Investors' Behavior in Cryptocurrency Market

[1], Stelios Arvanitis[1], Nikolas Topaloglou[2], and Georgios Tsomidis[3]

[1]Athens University of Economics and Business
[2]IPAG Business School and Athens University of Economics and Business
[3]Athens University of Economics and Business

Abstract

We investigate whether Cryptocurrencies enhance optimal portfolio performance for the most prominent investor personae in the Behavioral Finance literature, namely, the Cumulative Prospect Theory, the Markowitz and the Loss Averse types of investors. We frame our analysis on the grounds of risk aversion w.r.t. perceived returns, and thus remain consistent with Second order Stochastic Dominance. Using the Stochastic Spanning criterion, we construct optimal portfolios with and without cryptocurrencies, allowing however for local non-stationarities and bubbles in the dynamics of the returns process. Our out of sample comparative performance analysis indicates that investors impression of gains and losses affects significantly the aggregate performance of optimal portfolios and that cryptocurrencies are an attractive option for the examined investor types.

Keywords and phrases: Parametric and Non-parametric tests, Second Order Stochastic Dominance, Stochastic Spanning, Cumulative Prospect Theory, Loss Aversion, Markowitz Theory, Probability Weighting, Linear Programming, Portfolio Performance, Cryptocurrencies, Bubbles, Mildly Explosive dynamics.
JEL Classification: C12, C13, C15, C44, D81, G11, G14.

1 Introduction

The issue of whether cognitive biases in the perception of gains and losses actually affect the decision-making process when forming an investment strategy, is a point of disagreement between Classical and Behavioral Finance (henceforth BF). Under Classical Finance, the formation of portfolios is based on investors' preferences towards risk; the usual standard being global risk aversion. Under BF, attitude towards risk is the impact of investors' sentiment (behavioral trigger), as well as their perception of gains and losses. It is now experimentally recognized that when individuals face a prospect, they assess subjectively and hence differently (most likely) the potential outcomes (Kahneman and Tverksy, 1979). Given this subjectivity, their choices usually differ.

The normative classical approach is to regard each investor as rational (von Neumann and Morgenstern, 1953), and the Market as efficient (Bachelier, 1900). Thereby, cognitive biases, i.e. systematic deviations from rationality, are *de jure* neglected and investors are treated as perfectly tuned automata. They are never being carried away by their emotions, and their reasoning is always objective and grounded on the available information. Their goal is to maximize their portfolios' expected utility under their budget constraints and their dispositions towards risk (Markowitz, 1952b).

A large body of positive empirical evidence (among others Kahneman and Tversky, 1974; Kahneman, Knetsch & Thaler, 1991) has however revealed that cognitive biases are main drivers of financial decision-making. They are expressed via subjective *heuristics-decision shortcuts* (Camerer et al, 1998), that have the advantage of neutralizing emotions (doubts, etc.), in decision making.

We are interested in employing biases, under the form of heuristics, in an empirical application that involves traditional asset classes, such as stocks and bonds, along with modern fintech products such as cryptocurrencies. We are trying to assess whether the introduction of heuristics affects the construction and performance of optimal portfolios, when those are allowed to contain assets from the expanded asset class universe.

We employ heuristics in the form of three different investor personae on the cross section of the asset returns. Those arise from prominent theoretical frameworks in the BF literature; namely, Cumulative Prospect Theory (CPT, Tversky and Kahneman 1992; henceforth TK, 1992), Markowitz (Markowitz, 1952 a & b) and Loss Aversion (henceforth LA, Benartzi and Thaler, 1995; Barberis Huang and Santos, 2002). CPT and LA employ subjective probability transformations along with loss aversion; those are fundamental elements of both theories. In the case of the Markowitz persona, probability distortions are not part of the relevant theoretical context. A Markowitz investor exhibits risk aversion for losses and is a risk seeker for gains. A key behavioral aspect about LA is that it minimizes future regret, i.e. the frustration related to the belief that a different action would lead to a more desirable outcome.

For both CPT and LA, loss aversion and probability distortion stem from experimentally and empirically observed phenomena of choice (resulting from cognitive biases) such as narrow framing, overconfidence and mental accounting together with myopic investment decisions. For LA we follow the work of Benartzi and Thaler (1995), who argue that investors are primarily interested in avoiding losses and when they do, they search for ways to improve their investing performance. Availability of time also plays a significant role in decision making, especially when it comes to the discretion of a strategy to mature. When there is "time shortage", effort reasoning is disabled and heuristics take action. Heuristics are decision making shortcuts supported by sentimental factors such as anxiety, herd behavior etc. When there is abundance of time, temperance and prudence can be the main drivers of decision making, resulting into more sober choices and/or practises.

In general, short term gambles are supported by risk seeking behaviors (convex part in the value function), while long term gambles by risk averting (concave part in the value function). One way or another, there is strong evidence that loss aversion is always present, making the value function steeper near the origin in the losses domain. Loss aversion is a documented phenomenon of choice where potential or realized losses loom larger than gains. Thereby, while classical theory in Finance sets the risk averse type of investor as the norm, CPT, LA and Markowitz personae partly emerge descriptively.

For the CPT persona we deploy its whole theoretical framework, which is the combination of the value function with the relevant probability weighting functions (henceforth PWF), and we do it on the cross-section of returns. These PWF are actually subjective transformations of objective probability distributions and are used to form the so-called "capacities" (i.e. decision weights). These distortions result in over-weighting small probabilities of large gains and under-weighting moderate and large probabilities of losses (TK, 1992). Both the value function and the PWF are defined simultaneously over the losses and the gains domain while their combination gives rise to specific attitudes towards risk (Baucells and Heukamp, 2006), as well as to behavioral/cognitive biases. Some of the most common are: the disposition effect, mental accounting, narrow framing and overconfidence (Barberis and Thaler, 2003). We provide their meanings by the way we implicitly incorporate (i.e. embedding) them throughout the paper and discuss how they are likely to play a significant role in the investors' decisions and subsequently on their portfolios' performance. For example, the disposition effect is the observed phenomenon where investors tend to keep assets that have lost their value and sell the ones whose values have increased. This situation creates future lower returns, initiating stronger aversion to losses and enhancing risk seeking behaviors in order to mitigate portfolio's potential out-of-target performance. In Finance, due to the relation of the field with risk, PWF play a more important role than loss aversion. However, many scholars argue that loss aversion is more than enough in "capturing" real time financial behavior (Barberis and Thaler, 2003). Nevertheless, in our work we employ them both and assess their outcomes.

All three experiments focus on the formation of buy-and-hold portfolios, where weights are chosen on a business day basis (myopic framing). Re-balancing occurs at every point in time (every day), apart from some initial "training period". This does not necessarily reflect actual practises; in several cases re-balancing occurs once and annually, when investors take the report sheet of the annual performance of their investments (Barberis and Thaler, 2003). How investors *frame* gains and losses is plausibly influenced by the way relevant information is presented to them. One could argue that our approach introduces *myopic-loss aversion* (Benartzi and Thaler, 1995), because of the high frequency of valuation combined with loss aversion, in the CPT case. However, we point out that we are not interested in re-balancing itself, but rather on the behavior (actions) of the three investor types, through the choices of the weights in their portfolios.

The first portfolio universe (henceforth Traditional) contains only traditional assets (stocks and bonds). The expanded second portfolio universe (henceforth Augmented) additionally includes the four largest Cryptocurrencies by market capitalization (BTC: Bitcoin, ETH: Ethereum, LTC: Litecoin and XRP: Ripple). Cryptocurrencies tend to have an asymmetric risk profile because they exhibit an extremely high return volatility together with positive skewness and kurtosis (see Table 4 of descriptive statistics). Furthermore, their dynamic behavior is consistent with the existence of bubbles and mild explosivity (see Anyfantaki et al. (2021)-AAT21). Following the reasoning of Odean (1999), we analogously believe that cryptocurrencies have raised an attention effect that has resulted to excessive trading (i.e. overconfidence bias). Due to their extreme past performance they have caught the attention of investors. Despite this, it is also possible that their markets would be non-attractive for "rational" highly risk averting investors, due to the characteristics of the dynamics of their returns, their complicated technology and their connection to the not (yet?) mainstream decentralized Finance.

Thus, the comparative consideration of those portfolio universes is suitable for analysis that includes several cognitive biases like the above. For example "narrow framing" is triggered because both portfolio universes are treated in isolation from the rest of the market; we remain indifferent about the markets' aggregate performance. It is a fact that the market's aggregate performance affects investor's sentiment, and subsequently her actions, though the herd behavior mechanism. The taxonomy to Traditional and Augmented, could also introduce "mental accounting" because we classify their performance separately, irrespective of the fact that they both contain "traditional" base assets.

The optimal portfolios in each universe are constructed in the framework of Second order Stochastic Dominance (SSD) (Levy, 2006) and Stochastic Spanning (henceforth StSp) (Arvanitis et al. 2019), when those however are applied on the persona distorted value of the base assets returns. We thus work under the assumption that each persona decides portfolio weights optimally, exhibiting non satiation and risk aversion. Those however are not evaluated on the distribution of the original return process, but on its transformation according to the persona's cognitive biases. We work in the non-parametric framework of SSD since we want to remain agnostic on the non satiation and risk aversion characteristics of investors' preferences. Optimal weights are constructed for both portfolio types via the nested optimizations related to the construction of the functional that decides whether the traditional portfolio universe (empirically) stochastically spans the augmented one. Whether spanning is the case or not, the resulting optimal weights for both types can be perceived as optimal choices of an optimal risk averse utility.

In order to avoid conceptual and statistical problems with non-stationary returns, we use the AAT21 extension of the SSD and StSp, to the mildly explosive framework for logarithmic prices. This allows for the existence of multiple and possibly interconnected non-stationary bubbles, under appropriate limiting sparsity conditions that render them asymptotically negligible. Exploiting continuity properties of the involved value transformations, we show that the related statistical machinery for stochastic spanning is usable in the present setting.

Summarizing, for each personae we construct empirically optimal Augmented and Traditional portfolios using the StSp machinery, that are business-daily re-balanced, in a mildly explosive framework of asset returns. We assess their out-of-sample comparative performance, using a battery of non-parametric and parametric tests. To the best of our knowledge this combination of BF personae value transformations with stochastic spanning is novel. Our empirical results suggest that the Augmented portfolios outperform the optimal Traditional ones.

The rest of the paper is organized as follows: in the following section we describe in detail our methodology, derive the limit theory that supports it, and present our empirical findings. In the final section we conclude while also discussing some behavioral implications. In the appendices we provide the proofs of our asymptotic analysis, as well as further information related to our empirical work.

2 Methodology and Empirical Results

Our empirical application begins with the manipulation of the dataset, which covers the period from mid-August 2015 to end-of August 2021. Our dataset contains business days closing prices for the following "traditional" assets (stock and bond indices): The S&P 500 Index, the Barclays US Bond Index, the 1 Month T-Bill, the Russell 2000 Index, the Vanguard Value Index, the Vanguard Small-Cap Index, and the dynamic trading strategies SMB and HML. It additionally contains the aforementioned four cryptocurrencies. We first "align" the dataset by keeping the closing prices of cryptocurrencies for the business days only and afterwards we estimate the returns of all the assets. This is because the most common practise for investors is to consider the aggregate performance of specific assets, or groups of assets.

We apply on the dataset the distorted value transformation associated with each BF personae. We conduct both in-sample as well as out-of sample tests. In the out-of-sample analysis, we solve the analogous stochastic spanning optimization problem in a rolling window pattern - thus obtain the empirically optimal portfolios - and then assess their performance. At each stage we exemplify related theoretical concepts and subsequently present the relevant empirical results.

2.1 Distorted Value Function Transformation

We are "distorting" data according to cognitive biases that fall within any of the aforementioned personae. The data transformations are employed on the cross section of returns. A new dataset is thus obtained where now the returns have been transformed into the relevant "behavioral" values, at which optimal investment strategies are then decided and executed.

We consider, for analytical convenience, a general form of distorted value function transformation (DVT); $v : \mathbb{R}^d \to \mathbb{R}^d$, $v_j(x) := c_j x_j^{\alpha_j} \mathbb{I}(x_j \geq 0) + c_j^\star (L(x_j))^{\beta_j} \mathbb{I}(x_j < 0)$, where $c_j, c^\star_(j), \alpha_j, \beta_j$, $j = 1, \ldots, d$, are appropriate constants and $L : \mathbb{R} \to \mathbb{R}$ is a linear transformation. It encompasses all three transforms that we analytically present below, and thus facilitates derivations.

2.1.1 Cumulative Prospect Theory type of persona

CPT bypasses the drawbacks of Prospect Theory (henceforth PT) (Kahneman and Tversky, 1979). PT employs a monotonic transformation of outcome probabilities and this monotonic transformation can not be applied to prospects (i.e. portfolios) with any number of outcomes (is restricted to two). Also, PT does not always satisfy Stochastic Dominance (SD) and hence investors may choose dominated prospects (a mild form of irrationality).

Each of the d base assets is considered as a risky prospect. The set of risky projects is considered endowed with the discrete uniform distribution. Decision weights (or "capacities") which are non additive set functions that generalize the standard notion of probabilities, are applied to the prospects. Specifically, in the definitions below, the capacity π_i^+ mathematically exemplifies the assertion "the outcome is at least as good as x_i" minus "the outcome is strictly better than x_i". Analogously, π_i^- denotes that "the outcome is at least as bad as x_i" minus "the outcome is strictly worse than x_i", with x_i being the return of the i^{th} prospect. In order to specify the above, we introduce transformations w^+, w^- which are both non-decreasing, while being inverse S-shaped respectively with $w^+(0) = 0 = w^-(0)$ and $w^+(1) = 1 = w^-(1)$ such that $w^+, w^- : [0,1] \to [0,1]$.

Given the cross-section of returns, we initially rank the latter in increasing order for every t. Suppose that m of these returns are negative, while the remaining $n = d - m$ are positive. We label the sorted returns using r_{-m} for the minimum negative return and through to r_n for the maximum positive. The cross-sectional return distribution is then described by the array of pairs,

$$(r_{-m}, p_{-m}; r_{-m+1}, p_{-m+1}; ...; r_{-1}, p_{-1}; r_0, p_0; ...; r_{n-1}, p_{n-1}; r_n, p_n),$$

where p_j is the associated probability from the discrete uniform. Given w^+, w^-, the capacities are then defined by

$$\pi_n^+ = w^+(p_n) \quad \text{and} \quad \pi_{-m}^- = w^-(p_{-m})$$

$$\pi_i^+ = w^+(p_i + ... + p_n) - w^+(p_{i+1} + ... + p_n) \quad \text{where} \quad 0 \leq i \leq \ n$$

and

$$\pi_j^- = w^-(p_{-m} + ... + p_j) - w^-(p_{-m} + ... + p_{j-1}) \quad \text{where} \quad -m \leq j < 0$$

In our analysis the PWF used is,

$$w^+(p) = \frac{p^{0.61}}{[p^{0.61} + (1-p)^{0.61}]^{\frac{1}{0.61}}} \quad \textit{and} \quad w^-(p) = \frac{p^{0.69}}{[p^{0.69} + (1-p)^{0.69}]^{\frac{1}{0.69}}}.$$

This is justified by Ingersoll (2008), who provides a proof that for a range of values, the exponents in the PWF can induce non monotonicity. This can lead to negative decision weights and preference for first order stochastically dominated (FSD) prospects. As Ingersoll (2008) recomments, the problem is tackled if both exponents of w^+ and w^- are strictly above 0.279. Moreover, TK (1992) propose the following *S-shaped* functional form for the value function $u : \mathbb{R} \to \mathbb{R}$,

$$v_{\mathrm{P}}(x) = \begin{cases} \mathrm{x}^{\alpha} & \text{if} \quad x \geq 0 \\ \text{-}\lambda(-x)^{\beta} & \text{if} \quad x < 0 \end{cases},$$

where λ is the loss aversion coefficient ($\lambda = 2.25$, because losses 'hurt' twice as more than gains on average, TK 1992) and α, $\beta < 1$; where we use the same values as TK (1992), and set $\alpha = \beta = 0.88$.

We employ the estimates of TK (1992) for the parameter and the PWF exponent. Subsequent to TK (1992), several papers have used more sophisticated techniques, in conjunction with new experimental data, to estimate these parameters (Gonzalez and Wu 1999; Abdellaoui et al. 2007). Their estimates are similar to those obtained by TK (1992). Since in our experiments, we do not need to estimate the overall value of a mixed prospect but rather the individual "behavioral" value of every asset, the transformation $v_{\mathrm{P}} : \mathbb{R}^d \to \mathbb{R}^d$ is

$$v_{\mathrm{P}}(x) := (\pi_i^+ v_{\mathrm{P}}(x_i), -m \leq i < 0, \pi_j^- v_{\mathrm{P}}(x_j), 0 \leq j < n).$$

For $c_j := \pi_j^+$, $c_j^\star := -\lambda\pi_j^-$, $\alpha_j = a$, $\beta_j = b$ and $L(x) = -x$ our general DVT reduces to CPT.

2.1.2 Markowitz type of persona

We follow the work of Markowitz (1952a) who argues that the utility function of Friedman and Savage (1948) must have convex as well as concave segments near the point of origin. In this regard an inverse S-shaped value function emerges where the curvature changes at the point of origin which separates gains from losses. This type of investor is averting losses in the losses domain and seeking risk in the gains domain. Investors employ a reverse S-shaped value function, when it comes to bets whose outcomes are moderate and not too extreme. Under this context, no PWF is applied nor any loss aversion. We thus employ the following transformation $v_{\mathrm{M}} : \mathbb{R}^d \to \mathbb{R}^d$:

$$v_{\mathrm{M},j}(x) = \begin{cases} \mathrm{x}_j^\alpha & \text{if} \quad x_j \geq 0 \\ \text{-(-}\mathrm{x}_j)^\beta & \text{if} \quad x_j < 0 \end{cases}, \ j = 1, \ldots, d,$$

where α, $\beta > 1$; thus, in order to satisfy the curvature as described in theory, we are allowed to set $\alpha = \beta = 2$, without harming generality. For $c_j := 1$, $c_j^\star := -1$, $\alpha_j = a$, $\beta_j = b$ and $L(x) = -x$ our general DVT reduces to the Markowitz transform.

2.1.3 Loss averse type of persona

Following the work of Barberis, Huang and Santos (2002) we employ a value function that refers to investors who specialize in averting losses. The aforementioned authors argue that loss aversion is by itself enough in capturing and explaining major phenomena of choice, such as the disposition effect, i.e. the tendency to keep assets that have lost their value and sell assets that have increased their value, and that its magnitude depends on the size of prior gains and/or losses. Moreover, Benartzi and Thaler (1995) use the loss aversion behavioral bias to explain the equity premium puzzle. They find that investors evaluate the aggregate stock market by computing the PT value of its historical distribution. Loss aversion is regarded as the most important bias in the BF literature and the novelty of our approach stems from its use on the cross-section of returns. Avoiding losses is a major concern in investing because if they are realized, they affect expectations (developed endogenously) about future performance. These rational expectations serve as "anchors" (Kőszegi and Rabin, 2006), provide *wishful* targets to be met, and thus operate as bounds between gains and losses. We work with the following LA type of transformation $v_{\mathrm{LA}} : \mathbb{R}^d \to \mathbb{R}^d$::

$$v_{\mathrm{LA},j}(x) = \begin{cases} \mathrm{x}_j & \text{if} \quad x_j \geq 0 \\ 2.25\mathrm{x}_j & \text{if} \quad x_j < 0 \end{cases}, \ j = 1, \cdots, d.$$

Under this setting, the value function is piecewise with its lower set being convex. It is increasing, implying explicitly that investors are non-satiated, while its segmental linearity can implicitly imply risk aversion. It is kinked at the origin, where gains are zero which is the most often used boundary between gains and losses. For $c_j := 1$, $c_j^\star := \lambda$, $\alpha_j = 1$, $\beta_j = 1$ and $L(x) = x$ our general DVT reduces to LA.

2.2 Mildly explosive framework with multiple bubbles

In the next stage we employ the concept of spanning on the personae transformed dataset. The incorporation of cryptocurrencies' returns suggests that we may need to account for potentially parodically explosive dynamics in the associated DGP of base asset returns as far as the limiting properties of our statistical procedures are concerned. In this paragraph we describe our assumption framework that deals with this.

We work with a portfolio space defined as the set of positive convex combinations of the DVT transform on the returns of d base assets and represented by the $d-1$ simplex $\left\{\boldsymbol{\lambda} \in \mathbb{R}_+^d : \boldsymbol{\lambda}^T \mathbf{1}_d = 1\right\}$. The framework is broad enough to allow for base assets that are themselves constructed via complicated portfolio constraints on deeper underlying individual securities, like short sales, position limits, and restrictions on factor loadings. Some of the base assets are cryptocurrencies. Their empirical locally explosive behavior necessitates the considerations of (locally) non stationary return processes. We employ the framework of AAT21 of multivariate mildly explosive AR(1) processes for the stochastic processes that represent logarithmic prices. We employ almost exact versions of their Assumptions ME and AN. This is empirically justified in the following paragraph. In what follows and depending on the context, $\|\cdot\|$ denotes either the Euclidean norm on $\mathbb{R}^d$ or the Frobenius norm on the space of $d \times d$ real matrices. Moreover, c denotes a generic positive constant that may assume different values in different occurrences. We also denote with δ some constant greater than or equal to $\max(\max_j \max(\alpha_j, \beta_j), 1)$.

Assumption ME.

1. $(\boldsymbol{\varepsilon}_t)_{t\in\mathbb{N}}$ *which is an* $\mathbb{R}^d$*-valued stationary and strong mixing process with mixing coefficient sequence* $(\alpha_m)_{m\in\mathbb{N}}$ *that satisfy* $\alpha_m = O(m^r)$ *for some* $r > 1$*. Furthermore,* $\exists L, q > 0$ *such that for large enough* $t > 0$, $\mathbb{E}[\exp(t\|\boldsymbol{\varepsilon}_0\|)] \leq \exp(Lt^q)$.

2. *For the sample size* $T \in \mathbb{N}^\star$, $\{0, \ldots, T\}$ *is partitioned in* K *mild-explosivity periods* B_k, $k = 1, \ldots, K$ *and the remaining stationary periods* $\cap_{k=1}^K B_k^c$.

3. *The logarithmic prices* $\mathbb{R}^d$*-valued process sequence satisfies the recursion* $\boldsymbol{X}_t = \left(Id + \sum_{k=1}^K \frac{C_k}{M(T,k)}\mathbb{I}\{t \in B_k\}\right)\boldsymbol{X}_{t-1} + \boldsymbol{\varepsilon}_t$, $t > 0$*, where* C_k *is a positive* $d \times d$ *explosivity coefficient matrix at the* k^{th} *explosive period, and* $M(T,k) > 0$ *and diverging to infinity as* $T \to \infty$*, which represents the rate at which the* k^{th} *explosive behaviour vanishes as a function of* T*. The process is initiated by* $\boldsymbol{X}_0$ *and* $\exists L^\star, q^\star > 0$ *such that for large enough* $t > \delta$, $\mathbb{E}[\exp(t\|\boldsymbol{X}_0\|)] \leq \exp\left(L^\star t^{q^\star}\right)$.

Remark 1. ME.1 allows for a large variety of linear and/or conditionally heteroskedastic models typically used for the stationary parts of logarithmic returns in empirical finance (see, for example, Drost and Nijman, 1993; Basrak et al., 2002). This allows for models that exhibit stationary, ergodic and geometrically mixing temporal dynamics, along with innovation distributions that have densities (see AAT21 for examples). ME.2 allows for the existence of K sub-periods of non-stationary bubbles in parts of the base assets process. It is also allowed that $K \to \infty$ as $T \to \infty$. This implies that the number of bubbles need not asymptotically stabilize, though the following assumption will specify bounds on the intensity of the bubbles as T grows. In the third part, at each bubble period k the structure of the explosiveness coefficient matrix C_k is general enough to allow for intra-bubble dependence of currently explosive base assets on the dynamics of other currently and/or previously explosive assets, as well as on assets that are never explosive. Obviously for the latter the relevant blocks of C_k are zero for all k. The moment generating function conditions that appear in ME.(1),(3) are compatible with sub-Gaussian and sub-exponential distributions for the random variables involved (see, for example, Chapter 2 of Vershynin, 2018)-see AAT21 for a Gaussian example. They do not allow for multivariate distributions that do not possess moment generating functions; e.g. non-Gaussian stable distributions. Actually, due to the presence of δ in their formation, those conditions may be slightly stronger than the analogous ones in AAT21.

The base assets DVT of the locally non-stationary net return process is thus $v(\boldsymbol{R}_t) := v\left(\exp^\star\left(\sum_{k=1}^K \frac{C_k}{M(T,k)}\mathbb{I}\{t \in B_k\}\boldsymbol{X}_{t-1} + \boldsymbol{\varepsilon}_t\right) - \mathbf{1}\right)$, $t > 0$ with $\exp^\star : \mathbb{R}^d \to \mathbb{R}^d$ defined by $\exp^\star(\mathbf{y}) := (\exp(\mathbf{y}_1), \ldots, \exp(\mathbf{y}_d))^{\mathrm{T}}$ and $\mathbb{R}^d \ni \mathbf{1} := (1, \ldots, 1)^T$. $\mathcal{X}$ denotes the convex hull of the union of the supports of the elements of $v(\boldsymbol{R}_t)$. It is bounded from below by $\min_j c_j^\star L(-1)^{\beta_j}$.

Given that K is allowed to diverge, the following assumption prescribes restrictions between the singular values of the explosivity coefficient matrices C_k, the degree of the return to the random walk dynamics $M(T,k)$ and the maximal bubble time instance $\max_{t,k} B_k$. It is identical to the homonymous assumption of AAT21.

Assumption AN. $\exists c > 0$, $\epsilon > 0$ such that $\max_k \frac{\|C_k\|(d+\epsilon)^{\max_{t,k} B_k}}{M(T,k)} \leq c$ for all $T \geq 0$. As $T \to \infty$, $\frac{\max_{t,k} B_k}{\sqrt{T}} = o(1)$.

Remark 2. AN is satisfied when $\|C_k\|$ is bounded in k, $\max_{t,k} B_k \sim c \ln T$ and $M(T,k) = \delta_k T^{a_k}$, with $\min_k \delta_k, \min_k a_k > 0$ (such choices are compatible with the bubble duration conditions of Philips et al., 2015). The condition is not empirically identifiable and it concerns the future behaviour of bubbles. As remarked in AAT21, it allows for slowly diverging bubble durations compatible with future improvements in the technology and the regulation framework associated with the creation and circulation of cryptocurrencies. AAT21 point out that this is compatible with future improvements in the technology and the regulation framework associated with the creation and circulation of cryptocurrencies; e.g. stricter regulation as a result of series of severe bubbles, or the development of derivative markets on cryptocurrencies, as long as those deter investors from forming expectations via bubble producing sunspot processes, or force investors to correct parameters associated with fundamentals. AN also allows for more complicated behaviours; for example it is compatible with unbounded $\|C_k\|$ as k grows, thus intense intra-bubble feedbacks between explosive assets, mitigated by stronger degrees of the return to the random walk dynamics due to learning mechanisms like the above. In any case, AN prescribes the exact conditions on the explosive dynamics parameters of the logarithmic prices process, that imply asymptotic dominance of the stationary dynamics in the formation of the lower partial moments differentials employed in the SD relations.

2.2.1 In sample empirical results: evidence of bubbles and mild explosivity

Philips and Magdalinos (2007) argue that when asset prices exhibit explosive behavior this signals an underlying bubble behavior. Within this framework, we test the existence of multiple speculative bubbles in the crypto assets market by using the PSY methodology (Phillips et al., 2015).

The PSY test, relies on right-tailed Dickey-Fuller tests via a recursive estimation over rolling windows of increasing sizes. It is solely applied to cryptocurrencies since their market experienced several unique turmoils; the relevant literature concurs that this may be a quite strong indication for the presence of bubbles. The smallest window width fraction is manually chosen, in order to initiate computations, while the largest is the total sample size. We follow Phillips and Shi (2020) and use 10% as a starting fraction of the cryptocurrencies dataset. The test's null hypothesis assumes that logarithmic prices have a unit root random walk; the alternative assumes that there exists at least one subperiod exhibiting mild explosivity. When the null is rejected, the procedure is also used for date-stamping multiple bubbles.

The shaded green areas in the figures in Appendix 2 are the identified multiple bubble periods, obtained by using the the 95% bootstrap critical values, for BTC, ETH, XRP and LTC respectively. Moreover, Table 1 reports the total number of explosivity days that each crypto asset exhibits, which stems from the number of times the test statistic exceeds the critical value. BTC exhibits the highest number (365) of explosivity days; it has attracted most invested capital in the crypto market and has been also the starring cryptocurrency for many scams. The second is ETH (335), with LTC (230) and XRP (192) following. The bubble periods for all cryptos do not in general coincide, however they do intersect on numerous dates. On average, it seems that the period starting from early 2017 till end-of 2018 has been for all four cryptocurrencies the period with the highest frequency of multiple bubbles.

Table 1: Number of explosive days

	days
BTC	365
ETH	335
XRP	192
LTC	230

Entries report the number of days characterised by explosive price behaviour according to the BSADF test for the four analysed cryptocurrencies over the period from mid-of August 2015 to end-of August 2021, for a total of 1527 business days returns.

The bubble tests along with their time-stamping provide some evidence in favour of Assumption AN: given the large sample period under study it is seen that for all cryptocurrencies bubbles are sparse enough and of durations consistent with AN.

2.3 Stochastic dominance and stochastic spanning on distorted values

A Second Order Stochastic Dominance (SSD) approach (Levy, 2006) on the behaviorally modified resturns is used due to its non-parametric nature and its relation with risk aversion. To relate risk aversion (i.e. the traditional approach in Modern Portfolio Theory) with the aforementioned personae, we follow various works (Barberis, Mukherjee and Wang, 2016; Koszegi and Rabin 2006, among others) who argue that a better description of reality is the one where behavioral investors pay at least some attention to traditional factors. Additionally, it seems that a significant part of investors exhibit risk aversion. For example, the fact that the riskless interest rate is generally lower than the cost of capital of most firms, is an indication that investors are risk averse and require a risk premium (Levy, 2006).

SD ranks investments based on general regularity conditions (Hadar and Russel, 1969; Hanoch and Levy 1969; Rothchild and Stiglitz, 1970) for decision making under risk and it can been seen as a model-free alternative to Mean-Variance (M-V) dominance (Levy, 2006). SD is quite appealing because it accounts all moments of the returns' distribution without assuming any particular family of distributions nor specific preferences on behalf of agents. Because of its non-parametric nature, it is quite useful for assets with asymmetric risk profiles, like cryptocurrencies.

StSp (w.r.t. SSD) occurs if introducing new securities or relaxing investment constraints does not improve the investment opportunity set, uniformly over the class of increasing and concave utilities.

StSp can be seen as a model-free alternative to M-V spanning (Huberman and Kandel, 1987) that accounts for higher-order moment risk in addition to variance. StSp involves the comparison of two choice sets, not necessarily disjoint (e.g. nested). It evaluates all feasible portfolios, even the ones that include a relatively small number of assets and thus are more susceptible to higher moment risk. In other words, StSp is basically an SSD order-preserving reduction of the portfolio opportunity set.

2.3.1 Stochastic dominance in the mildly explosive framework

We employ SSD in our mild explosivity framework. We follow closely the constructions of AAT21. Our main difference lies in the fact that the associated lower partial moments that define the order, are w.r.t. the DVT of the net returns process, instead of the returns process per se. We thus modify the analogous definitions and notations of AAT21 as follows; for $z \in \mathcal{X}$ and $\boldsymbol{\kappa}, \boldsymbol{\lambda}$ elements of the unit $d-1$ simplex define, and for v the general form of DVT defined above:

$$\begin{gathered} D\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{R}_t\right)\right) := \left(z - \boldsymbol{\kappa}^{\mathrm{T}} v\left(\boldsymbol{R}_t\right)\right)_+ - \left(z - \boldsymbol{\lambda}^{\mathrm{T}} v\left(\boldsymbol{R}_t\right)\right)_+, \\ D_T\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{R}\right)\right) := \tfrac{1}{T} \sum_{t=1}^{T} D\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{R}_t\right)\right), \\ D^{\star}\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{R}\right)\right) := \lim_{T \to \infty} \mathbb{E}\left[D_T\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{R}\right)\right)\right]. \end{gathered}$$

Due to stationarity,

$$D^{\star}\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{\varepsilon}\right)\right) = \mathbb{E}\left[\left(z - \boldsymbol{\kappa}^{\mathrm{T}} v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_0\right) - \mathbf{1}\right)\right)_+\right] - \mathbb{E}\left[\left(z - \boldsymbol{\lambda}^{\mathrm{T}} v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_0\right) - \mathbf{1}\right)\right)_+\right],$$

which is the standard LPM differential employed in SSD, yet here w.r.t. the DVT of the stationary part of the returns.

As in AAT21, the assumptions above, imply an asymptotic negligibility property for the totality of bubble periods, with respect to some of the functionals above, when properly scaled. This is essentially represented by the following general result of central importance, the proof of which is quite similar to the proof of Proposition 3 of AAT21:

Proposition 1. *Suppose that Assumptions ME and AN hold. Then uniformly in $z, \boldsymbol{\lambda}$ as $T \to \infty$,*

$$\mathbb{E}\left[\left|\frac{1}{\sqrt{T}} \sum_{t \in \cup_{k=1}^{K} B_k} \left(z - \boldsymbol{\lambda}^{T} v\left(\boldsymbol{R}_t\right)\right)_+ - \frac{1}{\sqrt{T}} \sum_{t \in \cup_{k=1}^{K} B_k} \left(z - \boldsymbol{\lambda}^{T} v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_t\right) - \mathbf{1}\right)\right)_+\right|\right] = o\left(1\right).$$

Then, Assumption ME, Proposition 1 and dominated convergence imply that the functional $D^{\star}\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{R}\right)\right)$ is well defined, bounded and continuous in $(z, \boldsymbol{\kappa}, \boldsymbol{\lambda})$. We thus apply the AAT21 definition of SSD, which is compatible with the framework of multiple non-stationary bubbles (MESSD), for the DVT of the returns as follows:

Definition 1. $\boldsymbol{\kappa} \succeq_{v\text{-MESSD}} \boldsymbol{\lambda}$ iff $\forall z \in \mathcal{X}$, $D^{\star}\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{R}\right)\right) \leq 0$.

The Cezaro-limit based definition of $D^{\star}$ is similar to Definition 5.1 of Jin *et al.* (2017) that handles distributional heterogeneity in the context of forecast comparison. It corresponds to a limiting

Lebesgue–Stieltjes (discrete) integration (across time) of the LPM differentials that collapses to the standard definition of SSD under stationarity. Then, the auxiliary Proposition 2 (see Appendix) directly implies that:

Corollary 1. *Under Assumptions ME and AN,* $\boldsymbol{\kappa} \succeq_{v\text{-}MESSD} \boldsymbol{\lambda}$ *iff* $\forall z \in \mathcal{X},\ D^\star(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v(\boldsymbol{\varepsilon})) \leq 0$.

Thus, Assumption AN ensures that v-MESSD is essentially (DVT-) SSD between the DVTs of the stationary part of the returns. As in AAT21, the assumption forces the non-stationary contributions to $D_T(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v(\boldsymbol{R}))$ to asymptotically vanish.

2.3.2 Stochastic spanning in the mildly explosive framework

Consider two non-empty subsets of the general portfolio space, $\mathrm{K} \subset \Lambda$, which are also assumed to be closed and simplicial, to facilitate among others the invocation of properties of convex optimisation. We employ the concept of MESSD spanning of AAT21, to the DVTs of the associated returns, obtaining:

Definition 2. $K \succeq \Lambda$ iff $\forall \boldsymbol{\lambda} \in \Lambda, \exists \boldsymbol{\kappa} \in K : \boldsymbol{\kappa} \succeq_{v\text{-MESSD}} \boldsymbol{\lambda}$.

Arguments involving continuity and compactness imply that $K \succeq_{v\text{-MESSD}} \Lambda$ iff

$$\eta^\star := \sup_{\Lambda} \inf_{K} \sup_{\mathcal{X}} D^\star(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v(\boldsymbol{R})) = 0.$$

Then auxiliary Proposition 2 (see Appendix) implies that v-MESSD spanning equivalently holds iff $\eta := \sup_{\Lambda} \inf_{K} \sup_{\mathcal{X}} D(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v(\exp^\star(\boldsymbol{\varepsilon}_0) - \mathbf{1})) = 0$. Consider then statistically testing the hypothesis structure $\mathbf{H_0} : K \succeq_{v\text{-MESSD}} \Lambda$ vs. $\mathbf{H_1} : K \not\succeq_{v\text{-MESSD}} \Lambda$. Since under Assumption AN, the null hypothesis is equivalent to that $\eta^\star = 0$, auxiliary Proposition 3 (see Appendix) and the latency of the stationary part of the logarithmic returns process, $(\varepsilon_t)_t$, imply that, under Assumptions ME and AN, the spanning test statistic of AAT21 evaluated at the DVTs of the non-stationary returns sample is usable. We thus employ a scaled empirical analogue of $\eta^\star$, namely,

$$\eta_T^\star := \sup_{\Lambda} \inf_{K} \sup_{\mathcal{X}} \sqrt{T} D_T(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v(\boldsymbol{R})).$$

The asymptotic decision rule is to reject $\mathbf{H}_0$ in favor of $\mathbf{H}_1$ iff $\eta_T^\star > q(\eta_\infty^\star, 1-\alpha)$, which is the $(1-\alpha)$ quantile of the distribution of the null limiting distribution of the statistic at a significance level $\alpha \in (0,1)$. The quantile is expected to-among others-depend on latent parameters, like the dependence structure of $(\varepsilon_t)_t$ (see also Theorem 1 below and the auxiliary Proposition 4 in the Appendix). As in AAT21, we approximate the quantile via the use of a subsampling procedure.

A choice of the subsampling rate, $1 \leq b_T < T$, generates maximally overlapping subsamples $(R_s)_{s=t}^{t+b_T-1}$, $t = 1, \cdots, T - b_T + 1$. We evaluate the test statistic on each subsample, thereby obtaining $\eta^\star_{b_T;T,t}$ for $t = 1, \cdots, T - b_T + 1$, obtaining $q_{T,b_T}(1-\alpha)$, the $(1-\alpha)$ quantile of the empirical distribution of $\eta^{\star}_{b_T;T,t}$ across the subsamples. Using the above, the modified decision rule is to reject $\mathbf{H}_0$ in favor of $\mathbf{H}_1$ iff $\eta^{\star}_T > q_{T,b_T}(1-\alpha)$.

The following result exemplifies the first order limit theory of the procedure and asymptotically rationalizes the empirical results of the following section. Under Assumptions ME-AN and a standard subsampling rate restriction we obtain:

Theorem 1. *As $T \to \infty$, under Assumptions ME and AN, and if (b_T), possibly depending on $(R_t)_{t=1,\ldots,T}$, satisfies $\mathbb{P}(l_T \leq b_T \leq u_T) \to 1$, where (l_T) and (u_T) are real sequences such that $1 \leq l_T \leq u_T$ for all T, $l_T \to \infty$ and $\frac{u_T}{T} \to 0$ as $T \to \infty$:*

1. *Under $\boldsymbol{H}_0 : K \succeq_{v\text{-}MESSD} \boldsymbol{\Lambda}$,*

$$\eta^{\star}_T \rightsquigarrow \sup_{\Lambda} \inf_{K} \sup_{\mathcal{X}} \mathcal{L}_v\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}\right),\ (z, \boldsymbol{\kappa}, \boldsymbol{\lambda}) \in CS,$$

where the and the limiting Gaussian process $\mathcal{L}_v\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}\right)$ is defined in Proposition 4, and the contact set CS is defined by

$$CS := \left\{(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \Lambda, \boldsymbol{\kappa} \in \mathbf{K}, \boldsymbol{\kappa} \succeq_{v\text{-}MESSD} \boldsymbol{\lambda}, z \in \mathcal{X}, D^{\star}\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{\varepsilon}\right)\right) = 0\right\}.$$

2. *Under $\boldsymbol{H}_0 : K \succeq_{v\text{-}MESSD} \boldsymbol{\Lambda}$, and if $\exists\, (z^{\star}, \boldsymbol{\kappa}^{\star}, \boldsymbol{\lambda}^{\star}) \in CS:\ Var\left(\mathcal{L}_v\left(z^{\star}, \boldsymbol{\kappa}^{\star}, \boldsymbol{\lambda}^{\star}\right)\right) > 0$ then the testing procedure is asymptotically exact if $\alpha < 0.5$.*

3. *Under $\boldsymbol{H}_1 : K \not\succeq_{v\text{-}MESSD} \boldsymbol{\Lambda}$, the testing procedure is consistent.*

Theorem 1 is the direct analogue of Theorem 7 of AAT21, given that the testng procedure is applied on the DVTs of the locally non-stationary base return process, instead of the returns per se. The existence of a non trivial-in terms of asymptotic variance-condition in part 2 ensures exactness under every empirically plausible choice of significance level. The existence is ensured when there exists a pair $\boldsymbol{\kappa}^{\star}, \boldsymbol{\lambda}^{\star}$ such that the $\boldsymbol{\kappa}^{\star}$ DVT based portfolio may be chosen over $\boldsymbol{\lambda}^{\star}$ by any Russell-Seo (Russell and Seo, 1989) elementary utility, except for at least one utility that choses both, and this corresponds to a threshold level at the interior of intersection of the supports of the two portfolios. When the existence of a non trivial contact condition does not hold, arguments similar to Assumption 4.1.4 of Arvanitis *et al.* (2020) can be employed to ensure asymptotic conservatism.

2.3.3 Empirical results: in sample spanning tests

We test whether the traditional portfolio universe v-MESSD spans the augmented, for each investor persona, under the framework of mild explosivity developed previously. $\boldsymbol{\Lambda}$ denotes the portfolio universe that includes the cryptocurrencies as vertices, and $\boldsymbol{K}$ the one that only contains the traditional assets as vertices.

We employ the Linear Programming (LP) formulations as in Arvanitis et. al (2019), on the distorted dataset depending on the various investor types. We run all applications in Python 3.7

(Jupyter Notebook and PyCharm enviroments) with the Gurobi solver on a standard laptop with an Intel 8th Gen i7 processor and 16GB of RAM. The subsampling distribution of the test statistic is derived for subsample size $b_T \in [T^{0.6}, T^{0.7}, T^{0.8}, T^{0.9}]$.

Table 2 reports the test statistics n_T^*, the associated critical values q_T^{BC} together with the relevant decisions, for the three investor personas. To mitigate specification error of the subsample length and correct for bias in finite samples, we employ the bias-correction method of Arvanitis et. al. (2019), where the regression estimates q_T^{BC} are given for significance level $a = 0.05$.

The null hypothesis is rejected for the CPT and Loss Averse investor types. For the Markowitz type, the null hypothesis cannot be rejected. Thus, we can argue that for the chosen significance level, the performance of traditional portfolios can be improved with the inclusion of the four crypto assets for some CPT and Loss Averse investors; this does not seem to hold for the Markowitz type.

Table 2: In-sample performance with v-MESSD spanning test

	n_T^*	q_T^{BC}	Decision
CPT	0.00673	0.00173	Reject
Markowitz	0.00007	0.00017	Spanning
Loss Aversion	0.00802	0.00777	Reject

Entries report test statistics and critical values for stochastic spanning test of the augmented portfolios with respect to the traditional portfolio. The dataset spans the period from mid-of August 2015 to end-of August 2021, for a total of 1527 business days returns.

Irrespective of whether the spanning hypothesis is accepted or not, we continue our analysis by collecting the optimal $\boldsymbol{\lambda}$ and $\boldsymbol{\kappa}$ that result from the evaluation of the empirical spanning statistic in a rolling window framework as exemplified below.

2.4 Out of sample analysis

For each investor type, we are exploring the out-of-sample performance of her augmented optimal portfolio in comparison to the traditional one. We apply DVT, on the raw returns on each business day and hence we obtain a "behavioral" new dataset. We construct optimal portfolios based on the behavioral information up to time t, by evaluating the spanning statistic, and we then reap their actual returns, given the actual data, at time t+1. We record optimal portfolios separately for the Traditional and the Augmented opportunity sets, where the relevant optimisation problem is solved for the stochastic spanning test. The clock is set forward and we collect the realized returns of the optimal portfolios. This procedure is repeated for all subsequent business day returns till the end-of August 2021. In all cases we follow the reasoning that investors separate gains and losses as well as re-balance daily, after the first training year. We thus focus on the decisions investors would take about the optimal choice of portfolio weights, decisions that stem from the various behavioral elements that are being employed. We do not apply the DVTs to the returns of the optimal portfolios; we evaluate their performance using their realized actual returns.

The realized returns time series for all cases are presented in Figures 1 to 3. Despite the differences in aggregate returns, the traditional under-performs the augmented for the CPT and LA cases, while for the Markowitz case their performance is more or less the same, on average, except from a breakout in the performance of the Augmented in the final year. For the CPT and LA, the augmented portfolio performs at the end about two times higher than the traditional, while for Markowitz the spread between the augmented and the traditional is about 0.15%. Another interesting feature is that for all investor types, the augmented portfolios include relatively small weights for the Cryptocurrencies (see Table 6 in the Appendix).

Figure 1: CPT type of investor

Figure 2: Markowitz type of investor

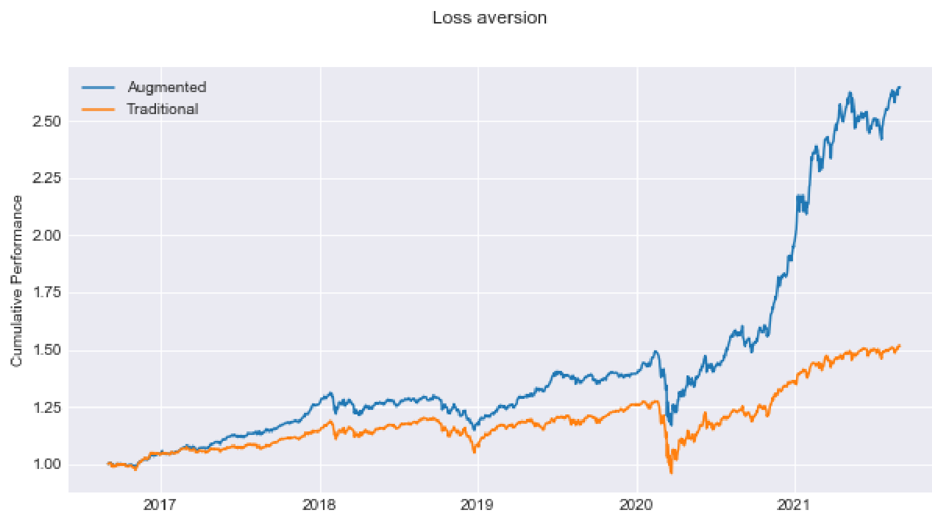


Figure 3: Loss averse type of investor

For all three personae, anticipated gains and losses play a tremendous role in the formation of portfolios, out-of-sample. Interestingly, we observe in Table 6 that all investor types prefer to invest lightly in cryptocurrencies, even though their average returns are higher compared to all other asset classes (Table 4). Thus, we could say that the main drivers of portfolio choice are not the high returns but low dispersion (S.D.), which is the highest in the case of cryptocurrencies. Hence, large standard deviations indicate abrupt ups-and-downs, or simply abrupt gains and losses. Although all three types exhibit different behaviors in gains and losses, they all depart from averting risk. That is why we observe high fund concentration on the safer asset classes. These asset classes demonstrate low S.D. and relative high mean. In our case, these are: the S&P 500 Index, the Barclays Bond Index and the 1M T-Bill. They attract an average capital of 71% in the traditional portfolios for the three personae. Subsequently, when cryptocurrencies come on the foreground this percentage climbs to an average of 76%. This 5% difference is a strong indication of skepticism towards the crypto-market, is the sense that they do want to benefit from the inclusion of cryptocurrencies but at the same time they are aware of the risk inherent in these exotic products. More specifically, the CPT and LA investors place a 6.6% and 7.8%, in total, on the cryptocurrencies group, while Markowitz investors stay at 1.6%.

The LA investor select a different approach and reduce significantly her funding on the two Vanguard indices (more than CPT and Markowitz combined), and remains the most exposed to cryptocurrencies. A possible explanation for that may be the fact that the two Vanguard indices exhibit negative skewness and high kurtosis, combined with high S.D., indicating extreme returns in the data. It is apparent that the LA is focused in avoiding extreme situations, especially when it comes to losses and data extremity can be reasonably perceived as signal for potential losses. In an analogous manner, the Russel 2000 index attracts very low weight percentages for all investor types.

CPT and Markowitz investors invest heavily on the 1M T-Bill, while LA prefers the S&P 500 Index. Note that, when cryptocurrencies are introduced, CPT and LA accumulate more funds on the 1M T-Bill while Makrowitz acts in the opposite direction. It seems that they are interested in taking advantage of the turbulence of the crypto market but at the same time reduce their exposure in the anticipation of certain losses. The certainty of losses stems form the fact that the specific market have suffered numerous times in the recent crisis. Regarding skewness and kurtosis, only Ripple (XRP) differs significantly and subsequently the relevant weights are the lowest. On the other hand, the Sharpe-ratios of the cryptocurrencies are the highest but we need to consider its drawback of the assumption of normal distribution.

Finally, the SMB and HML indices do not attract any type of investors because of the very low returns (HML exhibits a negative average return) and negative Sharpe ratios.

2.4.1 A conservative test for pairwise (non-) dominance

We perform a non-parametric out-of-sample performance comparative assessment of the optimal portfolios, using the Davidson and Duclos (2013) pairwise (non-) dominance test, as modified by AAT21 to allow for the mildly explosive framework. The modified test retains as null the less logically strict hypothesis of non-dominance, yet in a composite form concerning the comparison between every possible cluster point of the empirically optimal portfolio weights sequences.

The procedure allows for processes that appear in the context of Assumptions ME-AN, and for stochastic portfolio weights that may not be consistent or convergent at all. This is generally expected to be the case for the empirically optimal portfolio emerging from the stochastic spanning criterion; it need not have unique optimisers.

The test is applied exactly in its AAT21 form, on the returns of the empirically optimal portfolios, and not on their DVTs. Thereby, for brevity, we do not present the details on the hypothesis structure, the form of the test statistic and the rejection region, and its limiting properties. Those are discussed in AAT21 and the interested reader is referred there. We however point out the following:

The null hypothesis is that for any pair of cluster points between the associated portfolio, the cluster point of the first, does not MESSD the second; MESSD is now the dominance relation defined before for v equal to the identity. The alternative posits the existence of a pair for which the first dominates the second.

The test statistic is essentially a supremum of easily computable t statistics over $\mathcal{X}$, and the rejection region is based on the standard normal distribution. Thereby, the procedure is independent of the choice of numerical approximation parameters (like the subsampling length) as resampling based approximations of the limiting rejection region are avoided.

Under assumptions involving properties of the associated cluster points and the LPMs around them, and if a consistent estimator for long run covariances is used, the test is shown to be asymptotically conservative and consistent.

2.4.2 Out of sample empirical results: pairwise (non-) dominance

We apply our modification of the Davidson and Duclos (2013) pairwise (non-) dominance test on the two optimal portfolios derived in the previous rolling analysis.

Table 3 reports the quartile p-values from the distribution of daily portfolio returns, for the null hypothesis that the augmented portfolio does not stochastically dominate the traditional one by second order (see Davidson and Duclos, 2013). The results entail $T-1$ (1157) overlapping periods for the in-sample fitting of the two portfolios with corresponding out-of-sample comparisons. The $T-1$ p-values are considered from September 1, 2016 to August 31, 2021, using overlapping periods of 100 daily returns in all three cases. The quartile p-values from the distribution of the $T-1$ modified t-test statistics are computed.

We observe that, for the 25% and 50% (in 5%) quartile p-values, the null hypothesis that the augmented optimal portfolio does not stochastically dominate the traditional one by second order is rejected in all cases. Hence, the out-of-sample performance of optimal portfolios constructed by every investor persona, that include cryptocurrencies seems dominant to the performance of the corresponding benchmark.

Table 3: Out-of-sample performance: Non-parametric stochastic dominance test

	Traditional vs Augmented
CPT	
Quartile	
25% Rejection rate	47.13%
50% Rejection rate	65.31%
75% Rejection rate	84.19%
Markowitz	
Quartile	
25% Rejection rate	44.27%
50% Rejection rate	63.42%
75% Rejection rate	81.94%
Loss aversion	
Quartile	
25% Rejection rate	38.28%
50% Rejection rate	57.84%
75% Rejection rate	76.87%

Entries report quartile rejection rates from the distribution rejection rates across out-of-sample periods under the null hypothesis that the augmented cryptocurrencies optimal portfolio does not second order stochastically dominate the optimal traditional portfolio using a modification of the Davidson and Duclos (2013) test statistic, over the period from September 1, 2016 to August 31, 2021.

2.4.3 Further out of sample empirical results: parametric tests

We also apply a set of commonly used parametric performance measures: the Sharpe ratio, the downside Sharpe ratio (DS) (Ziemba, 2005), the upside potential (UP) and downside risk ratio

(Sortino and van den Meer, 1991), the opportunity cost (Simaan, 1993), the portfolio turnover (P.T.) and a measure of the portfolio risk-adjusted returns net of transaction costs (RL). Since assets' returns exhibit asymmetric return distributions, the downside Sharpe and UP ratios are more appropriate measures than the typical Sharpe ratio. For the compatibility of those with Assumptions ME-AN see AAT21.

For the DS ratio, we first evaluate the downside risk (downside variance) which is given by the formula:

$$\sigma^2_{P_-} = \frac{\sum_{t=1}^{T}(\min\left(x_t, 0\right))^2}{T-1}$$

where, x_t are those returns of portfolio P at day t below 0, i.e. those days with losses. To get the total variance we use: $2\sigma^2_{P_-}$, thus the DS ratio is,

$$S_P = \frac{\overline{R}_P - \overline{R}_f}{\sqrt{2\sigma^2_{P_-}}}$$

where, $\overline{R}_P$ is the average period return of portfolio P and $\overline{R}_f$ is the average risk free rate.

The DS ratio removes any effects of upward price movement on the standard deviation in order to focus on the distribution of the returns that are below a predefined threshold/target that is set by an investor (or fund) as a minimum required return. Its deviation from the Sharpe ratio lies in that it replaces the risk-free rate with the required return. In our experiments we assume that this required return is the average risk-free return of the whole period under examination.

The UP ratio compares the upside potential to the shortfall risk over a benchmark and is computed as follows. For R_t the realized daily return of portfolio P for $t = 1, ..., T$ of the backtesting period, where T is the number of experiments performed, and p_t the return of the benchmark (risk free rate), which in our case is the one month T-bill riskless asset for the same period:

$$\text{UP ratio} = \frac{\frac{1}{T_1}\sum_{t=1}^{T_1}\max\left(R_t - p_t, 0\right)}{\sqrt{\frac{1}{T_2}\sum_{t=1}^{T_2}(\max\left(p_t - R_t, 0\right))^2}} \quad , T = T_1 + T_2$$

The numerator of the above ratio is the average excess return over the benchmark and thus it reflects the upside potential. In the same sense, the denominator measures downside risk, i.e. shortfall risk over the benchmark.

Next, we evaluate the P.T. to get a feeling of the degree of rebalancing required to implement each one of the investment strategies under examination. For any portfolio strategy P, the portfolio turnover is defined as the average of the absolute change of weights over the T rebalancing points in time and across the d available assets, i.e.

$$\text{P.T.} = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{M}\left(\mid w_{Pi,t+1} - w_{Pi,t} \mid\right)$$

where $w_{Pi,t+1}, w_{Pi,t}$ are the optimal weights of asset i under strategy P (Traditional or Augmented) at time t and t+1, respectively.

We also evaluate the performance of the portfolios under the risk-adjusted returns measure, which is net of transaction costs, proposed by DeMiguel et al. (2009). It indicates the way that the proportional transaction cost, generated by the P.T., affects the portfolio returns. Let TrC be the proportional transaction cost, and $R_{P,t+1}$ the realized return of portfolio P at time t+1. The change in the net of transaction cost wealth NW_P of portfolio P through time is,

$$NW_{P,t+1} = NW_{P,t}\big(1 + R_{P,t+1}\big)\Big(1 - \text{TrC} \times \sum_{i=1}^{M}\left(\mid w_{P,i,t+1} - w_{P,i,t} \mid\right)\Big)$$

The portfolio return, net of transaction cost, is defined as,

$$RTC_{P,t+1} = \frac{NW_{P,t+1}}{NW_{P,t}} - 1$$

Let μ_{Tr}, μ_{Aug} be the out-of-sample mean of monthly RTC with the traditional and augmented opportunity set, respectively, and $\sigma_{Tr}, \sigma_{Aug}$ be the corresponding standard deviations. Then, the return-loss measure is,

$$R_{Loss} = \frac{\mu_{Aug}}{\sigma_{Aug}} \times \sigma_{Tr} - \mu_{Tr}$$

It evaluates the additional return needed so that the Traditional performs equally well with the Augmented. We follow the literature and use 35 basis points (bps), i.e. 0.35% , for the proportional transaction cost of stocks and bonds.

Finally, we use the concept of opportunity cost presented in Simaan (1993) to analyze the economic significance of the performance difference of the two optimal portfolios, in both experiments and for all three perspectives (i.e. investor types). Let R^i_{Tr} and R^i_{Aug} be the realized returns of the optimal Traditional and Augmented portfolio for every investor i. Then, the opportunity cost θ is defined as the return that needs to be added to (or subtracted from) R^i_{Tr}, so that the investor is indifferent (in utility terms) between the strategies imposed by the two different investment opportunity classes:

$$\mathbb{E}\Big[U\big(1 + R^i_{Tr} + \theta\big)\Big] = \mathbb{E}\Big[U\big(1 + R^i_{Aug}\big)\Big] \quad ; i = \text{CPT, M, LA},$$

where now $\mathbb{E}$ now denotes empirical expectation. A positive opportunity cost implies that an investor is better off if she includes additional assets in her portfolio, while a negative one implies that she would be worse off with the aforementioned inclusion. It is important to mention that the opportunity cost takes into account the entire probability distribution of portfolio returns and hence it is suitable to evaluate strategies even when the distribution is not normal. For the calculation of the opportunity cost we follow the literature and use the relevant *S-shaped* for CPT, *inverse S-shaped* for Markowitz (M), as well as the linear value function for the LA type.

Table 5 reports the parametric performance measures for the Traditional and the Augmented portfolios, for all investor types. All cases enrich the evidence obtained from the non-parametric SD measures. The higher the value of each one of these measures, the greater the investment opportunity for including cryptocurrencies. We follow the literature and use 35 bps for the transaction costs of stocks and bonds.

The results show that the inclusion of cryptocurrencies into the opportunity set increases all performance measures. We observe that since in all cases the UP ratio increases, all investors exhibit a benefit from the inclusion of cryptocurrencies. Furthermore, we observe that portfolios with only traditional assets induce less portfolio turnover than the ones with cryptocurrencies, which analogously creates more portfolio turnover for all types. Additionally, we can see that the return-loss measure is similarly positive for all types. Thus, the Traditional portfolio has to increase its return in order to perform equally "well" with the Augmented. Finally, in all cases we find positive opportunity costs θ for CPT, LA and Markowitz investors. Hence, one needs to give a positive return equal to θ to a CPT, LA and Markowitz investor who optimizes in the augmented opportunity set, so that she becomes willing not to include cryptocurrencies.

3 Discussion and future research

We examined how different investor personae "behave" when forming optimal portfolios of different asset classes. We were primarily interested in the performance of the three main decision-making personae in Behavioral Finance namely, CPT , Markowitz and LA type.

We worked with two classes of portfolio spaces: a traditional one, and one augmented with cryptocurrencies. We used a stochastic spanning methodology to test whether cryptocurrencies offer diversification benefits to some risk averse investors after returns are adjusted for the cognitive biases for each of the BF persona involved. We allowed for an empirically plausible framework of multiple - possibly interdependent - bubbles in cryptocurrencies. We conducted our analysis both in- and out-of-sample by constructing and comparing optimal portfolios derived from the two respective asset universes.

Even though the in-sample results suggest that for the Markowitz persona no investment opportunities exist in augmented portfolio space, compared to the traditional, the out of sample analysis suggest that optimal augmented portfolios stochastically dominate the respective traditional one for every persona, and generally outperform them w.r.t. every parametric criterion that we employed.

The asymptotic stationarity framework implied by our assumption on bubble sparsity is not generally empirically identifiable, and its validity is essentially based on backward looking historical arguments. If this does not hold and more generally the (essentially forward looking) condition is not valid, then our statistical inference becomes ambiguous. It may be thus of interest to extend the testing methodologies presented above in asymptotically persistent non-stationary frameworks for the returns, like the weakly non-stationary processes of Duffy and Kasparis (2021). This non trivial task is left for future research.

References

1. Abdellaoui, M., Bleichrodt, H., and Paraschiv, C., 2007, Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. Management Science 53(10):1659-1674.

2. Anyfantaki, S., Arvanitis, S. and Topaloglou, N., 2021, Diversification benefits in the cryptocurrency market under mild explosivity. European Journal of Operational Research, 295(1), pp.378-393.

3. Arvanitis, S. , Hallam, M. , Post, T. , and Topaloglou, N., 2019, Stochastic spanning. Journal of Business & Economic Statistics, 37 (4), 573–585.

4. Arvanitis, S., Scaillet, O., and N. Topaloglou, 2020. Spanning Tests for Markowitz Stochastic Dominance. Journal of Econometrics, 217(2), 291-311.

5. Arvanitis, S., and N. Topaloglou, 2017. Testing for Prospect and Markowitz Stochastic Dominance Efficiency. Journal of Econometrics 198(2), 253-270.

6. Bachelier, L., 1900, Theorie de la speculation. Annales scientifiques de l'école normale supérieure, 17, 21-86.

7. Barberis N, Huang M. and T. Santos, 2002, Prospect Theory and Asset Prices. Quarterly Journal of Economics 116, 1-53.

8. Barberis N, Mukherjee A, Wang B, Prospect Theory and Stock Returns: An Empirical Test, The Review of Financial Studies, Volume 29, Issue 11, November 2016, Pages 3068–3107

9. Barberis N. and Thaler R., 2003, A survrey on behavioral Finance. Handbook of the Economics of Finance, Volume 1, Part B, Pages i-xxv, 605-1246, I-1-I-25, Financial Markets and Asset Pricing.

10. Basrak, B. , Davis, R. A. , & Mikosch, T., 2002, Regular variation of GARCH processes. Stochastic Processes and Their Applications, 99 (1), 95–115.

11. Baucells M., Heukamp F., 2006, Stochastic Dominance and Cumulative Prospect Theory. Management Science, Vol. 52, No. 9, 1409-1423.

12. Benartzi S. and R.Thaler 1995, Myopic Loss Aversion and the Equity Premium Puzzle. Quarterly Journal of Economics 110, 73-92.

13. Camerer C., 1998, Prospect theory in the wild: Evidence from the field. Social Science Working Paper 1037.

14. Davidson, R. and J.-Y. Duclos, 2013, ”Testing for restricted stochastic dominance”, Econometric Reviews 32, 84-125.

15. DeMiguel, V., L. Garlappi and R. Uppal, 2009, Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?, Review of Financial Studies 22, 1915-1953.

16. Drost, F. C. , Nijman, T. E., 1993, Temporal aggregation of GARCH processes. Econometrica, 909–927.

17. Duffy, J.A. and Kasparis, I., 2021. Estimation and inference in the presence of fractional d= 1/2 and weakly nonstationary processes. The Annals of Statistics, 49(2), pp.1195-1217.

18. Friedman M. and Savage L, 1948, The utility analysis of choices involving risk. Journal of Political Economy 56, 279-304.

19. Gonzalez, Richard, and George Wu. 1999. “On the Shape of the Probability Weighting Function.” Cognitive Psychology 38(1): 129-166.

20. Hadar J. and W.R Russel, 1969, Rules for ordering uncertain prospects. AER 59, 234.

21. Hanoch, Giora and Levy, H., 1969, The Efficiency Analysis of Choices Involving Risk, Review of Economic Studies, 36, issue 3, p. 335-346.

22. Huberman G. and S. Kandel, 1987, Mean - variance spanning. Journal of Finance 42, 873-888.

23. Ingersoll J, 2008, Non-monotonicity of the Tversky-Kahneman Probability Weighting Function: A Cautionary Note. European Financial Management, Vol. 00, No. 0, 1-6.

24. Jin, S., Corradi, V. and Swanson, N.R., 2017, Robust forecast comparison. Econometric Theory, 33(6), pp.1306-1351.

25. Kahneman, D , and A. Tversky, 1979, Prospect theory: an analysis of decision under risk. Econometrica, 47:263-291.

26. Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." Journal of Economic Perspectives, 5 (1): 193-206.

27. Kőszegi, B., and Rabin, M., 2006, A Model of Reference-Dependent Preferences. The Quarterly Journal of Economics, 121(4), 1133–1165.

28. Levy H., 2006, Stochastic Dominance: Investment decision making under uncertainty, 3rd edition, Springer.

29. Markowitz H., 1952a, The Utility of Wealth. Journal of Political Economy, Vol. 60, No. 2 (Apr., 1952), pp. 151-158.

30. Markowitz H., 1952b, Portfolio Selection. The Journal of Finance, Vol. 7, No. 1 (Mar., 1952), pp. 77-91.

31. Odean, T., 1999, Do Investors Trade Too Much?. American Economic Review, 89 (5): 1279-1298.

32. Phillips, P. C. , & Magdalinos, T., 2007, Limit theory for moderate deviations from a unit root. Journal of Econometrics, 136 (1), 115–130.

33. Phillips, P. C. , Shi, S., 2020, Real time monitoring of asset markets: Bubbles and crises. In Handbook of statistics: vol. 42 (pp. 61–80). Elsevier.

34. Phillips, P. C. , Shi, S. , & Yu, J., 2015, Testing for multiple bubbles: Historical episodes of exuberance and collapse in the sp 500. International Economic Review, 56 (4), 1043–1078.

35. Rothchild M. and J.E. Stiglitz, 1970, "Increasing Risk: A definition", Journal of Economic Theory, 2(3), 225-243

36. Russel W.R. and Seo T.K., 1989, Representative sets for stochastic dominance rules in: T.B> Fomby and T.K. Seo (eds), Studies in the Economics of uncertainty: in honor of Josef Hadar, NY: Springer Verlag, 59-76.

    Perspectives 17, 83-104.

37. Simaan Y, 1993, What is the Opportunity Cost of Mean-Variance Investment Strategies?, Management Science, 39, (5), 578-587

38. Sortino, F. and R. van der Meer, 1991, Downside risk - capturing what is at stake in investment situations. Journal of Portfolio Management 17, 27-31.

39. Tversky, A., and D. Kahneman, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty 5, 297-323.

40. Von Neumann, J., and O. Morgenstern, 1953, Theory of Games and Economic Behavior.

41. Vershynin, R., 2018, High-dimensional probability: An introduction with applications in data science (Vol. 47). Cambridge University Press .

42. Ziemba, W., 2005, The symmetric downside risk Sharpe ratio. Journal of Portfolio Management 32, 108-122.

Appendix 1: Proofs and auxiliary results

The appendix contains the proofs of the main results, as well as the derivation of several auxiliary results used in the proofs.

Proof of Proposition 1. In what follows H. in. abbreviates the Holder inequality, CS in. the Cauchy-Schwarz inequality, tr. in. the triangle inequality and tr. in.* its dual form, D-R Mink. in. the dual form of the reverse Minkowski inequality, J. in. Jensen's inequality, and norms in. the bounding from above of the max-norm by a constant multiple of the Euclidean norm in $\mathbb{R}^d$. $\odot$ denotes the Hadamard product. c denotes a generic positive constant that may change its value at different occurrences. Notice first that

$$\begin{array}{c} \left\|\left(v\left(\exp^{\star}\left(\boldsymbol{X}_t\right)-\mathbf{1}\right)-v\left(\exp^{\star}\left(\varepsilon_t\right)-\mathbf{1}\right)\right)\right\| \\ \underset{\text{norms in.}}{\leq} c\max_i\left|c_i\right|(p)\sum_{i=1}^d\left|\begin{array}{c}\left|\exp\left(\boldsymbol{X}_{t(i)}\right)-\mathbf{1}\right|^{\alpha_i}\mathbb{I}\left(\left(\exp\left(\boldsymbol{X}_{t(i)}\right)-\mathbf{1}\right)\geq 0\right)\\ -\left|\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right|^{\alpha_i}\mathbb{I}\left(\left(\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right)\geq 0\right)\end{array}\right| \\ +c\max_i\left|c_i^{\star}\right|\sum_{i=1}^d\left|\begin{array}{c}\left|\exp\left(\boldsymbol{X}_{t(i)}\right)-\mathbf{1}\right|^{\beta_i}\mathbb{I}\left(\left(\exp\left(\boldsymbol{X}_{t(i)}\right)-\mathbf{1}\right)<0\right)\\ -\left|\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right|^{\beta_i}\mathbb{I}\left(\left(\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right)<0\right)\end{array}\right|. \end{array}$$

Then, for $\delta_i=\max\alpha_i,\beta_i$, the rhs of the previous display is then less than or equal to

$$\begin{array}{c} c\sum_{i=1}^d\left|\left|\exp\left(\boldsymbol{X}_{t(i)}\right)-\mathbf{1}\right|^{\alpha_i}-\left|\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right|^{\alpha_i}\right| \\ +c\sum_{i=1}^d\left|\left|\exp\left(\boldsymbol{X}_{t(i)}\right)-\mathbf{1}\right|^{\beta_i}-\left|\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right|^{\beta_i}\right| \\ +c\sum_{i=1}^d\left|\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right|^{\alpha_i}\left|\mathbb{I}\left(\left(\exp\left(\boldsymbol{X}_{t(i)}\right)-\mathbf{1}\right)\geq 0\right)-\mathbb{I}\left(\left(\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right)\geq 0\right)\right| \\ +c\sum_{i=1}^d\left|\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right|^{\beta_i}\left|\mathbb{I}\left(\left(\exp\left(\boldsymbol{X}_{t(i)}\right)-\mathbf{1}\right)<0\right)-\mathbb{I}\left(\left(\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right)<0\right)\right| \\ \underset{\text{D-R Mink. in. or tr.* in.}}{\leq} c\sum_{i=1}^d\left|\exp\left(\boldsymbol{X}_{t(i)}\right)-\exp\left(\varepsilon_{t(i)}\right)\right|^{\delta_i}+c\sum_{i=1}^d\left|\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right|^{\delta_i}, \end{array}$$

where the final inequality in the last display follows from the application of dual form if the Minkowski reverse inequality, or the dual form of the triangle inequality, depending on whether the relevant exponent is less or greater than or equal to one, respectively. Then, due to the Lipschitz continuity property of $x\to(x)_+$, and the above display, for $\delta:=\max_i\delta i$,

$$\begin{array}{c} \mathbb{E}\left[\left|\frac{1}{\sqrt{T}}\sum_{t\in\cup_{k=1}^K B_k}\left(z-\boldsymbol{\lambda}^{\mathrm{T}}v\left(\exp^{\star}\left(\boldsymbol{X}_t\right)-\mathbf{1}\right)\right)_+-\frac{1}{\sqrt{T}}\sum_{t\in\cup_{k=1}^K B_k}\left(z-\boldsymbol{\lambda}^{\mathrm{T}}v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_t\right)-\mathbf{1}\right)\right)_+\right|\right] \\ \leq\frac{1}{\sqrt{T}}\sum_{t\in\cup_{k=1}^K B_k}\mathbb{E}\left[\left|\boldsymbol{\lambda}^{\mathrm{T}}\left(v\left(\exp^{\star}\left(\boldsymbol{X}_t\right)-\mathbf{1}\right)-v\left(\exp^{\star}\left(\varepsilon_t\right)\right)-\mathbf{1}\right)\right|\right] \\ \underset{\text{CS in.}}{\leq}\frac{1}{\sqrt{T}}\sum_{t\in\cup_{k=1}^K B_k}\mathbb{E}\left[\left\|\left(v\left(\exp^{\star}\left(\boldsymbol{X}_t\right)-\mathbf{1}\right)-v\left(\exp^{\star}\left(\varepsilon_t\right)\right)-\mathbf{1}\right)\right\|\right] \\ \frac{c}{\sqrt{T}}\sum_{t\in\cup_{k=1}^K B_k}\sum_{i=1}^d\left[\mathbb{E}\left[\left|\exp\left(\boldsymbol{X}_{t(i)}\right)-\exp\left(\varepsilon_{t(i)}\right)\right|^{\delta_i}\right]+\mathbb{E}\left[\left|\exp\left(\varepsilon_{t(i)}\right)-\mathbf{1}\right|^{\delta_i}\right]\right] \end{array} \tag{1}$$

$$
\begin{aligned}
&\underset{\text{norms+tr.+J in.}}{\leq} \tfrac{c}{\sqrt{T}} \textstyle\sum_{t\in\cup_{k=1}^{K}B_k} \mathbb{E}\left[\left\|\exp^{\star}\left(\boldsymbol{X}_t\right)-\exp^{\star}\left(\varepsilon_t\right)\right\|^{\delta}\right] + \tfrac{c}{\sqrt{T}} \textstyle\sum_{t\in\cup_{k=1}^{K}B_k} \left(\sum_{i=1}^{d}[\mathbb{E}\left[\exp\left(\delta\varepsilon_{0(i)}\right)\right]] + d\right) \\
&\underset{\text{CS in.}}{\leq} \tfrac{1}{\sqrt{T}} \textstyle\sum_{t\in\cup_{k=1}^{K}B_k} \mathbb{E}\left[\left\|\left(\exp^{\star}\left(\delta \sum_{k=1}^{K} \frac{C_k}{M(T,k)} \mathbb{I}\left\{t\in B_k\right\} \boldsymbol{X}_{t-1}\right)-\mathbf{1}\right) \odot \exp^{\delta\star}\left(\varepsilon_t\right)\right\|\right] \\
&\qquad + \tfrac{c}{\sqrt{T}} \textstyle\sum_{t\in\cup_{k=1}^{K}B_k} \left(\sum_{i=1}^{d} \mathbb{E}\left[\exp\left(\delta\varepsilon_{0(i)}\right)\right] + d\right).
\end{aligned} \tag{2}
$$

Notice that due to Assumptions ME-AN,

$$
\begin{aligned}
\tfrac{c}{\sqrt{T}} \textstyle\sum_{t\in\cup_{k=1}^{K}B_k} \left(\sum_{i=1}^{d} \mathbb{E}\left[\exp\left(\delta\varepsilon_{0(i)}\right)\right] + d\right) &\leq \tfrac{cd}{\sqrt{T}} \textstyle\sum_{t\in\cup_{k=1}^{K}B_k} \left(\max_i \mathbb{E}\left[\exp\left(\delta\varepsilon_{0(i)}\right)\right] + 1\right) \\
&\leq c\frac{\max_{t,k} B_k}{\sqrt{T}} = o\left(1\right).
\end{aligned}
$$

The asymptotic negligibility of the first and remaining term in the final bound of the previous display follows from the proof of Proposition 3 of AAT21. □

Proof of Theorem 1. (1) and (3) follow exactly as in the proofs of Propositions 4 and B.2 respectively in Arvanitis et al. (2019) (see their Online Appendix) given Proposition 4. For (2) notice that if $\mathrm{Var}\left(\mathcal{L}_v\left(z^{\star},\boldsymbol{\kappa}^{\star},\boldsymbol{\lambda}^{\star}\right)\right) > 0$ and whenever $\mathcal{L}_v\left(z^{\star},\boldsymbol{\kappa}^{\star},\boldsymbol{\lambda}^{\star}\right) > 0$, then $\sup_{\Lambda}\inf_{K}\sup_{\mathcal{X}}\mathcal{L}_v\left(z,\boldsymbol{\kappa},\boldsymbol{\lambda}\right) \geq \mathcal{L}_v\left(z^{\star},\boldsymbol{\kappa}^{\star},\boldsymbol{\lambda}^{\star}\right)$. Due to zero mean Gaussianity this occurs with probability at least 0.5. The rest follows as in the proof of Proposition B.2 in Arvanitis et al. (2019) (see their Online Appendix). □

Auxiliary Results

Proposition 2. *Suppose that Assumptions ME and AN hold.*
Then $D^{\star}\left(z,\boldsymbol{\kappa},\boldsymbol{\lambda},v\left(\boldsymbol{R}\right)\right) = D^{\star}\left(z,\boldsymbol{\kappa},\boldsymbol{\lambda},v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_0\right)-\mathbf{1}\right)\right),\ \forall\left(z,\boldsymbol{\kappa},\boldsymbol{\lambda}\right)$.

Proof. Stationarity for the $\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_t\right)-\mathbf{1}\right)_t$ process implies stattionarity for $\left(v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_t\right)-\mathbf{1}\right)\right)_t$, and by Proposition 1: for any $z,\boldsymbol{\lambda}$,

$$
\begin{aligned}
&\left|\tfrac{1}{T}\textstyle\sum_{t=1}^{T}\mathbb{E}\left[\left(z-\boldsymbol{\lambda}^{\mathrm{T}}v^{\star}\left(\boldsymbol{R}_t\right)\right)_+\right] - \mathbb{E}\left[\left(z-\boldsymbol{\lambda}^{\mathrm{T}}v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_0\right)-\mathbf{1}\right)\right)_+\right]\right| \\
\leq\ &\mathbb{E}\left[\left|\tfrac{1}{T}\textstyle\sum_{t\in\cup_{k=1}^{K}B_k}\left(z-\boldsymbol{\lambda}^{\mathrm{T}}v\left(\exp^{\star}\left(\boldsymbol{X}_t\right)-\mathbf{1}\right)\right)_+ - \tfrac{1}{T}\sum_{t\in\cup_{k=1}^{K}B_k}\left(z-\boldsymbol{\lambda}^{\mathrm{T}}v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_t\right)-\mathbf{1}\right)\right)_+\right|\right] = o\left(\tfrac{1}{\sqrt{T}}\right).
\end{aligned}
$$

□

Proposition 3. *Suppose that Assumptions ME and AN hold. Then as* $T\to\infty$,

$$
\sup_{\Lambda,\mathrm{K},\mathcal{X}} \sqrt{T}\left|D_T\left(z,\boldsymbol{\kappa},\boldsymbol{\lambda},v\left(\boldsymbol{R}\right)\right) - D^{\star}\left(z,\boldsymbol{\kappa},\boldsymbol{\lambda},v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_0\right)-\mathbf{1}\right)\right)\right| = o_p\left(1\right).
$$

Furthermore,

$$
\left|\sup_{\Lambda}\inf_{K}\sup_{\mathcal{X}}\sqrt{T}D_T\left(z,\boldsymbol{\kappa},\boldsymbol{\lambda},v\left(\boldsymbol{R}\right)\right) - \sup_{\Lambda}\inf_{K}\sup_{\mathcal{X}}\sqrt{T}D^{\star}\left(z,\boldsymbol{\kappa},\boldsymbol{\lambda},v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_0\right)-\mathbf{1}\right)\right)\right| = o\left(1\right).
$$

Proof. The first result follows from Proposition 1 and the fact that L_1 convergence implies convergence in probability. The second from that the processes involved have almost surely bounded paths, from the first result and the CMT.□ □

Proposition 4. *Suppose that Assumptions ME and AN hold. Then as $T \to \infty$,*

$$\begin{array}{c} \sqrt{T}\left(D_T\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\boldsymbol{R}\right)\right) - D^{\star}\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}, v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_0\right) - \mathbf{1}\right)\right)\right) \rightsquigarrow \mathcal{L}_v\left(z, \boldsymbol{\kappa}, \boldsymbol{\lambda}\right) \\ := \int_{\mathbb{R}} \left(\left(z - \boldsymbol{\kappa}^T \mathbf{x}\right)_+ - \left(z - \boldsymbol{\lambda}^T \mathbf{x}\right)_+\right) d\mathcal{G}_{\mathrm{F}}\left(\mathbf{x}\right) \end{array},$$

in the space of $\mathbb{R}$ bounded functions on $K \times \Lambda \times \mathcal{X}$ equipped with the sup norm, where $\mathcal{G}$ is a centered Gaussian process with covariance kernel given by

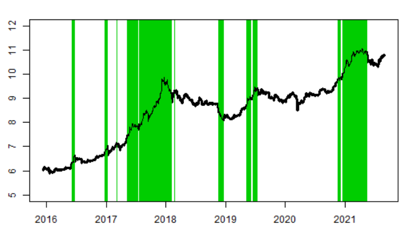
$$Cov\left(\mathcal{G}_{\mathrm{F}}\left(\mathbf{x}\right), \mathcal{G}_{\mathrm{F}}\left(\mathbf{y}\right)\right) = \sum_{t \in \mathbb{Z}} Cov\left(\mathbb{I}\left\{v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_0\right) - 1\right) \leq \mathbf{x}\right\}, \mathbb{I}\left\{v\left(\exp^{\star}\left(\boldsymbol{\varepsilon}_t\right) - 1\right) \leq \mathbf{y}\right\}\right)$$

and uniformly continuous sample paths on $\mathbb{R}^d$.

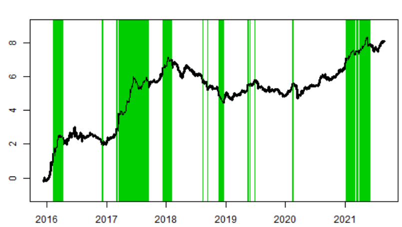
Proof. The result follows by the first part of Proposition 3, and Lemma A.1 in (the Technical Appendix of Arvanitis et. al., 2020). □

Appendix 2: Timestamping Cryptocurrencies' Bubbles

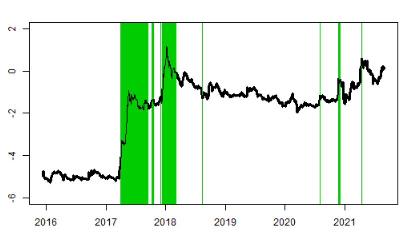
The shaded green areas in the following figure are the identified multiple bubble periods, obtained by using the the 95% bootstrap critical values for the cryptocurrencies involved.



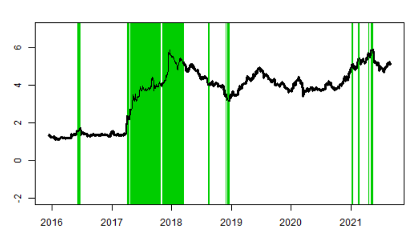
(a) Bitcoin



(b) Ethereum



(c) Ripple



(d) Litecoin

Appendix 3: Tables

Table 4: Descriptive statistics of business days returns

Asset	Mean	S.D.	Skewness	Kurtosis	Sharpe ratio
S&P 500 Index	0.000579	0.011754	-0.692944	19.285854	0.045633
Barclays bond Index	0.000139	0.002168	-0.934352	9.4163	0.044550
1M T-Bill	0.000043	0.000033	0.345543	-1.390497	-
Russel 2000	0.000526	0.014855	-0.966213	13.602642	0.032559
Vanguard Value	0.000418	0.01182	-0.622055	18.710689	0.031756
Vanguard Small-Cap	0.000523	0.014132	-0.217019	25.347711	0.033968
SMB	0.000032	0.006416	0.365433	5.124661	-0.001638
HML	-0.000204	0.009102	0.359444	6.115774	-0.027071
BTC	0.004476	0.047029	-0.016428	6.136154	0.094275
ETH	0.008007	0.080832	0.791753	13.348979	0.098535
XRP	0.00669	0.088847	3.683376	33.976011	0.074815
LTC	0.004677	0.069013	1.941846	18.103941	0.067149

Entries report the descriptive statistics on business daily returns from from mid-August 2015 to end-of-August 2021. The traditional set includes the: S&P 500 Index, the Barclays US Bond Index, the 1 Month T-Bill, the Russell 2000 Index, the Vanguard Value Index, the Vanguard Small-Cap Index, and the dynamic trading strategies SMB and HML. Bitcoin, Ethereum, Ripple and Litecoin US dollar closing prices are used to assess the cryptocurrency market. The average return, the standard deviation (S.D.), the skewness, the kurtosis, as well as the Sharpe ratio are reported.

Table 5: Out-of-sample performance: Parametric portfolio measures

Performance measures	CPT (T)	CPT (A)	Mark. (T)	Mark. (A)	LA (T)	LA (A)
Sharpe ratio	0.04169	0.10011	0.03595	0.08161	0.04153	0.09388
Downside Sharpe ratio	0.02722	0.06652	0.03145	0.05312	0.02702	0.06098
UP ratio	0.52837	0.60139	0.51309	0.57967	0.51930	0.57745
Portfolio Turnover	0.00926	0.01250	0.01212	0.02107	0.01819	0.02432
Return Loss	0.00050	-	0.00015	-	0.00031	-
Opportunity cost	0.00036	-	0.00005	-	0.00044	-

Entries report the performance measures (Sharpe ratio, Downside Sharpe ratio, UP ratio, Portfolio Turnover, Returns Loss and Opportunity Cost) for the traditional and the augmented optimal portfolios. The realised business daily returns cover the period from mid-August 2015 to end-of-August 2021. The traditional set includes the: S&P 500 Index, the Barclays US Bond Index, the 1 Month T-Bill, the Russell 2000 Index, the Vanguard Value Index, the Vanguard Small-Cap Index, and the dynamic trading strategies SMB and HML. The augmented portfolio includes additionally Bitcoin, Ethereum, Ripple and Litecoin. All values are rounded to the fifth decimal.

Table 6: Out-of-sample analysis: average portfolio composition

Asset	CPT (T)	CPT (A)	Mark. (T)	Mark. (A)	LA (T)	LA (A)
S&P 500 Index	0.167	0.149	0.168	0.189	0.267	0.293
Barclays bond Index	0.142	0.098	0.098	0.242	0.122	0.162
1M T-Bill	0.336	0.523	0.515	0.434	0.177	0.223
Russel 2000	0.058	0.013	0.020	0.015	0.040	0.021
Vanguard Value	0.096	0.079	0.079	0.059	0.138	0.125
Vanguard Small-Cap	0.125	0.116	0.088	0.041	0.155	0.114
SMB	0.041	0.014	0.019	0.012	0.058	0.040
HML	0.036	0.008	0.014	0.009	0.043	0.023
BTC		0.026		0.008		0.039
ETH		0.020		0.005		0.025
XRP		0.008		0.001		0.007
LTC		0.012		0.002		0.007

Entries report the average portfolio compositions for the full period from mid-August 2015 to end-of-August 2021. The traditional set includes the: S&P 500 Index, the Barclays US Bond Index, the 1 Month T-Bill, the Russell 2000 Index, the Vanguard Value Index, the Vanguard Small-Cap Index, and the dynamic trading strategies SMB and HML. The augmented portfolio includes additionally Bitcoin, Ethereum, Ripple and Litecoin.