

Clustering Investor Relative Performance and Crypto-Asset Return Dynamics

Guillaume Tchoffo Talla*

Université Laval

Abstract

Existing regret measures of investor behaviour typically define regret as a non-positive return shortfall of a held asset relative to the best available alternative. While this formulation is well suited to traditional equity markets, it becomes restrictive in crypto-asset markets, where extreme cross-sectional dispersion implies that relative performance can be meaningfully positive or negative across a broad set of benchmarks. We address this limitation by introducing the Investor Relative Performance (IRP) framework, which preserves both positive and negative realisations of relative performance across five benchmark classes: crypto-assets, equities, commodities, currencies, and fixed income. Using a sample of 122 crypto-assets, we construct a behavioural taxonomy via Gaussian Mixture Model (GMM) estimation in the IRP feature space, identifying three economically distinct investor types: a low IRP cluster dominated by stablecoins, a mid IRP cluster comprising mainstream crypto-assets, and a high IRP cluster of highly speculative tokens. Building on this classification, we construct a cluster-level Relative Performance Intensity Index (RII) and examine its relation to crypto-asset returns. The results document a pronounced asymmetry between positive and negative return states and show that return sensitivity to the RII is fully determined by cluster membership. Low and mid IRP assets exhibit positive and statistically significant sensitivities, whereas high IRP assets display uniformly negative sensitivities, a pattern that is difficult to reconcile with the compensation hypothesis implied by standard regret-based asset pricing frameworks. These findings are robust across alternative index specifications, after controlling for volatility, and throughout the conditional return distribution.

Keywords: Crypto-assets; Regret; Investor Relative Performance;

JEL Classification: G11; G12; C38; G41

*Corresponding author. Faculty of Business Administration, Université Laval, 2325 Rue de la Terrasse, Quebec City, Quebec G1V 0A6, Canada. Tel: +1 418 717 1054 E-mail: gutct@ulaval.ca

Clustering Investor Relative Performance and Crypto-Asset Return Dynamics

Abstract

Existing regret measures of investor behaviour typically define regret as a non-positive return shortfall of a held asset relative to the best available alternative. While this formulation is well suited to traditional equity markets, it becomes restrictive in crypto-asset markets, where extreme cross-sectional dispersion implies that relative performance can be meaningfully positive or negative across a broad set of benchmarks. We address this limitation by introducing the Investor Relative Performance (IRP) framework, which preserves both positive and negative realisations of relative performance across five benchmark classes: crypto-assets, equities, commodities, currencies, and fixed income. Using a sample of 122 crypto-assets, we construct a behavioural taxonomy via Gaussian Mixture Model (GMM) estimation in the IRP feature space, identifying three economically distinct investor types: a low IRP cluster dominated by stablecoins, a mid IRP cluster comprising mainstream crypto-assets, and a high IRP cluster of highly speculative tokens. Building on this classification, we construct a cluster-level Relative Performance Intensity Index (RII) and examine its relation to crypto-asset returns. The results document a pronounced asymmetry between positive and negative return states and show that return sensitivity to the RII is fully determined by cluster membership. Low and mid IRP assets exhibit positive and statistically significant sensitivities, whereas high IRP assets display uniformly negative sensitivities, a pattern that is difficult to reconcile with the compensation hypothesis implied by standard regret-based asset pricing frameworks. These findings are robust across alternative index specifications, after controlling for volatility, and throughout the conditional return distribution.

Keywords: Crypto-assets; Regret; Investor Relative Performance;

JEL Classification: G11; G12; C38; G41

1 Introduction

The regret literature generally assumes that investors experience disutility when the return on a held asset falls short of the best available alternative, and this shortfall is typically restricted to non-positive values because investors rarely outperform all benchmarks simultaneously (Qin; 2020; Arisoy et al.; 2024). In traditional equity markets, where cross-sectional return dispersion is relatively limited and broad declines across all asset classes are infrequent, this convention has little practical consequence. However, the crypto-asset market exhibits a markedly different return profile. As Figure 1 shows, cross-sectional return spreads can reach 217 percentage points within a single trading session, and during periods of market-wide stress a crypto-asset recording a modest loss may still outperform alternative assets, resulting in a positive relative return differential that the standard approach excludes. This paper therefore extends the conventional framework by introducing the Investor Relative Performance (IRP) measure, which retains both negative and positive realisations and evaluates the relative performance of each crypto-asset across five benchmark classes: equities, commodities, currencies, fixed income, and other crypto-assets.

The theoretical foundation of our approach is regret theory, originally formalised by Bell (1982) and Loomes and Sugden (1982) and extended to general choice sets by Quiggin (1994). In this framework, an investor’s utility depends not only on the realised outcome of the chosen action but also on the counterfactual utility that could have been obtained from the best available alternative. The resulting disutility is disproportionately large when the forgone gain is substantial, consistent with the loss-aversion asymmetry formalised by Tversky and Kahneman (1992) and documented in investor trading behaviour (Odean; 1998; Barberis and Xiong; 2009). Qin (2020) embeds this intuition in a formal asset pricing model, deriving a closed-form equilibrium in which a risky asset’s excess return is proportional to its exposure to aggregate investor regret. Arisoy et al. (2024) operationalise this concept empirically for U.S. equity markets, constructing a stock-level regret measure as the log-return shortfall relative to the best-performing stock in an investor’s reference portfolio and documenting a positive regret premium in the cross-section of equity returns. This paper extends this framework to the crypto-asset market, where substantially greater cross-sectional return dispersion may amplify the economic relevance of regret-driven dynamics relative to traditional asset classes.

Understanding where our contribution fits requires surveying how the existing literature approaches crypto-asset markets along three interconnected dimensions. The first concerns the theoretical underpinnings of crypto-asset valuation. Cong et al. (2020) develop a dynamic pricing model in which equilibrium token prices are determined by the aggregation of heterogeneous users’ transactional demand rather than discounted cash flows, while Pagnotta (2021) show that Bitcoin prices are jointly determined by transactional and resale values alongside the risk of systemic attacks. At the level of platform governance, Sockin and Xiong (2023) examine how tokenisation resolves conflicts between platforms and users, and Biais et al. (2023) offer a general equilibrium analysis in which equilibrium Bitcoin prices partly reflect extrinsic volatility arising from sunspot dynamics.

The second dimension concerns the empirical asset pricing of crypto-assets. Liu and Tsyvinski (2020) establish that crypto-asset returns are driven by network and attention factors specific to the asset class, while Liu et al. (2022) show that a three-factor model encompassing market, size, and momentum captures the cross-section of expected crypto-asset returns. Lan and Frömmel (2025) extend this factor zoo to eleven dimensions including supply dynamics, network activity, and behavioural attention, and Borri and Shakhnov (2022) provide a risk-based explanation of Bitcoin price differentials across exchanges, demonstrating that more expensive pairs carry higher liquidity

and sentiment risk.

The third dimension concerns market microstructure and investor behaviour. [Makarov and Schoar \(2020\)](#) document persistent cross-exchange arbitrage opportunities linked to capital controls, [Lehar and Parlour \(2025\)](#) characterise the equilibrium of decentralised exchanges under automated market making, and [Huang et al. \(2025\)](#) show that stablecoins serve as a transmission channel for risk between crypto-asset and dollar markets. At the investor level, [Kogan et al. \(2024\)](#) document that retail traders exhibit momentum-like rather than contrarian strategies in crypto-asset markets, while [Hackethal et al. \(2021\)](#) describe the portfolio biases and trading patterns of individual crypto-asset investors. Against this background, our paper introduces a regret-based perspective to the classification of crypto-assets by investor type and analyses its effect on return dynamics.

The broader behavioural asset pricing literature suggests that preference-based mechanisms of the type embedded in regret theory can have economically meaningful effects on the cross-section of returns, and this evidence motivates the empirical design of our IRP framework. [Barberis and Huang \(2001\)](#) show that loss aversion over individual stock fluctuations can generate high average returns, excess volatility, and a sizeable value premium in the cross-section. [Grinblatt and Han \(2005\)](#) demonstrate that prospect theory and mental accounting give rise to a capital-gains overhang variable that subsumes the momentum effect in stock returns, providing a behavioural explanation for why assets associated with shared unrealised losses may continue to attract investor demand. [Barberis et al. \(2016\)](#) examine whether investors evaluate stocks through the lens of prospect theory using historical return distributions and find that high-value stocks tend to earn lower subsequent returns, particularly among smaller firms where retail investors play a larger role in price formation. [Barberis et al. \(2021\)](#) develop a general equilibrium model incorporating elements of prospect theory and showing that it can account for a substantial share of 23 prominent stock market anomalies. These studies indicate that behavioural preference structures can influence return dynamics in ways not fully captured by conventional risk factors. The IRP framework is designed to identify and measure this channel within the crypto-asset market.

The link between regret and investor behaviour has been supported not only in equilibrium asset-pricing models but also in experimental and neuroeconomic studies, and this evidence motivates the utility structure adopted in our framework. [Frydman et al. \(2014\)](#) use fMRI evidence from an experimental stock market to show that brain regions associated with reward processing display activity consistent with realization utility, supporting the preference-based mechanism embedded in investor decision-making. [Frydman and Camerer \(2016\)](#) extend this evidence by showing that a regret-related neural response is triggered when investors observe positive returns on stocks they chose not to purchase, and that this response predicts suboptimal repurchase decisions. Their findings provide direct evidence that counterfactual comparisons influence investor choices in a manner consistent with regret theory.

The attention channel through which regret-related demand may aggregate across assets is also consistent with the findings of [Barber and Odean \(2007\)](#), who show that individual investors tend to purchase attention-grabbing stocks, including those with extreme recent returns, and that such attention-driven demand is associated with predictable return patterns. In crypto-asset markets, where retail investors represent a particularly important trader group ([Kogan et al.; 2024](#); [Hackethal et al.; 2021](#)), attention and regret effects may therefore be more pronounced than in traditional institutional markets, making the IRP framework particularly relevant in this setting.

The IRP measure introduces two extensions to the existing regret literature. The first is its

multi-benchmark structure. Whereas [Arisoy et al. \(2024\)](#) focus on comparisons within the equity market, crypto-asset investors are likely to evaluate opportunities across several financial markets simultaneously. The IRP is therefore constructed relative to five benchmark classes: other crypto-assets ($\mathcal{B}^{\text{CRY}}$), global equity sectors ($\mathcal{B}^{\text{STO}}$), commodities ($\mathcal{B}^{\text{COM}}$), currency pairs ($\mathcal{B}^{\text{CUR}}$), and fixed-income assets ($\mathcal{B}^{\text{BON}}$). The resulting five-dimensional IRP vector captures the relative cross-asset performance position that regret theory associates with investor comparisons, a dimension that has received limited attention in prior work.

The second extension is the retention of positive IRP realisations. Existing studies following the equity-market framework typically restrict the relative performance differential to non-positive values. In traditional equity markets, where daily return dispersion is relatively modest, this convention has limited practical implications. In crypto-asset markets, however, daily return movements can be substantially larger, such that during periods of broad market weakness a crypto-asset may still outperform equities, commodities, currencies, and fixed-income benchmarks simultaneously, generating a positive IRP realisation that the conventional specification would exclude. Retaining the full signed differential therefore allows the framework to capture a broader range of relative performance outcomes observed in crypto-asset markets.

Having established the theoretical motivation and construction of the IRP measure, we next use it to develop a behavioural taxonomy of crypto-assets. Specifically, we cluster the 122 assets in our sample into behaviourally similar groups by estimating a Gaussian Mixture Model (GMM) on the five-dimensional vector of time-series median IRP values for each asset. The GMM framework is preferred to alternative unsupervised methods because it allows clusters to differ in shape, orientation, and dispersion through flexible covariance structures, which is important given the heterogeneous nature of the IRP feature space.

The number of clusters is determined by maximising the average silhouette score across a range of candidate specifications, which identifies $K^* = 3$ as the optimal partition. The resulting clusters can be interpreted as representing increasing levels of relative underperformance intensity: a low-IRP cluster ($n = 20$) largely composed of stablecoins, a mid-IRP cluster ($n = 72$) containing mainstream crypto-assets, and a high-IRP cluster ($n = 14$) consisting primarily of speculative tokens. The stability of the clustering structure is supported by pairwise intra-cluster versus inter-cluster cohesion analysis using both RBF kernel similarity and Euclidean distance measures, while its economic interpretation is reflected in the systematic variation of IRP profiles and conditional volatility across the three groups. This behavioural classification contributes to the broader clustering literature in financial markets ([Mantegna; 1999](#); [Jiao and Zheng; 2026](#); [Malhotra and Malhotra; 2025](#)), where asset grouping has traditionally focused on statistical return co-movement rather than behavioural preference structures.

Building on this taxonomy, we distil the five IRP dimensions of each cluster into a single daily Relative Performance Intensity Index (RII) using Principal Component Analysis (PCA) ([Jolliffe and Cadima; 2016](#)), providing a parsimonious scalar measure of aggregate within-cluster relative underperformance intensity. The RII exhibits pronounced time-varying volatility, which we model using GARCH(1,1) ([Bollerslev; 1986](#)) to obtain a conditional standard deviation series that is included as a control in all return-sensitivity regressions. Two RII specifications are constructed, based on the first principal component alone and on a variance-weighted combination of the first two components, allowing us to assess the robustness of results to the dimensionality assumption embedded in the index construction. The high Pearson correlations between the two specifications confirm that both approaches yield nearly identical daily values, and the near-identical results across

specifications throughout the paper validate this robustness.

The empirical analysis delivers interesting findings that together provide a coherent behavioural interpretation of the crypto-asset market. The IRP asymmetry analysis shows that an asset's competitive position is systematically stronger on days of positive returns than on days of negative returns across all cluster–dimension combinations, with the magnitude of this asymmetry increasing monotonically from the low to the high IRP cluster. For the stock and currency dimensions, Cohen's d values exceed 1.5 in the mid and high IRP clusters, indicating economically meaningful differences by conventional benchmarks. This pattern is consistent with loss-aversion based behaviour (Tversky and Kahneman; 1992) and the individual investor evidence in Odean (1998) and Barberis and Xiong (2009), where losses tend to exert a stronger behavioural response than gains of comparable size.

Asset-level regressions of returns on the cluster RII suggest that the sign of return sensitivity is closely aligned with cluster membership across the sample. Low and mid IRP assets generally exhibit positive and statistically significant sensitivity to the RII, consistent with (Qin; 2020; Arisoy et al.; 2024). In contrast, high IRP assets tend to display negative and statistically significant sensitivities. While regret theory would typically predict higher expected returns as compensation for greater exposure to regret, the negative coefficients for high IRP assets suggest that periods of elevated relative underperformance intensity in this cluster are associated with lower contemporaneous returns. This result points to a potential tension between standard regret-based predictions and observed pricing dynamics in more speculative segments of the crypto market. In addition, quantile regression results indicate that this sign pattern is broadly stable across the conditional return distribution, from the 5th to the 95th percentile. This suggests that the negative sensitivity of high IRP assets is not confined to extreme market states but is present more generally across return outcomes.

Our paper makes contributions to four strands of the literature. Within the regret and behavioural asset pricing literature, it is directly related to Arisoy et al. (2024), whose equity-market framework we adapt and extend. The multi-benchmark structure, the retention of positive IRP realisations, and the GMM-based taxonomy represent methodological advances, each motivated by specific empirical properties of the crypto-asset class that have no direct counterpart in equity markets. Within the prospect theory and behavioural return literature, encompassing the foundational contributions of Barberis and Huang (2001), Grinblatt and Han (2005), Barberis et al. (2016), and Barberis et al. (2021), our paper introduces a regret-based dimension that operates through cross-asset relative comparison rather than through the distribution of a single asset's own past returns. The two channels are theoretically complementary but empirically distinct, and the IRP framework captures the daily competitive position of an asset relative to the global investment opportunity set. Within the crypto-asset pricing literature (Liu and Tsyvinski; 2020; Liu et al.; 2022; Lan and Frömmel; 2025; Borri and Shakhnov; 2022), we introduce a behaviourally motivated pricing dimension, demonstrating that relative performance dynamics account for a meaningful share of daily return variation in a setting where they have not previously been studied. Finally, the identification of the low IRP cluster as dominated by stablecoins and pegged instruments contributes to the literature on stablecoin economics (Huang et al.; 2025) by showing that regret dynamics are structurally muted for this class of instruments, providing a within-sample contrast that sharpens identification of regret-driven return patterns in the other two clusters.

The remainder of the paper is organised as follows. Section 2 describes the data, constructs the five-dimensional IRP measure, and develops the GMM-based taxonomy. Section 3 analyses the

asymmetry of IRP between up-days and down-days across the three clusters. Section 4 constructs the Relative Performance Intensity Index, estimates asset-level return sensitivities, and presents the quantile regression results. Section 5 concludes.

2 Data and preliminary analyses

The empirical analysis is built on a sample of 122 crypto-assets whose daily closing prices are retrieved from the CoinMetrics API. The sample spans January 4, 2016 through December 31, 2023, yielding an unbalanced panel in which individual assets enter the sample on their respective first trading dates as recorded by CoinMetrics. The entry dates vary considerably across assets, reflecting the staggered introduction of new crypto-assets to the market over the sample period. Only U.S.-dollar-denominated prices are employed in the construction of all variables, and returns are computed using log transformation. For fixed-income benchmarks, the return-like quantity is the arithmetic yield change $\Delta y_{j,t} = y_{j,t} - y_{j,t-1}$, where $y_{j,t}$ denotes the yield of bond instrument j on day t . This convention avoids the numerical difficulties that arise from applying the log transformation to yields that may approach zero or take negative values, and is standard in the fixed-income literature.

The five benchmark classes are constructed from a broad and internationally representative set of instruments obtained from Bloomberg, chosen to reflect the full opportunity set available to a sophisticated crypto-asset investor operating across global financial markets. The equity benchmark $\mathcal{B}^{\text{STO}}$ comprises nine sector indices drawn from the S&P Global 1200 Index: Energy (SGT), Materials (SGN), Industrials (SGH), Consumer Discretionary (SGU), Consumer Staples (SGD), Health Care (SGC), Financials (SGI), Information Technology (SGFS), and Utilities (SGM). The commodity benchmark $\mathcal{B}^{\text{COM}}$ includes eight actively traded commodity futures: Gold, Silver, Copper, Crude Oil (WTI), Brent Oil, Natural Gas, Wheat, and Soybeans, spanning precious metals, base metals, energy, and agricultural products. The currency benchmark $\mathcal{B}^{\text{CUR}}$ consists of seven major exchange rates and indices: EUR/USD, GBP/USD, JPY/USD, AUD/USD, CAD/USD, USD/CNY, and the U.S. Dollar Index (DXY). The fixed-income benchmark $\mathcal{B}^{\text{BON}}$ comprises ten U.S. Treasury constant-maturity yield series spanning the full maturity spectrum from 1-Month to 30-Year. Finally, the crypto-asset benchmark $\mathcal{B}^{\text{CRY}}$ consists of all sample crypto-assets other than asset i itself, with the self-exclusion preventing a trivially zero measure for the daily return leader.

To ensure a fully observed panel, any calendar date on which at least one asset or benchmark variable records a missing observation is excluded from the analysis in its entirety. This listwise date deletion preserves the cross-sectional integrity of the investor relative performance measures, which by construction require simultaneous price records across all assets and all benchmark classes on each retained trading day. The five-dimensional structure of the measure reflects the cross-market opportunity set facing a sophisticated crypto-asset investor who continuously monitors returns across asset classes (Duchin et al.; 2025; Trimborn et al.; 2019). Investor relative performance is inherently comparative: its magnitude and direction depend on which alternative the investor considers as the counterfactual, and different investors may anchor on different benchmarks depending on their prior asset allocation and investment mandate. By measuring relative performance simultaneously against five asset classes, the framework captures a richer and more realistic picture of investor experience than any single-dimension measure would allow.

Table 1 reports the distributional properties of daily log-returns and the five Investor Relative Performance dimensions across the full sample. Panel A documents the cross-sectional return

environment in which the IRP measures are constructed. Each row corresponds to one cross-sectional statistic computed on a given trading day, and each column reports the time-series distribution of that statistic across all days in the sample. The median daily cross-sectional standard deviation, reported in the row 50% and column Std, equals 0.0416, indicating that on a typical trading day the interquartile spread of returns across assets is substantial. At the 99th percentile of the cross-sectional maximum, the best-performing asset gained 0.9131 log-points, equivalent to a single-day return of approximately 149%. The time-series distribution of cross-sectional skewness ranges from -10.69 at its minimum to $+10.55$ at its maximum, confirming that the direction of distributional asymmetry varies substantially across trading days.

Excess kurtosis at the 95th percentile reaches 40.48 and at the 99th percentile reaches 79.44, documenting extreme leptokurtosis that is well outside the range typically observed in equity markets. At the extreme, the row Max and column Max cell records a single-day cross-sectional maximum of 2.7802 log-points, corresponding to a daily gain of approximately 1,513% for the best-performing asset on that day, while the column Min in the same row records a minimum of 0.0123, meaning that even the worst-performing asset on that day was positive. These statistics jointly establish that the crypto-asset market generates return differentials of an order of magnitude that has no equivalent in traditional asset classes, and they provide the direct empirical motivation for retaining both positive and negative realisations of the IRP measure: when the return spread between the best and worst asset on a single day can exceed 1,500 percentage points, a held asset that posts a modest loss may nonetheless represent a competitive choice within the global investment opportunity set.

2.1 IRP measure

The IRP measures developed in this paper are motivated by the regret-augmented utility framework of [Quiggin \(1994\)](#), who generalises the pairwise regret theory of [Bell \(1982\)](#) and [Loomes and Sugden \(1982\)](#) to arbitrary choice sets. In this framework, an agent's utility depends not only on the outcome of the chosen action but also on the counterfactual utility that could have been obtained from the best forgone alternative. [Qin \(2020\)](#) formalises this intuition in an asset pricing context, deriving a closed-form equilibrium pricing formula in which a risky asset's excess return is proportional to its regret beta, defined as its exposure to aggregate investor regret. [Arisoy et al. \(2024\)](#) operationalise this concept empirically for equity markets, measuring regret as the log-return shortfall of a stock relative to the best-performing stock in an investor's reference portfolio, and document a positive regret premium in the cross-section of U.S. equity returns. The neuroeconomic evidence of [Frydman and Camerer \(2016\)](#) further confirms that this counterfactual comparison is not merely a theoretical construct but a measurable cognitive and emotional response that influences investor decisions, providing independent support for the utility structure we adopt.

This paper adapts the [Arisoy et al. \(2024\)](#) framework to the crypto-asset market and extends it by considering five distinct cross-asset benchmark classes simultaneously. The central premise is that a crypto-asset investor continuously evaluates the performance of their holdings not only against other crypto-assets but also against the best available returns in equity, commodity, currency, and fixed-income markets on the same day. Each comparison defines a separate dimension of relative performance, and together the five dimensions characterise the full cross-asset investor relative performance profile of each asset. For an investor I at time t , utility is given by:

$$u(W_{I,t}) = v(W_{I,t}) + f\left(v(W_{I,t}) - \max_k \{v(W_{k,t})\}\right) \quad (1)$$

where $W_{I,t}$ denotes the wealth from investing in a given crypto-asset, and $\max_k \{v(W_{k,t})\}$ is the highest utility that could have been obtained over the same period from alternative assets $k = 1, \dots, n$. The term $v(W_{I,t}) - \max_k \{v(W_{k,t})\}$ captures the gap between realised and best attainable utility. The function $f(\cdot)$ is assumed to be increasing and convex in its argument, so that large shortfalls generate disproportionately large disutility, consistent with the loss-aversion asymmetry documented by [Tversky and Kahneman \(1992\)](#) and validated at the neurological level by [Frydman et al. \(2014\)](#).

Since current wealth is a direct function of past returns, the relevant measure can be expressed directly in terms of returns. We define the Investor Relative Performance of asset i with respect to benchmark class k on day t as:

$$\text{IRP}_{i,t}^k = R_{i,t} - \max_{j \in \mathcal{B}^k} R_{j,t} \quad (2)$$

where $R_{i,t}$ is the return on the crypto-asset held by investor I on day t , and $\max_{j \in \mathcal{B}^k} R_{j,t}$ is the highest return among the relevant alternative assets in class k over the same period. A more negative $\text{IRP}_{i,t}^k$ indicates that the investor would have obtained a substantially superior outcome by reallocating to class k , generating the psychological discomfort described by regret theory. A value close to zero indicates near-parity with the best available alternative. We interpret $\text{IRP}_{i,t}^k$ as a signed measure of relative performance within the full cross-asset opportunity set, where negative values correspond to underperformance relative to the best benchmark and positive values reflect periods in which the asset outperforms all benchmark classes simultaneously. This interpretation preserves information about both sides of the cross-sectional distribution that is directly relevant to investor evaluation of portfolio outcomes. In practice, negative realisations predominate across all five benchmark dimensions, as confirmed by the high percentage of negative values reported in Table 2, which is consistent with the theoretical motivation from regret theory as the primary driver of the empirical distribution.

Figure 2 reveals substantial cross-sectional heterogeneity in investor relative performance across all five benchmark dimensions. The CRY dimension exhibits the most negative and most tightly concentrated distribution, with a median of -0.1636 and a narrow kernel density reflecting the near-uniform underperformance of most crypto-assets relative to the daily crypto-asset market leader, while the STO, COM, CUR, and BON dimensions display considerably wider dispersions and less negative medians, ranging from -0.0114 for STO to -0.0527 for BON, indicating that the degree of relative underperformance against non-crypto benchmarks varies substantially across assets. The heterogeneity in distributional shape and location across the five panels confirms that the IRP dimensions are not redundant and that a single scalar measure would fail to capture the multi-dimensional nature of investor relative performance ([Borri and Shakhnov; 2022](#); [Lan and Frömmel; 2025](#)), providing the key motivation for the five-dimensional Gaussian Mixture Model clustering approach developed in the following section.

Panel B of Table 1 reports the distributional properties of the five IRP dimensions pooled across all asset-day observations. The median IRP values are negative for CRY (-0.163), COM (-0.030), and BON (-0.051), and approximately zero for STO (-0.010) and CUR (-0.006), indicating that on a typical trading day the held crypto-asset falls short of the best available alternative in most benchmark classes. The left tails are substantially heavier for CRY and BON, as reflected in their negative skewness of -4.67 and -0.24 , documenting episodic underperformance relative to the crypto-asset and fixed-income benchmarks during market stress. Critically, positive maximum

values are recorded across all five IRP dimensions, reaching 2.6442 for CRY, 2.7726 for STO, 2.7379 for COM, 2.7778 for CUR, and 2.7732 for BON, confirming that outperformance relative to every benchmark class occurs in the data.

Moreover, positive IRP realisations are not confined to isolated outliers: the 95th percentile of STO equals 0.077 and that of CUR equals 0.083, while the 75th percentile of STO equals 0.011 and that of CUR equals 0.015, indicating that on approximately one quarter of trading days the held crypto-asset simultaneously outperforms the best equity sector index and the best currency pair. The excess kurtosis values are large across all five dimensions, reaching 82.65 for STO and 77.81 for CUR, reflecting the heavy-tailed nature of the IRP distribution and confirming that extreme relative performance episodes are a recurrent feature of the data. This two-sided empirical distribution validates the methodological contribution of this paper: restricting the IRP measure to non-positive values, as implicitly assumed in the equity-market literature, would discard the positive realisation mass documented above the 75th percentile of STO and CUR, and would misrepresent the competitive return dynamics of the crypto-asset class.

2.2 GMM clustering

The objective of the clustering step is to identify groups of assets that share structurally similar IRP profiles across the five benchmark dimensions and to examine whether these groups differ in their return dynamics. The cross-asset investor relative performance profile of each asset is summarised by the vector of time-series medians of the five IRP dimensions:

$$\mathbf{f}_i = \left(\widetilde{\text{IRP}}_i^{\text{CRY}}, \widetilde{\text{IRP}}_i^{\text{STO}}, \widetilde{\text{IRP}}_i^{\text{COM}}, \widetilde{\text{IRP}}_i^{\text{CUR}}, \widetilde{\text{IRP}}_i^{\text{BON}} \right)^\top \quad (3)$$

where $\widetilde{\text{IRP}}_i^k$ denotes the time-series median of $\text{IRP}_{i,t}^k$ over all trading days for asset i . The median is preferred to the mean because the IRP distributions exhibit pronounced left skew and heavy tails, as documented in Panel B of Table 1, making the median a more robust location estimator under extreme observations (Milligan and Cooper; 1988). Prior to clustering, each dimension is standardised to zero mean and unit variance to prevent scale differences from artificially driving cluster separation.

Assets are partitioned by fitting a Gaussian Mixture Model (GMM) to the standardised feature vectors $\{\mathbf{f}_i\}_{i=1}^{122}$:

$$p(\mathbf{f}_i) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{f}_i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (4)$$

where $\pi_k > 0$ are mixture weights satisfying $\sum_{k=1}^K \pi_k = 1$. Parameters are estimated by the Expectation-Maximisation (EM) algorithm, and each asset is assigned to the cluster with the highest posterior probability:

$$\hat{c}_i = \arg \max_k \frac{\pi_k \mathcal{N}(\mathbf{f}_i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{\ell=1}^K \pi_\ell \mathcal{N}(\mathbf{f}_i; \boldsymbol{\mu}_\ell, \boldsymbol{\Sigma}_\ell)} \quad (5)$$

The GMM is preferred over k -means and hierarchical alternatives such as Ward's method (Ward Jr; 1963; Murtagh and Contreras; 2012) for two reasons. First, it allows clusters to differ in shape, orientation, and volume through unrestricted covariance matrices $\boldsymbol{\Sigma}_k$, in contrast to k -means which implicitly assumes spherical, equal-size clusters, an assumption that is inappropriate given the

heterogeneous geometry of the five-dimensional IRP space. Second, the GMM is likelihood-based, providing a principled probabilistic foundation for inference that is well-suited to the heavy-tailed nature of the IRP distributions documented in Table 1. This approach extends the clustering literature applied to financial markets (Mantegna; 1999; Jiao and Zheng; 2026; Malhotra and Malhotra; 2025) by grounding the feature space directly in investor psychology rather than in statistical return co-movement.

The number of clusters K is selected by maximising the average silhouette score over the grid $K \in \{2, 3, \dots, 8\}$, as illustrated in Figure 3. The silhouette score is the appropriate criterion here because it evaluates whether the partition constitutes meaningfully distinct investor types by measuring how much more similar each asset is to its own cluster than to the nearest alternative, directly in the five-dimensional IRP feature space. Unlike information criteria such as BIC, which reward goodness-of-fit to the Gaussian mixture irrespective of whether the clusters are economically interpretable, the silhouette score aligns directly with the paper’s objective of identifying behavioural groups that differ in their competitive return dynamics. The grid search selects $K^* = 3$ with a silhouette score of 0.38, the highest value across the entire grid, and cluster quality deteriorates for both smaller and larger values of K , confirming that the three-cluster solution reflects genuine structure in the IRP feature space rather than an artefact of a specific initialisation. This yields three clusters: a low IRP cluster (C_0 , $n = 20$), a mid IRP cluster (C_1 , $n = 72$), and a high IRP cluster (C_2 , $n = 14$), ordered by increasing composite relative underperformance intensity across all five dimensions.

Having identified the three clusters, we next examine how IRP profiles differ across them. Table 2 reports the distributional properties of the five IRP dimensions computed separately for each cluster. This step is important because the silhouette criterion ensures geometric separation but does not speak to the economic content of the clusters; the table establishes that the three groups correspond to distinct and interpretable investor types rather than statistically arbitrary partitions.

The three panels reveal a coherent and monotone gradient in relative underperformance intensity that validates the cluster ordering. Mean and median IRP values become more negative moving from C_0 to C_2 across all five dimensions: for BON the mean deteriorates from -0.057 to -0.058 to -0.063 , and similar progressions hold for COM (from -0.034 to -0.036 to -0.039) and CUR (from -0.005 to -0.007 to -0.010). The fraction of negative observations confirms that underperformance predominates across all clusters and dimensions, reaching 99.72% for CRY in C_0 , which is consistent with the regret-theoretic motivation of the IRP framework (Qin; 2020; Arisoy et al.; 2024).

At the same time, positive IRP realisations are present in every cluster and every dimension, as confirmed by the positive 95th percentile values for STO, CUR, and BON in all three panels: in C_2 the 95th percentile of STO reaches 0.104 and that of CUR reaches 0.110, indicating that on approximately 5% of trading days the high IRP cluster assets outperform the entire equity and currency benchmark panels simultaneously. The standard deviations increase monotonically from C_0 to C_2 for all non-CRY dimensions, with STO rising from 0.033 to 0.069 to 0.083, COM from 0.048 to 0.079 to 0.091, CUR from 0.033 to 0.071 to 0.084, and BON from 0.062 to 0.088 to 0.100, documenting that high IRP cluster assets face not only worse average relative performance but also greater day-to-day variability in their competitive position. The extreme kurtosis values for STO and CUR in C_0 (940.12 and 974.19 respectively) reflect the near-constant outperformance of stablecoin-dominated assets in these dimensions on most days, punctuated by rare but large

deviations, a pattern consistent with the structural role of stablecoins documented by [Huang et al. \(2025\)](#).

Figure 4 displays the full distribution of each IRP dimension and conditional volatility of crypto-asset IRP. Visualising the distributions is important because it reveals not only whether cluster means differ but also whether the three groups are genuinely separable, which is the condition required for the clustering to support distinct return analyses in subsequent sections. The crypto-asset IRP panel shows the clearest three-way separation, with C_0 concentrated between -0.180 and -0.145 , C_1 centred around -0.165 , and C_2 exhibiting a flatter and more dispersed distribution at less negative values, reflecting the lower and more variable crypto-asset-specific underperformance of high IRP cluster assets. The stock, commodity, currency, and bond panels reveal a consistent ordering in which C_0 assets, dominated by stablecoins, cluster tightly at the least negative end of each dimension, C_1 assets form the dominant mass of the distribution, and C_2 assets occupy the more negative tail. Importantly, the conditional volatility panel provides external validation of the cluster structure that is independent of the IRP feature space used for estimation: C_0 assets concentrate at near-zero volatility, C_1 assets display intermediate volatility centred around 0.05 , and C_2 assets exhibit the widest dispersion, confirming that the IRP-based partition also captures meaningful differences in underlying return risk.

While the distributional evidence in Table 2 and Figure 4 confirms that the three clusters differ economically, it is also necessary to verify that the partition is geometrically valid, that is, that assets within each cluster are genuinely more similar to one another than to assets in other clusters. Figure 5 addresses this by comparing pairwise intra-cluster and inter-cluster similarity using two independent metrics. Both metrics confirm that the clusters are well-separated and internally cohesive. The RBF cohesion ratios range from 2.64 for CRY to 4.38 for CUR, and the Euclidean cohesion ratios range from 2.01 for CRY to 2.91 for BON, all substantially above 1.0 across every dimension. For CRY, the intra-cluster RBF similarity mean of 0.580 is 2.64 times the inter-cluster mean of 0.219; for STO the ratio reaches 4.34 with an intra-cluster mean of 0.696 against an inter-cluster mean of 0.160; and for CUR the ratio peaks at 4.38. The bottom row corroborates these findings using Euclidean distance, with intra-cluster distances uniformly smaller than inter-cluster distances across all five dimensions. The joint evidence from Table 2, Figure 4, and Figure 5 establishes that the three-cluster GMM solution identifies economically meaningful and geometrically compact investor types whose IRP profiles differ substantially and systematically across all five benchmark dimensions, providing a solid foundation for the return-sensitivity analyses that follow.

3 IRP asymmetry analysis

A central prediction of regret theory is that investors experience greater psychological discomfort from underperformance on loss days than on gain days, because losses are evaluated against a higher reference point and are subject to the loss-aversion asymmetry formalised by [Tversky and Kahneman \(1992\)](#) and documented in investor trading behaviour by [Odean \(1998\)](#) and [Barberis and Xiong \(2009\)](#). Testing whether this prediction holds in the data serves two purposes. First, it provides empirical support for the regret-theoretic motivation underlying the IRP measure, connecting our framework to the neuroeconomic evidence of [Frydman and Camerer \(2016\)](#), who show that counterfactual comparison generates measurable psychological responses that can systematically influence investor decisions. Second, it reveals whether the degree of asymmetry varies across the three clusters, which would indicate that the behavioural types identified by the GMM differ

not only in their average competitive position but also in how their relative performance responds to market conditions. To conduct this test, we decompose the IRP measure into two conditional distributions based on the sign of the asset's own daily log-return. Formally, for each asset i in cluster c and each benchmark class k :

$$\overline{\text{IRP}}_{\text{up}}^k = \frac{1}{|\mathcal{T}_i^+|} \sum_{t \in \mathcal{T}_i^+} \text{IRP}_{i,t}^k, \quad (6)$$

$$\overline{\text{IRP}}_{\text{down}}^k = \frac{1}{|\mathcal{T}_i^-|} \sum_{t \in \mathcal{T}_i^-} \text{IRP}_{i,t}^k, \quad (7)$$

where $\mathcal{T}_i^+ = \{t : r_{i,t} > 0\}$ and $\mathcal{T}_i^- = \{t : r_{i,t} < 0\}$. All asset-day observations within a cluster are pooled to form cluster-level conditional distributions. The asymmetry gap is defined as:

$$\Delta^k = \overline{\text{IRP}}_{\text{up}}^k - \overline{\text{IRP}}_{\text{down}}^k, \quad (8)$$

so that $\Delta^k > 0$ indicates that the held asset's competitive position is better on gain days than on loss days. Statistical significance is assessed using a two-sided Welch t -test, and economic magnitude is quantified by Cohen's d .

Table 3 reports the conditional means, standard deviations, gaps, t -statistics, and Cohen's d values for each cluster-dimension combination, while Figure 6 displays the full conditional distributions. Together they provide a precise and visual account of how the competitive return position of each investor type shifts between market up-days and down-days, and whether that shift is consistent with the loss-aversion asymmetry that regret theory predicts (Qin; 2020; Arisoy et al.; 2024). Figure 6 reveals not only whether the means shift between up-days and down-days but also whether the entire shape of the distribution changes, which speaks to the breadth and uniformity of the asymmetry effect and return sign conditioning.

The gap Δ^k in Table 3 is positive for every cluster-dimension combination without exception, confirming that across the entire sample the held crypto-asset's competitive position is systematically better on up-days than on down-days. The asymmetry is largest for the stock, commodity, and currency dimensions in C_1 and C_2 : for STO in C_1 the gap reaches 0.086 with a t -statistic of 277.8 and a Cohen's d of 1.595, a large effect by conventional standards. All three clusters exceed the $d = 0.8$ threshold for STO, COM, and CUR. The asymmetry is weakest for the CRY dimension, where d reaches only 0.037 in C_0 and 0.330 in C_2 , reflecting the fact that the best crypto-asset tends to outperform the held asset by a wide margin on both up-days and down-days, leaving little room for the sign of the return to alter the competitive gap.

This pattern is broadly consistent with the attention-driven demand documented by Barber and Odean (2007), who show that extreme return realisations attract concentrated investor attention regardless of market direction, dampening the scope for asymmetric responses within the same asset class. Importantly, the gap escalates monotonically from C_0 to C_2 across most dimensions: the STO gap increases from 0.020 in C_0 to 0.086 in C_1 to 0.104 in C_2 , and the CUR gap from 0.022 to 0.090 to 0.107. This monotone progression indicates that the behavioural types identified by the GMM differ not only in their unconditional IRP levels but also in the degree to which their competitive position deteriorates under adverse market conditions, with high IRP cluster assets experiencing the sharpest relative losses on down-days. The result is consistent with the escalating underperformance

intensity documented in Table 2, and suggests that cluster membership is a stable determinant of how investors experience relative performance across market states.

On up-days the IRP distributions for all three clusters are concentrated at less negative values and exhibit lower dispersion than on down-days, consistent with the positive gaps in Table 3. The shift is most pronounced for the stock, commodity, and currency dimensions: on up-days the C_1 and C_2 distributions are centred near or above zero for STO and CUR, while on down-days the same distributions shift to the left, with C_1 and C_2 means reaching -0.052 and -0.061 for STO and -0.050 and -0.059 for CUR respectively. This leftward shift on down-days is consistent with what loss-aversion theory predicts (Tversky and Kahneman; 1992): when an asset declines in value, investors evaluate it against counterfactual alternatives that have risen, generating a disproportionately large regret response relative to the symmetric gain-day case. The crypto-asset dimension displays the smallest visual shift, corroborating the low Cohen's d values in the table. Across all five dimensions, the three cluster distributions maintain their relative ordering in both market conditions, with C_0 assets consistently recording the least negative IRP and C_2 assets the most negative. This stability of the cluster ordering across market regimes suggests that the behavioural types identified by the GMM reflect consistent structures, and it motivates the return-sensitivity analysis in the following section, where we examine whether the systematic differences in IRP profiles across clusters translate into corresponding differences in return dynamics.

4 Relative performance intensity index and return sensitivity

The IRP asymmetry results in the previous section establish that the three investor types identified by the GMM differ systematically in how their competitive return position responds to market conditions, with the asymmetry gap escalating monotonically from the low to the high IRP cluster. The natural next question is whether these differences in IRP profiles translate into corresponding differences in return dynamics. To answer this question, we first distil the five IRP dimensions into a single daily index for each cluster, then estimate how individual asset returns respond to fluctuations in that index, and finally test whether the return-sensitivity relationship holds uniformly across the conditional return distribution.

4.1 Construction of the RII

To distil the five IRP dimensions into a single daily scalar capturing the aggregate relative performance intensity within each cluster, a Relative Performance Intensity Index (RII) is constructed for each cluster $c \in \{C_0, C_1, C_2\}$ using Principal Component Analysis (Jolliffe and Cadima; 2016). The PCA approach is appropriate here because the five IRP dimensions are correlated within each cluster, as documented in Table 2, and a single index that optimally weights the common variation across dimensions is both more parsimonious and more interpretable than treating all five dimensions separately in the regressions. The construction proceeds in two steps.

In the first step, for each trading day t and cluster c , the cross-sectional mean of each IRP dimension is computed across all assets assigned to that cluster:

$$\overline{\text{IRP}}_{c,t}^k = \frac{1}{n_c} \sum_{i \in c} \text{IRP}_{i,t}^k, \quad k \in \{\text{CRY}, \text{STO}, \text{COM}, \text{CUR}, \text{BON}\} \quad (9)$$

where n_c is the number of assets in cluster c . Each of the five resulting time series is standardised to zero mean and unit variance and sign-flipped so that a higher RII value corresponds to

greater relative underperformance intensity, preserving the interpretive convention established in the asymmetry analysis of Section 3.

In the second step, PCA is applied to the $T \times 5$ matrix of standardised IRP dimension series for each cluster. Two specifications are considered. In the PC1 specification the RII is the score on the first principal component:

$$\text{RII}_{c,t}^{\text{PC1}} = \mathbf{w}_{c,1}^\top \overline{\text{IRP}}_{c,t}^* \quad (10)$$

In the PC1+PC2 specification it is a variance-weighted combination of the first two components:

$$\text{RII}_{c,t}^{\text{PC1+PC2}} = \frac{\lambda_{c,1}}{\lambda_{c,1} + \lambda_{c,2}} \mathbf{w}_{c,1}^\top \overline{\text{IRP}}_{c,t}^* + \frac{\lambda_{c,2}}{\lambda_{c,1} + \lambda_{c,2}} \mathbf{w}_{c,2}^\top \overline{\text{IRP}}_{c,t}^* \quad (11)$$

where $\lambda_{c,1} \geq \lambda_{c,2}$ are the two largest eigenvalues. The PC1+PC2 specification is motivated by the observation that the second principal component is dominated by the CRY dimension, capturing crypto-asset-specific relative performance variance orthogonal to the common cross-asset factor extracted by PC1. The sign of each eigenvector is corrected so that the loading on CRY is non-positive, ensuring that $\text{RII} \uparrow$ unambiguously corresponds to greater relative underperformance intensity. Reporting results under both specifications serves as a robustness check on the sensitivity of the return-IRP relationship to the dimensionality assumption embedded in the RII construction.

Table 4 summarises the PCA structure, RII distributional properties, and GARCH parameters. The first principal component explains between 47.1% of cross-dimension IRP variation in C_0 and 72.1% in C_2 , confirming that a common factor drives the bulk of co-movement within each cluster, with the degree of commonality increasing monotonically from the low to the high IRP cluster. This monotone pattern suggests that high IRP assets share a more uniform response to the benchmark opportunity set, while low IRP assets, dominated by stablecoins, have more idiosyncratic IRP profiles that a single factor captures less completely (Huang et al.; 2025). The PC1+PC2 combined specification explains substantially more variation, reaching 66.5% in C_0 and 91.6% in C_2 , justifying the inclusion of a second component particularly for the low IRP cluster. The PC1 loadings in Panel B reveal an interpretable factor structure: for C_0 and C_1 the first component loads positively on STO, COM, CUR, and BON while loading negatively on CRY, indicating that greater relative underperformance against non-crypto benchmarks defines the common factor, consistent with the cross-asset competitive dynamics documented by Borri and Shakhnov (2022) and Lan and Frömmel (2025). For C_2 all non-CRY loadings are negative, reflecting that high IRP assets underperform both crypto-asset and non-crypto benchmarks simultaneously. The PC2 loadings are dominated by CRY across all three clusters, with absolute values of 0.949, 0.998, and 0.994 for C_0 , C_1 , and C_2 respectively. The high Pearson correlations between the two RII specifications, ranging from 0.967 to 0.990 across clusters, confirm that both constructions yield nearly identical daily values, ensuring that subsequent results are robust to the dimensionality assumption.

The RII series exhibit pronounced time-varying volatility, as documented in Panel D of Table 4, a feature that is well established in crypto-asset return series more broadly (Liu and Tsyvinski; 2020; Bianchi et al.; 2022). Modelling this conditional heteroskedasticity is important because if unaddressed it could confound the estimated sensitivity of asset returns to relative performance intensity (Bollerslev; 1986; Brandt and Kang; 2004). A GARCH(1,1) model (Bollerslev; 1986) is therefore fitted separately to each cluster RII series. The ARCH-LM pre-test p -values are exactly zero for every cluster and specification, confirming conditional heteroskedasticity and warranting

GARCH estimation. The estimated total variance persistence $\hat{\alpha} + \hat{\beta}$ ranges from 0.975 to 0.990, indicating near-integrated variance dynamics: shocks to relative performance volatility dissipate very slowly, so that episodes of elevated IRP volatility persist for extended periods. The ARCH-LM post-test p -values are comfortably above conventional significance thresholds for all six cases, confirming that the GARCH(1,1) specification adequately removes conditional heteroskedasticity. The estimated conditional standard deviation $\hat{\sigma}_{c,t}$ is retained as a time-varying control variable in all return-sensitivity regressions.

4.2 Return sensitivity to the RII

Having constructed the RII for each cluster, we now estimate the sensitivity of individual asset returns to their cluster's RII. The central question is whether this sensitivity differs systematically across the three investor types identified by the GMM, and specifically whether the sign and magnitude of the return-IRP relationship depend on which cluster an asset belongs to. A finding that cluster membership determines the sign of the sensitivity would constitute evidence that the IRP-based taxonomy captures genuinely distinct return dynamics rather than merely grouping assets by a statistical criterion, complementing the broader evidence that behavioural preference structures shape return dynamics in economically meaningful ways (Barberis and Huang; 2001; Grinblatt and Han; 2005; Barberis et al.; 2021). We estimate two complementary specifications: an asset-level regression that characterises the full cross-sectional distribution of sensitivities within each cluster, and a cluster-level quantile regression that tests whether the return-IRP relationship holds uniformly across the conditional return distribution.

For each asset i in cluster c , the following regression is estimated with HAC standard errors:

$$r_{i,t} = \alpha_i + \beta_i \text{RII}_{c,t} + \gamma_i r_{i,t-1} + \delta_i \hat{\sigma}_{c,t} + \varepsilon_{i,t}, \quad (12)$$

where $r_{i,t-1}$ controls for short-term return autocorrelation, a feature well documented in crypto-asset markets (Liu and Tsyvinski; 2020; Liu et al.; 2022), and $\hat{\sigma}_{c,t}$ controls for the possibility that elevated relative performance intensity coincides mechanically with higher volatility (Brandt and Kang; 2004). A positive $\hat{\beta}_i$ indicates pro-cyclical behaviour, while a negative $\hat{\beta}_i$ indicates counter-cyclical behaviour. A parsimonious 1-factor specification retaining only the RII is also estimated as a robustness check:

$$r_{i,t} = \alpha_i + \beta_i \text{RII}_{c,t} + \varepsilon_{i,t}. \quad (13)$$

Comparing $\hat{\beta}_i$ across equations (12) and (13) reveals whether the return-IRP relationship is driven by relative performance intensity per se or is confounded by volatility or autocorrelation effects.

Table 5 reports the results. The findings are clear: the sign of $\hat{\beta}$ is entirely determined by cluster membership, without exception across both RII specifications and both model specifications. In C_0 the mean $\hat{\beta}$ under RII_{PC1} equals 0.007 with 90.3% of assets positive and 74.2% statistically significant; the one-sample t -statistic of 3.721 confirms that the cross-sectional mean differs from zero. In C_1 the mean rises to 0.026 with 100% of assets positive and significant and a mean R^2 of 0.491, indicating that the RII, after controlling for autocorrelation and volatility, explains a substantial fraction of daily return variation for the average mid IRP cluster asset. This pro-cyclical pattern is consistent with the attention-driven herding documented by Barber and Odean (2007) and the momentum-like behaviour of retail crypto-asset investors documented by Kogan et al. (2024): when cluster-level relative underperformance intensity rises, co-movement among assets sharing that experience tends to intensify. In C_2 , every asset records a negative $\hat{\beta}$ with mean -0.035 under RII_{PC1}

and -0.043 under $\text{RII}_{\text{PC1+PC2}}$, 100% statistically significant in both specifications, and t -statistics of -36.2 and -39.7 . Regret theory predicts that investors experiencing greater relative underperformance should require higher returns as compensation for emotional discomfort; the uniformly negative $\hat{\beta}$ values for high IRP assets suggest instead that extreme relative underperformance intensity is associated with lower contemporaneous returns, a pattern that sits in tension with the theory's prediction and may reflect dynamics specific to highly speculative tokens that are not fully captured by the regret framework as conventionally formulated. The close correspondence between full model and 1-factor estimates across all three clusters confirms that the return-IRP sensitivity pattern is not driven by autocorrelation or volatility effects but reflects a genuine relationship between asset returns and cross-asset relative performance dynamics.

Figures 7 and 8 reveal not only whether the cluster means differ but also whether the distributions are separated, which is the condition required to claim that cluster membership truly determines return sensitivity. Figure 7 presents scatter bubble plots of $\hat{\beta}$ against the HAC t -statistic for each cluster across all four specifications. The spatial separation is complete: all C_0 and C_1 assets lie in the positive half-plane and all C_2 assets in the negative half-plane, with no asset crossing the $\hat{\beta} = 0$ reference line in any specification. In C_0 most bubbles cluster near zero with a small number of non-stable assets such as BNB, GNT, and TRX reaching higher values; the low t -statistics for stablecoin assets confirm that their positive sensitivities are often statistically weak. In C_1 assets form a dense cloud with uniformly large t -statistics and deep red colouring indicating high R^2 , and the positive slope confirms that assets with higher $\hat{\beta}$ also tend to generate more precise estimates. In C_2 all 14 assets lie in the negative quadrant, with VTC reaching below -40 and uniform deep-red colouring confirming high R^2 throughout. Moving from PC1 to PC1+PC2 columns, the C_2 cloud shifts further left while C_0 and C_1 shift slightly rightward, consistent with the mean differences in Table 5.

Figure 8 presents the kernel density estimates and boxplots across all four specifications. The C_2 distribution is entirely to the left of zero, centred around -0.035 under RII_{PC1} and -0.043 under $\text{RII}_{\text{PC1+PC2}}$; the C_1 distribution forms a narrow peak consistent with the low cross-sectional dispersion in Table 5; and the C_0 distribution is wide and right-skewed, reflecting the heterogeneous composition of the low IRP cluster. The boxplots confirm that the C_1 interquartile range is the narrowest in all four specifications, the C_0 box extends well above zero, and the C_2 box is entirely below zero. The visual structure is essentially unchanged across all four columns, providing evidence that the cluster-determined sign pattern is robust to both the RII dimensionality assumption and the inclusion of return and volatility controls.

4.3 Quantile regression analysis

The regression results in the previous section establish that the sign of return sensitivity to the RII is entirely determined by cluster membership. However, regression at the conditional mean measures the average relationship, and it is possible that the return-IRP relationship is concentrated in specific regions of the return distribution, for instance in the crash tail, rather than holding uniformly across market states. This distinction matters for portfolio construction and risk management (Duchin et al.; 2025; Trimborn et al.; 2019): a relationship that operates only in specific market regimes is of limited value relative to one that is active throughout the return distribution. To investigate this, quantile regressions are estimated at the cluster level:

$$Q_{\tau}(r_{c,t} \mid \text{RII}_{c,t}, r_{c,t-1}, \hat{\sigma}_{c,t}) = \alpha(\tau) + \beta(\tau) \text{RII}_{c,t} + \gamma(\tau) r_{c,t-1} + \delta(\tau) \hat{\sigma}_{c,t}, \quad (14)$$

for $\tau \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$. If $\hat{\beta}(\tau)$ is stable across quantiles within each cluster, the return-IRP relationship is uniform throughout the distribution. If $\hat{\beta}(\tau)$ varies with τ , it indicates concentration in specific market regimes. A parsimonious 1-factor variant retaining only the RII is also estimated as a robustness check, following the same strategy applied to the full-model specifications in the previous subsection.

Table 6 reports the results. The sign of $\hat{\beta}(\tau)$ is uniform across all seven quantiles within each cluster, confirming that the return-IRP relationship holds throughout the conditional return distribution. In C_0 and C_1 the coefficients decline moderately from the lower to the upper tail: for C_1 under RII_{PC1} , $\hat{\beta}(0.05) = 0.027$ against $\hat{\beta}(0.95) = 0.023$, indicating slightly stronger sensitivity at the lower tail, but the values remain statistically significant throughout. The pseudo- R^2 for C_1 reaches 0.875 at $\tau = 0.05$, confirming that the RII explains a large fraction of return variation even at the extreme crash tail, a result that underscores the potential economic importance of regret dynamics during periods of market stress. Regarding C_2 , $\hat{\beta}(\tau)$ is uniformly negative with a near-flat profile across quantiles, ranging from -0.036 at $\tau = 0.10$ to -0.034 at $\tau = 0.95$ under RII_{PC1} . This stability across the full return distribution indicates that the negative sensitivity of high IRP cluster assets to the RII is not confined to crash conditions but is a persistent feature of their return dynamics across all market regimes. Rather than acting as a compensation for emotional discomfort as regret theory would predict, the RII appears to be associated with lower contemporaneous returns for speculative tokens across the entire return distribution. The robustness of all results to the PC1+PC2 specification, visible in Panels D–F, further confirms that the findings do not depend on the dimensionality assumption embedded in the RII construction, and that the cluster-determined sign pattern documented throughout this section represents a systematic feature of crypto-asset return dynamics.

5 Conclusion

This paper introduces the IRP measure, a five-dimensional framework that evaluates the daily competitive position of each crypto-asset relative to the best available alternative across five benchmark classes. Motivated by the regret utility framework, the IRP measure departs from the existing regret literature in two substantive respects. First, it extends the benchmark structure from a single asset class to five simultaneously monitored cross-asset dimensions, reflecting the choice architecture facing a sophisticated crypto-asset investor who continuously monitors return opportunities across global financial markets. Second, it retains both negative and positive realisations of the relative performance differential, a methodological necessity dictated by the extreme return environment of the crypto-asset class, where a held asset that posts a modest loss may nonetheless outperform every equity, commodity, currency, and fixed-income benchmark simultaneously on the same day.

Applying a Gaussian Mixture Model to the five-dimensional IRP feature space of 122 crypto-assets, we identify three behaviourally homogeneous investor types ordered by increasing relative underperformance intensity: a low IRP cluster dominated by stablecoins, a mid IRP cluster of mainstream crypto-assets, and a high IRP cluster of speculative tokens. The cluster structure is geometrically valid, as confirmed by pairwise intra-versus inter-cluster cohesion analysis using both RBF kernel similarity and Euclidean distance metrics, and economically interpretable. This behavioural taxonomy contributes to the clustering literature applied to financial markets by grounding the feature space directly in investor psychology rather than in statistical return co-movement.

The empirical analysis yields interesting findings. IRP exhibits a clear asymmetry between

up-days and down-days, with the held asset consistently performing more favourably on positive-return days across all cluster-dimension combinations. This asymmetry increases monotonically from low to high IRP clusters, with particularly large effects in the stock and currency dimensions, consistent with rising speculative intensity and stronger counterfactual comparison effects.

The relationship between returns and the cluster-level RII is fully determined by cluster membership across all assets and specifications. Low and mid IRP assets show positive and significant sensitivities, consistent with attention-driven herding, while high IRP assets display uniformly negative sensitivities, contradicting the standard regret-theoretic prediction of compensation for heightened regret exposure. This pattern is robust across different model specifications. Quantile regressions further confirm that the sign structure holds across the entire conditional return distribution, with high IRP assets exhibiting persistently negative and nearly flat sensitivities from the 5th to the 95th percentile, indicating a stable structural effect rather than one driven by specific market conditions.

These results suggest that IRP captures a persistent and economically meaningful form of cross-asset relative performance pressure that is not adequately described by standard return-based or volatility-based risk measures. The combination of strong clustering structure, consistent asymmetries across market states, and stable RII sensitivities indicates that investor behaviour in crypto-asset markets is systematically shaped by continuous counterfactual comparison rather than isolated price movements. The evidence points to a behavioural channel that is both cluster-dependent and distributionally robust, offering a coherent framework for understanding how relative performance evaluation influences return dynamics in highly speculative markets.

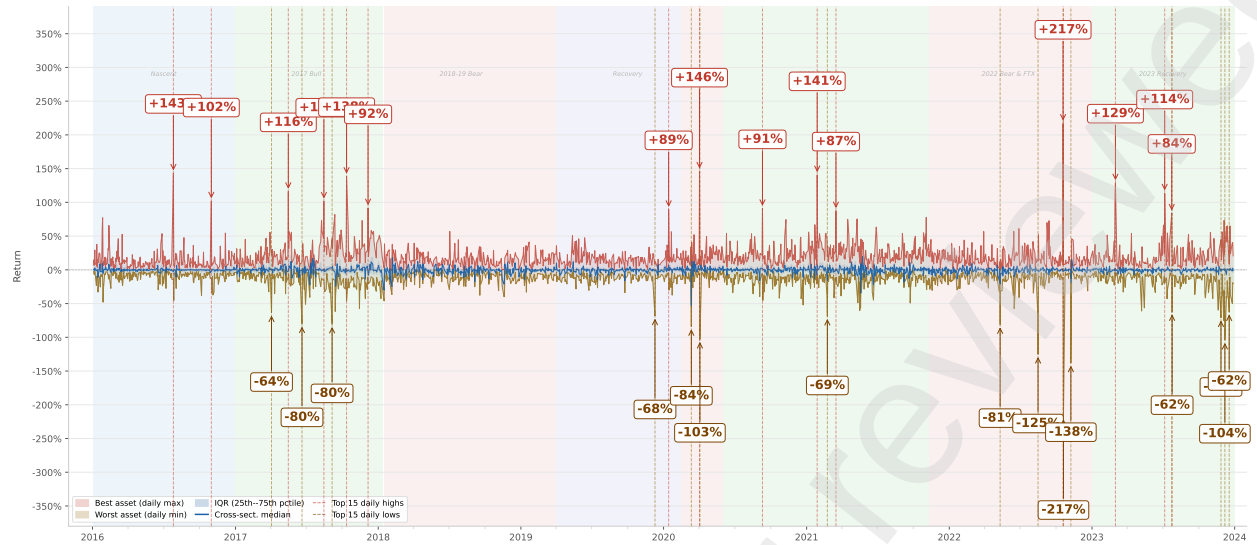
References

- Arisoy, Y. E., Bali, T. G. and Tang, Y. (2024). Investor regret and stock returns, *Management Science* **70**(11): 7537–7558.
- Barber, B. M. and Odean, T. (2007). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *The Review of Financial Studies* **21**(2): 785–818.
- Barberis, N. and Huang, M. (2001). Mental accounting, loss aversion, and individual stock returns, *The Journal of Finance* **56**(4): 1247–1292.
- Barberis, N., Jin, L. J. and Wang, B. (2021). Prospect theory and stock market anomalies, *The Journal of Finance* **76**(5): 2639–2687.
- Barberis, N., Mukherjee, A. and Wang, B. (2016). Prospect theory and stock returns: An empirical test, *The Review of Financial Studies* **29**(11): 3068–3107.
- Barberis, N. and Xiong, W. (2009). What drives the disposition effect? an analysis of a long-standing preference-based explanation, *The Journal of Finance* **64**(2): 751–784.
- Bell, D. E. (1982). Regret in decision making under uncertainty, *Operations Research* **30**(5): 961–981.
- Biais, B., Bisière, C., Bouvard, M., Casamatta, C. and Menkveld, A. J. (2023). Equilibrium bitcoin pricing, *The Journal of Finance* **78**(2): 967–1014.
- Bianchi, D., Babiak, M. and Dickerson, A. (2022). Trading volume and liquidity provision in cryptocurrency markets, *Journal of Banking & Finance* **142**: 106547.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**(3): 307–327.
- Borri, N. and Shakhnov, K. (2022). The cross-section of cryptocurrency returns, *The Review of Asset Pricing Studies* **12**(3): 667–705.
- Brandt, M. W. and Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent var approach, *Journal of Financial Economics* **72**(2): 217–257.
- Cong, L. W., Li, Y. and Wang, N. (2020). Tokenomics: Dynamic adoption and valuation, *The Review of Financial Studies* **34**(3): 1105–1155.
- Duchin, R., Solomon, D. H., Tu, J. and Wang, X. (2025). The cryptocurrency elephant in the room, *Review of Finance* **29**(6): 1721–1767.
- Frydman, C., Barberis, N., Camerer, C., Bossaerts, P. and Rangel, A. (2014). Using neural data to test a theory of investor behavior: An application to realization utility, *The Journal of Finance* **69**(2): 907–946.
- Frydman, C. and Camerer, C. (2016). Neural evidence of regret and its implications for investor behavior, *The Review of Financial Studies* **29**(11): 3108–3139.

- Grinblatt, M. and Han, B. (2005). Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* **78**(2): 311–339.
- Hackethal, A., Hanspal, T., Lammer, D. M. and Rink, K. (2021). The characteristics and portfolio behavior of bitcoin investors: Evidence from indirect cryptocurrency investments*, *Review of Finance* **26**(4): 855–898.
- Huang, Y., Ji, Y., Lin, J., Su, D. and Wang, P. (2025). The crumbling wall between crypto and non-crypto markets: Risk transmission through stablecoins, *Journal of Financial Markets* p. 101043.
- Jiao, W. and Zheng, X. (2026). Clustering-augmented reversal strategy improves return performance: Evidence from chinese stock market, *Pacific-Basin Finance Journal* **95**: 102996.
- Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review and recent developments, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374**(2065): 20150202.
- Kogan, S., Makarov, I., Niessner, M. and Schoar, A. (2024). Are cryptos different? evidence from retail trading, *Journal of Financial Economics* **159**: 103897.
- Lan, T. and Frömmel, M. (2025). Risk factors in cryptocurrency pricing, *International Review of Financial Analysis* **105**: 104389.
- Lehar, A. and Parlour, C. (2025). Decentralized exchange: The uniswap automated market maker, *The Journal of Finance* **80**(1): 321–374.
- Liu, Y. and Tsyvinski, A. (2020). Risks and returns of cryptocurrency, *The Review of Financial Studies* **34**(6): 2689–2727.
- Liu, Y., Tsyvinski, A. and Wu, X. (2022). Common risk factors in cryptocurrency, *The Journal of Finance* **77**(2): 1133–1177.
- Loomes, G. and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty, *The Economic Journal* **92**(368): 805–824.
- Makarov, I. and Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets, *Journal of Financial Economics* **135**(2): 293–319.
- Malhotra, R. and Malhotra, D. (2025). Clustering-based stock classification: A data-driven framework for investment., *Journal of Portfolio Management* **52**(2).
- Mantegna, R. N. (1999). Hierarchical structure in financial markets, *The European Physical Journal B-Condensed Matter and Complex Systems* **11**(1): 193–197.
- Milligan, G. W. and Cooper, M. C. (1988). A study of standardization of variables in cluster analysis, *Journal of classification* **5**(2): 181–204.
- Murtagh, F. and Contreras, P. (2012). Algorithms for hierarchical clustering: an overview, *Wiley interdisciplinary reviews: data mining and knowledge discovery* **2**(1): 86–97.
- Odean, T. (1998). Are investors reluctant to realize their losses?, *The Journal of Finance* **53**(5): 1775–1798.

- Pagnotta, E. S. (2021). Decentralizing money: Bitcoin prices and blockchain security, *The Review of Financial Studies* **35**(2): 866–907.
- Qin, J. (2020). Regret-based capital asset pricing model, *Journal of Banking & Finance* **114**: 105784.
- Quiggin, J. (1994). Regret theory with general choice sets, *Journal of Risk and Uncertainty* **8**(2): 153–165.
- Sockin, M. and Xiong, W. (2023). Decentralization through tokenization, *The Journal of Finance* **78**(1): 247–299.
- Trimborn, S., Li, M. and Härdle, W. K. (2019). Investing with cryptocurrencies—a liquidity constrained investment approach*, *Journal of Financial Econometrics* **18**(2): 280–306.
- Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* **5**(4): 297–323.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function, *Journal of the American statistical association* **58**(301): 236–244.

Figure 1: Cross-sectional daily log-return distribution across 122 crypto-assets



Note: Each day plots the cross-sectional distribution of log-returns across all 122 assets in the sample. The shaded bands show the full min-max range (light grey), the positive return zone (red, above zero), the negative return zone (amber, below zero), and the interquartile range (blue, 25th–75th percentile). The solid blue line is the cross-sectional median. Dashed vertical lines identify the 15 largest single-asset daily gains (red) and the 15 largest single-asset daily losses (amber) over the sample period; annotated values are the corresponding log-returns multiplied by 100. Background shading demarcates market regimes: green = bull market, red = bear market or crash episode. Data source: CoinMetrics.

Table 1: Distributional summary statistics

Panel A: Cross-sectional distribution of daily log-returns

	Min	1%	5%	25%	50%	75%	95%	99%	Max	Skew	Ex. Kurt.	Std
Min	-2.1735	-0.7975	-0.7034	-0.5882	-0.5258	-0.4105	-0.1143	-0.0154	-0.0154	-10.6923	-1.3357	0.0153
1%	-0.7421	-0.3500	-0.2375	-0.1742	-0.1419	-0.0894	-0.0009	0.0016	0.0034	-6.9233	-0.7987	0.0186
5%	-0.3549	-0.2129	-0.1486	-0.0974	-0.0708	-0.0329	0.0010	0.0220	0.0340	-2.5025	-0.2062	0.0229
25%	-0.1669	-0.1175	-0.0744	-0.0377	-0.0182	-0.0001	0.0255	0.0670	0.0959	-0.0951	1.8047	0.0324
50%	-0.1108	-0.0801	-0.0452	-0.0145	-0.0006	0.0104	0.0549	0.1082	0.1549	0.8236	4.5951	0.0416
75%	-0.0678	-0.0512	-0.0229	0.0000	0.0130	0.0335	0.0915	0.1677	0.2460	1.8223	11.1126	0.0569
95%	-0.0280	-0.0178	-0.0009	0.0195	0.0513	0.0837	0.1828	0.3087	0.4945	4.2444	40.4755	0.1021
99%	-0.0052	-0.0027	0.0008	0.0563	0.0913	0.1338	0.2781	0.4810	0.9131	7.5197	79.4351	0.1671
Max	0.0123	0.0125	0.0710	0.1447	0.1968	0.2478	0.5203	1.2173	2.7802	10.5517	113.5693	0.3712

Panel B: IRP summary statistics

	Min	1%	5%	25%	50%	75%	95%	99%	Max	Skew	Ex. Kurt.	Std
CRY	-3.0356	-0.9389	-0.5002	-0.2540	-0.1628	-0.1059	-0.0474	0.0056	2.6442	-4.6716	43.9256	0.1908
STO	-2.1745	-0.1799	-0.0984	-0.0348	-0.0101	0.0107	0.0766	0.1705	2.7726	1.1413	82.6524	0.0640
COM	-2.2172	-0.2522	-0.1425	-0.0631	-0.0302	-0.0058	0.0587	0.1540	2.7379	-0.0830	51.9311	0.0738
CUR	-2.1698	-0.1810	-0.0967	-0.0307	-0.0057	0.0151	0.0828	0.1760	2.7778	0.9077	77.8056	0.0653
BON	-2.3095	-0.2973	-0.1903	-0.0954	-0.0513	-0.0155	0.0517	0.1452	2.7732	-0.2425	33.4087	0.0840
RET	-2.1735	-0.1750	-0.0912	-0.0250	-0.0001	0.0202	0.0875	0.1813	2.7802	0.9490	78.4629	0.0651

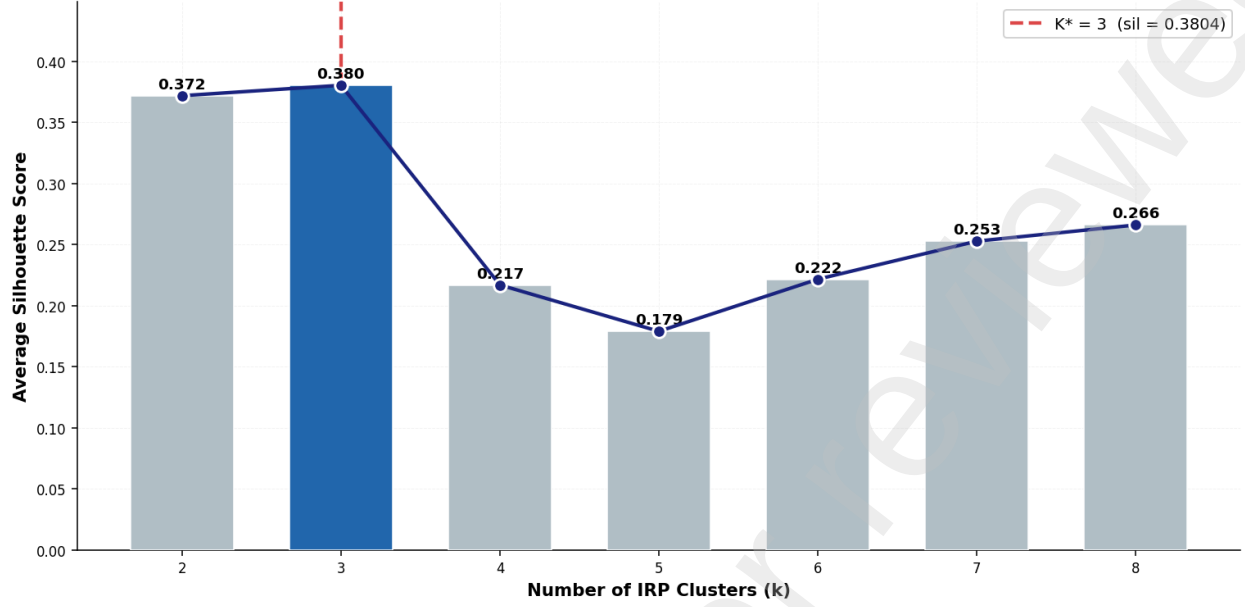
Note: Panel A reports the time-series distribution of cross-sectional summary statistics of daily log-returns. For each trading day t , a set of cross-sectional statistics is computed across all assets in the sample: the minimum, the 1st through 99th percentiles, the maximum, skewness, excess kurtosis, and standard deviation of the cross-sectional log-return distribution on that day. Each of these statistics is then summarised over all trading days in the sample, so that each row corresponds to one cross-sectional moment and each column corresponds to a quantile of that moment's time-series distribution. Panel B reports summary statistics for each of the five Investor Relative Performance (IRP) dimensions: Crypto (CRY), Stock (STO), Commodity (COM), Currency (CUR), and Bond (BON), as defined in equation (2), and the daily log-return (RET), computed by pooling all asset-day observations in the sample.

Figure 2: IRP distributions across crypto-assets



Note: Each panel corresponds to one of the five Investor Relative Performance (IRP) dimensions: Crypto (CRY), Stock (STO), Commodity (COM), Currency (CUR), and Bond (BON). The right half displays a kernel density estimate of the cross-sectional distribution of asset-level median IRP values. The left half shows a strip plot in which each dot represents one asset. Within each panel, the dashed horizontal line marks the cross-asset median and the solid horizontal line marks the cross-asset mean.

Figure 3: GMM model selection: silhouette score over the IRP feature space



Note: Average silhouette score for $k = 2, \dots, 8$ GMM clusters fitted on the five-dimensional IRP feature space. The blue bar and red dashed line identify the selected optimum $K^* = 3$ (silhouette = 0.3804). Cluster quality deteriorates for both smaller and larger values of k .

Table 2: Cluster-level IRP profiles

Panel A: Low IRP cluster (C_0 , $n = 20$)

Dimension	Mean	Median	Std	Skewness	Kurtosis	% Negative	Q5%	Q95%
CRY	-0.2070	-0.1590	0.1935	-4.9497	46.1370	99.72	-0.5009	-0.0454
STO	-0.0092	-0.0085	0.0332	1.6294	940.1181	79.41	-0.0450	0.0242
COM	-0.0338	-0.0265	0.0480	-2.4773	240.2532	92.68	-0.0966	0.0081
CUR	-0.0051	-0.0046	0.0329	1.2595	974.1934	80.29	-0.0379	0.0278
BON	-0.0574	-0.0466	0.0620	-1.3464	93.9438	93.02	-0.1589	0.0055

Panel B: Mid IRP cluster (C_1 , $n = 72$)

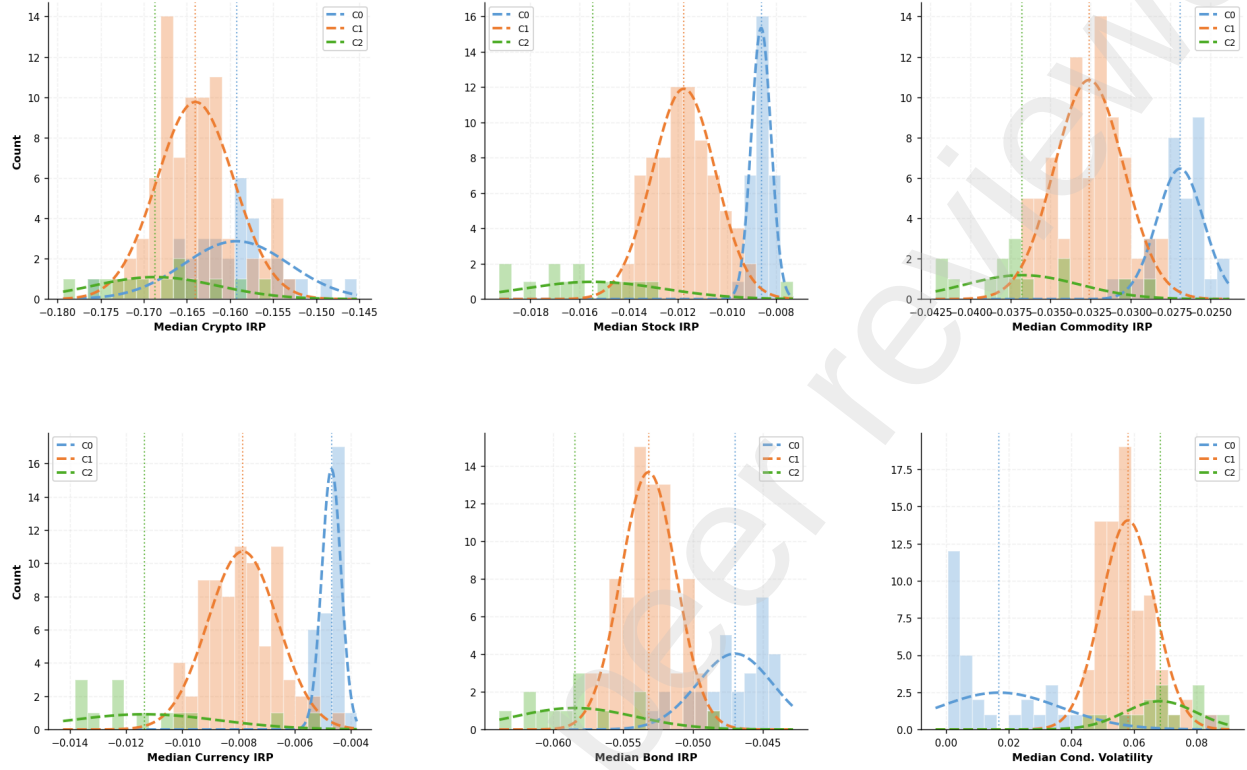
Dimension	Mean	Median	Std	Skewness	Kurtosis	% Negative	Q5%	Q95%
CRY	-0.2070	-0.1633	0.1898	-4.6658	44.2009	98.66	-0.4978	-0.0494
STO	-0.0112	-0.0116	0.0690	1.1299	65.6117	60.99	-0.1054	0.0856
COM	-0.0356	-0.0321	0.0785	0.0643	43.3085	74.58	-0.1493	0.0678
CUR	-0.0072	-0.0078	0.0706	0.8970	61.1437	57.27	-0.1040	0.0917
BON	-0.0580	-0.0530	0.0882	-0.1275	29.4014	80.58	-0.1946	0.0608

Panel C: High IRP cluster (C_2 , $n = 14$)

Dimension	Mean	Median	Std	Skewness	Kurtosis	% Negative	Q5%	Q95%
CRY	-0.2087	-0.1681	0.1910	-3.9872	36.1525	97.53	-0.5097	-0.0358
STO	-0.0141	-0.0152	0.0831	0.8083	25.5677	62.76	-0.1263	0.1037
COM	-0.0388	-0.0367	0.0907	0.2433	20.2596	74.96	-0.1675	0.0855
CUR	-0.0102	-0.0109	0.0844	0.6756	24.9570	59.78	-0.1255	0.1099
BON	-0.0625	-0.0582	0.0999	-0.0441	14.3821	80.19	-0.2121	0.0760

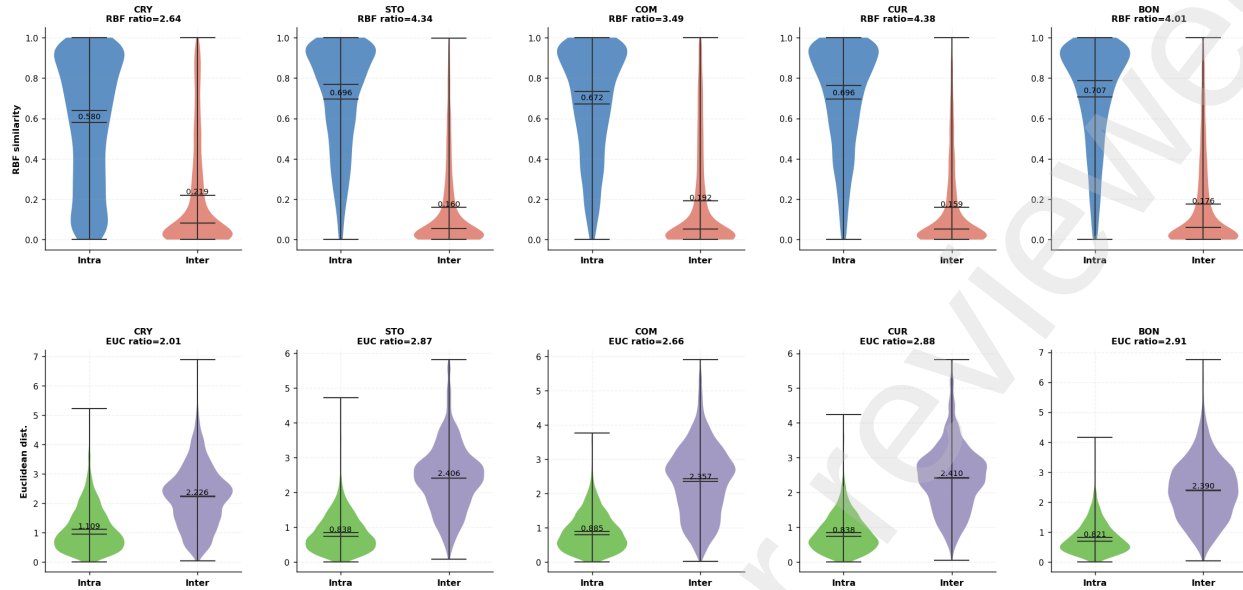
Note: Descriptive statistics of the five IRP dimensions computed separately for each GMM cluster by pooling all daily observations across assets within that cluster. Each $IRP_{i,t}^k$ measures the daily return of asset i relative to the best-performing asset within benchmark class k , as defined in equation (2). Mean and Median report location moments; Std is the sample standard deviation; Skewness and Kurtosis are standardised third and fourth central moments (excess kurtosis); % Negative is the fraction of observations strictly below zero; Q5% and Q95% are the 5th and 95th percentiles. The three panels reveal a monotone gradient in relative underperformance intensity from C_0 to C_2 , consistent with the cluster ordering established in Section 2.

Figure 4: Cluster-conditional distributions of IRP



Note: Each panel shows the distribution of one IRP dimension or conditional volatility separately for the three clusters (C_0 : blue; C_1 : orange; C_2 : green). Dotted vertical lines mark the cluster mean for each dimension. The three clusters exhibit distinct distributions across all five IRP dimensions, consistent with the interpretation that they represent different investor types. The conditional volatility panel is computed on the Crypto (CRY) IRP dimension.

Figure 5: Intra-cluster vs. inter-cluster cohesion in the IRP feature space



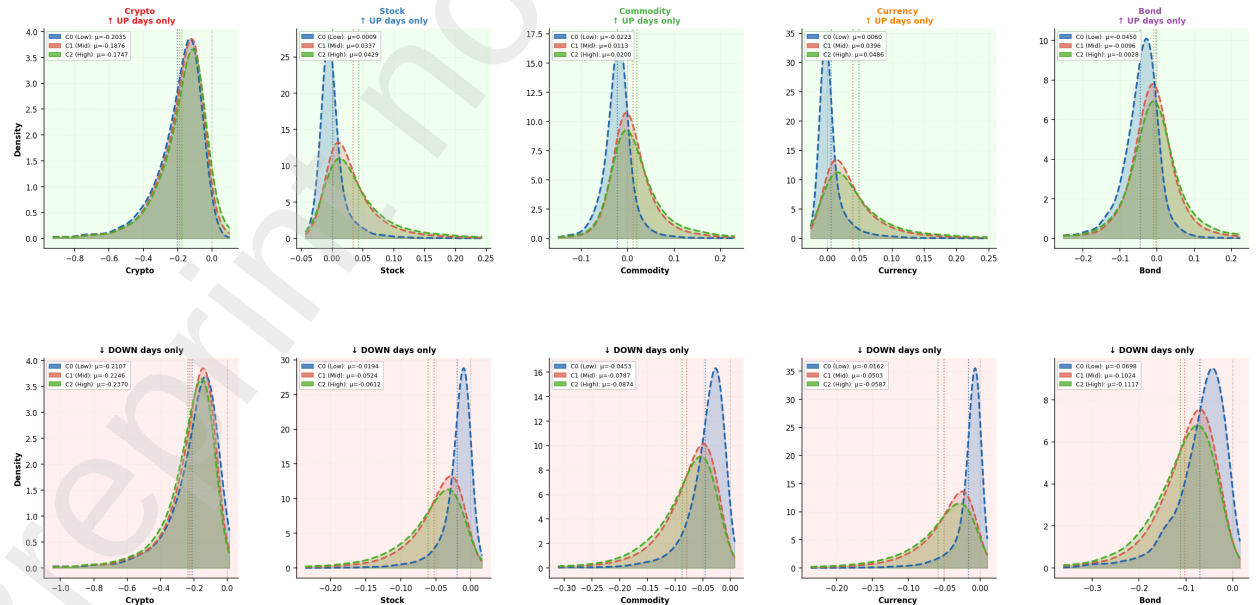
Note: Violin plots comparing pairwise intra-cluster scores (left violin in each panel, assets in the same cluster) and inter-cluster scores (right violin, assets in different clusters) across all five IRP dimensions. The top row uses RBF kernel similarity and the bottom row uses Euclidean distance, both computed on standardised IRP features. Annotated values report distribution means. A cohesion ratio above 1.0 indicates that assets within a cluster are more similar to each other than to assets outside it.

Table 3: IRP asymmetry between up-days and down-days

Cl.	Dim.	Up Days		Down Days		Gap	t	d
		Mean	Std	Mean	Std			
C_0	CRY	-0.2035	0.1726	-0.2107	0.2128	0.0072	3.879	0.037
	STO	0.0009	0.0329	-0.0194	0.0306	0.0203	66.612	0.640
	COM	-0.0223	0.0460	-0.0453	0.0474	0.0230	51.377	0.494
	CUR	0.0060	0.0318	-0.0162	0.0304	0.0222	74.186	0.713
	BON	-0.0450	0.0599	-0.0698	0.0617	0.0249	42.570	0.409
C_1	CRY	-0.1876	0.1783	-0.2246	0.1975	0.0370	34.491	0.196
	STO	0.0337	0.0573	-0.0524	0.0508	0.0861	277.827	1.595
	COM	0.0113	0.0659	-0.0787	0.0631	0.0900	244.115	1.397
	CUR	0.0396	0.0571	-0.0503	0.0521	0.0900	287.396	1.648
	BON	-0.0096	0.0757	-0.1024	0.0745	0.0928	216.220	1.236
C_2	CRY	-0.1747	0.1847	-0.2370	0.1920	0.0623	21.779	0.330
	STO	0.0429	0.0710	-0.0612	0.0602	0.1040	103.122	1.593
	COM	0.0200	0.0774	-0.0874	0.0704	0.1074	95.098	1.459
	CUR	0.0486	0.0707	-0.0587	0.0613	0.1074	105.953	1.633
	BON	-0.0028	0.0873	-0.1117	0.0816	0.1089	84.451	1.293

Note: IRP asymmetry between up-days (positive log-return) and down-days (negative log-return) for each cluster and IRP dimension, pooled across all asset-day observations within each cluster. $\text{Gap} = \Delta^k = \overline{\text{IRP}}_{\text{up}}^k - \overline{\text{IRP}}_{\text{down}}^k$ as defined in equation (8); a positive value indicates that the asset's relative competitive position is better on gain days than on loss days, consistent with the loss-aversion asymmetry predicted by regret theory. t is from a two-sided Welch t -test with unequal variances; d is Cohen's standardised effect size. Cl. = Cluster; Dim. = IRP Dimension.

Figure 6: IRP distributions on up-days and down-days across clusters



Note: Kernel density plots of each IRP dimension on up-days (top row, green background) and down-days (bottom row, red background). The three GMM clusters are overlaid in their respective colours (C_0 : blue; C_1 : orange; C_2 : green). Dotted vertical lines mark the cluster-conditional mean on each day type.

Table 4: RII index: PCA summary, descriptive statistics, and GARCH(1,1) parameters

Panel A: Variance explained				Panel B: PCA loadings							
Cluster	PC1%	PC2%	Comb.%	Cluster	PC	CRY	STO	COM	CUR	BON	
<i>PC1 specification</i>				<i>PC1 specification</i>							
C_0	47.1	—	—	C_0	PC1	−0.1718	0.5704	0.3969	0.5951	0.3652	
C_1	66.7	—	—	C_1	PC1	−0.0538	0.5272	0.4811	0.5319	0.4526	
C_2	72.1	—	—	C_2	PC1	−0.0999	−0.5146	−0.4892	−0.5171	−0.4674	
<i>PC1+PC2 specification</i>				<i>PC1+PC2 specification</i>							
C_0	47.1	19.4	66.5	C_0	PC1	−0.1718	0.5704	0.3969	0.5951	0.3652	
C_1	66.7	19.9	86.6		PC2	−0.9487	−0.1151	−0.2528	−0.0753	0.1309	
C_2	72.1	19.5	91.6	C_1	PC1	−0.0538	0.5272	0.4811	0.5319	0.4526	
					PC2	−0.9978	−0.0294	−0.0485	−0.0331	0.0062	
				C_2	PC1	−0.0999	−0.5146	−0.4892	−0.5171	−0.4674	
					PC2	−0.9938	0.0361	0.0360	0.0409	0.0897	
Panel C: RII descriptive statistics											
Spec	Cluster	Mean	Std	Min	Q1	Med.	Q3	Max	Skew	Ex.K	ρ
PC1	C_0	0.000	1.536	−11.233	−0.620	−0.003	0.621	10.524	−0.014	9.707	0.967
	C_1	0.000	1.827	−20.601	−0.855	0.084	1.013	7.835	−1.401	11.264	0.987
	C_2	0.000	1.899	−21.174	−0.821	−0.005	0.834	16.913	−0.468	18.738	0.990
PC1+PC2	C_0	0.000	1.126	−7.295	−0.505	−0.057	0.443	8.011	0.419	8.288	0.967
	C_1	0.000	1.426	−15.094	−0.710	0.036	0.784	5.967	−1.072	9.079	0.987
	C_2	0.000	1.510	−17.036	−0.646	−0.030	0.637	13.465	−0.476	19.308	0.990
Panel D: GARCH(1,1) parameters on cluster RII											
Spec	Cluster	ω	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha} + \hat{\beta}$	ARCH-LM _{pre}	ARCH-LM _{post}				
PC1	C_0	4.905	0.149	0.841	0.990	0.0000	0.2683				
	C_1	9.311	0.117	0.870	0.986	0.0000	0.9053				
	C_2	9.699	0.160	0.825	0.985	0.0000	0.7585				
PC1+PC2	C_0	4.903	0.165	0.809	0.975	0.0000	0.7731				
	C_1	7.098	0.111	0.865	0.976	0.0000	0.9357				
	C_2	6.108	0.153	0.831	0.984	0.0000	0.6893				

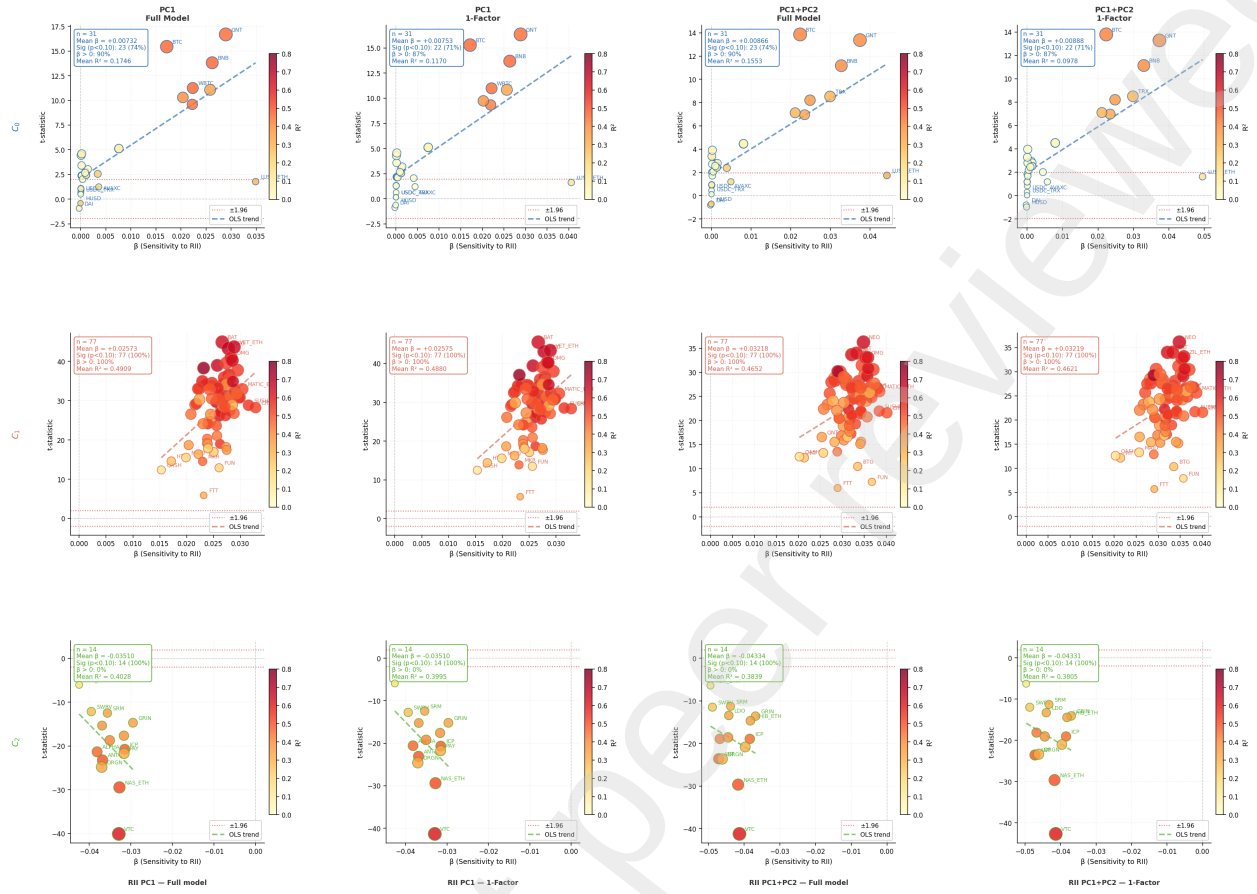
Note: Panel A reports the percentage of variance explained by PC1 and PC2 of the five standardised IRP dimensions, computed separately for each cluster. Panel B reports signed PCA loadings after sign correction so that $\text{RII}\uparrow$ corresponds to greater relative underperformance intensity (CRY loading constrained non-positive). Panel C reports descriptive statistics of the cluster-level daily RII; mean is zero by construction; ρ is the Pearson correlation between RII_{PC1} and $\text{RII}_{\text{PC1+PC2}}$; Med. = median; Ex.K = excess kurtosis. Panel D reports GARCH(1,1) parameter estimates fitted separately on the RII series. $\hat{\alpha} + \hat{\beta}$ is total variance persistence. ARCH-LM_{pre} is the p -value of the ARCH-LM test on the raw RII; rejection confirms conditional heteroskedasticity. ARCH-LM_{post} is applied to the standardised residuals; non-rejection confirms adequate model specification.

Table 5: Cross-sectional distribution of the asset-level return sensitivity coefficient $\hat{\beta}$

Panel A: RII _{PC1}									
Cluster	Distribution of $\hat{\beta}$					% $\hat{\beta} > 0$	% Sig	Mean R^2	$t \text{ (} H_0: \bar{\beta} = 0 \text{)}$
	$\bar{\beta}$	Med.	Std	Min	Max				
<i>Full model</i>									
C_0	0.007 32	0.000 96	0.010 96	−0.000 30	0.034 90	90.3	74.2	0.1746	3.721
C_1	0.025 73	0.026 07	0.003 06	0.015 29	0.032 84	100.0	100.0	0.4909	73.741
C_2	−0.035 10	−0.035 37	0.003 63	−0.042 43	−0.029 50	0.0	100.0	0.4028	−36.200
<i>1-Factor model</i>									
C_0	0.007 53	0.001 21	0.011 41	−0.000 26	0.040 58	87.1	71.0	0.1170	3.677
C_1	0.025 75	0.026 06	0.003 05	0.015 34	0.032 81	100.0	100.0	0.4880	73.978
C_2	−0.035 10	−0.035 21	0.003 62	−0.042 59	−0.029 66	0.0	100.0	0.3995	−36.232
Panel B: RII _{PC1+PC2}									
Cluster	Distribution of $\hat{\beta}$					% $\hat{\beta} > 0$	% Sig	Mean R^2	$t \text{ (} H_0: \bar{\beta} = 0 \text{)}$
	$\bar{\beta}$	Med.	Std	Min	Max				
<i>Full model</i>									
C_0	0.008 66	0.001 06	0.013 29	−0.000 26	0.044 27	90.3	74.2	0.1553	3.629
C_1	0.032 18	0.032 20	0.003 77	0.020 24	0.040 17	100.0	100.0	0.4652	74.926
C_2	−0.043 34	−0.044 02	0.004 09	−0.049 40	−0.036 80	0.0	100.0	0.3839	−39.696
<i>1-Factor model</i>									
C_0	0.008 88	0.001 26	0.013 75	−0.000 25	0.049 48	87.1	71.0	0.0978	3.595
C_1	0.032 19	0.032 20	0.003 75	0.020 31	0.040 07	100.0	100.0	0.4621	75.370
C_2	−0.043 31	−0.043 62	0.004 08	−0.049 68	−0.036 99	0.0	100.0	0.3805	−39.729

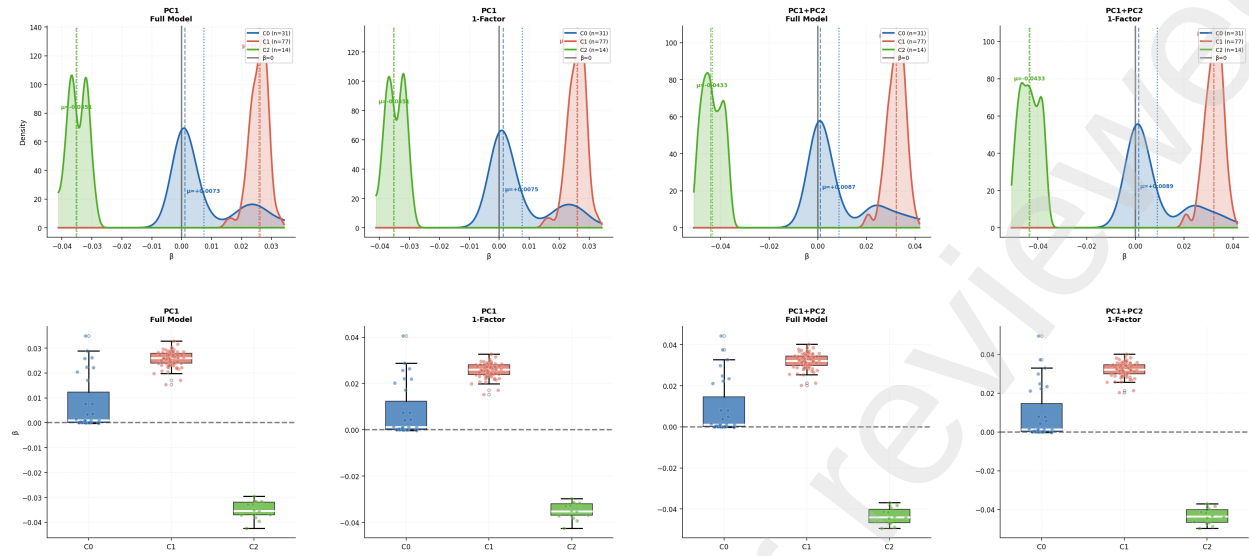
Note: Cross-sectional distribution of the asset-level return sensitivity coefficient $\hat{\beta}$ to the cluster RII, estimated by OLS with HAC standard errors using 5 Newey-West lags. Panel A uses RII_{PC1}; Panel B uses RII_{PC1+PC2}. $\bar{\beta}$ and Med. report the cross-sectional mean and median. % $\hat{\beta} > 0$ is the fraction of assets with a positive sensitivity coefficient; % Sig is the fraction significant at $p < 0.10$. Mean R^2 is the average in-sample fit. $t(H_0: \bar{\beta} = 0)$ tests whether the cross-sectional mean differs from zero. The Full model is estimated from equation 12; the 1-Factor model is estimated from equation 13.

Figure 7: Asset-level return sensitivity to the RII



Note: Each panel plots the asset-level $\hat{\beta}$ against its HAC t -statistic for one cluster and one model specification. Bubble size is proportional to the absolute t -statistic and colour indicates R^2 (yellow to red). The solid vertical line at $\hat{\beta} = 0$ separates assets with positive sensitivity ($\hat{\beta} > 0$) from those with negative sensitivity ($\hat{\beta} < 0$); dotted horizontal lines mark ± 1.96 . The dashed trend line is an OLS fit across assets within each panel. The stats box reports the number of assets, mean $\hat{\beta}$, significance rate, fraction with positive sensitivity, and mean R^2 .

Figure 8: Asset-level $\hat{\beta}$ kernel density across model specifications



Note: Top row: kernel density estimates of $\hat{\beta}$ overlaid across clusters (C_0 : blue; C_1 : orange; C_2 : green); dotted lines mark cluster means; dashed lines mark cluster medians; the solid vertical line at $\hat{\beta} = 0$ separates assets with positive sensitivity from those with negative sensitivity. Bottom row: boxplot with jittered strip overlay; the white horizontal line marks the median; the dashed line marks $\hat{\beta} = 0$. Columns correspond to the four model specifications: RII_{PC1} Full model, RII_{PC1} 1-Factor, RII_{PC1+PC2} Full model, and RII_{PC1+PC2} 1-Factor.

Table 6: Quantile regression: RII coefficient across the return distribution

Panel A: $\text{RII}_{\text{PC1}} - C_0$ (Low IRP)						
τ	Full Model			1-Factor		
	$\hat{\beta}$	t	Ps- R^2	$\hat{\beta}$	t	Ps- R^2
OLS	0.0104	46.999	0.8339	0.0104	48.335	0.8332
0.05	0.0115	69.203	0.7230	0.0117	73.411	0.7212
0.10	0.0115	116.165	0.6807	0.0116	108.503	0.6777
0.25	0.0113	125.728	0.5951	0.0113	126.140	0.5924
0.50	0.0106	109.180	0.5243	0.0106	108.402	0.5232
0.75	0.0099	87.725	0.5206	0.0100	88.062	0.5185
0.90	0.0090	47.562	0.5605	0.0100	44.682	0.5464
0.95	0.0083	31.234	0.5935	0.0099	38.017	0.5638

Panel B: $\text{RII}_{\text{PC1}} - C_1$ (Mid IRP)						
τ	Full Model			1-Factor		
	$\hat{\beta}$	t	Ps- R^2	$\hat{\beta}$	t	Ps- R^2
OLS	0.0257	119.450	0.9444	0.0257	120.074	0.9440
0.05	0.0269	228.003	0.8749	0.0269	226.140	0.8747
0.10	0.0268	260.019	0.8551	0.0268	259.746	0.8547
0.25	0.0264	241.899	0.8120	0.0264	237.077	0.8115
0.50	0.0260	196.079	0.7669	0.0260	196.877	0.7669
0.75	0.0253	157.808	0.7288	0.0255	154.116	0.7271
0.90	0.0242	79.101	0.6957	0.0246	81.924	0.6892
0.95	0.0229	40.504	0.6799	0.0238	45.692	0.6684

Panel C: $\text{RII}_{\text{PC1}} - C_2$ (High IRP)						
τ	Full Model			1-Factor		
	$\hat{\beta}$	t	Ps- R^2	$\hat{\beta}$	t	Ps- R^2
OLS	-0.0351	-172.538	0.9648	-0.0351	-178.225	0.9648
0.05	-0.0359	-301.594	0.8948	-0.0359	-305.440	0.8948
0.10	-0.0360	-308.777	0.8706	-0.0360	-307.215	0.8706
0.25	-0.0356	-275.817	0.8311	-0.0356	-274.702	0.8311
0.50	-0.0355	-270.718	0.7981	-0.0355	-270.757	0.7980
0.75	-0.0353	-184.248	0.7744	-0.0352	-172.096	0.7736
0.90	-0.0343	-97.466	0.7652	-0.0346	-90.108	0.7642
0.95	-0.0337	-57.437	0.7682	-0.0340	-62.306	0.7668

Panel D: $\text{RII}_{\text{PC1+PC2}} - C_0$ (Low IRP)						
τ	Full Model			1-Factor		
	$\hat{\beta}$	t	Ps- R^2	$\hat{\beta}$	t	Ps- R^2
OLS	0.0134	28.672	0.7449	0.0134	29.177	0.7428
0.05	0.0128	31.337	0.5253	0.0139	23.228	0.4894
0.10	0.0133	43.730	0.4964	0.0144	41.018	0.4710
0.25	0.0141	89.301	0.4500	0.0141	84.894	0.4391
0.50	0.0136	89.192	0.4240	0.0138	88.313	0.4209
0.75	0.0135	80.588	0.4632	0.0137	83.737	0.4589
0.90	0.0126	49.143	0.5429	0.0139	55.515	0.5235
0.95	0.0118	36.984	0.5952	0.0133	28.552	0.5499

Panel E: $\text{RII}_{\text{PC1+PC2}} - C_1$ (Mid IRP)						
τ	Full Model			1-Factor		
	$\hat{\beta}$	t	Ps- R^2	$\hat{\beta}$	t	Ps- R^2
OLS	0.0323	70.852	0.9092	0.0323	71.557	0.9087
0.05	0.0332	58.852	0.7311	0.0332	47.469	0.7282
0.10	0.0339	101.432	0.7378	0.0337	99.701	0.7346
0.25	0.0336	167.716	0.7291	0.0336	162.055	0.7254
0.50	0.0333	178.037	0.7102	0.0332	178.840	0.7089
0.75	0.0325	144.751	0.6923	0.0326	146.543	0.6910
0.90	0.0310	86.833	0.6763	0.0317	84.125	0.6696
0.95	0.0297	55.051	0.6697	0.0305	56.109	0.6554

Panel F: $\text{RII}_{\text{PC1+PC2}} - C_2$ (High IRP)						
τ	Full Model			1-Factor		
	$\hat{\beta}$	t	Ps- R^2	$\hat{\beta}$	t	Ps- R^2
OLS	-0.0435	-115.490	0.9362	-0.0436	-116.648	0.9361
0.05	-0.0450	-202.581	0.8592	-0.0450	-202.631	0.8592
0.10	-0.0449	-223.063	0.8268	-0.0448	-218.627	0.8266
0.25	-0.0448	-226.387	0.7755	-0.0448	-224.960	0.7753
0.50	-0.0446	-189.948	0.7281	-0.0445	-190.116	0.7281
0.75	-0.0436	-134.249	0.6969	-0.0437	-135.477	0.6968
0.90	-0.0423	-63.846	0.6852	-0.0430	-62.571	0.6822
0.95	-0.0411	-35.754	0.6878	-0.0418	-42.636	0.6852

Note: Quantile regression estimates of $\hat{\beta}(\tau)$ for $\tau \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$. The OLS row serves as a reference for the conditional mean. Panels A–C use RII_{PC1} ; Panels D–F use $\text{RII}_{\text{PC1+PC2}}$. The Full model specification is defined in equation 14; the 1-Factor variant retains only the RII. Ps- R^2 is the Koenker–Bassett pseudo- R^2 for quantile rows and the standard R^2 for the OLS row. The sign of $\hat{\beta}(\tau)$ is uniform across all quantiles within each cluster: positive for C_0 and C_1 , and negative for C_2 across the full return distribution.