and Methods

This section outlines the methodologies and data used in our research. First, we detail the composition of our sample and the source of our data. Next, we focus on the portfolios sorted by individual characteristics. Lastly, we describe the iterative factor selection procedure that identifies the minimum set of factors necessary to cover the full range of cryptocurrencies, including our stopping criteria and robustness tests across alternative specifications.

2.1. Sample and Data Sources

Our dataset merges information from CryptoCompare, CoinMarketCap, and IntoTheBlock, following the approach of Babiak and Bianchi (2024).[1] CryptoCompare provides volume-weighted open, high, low, and close prices, as well as trading volumes, which are collected from over 250 centralized exchanges. This gives an unbiased estimate of market prices. CoinMarketCap provides market-capitalization figures and the categorical tags necessary for our fixed exclusion rules, and IntoTheBlock contributes blockchain activity indicators. We use the two-step procedure described by Cakici et al. (2024) and Mercik et al. (2025) to align the observations across the three providers. First, cryptocurrencies are matched based on their full name and ticker symbol. Any residual mismatches are then checked manually. Unmatched tokens are then removed from the universe.

We then apply a two-stage refinement process to the raw panel, replicating the approach of Bianchi and Babiak (2022). In the fixed stage, we exclude the following, all of which are identified with CoinMarketCap classifications: (i) centrally managed or algorithmic stablecoins, such as USDT and DAI; (ii) metal-pegged tokens; (iii) collateral coins used by derivatives platforms, such as SNX; and (iv) wrapped assets, such as WBTC. In the dynamic stage, we eliminate non-positive price, volume, or market value observations, in line with Cong et al. (2022). We also discard days on which the trading volume-to-market capitalization ratio exceeds one, as Bianchi et al. (2022) suggest. Additionally, we remove coins whose market value falls below the $1 million threshold recommended by Liu et al. (2022). Finally, we winorize daily returns at -99% and $+500\%$, following Bianchi and Babiak (2022).

[1] CryptoCompare has been recognized for its reliability by Alexander and Dakos (2020) and is frequently cited in various studies (Borri, 2019; Lucchini et al., 2020; Bianchi et al., 2022; Borri & Shakhnov, 2022). IntoTheBlock as a source on-chain activity data was referenced by Cong et al. (2022), Bianchi and Babiak (2022), and Hoang & Baur (2022).

After applying these filters, the sample includes 565 unique cryptocurrencies observed from January 1, 2018, to July 25, 2024, at 12:00 a.m. UTC. Figure 1 plots the cross-sectional count over time. Though this set is small compared to the over 24,000 tokens listed on CoinMarketCap, it represents the vast majority of investable market value.

[Insert Figure 1 about here]

The study period encompasses the ICO surge and subsequent bust, the COVID-19 shock, Russia's invasion of Ukraine (Vidal-Tomás, 2022), the listing of Bitcoin and Ethereum futures on the Chicago Mercantile Exchange.[2] The launch of bitcoin exchange-traded funds[3]and China's successive bans on domestic crypto exchanges.

Following Liu et al. (2022) and Babiak and Bianchi (2024), we compiled 36 return-predictive variables identified in the literature, supplementing them with recent findings on crypto assets. These variables fall into six conceptual groups.

On-chain activity. This group includes the following variables: the new addresses (*new_add*), changes in new addresses (*new_add_ch*), active addresses (*active_add*), zero balance addresses (*zero_bal*), growth in balance addresses (*ba_growth*), network to market ratio (*bm*), new address to price ratio (*ap*), and active address to network value ratio (*aanv*). These variables are based on the research of Pagnotta & Buraschi (2018), Liu et al. (2021), Cong et al. (2022), and Liebi (2022).

Liquidity. This category comprises trading volume (*volume*), market value (*size*), bid-ask spread (*bidask*), illiquidity ratio (*illiq*), turnover (*turn*), detrended turnover (*dto*), turnover volatility (*std_dto*), trading volume volatility (*std_vol*), standardized abnormal turnover (*sat*), 30-day volume shocks (*volsh_30d*), and 60-day volume shocks (*volsh_60d*). The predictive power of these variables has been documented by Brauneis et al. (2021), Li et al. (2021), Zhang & Li (2023), Liu et al. (2022), Han (2023), Dong et al. (2022), Leirvik (2022), and Garfinkel et al. (2023).

Volatility and risk. Key measures in this group are realized volatility (*rvol*), capital asset pricing model (CAPM) beta (*beta*), idiosyncratic risk (*ivol*), and value-at-risk (*var*). These variables have been linked to future returns in studies by Jia et al. (2021), Zhang

[2] Bitcoin futures were first launched on December 11, 2017 (Liu et al., 2020) by the Chicago Board Options Exchange (CBOE), while a week later, the Chicago Mercantile Exchange (CME) launched its own bitcoin futures product (Fassas et al., 2020). These were significant events as they represented the introduction of Bitcoin to regulated financial markets.

[3] Regulatory restrictions on traditional funds have created over 180 bitcoin ETPs, enabling indirect crypto investment without ownership. Half of these have launched since late 2021, responding to strong demand from retail and institutional investors (Gemayel et al., 2023).

and Li (2023), Burggraf and Rudolf (2021), Zhang et al. (2021), and Dobrynskaya and Dubrovskiy (2023).

Historical return patterns. This group includes 7-day momentum (*r7_2*), 13-day momentum (*r13_2*), 23-day momentum (*r22_2*), 31-day momentum (*r31_2*), intermediate momentum (*r30_14*), daily reversal (*r2_1*), long-term reversal (*r360_31*), 90-day high (*90dh*), and CAPM alpha (*alpha*). These patterns align with the findings of Grobys & Sapkota (2019), Liu et al. (2020), Shen et al. (2020), Tzouvanas et al. (2020), Dobrynskaya (2023), Liu and Tsyvinski (2021), Liu et al. (2022) and Bianchi et al. (2022).

Return-distribution characteristics. This category includes skewness (*skew*), kurtosis (*kurt*), maximum daily return (*max*), and minimum daily return (*min*). Our approach follows the methodology of Bali et al. (2011) and its extensions to the cryptocurrency context by Grobys and Junttila (2021), Jia et al. (2021), Lin et al. (2021), Liu et al. (2021), and Ozdamar et al. (2021).

Other anomalies include nominal price (Miller & Scholes, 1982), cross-sectional seasonality (Keloharju et al., 2016), the salience measure of Cosemans and Frehen (2021), and chronological-order variables proposed by Mohrschladt (2021), denoted *prc*, *seas*, st, and *cro*, respectively. Zaremba et al. (2021), Cai and Zhao (2024), Chen et al. (2022), and Liu et al. (2022) support the relevance of these anomalies for crypto assets.

Table 1 describes each return-predictive variable and categorizes the 36 characteristics into six groups based on economic motivation. These variables were selected based on their documented predictive power in prior literature and their relevance to the unique structure of crypto markets. Table 2 presents summary statistics for returns, market values, and the full set of characteristics used in the analysis. The wide range of values across different categories reflects the diversity of cryptocurrencies in terms of trading activity, risk profiles, and blockchain usage.

[Insert Table 1 about here]

[Insert Table 2 about here]

To evaluate the profitability of the characteristics, we construct weekly long–short portfolios. At the end of each week, all cryptocurrencies are sorted by a given characteristic. The upper and lower quartiles form the long and short sides, respectively, of a value—or equal-weighted portfolio that is rebalanced after one week.

2.2. Methodology

We aim to identify the smallest number of factors necessary to capture all factor alphas. Inspired by Swade et al. (2024), we identify a factor model for cryptocurrencies that encompasses the full range of factors from an alpha perspective. We use a simple and efficient nested model approach. We iteratively add new factors to an expanding factor model until all remaining alphas become insignificant in the cross-section of equity factors. Starting with the CAPM, we add the factor that, when included in a two-factor model, most significantly reduces the remaining alphas, as indicated by the lowest GRS statistic. This selection criterion is equivalent to selecting the factor with the highest t-statistic for the current model. Once a factor is identified, it is permanently included in the factor model, and the procedure is repeated with the expanded factor models until no significant contributors remain. The selection strategy is as follows:

Step 1. Set $l := 0$ and start spanning the factor zoo using the one-factor model

$$f_i = \alpha_i + \beta_{CM} r_{CM} + \varepsilon_i \quad i = 1, \ldots, N \tag{1}$$

where r_m represents the excess market return and N denotes the count of factors beyond the crypto market factor.

Step 2. Test N - 1 different augmented factor models, each adding one of the remaining factors, labeled $f^{\,test}$, to the model from the previous iteration:

$$f_i = \alpha_i + \beta_{CM} r_{CM} + \sum_{k=1}^{l} \beta_k f_k + \beta^{test} f^{test} + \varepsilon_i \qquad i = 1, \ldots, N - l \tag{2}$$

Step 3. Rank the tested factor models according to their explanatory power, quantified by their GRS statistic, and select the highest statistic.

Step 4. Update $l := l + 1$ and determine the number of remaining significant factor alphas $n(\alpha)_{t>x}$ based on the augmented factor model as

$$n(\alpha)_{t>x} = |\{\, \alpha_i \mid t(\alpha_i) > x \,\}| \qquad i = 1, \ldots, N - l \tag{3}$$

where x is the selected significance threshold.

Step 5. If $n(\alpha)_{t>x} = 0$, meaning the remaining factors are statistically insignificant, stop the process. Otherwise, return to *Step 2.*

The stopping criterion for accurately determining the number of factors is simple: the total number of remaining significant factor alphas must be zero. Once a new factor model is identified, we test all remaining factors against this model and calculate the alphas for the remaining candidate factors. If the newly added factors are significant, the remaining significant factor alphas should decrease during the process. Alternative criteria could include the level of significance of the newly added factor based on statistical tests. If the new factor does not meet a significance threshold, it should not be considered a strong factor and should not be added to the model. Swade et al. (2024) recommend using higher statistical thresholds to address data mining concerns and potential misspecification. In line with Swade et al. (2024) and Harvey et al. (2016), we conduct our analysis using the standard thresholds of t > 1.96 and a more conservative threshold of t > 3.00.

We apply the GRS statistic from Gibbons et al. (1989) to test whether adding candidate factors improves a model's ability to explain expected returns by examining whether the intercepts (alphas) across test assets are jointly zero. Following Fama and French (2018), we express the test in terms of Sharpe ratios, where the maximum squared Sharpe ratio of the alphas is $Sh^2(\alpha) = \alpha^\top \Sigma^{-1} \alpha$, where $\Sigma = e^\top e/(\tau - K - 1)\ is$ theresidual covariance matrix from regressions with K factors, N assets, and τ time observations is represented. The maximum squared Sharpe ratio of the model's factors is $Sh^2(f) = \bar{f}^\top \Omega^{-1} \bar{f}$, where $\bar{f}$ is the average factor return and $\Omega = \left(f - \bar{f}\right)^\top \left(f - \bar{f}\right)/(\tau - 1)$ is the factor return covariance matrix. The GRS statistic is calculated as $F_{GRS} = \frac{\tau(\tau-N-K)}{N(\tau-K-1)} \frac{Sh^2(\alpha)}{1+Sh^2(f)}$, which effectively evaluates the ratio of unexplained return variation (alphas) to explained variation (factors) and follows a F distribution with N and τ − N − K degrees of freedom.

The objective in model selection is to minimize $Sh^2(\alpha)$, thereby ensuring that systematic components capture most of the return variation. The null hypothesis of the GRS test is that the alphas of all test assets are precisely zero. Suppose the GRS test statistic exceeds the critical value from the F-distribution at a given significance level. The null hypothesis is rejected, indicating that the factor model does not adequately explain the variation in test asset returns.

In addition to the main iterative selection procedure, we perform several robustness analyses to confirm that our findings are not the result of particular methodological choices. First, we examine how the results vary using different portfolio construction methods, such as alternative sorting depths (terciles, quartiles, and quintiles) and weighting schemes (equal versus value weighting). Second, to test the temporal stability of our factor selection, we split the sample into subperiods and repeated the analysis. Third, we investigate factor clustering and correlations to understand the economic relationships between the selected factors.

3. Empirical Findings

3.1. Descriptive Statistics for Long-Short Portfolios

Table 3 shows the performance of long-short portfolios formed using individual cryptocurrency characteristics. The predictability of cryptocurrency cross-sectional returns varies substantially across factor categories, with notable differences between equal- and value-weighted portfolio constructions. Several factors exhibit statistically significant predictive power, though their strength and direction depend on the weighting scheme. Among the top performers in equal-weighted portfolios, risk-based, behavioral, and liquidity-related variables dominate, while momentum and on-chain measures become more relevant in value-weighted portfolios.

[Insert Table 3 about here]

In equal-weighted portfolios, risk and behavioral factors are the most robust predictors of return. Idiosyncratic volatility has the highest t-statistic (3.39), followed by realized volatility and skewness. Salience theory (t = 3.19) and maximum and minimum daily returns also display strong statistical significance, indicating that attention biases and extreme return characteristics play notable roles. Liquidity factors, such as turnover volatility (t = 3.14), bid–ask spread (t = 2.19), and turnover (t = 2.07), are significant as well. However, their effects weaken or reverse in value-weighted portfolios, suggesting that liquidity effects strongly interact with size.

Value-weighted portfolios reveal the predictive power of momentum and blockchain-native metrics. The seven-day momentum factor achieves a t-statistic of 3.46 in value-weighted portfolios, compared to 1.97 in equal-weighted ones. This difference underscores a size-related distinction in trend-following behavior. On-chain variables, particularly the new address-to-price ratio (t = 2.97) and active addresses (t = 2.32), are also significant, especially for larger, more established tokens. In contrast, the crypto market factor yields a weekly return of 0.72% (t = 1.31), which is underperformed by many cross-sectional strategies.

3.2. Main Results: Iterative Factor Selection

Table 4 presents our key findings. Our iterative selection procedure reveals that only a few economically interpretable factors are needed to eliminate significant alphas in the cryptocurrency factor zoo. However, the composition of these factors differs notably between equal- and value-weighted portfolios. Equal-weighted strategies emphasize liquidity risk and behavioral anomalies, whereas value-weighted strategies prioritize trading frictions and momentum. These discrepancies reveal the structural segmentation of the crypto market, in which smaller tokens are more affected by attention-based and illiquidity effects. In comparison, larger tokens exhibit cost-related and trend-following behavior.

[Insert Table 4 about here]

Panel A shows the equal-weighted results. The first iteration identified turnover volatility as the most powerful factor in explaining cross-sectional alphas. Adding this liquidity risk measure to the market factor reduces unexplained anomalies to 10 ($t > 3$). Although the GRS statistic of 2.15 remains highly significant ($p < 0.001$), the substantial improvement demonstrates the important role of liquidity risk in cryptocurrency pricing. The average absolute alpha across the remaining factors decreases to 28 basis points per week, with a maximum of 66 basis points. The second iteration points to the salience theory variable as the next critical factor. The GRS statistic decreases to 1.67 ($p = 0.01$), and only three factors have t-statistics greater than 1.96. The third selected factor is the new address-to-price ratio, a unique on-chain measure for cryptocurrencies. Although all highly significant alphas have been eliminated, this factor improves the model's fit even more, reducing the GRS statistic to 1.47 (p = 0.05). Currently, only two factors remain marginally significant ($t > 1.96$), and the average weekly absolute alpha has stabilized at around 14 basis points.

The procedure continues with the bid-ask spread entering fourth, providing the first instance in which the GRS p-value exceeds our 10% threshold ($p = 0.11$). Selecting another liquidity measure despite already including turnover volatility suggests that these variables capture different aspects of trading frictions. Subsequent factors—new addresses, change in chronological return ordering, and intermediate momentum—continue to marginally improve model fit, with the GRS statistic eventually reaching 0.75 ($p = 0.81$) after ten factors.

Panel B displays a different pattern for value-weighted portfolios. Here, the bid-ask spread emerges as the primary factor, reducing significant alphas ($t > 3$) from the initial count of one to zero. The value-weighted procedure subsequently selects 7-day momentum, a sharp contrast to the weak performance of momentum in equal-weighted univariate sorts. This factor achieves a GRS p-value of 0.05, eliminating the last highly significant alpha. Skewness appears third in the value-weighted specification and represents distributional characteristics absent from the equal-weighted selection. With a GRS p-value of 0.15, skewness exceeds our significance threshold; however, the procedure continues to reveal the complete factor hierarchy. Subsequent factors include kurtosis, long-term reversal, and various liquidity measures.

Figure 2 exhibits the annualized average returns (in percent) for various long-short strategies categorized by economic rationale, such as risk, liquidity, past returns, distribution, on-chain measures, market factors, and other factors, as well as the crypto market portfolio (in black). Several factors that were important in univariate analyses do not appear in our multivariate selection. Despite their strong individual performance, risk measures such as idiosyncratic volatility and realized volatility appear redundant once liquidity and behavioral factors are included.

[Insert Figure 2 about here]

3.3. Alternative Research Design Choices

We examine how our results vary based on different methodological choices. This analysis determines whether our main findings reflect robust economic phenomena or methodological artifacts. We consider six specifications that vary in weighting schemes (equal versus capitalization) and sorting depths (terciles, quartiles, and quintiles). Figure 3 illustrates how our factor selection results are affected by alternative methodological choices.

[Insert Figure 3 about here]

The GRS statistics in the top-left panel show consistent patterns across all specifications. All six curves monotonically decline as factors are added, though at different rates. The equal-weighted specifications (green, blue, and red lines) cluster together and exhibit steeper initial declines. In contrast, the capitalization-weighted specifications (orange, purple, and yellow) demonstrate more gradual improvements. The p-value trajectories in the top-middle panel provide clearer evidence of when each specification achieves statistical adequacy. Using the $t > 3$ threshold in the bottom-right panel, the number of significant anomalies drops precipitously from eight to nine to nearly zero within just two to three factors across all specifications. The less stringent $t > 2$ threshold in the top-right panel presents more variation, with some specifications exhibiting small resurgences in significant factors after the initial decline. The adjusted squared Sharpe ratios (bottom-middle) increase monotonically for all specifications, which confirms that each selected factor contributes to the mean-variance efficiency of the factor model. The similar slopes of the curves suggest that marginal factor contributions remain stable across methodological choices. Capitalization-weighted specifications achieve slightly lower maximum Sharpe ratios, reflecting the difficulty of generating alpha in more liquid and efficiently priced cryptocurrencies.

[Insert Table 5 about here]

Table 5 demonstrates which factors emerge across the different specifications. Although the identity of the selected factors varies considerably—turnover volatility appears in equal-weighted specifications, whereas bid-ask spread dominates capitalization-weighted ones—clear patterns emerge at the factor group level. Every specification includes at least one liquidity measure and on-chain measures. Liquidity risk manifests through multiple channels—such as turnover volatility, bid-ask spreads, and volume shocks—but remains consistently important. Similarly, behavioral biases and network effects persist across specifications, even as their specific manifestations change (addresses with balance growth, new addresses, the new address-to-price ratio, or zero-balance addresses). Past returns are especially important factors in value-weighted specifications, such as 7-day momentum and long-term reversal.

3.4. Subperiod Analysis

To determine the stability of our factor selection results over time, we split our sample into halves and repeated the iterative selection procedure for each subperiod. Table 6 reports the results. This analysis addresses potential structural changes in the cryptocurrency market as it has matured and evolved through different market regimes, including the bull market of 2020–2021 and the subsequent correction.

[Insert Table 6 about here]

For equally weighted portfolios (Panel A), turnover volatility, chronological return ordering, trading volume, and market value appear in the top 10 of both periods. The first period (2018–2021) emphasizes extreme return characteristics, with minimum and maximum daily returns ranking first and third, respectively. The second period (2021-2024) unveils a shift toward market microstructure variables. Turnover volatility rises to first place, and the bid-ask spread enters second. This evolution suggests that liquidity risk became the dominant concern as the cryptocurrency market matured, potentially reflecting increased institutional participation and more sophisticated trading strategies.

The value-weighted results (Panel B) exhibit an even greater degree of historical variation. The first period displays a preference for traditional risk measures and distributional characteristics. The top performers are 7-day momentum, followed by idiosyncratic volatility, kurtosis, and skewness. This focus on the statistical properties of returns suggests that large-cap cryptocurrencies were heavily influenced by risk and behavioral factors during this period. The value-weighted results of the second period shift toward liquidity measures. Bid-ask spread ranks first, and trading volume volatility, trading volume, and the illiquidity ratio appear in the top five.

The limited overlap between periods, with only one factor persisting in each panel, raises important questions about the stability of cryptocurrency factor structures. There are several possible interpretations. First, as cryptocurrency markets evolved from a retail-dominated, speculative market to one with significant institutional presence, they may have undergone fundamental structural changes. The shift from distribution-based to liquidity-based factors supports this interpretation. Second, the changing landscape of factors may reflect the dynamic nature of the cryptocurrency ecosystem itself. The first period captured the aftermath of the ICO boom and the early emergence of DeFi, while the second period encompassed the rise of NFT markets, layer-2 scaling solutions, and regulatory clarity in major jurisdictions. Third, the instability may suggest that our relatively short subperiods (approximately three years each) are too short to identify persistent factors. Factors that show statistical significance in one period may represent temporary market inefficiencies rather than permanent risk factors.

Despite the instability of specific factors, patterns at the group level are more consistent. Liquidity-related factors are prominent in both periods, though they manifest differently. Behavioral factors are present through chronological return ordering and various momentum measures. This group-level stability contrasts with the instability of individual factors and reinforces our main finding. Additionally, factor selection in the early period appears more idiosyncratic, with unusual factors such as minimum daily returns and kurtosis playing significant roles. The later period emphasizes conventional liquidity and momentum.

4. Concluding remarks

This paper presents the first systematic application of factor zoo compression to cryptocurrency markets. By analyzing 36 factors across 565 cryptocurrencies from 2018 to 2024, we iteratively expanded a CAPM baseline by adding the factor that most reduced the GRS statistic. This process continued until all alphas lost significance. A compact set of coherent factors effectively explains cross-sectional returns.

For equal-weighted portfolios, key dimensions include liquidity (turnover volatility), investor biases (salience theory), and network activity (the new-address-to-price ratio). These are complemented by bid-ask spread, return ordering, and momentum. For value-weighted portfolios, residual anomalies are explained by bid-ask spread, seven-day momentum, and higher moments (skewness and kurtosis).

Three main insights emerge. First, liquidity and trading-cost proxies consistently outweigh traditional risk variables across weighting schemes and samples. Second, blockchain-native signals—such as address-based activity—retain explanatory power beyond market and microstructure effects. Third, factor themes remain cross-sectionally stable, though specific factor survival varies over time, reflecting the adaptive nature of crypto markets.

Compressed models have practical value for investors and risk managers. First, the compressed factor models allow for constructing sparse asset-pricing models that reduce noise and overfitting. Investors seeking to develop factor-based strategies in the cryptocurrency market can prioritize liquidity proxies, such as turnover volatility and bid-ask spreads, attention-related anomalies, and on-chain adoption metrics. Second, the unstable temporal persistence of individual factors suggests against relying on static models and supports adaptive investment frameworks that update factor relevance over time. Third, our approach provides insight into designing more robust benchmarking tools and smart beta indices tailored to the crypto asset class.

Of course, our study has limitations. The sample period is from 2018 to 2024, a brief but formative period in the history of cryptocurrency. The study concludes prior to the shift toward real-world asset tokens and spot ETF inflows that occurred after 2024. Factor payoffs remain volatile, and some characteristics (e.g., turnover volatility) entail high trading costs that could reduce net performance outside major exchanges. Additionally, the GRS-based procedure only considers linear relationships; therefore, it is possible that nonlinear or higher-order interactions could resurrect otherwise redundant signals.

References

Abdi, F., & Ranaldo, A. (2017). A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices. *The Review of Financial Studies*, *30*(12), 4437–4480. https://doi.org/10.1093/rfs/hhx084

Ahmed, S., Bu, Z., Symeonidis, L., & Tsvetanov, D. (2023). Which factor model? A systematic return covariation perspective. *Journal of International Money and Finance*, *136*, 102865. https://doi.org/10.1016/j.jimonfin.2023.102865

Alexander, C., & Dakos, M. (2020). A critical investigation of cryptocurrency data and analysis. *Quantitative Finance*, *20*(2), 173–188. https://doi.org/10.1080/14697688.2019.1641347

Anatolyev, S., & Mikusheva, A. (2022). Factor models with many assets: Strong factors, weak factors, and the two-pass procedure. *Journal of Econometrics*, *229*(1), 103–126. https://doi.org/10.1016/j.jeconom.2021.01.002

Babiak, M., & Bianchi, D. (2024). *Mispricing and Risk Compensation in Cryptocurrency Returns*. Available at SSRN: https://ssrn.com/abstract=3935934 or http://dx.doi.org/10.2139/ssrn.3935934

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, *99*(2), 427–446. https://doi.org/10.1016/j.jfineco.2010.08.014

Barillas, F., & Shanken, J. (2017). Which Alpha? *The Review of Financial Studies*, *30*(4), 1316–1338. https://doi.org/10.1093/rfs/hhw101

Bhambhwani, S. M., Delikouras, S., & Korniotis, G. M. (2023). Blockchain characteristics and cryptocurrency returns. *Journal of International Financial Markets, Institutions and Money*, *86*, 101788. https://doi.org/10.1016/j.intfin.2023.101788

Bianchi, D., & Babiak, M. (2022). On the performance of cryptocurrency funds. *Journal of Banking & Finance*, *138*, 106467. https://doi.org/10.1016/j.jbankfin.2022.106467

Bianchi, D., Babiak, M., & Dickerson, A. (2022). Trading volume and liquidity provision in cryptocurrency markets. *Journal of Banking & Finance*, *142*, 106547. https://doi.org/10.1016/j.jbankfin.2022.106547

Borri, N. (2019). Conditional tail-risk in cryptocurrency markets. *Journal of Empirical Finance*, *50*, 1–19. https://doi.org/10.1016/j.jempfin.2018.11.002

Borri, N., & Shakhnov, K. (2022). The Cross-Section of Cryptocurrency Returns. *The Review of Asset Pricing Studies*, *12*(3), 667–705. https://doi.org/10.1093/rapstu/raac007

Brauneis, A., Mestel, R., Riordan, R., & Theissen, E. (2021). How to measure the liquidity of cryptocurrency markets? *Journal of Banking & Finance*, *124*, 106041. https://doi.org/10.1016/j.jbankfin.2020.106041

Burggraf, T., & Rudolf, M. (2021). Cryptocurrencies and the low volatility anomaly. *Finance Research Letters*, *40*, 101683. https://doi.org/10.1016/j.frl.2020.101683

Cai, C. X., & Zhao, R. (2024). Salience theory and cryptocurrency returns. *Journal of Banking & Finance*, *159*, 107052. https://doi.org/10.1016/j.jbankfin.2023.107052

Cakici, N., Shahzad, S. J. H., Będowska-Sójka, B., & Zaremba, A. (2024). Machine learning and the cross-section of cryptocurrency returns. *International Review of Financial Analysis*, *94*, 103244. https://doi.org/10.1016/j.irfa.2024.103244

Chang, H.-L., Nie, W.-Y., Chang, L.-H., Cheng, H.-W., & Yen, K.-C. (2023). Cryptocurrency

Momentum and VIX premium. *Finance Research Letters*, *57*, 104196. https://doi.org/10.1016/j.frl.2023.104196

Chen, A. Y., & Zimmermann, T. (2021). Open Source Cross-Sectional Asset Pricing. *Finance and Economics Discussion Series*, *2021.0*(37), 1–66. https://doi.org/10.17016/feds.2021.037

Chen, R., Lepori, G. M., Tai, C.-C., & Sung, M.-C. (2022). Can salience theory explain investor behaviour? Real-world evidence from the cryptocurrency market. *International Review of Financial Analysis*, *84*, 102419. https://doi.org/10.1016/j.irfa.2022.102419

Chinco, A., Hartzmark, S. M., & Sussman, A. B. (2022). A New Test of Risk Factor Relevance. *The Journal of Finance*, *77*(4), 2183–2238. https://doi.org/10.1111/jofi.13135

Cochrane, J. H. (2011). Presidential Address: Discount Rates. *The Journal of Finance*, *66*(4), 1047–1108. https://doi.org/10.1111/j.1540-6261.2011.01671.x

Cong, L., Karolyi, G. A., Tang, K., & Zhao, W. (2022). Value Premium, Network Adoption, and Factor Pricing of Crypto Assets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3985631

Corwin, S. A., & Schultz, P. (2012). A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices. *The Journal of Finance*, *67*(2), 719–760. https://doi.org/10.1111/j.1540-6261.2012.01729.x

Cosemans, M., & Frehen, R. (2021). Salience theory and stock prices: Empirical evidence. *Journal of Financial Economics*, *140*(2), 460–483. https://doi.org/10.1016/j.jfineco.2020.12.012

De Nard, G., & Zhao, Z. (2023). Using, taming or avoiding the factor zoo? A double-shrinkage estimator for covariance matrices. *Journal of Empirical Finance*, *72*, 23–35. https://doi.org/10.1016/j.jempfin.2023.02.003

Dobrynskaya, V. (2023). Cryptocurrency Momentum and Reversal. *The Journal of Alternative Investments*, *26*(1), 65–76. https://doi.org/10.3905/jai.2023.1.189

Dobrynskaya, V. (2024). Is downside risk priced in cryptocurrency market? *International Review of Financial Analysis*, *91*, 102947. https://doi.org/10.1016/j.irfa.2023.102947

Dobrynskaya, V., & Dubrovskiy, M. (2023). Cryptocurrencies Meet Equities: Risk Factors and Asset-pricing Relationships. In S.-J. Kim (Ed.), *International Finance Review* (pp. 95–111). Emerald Publishing Limited. https://doi.org/10.1108/S1569-376720220000022006

Dong, X., Li, Y., Rapach, D. E., & Zhou, G. (2022). Anomalies and the Expected Market Return. *The Journal of Finance*, *77*(1), 639–681. https://doi.org/10.1111/jofi.13099

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E. F., & French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, *128*(2), 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012

Fassas, A. P., Papadamou, S., & Koulis, A. (2020). Price discovery in bitcoin futures. *Research in International Business and Finance*, *52*, 101116. https://doi.org/10.1016/j.ribaf.2019.101116

Feng, G., Giglio, S., & Xiu, D. (2020). Taming the Factor Zoo: A Test of New Factors. *The Journal of Finance*, *75*(3), 1327–1370. https://doi.org/10.1111/jofi.12883

Fieberg, C., Liedtke, G., & Zaremba, A. (2024). Cryptocurrency anomalies and economic constraints. *International Review of Financial Analysis*, *94*, 103218. https://doi.org/10.1016/j.irfa.2024.103218

Garfinkel, J. A., Hsiao, L., & Hu, D. (2023). Disagreement and the Cross Section of

Cryptocurrency Returns. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4345640

Gemayel, R., Franus, T., & Bowden, J. (2023). Price discovery between Bitcoin spot markets and exchange traded products. *Economics Letters*, *228*, 111152. https://doi.org/10.1016/j.econlet.2023.111152

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A Test of the Efficiency of a Given Portfolio. *Econometrica*, *57*(5), 1121. https://doi.org/10.2307/1913625

Grobys, K., & Junttila, J. (2021). Speculation and lottery-like demand in cryptocurrency markets. *Journal of International Financial Markets, Institutions and Money*, *71*, 101289. https://doi.org/10.1016/j.intfin.2021.101289

Grobys, K., & Sapkota, N. (2019). Cryptocurrencies and momentum. *Economics Letters*, *180*, 6–10. https://doi.org/10.1016/j.econlet.2019.03.028

Han, S. (2023). Is liquidity risk priced in cryptocurrency markets? *Applied Economics Letters*, *30*(17), 2481–2487. https://doi.org/10.1080/13504851.2022.2098235

Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the Cross-Section of Expected Returns. *Review of Financial Studies*, *29*(1), 5–68. https://doi.org/10.1093/rfs/hhv059

Hoang, L. T., & Baur, D. G. (2022). Loaded for bear: Bitcoin private wallets, exchange reserves and prices. *Journal of Banking & Finance*, *144*, 106622. https://doi.org/10.1016/j.jbankfin.2022.106622

Hou, K., Xue, C., & Zhang, L. (2020). Replicating Anomalies. *The Review of Financial Studies*, *33*(5), 2019–2133. https://doi.org/10.1093/rfs/hhy131

Jensen, T. I., Kelly, B., & Pedersen, L. H. (2023). Is There a Replication Crisis in Finance? *The Journal of Finance*, jofi.13249. https://doi.org/10.1111/jofi.13249

Jia, Y., Liu, Y., & Yan, S. (2021). Higher moments, extreme returns, and cross–section of cryptocurrency returns. *Finance Research Letters*, *39*, 101536. https://doi.org/10.1016/j.frl.2020.101536

Keloharju, M., Linnainmaa, J. T., & Nyberg, P. (2016). Return Seasonalities: Return Seasonalities. *The Journal of Finance*, *71*(4), 1557–1590. https://doi.org/10.1111/jofi.12398

Kozak, S., Nagel, S., & Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, *135*(2), 271–292. https://doi.org/10.1016/j.jfineco.2019.06.008

Lan, T., & Frömmel, M. (2025). Risk factors in cryptocurrency pricing. *International Review of Financial Analysis*, *105*, 104389. https://doi.org/10.1016/j.irfa.2025.104389

Leirvik, T. (2022). Cryptocurrency returns and the volatility of liquidity. *Finance Research Letters*, *44*, 102031. https://doi.org/10.1016/j.frl.2021.102031

Lettau, M., & Pelger, M. (2020). Factors That Fit the Time Series and Cross-Section of Stock Returns. *The Review of Financial Studies*, *33*(5), 2274–2325. https://doi.org/10.1093/rfs/hhaa020

Li, Y., Urquhart, A., Wang, P., & Zhang, W. (2021). MAX momentum in cryptocurrency markets. *International Review of Financial Analysis*, *77*, 101829. https://doi.org/10.1016/j.irfa.2021.101829

Liebi, L. J. (2022). Is there a value premium in cryptoasset markets? *Economic Modelling*, *109*, 105777. https://doi.org/10.1016/j.econmod.2022.105777

Lin, C.-H., Yen, K.-C., & Cheng, H.-P. (2021). Lottery-like momentum in the cryptocurrency market. *The North American Journal of Economics and Finance*, *58*, 101552.

https://doi.org/10.1016/j.najef.2021.101552
Liu, R., Wan, S., Zhang, Z., & Zhao, X. (2020). Is the introduction of futures responsible for the crash of Bitcoin? *Finance Research Letters*, *34*, 101259. https://doi.org/10.1016/j.frl.2019.08.007
Liu, W., Liang, X., & Cui, G. (2020). Common risk factors in the returns on cryptocurrencies. *Economic Modelling*, *86*, 299–305. https://doi.org/10.1016/j.econmod.2019.09.035
Liu, Y., & Tsyvinski, A. (2021). Risks and Returns of Cryptocurrency. *The Review of Financial Studies*, *34*(6), 2689–2727. https://doi.org/10.1093/rfs/hhaa113
Liu, Y., Tsyvinski, A., & Wu, X. (2021). Accounting for Cryptocurrency Value. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3951514
Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common Risk Factors in Cryptocurrency. *The Journal of Finance*, *77*(2), 1133–1177. https://doi.org/10.1111/jofi.13119
Lucchini, L., Alessandretti, L., Lepri, B., Gallo, A., & Baronchelli, A. (2020). From code to market: Network of developers and correlated returns of cryptocurrencies. *Science Advances*, *6*(51), eabd2204. https://doi.org/10.1126/sciadv.abd2204
Mclean, R. D., & Pontiff, J. (2016). Does Academic Research Destroy Stock Return Predictability?: Does Academic Research Destroy Stock Return Predictability? *The Journal of Finance*, *71*(1), 5–32. https://doi.org/10.1111/jofi.12365
Mercik, A., Będowska-Sójka, B., Karim, S., & Zaremba, A. (2025). Cross-sectional interactions in cryptocurrency returns. *International Review of Financial Analysis*, *97*, 103809. https://doi.org/10.1016/j.irfa.2024.103809
Miller, M. H., & Scholes, M. S. (1982). Dividends and Taxes: Some Empirical Evidence. *Journal of Political Economy*, *90*(6), 1118–1141. https://doi.org/10.1086/261114
Mohrschladt, H. (2021). The ordering of historical returns and the cross-section of subsequent returns. *Journal of Banking & Finance*, *125*, 106064. https://doi.org/10.1016/j.jbankfin.2021.106064
Ozdamar, M., Akdeniz, L., & Sensoy, A. (2021). Lottery-like preferences and the MAX effect in the cryptocurrency market. *Financial Innovation*, *7*(1), 74. https://doi.org/10.1186/s40854-021-00291-9
Pagnotta, E., & Buraschi, A. (2018). An Equilibrium Valuation of Bitcoin and Decentralized Network Assets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3142022
Sakkas, A., & Urquhart, A. (2024). Blockchain factors. *Journal of International Financial Markets, Institutions and Money*, *94*, 102012. https://doi.org/10.1016/j.intfin.2024.102012
Shen, D., Urquhart, A., & Wang, P. (2020). A three-factor pricing model for cryptocurrencies. *Finance Research Letters*, *34*, 101248. https://doi.org/10.1016/j.frl.2019.07.021
Swade, A., Hanauer, M. X., Lohre, H., & Blitz, D. (2024). Factor Zoo (.zip). *The Journal of Portfolio Management*, *50*(3), 11–31. https://doi.org/10.3905/jpm.2023.1.561
Tzouvanas, P., Kizys, R., & Tsend-Ayush, B. (2020). Momentum trading in cryptocurrencies: Short-term returns and diversification benefits. *Economics Letters*, *191*, 108728. https://doi.org/10.1016/j.econlet.2019.108728
Vidal-Tomás, D. (2022). Which cryptocurrency data sources should scholars use? *International Review of Financial Analysis*, *81*, 102061. https://doi.org/10.1016/j.irfa.2022.102061
Younus, M., & Naeem, M. A. (2024). Crash Risk and the Cross-Section of Cryptocurrency Returns. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4823589

Zaremba, A., Bilgin, M. H., Long, H., Mercik, A., & Szczygielski, J. J. (2021). Up or down? Short-term reversal, momentum, and liquidity effects in cryptocurrency markets. *International Review of Financial Analysis*, *78*, 101908. https://doi.org/10.1016/j.irfa.2021.101908

Zhang, W., & Li, Y. (2023). Liquidity risk and expected cryptocurrency returns. *International Journal of Finance & Economics*, *28*(1), 472–492. https://doi.org/10.1002/ijfe.2431

Zhang, W., Li, Y., Xiong, X., & Wang, P. (2021). Downside risk and the cross-section of cryptocurrency returns. *Journal of Banking & Finance*, *133*, 106246. https://doi.org/10.1016/j.jbankfin.2021.106246

Zhang, Z., & Zhao, R. (2023). Good volatility, bad volatility, and the cross section of cryptocurrency returns. *International Review of Financial Analysis*, *89*, 102712. https://doi.org/10.1016/j.irfa.2023.102712

Zhao, X., Wang, Y., & Liu, W. (2024). Someone like you: Lottery-like preference and the cross-section of expected returns in the cryptocurrency market. *Journal of International Financial Markets, Institutions and Money*, *91*, 101957. https://doi.org/10.1016/j.intfin.2024.101957

Figure 1. A Snapshot of the Research Sample

The diagram offers a view of the cryptocurrency sample utilized in our analysis. In Panel A, the market capitalization of the coins within our dataset is displayed in billions. Panel B shows the daily count of coins that have a complete dataset for strategy analysis. The dataset encompasses a total of 565 distinct cryptocurrencies, and the analysis period spans from January 1, 2018, to July 25, 2024.

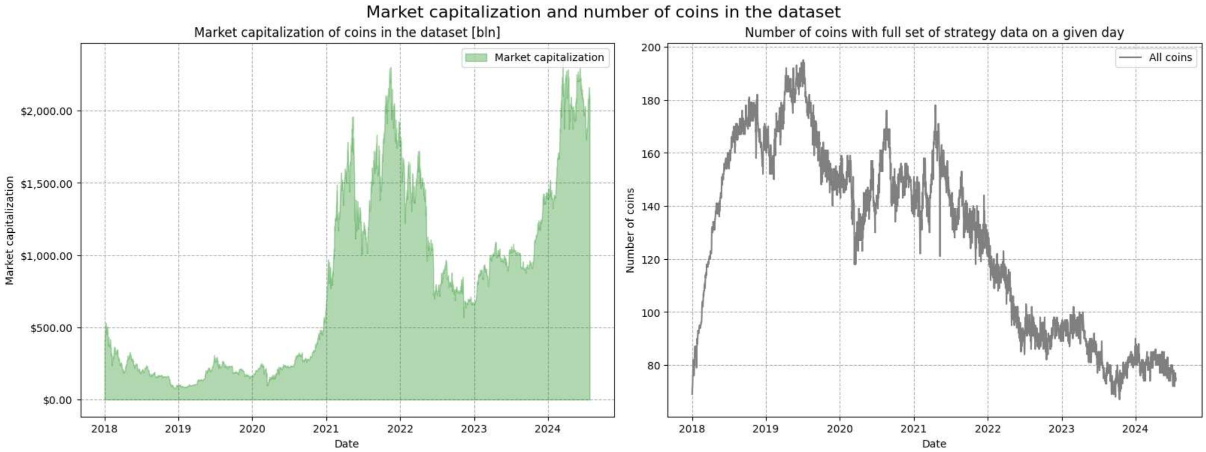
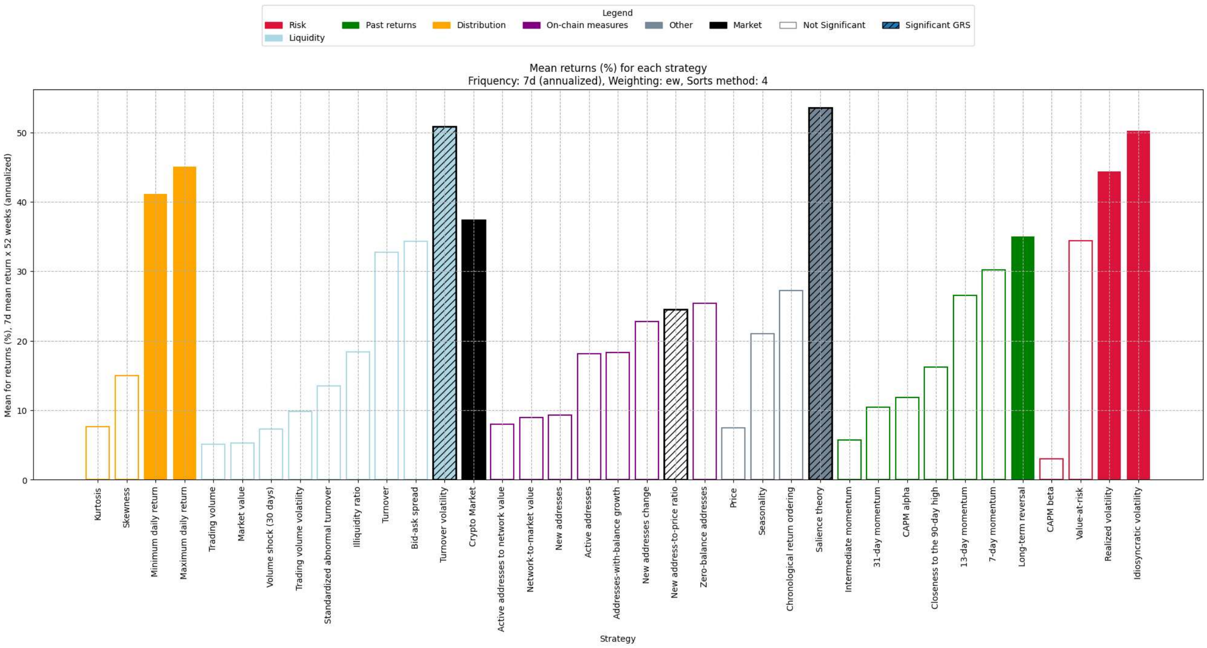


Figure 2. Mean Returns of Crypto Factors and Market Portfolio.

The figure shows the annualized average returns (in percent) for various factors (long-short strategies, which are categorized by their economic rationale: risk, liquidity, past returns, distribution, on-chain measures, the market, and other factors) and the crypto market portfolio (in black). The analysis uses a seven-day frequency. For each quartile, we calculate equally weighted returns (Panel A) and value-weighted returns, meaning we weight stocks by their market capitalization (Panel B). Strategies identified as statistically significant in univariate analyses under Benjamini and Hochberg's (1995) multiple-testing framework are indicated by filled bars, while those that are not significant are indicated by open bars. Factors identified as significant by the GRS test are marked by dashed bars. The sample includes 565 cryptocurrencies from January 1, 2018, to July 25, 2024.

Panel A: Equal-weighted



Panel B: Value-weighted

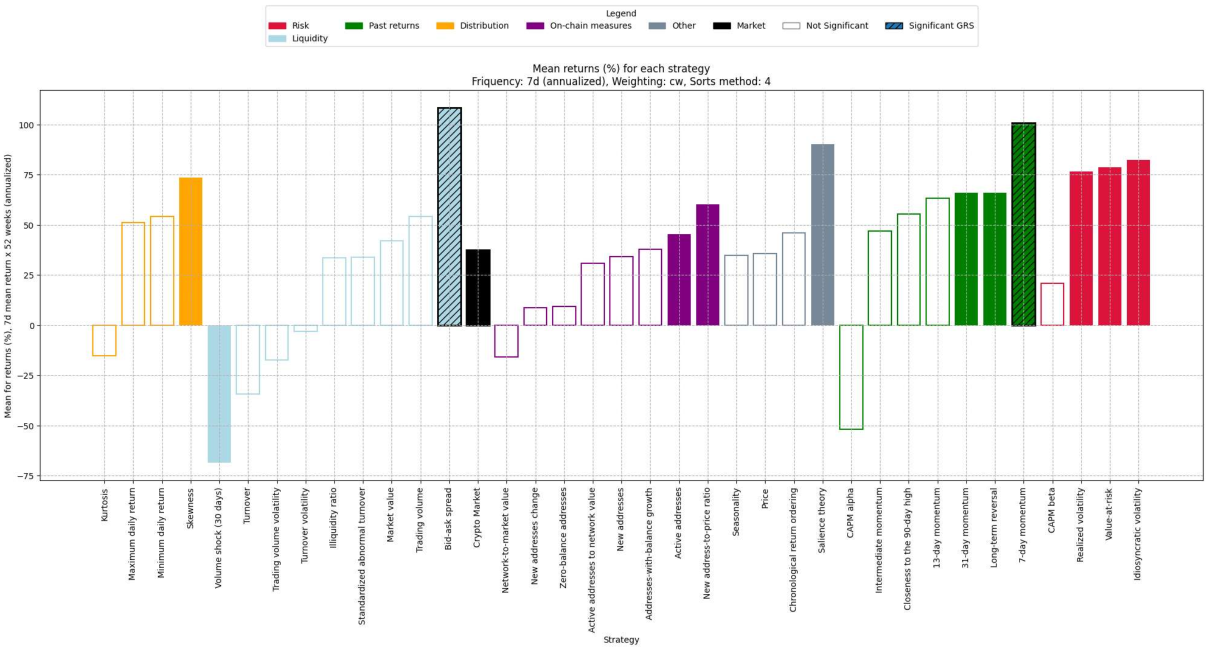


Figure 3. Robustness of Factor Selection Across Alternative Specifications

This figure displays the iterative factor selection results across six specifications varying weighting schemes (equal-weighted vs. capitalization-weighted) and sorting depths (terciles, quartiles, and quintiles). The top panels show the GRS statistic, p-value, and the number of factors with t-statistics exceeding 1.96. The bottom panels display average absolute alphas, adjusted squared Sharpe ratios, and the number of factors with t-statistics exceeding 3.00. Horizontal dashed lines in the p(GRS) panel indicate 5% and 10% significance levels. All specifications use weekly rebalancing with zero transaction costs. The sample includes 565 cryptocurrencies from January 1, 2018, to July 25, 2024.

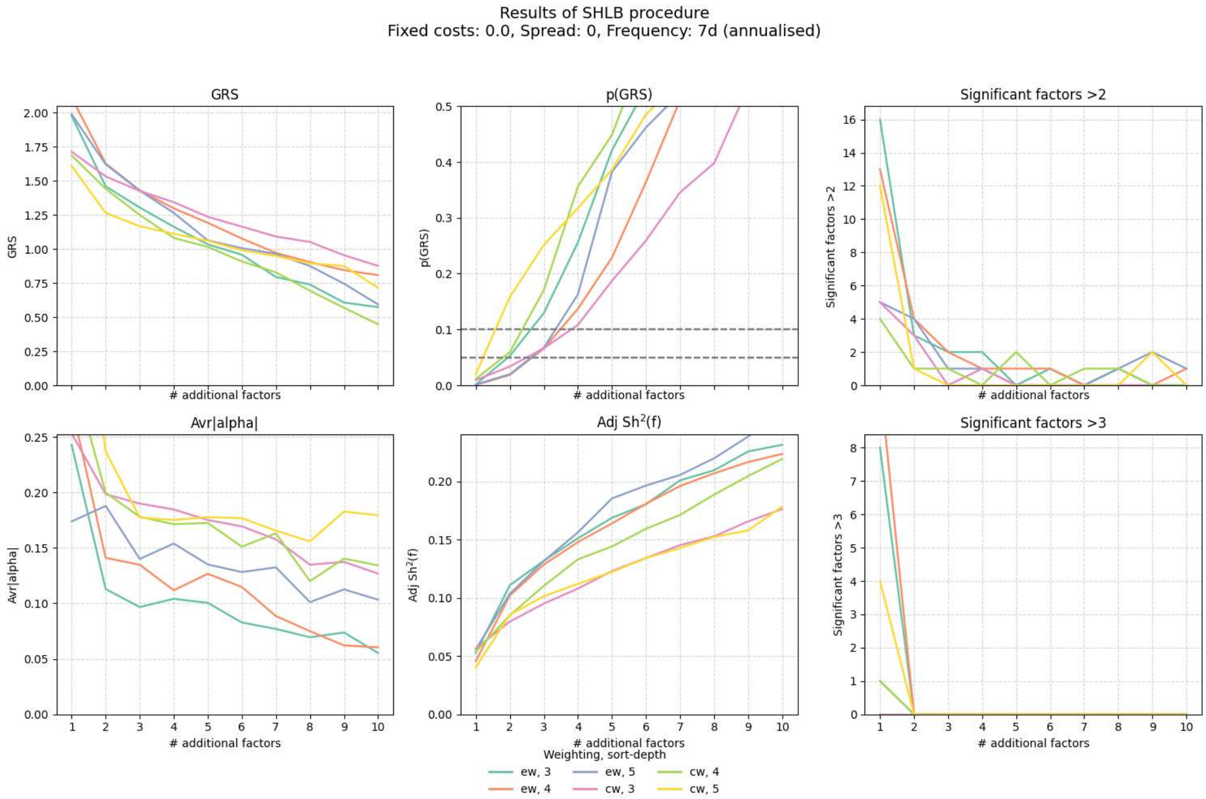


Table 1. Cryptocurrency Characteristics

The table presents the 36 cryptocurrency characteristics used as inputs to machine learning models. No. is the running number, and Symbol indicates the acronym used to denote the variable in the paper. The table spans three pages.

No.	Characteristic	Symbol	Definition
		Panel A: On-chain measures	
1	New addresses	*new_add*	The logged number of new addresses per coin, where a new address refers to a unique address appearing for the first time in a transaction of the native coin in the network (Liu et al., 2022).
2	New addresses change	*new_add_ch*	A weekly change in the logged number of all new addresses per coin over the last week versus the week before, where a new address refers to a unique address appearing for the first time in a transaction of the native coin in the network (Liu et al., 2022).
3	Active addresses	*active_add*	The logged number of unique addresses active in the network, either as a sender or receiver (Pagnotta & Buraschi, 2018). The calculations are limited to the addresses that were active in successful transactions.
4	Zero-balance addresses	*zero_bal*	The logged number of zero-balance addresses (Borri et al., 2022)
5	Addresses-with-balance growth	*ba_growth*	The weekly change in the logged value of addresses with balance (Cong et al., 2022).
6	Network-to-market value	*bm*	As in Pagnotta and Burashi (2018), the network-to-market value is calculated as a logged ratio of the cumulative number of unique addresses to the total market value (see Market value for the calculation details).
7	New address-to-price ratio	*ap*	A logged ratio of new addresses per coin over the last week to the coin market price (Liu et al., 2022).
8	Active addresses to network value	*aanv*	The average number of active addresses over the past 30 days to the total market value (Liebi, 2022).
		Panel B: Liquidity	
9	Trading volume	*volume*	The logged total dollar value of all native tokens transferred across wallets - both across and within centralized exchanges.
10	Market value	*size*	The logged market capitalization calculated by multiplying the total supply of coins or tokens by the current market price of each coin or token. (Liu et al., 2022)
11	Bid-ask spread	*bidask*	An estimation of the bid-spread calculated based on 30 days of OHLC data as a simple average of two approximations by Corwin and Schultz (2012) and Abdi and Ronaldo (2017).
12	Illiquidity ratio	*illiq*	The price impact measure of Amihud (2002) calculated as the average 90-day ratio of the absolute value of daily returns over the daily trading volume measured in USD.
13	Turnover	*turnover*	The last day's dollar trading volume (see Trading volume) over the market value (see Market Value) (Datar et al., 2018).
14	Turnover volatility	*std_dto*	Residuals' standard deviation from a regression of daily turnover on a constant using a 30-day estimation period.
15	Trading volume volatility	*std_vol*	The logged value of residuals' standard deviation from a regression of daily trading volume on a constant using a 30-day estimation period.

16	Standardized abnormal turnover	*sat*	Following Garfinkel et al. (2023), standardized abnormal turnover is calculated in two steps. First, we calculate the change in turnover as the difference between the last day's turnover (see Turnover for definition) and its average over the prior 30 days. Second, we divide the outcome by the 30-day standard deviation of daily turnover values.
17	Volume shock (30 days)	*volsh_30d*	Log-deviation of trading volume from its rolling 30-day average, as in Llorente et al. (2002) and Babiak et al. (2022).
		Panel C: Risk	
18	Realized volatility	*rvol*	Daily realized volatility calculated based on 30 days of OHLC prices using the estimator of Yang and Zhang (2000).
19	CAPM beta	*beta*	The market beta from the Capital Asset Pricing Model estimated using a trailing 30- day period. As in Levellen and Nagel (2006), the beta is calculated as the sum of two slope coefficients from the regression of daily cryptocurrency returns on the contemporaneous and one-day-lagged market excess returns. The market portfolio return is the value-weighted average return of all cryptocurrencies in the sample.
20	Idiosyncratic volatility	*ivol*	The standard deviation of the residuals from the regression of daily excess cryptocurency returns on the daily market portfolio excess returns estimated using a trailing volatility 30-day period. The market portfolio return is the value-weighted average return of all cryptocurrencies in the sample.
21	Value-at-risk	*var*	The historical empirical value-at-risk computed as the 5th percentile of daily returns over a rolling 90-day period.
		Panel D: Past returns	
22	7-day momentum	*r7_2*	Cumulative return from seven to two days before return prediction.
23	13-day momentum	*r13_2*	Cumulative return from 13 to two days before return prediction.
24	31-day momentum	*r31_2*	Cumulative return from 31 to two days before return prediction.
25	Intermediate momentum	*r30_14*	Cumulative return from 30 to 14 days before return prediction.
26	Long-term reversal	*r180_60*	Cumulative return from 360_31 days before return prediction (Cong et al., 2022).
27	Closeness to the 90-day high	*90dh*	Following the logic of George and Hwang (2004), the closeness to the 90-day high is the last day's price over the maximum price over the previous 90 days. The estimation period follows x and Bianchi (2022).
28	CAPM alpha	*alpha*	An intercept from the regression of daily excess cryptocurrency returns on the daily market portfolio excess returns based on a trailing 30-day period. The market portfolio return is the value-weighted average return on all cryptocurrencies in the sample.
		Panel E: Distribution	
29	Skewness	*skew*	Skewness of the daily return distribution calculated over a rolling 90-day period.
30	Kurtosis	*kurt*	Kurtosis of the daily return distribution calculated over a rolling 90-day period.
31	Maximum daily return	*max*	The maximum daily return over the last 30 days.
32	Minimum daily return	*min*	The minimum daily return over the last 30 days.
		Panel F: Other	

33	Salience theory	*st*	The salience theory variable is calculated closely following the multistep procedure in Cosemans and Frehen (2021) using a rolling 30-day estimation period. We use market portfolio return as the reference rate and set the parameters θ=0.1 and δ=0.7.
34	Chronological return ordering	*cro*	As in Mohrschladt (2021), the chronological return ordering variable is calculated as the correlation between daily returns over the last 30 days and the corresponding number of trading days until the end of the rolling 30-day estimation window.
35	Seasonality	*seas*	Average same-weekday return calculated over a rolling 20-week period.
36	Price	*prc*	Logged cryptocurrency price at the end of the previous day.

Table 2. Statistical Properties of Cryptocurrency Returns and Characteristics

The table presents the basic statistical properties of returns, market capitalization, and the 36 cryptocurrency characteristics. The explanation of variable symbols seen in the leftmost column is available in Table 1. The reported values are calculated using a pooled sample of all daily observations. The sample includes 565 cryptocurrencies from January 1, 2018, to July 25, 2024.

	Count	**Mean**	**Std. Dev.**	**Q1**	**Med.**	**Q3**
	Panel A: Returns and market value (log)					
Weekly simple returns	634 492	(0.00)	0.19	(0.09)	(0.01)	0.07
Weekly excess simple returns	634 492	(0.00)	0.19	(0.10)	(0.01)	0.07
Weekly log-returns	634 492	(0.02)	0.19	(0.10)	(0.01)	0.07
Lagged market value (mln USD)	634 492	3 059.3	39 471.5	8.4	36.4	159.6
	Panel B: On-chain measures					
Active addresses	397 118	10.85	2.32	9.76	10.55	11.49
Zero-balance addresses	630 409	7.46	4.14	3.53	8.85	10.19
New addresses	572 082	1.06	0.45	0.78	1.04	1.30
New addresses change	554 660	0.00	0.49	(0.21)	(0.00)	0.21
Adresses-with-balance growth	394 997	0.70	0.07	0.69	0.69	0.69
Network-to-market value	393 474	(6.41)	2.09	(7.49)	(6.24)	(5.13)
New address-to-price ratio	589 114	(12.98)	1.79	(13.99)	(12.87)	(11.84)
Active addressest to network value	605 785	(13.53)	1.88	(14.47)	(13.34)	(12.34)
	Panel C: Liquidity					
Trading volume	625 366	14.25	3.23	12.30	14.38	16.32
Market value (log)	625 366	17.59	2.23	15.96	17.42	18.90
Bid-ask spread	632 557	0.08	0.07	0.05	0.07	0.10
Illiquidity ratio	630 600	0.02	0.04	0.00	0.01	0.02
Turnover	633 532	0.00	0.16	0.00	0.00	0.00
Turnover volatility	625 366	0.11	0.16	0.01	0.05	0.13
Trading volume volatility	625 366	0.01	0.11	(0.04)	(0.00)	0.03
Standardized abnormal turnover	628 776	0.06	0.05	0.03	0.04	0.07
Volume shock (30 days)	630 600	0.06	0.06	0.01	0.03	0.09
	Panel D: Risk					
Realised volatility	632 557	0.08	0.07	0.05	0.07	0.10
CAPM beta	628 776	1.02	0.70	0.66	1.03	1.37
Idiosyncratic volatility	628 776	0.06	0.05	0.03	0.04	0.07
Value-at-risk	633 532	(0.10)	0.05	(0.12)	(0.09)	(0.07)
	Panel E: Past returns					
7-day momentum	630 140	0.00	0.18	(0.08)	(0.01)	0.06
13-day momentum	632 070	0.01	0.28	(0.12)	(0.02)	0.09
31-day momentum	632 795	0.04	0.61	(0.22)	(0.05)	0.14
Intermediate momentum	626 499	0.02	0.37	(0.15)	(0.03)	0.11
Long-term reversal	620 896	0.75	5.47	(0.73)	(0.38)	0.29
Closeness to the 90-day high	624 959	0.64	0.25	0.47	0.65	0.82
CAPM alpha	628 776	0.00	0.01	(0.01)	(0.00)	0.00
	Panel F: Distribution					
Maximum daily return	633 276	0.18	0.16	0.08	0.13	0.22
Minimum daily return	633 276	(0.14)	0.09	(0.17)	(0.12)	(0.08)
Skewness	632 989	0.55	1.19	(0.15)	0.38	1.07
Kurtosis	632 989	4.32	6.10	0.94	2.29	5.15
	Panel G: Other					
Seasonality	631 090	(0.00)	0.03	(0.01)	(0.00)	0.01
Price (log)	625 366	(2.01)	3.88	(4.18)	(2.06)	0.00
Chronological return ordering	628 776	0.00	0.18	(0.12)	0.00	0.12
Salience theory	628 776	(0.01)	0.00	(0.01)	(0.01)	(0.00)

Table 3. Performance of Portfolios

The table shows the performance of 36 zero-investment long-short portfolios, which are sorted by cryptocurrency characteristics and the market portfolio. For each characteristic, we report the full name, direction, number of weekly observations, and performance metrics for both the equal- and value-weighted portfolios. The direction indicates whether the strategy is long in cryptocurrencies with high characteristic values and short in cryptocurrencies with low characteristic values (1), or short in cryptocurrencies with high characteristic values and long in cryptocurrencies with low characteristic values (-1). Mean returns and standard deviations are expressed as weekly percentages. T-statistics are computed using Newey-West (1987) standard errors. Significant values under Benjamini and Hochberg's (1995) multiple-testing framework are underlined and bolded. The sample includes 565 cryptocurrencies from January 1, 2018, to July 25, 2024. Portfolios are rebalanced weekly using quartile sorts.

Full name of factor	Direction	Count	Equal-weighted			Value-weighted		
			Mean	T-stat	Std	Mean	T-stat	Std
Idiosyncratic volatility	-1	325	**0.97**	**3.39**	5.27	**1.58**	**2.59**	10.96
Salience theory	1	325	**1.03**	**3.19**	5.69	**1.73**	**2.93**	10.58
Maximum daily return	-1	325	**0.87**	**3.15**	5.02	0.98	1.88	9.63
Turnover volatility	-1	325	**0.98**	**3.14**	5.32	-0.06	-0.12	7.93
Minimum daily return	1	325	**0.79**	**2.89**	4.67	1.04	1.88	10.14
Realized volatility	-1	325	**0.85**	**2.62**	5.58	**1.47**	**2.45**	10.64
Long-term reversal	1	325	**0.67**	**2.61**	4.81	**1.27**	**2.81**	7.93
Bid-ask spread	-1	325	0.66	2.19	5.47	**2.09**	**3.65**	9.73
Value-at-risk	1	325	0.66	2.15	5.27	**1.51**	**2.51**	10.42
Turnover	-1	325	0.63	2.07	5.26	-0.66	-1.33	8.27
7-day momentum	1	325	0.58	1.97	5.23	**1.94**	**3.46**	10.67
New addresses change	-1	325	0.44	1.77	4.34	0.17	0.32	9.23
13-day momentum	1	325	0.51	1.75	5.34	1.22	2.02	11.27
Zero-balance addresses	1	325	0.49	1.73	5.01	0.18	0.4	7.86
Chronological return ordering	1	325	0.52	1.69	5.5	0.88	1.61	9.54
New address-to-price ratio	1	325	0.47	1.63	5.1	**1.16**	**2.97**	6.71
Seasonality	1	325	0.4	1.58	4.9	0.67	1.36	8.59
Addresses-with-balance growth	1	325	0.35	1.5	4.41	0.73	1.52	8.1
Crypto Market	1	325	0.72	1.31	9.54	0.72	1.31	9.54
Active addresses	1	325	0.35	1.24	5.02	**0.87**	**2.32**	6.42
Standardized abnormal turnover	-1	325	0.26	1.14	4.26	0.65	1.26	9.25
Skewness	-1	325	0.29	1.11	4.61	**1.42**	**3.31**	7.69
Closeness to the 90-day high	1	325	0.31	1	5.6	1.07	1.96	10.15
Illiquidity ratio	**-1**	**325**	**0.35**	**0.97**	**6.33**	**0.64**	**1.14**	**9.88**
New addresses	1	325	0.18	0.83	4.09	0.66	1.37	7.85
CAPM alpha	-1	325	0.23	0.73	5.66	-1	-1.8	10.03
31-day momentum	1	325	0.2	0.63	5.67	**1.27**	**2.3**	10.1
Network-to-market value	-1	325	0.17	0.59	5.47	-0.3	-0.53	10.42
Kurtosis	-1	325	0.15	0.58	4.59	-0.29	-0.59	8.63
Active addresses to network value	1	325	0.15	0.55	4.95	0.59	1.42	7.45
Volume shock (30 days)	-1	325	0.14	0.53	4.89	**-1.32**	**-2.78**	8.71
Trading volume volatility	-1	325	0.19	0.52	6.25	-0.33	-0.8	6.98
Price	1	325	0.14	0.36	7.02	0.69	1.25	9.52
Intermediate momentum	1	325	0.11	0.35	5.58	0.9	1.41	11.13
Trading volume	**1**	**325**	**0.1**	**0.28**	**5.95**	**1.04**	**2.11**	**8.4**
Market value	1	325	0.1	0.27	6.53	0.81	1.7	7.69
CAPM beta	-1	325	0.06	0.19	5.83	0.4	0.67	10.45

Table 4. Iterative Factor Selection Results

This table reports the iterative factor selection procedure following Swade et al. (2024). Starting with the cryptocurrency market factor, we sequentially add factors that produce the lowest GRS statistic. For each iteration, we report the tested factor name, GRS statistic, p-value, number of remaining factors with significant alphas at $t > 1.96$ and $t > 3.00$ thresholds, minimum and maximum absolute alphas among remaining factors, average absolute alpha (weekly %), and the squared Sharpe ratio of the factor model $Sh^2(f)$. Panel A presents results for equal-weighted portfolios, Panel B for value-weighted portfolios. The procedure continues until the GRS p-value exceeds 10% or no factors remain with significant alphas. The sample includes 565 cryptocurrencies from January 1, 2018, to July 25, 2024. Portfolios are rebalanced weekly using quartile sorts.

No	Tested factor	GRS	p(GRS)	t>2	t>3	Min(α)	*Avg*|α|	Max(α)	*Sh²(f)*
						Panel A: Equal-weighted			
1	**Turnover volatility**	**2.15**	**0**	13	10	-0.32	0.28	0.66	0.05
2	**Salience theory**	**1.67**	**0.01**	3	0	-0.14	0.14	0.39	0.1
3	**New address-to-price ratio**	**1.47**	**0.05**	2	0	-0.13	0.14	0.3	0.13
4	Bid-ask spread	1.34	0.11	1	0	-0.1	0.11	0.32	0.15
5	New addresses change	1.22	0.2	1	0	-0.13	0.13	0.35	0.17
6	Chronological return ordering	1.09	0.34	0	0	-0.11	0.12	0.32	0.18
7	Intermediate momentum	1	0.47	0	0	-0.08	0.09	0.25	0.2
8	Illiquidity ratio	0.91	0.6	0	0	-0.17	0.08	0.19	0.21
9	Market value	0.85	0.68	0	0	-0.06	0.09	0.26	0.22
10	Long-term reversal	0.75	0.81	0	0	-0.08	0.08	0.16	0.24
						Panel B: Value-weighted			
1	**Bid-ask spread**	**1.72**	**0.01**	4	1	-0.53	0.31	0.98	0.06
2	**7-day momentum**	**1.47**	**0.05**	1	0	-0.35	0.2	0.64	0.09
3	Skewness	1.27	0.15	1	0	-0.62	0.18	0.35	0.11
4	Kurtosis	1.1	0.33	0	0	-0.36	0.17	0.4	0.14
5	Long-term reversal	1.04	0.42	1	0	-0.44	0.17	0.38	0.15
6	Trading volume volatility	0.93	0.58	0	0	-0.4	0.15	0.32	0.16
7	Zero-balance addresses	0.85	0.69	1	0	-0.25	0.16	0.41	0.17
8	New address-to-price ratio	0.72	0.85	1	0	-0.29	0.12	0.29	0.19
9	Trading volume	0.59	0.95	1	0	-0.42	0.14	0.28	0.21
10	Price	0.46	0.99	0	0	-0.2	0.13	0.32	0.22

Table 5. Factor Selection Across Alternative Sorting Specifications

This table reports the top 10 factors selected by the iterative GRS procedure using different sorting depths. Factors are ranked by their order of selection, with Rank 1 representing the factor that most reduces the GRS statistic when added to the market factor. Panel A presents results for equal-weighted portfolios, Panel B for value-weighted portfolios. Bold typeface indicates factors that appear in the top 10 across multiple sorting specifications within each panel. Terciles divide the cross-section into three groups, quartiles into four groups, and quintiles into five groups. Factor names in **bold** appear as significant in whole-sample analyses. Underlined factors appear in the top ten in all three design choices (terciles, quartiles, and quintiles). The sample includes 565 cryptocurrencies from January 1, 2018, to July 25, 2024. Portfolios are rebalanced weekly using quartile sorts.

No	Terciles	Quartiles	Quintiles
		Panel A: Equal-weighted	
1	**Turnover volatility**	**Turnover volatility**	Minimum daily return
2	Idiosyncratic volatility	**Salience theory**	**Turnover volatility**
3	Addresses-with-balance growth	**New address-to-price ratio**	Addresses-with-balance growth
4	New addresses change	Bid-ask spread	New addresses change
5	New address-to-price ratio	Trading volume	New addresses
6	Intermediate momentum	Chronological return ordering	Standardized abnormal turnover
7	Chronological return ordering	Intermediate momentum	31-day momentum
8	Standardized abnormal turnover	Long-term reversal	Realized volatility
9	Turnover	Maximum daily return	Market value
10	**New address-to-price ratio**	Trading volume	Illiquidity ratio
		Panel B: Value-weighted	
1	New address-to-price ratio	**Bid-ask spread**	**7-day momentum**
2	**7-day momentum**	**7-day momentum**	**Bid-ask spread**
3	Skewness	Skewness	Long-term reversal
4	**Bid-ask spread**	Kurtosis	New address-to-price ratio
5	Zero-balance addresses	Long-term reversal	Turnover volatility
6	Kurtosis	Trading volume volatility	Turnover
7	Network-to-market value	Zero-balance addresses	Active addresses to network value
8	Salience theory	New address-to-price ratio	Standardized abnormal turnover
9	Minimum daily return	Trading volume	Trading volume
10	31-day momentum	Price	Trading volume volatility

Table 6. Factor selection's stability across subperiods

This table presents the top 10 factors selected by the iterative GRS procedure in two equal subperiods. The sample is split at the median date (March 8, 2021), creating a first period from January 1, 2018 to March 8, 2021 and a second period from March 8, 2021 to June 24, 2024. Panel A reports results for equal-weighted portfolios, Panel B for value-weighted portfolios. Factors are ranked by their order of selection in the iterative procedure, with Rank 1 representing the factor that most reduces the GRS statistic when added to the market factor. Underlined factors appear to be in the top 100 in both subperiods. **Bolded** factors are significant in whole-sample analyses. The sample includes 565 cryptocurrencies.

Rank	2018-01-01 to 2021-03-08	2021-03-08 to 2024-06-24
	Panel A: Equal-weighted	
1	Minimum daily return	**Turnover volatility**
2	**Turnover volatility**	Bid-ask spread
3	Maximum daily return	Intermediate momentum
4	New addresses change	CAPM alpha
5	Chronological return ordering	Chronological return ordering
6	Trading volume	Active addresses
7	**New address-to-price ratio**	Market value
8	Turnover	Long-term reversal
9	Market value	Trading volume
10	31-day momentum	7-day momentum
	Panel B: Value-weighted	
1	**7-day momentum**	**Bid-ask spread**
2	Idiosyncratic volatility	Trading volume volatility
3	Kurtosis	Trading volume
4	Skewness	Illiquidity ratio
5	13-day momentum	Turnover volatility
6	Long-term reversal	Standardized abnormal turnover
7	CAPM alpha	Price
8	Turnover volatility	Market value
9	New addresses	New address-to-price ratio
10	Illiquidity ratio	Zero-balance addresses