What is the implied upper bound of the Stochastic Discount Factor?

Eric J. Wilson*

November 16, 2025

Abstract

I use a comprehensive set of anomaly portfolios as test assets to determine their ability to span the Stochastic Discount Factor (SDF). I make two contributions. First, I estimate the SDF from the leading principal component (PC) factors extracted from a set of anomaly portfolios — the implied upper bound — and show that they do *not* span the SDF. Second, I introduce hedge fund portfolios as a new class of test assets and show that it gets us closer to spanning the SDF by generating a Sharpe Ratio that is 3.4x higher than the traditional anomaly portfolios. Importantly, I show that test assets built from hedge fund returns are robust to the following concerns regarding the use of hedge funds in managed portfolios: (i) Leverage, (ii) Short-selling, (iii) Capacity limits, (iv) Implementability out-of-sample and (v) Diverse strategy space beyond individual stocks.

JEL Classification: G10, G12

Keywords: Stochastic Discount Factor, Test Assets, Arbitrage, Stock Returns, Variances, Factor Structure, Return Predictability, Hedge Funds

*Wilfrid Laurier University, Lazaridis School of Business & Economics. Author's email: erwilson@wlu.ca.
I am grateful to my committee members Ronald Balvers (co-chair), Yoontae Jeon, Ken Li and Jiaping Qiu (co-chair) for their guidance and support. I also thank Vikas Agarwal, Chay Ornthanalai, Alberto Martín-Utrera (discussant), conference participants at the Bank of Canada Graduate Student Paper Award Workshop, Financial Management Association Annual Meeting, Midwest Finance Association and seminar participants at McMaster University and the University of Guelph for their helpful comments.

1 Introduction

By construction, our understanding of the cross-section of expected stock returns is based on a generally accepted set of test assets. This set of test assets is taken as a set of primitives that are used to build tradeable risk factors (or anomaly portfolios). The initial search for factors comes from either economic theory or data mining. More recently, there has been a growing concern regarding whether the large and growing set of factors (i.e. the "Factor Zoo" coined by Cochrane (2011)) is primarily a consequence of data mining efforts over the past several decades (see Chen et al. (2024); McLean and Pontiff (2016) for recent and prominent examples). It remains unclear whether we have identified a set of test assets that spans the Stochastic Discount Factor (SDF). This paper (i) documents new evidence that shows how the traditional set of test assets fails to span the SDF; and (ii) introduces an alternative approach to building test assets based on hedge fund portfolios that gets us closer to spanning the SDF. Figure 1 illustrates how the current set of anomalies makes up only a small portion of the relevant space of test assets needed to span the SDF. This stylized representation showcases how the extant set of anomalies fails to span the SDF, thereby motivating the research question in this paper.

The traditional approach to generating test assets is best exemplified by Fama and French (1993). In essence, they show considerable variation in stock returns along two firm characteristics: Size and Book-to-Market (B/M). As a consequence, they sort stocks into quintiles along these two dimensions thereby generating 25 Size-B/M portfolios (i.e. test assets). With this set of test assets they create tradable long/short (L/S) factors that buy (sell) the quantile with the highest (lowest) return: Size (SMB) and B/M (HML). Since Fama and French (1993), the empirical asset pricing literature has proposed 100+ factors to explain the cross-sectional variation in stock returns. However, McLean and Pontiff (2016) note how the out-of-sample (OOS) return predictability of the most notable factors has fallen dramatically over time (58% lower post publication). If these factors in fact span the SDF, then it follows that more than half of the return predictability in stock returns has disappeared over time as many of these anomalies have been traded away.

The initial motivating evidence indicating that the current set of test assets fails to span the SDF is shown in Figures 2 and 3, in which first moments are emphasized in Figure 2 whereas Figure 3 emphasizes second moments. Figure 2 Panel A displays the rolling ten-year Sharpe ratio (SR) of an equal-weighted combination portfolio of the Jensen et al. (2023) 13 factor themes. There is a steady decline in the risk-adjusted returns of anomalies since the 2000 Technology Bubble. Panel B of Figure 2 displays the rolling ten-year information ratio (IR) of daily CAPM alphas of an equal-weighted combination portfolio of the 13 factor themes versus the value-weighted index excess return. Similar to Panel A, there is also a notable downward trend in IRs over the same period. Together, both panels of Figure 2 reveal the basic insight of this paper that builds on the evidence from McLean and Pontiff (2016) documenting how the anomaly returns have declined steadily over time (as evidenced in Panel A). If the anomaly portfolios indeed span the SDF then the logical implication is that risk premia have declined over time across all major assets (including the market portfolio). However, the market's risk premium has increased over time *relative* to the anomaly portfolios, as evidenced by Panel B of Figure 2. Therefore, the anomaly portfolios do not span the SDF.

Figure 3 suggest that the anomaly portfolios have become less relevant over time. Panel A displays the realized volatility of the value-weighted market index (black line) in comparison to the realized volatility of an equal-weighted combination portfolio of the Jensen et al. (2023) 13 anomaly portfolios (red line). Panel B displays the 20-year rolling correlation between the two realized volatilities with a linear trendline. Realized volatility is useful because it indicates periods when risk premia are elevated. That is, when realized volatility of the aggregate market spikes upward, it is reasonable to assume that risk premia are elevated. If upward spikes in realized volatility of the market capture times when risk premia are elevated we should expect to see similar time-variation in the realized volatility of the anomalies portfolio. Surprisingly, we get a very different characterization of time-varying risk premia, with respect to the anomaly portfolios, beginning with the 2000 Tech Bubble (i.e. the correlation drops to 0.28 from 0.82 during this one period). While both panels indicate elevated risk premia during the first half (e.g. the 1929 Stock Market Crash and World War II), the set of anomalies suggest that the most recent periods of market turbulence (e.g. 2008 Financial Crisis) are pretty unexceptional historically in comparison to the 2000 Tech Bubble which generates realized volatility on the same order of magnitude as the 1929 Stock Market Crash and Great Depression. This leads to a significantly negative trendline in the correlation between the two time series of realized volatility. All together, the key takeaway is that not only are the anomalies telling us in their first moments (Figure 2) that they have lost their relevance since the 2000 Tech Bubble but we see the exact same story in terms of their second moments (Figure 3).

Ultimately, if return predictability exists, it is a direct reflection of changes in investors' subjective beliefs. So long as investor behavior has not changed over the past century we should expect predictable reactions (i.e. predictable changes in investors' subjective beliefs) to new and important sources of unforeseen events, manifested in large spikes in realized volatility. Hence, if the current set of test assets indeed spanned the SDF, then we should not see the significant departure in realized volatility between the market and the anomalies beginning with the 2000 Tech Bubble documented in Figure 3. This evidence suggest the current set of test assets does not adequately characterize risk premia (i.e. span the SDF) and it is incumbent on us to search for a new set of test assets.

Instead of using traditional test assets to span the SDF, we can use hedge fund portfolios.[1] Identifying expected returns indirectly via hedge funds can be viewed as analogous to something that is entirely unrelated to asset pricing that is often more sucessful than following a direct approach: looking for where the birds are to find fish. While this example may seem far afield from asset pricing, this observation offers us a useful analogy for identifying (unobservable) expected returns indirectly as opposed to the traditional method in empirical asset pricing of directly approximating expected returns with test assets based on firm characteristics (Fama and French, 1993). This indirect approach is implicitly based on the modus operandi of a hedge fund which is to harvest return predictability. Given that hedge funds are affected in a very real sense by their ability (or inability) to benefit from return predictability, they offer a promising alternative as test assets that span the SDF.

I propose the top three deciles of hedge fund returns (by month) as a parsimonious representation (i.e. factor model) of hedge fund test assets. I call this the implied upper bound of the SDF.[2] This upper bound can be viewed as analogous to the approach by Shiller (1981) in the following sense. Shiller (1981) constructed the implied price of the aggregate index based on perfect foresight of future dividends of the aggregate index. In our setting, the upper bound of the SDF can be viewed as equivalent to the perfect foresight SDF. The perfect foresight SDF is simply based on those factors that performed the best, ex-post, in any given period. While this implied upper bound of the SDF includes unpriced risk (i.e. noise), it avoids the issue of having to take a stance on the factors necessary to build expected returns. It is for this reason that this paper emphasizes the time-series dynamics of the implied upper bound which should mitigate the issue of unpriced risk as it is differenced out over time.

I argue that the implied upper bound of the SDF is useful for two main reasons. First, it provides an objective way of evaluating test assets with respect to whether they help us get closer to spanning the SDF. Second, it forces us to take more seriously the entire factor structure and it's dynamics. In doing so, it has the potential to add a more nuanced understanding to the strength of the factor structure and one of the most important economic

[1] This unconventional choice of test assets is in part motivated by the recent evidence of Koijen et al. (2023) that documents hedge funds have the largest impact on equity prices per dollar of wealth. I discuss the limitations of using hedge fund portfolios as test assets in Section 4.

[2] It is important to note that the implied upper bound of the SDF can be built using any set of test assets. Later in the paper I build an implied upper bound of the SDF based on a factor model comprised of the leading five principal components extracted from the Jensen et al. (2023) anomaly portfolios.

priors we have in asset pricing (i.e. the absence of near arbitrage opportunities).

In this paper, I apply the methodology of the implied upper bound to characterize expected stock returns. First, I determine what the implied upper bound of the SDF is based on the Jensen et al. (2023) factor theme anomaly portfolios and I examine it's factor structure and dynamics over the sample period. Second, I determine what the implied upper bound of the SDF is based on hedge fund decile portfolios and investigate it's factor structure and dynamics over it's respective sample period. This approach which takes the implied upper bound of the SDF seriously allows us to avoid the endless debate as to whether something is priced and instead focus our attention on how informative the test assets are.

In the cross-section of anomaly portfolios, I find that while the leading five principal components (PCs) are able to explain most of the variation in expected returns, the higher order PCs appear to also be a robust feature which is an important departure from the existing literature (see Kozak et al. (2018, 2020) for important studies who document the strong factor structure in the cross-section of stock returns). In a pseudo OOS analysis the PC1-5 maximum squared SR is 36% (24%) of the anomaly test asset's maximum squared SR during the IS (OOS) period. This is evidence that the PC factor model gets closer to the mean-variance frontier of anomaly portfolios during the IS period compared to the OOS period. In other words, the higher order PCs that are commonly excluded from factor models are a robust feature in explaining stock returns during both the IS and OOS period. This is in sharp contrast to previous studies that suggest the higher order PCs are no longer robust during the OOS period. The second interesting finding regarding the cross-section of anomaly portfolios is that the factor structure is largely dominated by only three of the 13 factor themes (Value, Low Leverage and Seasonality explain nearly half of the SDF's variance) with the Value Factor Theme contributing 23% to the SDF's variance on it's own. Importantly, this new evidence tells us that many well-known factors (e.g. Investment, Profitability, etc.) play a much more limited role in contributing to the maximum squared SR of the economy.

The time-series of the SDF's variance (based on anomaly portfolios) displays interesting dynamics. First, the variance of the SDF spikes upward during market dislocations (e.g. 2008 Financial Crisis).[3] The maximum squared SR has significant variation ranging between 1.67 to 46.62. This evidence is consistent with the original motivation in Figure 3 regarding elevated levels of risk premia during periods of market turbulence. The second finding which

[3] I refer to market dislocations as any period that is characterized by market turbulence (i.e. an upward spike in realized volatility of the aggregate market index) in which market prices deviate more from their fundamental values. The most recent market dislocations include the onset of the Russia-Ukraine War in early 2022, the COVID-19 Stock Market Crash in March 2020, the 2008 Financial Crisis and the 2000 Technology Bubble.

is perhaps more suprising is that the SDF's variance based on the factor structure is notably different than that based on the underlying test assets. In particular, the factor structure is largely a product of one single market dislocation: the 2000 Technology Bubble. More specifically, the maximum squared SR during the Tech Bubble is twice as large as that experienced during the onset of the Russia-Ukraine war in early 2022 and nearly three times as large as that experienced during the 2008 Financial Crisis. We learn from this evidence that our understanding with respect to the factor structure of stock returns based on the anomaly portfolios is in large part shaped by the return variation from two decades ago during the 2000 Tech Bubble. Put differently, the more recent periods of market turbulence (e.g. 2008 Financial Crisis and 2020 COVID-19 Crash) that created historic investment opportunities are significantly under-represented by the test assets that we often assume characterize investment opportunities.

Importantly, the evidence in this paper regarding the disproportionate impact of the 2000 Tech Bubble is consistent with both Jensen et al. (2023) and Chinco et al. (2021). In particular, one of the most salient features in Figure 2 of Jensen et al. (2023), documenting the anomalies portfolio cumulative CAPM alpha, is the significant outperformance relative to the market around the Tech Bubble. Roughly one third of the gains over their sample period from 1990 to 2020 are due to this market dislocation. The outsized influence on CAPM alphas from this event is entirely consistent with my findings regarding it's dominance in the factor structure of the anomaly portfolios. Moreover, Figure 1 of Chinco et al. (2021) displays the evolution of the prior variance (i.e. the ex-ante probability of witnessing a tradable anomaly dubbed the "anomaly base rate") over their sample period from 1978 to 2015 which *peaks* during the Tech Bubble. This evidence further reinforces the central message that our understanding of the cross-section of stock returns is shaped by this one period. A factor structure that is dominated by the Tech Bubble from two decades ago (and is to a large extent unaffected by the more recent bouts of market stress including the 2008 Financial Crisis and COVID-19 Crash in March 2020) illustrates that the anomaly portfolios have lost their relevance as test assets. All together, the skewed time-series and cross-sectional evidence of the SDF's Variance indicates that the anomaly test assets do not span the SDF.

When we instead use hedge fund portfolios as test assets we end up getting a very different picture of the SDF's variance both cross-sectionally and in the time-series. With respect to the cross-section of hedge fund portfolios, there is evidence of persistence in the top decile hedge funds whereby the OOS SR is 5.43 compared to an IS SR of 5.27, in sharp contrast to the anomaly test assets whose superior IS risk-adjusted returns disappear during the OOS period. This evidence documents that near arbitrage opportunities are persistent in contrast to Kozak et al. (2018). That is, we can no longer conclude that near arbitrage opportunities do not exist, since we have a set of test assets that show that they are robust during both the IS and OOS periods. It turns out that the evidence informing our economic prior regarding the absence of near arbitrage opportunities, is a product of the fact that the anomaly portfolios no longer span the SDF. Moreover, while the OOS maximum squared SR of the anomaly test assets is one fifth the size of it's IS counterpart, the hedge fund portfolio maximum squared SR more than doubles in the OOS period. Most importantly, using hedge fund portfolios as test assets gets us closer to spanning the SDF by producing a SR that is 3.4x higher than the traditional anomaly portfolios.

Consistent with anomaly portfolios as test assets, the time-variation of the SDF based on hedge fund portfolios also spikes upward during market dislocations. The key difference is that the Tech Bubble no longer dominates the factor structure of hedge fund portfolios. That is, the factor structure changes significantly across all of the most recent market dislocations. With that evidence we learn that near arbitrage opportunities occur from a more diverse set of periods characterized by market turbulence which is again consistent with the initial motivating piece of evidence in Figures 2 and 3 regarding time-varying risk premia. More specifically, time-varying risk premia reflect changes in investors' subjective beliefs, which are largely affected by new and important sources of exogenous events. Figure 3 makes clear that we have witnessed many of these events since the 2000 Tech Bubble that the anomaly portfolios fail to adequately capture in comparison to hedge fund portfolios. In essence, hedge fund portfolios better characterize the investment opportunity set relative to the existing set of anomaly test assets.

Lastly, I address the robustness concerns that arise from building managed portfolios, that serve as test assets, from hedge funds. Importantly, I show that the results based on hedge fund portfolios as a new class of test assets are qualitatively unchanged after accounting for leverage, inability to short-sell hedge funds, capacity limits, out-of-sample (OOS) implementability and restricting hedge fund strategies to those that only include individual stocks. After addressing all of these concerns jointly, I find that test assets based on hedge fund returns produce high risk-adjusted returns that persist over the entire sample period and reflect unique variation that contrasts sharply with the traditional anomaly portfolios.

This paper contributes to two strands of the literature. First, I contribute to the active literature that studies the factor structure of the cross-section of expected stock returns (see Bryzgalova et al. (2023a,b); Daniel et al. (2020); Feng et al. (2020); Giglio et al. (2023); Kelly et al. (2019); Kozak and Nagel (2024); Kozak et al. (2018, 2020); Lettau and Pelger (2020) for recent and important examples). I focus on two important papers which are most closely related to this study that have shaped how we think in empirical asset pricing. In a seminal paper, Kozak et al. (2018) provides robust evidence that we can use reduced form models from PCs of the covariance matrix of returns given their superior ability to capture the cross-sectional variation within stock returns. Moreover, they derive a closed-form expression (see Equation 4 of their paper) that shows the variance of the SDF is a function of the factor structure. All together, their empirical evidence coupled with their analytical result leads to the conclusion that near arbitrage opportunities do not exist in the cross-section of stock returns. In a subsequent and equally influential paper, Kozak et al. (2020) use a Bayesian approach to build an SDF from a larger set of 50 L/S anomaly portfolios. Given their earlier work, they adopt the novel prior that first and second moments should be linked and in fact find that the leading PCs do a good job explaining this higher-dimensional setting of anomalies. I build on their work by showing that while the larger set of anomalies also exhibits a strong factor structure over the full sample period (i.e. unconditionally), the factor structure of anomaly portfolios is highly *unstable* over time. Moreover, I show that their conclusion with respect to ignoring the lower ranked PCs is no longer warranted (during market dislocations) when we instead use hedge fund portfolios as test assets. Hence, this new evidence motivates us to adopt a more nuanced prior regarding the cross-section of stock returns. More specifically, we need to relax the strong prior regarding the absence of near arbitrage opportunities during periods of market stress.

Second, I contribute to the literature that builds (or chooses) test assets which help characterize the SDF (see Ahn et al. (2009); Bryzgalova et al. (2023c); Giglio et al. (2023) for important recent examples). In a pathbreaking paper, Bryzgalova et al. (2023c) revisits the important question as to whether we have identified portfolios that characterize the investment opportunity set (test assets) and offers an entirely novel approach called "Asset Pricing Trees (AP-Trees)" that can accomodate complex interactions and higher dimensionality among groups of stocks that conventional portfolio sorts are unable to incorporate. The AP-Trees method achieves a SR that is three times higher OOS compared to simple L/S factors which indicates that their approach gets us much closer to spanning the SDF compared to the traditional anomaly portfolios. I complement their study by documenting novel evidence that the SDF is not spanned by traditional test assets, thereby lending further justification regarding the introduction of alternative methods to construct test assets. In addition, I contribute to this literature by introducing an alternative approach to generating test assets (i.e. hedge fund portfolios) that are able to generate a SR that is 3.4x higher than the traditional anomaly portfolios.

The article proceeds as follows. Section 2 details the methodology and data used in the paper. The main results are presented in Section 3. Section 4 documents the findings regarding the new set of hedge fund test assets. Section 5 concludes.

2 Methodology, Data and Summary Statistics

This section details the construction of the SDF's upper bound, the data used in this paper and descriptive statistics which provide an overview of the data. The methodology follows Kozak et al. (2018) with respect to it's application to anomaly portfolio excess returns. The main departure is with respect to applying Equation 1 to hedge fund portfolios.

I estimate the minimum-variance SDF as in Hansen and Jagannathan (1991),

$$M = 1 - \mu' \Gamma^{-1} (R - \mu) \quad \Rightarrow \quad \mathbb{E}[M] = 1 \;\text{ and }\; \text{Var}[M] = \mu' \Gamma^{-1} \mu, \tag{1}$$

where the mean of portfolio excess returns is represented by $\mu = \mathbb{E}[R]$ and Γ represents the variance-covariance matrix of portfolio excess returns.

2.1 The Upper Bound of the SDF

One of the most important results in asset pricing is from Hansen and Jagannathan (1991) who derive the famous Hansen-Jagannathan Bound (HJB),

$$\frac{\sigma[M]}{\mathbb{E}[M]} \geq \text{Maximum SR in the Economy} \tag{2}$$

That is, with the SDF normalized to have mean equal to one, the variance of the SDF is equal to the maximum squared SR of the test assets.

Kozak et al. (2018) note how a factor can be both important in explaining the variance of returns and unimportant in explaining expected returns (i.e. the pricing performance of the factor). The pricing performance of a given factor depends on the covariance of a test assets' returns with a given factor. In essence, there is an important distinction between explaining return variances versus explaining expected returns. Said differently, there is some return variation that is not priced and should be viewed as noise (Daniel et al., 2020).

If the set of test assets does not span the SDF, these covariances are less informative. Most importantly, expected returns are inherently unobservable and depend crucially on the set of test assets. It is for this very reason that I rely instead on a factor's ability to explain return variances which avoids the empirical issue as to which part of return variances concerns expected returns. Hence, I call the implied upper bound of the SDF the reduced-form factor model implied by a set of test assets (e.g. the implied upper bound of the SDF can be the PC1-5 factor model whereby the leading five PCs are extracted from the variance-covariance matrix of anomaly excess returns). That is, the implied upper bound of the SDF is based on maximizing the explained variance from a set of test assets,

$$M^* \approx 1 - f'\Sigma^{-1}(F - f), \tag{3}$$

where F are the excess returns from the leading K principal components with $f = \mathbb{E}F_k$.

By taking the upper bound of the SDF (and it's time series dynamics) seriously, I am making the following implicit assumption: the proportion of the SDF's variance that is simply noise is relatively constant over time (i.e. homoskedastic). This assumption seems relatively mild since the proportion of the SDF's variance attributable to risk premia should be time-varying since it reflects some combination of investors' risk preferences and belief distortions. Hence, the time-varying nature of the SDF's variance should reflect time-variation in risk premia.

With the upper bound of the SDF defined, I now delineate how I estimate factor models with anomaly portfolios versus hedge fund portfolios.

2.2 Anomaly Portfolios

I begin with a large, comprehensive and widely accepted set of anomaly factor excess returns (i.e. test assets). Equation 1 is first applied to the original set of 13 factor themes which produces the SDF based on the test assets. As in Kozak et al. (2018), I also conduct Principal Component Analysis (PCA) to the anomaly portfolios to produce factor models ranging from the first leading PC to a five factor model comprised of the leading five PCs. With that, I estimate the factor-based SDF as before with the anomaly portfolios substituted for the PC factors.

2.3 Hedge Fund Portfolios

The standard approach to forming decile portfolio returns is to generate the decile portfolios over the entire sample. By computing the deciles cross-sectionally, it avoids the issue that explaining the set of test assets' returns is tautological. That is, we are explaining expected returns cross-sectionally, by construction. However, it is not obvious whether the cross-sectional explanatory power will have the same power along the time-series dimension. Given the paper's scope of interest (i.e. to estimate the upper bound of the SDF) I deviate from the convention by computing deciles each month which maximizes the ability to explain the set of hedge fund returns both cross-sectionally and in the time-series:[4]

[4] I choose to form decile portfolios (over quintile or tertile portfolios) to reflect the minority of hedge fund managers with skill.

$$D_i = \begin{cases} 1 & \text{if hedge fund return is in decile } i = 1 \\ 2 & \text{if hedge fund return is in decile } i = 2 \\ \vdots & \\ 10 & \text{if hedge fund return is in decile } i = 10 \end{cases} \tag{4}$$

With that, the implied upper bound based on this new set of test assets (hedge fund decile portfolios), I compute the implied upper bound in the following way,

$$M^* \approx 1 - d'\Sigma^{-1}(D - d), \tag{5}$$

where D are the excess returns excess returns from the top three deciles (D_8, D_9, and D_{10}) with $d = \mathbb{E}D_i$.

The decile by month approach is highly related to conducting PCA which maximizes the variance of the original set of test assets for a given number of PC factors. More specifically, PCA involves an eigenvalue decomposition of the variance-covariance matrix of portfolio excess returns which yields N orthogonal factors. Subsequent to this, the rotated test assets are then sorted from highest to lowest based on their eigenvalue (variance). Given the relatively short time series of hedge fund data relative to the number of hedge funds ($T << N$), the covariance matrix is unstable which is why we cannot apply PCA to the space of hedge fund returns. It is for this reason that I apply a similar method to PCA that is feasible (i.e. decile by month hedge fund portfolios) to maximize the variance explained both cross-sectionally and in the time-series of hedge fund returns.

2.4 Data

I consider a large publicly-available data set of factor excess returns from Jensen et al. (2023) ("JKP") who construct 153 factors spanning 93 countries from January 1926 to December 2023.[5] I have chosen to restrict the factor themes to the U.S. to make it more comparable to previous studies that have studied the factor structure within the cross-section of stock returns (Kozak et al., 2018, 2020). In Section 3, the time series is restricted to start from January 1952 to ensure that all 13 factor themes are represented in the SDF.

In addition to the large set of anomaly returns, I use individual monthly hedge fund excess returns from the combination of the two most widely used databases: (i) Lipper Trading Advisor Selection System Hedge Fund Commercial Database (hereafter TASS); and

[5] I have downloaded the daily U.S. capped value weighted factors across all 13 themes from https://jkpfactors.com/factor-returns.

(ii) Hedge Fund Research (HFR). The decision to use these two datasets is based on the findings from Joenväärä et al. (2021) who highlight how TASS and HFR are two of the highest research quality hedge fund databases. In particular, they note that TASS is the most widely used database in the hedge fund literature due to it's high quality earlier in the sample. However, it's coverage has deteriorated more recently. To compensate for this, I also rely on HFR which they highlight has a consistently high coverage of hedge fund data. The sample period is from January 1996 to June 2023 comprised of 315,646 hedge fund month observations. The TASS data represents the most widely studied hedge fund database (Joenväärä et al., 2021). I have cleaned the TASS data to address the known biases present in hedge fund commercial databases (Pedersen, 2015).[6] Hedge fund returns are net of fees and are in excess of the monthly risk-free rate.

Lastly, I have downloaded the 25 Size-B/M portfolios and Fama and French (1992) MKT, SMB and HML ("FF3") factors at a daily frequency from Ken French's website.[7] In addition, I have downloaded the daily Daniel et al. (2020) ("DMRS") five efficient factors that hedge out unpriced risk with respect to the original Fama and French (2015) ("FF5") from Kent Daniel's website.[8] The DMRS factor sample period begins in July 1963 and ends in March 2023.

2.5 Summary Statistics

Table 1 presents the descriptive statistics (mean, volatility, SR and skewness) of the two main datasets: L/S anomaly portfolios (Panel A) and hedge fund decile portfolios (Panel B). The mean and volatility of returns are expressed in percent and have been annualized.

The average anomaly returns range from 0.25% (Low Leverage) to 4.30% (Momentum) per annum. The spread in factor theme volatilities is consideraly larger ranging from 1.72% (Seasonality) to 10.98% (Low Risk) per annum. It is somewhat surprising that the low risk anomaly has the highest volatility of the 13 factor themes. The average SR is 0.57 with a minimum and maximum of 0.03 (Low Leverage) and 1.19 (Debt Issuance). The variation in SRs is primarily driven by the variation in volatility across factor themes. Ten of the 13 Factor themes have mild negative skewness with the remaining three (Profitability, Short Term Reversal and Value) having more significant positive skewness between 0.56 to 0.94.

[6] I have used the 3-step unsmoothing methodology proposed by Couts et al. (2024) to address the serial correlation present at both the strategy- and individual-level of hedge fund returns. The cleaned data has removed hedge funds whose primary strategy is labelled "Fund of Funds" given the high degree of autocorrelation in their returns. I have also removed any hedge funds whose primary strategy is labelled "Other" or "Undefined" consistent with Couts et al. (2024).

[7] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[8] http://www.kentdaniel.net/data.php

The average annualized hedge fund decile excess returns range from -92.9% to 107.0%, with an average of 6.1%. The SRs (Skewness) range between -3.91 (-3.18) to 5.18 (2.01), with an average SR (Skewness) of 0.53 (-0.73). Interestingly, the average SR is roughly the same for both the anomaly portfolios and hedge fund decile portfolios (0.57 vs. 0.53, respectively). In contrast, the average hedge fund portfolio is more negatively skewed compared to the average anomaly portfolio (-0.73 vs. -0.18, respectively). This is unsurprising given that the anomaly portfolios are highly diversified both within and across factor themes, whereas hedge fund returns are notorious for their more concentrated market downside risks (Kelly and Jiang, 2012) which is evident in the bottom six deciles.

3 Main Results

In this section, I start by showing that the more comprehensive set (cross-sectionally and in the time-series) of anomaly returns from Jensen et al. (2023) ends up following a strong factor structure over the full sample period from January 1952 to December 2023. After confirming this well-established stylized fact in the empirical asset pricing literature, I document a new property of the factor structure. That is, the factor structure of anomaly returns is *highly time-varying* and has particularly interesting dynamics around the Technology Bubble.

3.1 Establishing a Baseline: Revisiting Kozak et al. (2018)

Table 2 shows that the first five PCs are able to explain most of the cross-sectional variation in expected returns of the anomaly portfolios. Table 2 revisits the analysis conducted in Table II of Kozak et al. (2018) with a set of anomaly portfolios that has been extended both in the time series (1952 to 2023 versus 1966 to 2015) and the cross-section (Jensen et al. (2023) 153 characteristics clustered into 13 factor themes versus 15 L/S anomaly strategies from Novy-Marx and Velikov (2016)). In keeping with Kozak et al. (2018), for the purposes of comparison, we evaluate the anomaly returns alongside the Fama and French (1993) 25 Size-B/M portfolios. The upper panel displays the annualized mean returns of the factor themes along with their alphas. The alphas are estimated from time series regressions of the L/S anomaly return on the leading PCs extracted from the underlying anomaly test assets. The last five columns correspond to successively larger factor models beginning with one PC (PC1) and ending with five PCs (PC1-5). The results are consistent with Kozak et al. (2018) findings that the first PC is inadequate in explaining anomaly returns (e.g. momentum which has the largest mean return of 4.30% has an alpha of 4.2% per year). With five PCs the pricing error for momentum has shrunk to 0.1% per year.

The bottom panel of Table 2 displays the (ex-post) maximum squared SR of the factor theme portfolios (9.31) and the respective PC factors. It is interesting to note how the maximum squared SR of this updated set of anomalies is more than twice the size as the value reported in Kozak et al. (2018) of 4.23. This simply reflects the more volatile period in equity markets since their sample ended in 2015, beginning with the COVID-19 pandemic. Aside from the more volatile SDF in this updated sample, the results are consistent with Kozak et al. (2018): the maximum squared SR of the PC factors is well below that of the test assets (PC1-5 model has a maximum squared SR of 1.76). All five of the factor models reject the null hypothesis of no pricing errors at the 1% level of statistical significance. In spite of the large HJ-distance, the PC factor model outperforms the traditional Fama and French (1993) FF3 factor model (maximum squared SR of 0.71). Interestingly, it is not able to beat the DMRS factor model (maximum squared SR of 1.99) based on the anomaly set of test assets which helps showcase the importance of removing unpriced risk present in test assets.[9] The best performing model is the PC1-5 model based on the Fama and French (1993) 25 Size-B/M test assets that has a maximum squared SR of 2.07. All together, this updated evidence lends further support to the results in Kozak et al. (2018) that we can successfully build reduced-form factor models from PCA.

The last part of this subsection provides the final piece of evidence that confirms the baseline result first established by Kozak et al. (2018): the large HJ-distance observed in the bottom panel of Table 2 does not reflect persistent near-arbitrage opportunities. Figure 4 conducts the same pseudo OOS analysis from Figure 4 of Kozak et al. (2018) with our updated sample. That is, I divide the sample into two equal halves with the first (second) half treated as the IS (OOS) period. The two panels reflect scatterplots of the IS versus OOS SRs of the anomaly portfolios (Panel A) and the 25 Size-B/M Portfolios (Panel B). The OOS period begins December 31, 1987. Consistent with the findings in Kozak et al. (2018), near arbitrage opportunities are *not* persistent over time based on this analysis (i.e. nearly all of the test assets in both panels are well below the 45 degree line). The OOS SRs

[9] In a recent paper, Kozak and Nagel (2024) derive the conditions under which traditional portfolio sorts (e.g. Fama and French (1993)) span the SDF assuming that conditional expected stock returns are linear in firm characteristics. They find that unpriced risk must be orthogonal to the priced systematic risk for traditional portfolio sorts to span the SDF. That is, the DMRS factors that remove unpriced risk are more likely to span the SDF so long as stock returns are linear in firm characteristics. Table 2 shows that the maximum squared SR of the DMRS factors is roughly one fifth the size of the maximum squared SR generated from the Jensen et al. (2023) set of anomalies. Hence, this evidence suggests that the assumption that stock returns are linear in firm characteristics is likely not satisfied in the data. Therefore, we will not span the SDF by removing unpriced risk (via Daniel et al. (2020) or iterated hedging (Kozak and Nagel, 2024)) from a larger set of anomalies without accounting for important nonlinear dependencies among firm characteristics. The nonlinear dependencies that drive the cross-sectional variation of stock returns are more likely to be accounted for by the AP-Trees method of Bryzgalova et al. (2023c) or alternative methods to build test assets such as the hedge fund decile portfolios approach proposed in this paper.

are, on average, 41% (61%) as large as the IS SRs for the anomaly (25 Size-B/M) portfolios. This poor OOS evidence is consistent with the data mining concerns regarding anomalies (McLean and Pontiff, 2016). With that (in addition to further robustness checks), Kozak et al. (2018) reach the conclusion that the higher order PCs contributon to the higher IS SR is not a robust feature of expected returns. In essence, expected returns (i.e. the SDF) can be nicely characterized by a reduced-form factor model with a handful of the leading PCs (e.g. a five factor model).

Before reaching the conclusion that the lower ranked PCs are of no consequence to the cross-section of expected stock returns (i.e. near arbitrage opportunities do not exist), it seems like the most natural exercise would be to conduct the analysis in the bottom panel of Table 2 for the two halves of the sample (IS vs. OOS) and determine if we get a similarly large HJ-distance for the IS and OOS periods. Table 3 performs this analysis. The upper (bottom) panel is based on estimating the SDF during the IS (OOS) period for both the anomaly and the 25 Size-B/M portfolios. There are two important takeaways from this exercise. First, the maximum squared SR is notably higher in the IS period compared to the OOS period for both sets of test assets. However, the second and more interesting finding is that the HJ-distance is relatively larger during the OOS period compared to the IS period for *both* sets of test assets. A factor model with the five leading PCs gets closer to the mean-variance frontier during the IS period compared to the OOS period. More specifically, the PC1-5 maximum squared SR is 36% (24%) of the anomaly test assets' maximum squared SR during the IS (OOS) period. Moreover, the PC1-5 maximum squared SR is 58% (24%) of the 25 Size-B/M test assets' maximum squared SR during the IS (OOS) period. By performing this subsample analysis with respect to the maximum squared SR, I reach a different conclusion to Kozak et al. (2018): The higher order PCs appear to be a robust feature of the cross-section of expected stock returns (i.e. the higher order PCs appear to play an even larger role in the OOS period compared to the IS period).

After reporting the surprising findings of Table 3 that at first glance seem contradictory to the results in Figure 4, two natural questions arise: (i) What does the SDF's variance (i.e. maximum squared SR) look like over time?; and (ii) What does the SDF's factor structure look like over time? The next set of empirical results are guided by these two questions which will help reconcile the contradictory findings between a lower OOS SR (compared to IS SR) and a weaker OOS factor structure (compared to the IS factor structure). Answers to these questions will shed more light on this finding regarding the factor structure and in doing so, will provide new insights on the nature of the underlying test assets.

3.2 On the Dynamics of the SDF's Variance

Figure 5 displays the one-year rolling variance of the SDF based on the JKP anomaly portfolios (blue line) and the five leading PCs extracted from the test assets (red line). The variance is estimated over the full sample period from January 1952 to December 2023. There are two key takeaways from this figure. First, the SDF (based on the underlying test assets) reveals an important property that is intuitive from an asset pricing perspective: The SDF's variance is highly time-varying and spikes upward during market dislocations (e.g. the most recent episodes include the 2000 Technology Bubble, 2008 Financial Crisis, 2020 COVID-19 Stock Market Crash and 2022 Russia-Ukraine War). This is intuitive since we should expect the maximum squared SR to be larger when risk premia are larger. By definition, risk premia are larger during market dislocations (i.e when market prices deviate more from their fair values). The maximum squared SR experiences a wide range from 1.67 (February 1978) to 46.62 (April 2023). The ranking (with respect to the maximum squared SRs) of the most recent market dislocations from highest to lowest is: (i) 2022 Russia-Ukraine War (46.62), (ii) 2008 Financial Crisis (46.33), (iii) 2000 Technology Bubble (39.62); and (iv) 2020 COVID-19 Stock Market Crash (30.14). This exhibit showcases the heightened level of stock market volatility over this more recent period.

The second finding from Figure 5 illustrates the notably different dynamics of the SDF based on the PC1-5 factor model.[10] While this SDF is also time-varying,it peaks during the Technology Bubble of 2000 (one-year variance estimate reaches a maximum value of 20.38 in April 2001) and reaches a minimum value of 0.27 (December 1985). Most notably, the Tech Bubble plays a disproportionate role in characterizing the SDF's variance relative to any other market dislocation. The variance estimate during the Tech Bubble is roughly twice as large as that experienced during the 2022 Russia-Ukraine War and nearly three times as large as that experienced during the 2008 Financial Crisis. In essence, the factor structure (and it's properties) is in large part a product of the stock market dynamics during the Tech Bubble.

To help illustrate the dynamics of the entire factor structure over the sample period, Figure 6 displays a heatmap of the scree plot over time whereby PCA is estimated via a rolling window of the past 252 trading days. The warmer shades (yellow and orange) represent larger eigenvalues (stronger factor structure) whereas the cooler shades (blues and purples) represent smaller eigenvalues (weaker factor structure). This figure confirms the second finding of Figure 5. That is, the Tech Bubble is the brightest spot on this heat map. Interestingly, the heat map reveals that all of the eigenvalues shift upward indicated

[10] The correlation with the SDF (based on test assets) is 0.75 leaving a significant amount unexplained by the reduced form factor model.

by the warmer shades along the entire vertical axis. This happens to a lesser degree during the other market dislocations. All together, the factor structure is more *unstable* during market dislocations which coincides with the periods that experience the upward spikes in the variance of the SDF.

Figure 7 lends further support to the disproportionate impact of the Tech Bubble on the anomaly test assets' factor structure. This figure plots the first difference between the adjacent five leading PCs. For instance, the blue line represents PC1-$2_t = \Delta \text{PC}_{ij,t} = \text{PC}_{1,t} - \text{PC}_{2,t}$, where i, j represent adjacent PCs and $i < j$. In essence this figure zooms in on how the factor structure is changing among the leading PCs, thereby highlighting how the Tech Bubble is the dominant event in the dynamics of the leading PCs. This evidence is consistent with Chinco et al. (2021); Jensen et al. (2023). In particular, Figure 2 of Jensen et al. (2023) shows the cumulative CAPM alpha of an average of this set of anomaly portfolios. The most salient feature of this plot is the significant outperformance relative to the market around the Tech Bubble. Roughly one third of the gains over this period are due to this market dislocation. Moreover, Figure 1 of Chinco et al. (2021) displays the evolution of the prior variance over time (i.e. the ex-ante probability of witnessing a tradable anomaly dubbed the "anomaly base rate"). This anomaly base rate peaks during the Tech Bubble. This is entirely consistent with the evidence presented in this section showing the SDF's variance (based on the PC1-5 factor model) peaking during the Tech Bubble.

Lettau and Pelger (2020), who improve on PCA by adding regularization with respect to estimates of the test assets risk premia ("RP-PCA"), highlight the instability of the factor structure with respect to individual stocks. Moreover, they document the high generalized correlations (in excess of 0.9) of the leading PCA factors and conclude that the factor structure (based on portfolios) is relatively stable over time. However, the evidence presented in Figures 6 and 7 make clear how the factor structure is highly time-varying (i.e. experiences periods of instability) based on this set of anomaly portfolios. Moreover, this new evidence is also consistent with Kozak et al. (2018) Equation 4 which decomposes the SDF's variance providing an important insight regarding how the SDF's Variance is a function of the strength of the factor structure (i.e. a weaker factor structure coupled with a large spread in average returns implies a higher Variance of the SDF).

This section began by highlighting the seemingly contradictory evidence between both a relatively weaker factor structure and lower SR in the second half of the sample compared to the first half ("OOS period" versus "IS period," respectively). We can reconcile this evidence in light of the discrepancy in the SDF's variance estimated via test assets versus PC factors in Figure 5. That is, the time-variation in the PC1-5 SDF Variance is significantly different in the first half (IS) of the sample period in comparison to the latter half (OOS) of the sample. In particular, the average ten-year rolling correlation between the two estimates of the SDF Variance is 0.22 in the IS period compared to 0.74 during the OOS period. This evidence coupled with the insight of Kozak et al. (2018) Equation 4 tells us that we should expect higher SRs when the cross-sectional variation in average returns does not coincide with the leading PCs. In spite of the higher degree of comovement, the proportion of the SDF's variance explained by the PC factor model is lower in the OOS period as evidenced by Table 3.

All together, this section has documented the time-varying dynamics of the SDF's variance (based on the test assets) which spikes upward during market dislocations. More importantly, it is now clear how the SDF's variance (based on the factor structure of the underlying anomaly portfolios) displays notably different dynamics whereby the Tech Bubble has a disproportionate impact on the SDF's variance. Hence, we have a more nuanced understanding as to what time period is driving the dynamics of the SDF's variance. With that, the next logical step is to reveal how the underlying L/S anomaly portfolios contribute to the SDF's variance both cross-sectionally and in the time-series.

3.3 On the Composition of the SDF's Variance

Figure 8 displays the percent contribution of each respective factor theme to the SDF's Variance. More specifically, each factor theme is calculated in the following way,

$$\text{\% Contribution to SDF's Variance} = 1 - \frac{\text{Max SR}^2_{\text{w/o factor}}}{\text{Max SR}^2_{\text{w/ all factors}}} \tag{6}$$

By leaving the factor theme out one at a time and re-estimating the variance of the SDF, Figure 8 captures the cross-sectional impact of each factor theme on the maximum squared SR. It turns out that the Value Factor Theme plays a disproportionate role in the maximum squared SR (22.9% contribution to the SDF's variance). Moreover, three out of the 13 factor themes (Value, Low Leverage and Seasonality) explain nearly half of the variance 47.5%. Perhaps surprisingly, Investment, Profitability and Size have almost no influence on the SDF's variance (i.e. a combined contribution of 0.2%). It is important to note that this analysis by construction does not take into account any interaction effects between the factor themes (the interactions account for 18.2% of the SDF's variance). All together, this evidence highlights how the SDF's Variance is in large part a product of only a handful of the factor themes.

After having established the cross-sectional impact of each of the factor themes on the SDF's variance, I now display how each of the factor theme's percent contribution to the

SDF varies over time. Figure 9 displays the rolling one year variance of the difference between the SDF with all factors and the SDF with the respective factor removed. Given the outsized role of Value on the SDF's Variance, the Value Factor Theme's dynamics are highly similar to the SDF's variance based on the test assets with a similar ranking in peaks of the variance, with respect to market dislocations. In particular, value has the largest effect on the SDF's variance during the 2008 Financial Crisis. Most importantly, this figure provides a window into the underlying reason why the Tech Bubble has such a outsized role on the maximum squared SR. Six of the 13 factor themes (Debt Issuance, Low Leverage, Momentum, Profit Growth, Short-Term Reversal and Size) witness their peak contribution to the SDF's variance during the Tech Bubble. Of the remaining seven factor themes, the Tech Bubble is either second or third in relative peaks of it's contribution to the SDF's variance. It is this commonality across factor themes during a turbulent period in the stock market that results in the Tech Bubble playing such a large role in the factor structure of the anomaly test assets.

The evidence presented in this section highlights the need for a new set (and perhaps approach) to test assets if we are going to adequately characterize the cross-sectional variation in expected stock returns. In particular, the current set of test assets is dominated by one period (Tech Bubble) and one factor (Value). The next section takes up this challenge to get us one step closer to spanning the SDF by proposing hedge fund portfolios as a new set of test assets.

4 Expanding the Scope of Test Assets

Identifying test assets that help explain the cross-section of expected stock returns that remain robust OOS is a challenging enterprise. One of the main ideas of this paper is that we can potentially get closer to spanning the SDF if we identify expected returns *indirectly*. That is, if we are going to catch fish (i.e. expected returns) we shouldn't look for fish but instead we should look for the birds. In our setting, those birds are hedge funds. This section conducts the same empirical analysis as Section 3 with respect to hedge fund (in comparison to the Jensen et al. (2023) anomaly) portfolios as test assets.

4.1 Underlying Motivation to Use Hedge Fund Portfolios as Test Assets

In principle, hedge funds can be viewed as machines that harvest return predictability from markets. Given that this is arguably the modus operandi of any hedge fund it seems only natural that we might be more successful in identifying the maximum squared SR (i.e. spanning the SDF) among a group of managed portfolios comprised of hedge funds. The main obstacle that we face is that we don't get to observe the complete universe of existing hedge funds. Instead, we rely on a subset of the hedge fund universe that is reported in hedge fund commercial databases.

It is strictly voluntary whether or not a hedge fund chooses to report to a commercial database. The primary motivation for hedge funds to report (i.e. market) their returns to these databases is to raise more capital. Hence, the hedge funds with the best and worst performance will likely not be reflected in these databases. In fact, Barth et al. (2023); Brown et al. (2024) note how most large hedge funds with assets greater than US $ 1 billion do not report to commercial databases (i.e. in excess of US$ 2 trillion of combined AUM). Most importantly, the unobserved right tail of the distribution of hedge funds has superior performance (i.e. higher alphas) to those reported in commercial databases. In light of this fact, I aim to proxy for the very best performing hedge funds by constructing an *upper bound* of hedge fund returns by combining the top three deciles each month of hedge fund returns reported in TASS. This group of hedge fund returns corresponds to average SRs ranging from 2.48 to 5.18 (see Table 1).

4.2 On the Robustness of Hedge Fund Portfolios

In general, there are several issues with using hedge fund portfolios as test assets in empirical asset pricing. First, unlike asset pricing factors, hedge funds cannot be shorted. Second, hedge fund returns often involve leverage. Third, hedge funds have limited capacity and as a group are too small relative to the broader stock market. Fourth, hedge fund skill is in short supply and as a result it is difficult to achieve meaninful separation in hedge fund returns both ex-ante and ex-post. The scope of interest for this study is primarily focused on documenting the dynamics of the SDF's variance. While these limitations almost surely affect the level of the SDF's variance (maximum squared SR), they should not meaningfully alter the dynamics of the SDF's variance identified in this section. The only issue that poses a potential challenge to uncovering the SDF's dynamics concerns the issue of separation among hedge fund managers with respect to skill, which I turn to next.

It has been well documented in the hedge fund literature that a minority of hedge fund managers possess skill. Chen et al. (2017) estimates that 9% of hedge funds have superior performance which is consistent with Giglio et al. (2021) who also note the difficulty in predicting hedge fund performance given the small minority of hedge fund managers who possess skill. With that, it rules out the natural choice to form hedge fund test assets by hedge fund strategy since there are too few hedge fund managers within a given strategy that are able to achieve superior returns (i.e. hedge fund strategies produce similar returns). By construction, the methodology I propose (hedge fund deciles by month) addresses the lack of separation problem. Ultimately, the top decile hedge fund returns are intended to capture the most successful group of unobserved hedge funds that do not report to commercial databases (consistent with the findings of Barth et al. (2023); Brown et al. (2024)).

In a recent study, Koijen et al. (2023) note how hedge funds are the most price-sensitive (i.e. elastic) among institutional investors. In particular, they estimate the impact of hedge funds on stocks to be \$3.58 per dollar of wealth. In short, they conclude that hedge funds have the greatest impact on equity prices per dollar of wealth. This new evidence lends support to further investigating hedge funds as a potentially new set of test assets in spite of the limitations mentioned in the previous paragraph. Moreover, the scope of this study is in constructing an upper bound of the SDF. Hence, this study explicitly acknowldeges that the use of these new test assets is to construct an *upper bound*. Subsequent studies can later identify whether this upper bound is in fact the least upper bound (supremum) of the SDF implied by alternative sets of test assets.

Overall, I argue that hedge fund portfolios as test assets offer us a net benefit to achieving a more accurate representation of the investment opportunity set. Ultimately, the proof will be in the empirical results generated from using them as a new set of test assets.

4.3 Pricing Performance: Magnitude and Persistence

The first set of results reported in this section begins with Table 4. As in Section 3, this table follows the same analysis as Table 2 with the differences being: (i) the sample is now from January 1996 to June 2023 at a monthly frequency (to coincide with the available hedge fund data); and (ii) the maximum squared SR of hedge fund portfolios (in red) is reported in the bottom panel. The main results of the table are largely unchanged. That is, the upper panel again shows that the first five PCs are able to explain most of the cross-sectional variation in expected returns of the anomaly portfolios. In the bottom panel, the maximum squared SR of all anomalies rises considerably (from 9.31 to 24.52) thereby reflecting the shorter time window, sampled at a monthly frequency. That said, there remains a large distance between the maximum squared SR of the PC1-5 factor model and the test assets which is also confirmed with the 25 Size-B/M portfolios. The evidence to take note of in this table is that the maximum squared SR of hedge fund decile portfolios is larger (26.25) than any of the other underlying sets of test assets. Importantly, there is not too wide a gap betwen the Jensen et al. (2023) anomalies' maximum squared SR and that implied by hedge fund portfolios, which suggest that there is considerable overlap between the risk premia exposure of both sets of test assets. This initial piece of evidence is encouraging that hedge fund portfolios are worthwhile to investigate further as a new set of test assets.

Figure 10 conducts the same pseudo OOS analysis as in Figure 4 now with the 25 Size-B/M portfolios replaced by hedge fund decile portfolios. Given that the sample period now begins in January 1996, the OOS period has a start date of September 2009. As before, near arbitrage opportunities are *not* persistent over time with respect to the anomaly portfolios. More specifically, the OOS SRs are, on average, 43% as large as the IS SRs for the anomaly portfolios. In sharp contrast to the factor theme anomaly test assets, there is evidence of persistence in the near arbitrage opportunities of hedge fund portfolio test assets. In particular, the highest decile hedge funds (D10) is above the 45 degree line (i.e. an OOS SR of 5.43 compared to an IS SR of 5.27). This is evidence that near arbitrage opportunities are not as transient as previously considered (Kozak et al., 2018). It is important to note that while the remaining nine decile portfolios fall below the 45 degree line, this is to be expected given that there is a small minority of hedge fund managers with skill (i.e. less than 10% based on the estimate of Chen et al. (2017)).

As mentioned previously, the most important exercise involves the same pseudo OOS analysis with respect to the maximum squared SR (Table 3) in order to determine whether or not the lower ranked PCs (i.e. near arbitrage opportunities) matter to expected returns. Table 5 conducts this exercise with respect to anomaly versus hedge fund portfolios. As before the maximum squared SR of the anomalies is significantly lower in the OOS period (i.e. one fifth the size of the IS period). This is consisent with the results from Figure 10. All of this changes when we now look at the maximum squared SR with respect to hedge fund portfolios: the maximum squared SR more than doubles in the OOS period (48.98) compared to the IS period (22.46). Most importantly, hedge fund portfolios get us closer to spanning the SDF by generating a SR that is 3.4x higher than the traditional anomaly portfolios during the OOS period. In contrast to our existing understanding of the cross-section of expected returns, this piece of evidence confirms two new important facts: near arbitrage opportunities are both *larger* and more *persistent* than previously documented. With that, it becomes less clear that we can rule out near arbitrage opportunities in markets. Having established the cross-sectional results of the hedge fund portfolios, we now turn towards the time-series dynamics of this new set of test assets.

4.4 Variance Dynamics of the SDF's Upper Bound (Implied by Hedge Fund Portfolios)

Figure 11 displays the one-year rolling variance of the SDF based on the hedge fund decile portfolios (blue line) versus the upper bound (red line). Recall that the upper bound is simply a three-factor model comprised of hedge fund deciles eight to ten (D8-D10). Consistent with the SDF estimated via the anomaly portfolios, the variance is highly time-varying with upward spikes during market dislocations (e.g. the most recent episodes include the 2000 Technology Bubble, 2008 Financial Crisis, 2020 COVID-19 Stock Market Crash and 2022 Russia-Ukraine War). However, the noteworthy departure is that we now see significant factor structure dynamics across each of the most recent market dislocations in contrast to Figure 5 which exhibited a disproportionate impact of the Technology Bubble with respect to the SDF's Variance based on PC1-5 factor model. Recall that decile portfolios (especially the upper bound) are highly related to portfolios estimated via PCA. In other words, the decile portfolios exhibit a more balanced representation of different market dislocations.

Figure 12 better showcases the factor structure dynamics (of hedge fund portfolios) by displaying a heatmap of the scree plot over time estimated via a five-year rolling window with monthly observations. Recall, the warmer shades (yellow and orange) represent larger eigenvalues (stronger factor structure) whereas the cooler shades (blues and purples) represent smaller eigenvalues (weaker factor structure). The key takeaway from this figure is that the factor structure of hedge fund test assets shows that each of the most recent market dislocations has a significant impact on the factor structure in contrast to the previous heatmap which showed the Tech Bubble dominated the factor structure. This is what we should expect of test assets that span the SDF. That is, we should expect risk premia (including near arbitrage opportunities) to be large during periods of market turbulence. Lastly, Figure 13 plots the first difference between the adjacent five leading PCs (with respect to hedge fund test assets) as shown previously with Figure 7. This evidence further highlights the third new fact regarding expected returns: near arbitrage opportunities occur from a more diverse set of periods characterized by market turbulence. That is, the factor structure is no longer the product of just the Tech Bubble.

4.5 Decomposing the Upper Bound of Hedge Fund Returns

After having presented both the cross-sectional and time-series results with respect to hedge fund test assets, we are left asking what exactly do these portfolios represent. Table 6 reports the results of univariate time series regressions of the hedge fund upper bound on the Jensen et al. (2023) factor theme anomalies (Panel A) and hedge fund strategies (Panel B). This table documents two interesting findings. First, the hedge fund upper bound has significant loadings across a number of factor theme anomaly portfolios. The only anomalies which the upper bound does not load on are: Momentum, Profit Growth and Short Term Reversal. The Low Risk Anomaly has the highest R^2 of 11%. When we include all 13 factor themes, the R^2 rises to 16%. Importantly, the hedge fund upper bound has highly significant alphas across all of the factor theme anomaly portfolios which is unchanged when we include them all in a multivariate regression.

Second, in Panel B of Table 6 we see that again the upper bound loads significantly on a wide array of hedge fund strategies. In particular only three of the 11 strategies are not significant: Fixed Income Arbitrage, Managed Futures and the Options Strategy. The L/S Equity Hedge has the highest R^2 of 14%, which is reassuring given that we are interested in spanning the cross-section of expected stock returns. When we include all 11 hedge fund strategies, the R^2 rises to 25%. Importantly, the hedge fund upper bound has highly significant alphas across all of the hedge fund strategies which is left unchanged when we include them all in a multivariate regression. All together, this evidence confirms that the hedge fund upper bound portfolio returns is comprised of a wide array of exposures with respect to factor theme anomalies and hedge fund strategies. In spite of these exposures, the vast majority of the upper bound's return variation is left unexplained. This is a necessary feature of a set of test assets that arguably gets us closer to spanning the SDF.

4.6 Robustness

In this subsection I address the robustness concerns with respect to building managed portfolios (that serve as test assets) from hedge fund returns. There are five main issues regarding hedge funds: (i) The universe of hedge funds is too diverse and will include hedge funds trading instruments other than individual stocks; (ii) Hedge fund returns often include leverage; (iii) Hedge funds cannot be shorted (i.e. cannot create zero-cost L/S portfolio); (iv) Hedge funds face AUM (or capacity) limits; and (v) Managed portfolios need to be implementable which requires identifying hedge funds with skill ex-ante. In the following analysis I address each of these potential issues and find that the results are qualitatively unchanged.

To address the first two issues raised, I restrict the sample of hedge funds to those that only trade individual stocks (Equity Market Neutral and Long/Short strategies) and those that do not use leverage. Among this smaller set of hedge fund returns I then identify hedge fund managers with skill based on the early insight of Kraus and Litzenberger (1976) that was more recently highlighted by Back et al. (2018): Fund managers that lack skill will not be able to produce positive alpha without also producing negative market timing.

Said differently, hedge fund managers with skill will produce positive alpha without negative market timing. To that end, I estimate the following classic market timing test regression of Treynor and Mazuy (1966) with respect to hedge fund returns:

$$R_t - R_{t,f} = a + b_1 \left(R_{m,t} - R_{t,f}\right) + b_2 \left(R_{m,t} - R_{t,f}\right)^2 + \varepsilon, \tag{7}$$

where $R_t - R_{t,f}$ represent excess hedge fund returns and $R_{m,t} - R_{t,f}$ are excess market returns.

Based on Equation 7, I estimate hedge fund returns with skill as those hedge funds that have (i) a positive intercept (a) that is significant at the 1% level and (ii) a positive b_2 coefficient loading. This latter condition avoids the unskilled managers that produce positive alpha by taking on strategies that have undesirable market timing (i.e. more extreme negative returns when the market is also experiencing a downturn). I form an equal-weighted portfolio of those hedge funds identified with skill based on three-year rolling regressions with an annual rebalancing every December. This ends up producing an OOS hedge fund portfolio with a Sharpe Ratio of 1.03 with skewness of 0.54.

The final issue addresses the reality that hedge funds cannot be shorted. In aggregate, hedge fund returns are often viewed as producing returns that resemble shorting volatility (Agarwal and Naik, 2004; Jurek and Stafford, 2015). In light of this stylized fact, I use CBOE VIX futures as a tradable asset that can easily be shorted (does not require dynamic hedging unlike options), thereby resembling average hedge fund returns. VIX futures have the added benefit that they comprise a deep and liquid market (relative to shorting SPX put options that are bifurcated into moneyness and maturity), which addresses the issue that hedge funds face capacity limits.[11] I then build a factor-mimicking portfolio of hedge fund returns identified with skill using the 2nd dated VIX futures returns. In particular, I run the following three-year rolling regression to construct the hedge fund FMP:

$$R_t - R_{t,f} = \gamma + \beta \left(R_{\text{VIX},t}\right) + \epsilon, \tag{8}$$

where $R_t - R_{t,f}$ represent equal-weighted portfolio of skilled hedge fund returns and $R_{\text{VIX},t}$ are VIX futures returns which are in excess of the risk-free rate.

After estimating Equation 8, I then construct the hedge fund FMP as:

$$R_t^{\text{FMP}} = \hat{\beta}_{t-36} \times \left(R_{\text{VIX},t}\right), \tag{9}$$

where I use the coefficient loading on the VIX futures from three-years ago. This produces a FMP that serves as the test asset of hedge fund returns with a SR of 1.02 and skewness of

[11] Since VIX futures began trading in March 2004, I splice the synthetic VIX futures returns as used in Johnson (2017) (from Travis Johnson's website) to cover the period of VIX returns prior to March 2004.

0.74. It is important to note that this test asset is now fully implementable and is zero-cost.

Figures 14 and 15 display the new results using the hedge fund FMP which confirms that the results are largely unchanged. In Figure we see that the pseudo OOS SR is higher than the pseudo IS SR for the FMP which confirms that the relatively high risk-adjusted returns are not a transient feature of the sample period but persists over the entire sample period. Lastly, Figure 15 highlights how the the maximum squared SR increases significantly when the FMP is added as a new test asset. In particular the variance of the SDF increases by 50%. Moreover, the variation is significantly different from the JKP portfolios and experiences much more dramatic and frequent spikes upward in volatility. All together, this evidence indicates that hedge fund returns offer a new window into capturing risk premia that was previously unexplored following the traditional approach that relies on combinations of firm characteristics. Hedge fund managed portfolios reflect return variation that may reflect other variables or firm characteristics that have not yet been identified in the literature on the cross-section of stock returns.

5 Conclusion

I document that a large and comprehensive set of test assets does not span the SDF. In particular, I show that the factor structure of the test assets is dominated in the time-series by the Tech Bubble and is dominated in the cross-section by only three (out of 13) factor themes: Value, Low Leverage and Seasonality.

After having established that the existing set of anomalies fails to span the SDF, I offer an alternative set of test assets: hedge fund portfolios. By adopting hedge fund portfolios as test assets, I show that near arbitrage opportunities are larger, more persistent and are the product of a wider array of periods characterized by market turbulence. All together, this new evidence forces us to revist our current prior regarding the absence of near arbitrage opportunities. While this asset pricing prior is likely satisfied on an unconditional basis, it is far less clear whether it is not violated during market dislocations when the factor structure is highly unstable. Updating our economic prior to reflect the dynamics of the factor structure will undoubtedly lead to better OOS prediction in the important and emerging field of economic regularization.

References

Agarwal, Vikas, and Narayan Y. Naik, 2004, Risks and portfolio decisions involving hedge funds, *The Journal of Finance* 17, 63–98.

Ahn, Dong-Hyun, Jennifer Conrad, and Robert F. Dittmar, 2009, Basis assets, *The Review of Financial Studies* 22, 5133–5174.

Back, Kerry, Alan D. Crane, and Kevin Crotty, 2018, Skewness consequences and seeking alpha, *The Review of Financial Studies* 31, 4720—-4761.

Barth, Daniel, Juha Joenväärä, Mikko Kauppila, and Russ Wermers, 2023, A bias-free assessment of the hedge fund industry: A new evaluation of total assets, alphas, and the flow-performance relation, Working Paper.

Brown, Gregory, Christian Lundblad, and William Volckmann, 2024, What do we know about institutional-quality hedge funds?, Institute for Private Capital.

Bryzgalova, Svetlana, Victor DeMiguel, Sicong Li, and Markus Pelger, 2023a, Asset-pricing factors with economic targets, Working Paper.

Bryzgalova, Svetlana, Jiantao Huang, and Christian Julliard, 2023b, Bayesian solutions for the factor zoo: We just ran two quadrillion models, *The Journal of Finance* 78, 487–557.

Bryzgalova, Svetlana, Markus Pelger, and Jason Zhu, 2023c, Forest through the trees: Building cross-sections of stock returns, Journal of Finance, Forthcoming.

Chen, Andrew Y., Alejandro Lopez-Lira, and Tom Zimmermann, 2024, Does peer-reviewed research help predict stock returns?, Working Paper.

Chen, Yong, Michael Cliff, and Haibei Zhao, 2017, Hedge funds: The good, the bad, and the lucky, *The Journal of Financial and Quantitative Analysis* 52, 1081–1109.

Chinco, Alex, Andreas Neuhierl, and Michael Weber, 2021, Estimating the anomaly base rate, *Journal of Financial Economics* 140, 101–126.

Cochrane, John H., 2011, Presidential address: Discount rates, *The Journal of Finance* 66, 1047–1108.

Couts, Spencer J., Andrei S. Gonçalves, and Andrea Rossi, 2024, Unsmoothing returns of illiquid funds, *The Review of Financial Studies* 37, 2110–2155.

Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos, 2020, The cross-section of risk and returns, *The Review of Financial Studies* 33, 1927–1979.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2020, Taming the factor zoo: A test of new factors, *The Journal of Finance* 75, 1327–1370.

Giglio, Stefano, Yuan Liao, and Dacheng Xiu, 2021, Thousands of alpha tests, *The Review of Financial Studies* 34, 3456–3496.

Giglio, Stefano, Dacheng Xiu, and Dake Zhang, 2023, Test assets and weak factors, Journal of Finance, Forthcoming.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen, 2023, Is there a replication crisis in finance?, *The Journal of Finance* 78, 2465–2518.

Joenväärä, Juha, Mikko Kauppila, Robert Kosowski, and Pekka Tolonen, 2021, Hedge fund performance: Are stylized facts sensitive to which database one uses?, *Critical Finance Review* 10, 271–327.

Johnson, Travis L., 2017, Risk premia and the vix term structure, *The Journal of Financial and Quantitative Analysis* 52, 2461–2490.

Jurek, Jakub W., and Erik Stafford, 2015, The cost of capital for alternative investments, *The Journal of Finance* 70, 2185–2226.

Kelly, Bryan T., and Hao Jiang, 2012, Tail risk and hedge fund returns, Working Paper.

Kelly, Bryan T., Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* 134, 501–524.

Koijen, Ralph S.J., Robert J. Richmond, and Motohiro Yogo, 2023, Which investors matter for equity valuations and expected returns?, *The Review of Economic Studies* 91, 2387–2424.

Kozak, Serhiy, and Stefan Nagel, 2024, When do cross-sectional asset pricing factors span the stochastic discount factor?, Working Paper.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *The Journal of Finance* 73, 1183–1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.

Kraus, Alan, and Robert H. Litzenberger, 1976, Skewness preference and the valuation of risk assets, *The Journal of Finance* 31, 1085–1100.

Lettau, Martin, and Markus Pelger, 2020, Factors that fit the time series and cross-section of stock returns get access arrow, *The Review of Financial Studies* 33, 2274–2325.

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *The Journal of Finance* 71, 5–32.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *The Review of Financial Studies* 29, 104–147.

Pedersen, Lasse Heje, 2015, *Efficiently Inefficient: How Smart Money Invests and Market Prices are Determined* (Princeton University Press).

Shiller, Robert, 1981, Do stock prices move too much to be justified by subsequent changes in dividends?, *American Economic Review* 71, 421–436.

Treynor, Jack, and Kay Mazuy, 1966, Can mutual funds outguess the market?, *Harvard Business Review* 44, 131–136.

Table 1: Descriptive Statistics of Anomaly and Hedge Fund Portfolios
The table presents the summary statistics for the factor theme anomaly portfolios (Panel A) and hedge fund decile portfolios (Panel B). The sample periods reflect the full sample that is available for each respective dataset. Anomaly (Hedge Fund) portfolio return statistics are annualized from their daily (monthly) returns. The mean and volatility of returns are expressed in percent.

Panel A: Anomaly Portfolios (Jan. 1952 to Dec. 2023)

	Mean Return	Volatility	Sharpe Ratio	Skewness
Accruals	2.90	3.10	0.93	−0.08
Debt Issuance	2.46	2.07	1.19	−0.07
Investment	3.00	5.97	0.50	−0.34
Low Leverage	0.25	7.79	0.03	−0.05
Low Risk	1.08	10.98	0.10	−0.15
Momentum	4.30	8.57	0.50	−1.44
Profit Growth	2.04	3.10	0.66	−1.07
Profitability	2.17	5.43	0.40	0.56
Quality	3.24	4.56	0.71	−0.86
Seasonality	1.68	1.72	0.98	−0.28
Short Term Reversal	2.05	3.45	0.59	0.95
Size	2.18	5.95	0.37	−0.10
Value	3.67	8.49	0.43	0.64

Panel B: Hedge Fund Decile Portfolios (Jan. 1996 to Jun. 2023)

	Mean Return	Volatility	Sharpe Ratio	Skewness
D1	−92.90	23.75	−3.91	−3.18
D2	−36.63	16.42	−2.23	−2.77
D3	−19.51	13.60	−1.43	−2.46
D4	−8.53	12.12	−0.70	−2.01
D5	0.93	11.43	0.08	−1.31
D6	9.96	11.37	0.88	−0.48
D7	19.84	11.78	1.68	0.31
D8	31.51	12.72	2.48	1.01
D9	49.24	14.94	3.30	1.55
D10	106.96	20.63	5.18	2.01

Table 2: Revisiting Anomaly Portfolios with Principal Component Factors
This table revisits the analysis conducted in Table II of Kozak et al. (2018) with the following modifications. The long-short strategy daily returns have been replaced by a larger and more comprehensive set of long-short strategy daily returns from Jensen et al. (2023) that span 13 themes comprised of 153 factors. In addition, I have added the Daniel et al. (2020) ("DMRS") factors, indicated by an asterisk (*), that hedge out unpriced risk. All table entries have been annualized. The sample period for all of the analysis, except for the DMRS factors, is January 1952 to December 2023. The DMRS factors uses data from July 1963 to March 2023.

		PC Factor-Model Alphas				
	Mean Return	PC1	PC1-2	PC1-3	PC1-4	PC1-5
Accruals	2.90	2.88	2.97	2.65	2.66	3.60
Debt Issuance	2.46	2.51	2.43	2.27	2.25	1.90
Investment	3.00	1.83	2.10	1.62	1.50	1.66
Low Leverage	0.25	1.95	1.73	1.61	1.80	1.94
Low Risk	1.08	−1.23	−2.14	−2.35	−2.03	−0.93
Momentum	4.30	4.24	1.73	0.94	0.76	0.11
Profit Growth	2.04	2.31	1.73	1.84	1.84	1.35
Profitability	2.17	1.50	0.89	1.92	1.90	0.28
Quality	3.24	3.39	2.49	3.04	3.21	1.37
Seasonality	1.68	1.57	1.46	1.36	1.38	1.18
Short Term Reversal	2.05	1.91	2.05	2.17	2.26	2.96
Size	2.18	2.11	3.16	2.13	2.29	−0.51
Value	3.67	1.88	2.78	2.99	2.84	1.90

		PC Factors' Max. Squared SR				
	Max SR^2	PC1	PC1-2	PC1-3	PC1-4	PC1-5
All Anomalies	9.31	0.07	0.22	0.32	0.33	1.76
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
For Comparison:						
25 Size-B/M	3.83	0.89	1.00	1.22	1.89	2.07
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MKT, SMB and HML	0.71	n/a	n/a	n/a	n/a	n/a
MKT*, SMB*, HML*	1.99	n/a	n/a	n/a	n/a	n/a

Table 3: Anomaly Portfolios with Principal Component Factors (IS versus OOS)
This table repeats the empirical work in Table 2 except now for the two equal halves as in Figure 4 whereby the first half (upper panel) represents the IS period (January 1952 to December 1987) and the second half (bottom panel) represents the OOS period (January 1988 to December 2023). All table entries have been annualized.

	Max SR^2	PC Factors' Max. Squared SR (IS)				
		PC1	PC1-2	PC1-3	PC1-4	PC1-5
All Anomalies	20.77	0.25	1.73	2.19	2.24	7.58
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
25 Size-B/M	8.13	1.59	1.59	4.65	4.70	4.70
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

	Max SR^2	PC Factors' Max. Squared SR (OOS)				
		PC1	PC1-2	PC1-3	PC1-4	PC1-5
All Anomalies	6.69	0.04	0.10	0.11	0.12	1.61
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
25 Size-B/M	4.76	0.60	0.67	0.67	1.03	1.14
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Table 4: Anomaly Portfolios with PC Factors (Jan. 1996 to Jun. 2023)
This table repeats the empirical work in Table 2 except for a shorter sample period from January 1996 to June 2023. The shorter time series is used here to accomodate the more limited hedge fund data. As previously, the long-short strategy daily returns are from Jensen et al. (2023). The Daniel et al. (2020) ("DMRS") factors are indicated by an asterisk (*). All table entries have been annualized.

	Mean Return	PC Factor-Model Alphas				
		PC1	PC1-2	PC1-3	PC1-4	PC1-5
Accruals	1.15	1.34	1.42	1.38	1.38	2.15
Debt Issuance	2.15	2.23	2.16	2.12	2.11	1.40
Investment	2.99	1.51	1.84	1.70	1.74	1.50
Low Leverage	0.07	2.42	2.06	2.09	1.87	1.91
Low Risk	1.66	−1.15	−2.13	−2.11	−2.54	−0.25
Momentum	4.12	3.87	1.34	1.17	1.35	0.09
Profit Growth	0.57	0.87	0.30	0.35	0.36	−0.38
Profitability	3.59	2.40	2.00	2.18	2.25	0.09
Quality	3.74	3.78	2.91	3.05	2.80	0.79
Seasonality	1.42	1.27	1.14	1.12	1.10	1.03
Short Term Reversal	1.43	1.31	1.40	1.43	1.40	2.23
Size	1.44	1.54	2.49	2.33	1.88	−0.73
Value	3.84	1.48	2.53	2.53	2.72	1.18

	Max SR^2	PC Factors' Max. Squared SR				
		PC1	PC1-2	PC1-3	PC1-4	PC1-5
All Anomalies	24.52	0.27	0.58	0.59	0.68	6.81
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
For Comparison:						
25 Size-B/M	19.70	1.99	2.08	2.09	3.54	4.15
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MKT, SMB and HML	1.46	n/a	n/a	n/a	n/a	n/a
MKT*, SMB*, HML*, RMW* and CMA*	8.23	n/a	n/a	n/a	n/a	n/a
Hedge Fund Decile Portfolios	26.25	n/a	n/a	n/a	n/a	n/a

Table 5: HF and Anomaly Portfolios with Principal Component Factors (IS versus OOS) This table repeats the empirical work in Table 3 except now for the hedge fund sample period whereby the first half (upper panel) represents the IS period (January 1996 to September 2009) and the second half (bottom panel) represents the OOS period (October 2009 to June 2023). All table entries have been annualized.

		PC Factors' Max. Squared SR (IS)				
	Max SR^2	PC1	PC1-2	PC1-3	PC1-4	PC1-5
All Anomalies	22.28	0.08	0.14	0.21	2.78	3.79
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Hedge Fund Decile Portfolios	22.46	n/a	n/a	n/a	n/a	n/a

		PC Factors' Max. Squared SR (OOS)				
	Max SR^2	PC1	PC1-2	PC1-3	PC1-4	PC1-5
All Anomalies	4.22	0.10	0.14	0.16	0.71	0.72
χ^2 p-value for zero pricing errors		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Hedge Fund Decile Portfolios	48.98	n/a	n/a	n/a	n/a	n/a

Table 6: Decomposition of Hedge Fund Upper Bound Portfolio Returns
The table presents univariate regressions of the hedge fund upper bound portfolio returns (equal weighted across D8-D10) spanned on factor theme anomaly portfolios (Panel A) and hedge fund strategy returns (Panel B). Coefficient estimates (in decimal form) of the intercept and coefficient are reported along with their respective Newey and West (1987) t-statistics and R^2. The sample period is monthly returns from January 1996 to June 2023. Asterisks denote the levels of statistical significance of the coefficient: 10% level (*), 5% level (**) and 1% level (***).

Panel A: Anomaly Portfolios

	Intercept	t-stat	Coefficient	t-stat	R^2
Accruals**	0.155	14.21	1.743	2.27	0.03
Debt Issuance**	0.152	15.01	2.521	1.96	0.04
Investment*	0.158	13.09	−0.605	−1.79	0.02
Low Leverage***	0.157	14.10	0.748	2.58	0.08
Low Risk***	0.157	14.59	−0.783	−3.16	0.11
Momentum	0.157	13.75	−0.203	−0.57	0.01
Profit Growth	0.156	13.47	0.026	0.03	0.00
Profitability***	0.159	13.58	−1.040	−3.17	0.07
Quality*	0.158	13.27	−0.691	−1.66	0.01
Seasonality*	0.158	13.36	−1.389	−1.81	0.01
Short Term Reversal	0.157	13.11	−0.798	−1.44	0.01
Size***	0.156	13.68	0.951	2.96	0.04
Value**	0.158	13.55	−0.578	−2.03	0.05
All Anomalies	0.151	15.52	n/a	n/a	0.16

Panel B: Hedge Fund Strategy Returns

	Intercept	t-stat	Coefficient	t-stat	R^2
Convertible Arbitrage***	0.154	13.66	0.674	4.94	0.06
Dedicated Short Bias***	0.160	11.70	−0.271	−3.96	0.06
Emerging Markets***	0.154	13.82	0.401	3.43	0.08
Equity Market Neutral**	0.153	13.84	0.876	2.37	0.03
Event Driven***	0.152	14.04	0.727	2.97	0.06
Fixed Income Arbitrage	0.155	12.92	0.448	1.03	0.01
Global Macro**	0.155	13.14	0.396	2.21	0.02
L/S Equity Hedge***	0.152	15.75	0.771	4.90	0.14
Managed Futures	0.134	12.96	0.015	0.17	0.00
Multi Strategy***	0.153	14.24	0.770	2.64	0.06
Options Strategy	0.135	17.46	−0.020	−0.52	0.00
All Strategies	0.120	11.13	n/a	n/a	0.25

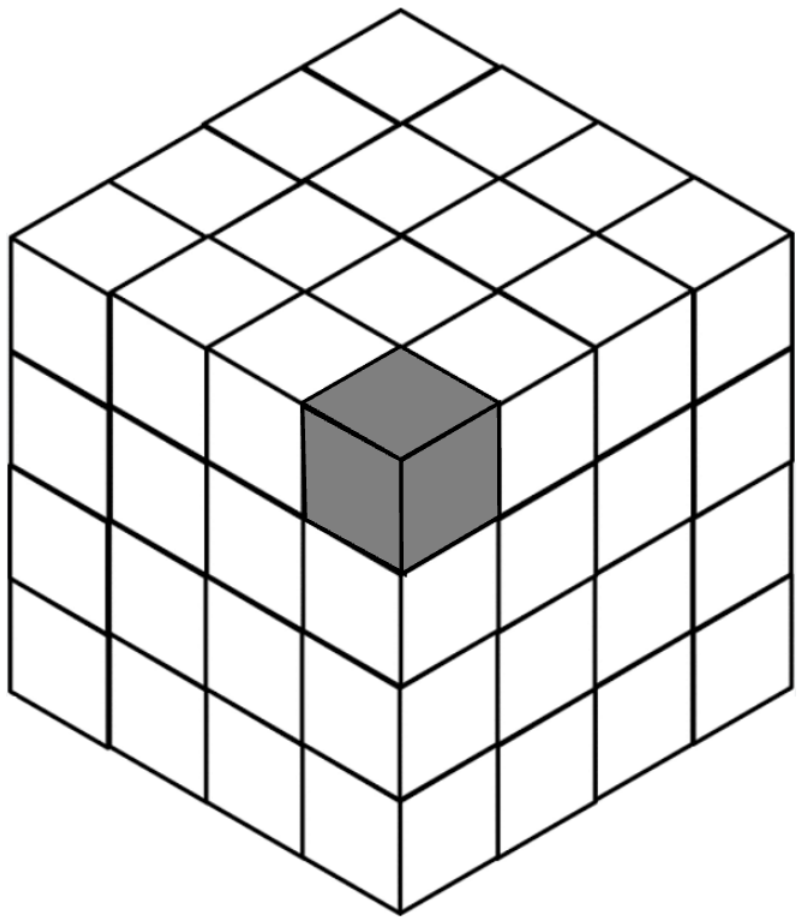
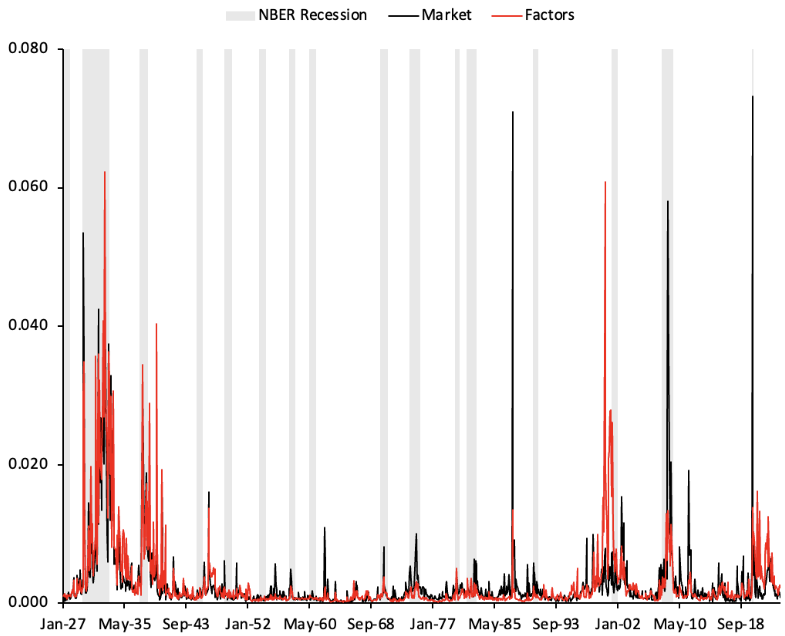


Figure 1: The Space of Relevant Test Assets
The figure displays a cube partitioned into smaller cubes to represent the set of relevant test assets required to span the SDF. Each cube represents a test asset that possesses a unique source of return variation required to span the SDF. The grey shaded cube represents the extant set of anomaly portfolios used to span the SDF.

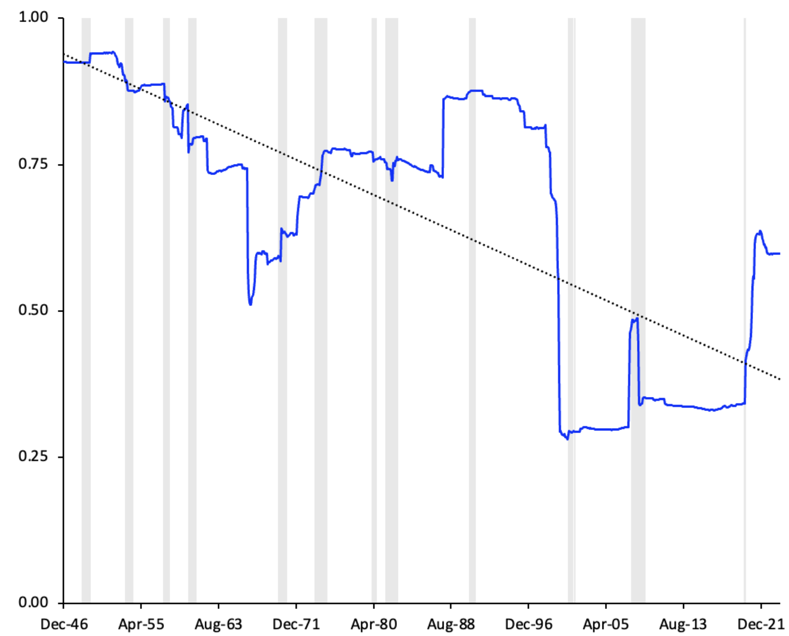
(a) Sharpe Ratio (SR)

(b) Information Ratio (IR)

Figure 2: Anomaly Portfolios (versus the Market) over Time
The figure displays the time-varying dynamics of Jensen et al. (2023) anomaly portfolios compared to the CRSP value-weighted index. Panel A displays the ten-year rolling Sharpe ratio (SR) of an equal-weighted combination portfolio of factor themes. Panel B shows the ten-year rolling information ratio (IR) of daily CAPM alphas (with Newey and West (1987) standard errors) of an equal-weighted combination portfolio of factor themes versus the value-weighted index excess return. The sample period is from January 1952 to December 2023 to ensure that all 13 factor themes are represented in the combination portfolio.

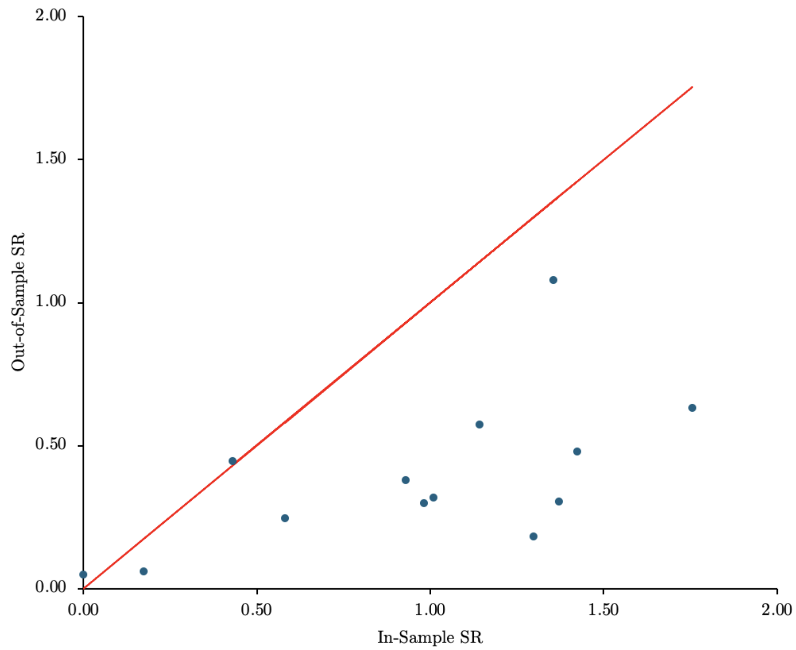


(a) Realized Volatility

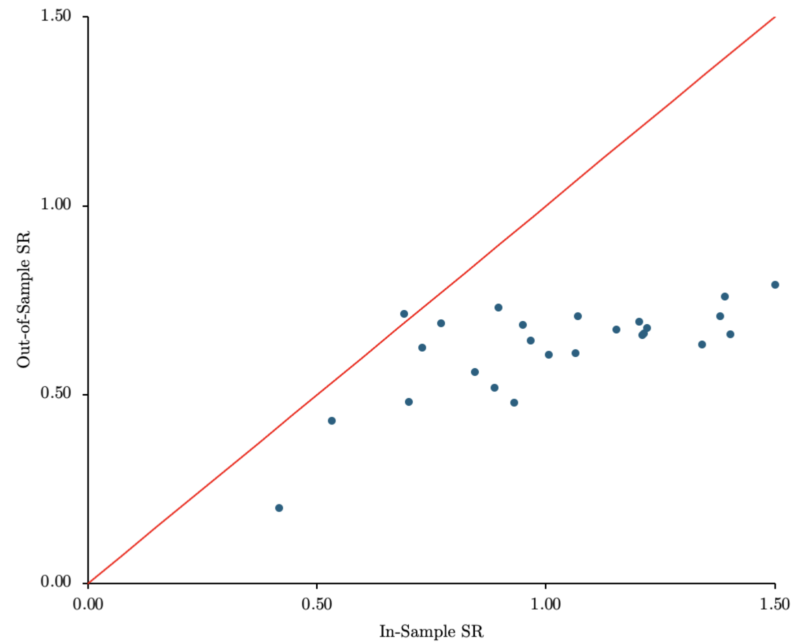


(b) Rolling 20-Year Correlation

Figure 3: Realized Volatility of Aggregate Market vs. Anomalies
Panel A displays the monthly realized volatility (based on daily returns) of the aggregate value-weighted market index (black line) versus the monthly realized volatility of an equal-weighted combination portfolio of the realized volatility from the Jensen et al. (2023) 13 Anomaly Portfolios (red line). Panel B displays the rolling 20-year correlation between the two realized volatilities along with the respective trendline. The combination portfolio is constructed in the following way. First, the anomaly returns are scaled to have 1/13th the annualized volatility of the aggregate market index. Second, the scaled anomaly returns are each squared and then summed into a combination portfolio at a daily frequency. The last step computes the monthly realized volatility based on the combination of the scaled daily squared anomaly excess returns. NBER recessions are indicated by the grey shaded regions. The sample period is from January 1927 to December 2023.



(a) Anomaly Portfolios



(b) 25 Size-B/M Portfolios

Figure 4: IS and OOS Sharpe Ratios (Anomaly vs. Size-B/M Portfolios)
The figure applies the pseudo OOS analysis from Figure 4 of Kozak et al. (2018). That is, I divide the full sample into two equal halves and treat the first half as the IS period whereas the second half of the sample is the OOS period with a start date of December 31, 1987. Panel A performs this analysis on the 13 factor themes from Jensen et al. (2023) whereas Panel B uses the 25 Size-B/M portfolios. The full sample period uses daily data from January 1952 to December 2023.

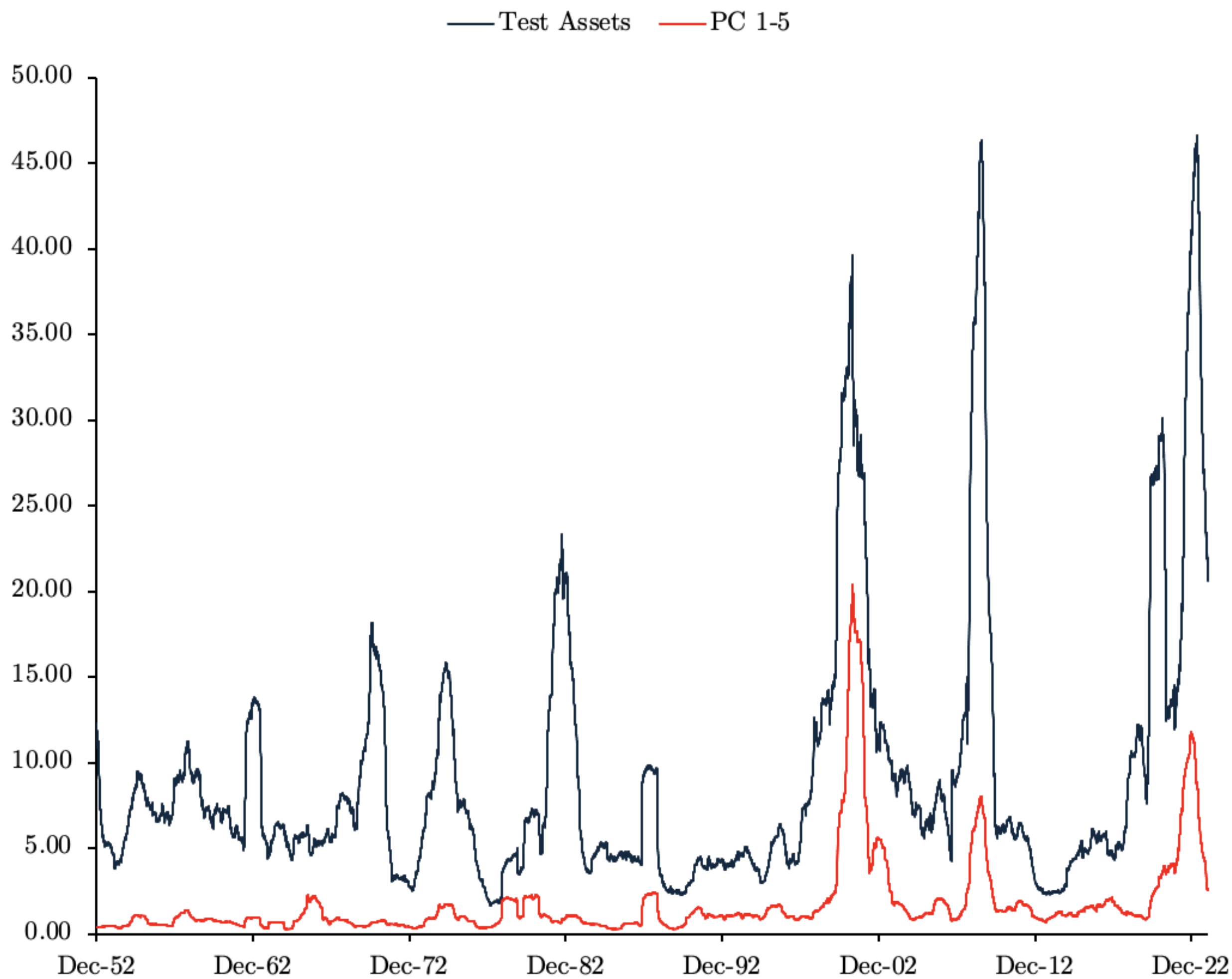


Figure 5: SDF Variance Over Time (Anomaly versus PC Portfolios)
The figure displays the one-year rolling variance estimates of the SDF based on the test assets (dark blue line) and the five leading PCs (PC 1-5) extracted from the test assets (red line). The test assets correspond to the 13 factor theme portfolios from Jensen et al. (2023). The sample period uses daily data from January 1952 until December 2023.

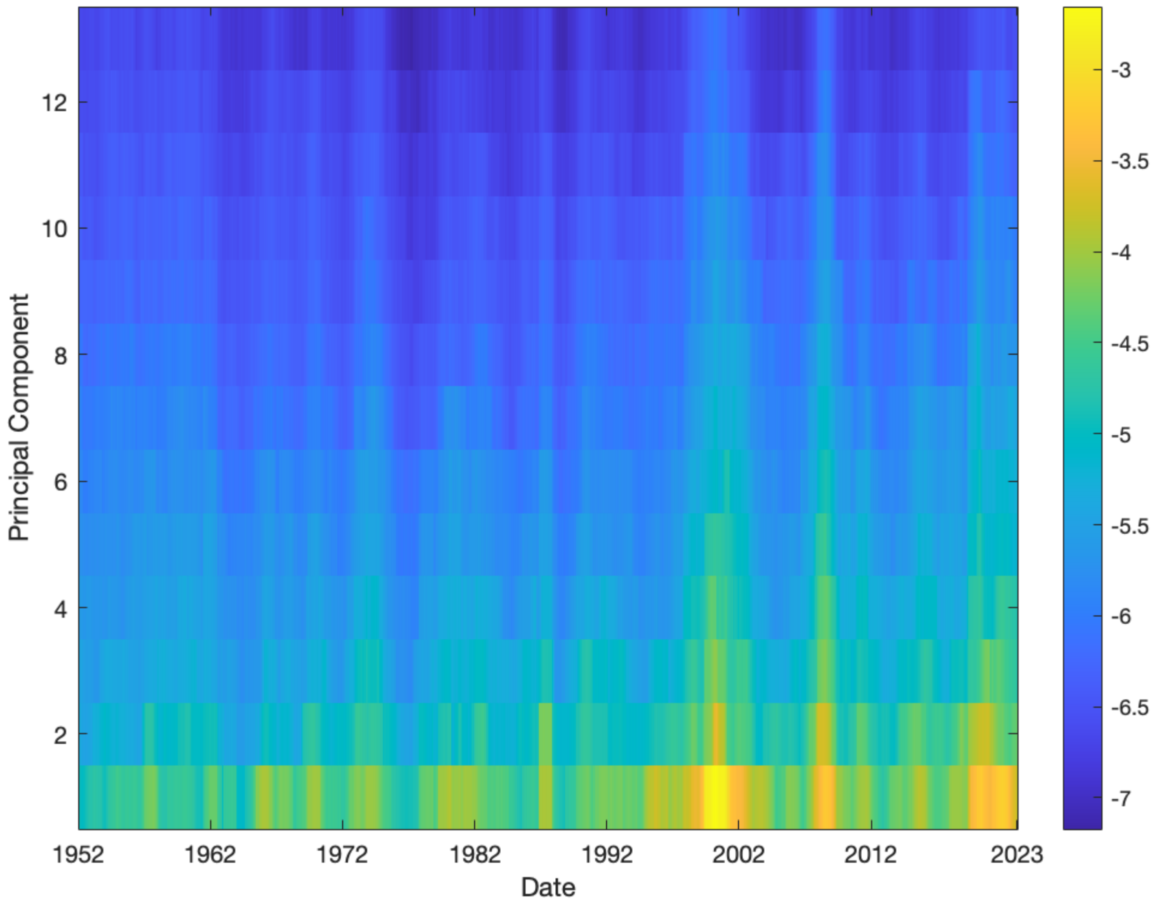


Figure 6: Factor Structure Over Time (Anomaly Portfolios)
The figure displays the time-varying factor structure of the anomaly portfolios. The colorbar in the two-dimensional view of the scree plot over time represents the magnitude of the eigenvalues on a logarithmic scale. The warmer shades (yellow and orange) represent larger eigenvalues (i.e. stronger factor structure). The cooler shades (blue and purple) represent smaller eigenvalues (i.e. weaker factor structure). Higher values of the vertical axis should correspond to decreasing importance of the higher order PCs (warmer to cooler shades). The factor structure is estimated via rolling PCA over the prior year. The sample period uses daily data from January 1952 until December 2023.

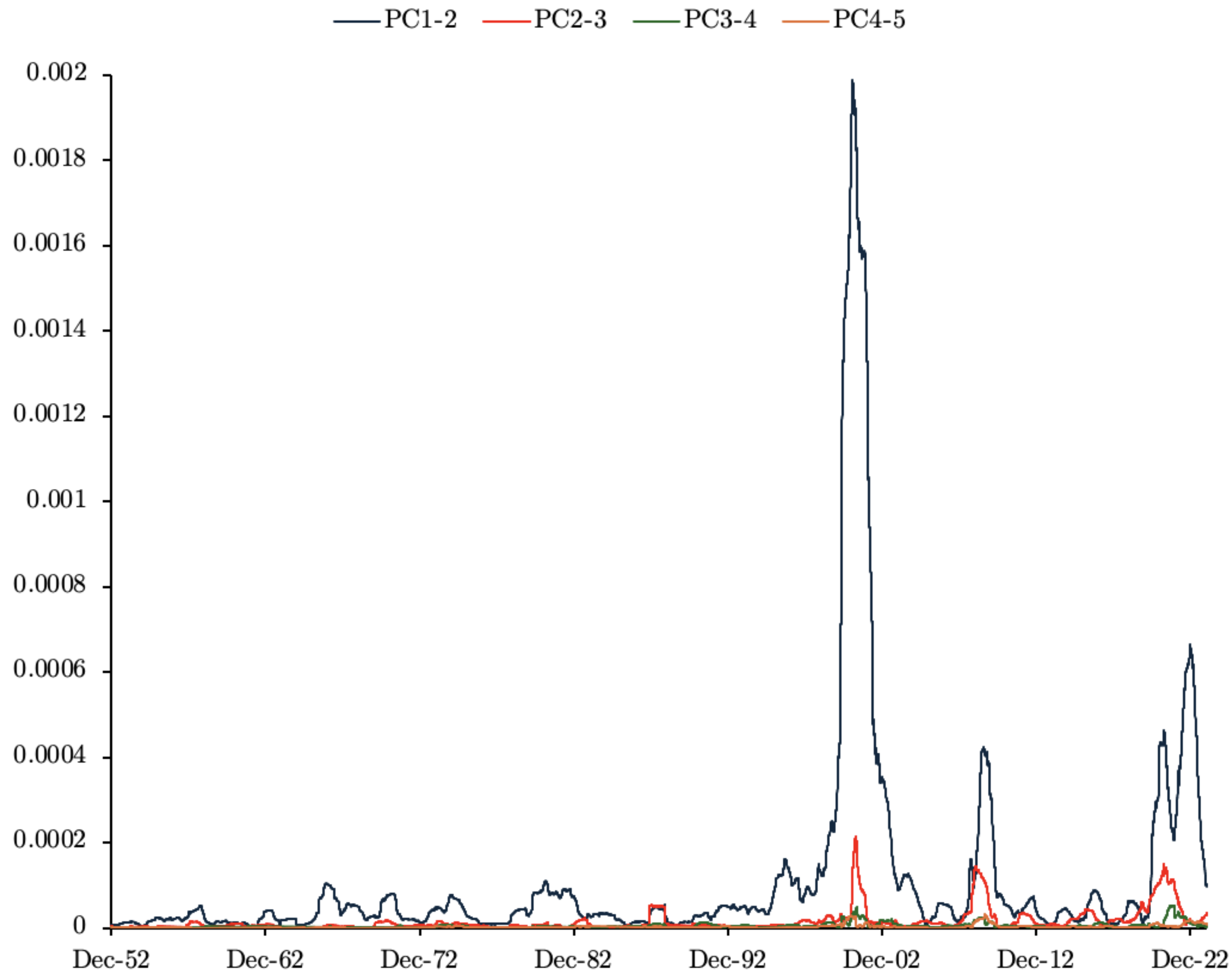


Figure 7: Δ PC Over Time (Anomaly Portfolios)
The figure displays the difference between the leading five principal components (based on the anomaly portfolios) over time. To estimate the time-varying dynamics of the factor structure, I estimate PCA using a rolling window of 252 trading days. PC1-2 (black line) refers to the difference between the leading two principal components: $PC1 - 2 = PC_{1,t} - PC_{2,t}$. PC2-3 (red line), PC3-4 (green line) and PC4-5 (orange line) follow the same calculation with their corresponding PCs. The factor structure is estimated via rolling PCA over the prior year. The sample period uses daily data from January 1952 until December 2023.

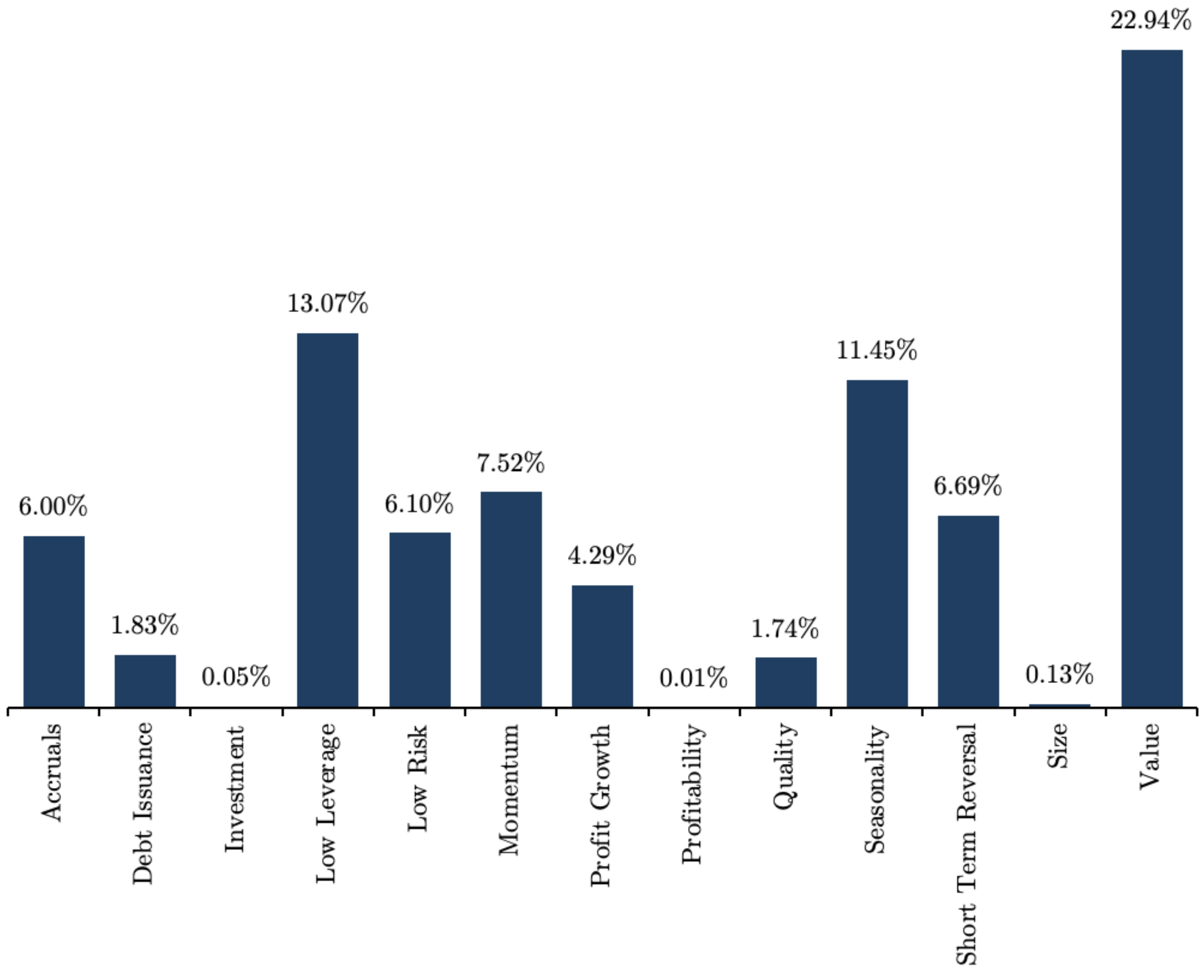


Figure 8: Factor Theme (%) of SDF Variance (Anomaly Portfolios)
The figure displays the percentage of each respective factor theme's contribution of the total SDF variance (maximum SR). That is, each factor theme percentage is calculated as $1 - \frac{\text{Max SR}^2_{\text{w/o factor}}}{\text{Max SR}^2_{\text{w/ all factors}}}$. The sample period uses daily data from January 1952 until December 2023.

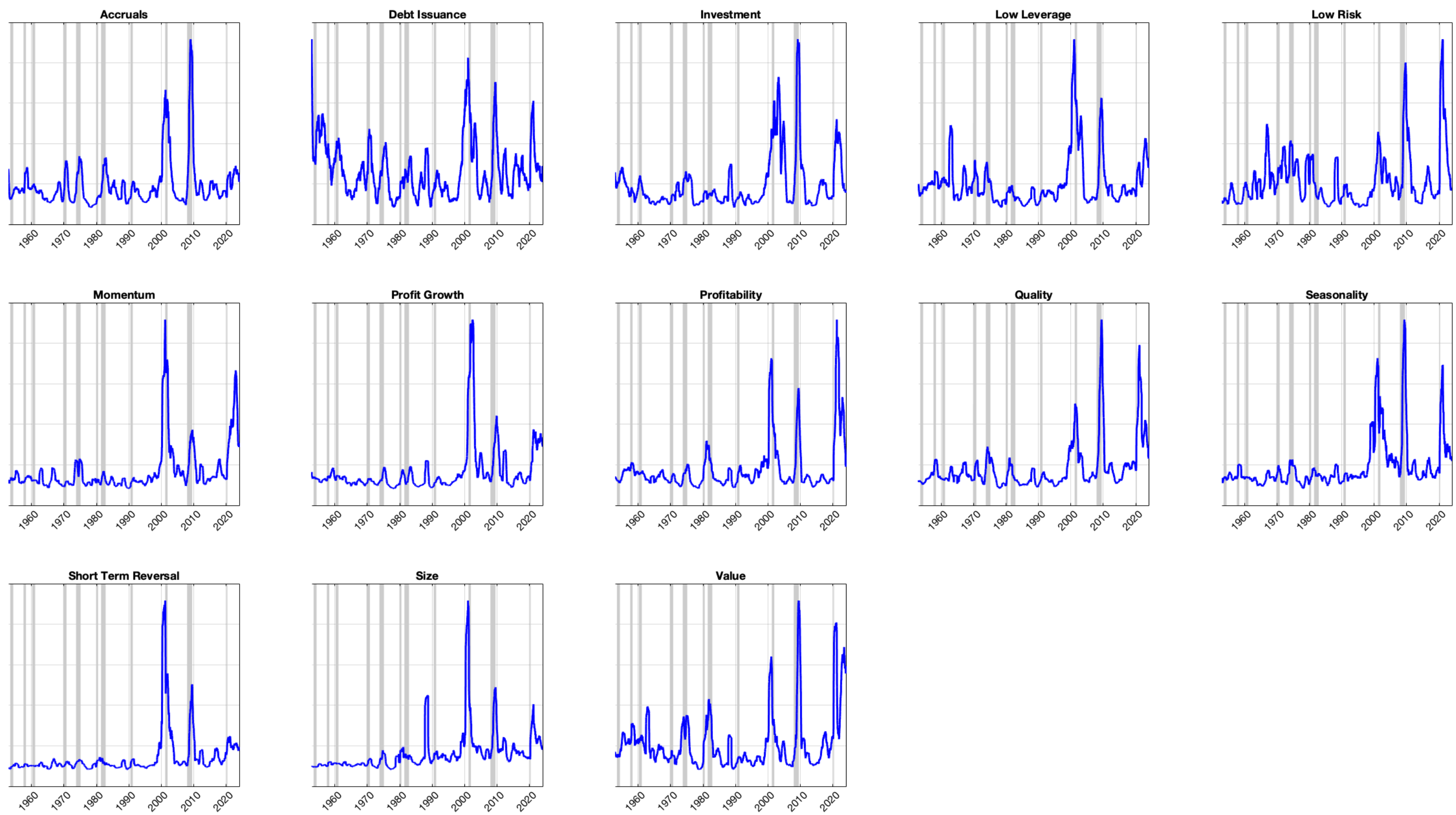
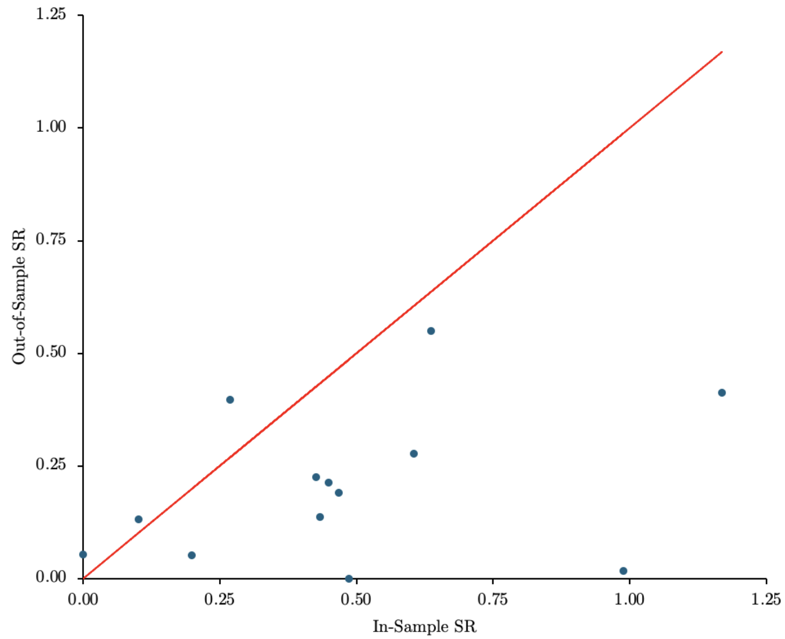
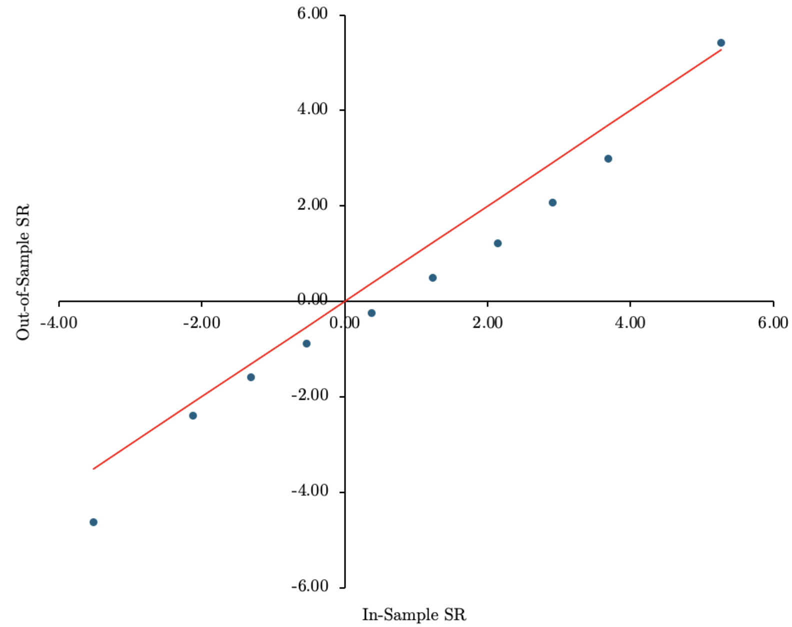


Figure 9: Factor Theme Contribution to SDF Variance Over Time (Anomaly Portfolios)
The figure displays the rolling one-year variance of the difference between the SDF with all factors and the SDF with the respective factor removed. Each respective factor theme is displayed as it's own subfigure. NBER shaded recessions are the grey shaded bars. The sample period uses daily data from January 1952 until December 2023.



(a) Anomaly Portfolios



(b) Hedge Fund Portfolios

Figure 10: IS and OOS Sharpe Ratios (Anomaly vs. Hedge Fund Portfolios)
The figure applies the pseudo OOS analysis as Figure 4 over a smaller sample period to account for the more limited time series of hedge fund data. The OOS period begins September 2009. Panel A performs this analysis on the 13 factor themes from Jensen et al. (2023) whereas Panel B now uses the ten hedge fund decile portfolios. The full sample period uses monthly data from January 1996 to June 2023.

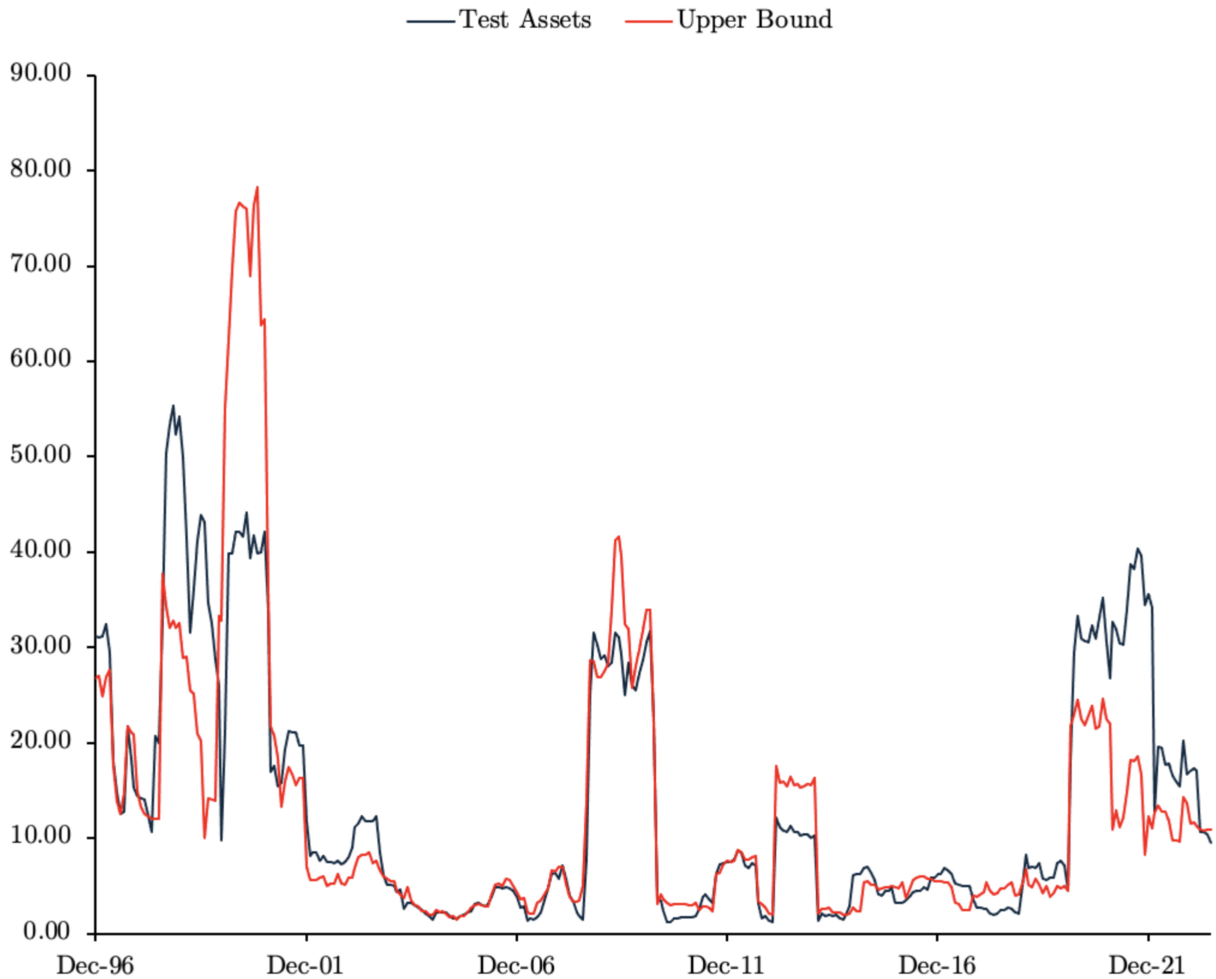


Figure 11: SDF Variance Over Time (Hedge Fund Deciles versus Upper Bound Portfolios) The figure displays the one-year rolling variance estimates of the SDF based on the test assets (dark blue line) and the upper bound (Deciles 8-10) extracted from the test assets (red line). The test assets correspond to the ten hedge fund decile portfolios. The sample period uses monthly data from January 1996 until June 2023.

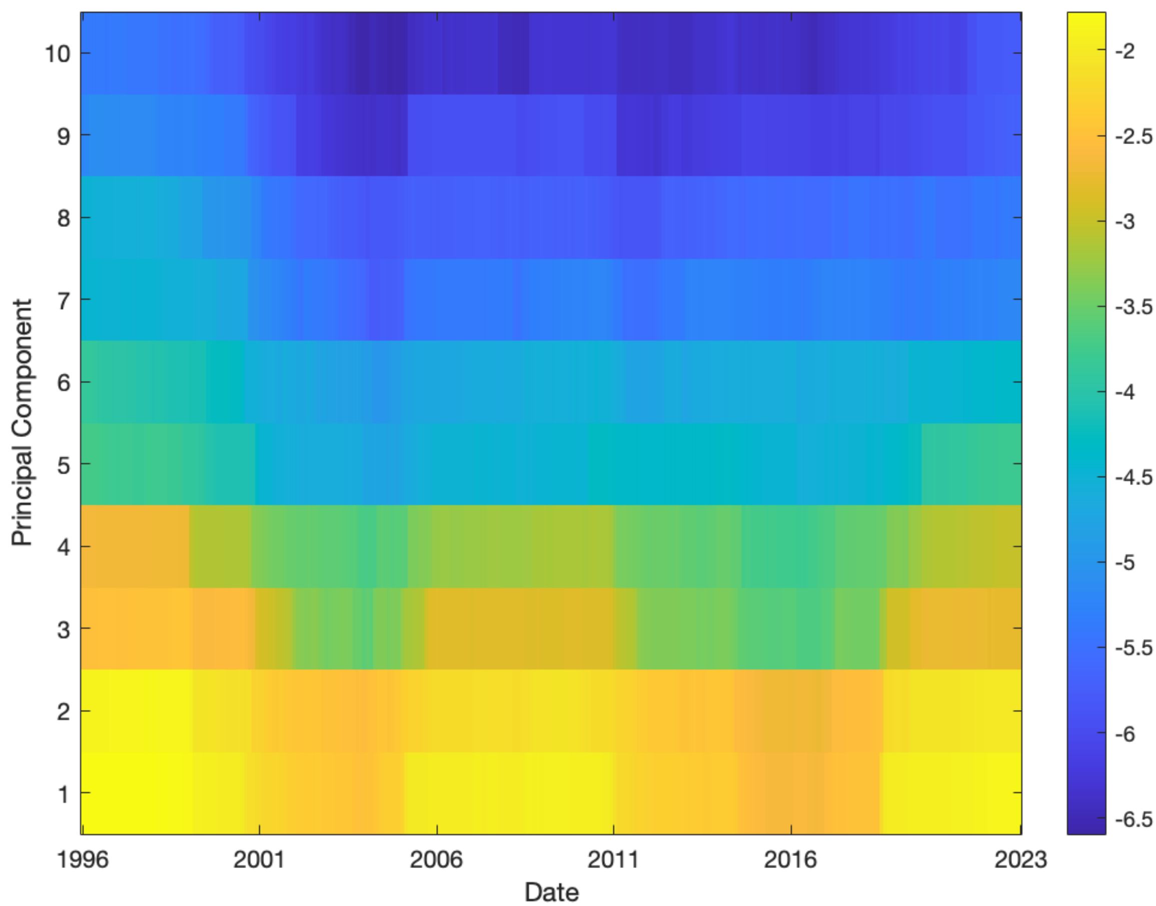


Figure 12: Factor Structure Over Time (Hedge Fund Portfolios)
The figure displays the time-varying factor structure of the hedge fund portfolios. The colorbar in the two-dimensional view of the scree plot over time represents the magnitude of the eigenvalues on a logarithmic scale. The warmer shades (yellow and orange) represent larger eigenvalues (i.e. stronger factor structure). The cooler shades (blue and purple) represent smaller eigenvalues (i.e. weaker factor structure). Higher values of the vertical axis should correspond to decreasing importance of the higher order PCs (warmer to cooler shades). The factor structure is estimated via rolling PCA over the last five years. The sample period uses monthly data from January 1996 until June 2023.

Figure 13: Δ PC Over Time (Hedge Fund Portfolios)
The figure displays the difference between the leading five principal components (based on the hedge fund portfolios) over time. To estimate the time-varying dynamics of the factor structure, I estimate PCA using a rolling window of five years of monthly data. PC1-2 (black line) refers to the difference between the leading two principal components: $PC1-2 = PC_{1,t} - PC_{2,t}$. PC2-3 (red line), PC3-4 (green line) and PC4-5 (orange line) follow the same calculation with their corresponding PCs. The sample period is from January 1996 until June 2023.

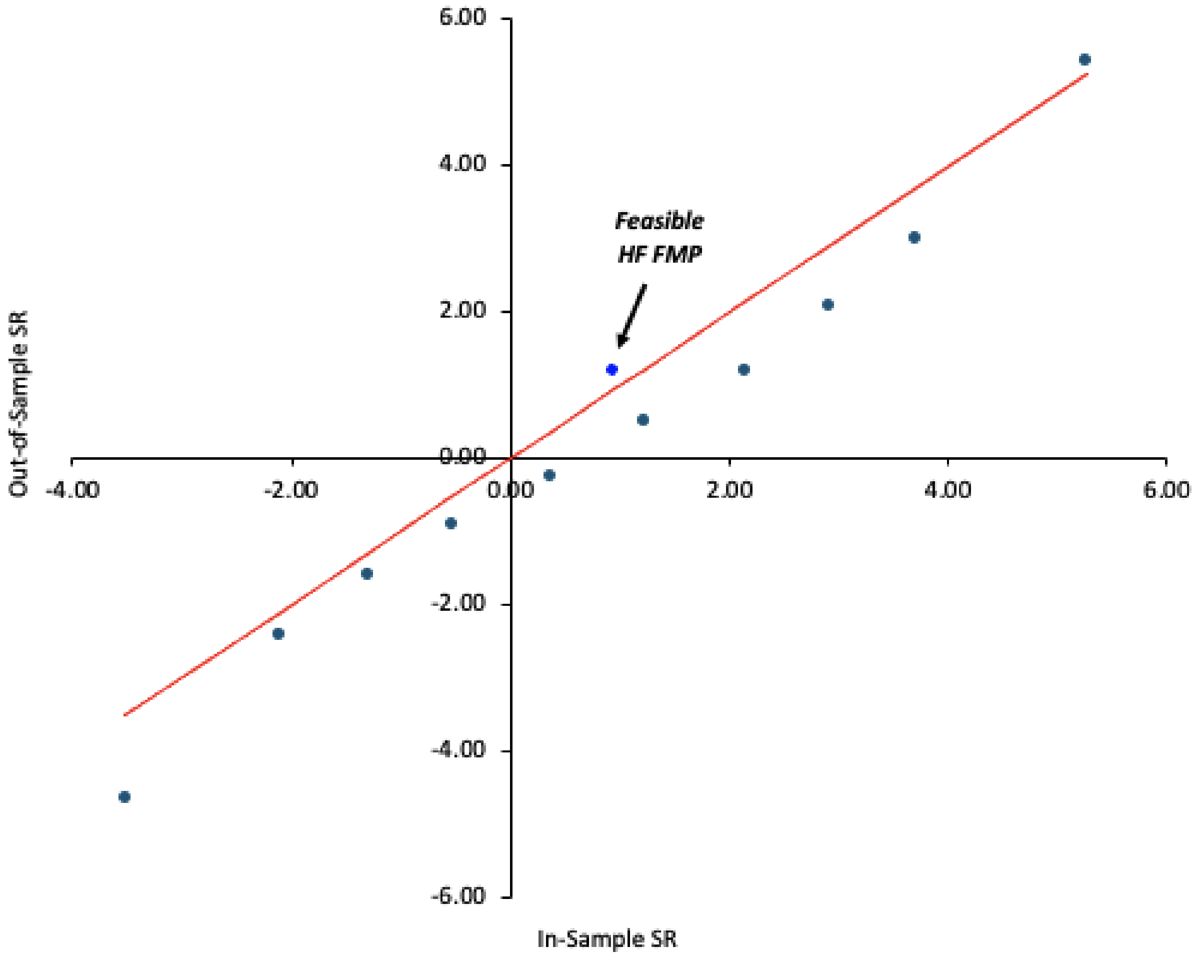


Figure 14: Robust Hedge Fund Managed Portfolio
The figure applies the pseudo OOS analysis as Figure 4 over a smaller sample period to account for the more limited time series of hedge fund data and estimation required involving the FMP. The figure now plots the implementable hedge fund FMP along next to the ten hedge fund decile portfolios. The full sample period uses monthly data from January 1996 to June 2023.

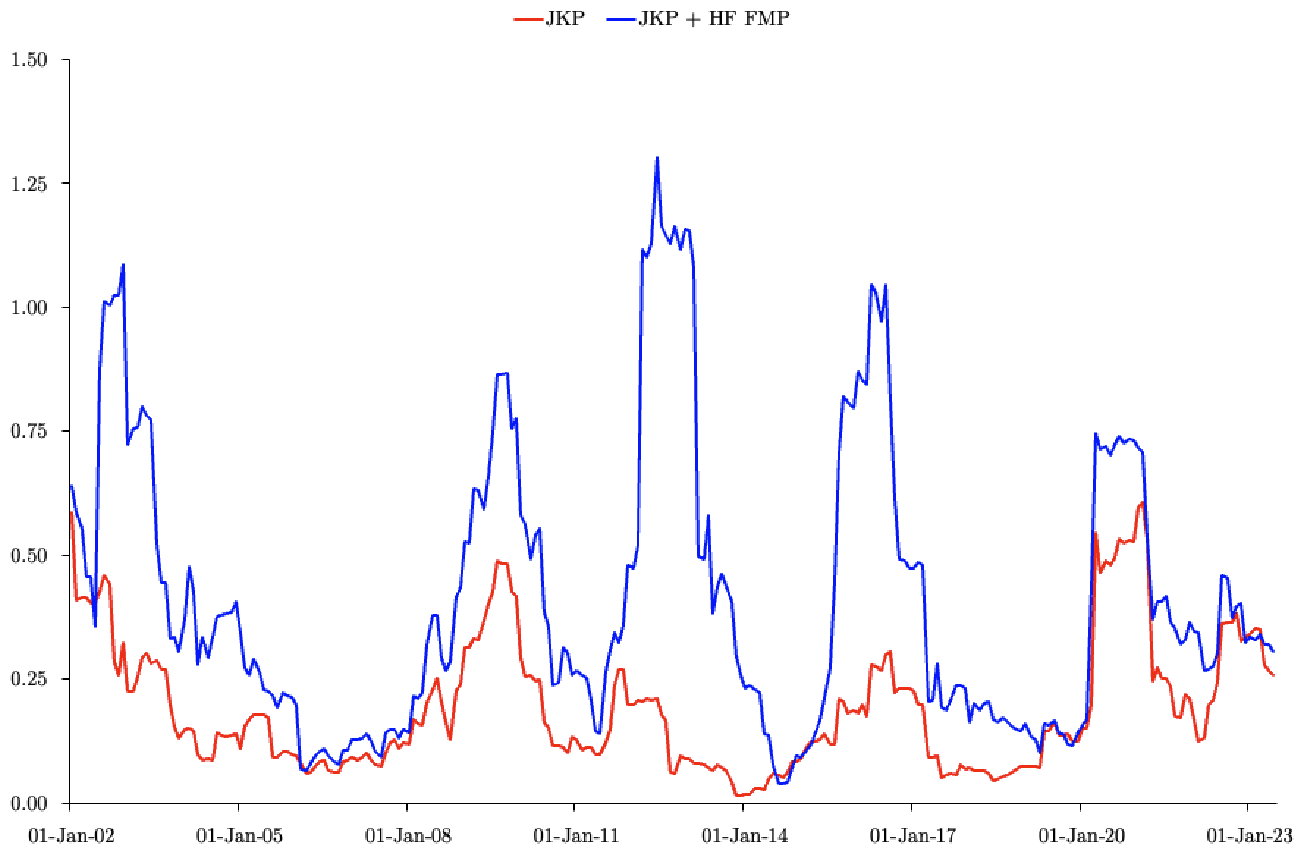


Figure 15: SDF Variance Over Time (with OOS HF Managed Portfolio)
The figure plots the one-year rolling variance estimates of the SDF based on the JKP anomaly test assets in combination with the new hedge fund FMP (blue line) in comparison to the JKP anomaly test assets (red line). The sample period uses monthly data from January 1996 until June 2023.