Spanning analysis of stock market anomalies under Prospect Stochastic Dominance

Stelios Arvanitis[a], Olivier Scaillet[b], Nikolas Topaloglou[a,c] *

February 2023

Abstract

We develop and implement methods for determining whether introducing new securities or relaxing investment constraints improves the investment opportunity set for prospect investors. We formulate a new testing procedure for prospect spanning for two nested portfolio sets based on sub-sampling and Linear Programming. In an application, we use the prospect spanning framework to evaluate whether well-known anomalies are spanned by standard factors. We find that of the strategies considered, a few of them expand the opportunity set of the prospect type investors, thus have real economic value for them, and involve absence of loss aversion. Those are the Net Stock Issue anomaly under the FF-5 model, the Momentum and Net Stock Issue anomalies under the M-4 model, and the Momentum anomaly under the q model. In-sample and out-of-sample results prove remarkably consistent in identifying genuine anomalies for prospect investors.

Keywords and phrases: Nonparametric test, prospect stochastic dominance efficiency, prospect spanning, market anomaly, Linear Programming, absence of loss aversion.

JEL Classification: C12, C14, C44, C58, D81, G11, G40.

[a] Athens University of Economics and Business, [b] University of Geneva and Swiss Finance Institute: corresponding author; GFRI, Bd du Pont d'Arve 40, 1211 Geneva, Switzerland; +41 22 379 88 16; olivier.scaillet@unige.ch, [c] IPAG Business School and Athens University of Economics and Business.

* We would like to thank the editor, the associate editor, and the referees for constructive criticism and numerous suggestions that have led to substantial improvements over previous versions of the paper.

1 Introduction

Traditional models in economics and finance assume that investors evaluate portfolios according to the expected utility framework. The theoretical motivation for this goes back to Von Neumann and Morgenstern (1944). Nevertheless, experimental and empirical work has shown that people systematically violate expected utility theory when choosing among risky assets. Prospect theory, first described by Kahneman and Tversky (1979) (see also Tversky and Kahneman (1992)), is widely viewed as a better description of how people evaluate risk in experimental settings. While the theory contains many remarkable insights, it has proven challenging to apply these insights in asset pricing, and it is only recently that there has been real progress in doing so (Barberis et al. (2021)). Barberis and Thaler (2003) and Barberis (2013) are excellent reviews on behavioral finance and prospect theory.

Stock market anomalies are key drivers of innovation in asset pricing. These are tradable portfolio strategies, usually constructed as long-short portfolios based on the top and bottom deciles of sorted stocks, according to specific characteristics (anomalies). Under the standard Mean-Variance (M-V) paradigm, establishing a cross-sectional return pattern as an anomaly involves testing for pricing based on a factor model. If factors are traded, spanning regressions relate to M-V criterion. Arbitrage pricing stipulates that a portfolio of factors is M-V efficient and no other portfolio can achieve a higher Sharpe Ratio (SR). In that sense, an anomaly is a strategy that exhibits higher SR and should be traded away. However, we can question M-V spanning for portfolio selection if returns do not follow elliptical distributions, or investor preferences depend on more than the first two moments of the return distribution. Chalamandaris et al. (2020) compare the M-V spanning and second-order stochastic spanning of 13 standard empirical asset pricing anomalies under various factor models using the Huberman-Kandel M-V spanning test (Huberman and Kandel (1987)). They show that the M-V spanning tests do not reconcile in-sample and out-of-sample. In the in-sample tests, almost all these anomalies reject the M-V spanning, but out-of-sample only a few of them are genuine anomalies. On the contrary, the results under second-order stochastic spanning match in-sample and out-of-sample. Our paper shows that this finding holds as well for prospect stochastic spanning on those 13 anomalies and 5 additional ones used by Barberis et al. (2021). Besides, experimental evidence (Baucells and Heukamp (2006)) suggests that investors do not always act as risk averters. Instead, they behave in a much more complex fashion, exhibiting characteristics of both risk-loving and risk-averting. They behave differently on gains and losses, and they are more sensitive to losses than to gains (loss aversion, i.e., the tendency to prefer avoiding losses to acquiring equivalent gains). The relevant utility function could be concave for gains and convex for losses (S-Shaped).

The present study contributes to this literature by introducing, operationalizing and applying new prospect spanning tests for portfolio analysis. The general research question is whether a given investment possibility set $\mathbb{K}$, namely the benchmark set, contains portfolios which prospect dominate all alternatives in an expanded investment possibility set $\mathbb{L}$. Imposing less restrictive assumptions and allowing for risk-seeking preferences, prospect spanning tests may include portfolios in the efficient set that the M-V criterion may exclude. On the other hand, the less informationally demanding M-V criterion may include portfolios that the stochastic spanning criterion may exclude. Therefore, efficient portfolio sets under the stochastic spanning and M-V criterion may be non-nested, in the sense that neither of them is a subset of the other. The use of a M-V criterion by a prospect investor might induce an opportunity cost coming from an expected utility loss caused by wrong decision making in terms of investments.

Stochastic spanning (Arvanitis et al. (2019)) is a model-free alternative to M-V spanning of Huberman and Kandel (1987) (see also Jobson and Korkie (1989), De Roon et al. (2001)). Spanning occurs if introducing new securities or relaxing investment constraints does not improve the investment possibility set for a given class of investors. M-V spanning checks if the M-V frontier of a set of assets is identical to the M-V frontier of a larger set made of those assets plus additional assets (Kan and Zhou (2012), Penaranda and Sentana (2012)). Here, we investigate such a problem for investors with prospect type preferences which are interested in the whole return distributions generated by two sets of assets, namely we study stochastic dominance.

As in Arvanitis et al. (2019), we have that any prospect spanning set provides an outer approximation of the set of prospect efficient portfolios. It is useful in at least two ways. First, if the spanning set is small enough, the problem of optimal choice is reduced to a potentially simpler problem. Indeed, a spanning set is a reduction of the original portfolio set without loss of investment opportunities for any investor with S-shaped preferences. Second, if an algorithm for the choice of non-trivial candidate spanning sets is available, we can use it to construct decreasing sequences of prospect spanning sets that ensure the convergence to the efficient set. Given the complexity of the prospect efficient set (see for example Ingersoll (2016)), such an approach can be useful for the determination of its properties.

Moreover, the rejection of prospect spanning admits an interpretation in conjunction with SSD spanning. Our theory indicates that whenever SSD spanning holds, rejection of its prospect counterpart with an optimal utility threshold large enough implies investment opportunities for some prospect-type investors that at least exhibit absence of loss aversion. This association of loss aversion to the combination of SSD and PSD spanning is, to our knowledge, new in the literature and is used here to empirically study market anomalies w.r.t. standard factor models.

Thus, the second contribution of the paper is to examine if well-known stock market anomalies expand the investment opportunity set for prospect investors. To do so, we test if trading strategies are genuine violations of standard factor models. More precisely, in our in-sample analysis, we use the prospect spanning test in order to check whether a portfolio set originating from a standard factor model, $\mathbb{K}$, spans the same set augmented with a market anomaly, $\mathbb{L}$. If the hypothesis of prospect spanning holds, the particular market anomaly can be explained by the factor model in the framework of S-shaped utilities. Then, the trading strategy that is identified in the literature as market anomaly may not be an attractive investment opportunity for prospect investors. On the contrary, if the hypothesis is not true, the anomaly expands the opportunity set for prospect investors, and is useful to that extent. As discussed above, if SSD spanning is additionally not rejected, the anomaly appears attractive to some prospect investors exhibiting absence of loss aversion. We thus also test for SSD spanning, and focus particularly in the anomalies that reject PSD spanning but not SSD spanning. In those cases, we also identify smoothed versions of the empirically optimal S-shaped utilities, their associated curvature properties and their local levels of loss aversion. We are the first in the literature to unveil those empirical features associated to a PSD setting.

In our out-of-sample analysis, we examine whether the cross-sectional patterns that are found to expand the set of factors in-sample, maintain their abnormal returns out-of-sample. Therefore, we use out-of-sample back-testing experiments as an independent criterion for robustness of in-sample test results (Harvey et al. (2016)). It turns out that prospect spanning tests produce remarkably consistent results both in- and out-of-sample in identifying trading strategies as genuine market anomalies for prospect investors. We complement the out-of-sample analysis for the cases where SSD-spanning is not empirically rejected, yet prospect spanning cannot be accepted, by performing a Rosenberg and Engle (2002) pricing kernel analysis. It relies on conditional distributions (GARCH model) and a modification of the pricing kernel to suit our S-shaped utility setting. We report the out-of-sample time variation of the associated risk aversion/seeking coefficients that appear in the pricing kernel in conjunction to the business cycle. We further show that there are asymmetries for the links between risk aversion, risk seeking, and consumer confidence measured by the US index of Consumer Sentiment (UMCSENT).

The third contribution of the paper is to compare prospect spanning with M-V spanning both in- as well as out-of-sample. That comparison reveals several differences, for example, in terms of portfolio performance and weight allocations. Here, we opt for the portfolio perspective and contribute to the anomalies literature by asking whether the so-called anomalies are a good investment opportunity for prospect investors. Prospect theory is a valid alternative to M-V for building portfolios (and judging anomalies as investment opportunities). Andrew (2014), as well as Cochrane's NBER keynote speech (2021), criticise M-V optimizers as too sensitive on mean and covariance estimates. The portfolio optimization using average return and covariance matrix estimates could be devilishly unstable. They both argue for using alternative utility functions when building portfolios.

Let us now briefly review applications of prospect theory in finance. Benartzi and Thaler (1995) utilize prospect theory to present an approach called myopic loss aversion which consists of two behavioural concepts, namely loss aversion and mental accounting. Barberis et al. (2001) study asset prices in an economy where investors derive direct utility not only from consumption but also from fluctuations in the value of their financial wealth. They are loss averse over these fluctuations and how loss averse they are depends on their prior investment performance. The design of their model is influenced by prospect theory. Barberis and Huang (2008) study the pricing of financial securities when investors make decisions according to cumulative prospect theory. Several other papers confirm that positively skewed stocks have lower average returns (Boyer et al. (2010), Bali et al. (2011), Kumar (2009), Conrad et al. (2013)). Barberis and Xiong (2009, 2012) and Ingersoll and Jin (2013) show that theoretical investment models based on S-Shape utility maximisers help to understand the disposition effect found empirically in many studies (see, e.g., Odean (1988), Grinblatt and Han (2005), Frazzini (2006), Calvet et al. (2009)). Kyle et al. (2006) provide a formal framework to analyze the liquidation decisions of economic agents under prospect theory. He and Zhou (2011) study the impact of prospect theory on optimal risky exposures in portfolio choice through an analytical treatment. Ebert and Strack (2015) set up a general version of prospect theory and prove that probability weighting implies a strong skewness preference. Barberis et al. (2016) test the hypothesis that, when thinking about allocating money to a stock, investors mentally represent the stock by the distribution of its past returns and then evaluate this distribution in the way described by prospect theory. Moreover, Barberis et al. (2021) present a model of asset prices in which investors evaluate risk according to prospect theory and examine its ability to explain prominent stock market anomalies.

The paper is organised as follows. Sections 2 and 3 include our methodological, statistical, and numerical contributions. Specifically, in Section 2 we review the definition of prospect stochastic dominance relation and define the concept of prospect spanning. We provide a new representation of the dominance relation based on a class of S-shaped utility functions constructed as convex mixtures of appropriate "ramp functions", in the spirit of Russel and Seo (1989). It avoids the differentiability assumption in the representation of Levy and Levy (2002) and constitutes a convenient setting for numerical analysis as well as economic interpretation in terms of expected utility. We explain how no rejection of SSD spanning together with rejection of PSD Spanning points to loss aversion. Using an empirical approximation of the representation of the dominance relation, we construct a test for the null hypothesis of spanning based on sub-sampling. We also derive the limiting power of the test under particular local alternatives in the Online Appendix.

In Section 3 we provide a numerical approximation of the statistic based on the utility representation derived before. For every such utility representation, we solve two embedded linear maximization problems. It is an improvement over the implementation in Arvanitis and Topaloglou (2017) and Arvanitis et al. (2020) where they formulate tests in terms of Mixed-Integer Programming (MIP) problems. MIP problems are known in Operations Research to be NP-complete, and far more difficult to solve. Our numerical approximations are simple and fast since they are based on standard LP. They better suit re-sampling methods, which otherwise quickly become computationally demanding in empirical applications. We also show that the numerical approximation converges to the test statistic for each sample size when the number of piece-wise linear components of the utilities approaches infinity.

In Section 4, we perform an empirical application where we use the prospect spanning tests to evaluate stock market anomalies using standard factor models. We consider three such models that build on the pioneer three-factor model of Fama and French (1993): the four-factor model of Hou et al. (2015), the five-factor model of Fama and French (2015), and the four-factor model of Stambaugh and Yuan (2017). Given the extensive set of results produced under alternative spanning criteria, the analysis is confined to 11 well-known strategies used to construct Stambaugh-Yuan factors, along with 7 extra (18 overall) that attracted significant attention, namely Betting against Beta, Quality minus Junk, Size, Growth Option, Value (Book-to-Market), Idiosyncratic Volatility and Profitability. The 11 anomalies used in Stambaugh and Yuan (2017) are realigned appropriately to yield positive average returns. In particular, anomaly variables that relate to investment activity (Asset Growth, Investment to Assets, Net Stock Issues, Composite Equity Issue, Accruals) are defined low-minus-high decile portfolio returns, rather than high-minus-low. All the other anomalies are constructed as high-minus-low decile portfolio returns. We emphasize that this paper is not intended to compare factor models in terms of their ability to capture the cross-section of expected returns under prospect preferences. Each factor model is our initial system of investment coordinates which we take as a granted opportunity set, without questioning its asset pricing validity. We view here the factors solely as investable assets (since they correspond to tradable strategies based on asset portfolios), and similarly for the anomalies. The anomalies might be labelled by other authors as factors if indeed priced in the cross-section, but we do not address such a research question in this paper. In the empirical analysis under our portfolio perspective, we also investigate the post-publication period for each anomaly as in Chinco et al. (2021) to test which anomalies survive after publication. We conduct the empirical analysis described above both in- and out-of-sample. In-sample, we also focus on how we can use our smooth approximation approach to find strong empirical evidence of absence of loss aversion, and identify the underlying S-Shaped utilities for the investors that have investment opportunities. We find empirical evidence of investment opportunities combined with absence of loss aversion for the Net Stock Issue anomaly under the FF-5 model, the Momentum and

Net Stock Issue anomalies under the M-4 model, and the Momentum anomaly under the q model. Out-of-sample, we focus on the representative prospect theory investor and derive the time-variation properties of the associated pricing kernels.

Finally, Section 5 concludes the paper. In Appendix A, we provide a short description of the stock market anomalies used in the empirical application. In Appendix B, we also provide a short description of the performance measures used in the out-of-sample analysis. We give in a separate Online Appendix: i) the limiting properties of the testing procedures under sequences of local alternatives, ii) Monte Carlo studies of the finite sample properties of the test, iii) the proofs of the main results, as well as several auxiliary lemmata and their proofs, iv) summary statistics of the factor and anomaly returns over our sample period from January 1974 to December 2016, additional empirical results on the v) in- and vi) out-of-sample analysis of market anomalies, and vii) outline of the pricing kernel approach.

2 Prospect Stochastic Dominance and Stochastic Spanning

The theory of stochastic dominance (SD) gives a systematic framework for analyzing investor behavior under uncertainty (see Chapter 4 of Danthine and Donaldson (2014) for an introduction oriented towards finance). Stochastic dominance ranks portfolios based on general regularity conditions for decision making under risk (see Hadar and Russell (1969), Hanoch and Levy (1969), and Rothschild and Stiglitz (1970)). SD uses a distribution-free assumption framework which allows for nonparametric statistical estimation and inference methods. We can see SD as a flexible model-free alternative to M-V dominance of Modern Portfolio Theory (Markowitz (1952)). The M-V criterion is consistent with expected utility for elliptical distributions such as the normal distribution (Chamberlain (1983), Owen and Rabinovitch (1983), Berk (1997)), but has limited economic meaning when we cannot completely characterize the probability distribution by its location and scale. Simaan (1993), Athayde and Flores (2004), and Mencia and Sentana (2009) develop a mean-variance-skewness framework based on generalizations of elliptical distributions that are fully characterized by their first three moments. SD presents a further generalization that accounts for all moments of the return distributions without necessarily assuming a particular family of distributions.

Inspired by previous work, Levy and Levy (2002) formulate the notions of prospect stochastic dominance (PSD) (see also Levy and Wiener (1998), Levy and Levy (2004)) and Markowitz stochastic dominance (MSD). Those notions extend the well-known first-degree stochastic dominance (FSD) and second-degree stochastic dominance (SSD). PSD and MSD investigate choices by investors who have S-shaped utility functions and reverse S-shaped utility functions. Arvanitis and Topaloglou (2017) develop consistent tests for PSD and MSD efficiency which is an extension to the case where full diversification is allowed. Arvanitis et al. (2020) investigate MSD spanning. This paper extends those works to prospect spanning, which is consistent with prospect preferences.

2.1 Stochastic Spanning for Prospect Dominance and Analytical Representation

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, suppose that F denotes the cdf of some probability measure on $\mathbb{R}^n$. Let $G(z, \lambda, F)$ be $\int_{\mathbb{R}^n} 1_{\{\lambda^T u \leq z\}} dF(u)$, i.e., the cdf of the linear transformation $x \in \mathbb{R}^n \to \lambda^T x$ where λ assumes its values in $\mathbb{L}$, which denotes the portfolio space. We suppose that the portfolio space is a closed non-empty subset of $\mathbb{S} = \{\lambda \in \mathbb{R}^n_+ : \mathbf{1}^T \lambda = 1\}$, possibly formulated by further economic, legal restrictions, etc. In many applications, we have that $\mathbb{L} = \mathbb{S}$. Let us further explain the meaning of the weight constraints $\lambda \in \mathbb{R}^n_+$ and $\mathbf{1}^T \lambda = 1$ in economic terms when we work with excess returns and long-short portfolios as in our empirical application. The positivity constraints means that we impose to keep the sign of the long leg and short leg of the long-short portfolio tracking the anomaly. It does not mean that we do not allow short positions. The sum to one implies a constraint on the gross leverage. Gross leverage adds the short and long positions in securities, divided by asset under management, while net leverage is the difference between long and short positions in securities, divided by asset under management. In our setting, the sum to one implies that the gross leverage of the portfolio is 200% since we have 100% invested in long legs and 100% invested in short legs, while the net leverage is zero since the long and short legs compensate. On the contrary, when we work with returns and long-only portfolios, the sum to one implies a gross leverage of 100% and a net leverage of 100%, and the economic interpretation is different.

We denote a distinguished sub-collection of $\mathbb{L}$ by $\mathbb{K}$ and generic elements of $\mathbb{L}$ by λ, κ, etc. We assume that $\mathbb{L}$, $\mathbb{K}$ are convex (which is needed for the proof of Lemma 4 below). This assumption is in line with our empirical applications where the associated portfolio spaces are constructed as convex hulls of base assets. The base assets are simply considered as the vertices of the underlying space, hence need not necessarily be individual securities. The assumption thus allows for base assets that are themselves constructed via complicated portfolio constraints on the underlying individual securities; e.g. short sales, position limits, restrictions on factor loadings, etc. Hence it is a mild assumption. In order to define the concepts of PSD and subsequently of stochastic spanning, we consider $\mathcal{J}(z_1, z_2, \lambda; F) := \int_{z_1}^{z_2} G(u, \lambda, F)\, du$.

Definition 1. *κ weakly Prospect-dominates λ*, written as $\kappa \succcurlyeq_P \lambda$, iff we have the system of inequalities $P_1(z, \lambda, \kappa, F) := \mathcal{J}(z, 0, \kappa, F) - \mathcal{J}(z, 0, \lambda, F) \geq 0$, $\forall z \in \mathbb{R}_-$ and $P_2(z, \lambda, \kappa, F) := \mathcal{J}(0, z, \kappa, F) - \mathcal{J}(0, z, \lambda, F) \leq 0$, $\forall z \in \mathbb{R}_{++}$.

The expected utility representation of the relation in Levy and Levy (2002), as well as the Russel and Seo (1989) ramp function construction, show that the conditions on P_2 are consistent with concave preferences on the positive domain, while the conditions on P_1 are consistent with convex preferences on the negative domain. PSD is thus associated with non-global dispositions towards risk: risk loving preferences in the negative domain and risk aversion in the positive domain. As such, it is fundamentally different from SSD which involves the class of utilities that represent global risk aversion. The following simple example demonstrates the discrepancy between the two forms of stochastic dominance.

Consider the pair κ, λ, with

$$G(u, \kappa, F) := \begin{cases} 0, & \text{for} \quad u < -1, \\ \frac{1}{8}, & \text{for} \quad -1 \leq u < 0, \\ \Phi(u), & \text{for} \quad 0 \leq u, \end{cases} \quad G(u, \lambda, F) := \begin{cases} 0, & \text{for} \quad u < -1, \\ \frac{1}{4}(u+1), & \text{for} \quad -1 \leq u < 0, \\ \Phi(u), & \text{for} \quad 0 \leq u, \end{cases}$$

where Φ denotes the standard Normal cdf. Here, while λ dominates κ w.r.t. SSD, we have on the contrary that $\kappa \succcurlyeq_P \lambda$. On the negative domain, every risk loving agent chooses κ over λ due to the behavior of the two cdf on the interval $[-1, 0)$. That behavior also implies that κ is not chosen over λ by any risk averse agent on that interval. However, whenever the base assets vector is entirely supported on the positive real line, then Prospect dominance is equivalent to second-order dominance. If the joint distribution is further elliptical, it is then also equivalent to M-V dominance (see Arvanitis et al. (2018)).

Given the stochastic dominance relation above, stochastic spanning occurs when augmentation of the portfolio space does not enhance investment opportunities, or equivalently, investment opportunities are not lost when the portfolio space is reduced. The following definition clarifies the concept w.r.t. the Prospect dominance relation.

Definition 2. $\mathbb{K}$ Prospect-spans $\mathbb{L}$ ($\mathbb{K} \succcurlyeq_P \mathbb{L}$) iff for any $\lambda \in \mathbb{L}$, $\exists \kappa \in \mathbb{K} : \kappa \succcurlyeq_P \lambda$. If $\mathbb{K} = \{\kappa\}$, the element κ of the singleton $\mathbb{K}$ is termed as Prospect super-efficient.

The efficient set of the dominance relation is the subset of $\mathbb{L}$ that contains the maximal elements. The efficient set is a spanning subset of the portfolio space. Thereby, any superset of the efficient set is also a spanning subset of $\mathbb{L}$. We can consider a spanning set as an outer approximation of the efficient set. Given a candidate spanning set exists, the question is whether it actually spans the portfolio space. If a method for answering such a question also exists, we can accurately approximate the efficient set via the choice of finer spanning subsets of the portfolio space. It helps in understanding decision making and investment choice.

Hence, the question we address here is: given a candidate $\mathbb{K}$, is $\mathbb{K} \succcurlyeq_P \mathbb{L}$? The following lemma provides an analytical characterization by means of nested optimizations, which, along with the economic representation of the relation in terms of expected utility in the following section, is key for a numerical implementation on real data and statistical inference. Its proof (see the Online Appendix) is based on parameter continuity arguments for the functionals involved. It is of similar form to the analogous functional employed in the SSD spanning by Arvanitis et al. (2017), yet containing an additional layer of optimization, a supplementary complexity to be handled in the proof. It is due to the two component system of inequalities that appears in the definition of the relation and it is analogous to the representation of spanning w.r.t. the Markowitz dominance relation in Arvanitis et al. (2020).

Lemma 3. *Suppose that $\mathbb{K}$ is closed. Then $\mathbb{K} \succcurlyeq_P \mathbb{L}$ iff we get the condition*

$\rho(F) := \max_{i=1,2} \sup_{\lambda \in \mathbb{L}} \sup_{z \in A_i} \inf_{\kappa \in \mathbb{K}} P_i(z, \lambda, \kappa, F) = 0$, *where* $A_1 = \mathbb{R}_-$, $A_2 = \mathbb{R}_{++}$.

2.2 Representation By Utility Functions

Here, we provide an expected utility characterization of spanning. First, it generalizes the utility characterization of PSD in Levy and Levy (2002), in that it does not require two-sided differentiability of the utilities involved. Local representations of the convex/concave components of the utilities involved imply that they have locally integrable one sided derivatives (see Appendix C.3 of Pollard (2002)), but need not have integrable derivatives. Hence, our representation enriches the class of functions involved. It allows us to analyze markets that contain investors with non-smooth preferences. Such preferences can be associated with situations in which information about the optimality of the equilibrium allocation is not fully characterized by equilibrium prices (see Ohtaki (2019)). Hence, our extension permits market conditions potentially involving ambiguity about asset equilibrium allocations. Second and foremost, our approach is in the spirit of the Russel and Seo (1989) representations for SSD. We rely on utilities represented as unions of graphs of convex mixtures of appropriate "ramp functions" on each half-line. Aside its economic interpretation, and given Lemma 3, it is key to the numerical LP implementation of the inferential procedures. In the next section, it enables a finite dimensional approximation of the utility class by functions with piece-wise linear components (see Arvanitis et al. (2017) for a simpler construction for SSD).

To this end, we denote with $\mathcal{W}_-, \mathcal{W}_+$, the sets of Borel probability measures on the real line with supports that are closed subsets of $\mathbb{R}_-$ and $\mathbb{R}_+$, respectively, with existing first moments and uniformly integrable. The latter requirement is convenient yet harmless since orderings are invariant to utility re-scaling. Those sets are convex, closed w.r.t. the topology of weak convergence and their union contains the set of degenerate measures.

Define $V_- := \left\{v_w : \mathbb{R}_- \to \mathbb{R},\ v_w(u) = \int_{\mathbb{R}_-} \left[z1_{u\leq z} + u1_{z\leq u\leq 0}\right] dw(z),\ w \in \mathcal{W}_-\right\}$, and $V_+ := \left\{v_w : \mathbb{R}_+ \to \mathbb{R},\ v_w(u) = \int_{\mathbb{R}_+} \left[u1_{0\leq u\leq z} + z1_{z\leq u<+\infty}\right] dw(z),\ w \in \mathcal{W}_+\right\}$. Every element of V_+ is increasing and concave, and dually every element of V_- is increasing and convex. Furthermore, any function defined by the union of the graph of an arbitrary element of V_+ with the graph of an arbitrary element of V_- is the graph of an S-shaped utility function as defined by Levy and Levy (2002). Such a utility function is concave for gains and convex for losses. Denote the set of S-shaped utility functions obtained by such graph unions as V. Thereby,

$$V := \left\{ v : \mathbb{R} \to \mathbb{R},\ v(u) = \begin{cases} v_w^-(u), & u \leq 0 \\ v_w^+(u), & u \geq 0 \end{cases},\ \text{where } v_w^- \in V_-, v_w^+ \in V_+ \right\}.$$

Lemma 4. *We have* $\rho(F) = \max_{i=1,2} \sup_{v_w \in V_i} \left[\sup_{\lambda \in \mathbb{L}} \mathbb{E}_\lambda \left[1_{u \in A_i} v_w(u) \right] - \sup_{\kappa \in \mathbb{K}} \mathbb{E}_\kappa \left[1_{u \in A_i} v_w(u) \right] \right],$ *where* $\mathbb{E}_\lambda$ *denotes expectation w.r.t.* $G(z, \lambda, F)$*. If the hypotheses of Lemma 3 hold, then* $\mathbb{K} \succcurlyeq_P \mathbb{L}$ *iff,* $\sup_{v \in V} \left[\sup_{\lambda \in \mathbb{L}} \mathbb{E}_\lambda [v] - \sup_{\kappa \in \mathbb{K}} \mathbb{E}_\kappa [v] \right] = 0.$

The first part of the lemma connects the functional that represents spanning to the aforementioned classes of utilities. We exploit it below in order to obtain feasible numerical formulations based on LP. Those formulations are reminiscent of the LP programs developed in the early papers of testing for SSD efficiency of a given portfolio by Post (2003) and Kuosmanen (2004). The second part of Lemma 4 crystalizes the economic characterization of spanning w.r.t. investment opportunities. It states that spanning holds if and only if the reduction of investment opportunities from $\mathbb{L}$ to $\mathbb{K}$ does not reduce optimal choices uniformly w.r.t. this class of preferences. Equivalently, it essentially states that when spanning does not hold, the restriction of the portfolio possibilities from $\mathbb{L}$ to $\mathbb{K}$ results to the maximal expected utility loss given by $\sup_{v \in V} \left[\sup_{\lambda \in \mathbb{L}} \mathbb{E}_\lambda [v] - \sup_{\kappa \in \mathbb{K}} \mathbb{E}_\kappa [v] \right] > 0$. The reason is that $\mathbb{K}$ misses some PSD efficient elements. The proof (see the Online Appendix) depends on standard integration by parts for Lebesgue-Stieljes integrals, and a min-max theorem.

2.3 Link Between Rejection of PSD Spanning and Loss Aversion

Lemma 4 implies a characterization of no-spanning by an optimal portfolio w.r.t. an optimal S-shaped utility. Specifically, $\mathbb{K}$ does not span $\mathbb{L}$ if and only if there exists an S-shaped utility, and an element of $\mathbb{L}$ that is strictly preferred to every portfolio in $\mathbb{K}$ by this utility. It implies that the preferred portfolio lies necessarily in $\mathbb{L} - \mathbb{K}$, and is optimal, i.e., it must be the best choice of the particular utility in comparison to every element in $\mathbb{K}$ (and also in $\mathbb{L}$). The utility function is optimal in the sense that it yields the maximal expected utility differences between the optimal portfolio and every element of $\mathbb{K}$.

Furthermore, the arguments in the proof of Lemma 4 and the linearity of the elements of V_+, for

$w \in \mathcal{W}_+$, and V_-, for $w \in \mathcal{W}_-$, imply that the optimal utility made of v_w^- for negative returns and v_w^+ for positive returns corresponds to degenerate mixing measures, i.e., the negative part corresponds to a probability measure w that is degenerate (point mass) at a supremum negative threshold z_-, and the positive part corresponds to a probability measure w that is degenerate to a supremum positive threshold z_+. It means that the optimal utility $u^\star$ is a member of the set of extreme points of V, and is characterized by a negative threshold z_- and a positive threshold z_+.

Under further conditions, we can associate the optimal S-shaped utility with absence of loss aversion. We employ a slightly weaker localized-form of the Kahneman and Tversky (1979) definition of loss aversion, in order to account for the specific piece-wise linear form of the extreme points of V: an increasing S-shaped u exhibits local loss aversion if and only if (i) $u(x) \leq -u(-x)$, for all $x > 0$, with at least some strict inequality; it exhibits loss aversion if and only if (i) holds with strict inequalities, and, if the utility is also differentiable it exhibits strong loss aversion (see Wakker and Tversky (1993)) if it exhibits loss aversion and (ii) $u'(x) < u'(-x)$, for all $x > 0$. A sufficient condition for absence of loss aversion is thus $u(x) > -u(-x)$, for some $x > 0$. Local loss aversion reversal holds when (i') $u(x) \geq -u(-x)$, for all $x > 0$, with at least one strict inequality, loss aversion reversal holds when (i') holds strictly, and upon differentiability, strong loss aversion reversal holds when (i') holds strictly and moreover (ii') $u'(x) > u'(-x)$, for all $x > 0$. The following lemma establishes the connection between rejection of spanning and absence of loss aversion:

Lemma 5. *Suppose that, (a) the supports of the base assets in $\mathbb{L}$ are bounded below by $-\underline{z} < 0$, (b) PSD spanning does not hold, (c) SSD spanning holds, and (d) $z_+ \geq \underline{z}$. Then the optimal utility $u^\star$ exhibits absence of loss aversion.*

The proof of the lemma is straightforward. As mentioned above, non PSD spanning implies the existence of the optimal S-shaped $u^\star$ as an extreme point of V. The lower bound of the excess returns is $-\underline{z}$. Hence, the negative threshold cannot be less than that. It is due to that, if $z_- < -\underline{z}$, then the expected optimal utility would be concave w.r.t. to the joint distribution of the returns, and thus since SSD spanning holds, the optimal portfolio in $\mathbb{L}$ would not be strictly preferred over every portfolio in $\mathbb{K}$, and it would yield a contradiction. Hence, $z_- > -\underline{z}$, and since $z_+ \geq \underline{z}$, we have that $|z_-| < z_+$, and thereby $u^\star(x) > -u^\star(-x)$, for all $x \geq |z_-|$.

The lower bound hypothesis is mild in our empirical finance framework. The returns for the long legs are naturally bounded from below by -1. In theory, the short legs could lead to an unbounded negative excess return if prices of the short leg portfolios are unbounded. In practice, we have doubts about this happening and assuming a finite lower bound for excess returns induced by bounded negative prices does not seem restrictive for our empirics. The minimum of the empirical minima of every time series of weekly excess returns in our analysis is not lower than -12%; see Table 5 in the Online Appendix.

As with loss aversion, absence of loss aversion is also documented in psychological and behavioral experiments; see Yechiam and Hochman (2014) and Walasek and Stewart (2015). Our result establishes that it can emerge as a characteristic of an optimal utility via a combination of spanning w.r.t. different stochastic dominance relations under restrictions on the concave part.

Whenever absence of loss aversion is obtainable for the optimal utility, its piece-wise linear form cannot ascertain whether it is due to (strong) loss aversion reversal. One way to obtain information on whether absence of loss aversion or (strong) loss aversion reversal is the case at hand, is to smooth the optimal utility by approximating it by an S-shaped utility with adequate curvature. In our empirical analysis, we perform smoothing using the prospect theory exponential utilities of Koebberling and Wakker (2005). Those S-shaped utilities exhibit a well-defined index of loss aversion and enough curvature so as to ascertain whether conditions (i') and (ii") above hold. We focus on cases where statistical inference does not accept PSD spanning, cannot reject SSD spanning, and the empirical optimal positive threshold is large enough, with some predefined level of statistical confidence. We then perform smoothing and test whether (i') (and (ii')) holds or not.

2.4 An Asymptotically Exact and Consistent Test for Spanning

We cannot directly rely on Lemma 3 for empirical work if F is unknown and/or the optimizations are infeasible. We construct a feasible statistical test for the null hypothesis of $\mathbb{K} \succcurlyeq_P \mathbb{L}$ by utilizing an empirical approximation of F and by building feasible and fast optimizations with LP. The null and alternative hypotheses take the following forms: $\mathbf{H_0} : \rho(F) = 0$, and $\mathbf{H_a} : \rho(F) > 0$. In the special case of singleton $\mathbb{K}$, the hypotheses write as in Arvanitis and Topaloglou (2017).

We consider a process $(Y_t)_{t \in \mathbb{Z}}$ taking values in $\mathbb{R}^n$. $Y_{i,t}$ denotes the i^{th} element of Y_t. The sample path of size T is the random element $(Y_t)_{t=1,\ldots,T}$. In our empirical finance framework, it represents returns of n financial assets upon which we can construct portfolios via convex combinations. F is the cdf of Y_0 and F_T is the empirical cdf associated with the random element $(Y_t)_{t=1,\ldots,T}$. Under our assumptions below, F_T is a consistent estimator of F, so we consider the following Kolmogorov-Smirnov type test statistic $\rho_T := \sqrt{T}\rho(F_T) = \sqrt{T}\max_{i=1,2}\sup_{\lambda\in\mathbb{L}}\sup_{z\in A_i}\inf_{\kappa\in\mathbb{K}} P_i(z,\lambda,\kappa,F_T)$, which is the scaled empirical analog of $\rho(F)$. The consideration of empirical analogues of the functionals that represent SD spanning properties, instead of traditional M-V spanning tests (Huberman and Kandel (1987), Jobson and Korkie (1989), De Roon et al. (2001)) in the context of SSD, is explained in Section 4 of Arvanitis et al. (2017) upon deviations from ellipticity. A fortiori, in our context of S-shaped preferences, traditional M-V spanning tests are generally non-admissible due to the fundamental differences between the M-V and Prospect dominance relations when the supports include components of the negative domain as explained and shown with a counterexample in Section 2.1. In addition to the economic interpretation of an estimated maximal expected utility loss from the discussion after Lemma 4, we prefer to use a Kolmogorov-Smirnov type statistic, instead of, say, a Cramer-von Mises type statistic, because of the availability and tractability of the numerical approximation through LP that is exemplified in the next section.

The following assumption enables the derivation of the limit distribution of ρ_T under $\mathbf{H_0}$ and is weaker than Assumption 2 in Arvanitis et al. (2020).

Assumption 6. *F is absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}^n$ with convex support, and for some $0<\delta$, $\mathbb{E}\left[\|Y_0\|^{2+\delta}\right]<+\infty$. $(Y_t)_{t\in\mathbb{Z}}$ is a-mixing with mixing coefficients $a_T = O(T^{-a})$ for some $a>1+\frac{2}{\eta}$, $0<\eta<2$, as $T\to\infty$.*

The mixing part is readily implied by concepts such as geometric ergodicity which holds for many stationary models used in the context of financial econometrics under parameter restrictions and restrictions on the properties of the underlying innovation processes. Examples are the strictly stationary versions of (possibly multivariate) ARMA or several GARCH and stochastic volatility type of models (see Francq and Zakoian (2011) for several examples). The moment condition is established in the aforementioned models via restrictions on the properties of building blocks and the parameters of the processes involved.

For the derivation of the limit theory of ρ_T under the null hypothesis, we consider the contact sets $\Gamma_i = \left\{\lambda \in \mathbb{L}, \kappa \in \mathbb{K}_{\lambda}^{\succeq}, z \in A_i : P_i(z, \lambda, \kappa, F) = 0\right\}$, where $\mathbb{K}_{\lambda}^{\succeq} := \{\kappa \in \mathbb{K} : \kappa \not\succeq_P \lambda\}$ which under the null contains elements different from λ for any element of $\mathbb{L} - \mathbb{K}$. For any i, the set Γ_i is non empty since $\Gamma_i^{\star} := \{(\kappa, \kappa, z), \kappa \in \mathbb{K}, z \in A_i\} \subseteq \Gamma_i$. Furthermore, $(\lambda, \kappa, 0) \in \Gamma_1, \forall \lambda, \kappa$. Since $\underline{z} := \inf_{\lambda, Y_0} \lambda' Y_0$ exists from Assumption 6, we have that for all $z \leq \underline{z}$, $(\lambda, \kappa, z) \in \Gamma_i$, $\forall \lambda \in \mathbb{L}, \kappa \in \mathbb{K}_{\lambda}^{\succeq}$ for the i that corresponds to the sign of $\underline{z}$. In what follows, we denote convergence in distribution by $\rightsquigarrow$.

Proposition 7. *Suppose that $\mathbb{K}$ is closed, Assumption 6 holds and that $\mathbf{H_0}$ is true. Then as $T \to \infty$, $\rho_T \rightsquigarrow \rho_\infty$, where $\rho_\infty := \max_{i=1,2} \sup_\lambda \sup_z \inf_\kappa P_i(z, \lambda, \kappa, \mathcal{G}_F)$, $(\lambda, z, \kappa) \in \Gamma_i$, and $\mathcal{G}_F$ is a centered Gaussian process with covariance kernel given by $Cov(\mathcal{G}_F(x), \mathcal{G}_F(y)) = \sum_{t \in \mathbb{Z}} Cov\left(1_{\{Y_0 \leq x\}}, 1_{\{Y_t \leq y\}}\right)$ and $\mathbb{P}$ almost surely uniformly continuous sample paths defined on $\mathbb{R}^n$.*

The limiting random variable ρ_∞ obtained by optimization on a functional of a Gaussian process is well defined since the following inequalities hold
$\int_0^{+\infty} \sum_{t \in \mathbb{Z}} \text{Cov}\left(1_{\{\lambda^T Y_0 \leq u\}}, 1_{\{\lambda^{Tr} Y_t \leq u\}}\right) du \leq 2 \sum_{t=0}^{\infty} \sqrt{a_T} \int_0^{+\infty} \sqrt{1 - G(u, \lambda, F)} du < +\infty$, and
$\int_{-\infty}^0 \sum_{t \in \mathbb{Z}} \text{Cov}\left(1_{\{\lambda^T Y_0 \leq u\}}, 1_{\{\lambda^{Tr} Y_t \leq u\}}\right) du \leq 2 \sum_{t=0}^{\infty} \sqrt{a_T} \int_{-\infty}^0 \sqrt{G(u, \lambda, F)} du < +\infty$, where the first inequalities in each of the previous expressions follow from inequality 1.12b in Rio (2000), and the second ones follow from Assumption 6 (see also p. 196 of Horvath et al. (2006)).

Since F and Γ_i are unknown in practice, we use the results of the previous lemma to construct a decision procedure based on sub-sampling, in the spirit of Linton et al. (2014) (see also Linton et al. (2005)).[1]

Algorithm 8. *It consists of the following steps:*

1. Evaluate ρ_T at the original sample value.

2. For $0 < b_T \leq T$, generate subsample values

   from the original observations $(Y_l)_{l=t,\ldots t+b_T-1}$ for all $t = 1, 2, \ldots, T - b_T + 1$.

[1] The partitioning used to get the results in Proposition 7 directly leads to the consideration of sub-sampling as a re-sampling procedure. A testing procedure based on (block) bootstrap as in Scaillet and Topaloglou (2010), can, due to the form of the recentering, be consistent, but can be too conservative asymptotically, and thereby suffer from a lack of power compared to the sub-sampling under particular local alternatives (see also the relevant discussion in Arvanitis et al. (2019)). The potential of asymptotic exactness for the sub-sampling test justifies the particular re-sampling choice for inference.

3. Evaluate the test statistic on each sub-sample value
   thereby obtaining $\rho_{T,b_T,t}$ for all $t = 1, 2, \ldots, T - b_T + 1$.

4. Approximate the cdf of the asymptotic distribution under the null of ρ_T
   by $s_{T,b}(y) = \frac{1}{T-b_T+1} \sum_{t=1}^{T-b_T+1} 1\left(\rho_{T,b_T,t} \leq y\right)$ and calculate its $1-\alpha$ quantile
   $q_{T,b_T}(1-\alpha) = \inf_y \left\{s_{T,b}(y) \geq 1-\alpha\right\}$, for the significance level $0 < \alpha < .5$.

5. Reject the null hypothesis $\mathbf{H_0}$ if $\rho_T > q_{T,b_T}(1-\alpha)$.

In order to derive the limit theory for the testing procedure, namely its asymptotic exactness and consistency stated in the next theorem, we first use the following standard assumption that restricts the asymptotic behaviour of b_T governing the size $b_T + 1$ of each subsample.

Assumption 9. *Suppose that (b_T), possibly depending on $(Y_t)_{t=1,\ldots,T}$, satisfies the condition $\mathbb{P}\left(l_T \leq b_T \leq u_T\right) \to 1$, 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 $T \to \infty$.*

Theorem 10. *Suppose Assumptions 6 and 9 hold. For the testing procedure described in Algorithm 8, we have that*

1. *If $\mathbf{H_0}$ is true, and for $\lambda \in \mathbb{L} - \mathbb{K}$, $\inf_{Y_0} \lambda^{Tr} Y_0 \leq 0$ there exists $(\kappa, z) \in \mathbb{K}_\lambda^{\succ} \times \mathbb{R}_{++}$ with $(\lambda, \kappa, z) \in \Gamma_2$ and that if $(\lambda, \kappa^\star, z^\star) \in \Gamma_2$ for $\kappa^\star \neq \kappa$ then $z^\star \neq z$, then for all $\alpha \in (0, .5)$ $\lim_{T\to\infty} \mathbb{P}\left(\rho_T > q_{T,b_T}(1-\alpha)\right) = \alpha$.*

2. *If $\mathbf{H_a}$ is true then $\lim_{T\to\infty} \mathbb{P}\left(\rho_T > q_{T,b_T}(1-\alpha)\right) = 1$.*

When, for $\lambda \in \mathbb{L} - \mathbb{K}$, $\inf_{Y_0} \lambda^{Tr} Y_0 \leq 0$, then, due to Assumption 6 for any contact triple $(\lambda, \kappa, z) \in \Gamma_2$, we have that $P_2\left(z, \lambda, \kappa, \mathcal{G}_F\right)$ must be non-degenerate. Whenever z corresponds solely to the particular κ, we obtain that ρ_∞ is non-degenerate and if its cdf jumps at the infimum of its support, then the jump magnitude is bounded above by .5. Hence, the test is asymptotically exact for all the usual choices of the significance level since the probability of rejection under the null hypothesis, i.e., the size of the test, reaches α in large samples. We combine Proposition 6 above and Theorem 3.5.1 of Politis et al. (1999) in the proof of the exactness statement, namely point 1 of Theorem 10. To get exactness, the condition imposed on $\mathbb{L} - \mathbb{K}$ is significantly weaker than the assumption on the relation between the extreme points of $\mathbb{L}$ and $\mathbb{K}$ adopted by Arvanitis et al. (2020). It amounts to the existence of a spanned portfolio whose support is not strictly positive and so that, in the event of positive returns, there exists an elementary increasing and concave utility for positive returns and a unique portfolio such that the latter dominates the former and we are indifferent between the two portfolios with this particular utility. Besides, the test is also consistent since the probability of rejection under the alternative hypothesis, i.e., the power of the test, reaches 1 in large samples. We show in the proof of the consistency statement, namely point 2 of Theorem 10, that the test statistic diverges to $+\infty$ under the alternative hypothesis when T goes to $+\infty$.

We opt for the "bias correction" regression analysis of Arvanitis et al. (2019) to reduce the sensitivity of the quantile estimates $q_{T,b_T}(1-\alpha)$ on the choice of b_T in empirically realistic dimensions for n and T (see also Arvanitis et al. (2020) for further evidence on its better finite sample properties). Specifically, given α, we compute the quantiles $q_{T,b_T}(1-\alpha)$ for a "reasonable" range of b_T. In the empirical section, we use $b_T \in \{T^{0.6}, T^{0.7}, T^{0.8}, T^{0.9}\}$. Next, we estimate the intercept and slope of the following regression line by OLS: $q_{T,b_T}(1-\alpha) = \gamma_{0;T,1-\alpha} + \gamma_{1;T,1-\alpha}(b_T)^{-1} + \nu_{T;1-\alpha,b_T}$. Finally, we estimate the bias-corrected $(1-\alpha)$-quantile as the OLS predicted value for $b_T = T$: $q_T^{BC}(1-\alpha) := \hat{\gamma}_{0;T,1-\alpha} + \hat{\gamma}_{1;T,1-\alpha}(T)^{-1}$. Since $q_{T,b_T}(1-\alpha)$ converges in probability to $q(\rho_\infty, 1-\alpha)$ and $(b_T)^{-1}$ converges to zero as $T \to \infty$, $\hat{\gamma}_{0;T,1-\alpha}$ converges in probability to $q(\rho_\infty, 1-\alpha)$ and the asymptotic properties are not affected. In the Online Appendix, we also show that, under further assumptions, the test is asymptotically locally unbiased under given sequences of local alternatives. Besides, the Monte Carlo analysis reported in the Online Appendix shows that the test performs well with an empirical size close to 5% and an empirical power above 90% for a significance level $\alpha = 5\%$.

3 Numerical Implementation

We exploit the results of Lemma 4 in order to provide with a finitary approximation of the test statistic. We rely on this approximation to provide with a numerical implementation based on LP below. We approximate the S-shaped utility function in the risk-seeking area using a finite set of increasing and convex piece-wise-linear functions, while we approximate the risk-averse part by increasing and concave piece-wise-linear functions. The latter is also used in Arvanitis et al. (2019) for SSD spanning. We denote expectation w.r.t. the empirical measure by $\mathbb{E}_{F_T}$. Let $\mathcal{R}^-$ denote $\max_{i=1,\dots,n} \text{Range}\left(Y_{i,t}1_{Y_{i,t}\leq 0}\right)_{t=1,\dots,T} = [\underline{x}, 0]$. Partition $\mathcal{R}^-$ into n_1 equally-spaced values as $\underline{x} = z_1 < \cdots < z_{n_1} = 0$, where $z_n := \underline{x} - \frac{n-1}{n_1-1}\underline{x}$, $n = 1, \cdots, n_1$; $n_1 \geq 2$. Furthermore, partition the interval $[0,1]$, as $0 < \frac{1}{n_2-1} < \cdots < \frac{n_2-2}{n_2-1} < 1$, $n_2 \geq 2$. Similarly, $\mathcal{R}^+ := \max_{i=1,\dots,n} \text{Range}\left(Y_{i,t}1_{Y_{i,t}\geq 0}\right)_{t=1,\dots,T} = [0, \overline{x}]$. Partition $\mathcal{R}^+$ into p_1 equally-spaced values as $0 = z_1 < \cdots < z_{p_1} = \overline{x}$, where $z_p := \frac{p-1}{p_1-1}\overline{x}$, $n = 1, \cdots, p_1$; $p_1 \geq 2$, and again partition the interval $[0,1]$, as $0 < \frac{1}{p_2-1} < \cdots < \frac{p_2-2}{p_2-1} < 1$, $p_2 \geq 2$. Using the above, we consider the test statistic:

$$\rho_T^\star := \sqrt{T} \max_{i=1,2} \sup_{v \in V_i^\star} \left[\sup_{\lambda \in \mathbb{L}} \mathbb{E}_{F_T}\left[v\left(\lambda^T Y\right)\right] - \sup_{\kappa \in \mathbb{K}} \mathbb{E}_{F_T}\left[v\left(\kappa^T Y\right)\right] \right], \tag{1}$$

where the set of utility functions for negative returns is:

$$V_-^\star := \left\{ v : v(u) = \sum_{n=1}^{n_1} w_n \left[z_n 1_{\underline{x} \leq u \leq z_n} + u 1_{z_n \leq u \leq 0} \right], \ (w_1, \dots, w_{n_1}) \in \mathrm{W}^- \right\},$$

$$\mathrm{W}^- := \left\{ (\mathrm{w}_1, \dots, \mathrm{w}_{\mathrm{n}_1}) \in \left\{ 0, \frac{1}{\mathrm{n}_2 - 1}, \cdots, \frac{\mathrm{n}_2 - 2}{\mathrm{n}_2 - 1}, 1 \right\}^{\mathrm{n}_1} : \sum_{\mathrm{n}=1}^{\mathrm{n}_1} \mathrm{w}_\mathrm{n} = 1 \right\},$$

and the set of utility functions for positive returns is:

$$V_+^\star := \left\{ v : v(u) = \sum_{p=1}^{p_1} w_p \left[u 1_{0 \leq u \leq z_p} + z_p 1_{z_p \leq u \leq \overline{x}} \right], \ (w_1, \dots, w_{p_1}) \in \mathrm{W}^+ \right\},$$

$$\mathrm{W}^+ := \left\{ (\mathrm{w}_1, \dots, \mathrm{w}_{\mathrm{p}_1}) \in \left\{ 0, \frac{1}{\mathrm{p}_2 - 1}, \cdots, \frac{\mathrm{p}_2 - 2}{\mathrm{p}_2 - 1}, 1 \right\}^{\mathrm{p}_1} : \sum_{\mathrm{p}=1}^{\mathrm{p}_1} \mathrm{w}_\mathrm{p} = 1 \right\}.$$

We obtain the following result on the approximation of ρ_T by $\rho_T^\star$.

Proposition 11. *As $n_1, n_2, p_1, p_2 \to \infty$, we have $\rho_T^\star \to \rho_T$, $\mathbb{P}$ a.s.*

Our feasible computational strategy builds on LP formulations for the numerical evaluation using the previous finitary approximation of the test statistic. We have a set of increasing and convex utility functions: $v(u) = \sum_{n=1}^{n_1} w_n \max(u, z_n)$ for the negative part, namely the risk-seeking area. For every $v \in V_-^\star$, we have at most n_2 line segments with knots at n_1 possible outcome levels. Then, we can enumerate all $n_3 = \frac{1}{(n_1-1)!}\prod_{i=1}^{n_1-1}(n_2+i-1)$ elements of $V_-^\star$. Our application in Section 4 uses $n_1 = 10$, and $n_2 = 5$, which gives $n_3 = 715$ distinct utility functions, and a total of 1430 small LP problems for the two embedded maximization problems in (1). Solving (1) yields simultaneously the optimal factor portfolio κ, and the optimal augmented portfolio λ that maximize the expected utility. Below, we give the mathematical formulation for the first optimization problem $\sup_{\lambda\in\Lambda}\mathbb{E}_{F_N}\left[u\left(\lambda^T Y\right)\right]$, that yields the optimal augmented portfolio λ. The same formulation is used for the second optimization $\sup_{\kappa\in\kappa}\mathbb{E}_{F_N}\left[u\left(\kappa^T Y\right)\right]$. Let us define: $c_{0,n} := \sum_{m=n}^{n_1}\left(c_{1,m} - c_{1,m+1}\right)z_m$, $c_{1,n} := \sum_{m=n}^{n_1} w_m$, and $\mathcal{N} := \{n = 1, \cdots, n_1 : w_n > 0\}\bigcup\{n_1\}$. For any given $u \in V_-$, $\sup_{\lambda\in\Lambda}\mathbb{E}_{F_N}\left[u\left(\lambda^T Y\right)\right]$ is the optimal value of the objective function of the following LP problem in canonical form: $\max T^{-1}\sum_{t=1}^{T} y_t$ s.t. $y_t \le \lambda^T Y_t c_{1,n} + Q_t^- + Q_t^+$, $y_t \le c_{0,n} + Q_t^- + Q_t^+$, $Q_t^- \ge c_{0,n} - \lambda^T Y_t c_{1,n}$, $Q_t^+ \ge \lambda^T Y_t c_{1,n} - c_{0,n}$, $Q_t^- \ge 0$, $Q_t^+, \ge 0$, $\sum_{i=1}^{M}\lambda_i = 1$, $\lambda_i \ge 0$, and y_t being free, for $t = 1, \cdots, T$, $n \in \mathcal{N}$, $i = 1, \cdots, M$,

For the positive part, namely the risk-averse area, we take a set of concave utility functions: $v(u) = \sum_{p=1}^{p_1} w_p \min(u, z_p)$. Again, for every $v \in V_+^\star$, we have at most p_2 line segments with knots at p_1 possible outcome levels. As before, the number of elements of $V_+^\star$ is $p_3 = \frac{1}{(p_1-1)!}\prod_{i=1}^{p_1-1}(p_2+i-1) = 715$, for $p_1 = 10$ and $p_2 = 5$. Let us define: $c_{0,p} := \sum_{m=p}^{p_1}\left(c_{1,m} - c_{1,m+1}\right)z_m$, $c_{1,p} := \sum_{m=p}^{p_1} w_m$, and $\mathcal{P} := \{p = 1, \cdots, p_1 : w_p > 0\}\bigcup\{p_1\}$. For any given $u \in V_+$, $\sup_{\lambda\in\Lambda}\mathbb{E}_{F_N}\left[u\left(\lambda^T Y\right)\right]$ is the optimal value of the objective function of the following LP problem in canonical form: $\max T^{-1}\sum_{t=1}^{T} y_t$ s.t. $y_t \le \lambda^T Y_t c_{1,p}$, $y_t \le c_{0,p}$, $\sum_{i=1}^{M}\lambda_i = 1$, $\lambda_i \ge 0$, and y_t being free, for $t = 1, \cdots, T$, $n \in \mathcal{P}$, $i = 1, \cdots, M$. The total run time for each computation does not exceed one minute when we use a desktop PC with a 3.6 GHz, 6-core Intel i7 processor, with 16 GB of RAM, using MATLAB and GAMS with the Gurobi optimization solver.

4 Empirical Application

We examine whether we can explain well-known stock market anomalies by standard factors within a new breed of asset pricing models, for prospect type investor preferences. For this purpose, we use the prospect spanning tests, both in-sample and out-of-sample. We also test in-sample for SSD spanning and we focus on the anomalies where prospect spanning is rejected while SSD spanning is not rejected, and the empirical positive threshold of the optimal utility is large enough, since in these cases rejection is associated to absence of loss aversion.

4.1 Factor Models and Anomalies

We consider three models that build on the pioneer three-factor model of Fama and French (1993): the four-factor model of Hou et al. (2015), the five-factor model of Fama and French (2015), and the four-factor model of Stambaugh and Yuan (2017). Fama and French (1993) aim to capture the part of average stock returns left unexplained in CAPM of Sharpe (1964) and Lintner (1965) by including, in addition to the market factor, two extra risk factors relating to size (measured by market equity) and the ratio of book-to-market equity. In addition to the market excess return, the influential three-factor model of Fama and French (1993) includes a book-to-market or "value" factor, HML, and a size factor, SMB, based on market capitalization. Motivated by Miller and Modigliani (1961), Fama and French (2015) five-factor model (henceforth, FF-5) augments the original Fama-French three-factor model by two extra factors, one for profitability and another for investment. Hou et al. (2015) consider a four-factor model (dubbed the q factor model) that includes the original market and size factors of Fama and French (1993) augmented by a profitability and investment factor. Stambaugh and Yuan (2017) consider a four-factor model (henceforth, M-4) including the standard market and size factors along with two composite factors for investment and profitability. To construct the composite factors, they combine information from 11 market anomalies relating to investment and profitability measures. We use alternative factor models as a robustness check, namely for testing the consistency of in- and out-of-sample results under the prospect preferences, and not for a horse race in cross-sectional asset pricing.

The stock market anomalies have a long history in the relevant literature. A common theme in the original papers that first highlighted these patterns is that they all challenge the rational asset pricing paradigm as they exhibit returns that are not in line with the risks taken. However, notwithstanding whether they are caused by sentiment (a catch-all term that stands for all kinds of irrational decision-making) and/or by market frictions (e.g., margin requirements), it is also acknowledged that most of them persist because they cannot be “arbitraged” away. From the perspective of the Arbitrage Pricing Theory, it implies that arbitrageurs cannot trade against them without exposing themselves to significant risks. The anomalies are realigned appropriately to yield positive average returns.

In particular, anomaly variables that relate to investment activity (Asset Growth, Investment to Assets, Net Stock Issues, Composite Equity Issues, Accruals) are defined low-minus-high decile portfolio returns, rather than high-minus-low, as in Hou et al. (2015). All the other anomalies are constructed as high-minus-low decile portfolio returns. A short description of the 18 market anomalies that we study in the paper is given in Appendix A (see Stambaugh and Yuan (2017) for further details). 12 of these anomalies are used in Barberis et al. (2021). Returns of the Fama and French 5 factors were downloaded from Kenneth French's site. The dataset consists of all weekly observations from January 1974 until December 2016, a total of 2236 weekly returns. M-4 factor returns and anomaly spread return series were downloaded from the websites of Robert Stambaugh and AQR. In the Online Appendix, we report summary statistics of the factor and anomaly returns over our sample period.

4.2 In-Sample Analysis

The hypothesized portfolio manager with prospect preferences forms optimal portfolios from two separate asset universes: the first universe, the set $\mathbb{K}$, is solely formed of basis factors from a factor model (FF-5, M-4, q). The second universe, the set $\mathbb{L}$, is formed when the respective set of basis factors is augmented by a single trading (spread) strategy. Portfolio managers are assumed to solve portfolio optimization problems, effectively looking for a portfolio picked from the augmented universe $\mathbb{L}$ that dominate all portfolios of the respective factor universe $\mathbb{K}$ w.r.t. some S-shaped utility. Such a portfolio is empirically approximated by the the optimal portfolio λ that produces ρ_T for the particular sample value, when the prospect spanning hypothesis is rejected. By construction this portfolio is empirically efficient (see Definition 2.1 in Linton et al. (2014) for the SSD case which we can be easily adapted to the PSD case).

Many of the anomalies employed were published post 1974 and since they build on CRSP/Compustat data, they typically work at least from the mid 60s until the publication year. To account for that, we repeat the analysis for each anomaly only post publication as in Chinco et al. (2021), to test which anomalies survive after publication. We use weekly data to have enough observations. We also report tests of stochastic spanning for risk averse investors (Arvanitis et al. (2019) to check whether rejection occurs solely due to the local risk-seeking behaviour of the investors. We additionally test for M-V spanning using the Huberman and Kandel (1987) test and compare the results.

4.2.1 PSD and SSD Spanning Tests

We compute the sub-sampling distribution of the test statistic for sub-sample size $b_T \in \{T^{0.6}, T^{0.7}, T^{0.8}, T^{0.9}\}$. Using OLS regression on the empirical quantiles $q_{T,b_T}(1-\alpha)$ for a significance level $\alpha = 5\%$, we get the estimate q_T^{BC} for the bias-corrected critical value. We reject spanning if the test statistic $\rho_T^{\star}$ is higher than the regression estimate q_T^{BC}. Tables 6-8 in the online Appendix report the test statistics $\rho_T^{\star}$ as well as the regression estimates q_T^{BC} when we test for prospect spanning of the alternative factor models w.r.t. each one of the 18 market anomalies, both in the full period as well as in the post publication period. We additionally report the test statistics $\eta_T^{\star}$ as well as the regression estimates g_T^{BC} for SSD spanning, to check whether there are anomalies that reject spanning due to risk-seeking only behaviour. Chalamandaris et al. (2021) test for SSD spanning of market anomalies. We repeat their tests and we also test for SSD spanning of the additional market anomalies that we use.

For the full sample, we observe that the FF-5 model spans 6 out of 18 market anomalies, that is, Accruals, Asset Growth, Return on Assets, Size, Growth Option, and Profitability. The M-4 model spans the same 6 market anomalies, while the q model spans Return on Assets, Betting against Beta, Size, and Profitability. Thus, in most cases, optimal portfolios based on the investment opportunity set that includes a market anomaly is not spanned by the corresponding optimal portfolio strategies based on the original factors. We also observe that Return on Assets, Size, and Profitability are spanned by all the factor models, indicating the robustness of these characteristics being not considered as genuine market anomalies by prospect investors. The lack of rejection for anomalies such as size and profitability is not surprising given that these anomalies are related to factors in the original investment sets.

In contrast to the full period, only 5 out of the 18 characteristics are not spanned by any factor model in the post publication period. These genuine anomalies are the Composite Equity Issue, Momentum, Betting Against Beta, Value and Idiosyncratic Volatility. Although Net Stock Issues is not spanned by the first two factor models, it is spanned by the q model. The rejection for the value anomaly in the FF-5 model is not surprising given that, with the addition of profitability and investment factors, the value factor of the FF-3 model becomes redundant for describing US average returns (Fama and French (2015)). The descriptive statistics of Table 5 in the Online Appendix show that the value anomaly and the HML factor have a different behaviour in terms of kurtosis, Sharpe ratio, and minimum return. The observed differences in terms of their distributional features come from different re-balancing schemes. Fama and French (1992) calculate B/M for each stock and form the HML strategy by updating value once a year on June 30, using book and price as of the prior December 31. They hold these values constant until re-balancing the portfolio on the following June. We use the Asness and Frazzini (2013) Value strategy where the HML is constructed using current prices instead of lagged fiscal year-end prices.

4.2.2 M-V Spanning Tests

Following Huberman and Kandel (1987), the set of factors M-V spans the augmented set with the anomaly if the minimum-variance frontier of the factors is identical to the minimum-variance frontier of the augmented set. For each anomaly i, we use their joint test of an intercept restriction $\alpha_i = 0$ and slope restriction $\delta_i = 1 - \beta_i e = 0$, where e is a vector of ones, and α_i and β_i are the intercept and factor loadings. The second restriction is consistent with the PSD spanning tests with portfolio weights summing to one.

Tables 9-11 in the online Appendix report M-V spanning test results of each market anomaly relative to FF-5, M-4 and q factors, respectively, using three test statistics. The first and second columns in each table display the value and heteroskedasticity-adjusted t-statistic of alpha coefficient. The third and fourth columns relate to the regression-based test of Huberman and Kandel (1987) showing, respectively, the Lagrange Multiplier (LM) and Likelihood Ratio (LR) statistics, both of which are suitable for asymptotic tests (Kan and Zhou (2008)). To be consistent with the prospect spanning tests which use constrained portfolios, we perform the M-V spanning tests but we constrain $\boldsymbol{\beta}$ to be non-negative in the optimization problem. M-V spanning tests broadly agree that most strategies are market anomalies, irrespective of the factor set. It differs markedly from PSD spanning. M-V spanning LM and LR tests agree that 18 out of 18 strategies represent anomalies (reject spanning) with respect to FF-5 factors in the full sample, and 13 out of 18 in the post publication period. We observe qualitatively similar results when assessing trading strategies against M-4 factors. In this case, test statistics agree that 15 out of 18 strategies represent market anomalies in the full sample, and again 12 out of 18 post publication. Similar results obtain with respect to q factors, where 16 out of 18 strategies represent market anomalies in the full sample, and 11 out of 18 post publication. Although in the post publication period the number of anomalies is reduced, still many strategies are found to be anomalies. The strategies that seem to be explained by the factor models after publication are the Composite Issue, Gross Profitability, Investment/Assets, Net Operating Assets, Return on Assets and Profitability. In-sample prospect spanning tests agree less often that strategies are market anomalies in the post publication period than M-V spanning. Tables 1-3 below summarise the results of the In-sample analysis.

Table 1: Fama and French (FF-5) Factors

Variable	SSD Spanning	M-V Spanning	PSD Spanning
Panel (a): Full period			
Accruals	-	Reject	-
Asset Growth	-	Reject	-
Composite Equity Issue	-	Reject	Reject
Distress	-	Reject	Reject
Growth Profitability Premium	-	Reject	Reject
Investment to Assets	Reject	Reject	Reject
Momentum	Reject	Reject	Reject
Net Operating Assets	-	Reject	Reject
Net Stock Issues	-	Reject	Reject
O-Score	-	Reject	Reject
Return on Assets	-	Reject	-
Betting against Beta	Reject	Reject	Reject
Quality minus Junk	Reject	Reject	Reject
Size	-	Reject	-
Growth Option	-	Reject	-
Value (Book to Market)	Reject	Reject	Reject
Idiosyncratic Volatility	Reject	Reject	Reject
Profitability	-	Reject	-
Panel (b): Post publication period			
Accruals	-	Reject	-
Asset Growth	-	Reject	-
Composite Equity Issue	-	-	Reject
Distress	-	Reject	-
Growth Profitability Premium	-	-	-
Investment to Assets	-	Reject	-
Momentum	Reject	Reject	Reject
Net Operating Assets	-	Reject	-
Net Stock Issues	-	-	Reject
O-Score	-	Reject	-
Return on Assets	-	-	-
Betting against Beta	Reject	Reject	Reject
Quality minus Junk	-	Reject	-
Size	-	Reject	-
Growth Option	-	Reject	-
Value (Book to Market)	Reject	Reject	Reject
Idiosyncratic Volatility	Reject	Reject	Reject
Profitability	-	-	-

Entries report the overall results of rejection for SSD, M-V and PSD spanning for the Fama and French (FF-5) model with respect to each one of the 18 market anomalies. Panel (a) uses the full period of weekly returns while Panel (b) uses the post publication period.

Table 2: Stambaugh-Yuan (M-4) Factors

Variable	SSD Spanning	M-V Spanning	PSD Spanning
Panel (a): Full period			
Accruals	-	Reject	-
Asset Growth	-	Reject	-
Composite Equity Issue	Reject	Reject	Reject
Distress	-	Reject	Reject
Growth Profitability Premium	-	Reject	Reject
Investment to Assets	Reject	-	Reject
Momentum	Reject	-	Reject
Net Operating Assets	-	Reject	Reject
Net Stock Issues	-	Reject	Reject
O-Score	-	Reject	Reject
Return on Assets	-	Reject	-
Betting against Beta	Reject	-	Reject
Quality minus Junk	-	Reject	Reject
Size	-	Reject	-
Growth Option	-	Reject	-
Value (Book to Market)	Reject	Reject	Reject
Idiosyncratic Volatility	Reject	Reject	Reject
Profitability	-	Reject	-
Panel (b): Post publication period			
Accruals	-	Reject	-
Asset Growth	-	Reject	-
Composite Equity Issue	Reject	-	Reject
Distress	-	Reject	-
Growth Profitability Premium	-	-	-
Investment to Assets	-	-	-
Momentum	-	Reject	Reject
Net Operating Assets	-	-	-
Net Stock Issues	-	Reject	Reject
O-Score	-	Reject	-
Return on Assets	-	-	-
Betting against Beta	Reject	Reject	Reject
Quality minus Junk	-	Reject	-
Size	-	Reject	-
Growth Option	-	Reject	-
Value (Book to Market)	Reject	Reject	Reject
Idiosyncratic Volatility	Reject	Reject	Reject
Profitability	-	-	-

Entries report the overall results of rejection for SSD, M-V and PSD spanning for the Stambaugh-Yuan (M-4) model with respect to each one of the 18 market anomalies. Panel (a) uses the full period of weekly returns while Panel (b) uses the post publication period.

Table 3: Hou-Xue-Zhang (q) Factors

Variable	SSD Spanning	M-V Spanning	PSD Spanning
Panel (a): Full period			
Accruals	-	Reject	Reject
Asset Growth	-	-	Reject
Composite Equity Issue	-	Reject	Reject
Distress	-	Reject	Reject
Growth Profitability Premium	-	Reject	Reject
Investment to Assets	-	Reject	Reject
Momentum	Reject	Reject	Reject
Net Operating Assets	-	Reject	Reject
Net Stock Issues	Reject	Reject	Reject
O-Score	-	Reject	Reject
Return on Assets	-	-	-
Betting against Beta	Reject	Reject	Reject
Quality minus Junk	-	Reject	Reject
Size	-	Reject	-
Growth Option	-	Reject	Reject
Value (Book to Market)	Reject	Reject	Reject
Idiosyncratic Volatility	Reject	Reject	Reject
Profitability	-	Reject	-
Panel (b): Post publication period			
Accruals	-	Reject	-
Asset Growth	-	Reject	-
Composite Equity Issue	-	-	Reject
Distress	-	Reject	-
Growth Profitability Premium	-	-	-
Investment to Assets	-	-	-
Momentum	-	Reject	Reject
Net Operating Assets	-	-	-
Net Stock Issues	-	-	-
O-Score	-	Reject	-
Return on Assets	-	-	-
Betting against Beta	-	Reject	Reject
Quality minus Junk	-	Reject	-
Size	-	Reject	-
Growth Option	-	Reject	-
Value (Book to Market)	Reject	Reject	Reject
Idiosyncratic Volatility	Reject	Reject	Reject
Profitability	-	-	-

Entries report the overall results of rejection for SSD, M-V and PSD spanning for the Hou-Xue-Zhang (q) model with respect to each one of the 18 market anomalies. Panel (a) uses the full period of weekly returns while Panel (b) uses the post publication period.

4.2.3 Performance Measures

We compute a number of commonly used performance measures: the average return (Mean), the standard deviation (SD), the Skewness and the Kurtosis of returns, the Sharpe ratio, the downside Sharpe ratio (D. Sharpe ratio) of Ziemba (2005), the upside potential and downside risk (UP) ratio of Sortino and van der Meer (1991), the opportunity cost of Simaan (2013), and a measure of the portfolio risk-adjusted returns net of transaction costs (Return Loss) of DeMiguel et al. (2009). The downside Sharpe and UP ratios are considered to be more appropriate measures of performance than the typical Sharpe ratio given the asymmetric return distribution of the anomalies. For the calculation of the opportunity cost, we use the following utility function which satisfies the curvature of prospect theory (S-shaped): $v(x) = x^a$ if $x \geq 0$ or $-\gamma(-x)^b$ if $x < 0$, where γ is the coefficient of loss aversion and $a, b < 1$ (usually $\gamma = 2.25$, $0 < a = b = 0.88$; see Tversky and Kahneman (1992)). We provide a short description of those performance measures in Appendix B. In the next lines, we only detail the results of the in-sample tests for the Momentum market anomaly in the post publication period. The latter is well documented on diverse markets and asset classes (Asness et al. (2013)). In the Online Appendix, we report the performance measures for the 5 Fama and French, the 4 Stambaugh and Yuan and the 4 Hou-Xue-Zhang optimal factor portfolios, and the optimal augmented portfolios for all the other market anomalies that we test, in the post publication period.

Table 4 reports the performance measures for the Momentum anomaly under each factor model (Panels A, B and C, respectively). The results are consistent with the prospect and the M-V spanning tests. All the performance measures of the augmented portfolios are improved w.r.t. the optimal factor portfolios. In addition, we compare the M-V optimal portfolio to the prospect portfolio in terms of prospect utility loss. We measure it by the opportunity cost θ, which is the return that needs to be added to the M-V portfolio so that the investor is indifferent in prospect utility terms. In Table 4, we see that the opportunity cost is always positive, indicating that the prospect investor is better off compared to the M-V investor. In the Online Appendix, we present analogous Tables for the other market anomalies, and similar remarks hold.

Table 4: Performance measures of the optimal in-sample spanning portfolios. The Momentum anomaly.

PSD spanning						
	Panel A		Panel B		Panel C	
	FF-5	+ anom.	M-4	+ anom.	q	+ anom.
Mean	0.0026	0.0042	0.0021	0.0051	0.0028	0.0053
SD	0.0192	0.0194	0.0168	0.0204	0.0196	0.0217
Skewness	-0.5644	-0.4576	-0.8799	-0.7566	-0.4733	-0.9655
Kurtosis	3.184	4.8577	2.6675	4.8555	2.9985	3.9066
Sharpe ratio	0.1354	0.2165	0.1250	0.2500	0.1429	0.2442
D. Sharpe ratio	0.3874	0.4692	0.4445	0.5017	0.3954	0.4761
UP ratio	0.7388	0.9456	0.8951	1.0115	0.7966	0.9601
Return Loss		0.056%		0.039%		0.028%
M-V spanning						
	Panel A		Panel B		Panel C	
	FF-5	+ anom.	M-4	+ anom.	q	+ anom.
Mean	0.0022	0.0032	0.0019	0.0042	0.0022	0.0044
SD	0.0172	0.0170	0.0160	0.0181	0.0182	0.0197
Skewness	-0.2855	-0.3665	-0.9855	-0.8966	-0.8655	-0.9755
Kurtosis	2.9555	3.385	2.9556	3.3488	3.0622	3.3422
Sharpe ratio	0.1279	0.1882	0.1188	0.2320	0.1209	0.2234
D. Sharpe ratio	0.3906	0.4809	0.4500	0.4788	0.4003	0.4903
UP ratio	0.8332	0.9686	0.9059	0.9646	0.8062	0.9882
Return Loss		0.017%		0.019%		0.014%
Prospect utility loss (Opportunity Cost)						
$\alpha = \beta = 0.2$	0.142%	0.245%	0.165%	0.258%	0.134%	0.185%
$\alpha = \beta = 0.4$	0.113%	0.201%	0.123%	0.217%	0.102%	0.160%
$\alpha = \beta = 0.6$	0.102%	0.194%	0.110%	0.185%	0.083%	0.125%

Entries report the performance measures (Mean, Standard Deviation, Sharpe ratio, Downside Sharpe ratio, UP ratio and Returns Loss), Skewness and Kurtosis for the weekly realised returns of the factor optimal portfolios, and the augmented with the Momentum optimal portfolio under PSD spanning and M-V spanning. Panel A reports measures for the case of the FF-5 factors, Panel B for the case of the M-4 factors, and Panel C for the case of the q factors. Table also shows the prospect utility loss of the M-V portfolios over the prospect portfolios.

4.2.4 Optimal Utilities: Evidence of Absence of Loss Aversion

In line with Paragraph 2.3, we now focus on anomalies where, in the post publication period, we reject PSD spanning but cannot reject SSD spanning. Tables 1-3 suggest that these are the Composite Equity Issue and Net Stock Issues for the FF-5 model, the Momentum and Net Stock Issues for the M-4 model, and the Composite Equity Issue, Momentum and Betting against Beta for the q model. We proceed by assuming a lower bound $-\underline{z}$ on weekly excess returns (see the discussion in Section 2.3). Table 5 in the Online Appendix shows that the minimum among all factors and anomalies on the large sample from January, 1974 to December, 2016, that we study, does not go below -12%. Since the minimum observed excess return is above -12%, we use a conservative value for the lower bound on weekly excess returns equal to 10 times that value, namely we set $\underline{z} = 120\%$.

Table 5 exhibits the estimates of the thresholds of the optimal utilities for the above mentioned anomalies, along with their bootstrap 95% confidence intervals (in brackets) evaluated from block bootstrapping the original returns. The number of bootstrap samples is 100. For this significance level, we observe that the lower bound of the confidence interval lies above $\underline{\boldsymbol{z}}$ in all cases, except for the Composite Equity Issue for the FF-5 model, and the Composite Equity Issue and Betting against Beta for the q model anomalies. It indicates that, for the aforementioned level of significance and our conservative choice of $\underline{\boldsymbol{z}}$, the Net Stock Issues for the FF-5 model, the Momentum and Net Stock Issues for the M-4 model, and the Momentum for the q model all lie within the scope of Lemma 5 and absence of loss aversion is expected. For comparison purposes, we provide thresholds for the other anomalies in the Online Appendix.

Table 5: Evidence of Absence of Loss Aversion

Variable	Factor Model	z_+	z_-
	FF-5		
Composite Equity Issue		1.5708	-0.085
		$[1.1716, 2.0015]$	$[-0.1236, -0.0537]$
Net Stock Issues		3.3879	-0.1701
		$[2.5756, 4.4016]$	$[-0.3419, -0.0374]$
	M-4		
Momentum		2.0366	-0.1138
		$[1.6392, 2.5409]$	$[-0.1681, -0.0541]$
Net Stock Issues		3.7391	-0.1772
		$[2.8714, 4.3582]$	$[-0.2891, -0.0333]$
	q		
Composite Equity Issue		1.0157	-0.0328
		$[0.7807, 1.3277]$	$[-0.0432, -0.0229]$
Momentum		2.1374	-0.1261
		$[1.5886, 2.5693]$	$[-0.182, -0.0687]$
Betting against Beta		0.0621	-0.0022
		$[0.0123, 0.1284]$	$[-0.0029, -0.0013]$

Entries report the estimated thresholds z_+ and z_- of the optimal utility function and their bootstrap confidence intervals at 95%, for the anomalies for which PSD spanning is rejected while SSD spanning is not rejected. The number of block bootstrap samples of weekly returns is 100.

4.2.5 Optimal Utilities: Smoothing and Absence of Loss Aversion

For the aforementioned four anomalies, we perform smoothing of the piece-wise linear optimal utilities using the prospect theory exponential utility class of Koebberling and Wakker (2005). The members of this class are represented by $v^+(x) = (1 - e^{-\delta_1^+ x})/\delta_1^+$, for $x \geq 0$, and $v^-(x) = -\delta_0^-(1 - e^{-\delta_1^-(-x)})/\delta_1^-$, for $x < 0$, with $\delta_0^- > 0$, $\delta_1^+ > 0$, and $\delta_1^- > 0$. They deliver well-defined indices of loss aversion as discussed by Koebberling and Wakker (2005) since they avoid $v'(0) = 0$ exhibited by the often used prospect theory power utility function. We have that $v^{+\prime}(x) = e^{-\delta_1^+ x}$, for $x \geq 0$, and $v^{-\prime}(x) = \delta_0^- e^{-\delta_1^-(-x)}$, for $x < 0$. For $\delta_0^- \geq 1$ and $\delta_1^- < \delta_1^+$, we get loss aversion since $-v(-x) > v(x)$, for all $x > 0$, and even its strong form (Wakker and Tversky (1993)) since $v'(-x) > v'(x)$, for all $x > 0$.

For a grid of return points, namely 1200 points, the parameters $\delta_0^-, \delta_1^+, \delta_1^-$ are optimally selected via a Nonlinear Least Squares regression of the optimal utility levels paired with the grid points and obtained from the spanning analysis. Given the parameter choice, the absence of loss aversion reversal index $v^+(x) - (-v^-(-x))$ and the strong loss aversion reversal index $v^{+\prime}(x) - v^{-\prime}(-x)$ are computed in the positive part of the grid. Confidence bands for the smoothed utilities and the aforementioned indices are obtained via the block bootstrapping method of the previous paragraph.

Figure 1 shows the propect theory exponential utility, the loss aversion reversal index $v^+(x) - (-v^-(-x))$, and the strong loss aversion reversal index $v^{+\prime}(x) - v^{-\prime}(-x)$, as well as the corresponding confidence bands at 95% for the Momentum anomaly under the M-4 model. The optimal parameters are $\delta_0^- = 1.8502$, $\delta_1^- = 16.1202$, and $\delta_1^+ = 0.3457$, and their respective 95% bootstrap confidence intervals are $[1.8345, 1.8714]$, $[10.7578, 34.3578]$, and $[0.2222, 0.4882]$. We face a strong asymmetry between the two values of risk aversion and risk seeking and it contradicts the usual choice of taking a common value for both such as the values $a = b = .88$ found by Tversky and Kahneman (1992) for a prospect theory power utility function. The figure corroborates the absence of loss aversion established in the previous paragraph. It moreover establishes that the optimal utility exhibits no loss aversion reversal, neither in the weak nor, obviously, in the strong form for our 95% significance level since we observe crossings of the zero horizontal line for low values of x. We also see in Figure 1 that the difference at $x = 0$ is such that $v^{+\prime}(0) - v^{-\prime}(0) < 0$ and so the index $v^{+\prime}(0)/v^{-\prime}(0)$ advocated by Koebberling and Wakker (2005) is below one and would conclude to presence of loss aversion locally around zero. The results for the other anomalies are presented in the Online Appendix and the results are analogous. Such a novel evidence in favour of absence of loss aversion and against loss aversion or its reversal is a direct by-product of our testing procedure. It also clarifies that $\delta_0^- > 1$ does not necessarily yield loss aversion since there is an interplay with the δ_1^- and δ_1^+ values and it is why we need to look at Figure 1. It would be the case if $\delta_1^- = \delta_1^+$, but we do not get that empirically.

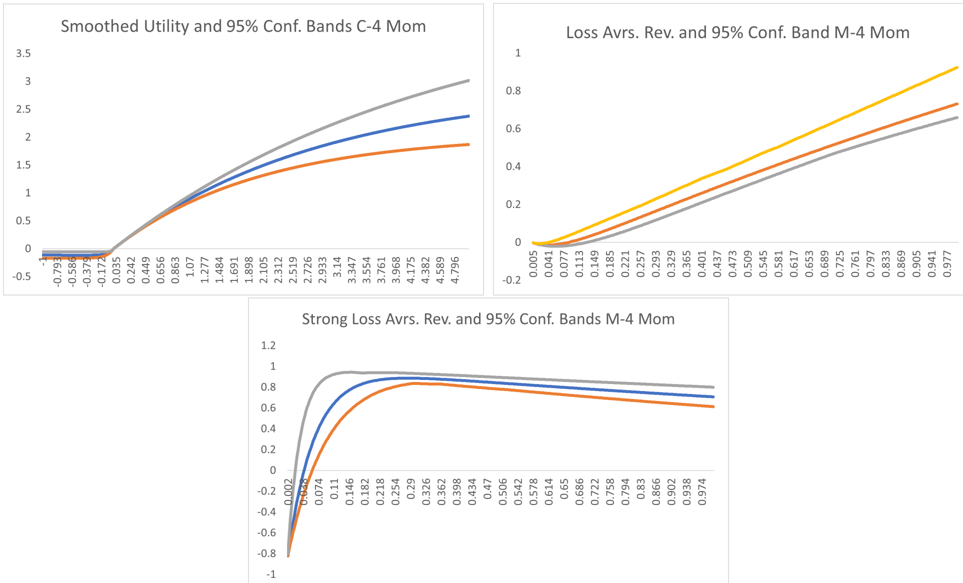

Figure 1: The left upper Figure is the prospect theory exponential utility, the right upper Figure is the loss aversion reversal index $v^{+}(x) - (-v^{-}(-x))$, and the lower Figure is the strong loss aversion reversal index $v^{+\prime}(x) - v^{-\prime}(-x)$ for the Momentum anomaly under the M-4 model. The confidence bands at 95% are computed by bootstrap.

4.3 Out-of-Sample Analysis

We conduct out-of-sample backtesting experiments using both prospect and M-V spanning in the post publication period, to examine whether the inclusion of a market anomaly in the investment opportunity set benefits to investors out-of-sample. Those experiments allow us to check the robustness of in-sample results. Although we reject the null hypothesis of prospect spanning in some cases for in-sample tests, it is not known a priori whether an optimal augmented portfolio also outperforms an optimal portfolio made of factors only in an out-of-sample analysis. It is because by construction we form these portfolios at time t, based on the information prevailing at time t, while we reap the portfolio returns over $[t, t+1]$ (next week). The out-of-sample test is a real-time exercise avoiding a potential look-ahead bias and mimicking the way that a real-time investor acts in practice.

We resort to back-testing experiments on a rolling horizon basis. The rolling windows cover the period from 01/1974 to 12/2016. First, we specify the publication date of each anomaly. At each week, we use the data from the previous 4 years (208 weekly observations) to calibrate the procedure. We solve the resulting optimization problem for the prospect stochastic spanning test and record the optimal portfolios. The clock is advanced and we determine the realized returns of the optimal portfolios from the actual returns of the various assets. Then, we repeat the same procedure for the next time period and we compute the ex post realized returns over the entire period for both portfolios.

Finally, we implement a Rosenberg and Engle (2002) pricing kernel analysis based on a prospect type representative agent framework. Specifically, for each rolling window, we use a GARCH model framework and estimate the pricing kernel emerging from a prospect theory exponential utility of the aforementioned Koebberling and Wakker (2005) utility class. Thus, we assess the time variation of the local risk seeking and risk aversion dispositions of the associated investors. We contrast those series to recession periods. For this analysis, we use conditional distributions as opposed to the in-sample analysis which by construction concerns a one-period ahead investment horizon and is based on unconditional distributions. It is due to the conditional setting of the pricing relations.

4.3.1 Performance Measures

Table 6 reports the performance measures for the Momentum anomaly under each factor model (Panels A, B and C, respectively). These performance measures supplement the evidence obtained from the in-sample analysis. For PSD spanning, we observe that the Mean, the Sharpe ratio, downside Sharpe ratio and UP ratio of the optimal augmented portfolio are improved with respect to the optimal factor portfolio. Although these measures are based on the first two moments, they support the in-sample result that the set enlarged with the momentum anomaly is not spanned by any factor model. Moreover, Skewness is negative, while Kurtosis is higher. The Return Loss is always positive. We observe that augmenting the factors by Momentum increases the performance of the optimal portfolio with respect to each factor model. These results are consistent with the in-sample tests.

In contrast, the out-of-sample results for M-V spanning are not consistent with in-sample tests. Although in-sample the Momentum is found to be an anomaly in the M-V framework, out-of sample we observe that the augmented portfolio is not improved compared to the factors optimal portfolio. Thus, we confirm the robustness of in-sample and out-of sample results for PSD spanning but not for M-V spanning. One possible reason for this inconsistency is that M-V optimization average return and covariance matrix estimates over each rolling window could be unstable. Moreover, we compare the M-V optimal portfolio with the prospect portfolio in terms of prospect utility loss. We measure it by the opportunity cost θ, which is the return that needs to be added to the M-V portfolio so that the investor is indifferent in prospect utility terms. In Table 6, we see that the opportunity cost is always positive, indicating that the prospect investor is better off compared to choices under an M-V criterion. The Momentum characteristic is negatively skewed and leptokurtic, which makes it attractive for prospect investors, and we can see in Table 6 that the kurtosis is larger after including this anomaly under PSD spanning. In the Online Appendix, we present analogous Tables for the other market anomalies. They indicate that the Composite Equity Issue, Momentum, Betting against Beta, Value, and Idiosyncratic Volatility emerge as unambiguously genuine market anomalies under all factor sets, both in-sample and out-of-sample.

Prospect investors would benefit from including these characteristics in their portfolios, expanding the investment opportunity set offered by factor portfolios. They can also prefer strategies that can produce opportunities with low skewed returns. All these anomalies have high Sharpe ratios, and the skewness is low as expected. We stress that the PSD spanning approach is particularly robust in- and out-of-sample. This remarkable consistency offers good incentives for adopting such an approach when exploring instances of apparent market inefficiency. In the M-V framework, the Composite Issue, Distress, Gross Profitability, Net Operating Assets, Return on Assets, Net Stock Issues, Betting against Beta, Value and Idiosyncratic Volatility improve the opportunity set of M-V investors out-of-sample. These characteristics exhibit high Sharpe ratio, which makes them attractive for investors that take into account the first two moments only. Finally, we observe that the prospect utility loss is almost in all cases positive, indicating the superiority of the prospect portfolios in prospect utility terms compared to the M-V portfolios. It means that one needs to give a positive return equal to θ to a M-V investor so that she becomes as happy as a prospect investor in expected utility terms.

Tables 47-49 in the Online Appendix report the average weight allocation of the optimal augmented portfolios under the PSD spanning and the M-V spanning, to understand better how the choices of a prospect investor and a M-V investor differ. Those vary substantially over time; prospect investors often include a much higher proportion of their wealth in the anomaly than the M-V investors, especially in the anomalies that provide clear portfolio improvements. They take concentrated positions in characteristics with joint low skewness and high kurtosis. The optimal weight of Momentum is between 72.6% and 80.3% for the prospect investors, indicating the superior performance of this anomaly for the prospect investor. For M-V investors, the weight of Momentum is only between 0.2% and 18.3%.

Table 6: Performance measures of the optimal spanning portfolios. The Momentum anomaly.

PSD spanning						
	Panel A		Panel B		Panel C	
	FF-5	+ anom.	M-4	+ anom.	q	+ anom.
Mean	0.0020	0.0216	0.0021	0.0214	0.0028	0.0150
SD	0.0231	0.1938	0.0167	0.2367	0.0166	0.1270
Skewness	-0.4319	-0.1734	-0.7125	-0.4311	-0.3655	-0.8107
Kurtosis	3.8235	5.8714	2.4895	4.2840	2.2883	5.7100
Sharpe ratio	0.08658	0.1114	0.1257	0.0904	0.1687	0.1181
D. Sharpe ratio	0.1408	0.6086	0.4433	0.0535	0.2468	0.4055
UP ratio	0.7435	1.2393	0.9736	1.2533	0.9946	1.1983
Return Loss		0.085%		0.024%		0.008%
M-V spanning						
	Panel A		Panel B		Panel C	
	FF-5	+ anom.	M-4	+ anom.	q	+ anom.
Mean	0.0016	0.0017	0.0019	0.0020	0.0027	0.0027
SD	0.0134	0.0140	0.0130	0.0141	0.0151	0.0167
Skewness	-0.3618	-0.2596	-0.4854	-0.3591	-0.4563	-0.4475
Kurtosis	2.8533	3.8774	2.0232	3.9947	2.9675	3.9504
Sharpe ratio	0.1194	0.1214	0.1465	0.1418	0.1788	0.1616
D. Sharpe ratio	0.1102	0.1087	0.1653	0.1659	0.2812	0.2381
UP ratio	0.6120	0.5999	0.6650	0.6494	0.8010	0.7272
Return Loss		-0.003%		-0.004%		-0.020%
Prospect utility loss						
(Opportunity Cost)						
$\alpha = \beta = 0.2$	0.162%	1.139%	0.287%	0.491%	0.159%	0.839%
$\alpha = \beta = 0.4$	0.146%	1.025%	0.228%	0.442%	0.143%	0.775%
$\alpha = \beta = 0.6$	0.131%	0.923%	0.175%	0.398%	0.129%	0.617%

Entries report the performance measures, Skewness and Kurtosis for the weekly realised returns of the factor optimal portfolios, and the augmented with the Momentum optimal portfolio under PSD spanning and M-V spanning. Panel A reports measures for the case of the FF-5 factors, Panel B for the case of the M-4 factors, and Panel C for the case of the q factors. The table also shows the prospect utility loss of the M-V portfolios over the prospect portfolios.

In the in-sample tests, we find that we reject prospect spanning far less often than we reject M-V spanning. The M-V model results in optimal portfolios under the assumption that investors have quadratic utility functions. In contrast, the stochastic dominance framework is valid under less restrictive assumptions on investor preferences but is informationally more demanding than the M-V framework as it depends on the whole distribution of returns, rather than on the two first moments only. Investors do not necessarily adhere to a M-V view, but have instead more complicated risk preferences that may pay attention to higher-order moments. On one hand, as a result of imposing less restrictive assumptions on investor preferences, SD spanning tests may include portfolios in the efficient set that the MV criterion may exclude. On the other hand, the MV criterion, which requires less information from the data, may include portfolios that PSD criterion may exclude from the efficient set. Overall, we believe that the observed discrepancies between in-sample and out-of-sample inference support the hypothesis that in-sample MV-spanning tests tend to classify strategies as market anomalies far too easily.

4.3.2 Pricing Kernel Analysis and Time-Varying Risk Aversion/Seeking

We evaluate the empirical risk aversion (in periods of gains) as well as the empirical risk seeking (in periods of losses) coefficients each week. We do so by modifying the approach of Rosenberg and Engle (2002) (see also Ait-Sahalia and Lo (2000), Jackwerth (2003)) regarding the form of the pricing kernel. Specifically, we assume that the pricing kernel at the end of each rolling window is:

$$M^*(r_{t+1};\theta_t) = \begin{cases} \theta_{0,t} e^{-\theta^+_{1,t} r_{t+1}}, & r_{t+1} \geq 0, \\ \theta_{0,t} e^{-\theta^-_{1,t}(-r_{t+1})}, & r_{t+1} < 0. \end{cases} \tag{2}$$

In order for such a pricing kernel to be compatible with S-shaped preferences, $\theta^+_{1,t}$ and $\theta^-_{1,t}$ must be strictly positive. The value $\theta^+_{1,t}$ coincides with the Arrow-Pratt measure of absolute risk aversion. Analogously, we can interpret $\theta^-_{1,t}$ as an absolute risk seeking measure. In the Online Appendix, we explain the link between the pricing kernel (2) and the Koebberling and Wakker (2005) time-varying utility function v_t characterized by $v^+_t(x) = (1 - e^{-\delta^+_{1,t} x})/\delta^+_{1,t}$ for $x \geq 0$, and $v^-_t(x) = -\delta^-_{0,t}(1 - e^{-\delta^-_{1,t}(-x)})/\delta^-_{1,t}$ for $x < 0$, with $\delta^-_{0,t} > 0$, $\delta^+_{1,t} > 0$, and $\delta^-_{1,t} > 0$. The parameter $\theta_{0,t} = \zeta_t e^{-r_{f,t+1}}$ is related to the discount factor, where $r_{f,t+1}$ is the risk-free rate and ζ_t is a strictly positive time-varying scale coefficient independent of the return level r_{t+1}. The multiplicative loss aversion coefficient $\delta^-_{0,t}$ in the exponential S-shape utility function is not identifiable from the pricing kernel; $\delta^-_{0,t}$ cancels out from the ratio of marginal utilities at negative returns. Using the prices of the anomaly portfolio and its associated optimal prospect portfolio, we estimate at each window the values of the aforementioned measures keeping it the same over a month. We provide details of the estimation method based on specification (2) of the pricing kernel in the Online Appendix. We also explain there why a standard power S-shaped utility function cannot rationalize a power pricing kernel (Rosenberg and Engle (2002)): such a utility function has $v_t'(0) = 0$ and yields an infinite pricing kernel when we compute the ratio of marginal utilities. We report the risk aversion coefficient only in periods of positive returns, and the risk seeking coefficient only in periods of negative returns.

Figure 2 depicts the weekly empirical risk aversion or risk seeking, calculated each month for the out-of-sample period for the Momentum anomaly under the M-4 model. The grey areas are the NBER recession periods. Empirical evidence of time-varying risk seeking is new in the literature, in particular the wide presence of risk seeking behaviour in recession times. Additionally, the risk aversion is lower in the recession periods, and higher in boom periods. Over the sample period, empirical risk aversion averages 4.3. However, the level fluctuates substantially, ranging from 1.16 to 9.3. Finally, the risk seeking is higher in the recession periods and lower in the boom periods. Over the sample period, empirical risk seeking averages 5.0. The level ranges from 1.42 to 9.72. In the Online Appendix, we present analogous Figures for the other market anomalies for which prospect spanning empirically disagrees with SSD spanning.

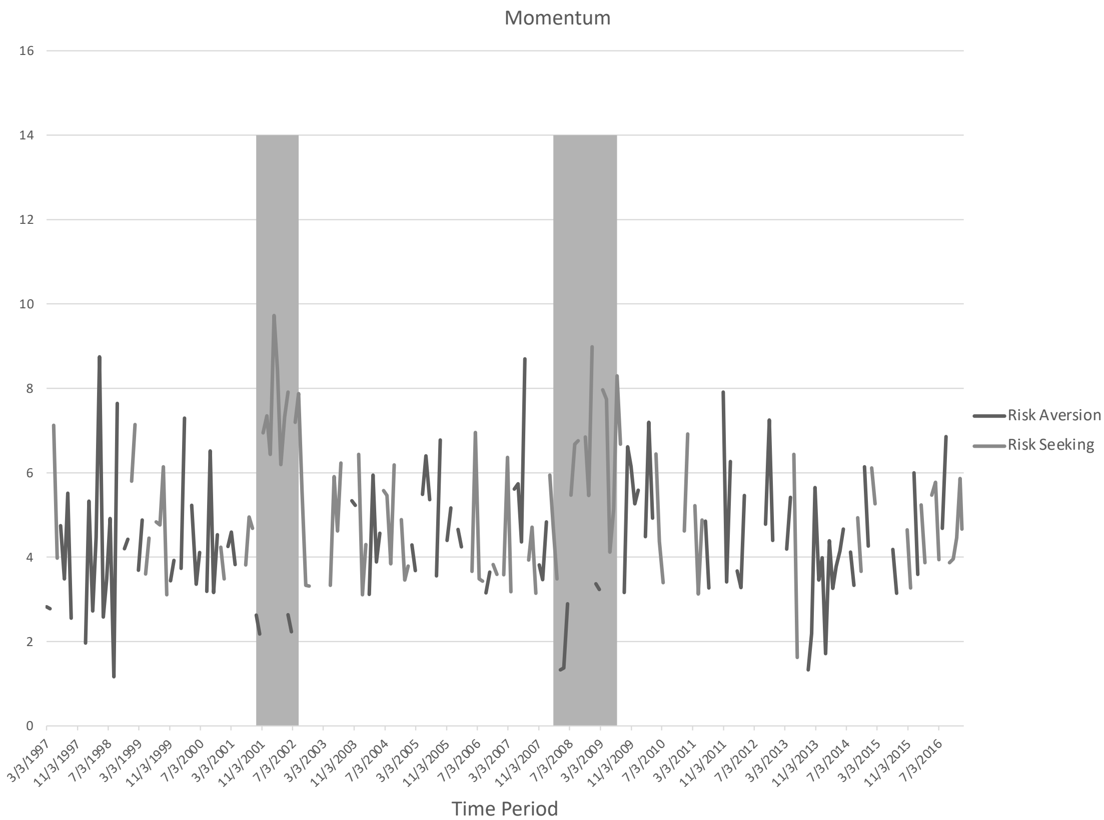

Figure 2: The Figure presents the monthly empirical risk aversion $\theta^{+}_{1,t}$, or risk seeking $\theta^{-}_{1,t}$, obtained by calibration of weekly observations for the period from March 1997 to December 2016, for the momentum anomaly under the M-4 model.

Finally, we observe in Figure 2 that the risk aversion is low at times outside recession windows. A potential explanation is an uptrend in the US index of Consumer Sentiment (UMCSENT) during those periods. This index builds on the results of the University of Michigan's monthly Survey of Consumers, and is used to estimate future spending and saving. As we observe from Table 7, the monthly changes in UMCSENT are negatively correlated with changes in the monthly risk aversion (apart from the BaB anomaly for the years 2015-2016) during periods with positive returns. When we have a marked improvement in sentiment, consumers show signs of more certainty over the trajectory of the US economy, and become less risk averse. Similarly the correlation coefficients of monthly changes in UMCSENT with the changes in monthly risk seeking during periods with negative returns are positive in most of the cases, but not always, indicating that there are asymmetries for the links between risk aversion, risk seeking, and consumer confidence.

The coefficient values in Figure 2 show empirical evidence of time-varying risk aversion and risk seeking, but they do not need to agree with the values obtained in the unconditional in-sample analysis of Section 4.2.5. Here, we get coefficients estimates from a conditional out-of-sample analysis. We use rolling windows, but not the full sample, to calibrate a conditional distribution (GARCH model) and generate out-of-sample prices to calibrate the empirical pricing kernel (see the Online Appendix for details).

Table 7: Correlation coefficients

	Corr with risk aversion	Corr with risk seeking
Fama and French (FF-5) Factor Model		
Composite Equity Issue	-0.1081	-0.0102
Net Stock Issues	-0.3812	0.4044
Stambaugh-Yuan (M-4) Factor Model		
Momentum	-0.0244	-0.0845
Net Stock Issues	-0.1065	-0.0756
Hou-Xue-Zhang (q) Factor Model		
Composite Equity Issue	-0.2213	0.1168
Momentum	-0.0197	0.0229
Betting against Beta	0.1142	0.1836

Entries report correlation coefficients of monthly changes in the US index of Consumer Sentiment (UMCSENT) with changes in the monthly risk aversion during periods with positive returns and changes in monthly risk seeking during periods with negative returns.

5 Conclusions

In this paper, we develop and implement methods for determining whether introducing new securities or relaxing investment constraints improves the investment opportunity set for prospect investors. We develop a testing procedure for prospect spanning for two nested portfolio sets based on subsampling and standard LP. In the empirics, we apply the prospect spanning framework to asset prices in which investors evaluate risk according to prospect theory and examine its ability to explain 18 well-known stock market anomalies. We find that of the strategies considered, a few expand the opportunity set of prospect investors, and thus have real economic value for them. Moreover, when they are additionally spanned by factor models in the global risk aversion framework (SSD spanning), they present investment opportunities for prospect investors, associated to the pronounced absence of loss aversion. Those are the Net Stock Issue anomaly under the FF-5 model, the Momentum and Net Stock Issue anomalies under the M-4 model, and the Momentum anomaly under the q model. Out of sample, we find through a pricing kernel approach that prospect representative agents have time-varying dispositions towards risk that seem more pronounced on the downturn periods of the business cycle. We also find that there are asymmetries for the links between risk aversion, risk seeking, and consumer confidence measured by the US index of Consumer Sentiment (UMCSENT). Those empirical features for a PSD setting were not documented in the previous literature. Finally, we show that the prospect spanning approach is particularly robust between in- and out-of-sample analysis. We also compare the prospect spanning with M-V spanning both in- as well as out-of-sample. We observe that M-V spanning results are not that robust in- and out-of-sample. Moreover, in most cases, the prospect investors are better off compared to choices under an M-V criterion, as measured by the prospect utility loss.

The paper contributes to a current strand of literature aiming to reevaluate published anomalies and discern those with real economic content for prospect investors. From a practitioner perspective, this robust framework for establishing investment opportunities for prospect investors can be of real value, especially in the case of quantitative investment funds that combine talent, capital and computational power to the purpose of exploiting the existing anomalies and discovering new ones.

References

[1] Ait-Sahalia, Y., and A. Lo, 2000. Nonparametric Risk Management and Implied Risk Aversion. Journal of Econometrics, 94, 9-51.

[2] Anderson C., and L. Garcia-Feijoo, 2006. Empirical Evidence on Capital Investment, Growth Options, and Security Returns. Journal of Finance, 61(1), 171-194.

[3] Ang, A., Hodrick, R., Xing, Y., and X. Zhang, 2006. The Cross-Section of Volatility and Expected Returns. Journal of Finance, 61(1), 259-299.

[4] Ang, A. 2014. Asset Management: A systematic Approach to Factor Investing, Oxford University Press.

[5] Arvanitis, S., Hallam, M. S., Post, T., and N. Topaloglou, 2019. Stochastic spanning. Journal of Business and Economic Statistics, 37(4), 573-585.

[6] Arvanitis, S., Post, T., and Topaloglou, N., 2021. Stochastic Bounds for Reference Sets in Portfolio Analysis. Management Science, 67(12), 7737-7754.

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

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

[9] Asness, C., and A. Frazzini, 2013. The devil in HML details. Journal of Portfolio Management, 39, 49-68.

[10] Asness, C., Frazzini, A., and L.H. Pedersen, 2019. Quality minus junk. Review of Accounting Studies, 24(1), 34-112.

[11] Asness, C., Moskowitz, T., and L.H. Pedersen, 2013. Value and Momentum Everywhere. Journal of Finance, 68(3), 929-985.

[12] Athayde G., and R. Flores, 2004. Finding a Maximum Skewness Portfolio: A General Solution to Three-Moments Portfolio Choice. Journal of Economic Dynamics and Control, 28, 1335-1352.

[13] Bali, T., Cakici, N., and R. Whitelaw, 2011. Maxing out: Stocks as Lotteries and the Cross-Section of Expected Returns. Journal of Financial Economics, 99(2), 427-446.

[14] Barberis, N., 2013. Thirty years of prospect theory in economics: A review and assessment. Journal of Economic Perspectives, 27, 173-196.

[15] Barberis, N., and M. Huang, 2008. Stocks as Lotteries: The implications of Probability Weighting for Security Prices. American Economic Review, 98(5), 2066–2100

[16] Barberis, N., Huang, M., and T. Santos, 2001. Prospect Theory and Asset Prices. Quarterly Journal of Economics, 116(1), 1-53.

[17] Barberis, N., Jin, L., and B. Wang, 2021. Prospect Theory and Stock Market Anomalies. Journal of Finance, 76(5), 2639-2687.

[18] Barberis, N., Mukherjee, A., and B. Wang, 2016. Prospect Theory and Stock Returns: An Empirical Test. Review of Financial Studies, 29, 3068-3107.

[19] Barberis, N., and R. Thaler, 2003. A Survey of Behavioral Finance. Handbook of the economics of finance, G. Constantinides, R. Stulz, M. Harris eds., North Holland.

[20] Barberis, N., and W. Xiong, 2009. What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation. Journal of Finance, 64(2), 751-784.

[21] Barberis, N., and W. Xiong, 2012. Realization Utility. Journal of Financial Economics, 104, 251-71.

[22] Baucells, M., and F. Heukamp, 2006. Stochastic Dominance and Cumulative Prospect Theory. Management Science, 52, 1409-1423.

[23] Benartzi, S., and R. Thaler, 1995. Myopic Loss Aversion and the Equity Premium Puzzle. Quarterly Journal of Economics, 110(1), 73-92.

[24] Berk, J., 1997. Necessary Conditions for the CAPM. Journal of Economic Theory, 73, 245-257.

[25] Berk, J., Green, R., and V. Naik, 1999. Optimal Investment, Growth Options, and Security Returns. Journal of Finance, 54(5), 1553-1607.

[26] Black, F., Jensen, M., and M. Scholes, 1972. The Capital Asset Pricing Model: Some Empirical Tests. In Jensen, M., editor, Studies in the Theory of Capital Markets. Praeger Publishers Inc.

[27] Boyer, B., Mitton, T., and K. Vorkink, 2010. Expected idiosyncratic Skewness. Review of Financial Studies, 23(1), 169-202.

[28] Calvet, L. , Campbell, J., and P. Sodini, 2009. Fight or Flight? Portfolio Rebalancing by Individual Investors. Quarterly Journal of Economics, 124, 301-48.

[29] Campbell, J., Hilscher, J., and J. Szilagyi, 2008. In Search of Distress Risk. Journal of Finance, 63(6), 2899-2939.

[30] Carhart, M., 1997. On Persistence in Mutual Fund Performance. Journal of Finance, 52(1), 57-82.

[31] Chalamandaris G., Pukthuanthong K., and N. Topaloglou, 2021. Are stock-market anomalies anomalous after all?. Working paper.

[32] Chamberlain, G., 1983. A Characterization of the Distributions that Imply Mean- Variance Utility Functions. Journal of Economic Theory, 29, 185-201.

[33] Chen, L., Novy-Marx, R., and L. Zhang, 2010. An Alternative Three-Factor Model, Working paper, CKGSB, University of Rochester, and Ohio State University.

[34] Chinco, A., A. Neuhierl, and M. Weber, 2021. Estimating the anomaly base rate. Journal of Financial Economics, 140(1), 101-126.

[35] Cochrane J., 2021. Portfolios for long-term investors. NBER Keynote speech paper.

[36] Conrad, J., Dittmar, R., and E. Ghysels, 2013. Ex-Ante Skewness and Expected Stock Returns. Journal of Finance, 68(1), 85-124.

[37] Cooper, M., Gulen, H., and M.J. Schill, 2008. Asset Growth and the Cross-Section of Stock Returns. Journal of Finance, 63(4), 1609-52.

[38] Daniel, K., and S. Titman, 2006. Market Reactions to Tangible and Intangible Information. Journal of Finance, 61(4), 1605-43.

[39] Danthine, J.P., and J. Donaldson, 2014. Intermediate Financial Theory. 3rd edition, North-Holland.

[40] DeMiguel, V., Garlappi, L., and R. Uppal, 2009. Optimal versus Naive Diversification: How Inefficient is the 1/n Portfolio Strategy? Review of Financial Studies, 22(5), 1915- 1953.

[41] De Roon, F. , Nijman, T., and B. Werker, 2001. Testing for Mean-Variance Spanning with Short Sales Constraints and Transaction Costs: the Case of Emerging Markets. Journal of Finance, 56(2), 721-742.

[42] Ebert, S., and P. Strack, 2015. Until the Bitter End: On Prospect Theory in a Dynamic Context. American Economic Review, 2015, 105(4), 1618-1633.

[43] Fama, E., and K. French, 1992. The Cross-Section of Expected Stock Returns. Journal of Finance, 47(2), 427-465.

[44] Fama, E., and K. French, 1993. Common Risk Factors in the Returns on Stocks and Bonds. Journal of Financial Economics, 33(1), 3-56.

[45] Fama, E., and K. French, 2008. Dissecting Anomalies. Journal of Finance, 63(4), 1653-1678.

[46] Fama, E., and K. French, 2015. A Five-Factor Asset Pricing Model. Journal of Financial Economics, 116(1), 1-22.

[47] Francq, C., and J.M. Zakoian, 2011. GARCH Models: Structure, Statistical Inference and Financial Applications. John Wiley & Sons.

[48] Frazzini, A., 2006. The Disposition Effect and Underreaction to News. Journal of Finance, 61(4), 2017-46.

[49] Frazzini, A., and L.H. Pedersen, 2014. Betting Against Beta. Journal of Financial Economics, 111(1), 1-25.

[50] Grinblatt, M., and B. Han, 2005. Prospect Theory, Mental Accounting, and Momentum. Journal of Financial Economics, 78(2), 311-339.

[51] Hadar, J., and W. Russell, 1969. Rules for Ordering Uncertain Prospects. American Economic Review, 59, 2-34.

[52] Hanoch, G., and H. Levy, 1969. The Efficiency Analysis of Choices Involving Risk. Review of Economic Studies, 36, 335-346.

[53] Harvey, C., Liu, Y., and H. Zhu, 2016, ... and the Cross-Section of Expected Returns. Review of Financial Studies, 29(1), 5-68.

[54] He, X., and X. Zhou, 2011. Portfolio Choice Under Cumulative Prospect Theory: An Analytical Treatment. Management Science, 57(2), 315-331.

[55] Hirshleifer, D., Hou, K., Teoh, S.H., and Y. Zhang, 2004. Do Investors Overvalue Firms with Bloated Balance Sheets? Journal of Accounting and Economics, 38, 297-331.

[56] Horvath, L., Kokoszka, P., and R. Zitikis, 2006. Testing for Stochastic Dominance Using the Weighted McFadden-Type Statistic. Journal of Econometrics, 133, 191- 205.

[57] Hou, K., Xue, C., and L. Zhang, 2015. Digesting Anomalies: An Investment Approach. Review of Financial Studies, 28(3), 650-705.

[58] Huberman, G., and S. Kandel, 1987. Mean-Variance Spanning. Journal of Finance, 42(4), 873-888.

[59] Ingersoll Jr, J., 2016. Cumulative Prospect Theory, Aggregation, and Pricing. Critical Finance Review, 5(2), 305-350.

[60] Ingersoll, J. , and L. Jin, 2013. Realization Utility with Reference-Dependent Preferences. Review of Financial Studies, 26(3), 723-767.

[61] Jackwerth, J. C., 2003. Recovering Risk Aversion from Option Prices and Realized Returns. Review of Financial Studies, 13(2), 433-451.

[62] Jegadeesh, N., and S. Titman, 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. Journal of Finance, 48(1), 65-91.

[63] Jobson, J., and B. Korkie, 1989. A Performance Interpretation of Multivariate Tests of Asset Set Intersection, Spanning and Mean-Variance Efficiency. Journal of Financial and Quantitative Analysis, 24, 185-204.

[64] Kahneman, D., and A. Tversky, 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(2), 263-291.

[65] Kan, R., and G. Zhou, 2012. Tests of Mean-Variance Spanning. Annals of Economics and Finance, 13-1, 145-193.

[66] Kester, W., 1984. Today's Options for Tomorrow's Growth. Harvard Business Review, 62, 153-165.

[67] Koebberling, V. and P. Wakker, 2005. An Index of Loss Aversion. Journal of Economic Theory, 122, 119-131

[68] Kumar, A., 2009. Who Gambles in the Stock Market? Journal of Finance, 64(4), 1889-1933.

[69] Kuosmanen, T., 2004. Efficient Diversification According to Stochastic Dominance Criteria. Management Science, 50(10), 1390-1406.

[70] Kyle, A., Ou-Yang, H., and W. Xiong, 2006. Prospect Theory and Liquidation Decisions. Journal of Economic Theory, 129, 273-288.

[71] Levy, M., and H. Levy, 2002. Prospect Theory: Much Ado about Nothing? Management Science, 48(10), 1334-1349.

[72] Levy, H., and M. Levy, 2004. Prospect Theory and Mean-Variance Analysis. Review of Financial Studies, 17(4), 1015-1041.

[73] Levy, H., and Z. Wiener, 1998. Stochastic Dominance and Prospect Dominance with Subjective Weighting Functions. Journal of Risk and Uncertainty, 16, 147-163.

[74] Lintner, J., 1965. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. Review of Economics and Statistics, 47, 13-37.

[75] Linton, O., Maasoumi, E., and Y.J. Whang, 2005. Consistent Testing for Stochastic Dominance under General Sampling Schemes. Review of Economic Studies, 72, 735-765.

[76] Linton, O., Post, T., and Y.J. Whang, 2014. Testing for the stochastic dominance efficiency of a given portfolio. The Econometrics Journal, 17(2), 59-74.

[77] Loughran. T., and J. Ritter, 1995. The New Issues Puzzle. Journal of Finance, 50(1), 23-51.

[78] Markowitz, H., 1952. Portfolio Selection. Journal of Finance, 7(1), 77-91.

[79] Mencia, J., and E. Sentana, 2009. Multivariate Location-Scale Mixtures of Normals and Mean-Variance-Skewness Portfolio Allocation. Journal of Econometrics, 153, 105-121.

[80] Miller, M., and F. Modigilani, 1961. Dividend Policy, Growth, and the Valuation of Shares. Journal of Business, 34(4), 411-433.

[81] Novy-Marx, R., 2013. The Other Side of Value: The Gross Profitability Premium. Journal of Financial Economics 108(1), 1-28.

[82] Novy-Marx, R., and M. Velikov, 2015. A taxonomy of anomalies and their trading costs. Review of Financial Studies, 29(1), 104-147.

[83] Odean, T., 1998. Are Investors Reluctant to Realize their Losses? Journal of Finance 53(5), 1775-98.

[84] Ohlson, J., 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research 18(1), 109-31.

[85] Ohtaki, E., 2020. Optimality in an OLG model with nonsmooth preferences. TCER working paper E-145.

[86] Owen, J., and R. Rabinovitch, 1983. On the Class of Elliptical Distributions and their Applications to the Theory of Portfolio Choice. Journal of Finance 38(3), 745-752.

[87] Penaranda, F., and E. Sentana, 2012. Spanning Tests in Portfolio and Stochastic Discount Factor Mean-Variance Frontiers: A Unifying Approach. Journal of Econometrics 170, 303-324.

[88] Pollard, D., 2002. A user's guide to measure theoretic probability (Vol. 8). Cambridge University Press.

[89] Politis, D., Romano, J., and M. Wolf, 1999. Subsampling. Springer New York.

[90] Post, T., 2003. Empirical Tests for Stochastic Dominance Efficiency. Journal of Finance 58(5), 1905-1932.

[91] Rio, E., 2000. Theorie Asymptotique des Processus Aleatoires Faiblement Dependants. Mathematiques et Applications, Vol. 31, Berlin: Springer-Verlag.

[92] Ritter, J.,1991. The Long-run Performance of Initial Public Offerings. Journal of Finance, 46(1), 3-27.

[93] Rothschild, M., and J. Stiglitz, 1970. Increasing Risk: I. A Definition. Journal of Economic Theory, 2(3), 225-243.

[94] Rosenberg, J. V., and R. F. Engle. 2002. Empirical Pricing Kernels. Journal of Financial Economics, 64(3), 341-372.

[95] Russell, W., and T. Seo, 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, New York: Springer Verlag, 59-76.

[96] Scaillet, O., and N. Topaloglou, 2010. Testing for Stochastic Dominance Efficiency. Journal of Business and Economic Statistics 28(1), 169-180

[97] Sharpe, W., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. Journal of Finance 19(3), 425-442.

[98] Simaan, Y., 1993. Portfolio Selection and Asset Pricing-Three-Parameter Framework. Management Science 39(5), 568-577.

[99] Sloan, R., 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? The Accounting Review 71(3), 289-315.

[100] Sortino, F., and R. van der Meer 1991. Downside Risk - Capturing What's at Stake in Investment Situations. Journal of Portfolio Management 17, 27-31.

[101] Stambaugh, R., and Y. Yuan, 2017. Mispricing Factors. Review of Financial Studies 30(4), 1270-1315.

[102] Titman, S., Wei, K., and F. Xie, 2004. Capital Investments and Stock Returns. Journal of Financial and Quantitative Analysis 39(4), 677-700.

[103] Tversky, A., and D. Kahneman, 1992. Advances in Prospect Theory: Cumulative representation of uncertaint. Journal of Risk and Uncertainty 5, 297-323.

[104] Von Neumann, J., and O. Morgenstern, 1944. Theory of Games and Economic Behavior. Princeton University Press.

[105] Walasek, L., and N. Stewart, 2015. How to Make Loss Aversion Disappear and Reverse: Tests

of the Decision by Sampling Origin of Loss Aversion. Journal of Experimental Psychology: General, 144(1), 7-11.

[106] Yechiam, E., and G. Hochman, 2014. Loss Attention in a Dual Task Setting. Psychological Science, 25, 494-502

[107] Ziemba, W., 2005. The Symmetric Downside Risk Sharpe Ratio. Journal of Portfolio Management, 32(1), 108-122.

APPENDIX A: Description of Stock Market Anomalies

Below we provide the origin and a short description of the 18 market anomalies used in the empirical application.

1. Accruals: Sloan (1996) argues that investors tend to overestimate in their earnings expectations the persistence of the earnings' component that is due to accruals. As a result, firms with low accruals earn on average abnormally higher returns than firms with high accruals.

2. Asset Growth: Cooper, Gulen, and Schill (2008) maintain that investors tend to overreact positively right after asset expansions. According to the authors, this behavior causes firms with high growth in their total assets to exhibit relatively lower returns over the subsequent fiscal years.

3. Composite Equity Issues: Daniel and Titman (2006) base their analysis on a measure of equity issuance that they devised finding that equity issuers tend to under-perform non-issuer firms.

4. Distress: Campbell, Hilscher, and Szilagyi (2008) find that firms with high default probability tend to exhibit lower subsequent returns. This pattern is counter-intuitive in the context of rational asset pricing, given that according to the standard models high risk entails high expected return and vice versa.

5. Gross Profitability Premium: Novy-Marx (2013) argues that gross profit is the most objective profitability metric. As a result, firms with the strongest gross profit have on average higher returns than the less profitable ones.

6. Investment to Assets: Titman, Wei, and Xie (2004) argue that investors are put off by empire-building managers who over-invest. For this reason, firms showing a significant increase in gross property, plant, equipment or inventories tend to under-perform the market.

7. Momentum: Momentum (Jegadeesh and Titman (1993)) is perhaps the most cited anomaly in asset pricing. Since Carhart factor model (1997), it has been included in various reduced-form models of the Stochastic Discount Factor as a factor. The momentum effect is attributed to sentiment and describes the pattern of "winner" stocks gaining higher subsequent returns and "loser" stocks relatively lower.

8. Net Operating Assets: Hirshleifer et al. (2004) suggest that investors often neglect information about cash profitability and focus instead on accounting profitability. Because of this bias, firms with high net operating assets (measured as the cumulative difference between operating income and free cash flow) get to have negative long-run stock returns.

9. Net Stock Issues: Ritter (1991) and Loughran and Ritter (1995) indicate that equity issuers underperform non-issuers with similar characteristics. Fama and French (2008) demonstrate that net stock issues are negatively correlated with subsequent returns.

10. O-Score: This anomaly coincides with the distress anomaly we mentioned earlier. In this case, the spread portfolios are constructed from stock ranking based on the O-score (Ohlson (1980)) to measure distress likelihood.

11. Return on Assets: Chen, Novy-Marx, and Zhang (2010) associate high past return on assets with abnormally high subsequent returns. Return on assets is measured as the ratio of quarterly earnings to last quarter's assets.

12. Betting against Beta: Black, Jensen and Scholes (1972) showed that low (high) beta stocks have consistently positive (negative) risk-adjusted returns. Frazzini and Pedersen (2014) propose an investment strategy ("betting-against-beta" (BAB)) that exploits this anomaly by buying low-beta stocks and shorting high-beta stocks. Because of its robustness, this anomaly is currently one of the most widely examined APT violations.

13. Quality minus Junk: Asness, Frazzini and Pedersen (2013) show that high-quality stocks (safe, profitable, growing, and well managed) exhibit high risk-adjusted returns. The authors attribute this pattern to mispricing.

14. Size: The market capitalization is computed as the log of the product of price per share and number of shares outstanding, computed at the end of the previous month.

15. Growth Option: Growth Option measure represents the residual future-oriented firm growth

potential. This future (yet-to-be exercised) growth option measure is calculated as the % of a firm's market value (V) arising from future-oriented growth opportunities (PVGO/V). It is inferred by subtracting from the current market value of the firm (V) the perpetual discounted stream of expected operating cash flows under a no-further growth policy (see, e.g., Kester (1984), Anderson and Garcia-Feijoo (2006), Berk, Green, and Naik (1999)).

16. Value (Book to market): The log of book value of equity scaled by market value of equity, computed following Asness and Frazzini (2013); firms with negative book value are excluded from the analysis.

17. Idiosyncratic Volatility: Standard deviation of the residuals from a firm-level regression of daily stock returns on the daily Fama-French three factors using data from the past month. See Ang et al. (2006).

18. Profitability.: It is measured as revenue minus cost of goods sold at time t, divided by assets at time t-1. Stocks with high profitability ratios tend to outperform on a risk-adjusted basis (Novy-Marx (2013), Novy-Marx and Velikov (2015)). Recent research suggests that profitability is one of the stock return anomalies that has the largest economic significance (see Novy-Marx (2013)).

APPENDIX B: Description of Performance Measures

For the downside Sharpe ratio, first we need to calculate the downside variance (or more precisely the downside risk), $\sigma^2_{P_-} = \frac{\sum_{t=1}^{T}(x_t - \bar{x})^2_-}{T-1}$, where the benchmark $\bar{x}$ is zero, and the x_t taken are those returns of portfolio P at month t below $\bar{x}$, i.e., those t of the T months with losses. To get the total variance, we use twice the downside variance namely $2\sigma^2_{P_-}$ so that the downside Sharpe ratio is, $S_P = \frac{\bar{R}_p - \bar{R}_f}{\sqrt{2}\sigma_{P-}}$, where $\bar{R}_p$ is the average period return of portfolio P and $\bar{R_f}$ is the average risk free rate. The UP ratio compares the upside potential to the shortfall risk over a specific target (benchmark) and is computed as follows. Let R_t be the realized monthly return of portfolio P for $t = 1, ..., T$ of the backtesting period, where $T = 216$ is the number of experiments performed and let ρ_t be respectively the return of the benchmark (risk free rate) for the same period. Then, we have UP ratio $= \frac{\frac{1}{K}\sum_{t=1}^{K}\max[0, R_t - \rho_t]}{\sqrt{\frac{1}{K}\sum_{t=1}^{K}(\max[0, \rho_t - R_t])^2}}$. It is obvious that the numerator of the above ratio is the average excess return over the benchmark and so reflects upside potential. In the same way, the denominator measures downside risk, i.e., shortfall risk over the benchmark.

Next, we use the concept of opportunity cost presented in Simaan (2013) to analyse the economic significance of the performance difference of the two optimal portfolios. Let R_{Aug} and R_F be the realized returns of the optimal augmented and the optimal factors portfolios, respectively. Then, the opportunity cost θ is defined as the return that needs to be added to (or subtracted from) the optimal factors portfolio return R_F, so that the investor is indifferent (in utility terms) between the strategies imposed by the two different investment opportunity sets, i.e., $E[U(1+R_F+\theta)] = E[U(1+R_{Aug})]$.

A positive (negative) opportunity cost implies that the investor is better (worse) off if the investment opportunity set allows for the market anomaly factor prospect type investing. The opportunity cost takes into account the entire probability distribution of asset returns and hence it is suitable to evaluate strategies even when the asset return distribution is not normal. For the calculation of the opportunity cost, we use the following utility function which satisfies the curvature of prospect theory (S-shaped): $U(R) = R^\alpha$ if $R \geq 0$ or $-\gamma(-R)^\beta$ if $R < 0$, where γ is the coefficient of loss aversion (usually $\gamma = 2.25$) and $\alpha, \beta < 1$.

Finally, we evaluate the performance of the two portfolios under the risk-adjusted (net of transaction costs) returns measure, proposed by DeMiguel et al. (2009) which indicates the way that the proportional transaction cost, generated by the portfolio turnover, 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, $NW_{P,t+1} = NW_{P,t}(1+R_{P,t+1})[1 - trc \times \sum_{i=1}^{N}(|w_{P,i,t+1} - w_{P,i,t}|)$. The portfolio return, net of transaction costs is defined as $RTC_{P,t+1} = \frac{NW_{P,t+1}}{NW_{P,t}} - 1$. Let μ_F and μ_{Aug} be the out-of-sample mean of monthly RTC factors and the Augmented optimal portfolio, respectively, and σ_F and σ_{Aug} be the corresponding standard deviations. Then, the return-loss measure is $R_{Loss} = \frac{\mu_{Aug}}{\sigma_{Aug}} \times \sigma_F - \mu_F$, i.e., the additional return needed so that the factors perform equally well with the optimal augmented with the market anomaly portfolio. We follow the literature and use 35 bps for the transaction costs.